International Journal of Game Theory (1990) 19: 191–217

Nash Equilibria of n-Player Repeated Games With Semi-Standard Information

By E. Lehrer[1]

Abstract: The folk theorem is extended here to the case where after each stage of the repeated game each player is informed only about the equivalence classes of the pure actions which were used by the other players. The sets of upper equilibrium payoffs and of lower equilibrium payoffs are characterized here, and they are found to be different.

1 Introduction

In this paper we deal with n-player repeated games with imperfect monitoring, i.e., games in which a player is not informed necessarily about all other players' actions. Such situations are most often found in economic repetitive conflicts. For example, consider a repetitive interaction between firms practicing two sorts of actions: external, which are observable, and internal, which are nonobservable by other firms.

The objective of a player is to maximize the long run outcome of the game, and he may rely on the information he has collected previously during the game, when he decides what to do at a current turn. However, this information does not contain necessarily the payoffs during the game. For instance, in a case where interactions are frequently repeated and firms are large and spread out, the decision makers in the firms are not up-to-date with the financial state of their own firms. A player has to consider the payoffs during the game, because they affect the long run outcome (which will become known after the game terminates) but he cannot rely on them when he makes a decision.

At each stage of the game a player cannot observe his opponent's actions or their effects. This leads to a situation by which certain contracts become non-enforceable, and as a result certain efficient interactions may become excluded. Facing such a situation, one may look for the outcomes that can be sustained by self-enforcing agreements, and in particular for the most efficient ones. For this purpose we are looking at the model of n-player undiscounted repeated games, with imperfect monitoring, where the information structure is the following.

[1] Ehud Lehrer, Department of Managerial Economics and Decision Sciences, J. L. Kellogg Graduate School of Management, and Department of Mathematics, Northwestern University, 2001 Sheridan Road, Evanston, Illinois 60208.

0020–7276/90/2/191–217 $2.50

Let Σ_i be player i's set of actions and let $\overline{\Sigma}_i$ be a partition of it. In case player i takes action a of Σ_i, all other players are informed about the correspondent class $\bar{a}$, which is interpreted in our former example as the set of all the pairs (external-internal actions of the firm) that involve the same external action as does a. We call this information structure a semi-standard information.

It is worth mentioning that the above information structure covers also the case in which the payoffs are observable and the internal actions of a firm have no effect upon other firms. In particular, it includes as a special case the two-sided moral hazard, where each of the players is informed about the set that includes his opponent's action, but not about the action itself.

The well-known folk theorem (see Aumann [A2]) deals with undiscounted repeated games with standard information, which means that at every stage of the game each player is informed about the actions that took place at that stage. Radner [R1] and Rubinstein-Yaari [RY] have studied undiscounted repeated one-sided moral hazard games. In these games the principal cannot observe the action or a part of the action taken by the agent. The agent is fully informed, while the principal is informed only about the outcome, which is stochastically determined by the agent's action. In other words, the outcome is picked according to a distribution that depends on that action.

Fudenberg-Maskin [FM1, FM2], Radner [R2], Green-Porter [GP], and Abreu-Pearce-Stachetti [APS]have investigated games in which all the players are informed only about an outcome that depends (stochastically) on their actions. In these models the payoff of a player depends on this outcome and on the player's action.

Considering Nash equilibria, we are investigating the set of all the payoffs that can be sustained by equilibrium strategies. In particular, we get a full description of all the efficient outcomes in terms of the one-shot game.

Unilateral deviation from a prescribed strategy may cause the sequence of average payoffs of a player to be a divergent sequence, and the question is how to define a profitable deviation. Three ways to treat this problem are found in the literature. The first one (see [A2]) defines the profitability by the liminf. In other words, a player would be willing to deviate only if from a certain period the sequence (of the averages) is greater than the prescribed payoff. Using this pessimistic notion of profitability we define the lower equilibrium. The second way to address this problem (see [A2], [R1], [RY], [S1]) uses the limsup instead of the liminf. By this definition it is enough for the deviator to achieve from time to time, but infinitely many times, an average payoff above the prescribed payoff. Using this optimistic notion, we define the upper equilibrium.

The third way evaluates any bounded sequence by an extended limit notion, called a Banach limit. Using the Banach limit we can relate to any bounded sequence of average payoffs as if it has a limit and define the profitability notion in the natural way. Hart [H] uses the Banach limit in his characterization of the Nash equilibrium payoffs in repeated games with one-sided incomplete information.

We explore the Nash equilibrium payoffs that are correspondent to all these alternative definitions. We find, in contrast to Aumann [A2], that the set of lower equilibrium payoffs differs from the set of upper equilibrium payoffs. Furthermore, the latter coincides with the set of the Banach-equilibrium payoffs. The characterizations of these sets are given in Section 3 in terms of the sets $D_1, D_2, \dots, D_n$. D_i is the set of all the one-shot game joint actions $(p_1, \dots, p_n)$, where p_i is player i's best response versus the other players' actions, among all his actions that preserve the distribution over player i's classes, that are induced by p_i. In other words, if $(p_1, \dots, p_n)$ is played repeatedly, player j can profit only by detectable strategies.

The folk theorem characterizes the set of equilibrium payoffs as the set of all payoffs that own two properties: first, they are individually rational and second, they are feasible. Our characterizations have a similar formulation: the set of the upper equilibrium payoffs is the set of payoffs that are first, individually rational and second, sustained by joint actions that are members in all the D_j simultaneously (i.e., no player can profit by a nondetectable deviation).

The set of lower equilibrium payoffs is characterized as the individually rational payoffs that can be sustained by n (possibly different) joint actions, where the i-th joint action is contained in D_i and not necessarily in other D_j's. In other words, playing the i-th joint action, player i cannot deviate to a nondetectable action and still gain, but all other players may have profitable and nondetectable deviations.

Section 6 is devoted to concluding remarks. We refer to the Banach equilibrium and we state that the proof concerning upper equilibria holds also for Banach equilibria. The notion of uniform equilibrium is defined in that section (see [S2] for more extensive study). It turns out that the set of all uniform equilibrium payoffs coincides with the upper equilibrium payoff set. We refer also to the possibility of extending our results to a game in which the payoffs are stochastically dependent on the actions. Finally, other results of the author in related topics are mentioned, and the paper terminates with a few open questions.

2 The Model

Definition 2.1: An n-player repeated game with nonobservable actions consists of:

(i) n finite sets of actions: $\Sigma_1, \Sigma_2, \dots, \Sigma_n$.

(ii) n payoff functions $h_1, \dots, h_n$, where $h_i : X_{j=1}^n \Sigma_j \to \mathbb{R}$, $i = 1, \dots, n$.

(iii) n information functions $\lambda_1, \dots, \lambda_n$, and n information sets $L_1, \dots, L_n$, where $\lambda_i : X_{j=1}^n \Sigma_j \to L_i$, $i = 1, \dots, n$.

We constrain ourselves to a special kind of information function.

Definition 2.2: The information is *semi-standard* if for every $1 \le i \le n$ there is a partition of $\Sigma_i : \overline{\Sigma}_i$ s.t. for every $(x_1, \dots, x_n) \in X_{j=1}^n \Sigma_j$

$$\lambda_i(x_1, \dots, x_n) = (\overline{x}_1, \dots, \overline{x}_{i-1}, x_i, \overline{x}_{i+1}, \dots, \overline{x}_n).$$

where $\overline{x}_j$ is the class in $\overline{\Sigma}_i$ which includes x_j, $j = 1, \dots, n$.

In the case of semi-standard information

$$L_i = \overline{\Sigma}_i \times \ldots \times \overline{\Sigma}_{i-1} \times \Sigma_i \times \overline{\Sigma}_{i+1} \times \ldots \times \overline{\Sigma}_n.$$

In other words, in the case where player j acted $x_j \in \Sigma_j$ at the previous stage, player i ($i \neq j$) is informed about the equivalence class of x_j, which is $\overline{x}_j$.

Definition 2.3: The set of pure strategies of player i in the repeated game is:

$$\Sigma_i^* = \{(f_i^1, f_i^2, f_i^3, \ldots);\ \text{for each } n \in \mathbb{N}, f_i^n : L_i^{n-1} \to \Sigma_i\}$$

where L_i^0 is a singleton.

A mixed strategy of player i is a probability measure μ_i on Σ_i^*. The sets of mixed strategies of player i in the one-shot game (henceforth, mixed actions) and in the repeated game will be denoted by $\Delta(\Sigma_i)$ and $\Delta(\Sigma_i^*)$, respectively. An n-tuple of strategies is called a joint strategy.

Every joint pure strategy f induces two strings:

$(s_1^t(f),\ldots,s_n^t(f))_{t=1}^{\infty}$ and $(a_1^t(f),\ldots,a_n^t(f))_{t=1}^{\infty}$, where $s_i^t(f)$ and $a_i^t(f)$ are, respectively, the signal and the payoff that player i receives at stage t.

Definition 2.4: Let $\mu = (\mu_1,\mu_2,\ldots,\mu_n) \in X_{i=1}^n \Delta(\Sigma_i^*)$ and $T \in \mathbb{N}$

(i) $H_i^T(\mu) = \mathrm{Exp}_\mu(1/T\, \Sigma_{t=1}^T\, a_i^t(f)),\ i = 1,\ldots,n.$

$H_i^T(\mu)$ is the expectation of the average payoff of player i at the first T stages when μ_j is the mixed strategy played by player j, $j = 1,\ldots,n$.

(ii) $a_i^T(\mu) = \mathrm{Exp}_\mu(a_i^T(f)).$

Definition 2.5: Let $\mu = (\mu_1,\ldots,\mu_n) \in X_{i=1}^n \Delta(\Sigma_i^*)$.

(i) $H_i^*(\mu) = \lim_T H_i^T(\mu)$ if the limit exists, $i = 1,\ldots,n$.

(ii) $H^*(\mu) = (H_1^*(\mu),\ldots,H_n^*(\mu))$ if all $H_i^*(\mu)$ are defined.

Definition 2.6: Let $\mu \in X_{i=1}^n \Delta(\Sigma_i^*)$.

(i) μ is an *upper equilibrium* if

(a) $H^*(\mu)$ is defined, and
(b) for every $1 \leq i \leq n$ and $\overline{\mu}_i \in \Delta(\Sigma_i^*)$

$$\mathrm{limsup}_T\, H_i^T(\mu_1,\ldots,\mu_{i-1},\overline{\mu}_i,\mu_{i+1},\ldots,\mu_n) \leq H_i^*(\mu).$$

(ii) μ is a *lower equilibrium* if

(a) $H^*(\mu)$ is defined, and

(b) for every $1 \leq i \leq n$ and $\overline{\mu}_i \in \Delta(\Sigma_i^*)$
$\liminf_T H_i^T(\mu_i,\ldots,\mu_{i-1},\overline{\mu}_i,\mu_{i+1},\ldots,\mu_n) \leq H_i^*(\mu)$.

Notation 2.7: (1) UEP $= \{H^*(\mu) \mid \mu$ is an upper equilibrium$\}$;

(2) LEP $= \{H^*(\mu) \mid \mu$ is a lower equilibrium$\}$.

Our main task is to characterize UEP and LEP.

Remark 2.8: A repeated game with semi-standard information has a perfect recall. Therefore, by the Kuhn Theorem ([A1], [K]) we can consider behavior strategies, whenever this is convenient.

Notation 2.9: Let $1 \leq i \leq n$, $p \in \Delta(\Sigma_i)$ and $\overline{x} \in \overline{\Sigma}_i$ then $p(\overline{x}) = \Sigma_{w \in \overline{x}} p_w$.

We can relate to the classes of $\overline{\Sigma}_i$ as equivalence classes. We will say that $x, y \in \Sigma_i$ are *equivalent* ($x \sim_i y$) if $\overline{x} = \overline{y}$. This equivalence relation can be extended to $\Delta(\Sigma_i)$ in the following way:

Definition 2.10: Let $p, q \in \Delta(\Sigma_i)$ for some $1 \leq i \leq n$, then $p \sim_i q$ if for every class $\overline{x} \in \overline{\Sigma}_i$, $p(\overline{x}) = q(\overline{x})$.

The equivalence relations will play an important role in the characterization of UEP and of LEP.

3 The Main Theorems

The characterizations of LEP and of UEP are done by the sets D_i of joint mixed actions. For denoting the tuple $(\alpha_1,\ldots,\alpha_{i-1},\beta,\alpha_{i+1},\ldots,\alpha_n)$ we will use the notation (α_{-i},β).

Notation 3.1:

$$D_i = \{(p_1,\ldots,p_n) \in X_{i=1}^n \Delta(\Sigma_i) \mid h_i(p_i,\ldots,p_n) = \mathrm{Max}_{p \sim_i p_i} h_i(p_{-i},p)\},\ i = 1,\ldots,n.$$

In words, the set D_i contains all the joint mixed actions in which player i plays his best response among all the mixed actions which preserve the same distribution of other players' signals. The intuition is that if player i wants to increase his payoff by deviating to another mixed action he can do that solely by a detectable way-namely, by changing the probability distribution of his opponents' signals.

Notation 3.2: Let $i \in \{1,\dots,n\}$.

(i) $d_i = \mathrm{Min}_{(p_j)_{j \neq i} \in X_{j \neq i} \Delta(\Sigma_j)} \mathrm{Max}_{p_i \in \Delta(\Sigma_i)} h_i(p_{-i}, p_i)$.

(ii) σ_j^i is one of player j's mixed actions ($j \neq i$) that satisfies

$$d_i = \mathrm{Max}_{p \in \Delta(\Sigma_i)} h_i(\sigma_{-i}^i, p).$$

(iii) $\mathrm{IR} = \{(r_1,\dots,r_n) \in \mathbb{R}^n \mid r_j \geq d_j\,, j = 1,\dots,n\}$
$=$ the individually rational payoffs.

Theorem 3.3: In an n-player repeated game with semi-standard information

$$\mathrm{LEP} = \cap_{j=1}^n \operatorname{conv} h(D_j) \cap \mathrm{IR}.$$

Theorem 3.4: In an n-player repeated game with semi-standard information

$$\mathrm{UEP} = \operatorname{conv} h(\cap_{j=1}^n D_j) \cap \mathrm{IR}.$$

Example 3.5: Standard information.
The partition $\overline{\Sigma}_i$ is discrete:

$$\overline{\Sigma}_i = \{\{s\} \mid s \in \Sigma_i\},\ i = 1,\dots,n.$$

Hence, for every $1 \leq i \leq n$, $D_i = X_{i=1}^n \Delta(\Sigma_i)$ and, therefore,

$$\mathrm{UEP} = \mathrm{LEP} = \operatorname{conv} h(X_{i=1}^n \Delta(\Sigma_i)) \cap \mathrm{IR},$$

which is the content of the folk theorem.

Example 3.6: Composed prisoner's dilemma.
Each one of the two players has three actions: A_1,A_2 and B. A_1,A_2 are two actions of cooperation. A_1 is a strong cooperation and A_2 is a weak one. A player cannot distinguish between A_1 and A_2 of his opponent. Let the payoff matrix be:

	B	A_1	A_2
B	1, 1	4,0	6,-1
A_1	0,4	3,3	5,2
A_2	-1,6	2,5	4,4

Here $\overline{\Sigma}_1 = \overline{\Sigma}_2 = \{\{B\}, \{A_1,A_2\}\}$. By direct computation we get:

$$\text{UEP} = \text{LEP} = \text{Conv}\,\{(1,1),(1,11/3),\ (11/3,1),(3,3)\}.$$

The weak cooperation payoff (3,3) is an equilibrium payoff, but the strong cooperation payoff (4,4) is not.

Example 3.7: The repeated game of:

	A_1	A_2	B
A_1	0,0	1,0	3,0
A_2	0,1	1,1	2,2
B	0,3	2,2	1.75,1.75

where $\Sigma_1 = \overline{\Sigma}_2 = \{\{B\},\{A_1,A_2\}\}$.

$\text{conv}\, h(D_1) \cap \text{conv}\, h(D_2) = \text{conv}\,\{(0,0),(3,0),(0,3),(2,2)\}$ and
$\text{conv}\, h(D_1 \cap D_2) = \text{conv}\,\{(0,0),(3,0),(0,3),(1.75,1.75)\}$.

Thus, $\text{UEP} \subsetneq \text{LEP}$ (see Figure 1).

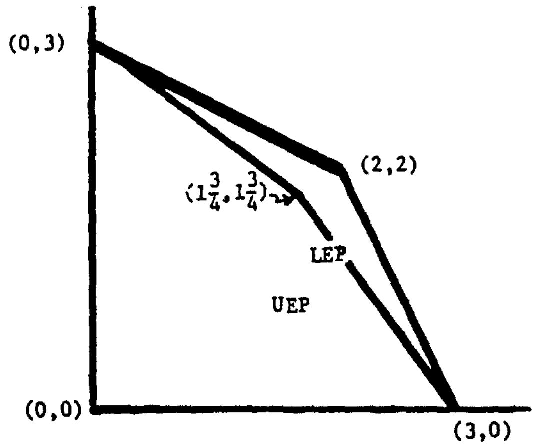

Fig. 1.

Corollary 3.8: Let $G_1^* = ((\Sigma_i, h_i, \overline{\Sigma}_i^1)_{i=1}^n)$ and $G_2^* = ((\Sigma_i, h_i, \overline{\Sigma}_i^2)_{i=1}^n)$ be two n-player repeated games with semi-standard information, where $\overline{\Sigma}_i^2$ refines $\overline{\Sigma}_i^1$ for every $i = 1,\ldots,n$. Then

(i) $\mathrm{LEP}(G_1^*) \subseteq \mathrm{LEP}(G_2^*)$;

(ii) $\mathrm{UEP}(G_1^*) \subseteq \mathrm{UEP}(G_2^*)$.

Proof: $\overline{\Sigma}_i^2$ refines $\overline{\Sigma}_i^1$ means that $D_i(G_1^*) \subseteq D_i(G_2^*)$ for each $i = 1,\ldots,n$. Therefore the proof is clear. //

The intuitive meaning of Corollary 3.8 is that the players, by knowing more about their opponents, can enlarge the sets of equilibrium payoffs since there are less non-detectable deviations.

4 Proof of Theorem 3.3

Step 1: LEP $\subseteq$ IR.

If **LEP** $\not\subseteq$ IR, then there is $\alpha = (\alpha_1,\ldots,\alpha_n) \in$ LEP s.t. $\alpha_i < d_i$ for some i. W.l.o.g. $i = 1$. Let $f = (f_1,\ldots,f_n)$ be a lower equilibrium strategy with $H^*(f) = \alpha$. We will define a deviation strategy g_1 of player 1 which increases his payoff. By Remark 2.8 we can relate to $(f_1,\ldots,f_n)$ as joint behavior strategy.

Notation 4.1: If $x^t = (x_1^t,\ldots,x_n^t) \in X_{i=1}^n \Sigma_i^t$ for some $1 \leq i \leq n$ and $t \in \mathbb{N}$ then $\bar{x}^t = (\bar{x}_1^t, \bar{x}_2^t,\ldots,\bar{x}_n^t) = \{(z_1^t,\ldots,z_n^t) | \bar{z}_j^t = \bar{x}_j^t, j = 1,\ldots,n\}$.

Lemma 4.2: Fix a joint strategy $\sigma \in X_{i=1}^n \Delta(\Sigma_i^*)$. For every $x^t = (x_1^t,\ldots,x_n^t) \in X_{i=1}^n \Sigma_i^t$.

(i) $pr(\bar{x}^t) = pr(\bar{x}^{t-1}) \Pi_{j=1}^n pr(\bar{x}_j^t | \bar{x}^{t-1})$.

(ii) $pr(x^t) = pr(\bar{x}^t) \Pi_{j=1}^n pr(x_j^t | \bar{x}^t)$,

where all the probabilities are those induced by σ.

Proof: (i) Denote by y_j^t the last coordinate of x_j^t and the first $t-1$ ones by x_j^{t-1}.

$$pr(\bar{x}^t) = \Sigma_{w_j^t \in \bar{y}_j^t}\ pr(\bar{x}^{t-1})\, pr(w^t \mid \bar{x}^{t-1})$$

$$= pr(\bar{x}^{t-1}) = \Sigma_{w_j^t \in \bar{y}_j^t}\ \Pi_{j=1}^n pr(w_j^t \mid \bar{x}^{t-1})$$

$$= pr(\bar{x}^{t-1}) = \Pi_{j=1}^n\ \Sigma_{w_j^t \in \bar{y}_j^t}\ pr(w_j^t \mid \bar{x}^{t-1})$$

$$= pr(\bar{x}^{t-1}) \Pi_{j=1}^n pr(\bar{y}_j^t \mid \bar{x}^{t-1}).$$

(ii) By induction on t.

By the definition of the probabilities,

$$pr(x^t) = pr(x^{t-1}) \Pi_{j=1}^n pr(y_j^t \mid x^{t-1}). \tag{1}$$

By the induction hypothesis,

$$pr(x^{t-1}) = pr(\bar{x}^{t-1}) \Pi_{j=1}^n pr(x_j^{t-1} \mid \bar{x}^{t-1}). \tag{2}$$

Combine (1) and (2) to get:

$$pr(x^t) = pr(\bar{x}^{t-1}) \Pi_{j=1}^n pr(x_j^{t-1} | \bar{x}^{t-1}) \Pi_{j=1}^n pr(y_j^t | x^{t-1}). \tag{3}$$

Since the information is semi-standard,

$$pr(y_j^t | x^{t-1}) = pr(y_j^t | \bar{x}_{-j}^{t-1}, x_j^{t-1})$$

for every $1 \leq j \leq n$. By (3) and (i):

$$pr(x^t) = pr(\bar{x}^{t-1})\,\Pi_{j=1}^n\, pr(x_j^t\,|\,\bar{x}^{t-1})$$

$$= pr(\bar{x}^t)\,\Pi_{j=1}^n\, pr(x_j^t\,|\,\bar{x}_{-j}^{t-1},\, \bar{x}_j^t)$$

$$= pr(\bar{x}^t)\,\Pi_{j=1}^n\, pr(x_j^t\,|\,\bar{x}^t). \qquad //$$

Now the deviation $g_1 = (g_1^1, g_1^2, \ldots)$ can be defined by induction. g_1^1 is the mixed action in $\Delta(\Sigma_1)$ which ensures player 1 his minmax payoff. Namely, $h_1(g_1^1, f_2^1, \ldots, f_n^1) \geq d_1$.

Notation 4.3: Let $t \in \mathbb{N}$ and $x^t \in X\Sigma_i^t$ and $1 \leq j \leq n$, then

$$\bar{f}_j^{t+1}(\bar{x}^t) = \Sigma_{z_j^t \in \Sigma_j^t}\, pr(z_j^t\,|\,\bar{x}^t)\, f_j^{t+1}(\bar{x}_{-j}^t,\, z_j^t).$$

Assume that $g_1^2, g_1^3, \ldots, g_1^t$ have been defined in a way that for every $x^{s-1} \in X\Sigma_i^{s-1}$ $(2 \leq s \leq t)$ the following holds:

$$h_1(\bar{f}_{-1}^{\,s}(\bar{x}^{s-1}), g_1^s(\bar{x}_{-1}^{s-1}, x_1^{s-1})) \geq d_1.$$

Let $x^t \in X\Sigma_i^t$. Define $g_1^{t+1}(\bar{x}_{-1}^t, x_1^t)$ to be a certain mixed action in $\Delta(\Sigma_1)$ which ensures at least d_1 for player 1, when each player $j \neq 1$ plays $\bar{f}_j^{t+1}(\bar{x}^t)$. Note that $\bar{g}_1^{t+1}(\bar{x}^t) = g_1^{t+1}(\bar{x}_{-1}^t, x_1^t)$ for all x_1^t.

We will prove that $\inf_T H_1^T(f_{-1}, g_1) \geq d_1$. The expected payoff of player 1 at stage $t+1$ is

$$\Sigma_{x^t \in X_{i=1}^n \Sigma_i^t}\, pr(x^t)\, h_1(g_1^{t+1}(\bar{x}_{-1}^t, x_1^t), f_2^{t+1}(\bar{x}_{-2}^t, x_2^t), \ldots, f_n^{t+1}(\bar{x}_{-n}^t, x_n^t)).$$

By Lemma 4.2 (ii) this is equal to

$$\Sigma_{x^t \in X_{i=1}^n \Sigma_i^t}\, pr(\bar{x}^t)\Pi_{j=1}^n\, pr(x_j^t\,|\,\bar{x}^t)$$

$$\times\, h_1(g_1^{t+1}(\bar{x}_{-1}^t, x_1^t), f_2^{t+1}(\bar{x}_{-2}^t, x_2^t), \ldots, f_n^{t+1}(\bar{x}_{-n}^t, x_n^t))$$

$$= \Sigma_{\bar{x}^t}[\, pr(\bar{x}^t)\, \Sigma_{x^t \in \bar{x}^t}\Pi_{j=1}^n\, pr(x_j^t\,|\,\bar{x}^t)$$

$$\times\, h_1(g_1^{t+1}(\bar{x}_{-1}^t, x_1^t), f_2^{t+1}(\bar{x}_{-2}^t, x_2^t), \ldots, f_n^{t+1}(\bar{x}_{-n}^t, x_n^t))].$$

Since h_1 is multilinear this is equal to

$$\Sigma_{\bar{x}^t}\, pr(\bar{x}^t)\, h_1(\bar{f}_{-1}^{\,t+1}(\bar{x}^t), \bar{g}_1^{t+1}(\bar{x}^t))$$

(recall Notation 4.3).

By the definition of g_1^{t+1}, this is greater or equal to

$\Sigma_{\bar{x}^t}\ pr(\bar{x}^t)d_1 \ \geq\ d_1.$

Thus, $H_1^T(f_{-1},g_1) \geq d_1$ for all $T \in \mathbb{N}$.

Step 2: LEP $\subseteq \cap_{i=1}^n$ conv $h(D_i)$.

If LEP $\subseteq \cap_{i=1}^n$ conv $h(D_i)$, then w.l.o.g. there is some lower equilibrium strategy $(f_1,...,f_n)$ s.t. $H^*(f_1,...,f_n) = (\alpha_1,...,\alpha_n) \notin$ conv $h(D_1)$.

Define a strategy $g_1 = (g_1^1,g_1^2,...)$ by induction. g_1^1 is a mixed action $\bar{p}$ in $\Delta(\Sigma_1)$ which satisfies

(i) $f_1^1 \sim_1 \bar{p}$ and

(ii) $h_1(f_{-1}^1,\bar{p}) = \mathrm{Max}_{p \sim_1 f_1^1} h_1(f_{-1}^1,p)$.

Assume that g_1^{r+1} had been defined for all $r < t$ in such a way that for all $x^r \in X_{j=1}^n \Sigma_j^r$ the following holds:

$$g_1^{r+1}(x_{-1}^r,x_1^r) \sim_1 \bar{f}_1^{r+1}(\bar{x}^r) \tag{5}$$

and

$$g_1^{r+1}(\bar{x}_{-1}^r,\ x_1^r) \text{ is the best response versus} \tag{6}$$

$$(\bar{f}_2^{r+1}(\bar{x}^r),\bar{f}_3^{r+1}(\bar{x}^r),...,\bar{f}_n^{r+1}))$$

among all the mixed actions in $\Delta(\Sigma_1)$ which are equivalent to the right side of (5).

Now define g_1^{t+1} to be the strategy which satisfies (5) and (6) for all $x^t \in X_{j=1}^n \Sigma_j^t$. Define g to be (f_{-1},g_1).

It will be shown (in Lemma 4.4) that by playing g_1, player 1 does not affect the probability of signals that the other players get during the game. Therefore, the mixed actions played by other players after every history are also retained. In other words, the deviation of player 1 is not detectable. Furthermore, we will prove at Lemma 4.10, using the previous Lemmata, that at all the stages in a relatively big set of stages player 1 profits at least by $\epsilon > 0$ with probability of at least $\delta > 0$, which provides the desired proof.

Lemma 4.4: For any $j \neq 1$, $t \in \mathbb{N}$ and $x^t \in X_{j=1}^n \Sigma_j^t$,

(i) $pr_g(\bar{x}^t) = pr_f(\bar{x}^t)$.

(ii) $pr_g(x_j^t\,|\bar{x}^t) = pr_f(x_j^t\,|\bar{x}^t)$.

Proof: By (i) of Lemma 4.2,

$$pr_g(\overline{x}^t) = pr_g(\overline{x}^{t-1})\, \Pi_{j=1}^{n}\, pr_g(x_j^t\,|\,\overline{x}^{t-1}).$$

The induction hypothesis will be both that $pr_g(x_j^{t-1}|\overline{x}^{t-1}) = pr_f\,(x_j^{t-1}|\overline{x}^{t-1})$, and that $pr_g(\overline{x}^{t-1}) = pr_f\,(\overline{x}^{t-1})$, for every $j = 2,\ldots,n$ and $x^{t-1} \in X\Sigma_i^{t-1}$. In order to prove (i) we have to prove that for every $1 \leq j \leq n$, $pr_g(\overline{x}_j^t\,|\overline{x}^{t-1}) = pr_f\,(\overline{x}_j^t\,|\overline{x}^{t-1})$.

Letting y_j^t be the last coordinate of x_j^t, we obtain

$$pr_g(\overline{x}_j^t\,|\overline{x}^{t-1}) = pr_g(\overline{y}_j^t\,|\overline{x}^{t-1})$$

$$= \Sigma_{z_j^{t-1} \in \overline{x}_j^{t-1}}\, pr_g(z_j^{t-1}|\overline{x}^{t-1})\, pr_g(\overline{y}_j^t\,|\overline{x}_{-j}^{t-1}, z_j^{t-1}) = (*).$$

By the induction hypothesis and because the strategy of player j ($j > 1$) in g is f_j, namely, because

$$pr_g(\overline{y}_j^t\,|\overline{x}_{-j}^{t-1},\, z_j^{t-1}) = pr_f\,(\overline{y}_j^t\,|\overline{x}_{-j}^{t-1},\, z_j^{t-1}),$$

whenever $j > 1$, we can write (*) also substituting g by f. If $j = 1$ then, by (5), $pr_g(\overline{y}_j^t\,|\overline{x}^{t-1}) = pr_f(\overline{y}_j^t\,|\,\overline{x}^{t-1})$. This concludes the proof of (i).

(ii) remains to be proven. By direct computation of conditional probabilities we get for every strategy $\sigma \in X_{i=1}^n\, \Delta(\Sigma_i^*)$, and $x^t \in X_{i=1}^n\, \Sigma_i^t$ the following:

$$pr_\sigma(x_j^t\,|\overline{x}^t) = pr_\sigma(x_j^t\,|\overline{x}_{-j}^{t-1}, \overline{x}_j^t)$$

$$= pr_\sigma(y_j^t\,|\overline{x}_{-j}^{t-1}, x_j^{t-1})\, pr_\sigma(x_j^{t-1}|\overline{x}^{t-1})/$$

$$/pr_\sigma(\overline{y}_j^t\,|\overline{x}^{t-1}) \text{ provided that } pr_\sigma(\overline{x}_{-j}^t, x_j^t) > 0.$$

Combine this and the following three equalities in order to get the proof of (ii). The first one,

$$pr_g(y_j^t\,|\overline{x}_{-j}^{t-1}, x_j^{t-1}) = pr_f\,(y_j^t\,|(\overline{x}_{-j}^{t-1},\, x_j^{t-1}),$$

holds because the strategies of player j ($j \neq 1$) in g and in f are equal. The second one has already been shown, $pr_g(\overline{y}_j^t\,|\overline{x}^{t-1}) = pr_f(\overline{y}_j^t\,|\,\overline{x}^{t-1})$, and the third follows from the induction hypothesis, $pr_g(x_j^{t-1}|\overline{x}^{t-1}) = pr_f\,(x_j^{t-1}|\overline{x}^{t-1})$. //

Lemma 4.5: $H^t(g) \in \text{conv}\, h(D_1)$ for all t.

Proof: It is ensured by (5) and (6). //

Let L be a hyperplane in $\mathbb{R}^n$ dividing it into two parts, L^- and L^+, and separating $(\alpha_1,\dots,\alpha_n)$ and conv $h(D_1)$ in such a way that:[2]

(a) $\text{dist}((\alpha_1,\dots,\alpha_n), L) = \text{dist}(\text{conv } h(D_1), L) = \gamma > 0$ and

(b) $(\alpha_1,\dots,\alpha_n) \in L^-$.

Lemma 4.6: There is an $\epsilon > 0$ s.t. if $(p_1,p_2,\dots,p_n) \in X_{j=1}^n \Delta(\Sigma_j)$ and $h(p_1,\dots,p_n) \in L^-$ then there is a $q_1 \sim_1 p_1$ which satisfies $h_1(p_{-1},q_1) > h_1(p_1,p_2,\dots,p_n) + \epsilon$.

Proof: Otherwise there is a sequence $\{(p_1^s,p_2^s,\dots,p_n^s)\}_{s=1}^{\infty} \in X_{j=1}^n \Delta(\Sigma_j)$ s.t. for every $q_1^s \sim_1 p_1^s$, $h_1(p_1^s,\dots,p_n^s) + 1/s > h_1(q_1^s,p_2^s,\dots,p_n^s)$ and $h_1(p_1^s,\dots,p_n^s) \in L^-$. By compactness there is an accumulation point of the sequence. Denote this point by $(p_1,\dots,p_n)$. Now, $h(p_1,\dots,p_n) \in L^-$ and for every $q_1 \sim_1 p_1$, $h_1(q_1,p_2,\dots,p_n) \le h_1(p_1,p_2,\dots,p_n)$. This means that $(p_1,\dots,p_n) \in D_1$ and $h(p_1,\dots,p_n) \in \text{conv } h(D_1) \subseteq L^+$, a contradiction. //

Definition 4.7: A set $M \subseteq \mathbb{N}$ has a (upper) (lower) density η if

$$\lim_t |M \cap \{1,\dots,t\}|/t = \eta$$

$$(\limsup |M \cap \{1,\dots,t\}|/t = \eta)$$

$$(\liminf |M \cap \{1,\dots,t\}|/t = \eta)$$

Remark 4.8: The set $M = \{t \mid (a_1^t(f),\dots,a_n^t(f)) \in L^-$ and $\text{dist}((a_1^t(f),\dots,a_n^t(f)), L) > \gamma/2\}$ has a positive lower density η.

Lemma 4.9: There is a positive constant $\delta > 0$ s.t. for every $t \in M$,

$$pr_f\,\{\bar{x}^{t-1} \mid h(\bar{f}_1^{\,t}(\bar{x}^{t-1}), \bar{f}_2^{\,t}(\bar{x}^{t-1}),\dots,\bar{f}_n^{\,t}(\bar{x}^{t-1})) \in L^-\} > \delta.$$

Proof: The lemma holds because the set of feasible payoffs is bounded, and because of the definition of M. //

Lemma 4.10: $\liminf_T H_1^T(g) > H_1^*(f)$.

Proof: For every integer t, $\text{Exp}_f\,(h_1(y^{t+1})) =$

$$= \Sigma_{x^t \in X\Sigma_j^t}\; pr_f\,(x^t)\; h_1(f_1^{\,t+1}(\bar{x}_{-1}^t,x_1^t), f_2^{\,t+1}(\bar{x}_{-2}^t,x_2^t),\dots,f_n^{\,t+1}(\bar{x}_{-n}^t,x_n^t)).$$

---

[2] For $x \in \mathbb{R}^n$ and $K \subseteq \mathbb{R}^n$, $\text{dist}(x,K) = \inf_{y \in K} \|x - y\|_1$.

By (ii) of Lemma 4.2 this is equal to

$$\Sigma_{\bar{x}^t \in \Sigma^t} \Sigma_{w_j^t \in \bar{x}_j^t} \; pr_f(\bar{x}^t) \, \Pi_{j \geq 1} \, pr_f(w_j^t \mid \bar{x}^t)$$
$$\cdot \; h_1(f_1^{t+1}(\bar{x}_{-1}^t, w_1^t), f_2^{t+1}(\bar{x}_{-2}^t, w_2^t), \ldots, f_n^{t+1}(\bar{x}_{-n}^t, w_n^t)).$$

h_1 is multilinear; therefore, this is equal to

$$\Sigma_{\bar{x}^t \in \Sigma^t} \; pr_f(\bar{x}^t) \, h_1(\bar{f}_1^{t+1}(\bar{x}^t), \bar{f}_2^{t+1}(\bar{x}^t), \ldots, \bar{f}_n^{t+1}(\bar{x}^t)).$$

However, by (ii) of Lemma 4.4, for every $j > 1$, $\bar{f}_j^{t+1}(\bar{x}^t)$ remains unchanged both when it is defined by $pr_f(\cdot)$ or by $pr_g(\cdot)$. Thus, by (6) player 1, by playing g_1, can achieve in each stage at least what he could achieve by playing f_1. Furthermore, according to Lemmata 4.6, 4.9, and Remark 4.8, player 1, by playing g_1, profits at least by $\epsilon > 0$ with probability of at least $\delta > 0$ at each stage of a set M, which has a lower density $\eta > 0$. Thus

$$\liminf_T H_1^T(g) > H_1^*(f) + \eta\epsilon\delta. \quad //$$

Step 3: $\cap_{j=1}^n \operatorname{conv} h(D_j) \cap \mathrm{IR} \subseteq \mathrm{LEP}$.

From here on we assume that $h_i \geq 0$ for all i. W will denote the greatest payoff appearing in the game.

Lemma 4.11: The extreme points of conv $h(D_j)$ are of the form $h(p_1,\ldots,p_n)$, where p_j is a pure strategy.

Proof: Let $p = (p_1,\ldots,p_n) \in D_j$. $p \in \Delta(\Sigma_j)$ means that[3] $p_j = \Sigma_{s \in \Sigma_j} \alpha_s \delta_s$. However, $p \in D_j$ implies that $(p_{-j}, \delta_s) \in D_j$ for every s with $\alpha_s > 0$. Thus, $h(p) = \Sigma_{s \in \Sigma_j} \alpha_s \, h(p_{-j}, \delta_s) \in \operatorname{conv} h(D_j)$. //

In order to define a lower equilibrium strategy f, it is enough to ensure that for each player there are infinitely many stages in which his average payoff cannot exceed his prescribed payoff. For this purpose we divide the set of stages into consecutive blocks. The first one will be devoted to player 1, the second to player 2, the n-th block to player n, the $n + 1$-th block to player 1, and so on. A block devoted to player i is called an i-block. In the stages of the i-blocks player i will play a *pure action* which is a best response without being detected. In other words, in these blocks player i will not be able to increase his average payoff without being discovered by other players. This means that a player is forced, by the threat of punishment, not to deviate in the blocks that are devoted to him.

[3] δ_s is the mixed action that assigns probability 1 to the action s.

The blocks will be defined in such a way that the number of stages preceding a block is very small compared to its length. Therefore, the average payoff of player i immediately after an i-block terminates cannot exceed by much his prescribed payoff.

A punishment of player i is executed by all other players when they observe, at a stage of an i-block, that a signal of player i differs from the signal they expected to observe (recall that at these stages player i plays only pure strategies). In this case the players can be sure that player i had deviated and they punish him from the moment of deviation on forever.

We remark here that such "grim" strategy could not be defined at the next section, since there, when players come to a conclusion that a player had deviated they may, with a positive probability, be wrong.

Let $\alpha = (\alpha_1,\ldots,\alpha_n) \in \cap_{j=1}^{n} \operatorname{conv} h(D_j) \cap \mathrm{IR}$. By Lemma 4.11 and by the Caratheodory theorem, for each $1 \le j \le n$ there are $n + 1$ joint mixed actions $q_j^m = (q_j^m(1),\ldots,q_j^m(n))$, $m = 1,\ldots,n + 1$ in D_j (where $q_j^m(j) \in \Sigma_j$) and $n + 1$ positive constants $(\gamma_j^m)_{m=1}^{n+1}$ with total sum 1 s.t. $\alpha = \Sigma_{m=1}^{n+1} \gamma_j^m h(q_j^m)$.

Divide $\mathbb{N}$, the set of stages, into blocks: $B_1, B_2, \ldots$, with the following properties:

$$|B_1| = 1 \tag{7a}$$

$$|B_j| = j\, \Sigma_{i<j} |B_i| \tag{7b}$$

$$\operatorname{Max} B_j + 1 = \operatorname{Min} B_{j+1} \tag{7c}$$

Each one of the blocks B_j is divided into $n + 1$ subsets, $B_j^1, B_j^2, \ldots, B_j^{n+1}$, in such a way that

$$\left|\, |B_j^m \cap T|/j - \gamma_{j(\mathrm{mod}\, n)}^m \right| < 1/j\,, \tag{8}$$ [4]

for every segment T of B_j with length j, and for every $1 \le m \le n + 1$.

Denote $j(\mathrm{mod}\ n)$ by $j(n)$. The strategy f will be defined as follows. At all the stages of B_j^m, player i will play the strategy $q_{j(n)}^m(i)$ unless he gets in some stage of B_j^m a signal which points out that player $j(n)$ did not play the pure strategy he should have played; to be precise, unless player i does not get a signal from the set[5]

$$\{(\bar{x}_{-i}, x_i) \,|\, x_i \in \Sigma_i,\ \bar{x}_\ell \in \bar{\Sigma}_\ell,\ \ell \ne j(n), \text{ and } \bar{x}_{j(n)} = \bar{q}_{j(n)}^m(j(n))\}.$$

In this case player i will play $\sigma_i^{j(n)}$ from that stage on forever (recall Notation 3.2(ii)).

4 For simplicity, if j is divided by n ($n\,|\,j$) then $j(\mathrm{mod}\ n)$ will denote n and not zero as usual.

5 $i = j(\mathrm{mod}\ n)$ if $j - i$ is divided by n.

Lemma 4.12: $H^*(f) = (\alpha_1,\alpha_2,...,\alpha_n)$.

Proof: Denote $a^t = (a_1^t,...,a_n^t) = (a_1^t(f),...,a_n^t(f))$ (recall Definition 2.4(ii)). Let $t \in B_j^m$. By the definition of f each player i plays $q_{j(n)}^m(i)$, thus $a_i^t = h_i(q_{j(n)}^m)$. The average of the expectations on a segment T of length j in B_j is:

$$1/j\, \Sigma_{t\in B_j \cap T}\, a^t = 1/j\, \Sigma_{m=1}^{n+1} |\, B_j^m \cap T| h(q_{j(n)}^m).$$

By (8):

$$\|1/j\, \Sigma_{t\in T\cap B_j}\, a^t - \alpha\|_\infty \leq (1/j)W. \tag{9}$$

Hence,

$$\|1/|B_j|\, \Sigma_{t\in B_j}\, a^t - \alpha\|_\infty \leq (1/j)W. \tag{10}$$

(9), (10) and (7b) give the desired proof. //

Lemma 4.13: f is a lower equilibrium strategy.

Proof: Let g_k be a behavior strategy of player k in the repeated game. Let t be a stage in B_s^t where $k \equiv s(\text{mod } n)$, and denote by A_t the event in which player k does not play an action from the class $\overline{q}_k^m(k)$ at stage t. Denote $B_s = A_s \setminus \cup_{j<s} A_j$, $C_s = \cup_{j<s} A_j$, $b_s = pr(B_s)$, and $c_s = pr(C_s)$.

By the definition of f, the expected payoff of player k at stage t is less than

$$c_t\, d_k + b_t W + (1 - c_{t+1}) h_k\, (q_k^m) \tag{11}.$$

The first term is the probability of k to be detected, multiplied by his minmax payoff d_k. The second term is the probability of player k to defect at stage t for the first time[6], multiplied by the bound of its profit W. The third term is the probability not to act outside the set of actions $\overline{q}_k^m(k)$ which is $1 - c_{t+1}$, multiplied by the maximum payoff of player k (because $q_k^m \in D_k$) when he is playing an action in $\overline{q}_k^m(k)$ and all the rest are playing $q_k^m(i)$, $i \neq k$, which is $h_k(q_k^m)$.

By (11), whenever $k = s(n)$

$$1/|B_s^m|\, \Sigma_{t\in B_s^m} a_k^t \leq$$

$$1/|B_s^m|\, \Sigma_{t\in B_s^m} [c_t d_k + b_t W + (1 - c_{t+1}) h_k(q_k^m)]. \tag{12}$$

6 Without including defections outside blocks in which player k is checked up, namely, outside $\cup_{k=w(n)} B_w$.

Since $c_{t+1} = c_t + b_t$ and $h_k(q_k^m) \geq d_k$, the right side of (12) is less or equal to

$$1/|B_s^m| \; \Sigma_{t\in B_s^m} \, [(1 - c_t)h_k(q_k^m) + b_t W], \tag{13}$$

for every $m = 1,\ldots,n + 1$. Because of (8) and since $b_t \to 0$,

$$1/|B_s| \; \Sigma_{m=1}^{n+1} \Sigma_{t\in B_s^m} \, [(1 - c_t)h_k(q_k^m) + b_t W] \leq \alpha_k + \epsilon(s), \tag{14}$$

where $\epsilon(s) \to_{s\to\infty} 0$. (14) and (7b) imply that whenever $k \equiv s(n)$

$$1/\mathrm{Max}\, B_s \, \Sigma_{t=1}^{\mathrm{Max}\, B_s} a_k^t \leq W/(s + 1) + \alpha_k + \epsilon(s) \underset{s\to 0}{\to} \alpha_k\,, \tag{15}$$

which concludes the proof. //

5 Proof of Theorem 3.4

We know that LEP $\subseteq$ IR. Since UEP $\subseteq$ LEP we get UEP $\subseteq$ IR.

Step 2: UEP $\subseteq$ conv $h(\cap_{j=1}^n D_j)$.

Proof: Let $\alpha \notin$ conv $h(\cap_{j=1}^n = D_j)$. We will show that $\alpha \notin$ UEP, by showing that if $f = (f_1, f_2,\ldots,f_n) \in X\,\Delta(\Sigma_i^*)$ and $H^*(f) = \alpha$, then f is not an upper equilibrium strategy.

Let L be a hyper-plan that separates conv $h(\cap_{j=1}^n D_j)$ and α with the following properties:

$$\mathrm{dist}(\alpha,L) = \mathrm{dist}(\mathrm{conv}\; h(\cap_{j=1}^n D_j),\, L) = \gamma > 0, \tag{16a}$$

where dist$(\cdot,\cdot)$ is the distance induced by the L_1 norm and,

$$a \in L^-. \tag{16b}$$

By Remark 4.8 there is a set of stages M, with a positive lower density, such that at each stage $t \in M$, $(a_1^t(f),\ldots,a_n^t(f)) \in L^-$. Furthermore, like in Lemma 4.9 we can find $a\ \delta > 0$ such that

$$pr\{\bar{x}^{t-1} \,|\, h(\bar{f}_1^t(\bar{x}^{t-1}),\ldots,\bar{f}_n^t(\bar{x}^{t-1})) \in L^- \} > \delta$$

for every $t \in M$.

By (16a) and (16b), and since h is multilinear, there is $\beta_1 > 0$ for which[7]

$$pr\,\{\bar{x}^{t-1}\,|\,\mathrm{dist}((\bar{f}_1^t(\bar{x}^{t-1}),\ldots,\bar{f}_n^t(\bar{x}^{t-1})),\ \mathrm{conv}\ \cap_{j=1}^n D_j) > \beta_1\} > \delta.$$

Lemma 5.1: For any $\beta_1 > 0$ there is $\beta_2 > 0$ such that if $\mathrm{dist}((p_1,\ldots,p_n),\ \mathrm{conv}\cap_{j=1}^n D_j) > \beta_1$ then there is a $\mathrm{dist}((p_1,\ldots,p_n),D_i) > \beta_2$ for at least one i $(1 \le i \le n)$.

Proof: Clear. //

Define now a partition of M into n subsets, $M_1M_2,\ldots,M_n$, as follows: $t \in M_j$ if and only if.

(i) $t \notin M_m$ for all $m < j$.

(ii)[8] $pr\,\{\bar{x}^{t-1}\,|\,\mathrm{dist}((\bar{f}^t(\bar{x}^{t-1}),D_j) > \beta_2\} \ge \delta/n$.

Lemma 5.2: $\{M_1,\ldots,M_n\}$ is a partition of M.

Proof: $M_j \cap M_i = \emptyset$ for all $j \neq i$, this is implied by (i). The fact that $\{M_1,\ldots,M_n\}$ is not a partition of M means that there is some $t \in M \setminus \cup M_j$, namely:

$$pr\,\{\bar{x}^{t-1}\,|\,\mathrm{dist}((\bar{f}^t(\bar{x}^{t-1}),\ D_j) > \beta_2\} < \delta/n. \tag{17}$$

for all $1 \le j \le n$. By Lemma 5.1

$$\{\bar{x}^{t-1}\,|\,\mathrm{dist}(\bar{f}^t(\bar{x}^{t-1}),\ \mathrm{conv}\ \cap D_j) > \beta_1\} \subseteq$$
$$\subseteq \cup_{j=1}^n \{\bar{x}^{t-1}\,|\,\mathrm{dist}(\bar{f}^t(\bar{x}^{t-1}),\ D_j) > \beta_2\}.$$

Thus, the probability of the first set is less than $(\delta/n)n = \delta$. This is in contradiction to the fact that $t \in M$. //

Lemma 5.3: If $M \subseteq \mathbb{N}$ is of positive lower density and $\{M_1,\ldots,M_n\}$ is a partition of M, then there is some M_j with a positive upper density.

Proof: $M \cap \{1,\ldots,m\} = \cup_{j=1}^n M_j \cap \{1,\ldots,m\}$ for all $m \in \mathbb{N}$.

Denote $\beta^m = |\,M \cap \{1,\ldots,m\}\,|$ and $\beta_j^m = |M_j \cap \{1,\ldots,m\}|$, $\beta^m = \Sigma_{j=1}^n \beta_j^m$. Now we have $\Sigma_{j=1}^n \limsup_m (\beta_j^m/m) \ge \limsup_m \Sigma_{j=1}^n (\beta_j^m/m) = \limsup_m (\beta^m/m) \ge \liminf_m (\beta^m/m) > 0$. //

---

7 Here also $\mathrm{dist}(\bar{p},A)$ where $\bar{p} \in X_{i=1}^n \Delta(\Sigma_i)$ and $A \subseteq X_{i=1}^n \Delta(\Sigma_i)$ is the distance induced by the L_1 norm.

8 $\bar{f}^t(\bar{x}^{t-1}) = (\bar{f}_1^t(\bar{x}_1^{t-1},\ldots,\bar{x}_n^{t-1}),\ldots,\bar{f}_n^t(\bar{x}_1^{t-1},\ldots,\bar{x}_n^{t-1}))$.

W.l.o.g., M_1 has a positive upper density, say, $\eta' > 0$. Define g_1, the deviation of player 1, as it was defined at Step 2 of Section 4, and denote $g = (g_1, f_2, ..., f_n)$. D_1 is closed. Thus, by a similar argument to that of Lemma 4.6, we can prove that there is a positive constant $\gamma > 0$ such that whenever, $\text{dist}((\bar{f}^t(\bar{x}^{t-1}), D_1) > \beta_2$ there is a mixed action $q_1(\bar{x}^{t-1}) \sim_1 \bar{f}_1^t(\bar{x}^{t-1})$, which satisfies

$$h_1(\bar{f}^t(\bar{x}^{t-1})) < h_1((q_1, \bar{f}_2^t, ..., \bar{f}_n^t)(\bar{x}^{t-1})) - \gamma.$$

The left side of the inequality is equal to $E_f(h_1(y^t)|\bar{x}^{t-1})$ because of the multilinearity of h and by Lemma 4.2(ii). We come to the conclusion that, by the deviation g_1, player 1 can profit at least by $\gamma > 0$ at each stage of the set M_1, which has an upper positive density η', with probability of at least δ/n. We already know that $a_1^t(f) \leq a_1^t(g)$ for every $t \notin M_1$. Thus,

$$\text{limsup}_t H_1^t(g_1, f_2, ..., f_n) > H_1^*(f) + (\delta/n)\gamma\eta'.$$

Step 3: IR $\cap$ conv $h(\cap_{j=1}^n D_j) \subseteq$ UEP.

In order to define an upper equilibrium f it is necessary to ensure that any player will be able to profit by a deviation only at finitely many stages. Thus the strategy f will be defined in such a way that all the times all the players will play their best response among their nondetectable actions (which are not necessarily their best response.) In other words, it is possible for a certain player to deviate and to gain (only at the long run). However, this deviation is detectable. A repeated deviation will be reflected (with high probability) in the frequency of the appearance of the various signals. So, the players have to check all the time the relative frequency of the various signals they previously got. In case where this relative frequency is far from the expected one, the players punish the player responsible for these "bad-behaving" signals (if there are several such players, punish the one whose index is smallest). However, even if all the players do not deviate and play according to the prescribed mixed strategies, there is a positive probability that the frequency of a signal will be "bad-behaving". For this reason the players cannot punish the deviator (or the player who is referred to as the deviator) from the deviation moment on forever. They have to punish the deviator for a while, and then return to the master plan.

Let $\alpha \in$ IR $\cap$ conv $h(\cap_{j=1}^n D_j)$. We will show that $\alpha \in$ UEP, by defining an upper equilibrium strategy f s.t. $H^*(f) = \alpha$. $\alpha \in$ conv $h(\cap_{j=1}^n D_j)$, so there are $n + 1$ strategies, $(q_1^m, ..., q_n^m) = q^m \in \cap_{j \neq 1}^n D_j$, $m = 1, ..., n + 1$, and $n + 1$ constants $\gamma^m > 0$, $(m = 1, ..., n + 1)$ with total sum 1 which satisfy:

$$\alpha = \Sigma_{m=1}^{n+1} \gamma^m h(q^m). \tag{18}$$

Divide $\mathbb{N}$ into blocks: $B_1, B_2, \ldots$, and divide each block B_k into $n+1$ parts: $B_k^1, \ldots, B_k^{n+1}$ s.t

$$|B_1| = 1. \tag{19a}$$

$$\text{Min } B_k = \text{Max } B_{k-1} + 1, \; |B_k| = k^{10}, \; k = 2,3,\ldots \tag{19b}$$

For every segment $S \subseteq B_k$ of length k (19c)

$$||B_k^m \cap S|/k - \gamma^m| < 1/k, \; m = 1,\ldots,n+1.$$

Unless he finds a deviation at the previous block, player i has to play at stage $t \in B_k^m$ the strategy q_i^m. If player i finds a deviation at block B_k, he has to punish at blocks $B_{k+1}, B_{k+2}, \ldots, B_{k^2}$ and from B_{k^2+1} he has to play again q_i^m, and so on.

How does a player recognize a deviation and who does he punish? Player i counts the number of times he got the signal $\bar{x}$ from player j on the part B_k^m of B_k. Denote this number of times by $OB_k^m(\bar{x}, j)$. Note that this number is common knowledge. Then he checks the relative frequency of $\bar{x}$ in B_k^m to see whether it is far from the expected number or not. Namely, whether

$$|OB_k^m(\bar{x}, j)/|B_k^m| - q_j^m(\bar{x})| > 1/2k \tag{20}$$

or not (recall Notation 2.9). In a case where player i finds that (20) holds for some $1 \le j \le n$ and $\bar{x} \in \bar{\Sigma}_j$, he comes to the conclusion that player j has deviated at block B_k. Player i will punish the player with the smallest index who has deviated at block B_k. Again, if j is the player with the smallest index who was found to be a deviator, then player i will play σ_i^j in the blocks $B_{k+1}, \ldots, B_{k^2}$, where σ_i^j is any action.

Lemma 5.4: $H^*(f) = \alpha$.

Proof: Let $t \in \mathbb{N}$, and denote $a^t(f)$ by a^t. At stage $t \in B_k^m$ either player i plays q_i^m or he punishes someone. The probability that player i will punish someone is the probability that he has found a deviation at one of the $[\sqrt{k}] + 1$ previous blocks.[9] For every $1 \le j \le n$ and $\bar{x} \in \Sigma_j$, $OB_k^m(\bar{x}, j)$ is common knowledge. So, by the definition of f, whenever one player punishes player i, all the players punish him as well. Furthermore, f is defined in such a way that whenever player j punishes someone at block B_k he does not check whether some other player defects in that block or not. For these reasons, coming to the conclusion that player j deviated at block B_k is equivalent to finding $\bar{x} \in \bar{\Sigma}_j$ and B_k^m such that (20) holds, while player j was actually playing q_j^m at B_k^m (and not σ_j^ℓ for some ℓ). The probability to find such a thing is, by the Chebyshev inequality.

9 It is convenient to define $[x] = \text{Max}\{n \in N | n < x\}$ for every $x \in R$.

$$pr\{|OB_k^m(\bar{x},j)/|B_k^m| - q_j^m(\bar{x})| > 1/2k\} < V_j^m(\bar{x})\,(2k)^2/|B_k^m| \tag{21}$$

where $V_j^m(\bar{x})$ is the corresponding variance. However, by (19b) and (19c), whenever k is big enough we have.

$$\frac{V_j^m(\bar{x})(2k)^2}{|B_k^m|} < \frac{V_j^m(\bar{x})(2k)^2}{(\gamma_k^m - 1/k)k^{10}} < \frac{V_j^m(\bar{x})}{\gamma_k^m\, k^6} < \frac{1}{k^5} \tag{22}$$

Hence, the probability of finding a deviation at block B_k is less than $\mathrm{Max}_j\,|\bar{\Sigma}_j|\,n(n+1)/k^5$ ($\mathrm{Max}_j\,|\bar{\Sigma}_j|\,n(n+1)$ stands for all the possibilities of $\bar{x}$, j and m) when k is big enough. Thus, the probability of finding a deviation at one of the blocks $B_{[\sqrt{k}]+1},\ldots,B_{k-1}$ is less than

$$\sum_{\ell=[\sqrt{k}]+1}^{k-1} \mathrm{Max}\,|\bar{\Sigma}_j|\,n(n+1)/\ell^5 < (k-1)/\sqrt{k}^5\,\mathrm{Max}\,|\bar{\Sigma}_j|\,n(n+1) < 1/k.$$

So, if $t \in B_k^m$ and t is big enough, then

$$\|a^t - h(q^m)\|_\infty < W/k. \tag{23}$$

(18), (19) and (23) give the desired result. //

Lemma 5.5: $f = (f_1,\ldots,f_n)$ is an upper equilibrium strategy.

Proof: We refer now to a_j^t, $1 \le j \le n$, $t \in \mathbb{N}$ as random variables. Let g_j be a mixed strategy of player j. $g = (f_{-j}, g_j)$ defines a measure μ on $X_{i=1}^n\, \Sigma_i^*$. We will show that

$$\mathrm{limsup}_t\, (a_j^1 + \ldots + a_j^t)/t \le \alpha_j\ \mu\text{-}a.s.,$$

and this implies

$$\mathrm{limsup}_t\, E_\mu(a_j^1 + \ldots + a_j^t)/t \le \alpha_j\,.$$

For this purpose we need the following probabilistic statement.

Lemma 5.6: let $R_1,\ldots,R_n$ be a sequence of identically distributed Bernoulli random variables, with parameter p. Let $Y_1,\ldots,Y_n$ be a sequence of Bernoulli random variables such that for each $1 \le m \le n$, R_m is independent of $R_1,\ldots,R_{m-1},Y_1,\ldots,Y_m$. Then

$$pr\{|\frac{R_1Y_1 + ... + R_nY_n}{n} - p\frac{Y_1 + ... + Y_n}{n}| \geq \epsilon\} \leq \frac{1}{n\epsilon^2} \tag{25}$$

for every $\epsilon > 0$.

Proof: Define $\mathscr{F}_m$ to be the field generated by $R_1,...,R_{m-1},Y_1,...,Y_m$ and define $Z_m = R_mY_m - pY_m$. $E(Z_m | \mathscr{F}_s) = 0$ a.s. for every $s < m$. Furthermore, if $s < t$ then

$$E(Z_sZ_t) = E(E(Z_sZ_t | \mathscr{F}_s)) = E(Z_sE(Z_t | \mathscr{F}_s)) = E(Z_s0) = 0.$$

Denote $S_n = \Sigma_{i=1}^n Z_i$. By Chebyshev inequality,

$$pr\{|S_n| \geq n\epsilon\} \leq E(S_n^2)/(n\epsilon)^2 = (\Sigma_{i=1}^n E(Z_i^2))/n^2\epsilon^2 \leq 1/n\epsilon^2. \quad //$$

Fix k and m $(1 \leq m \leq n + 1)$, and define for each $x \in \Sigma_j$, $w \in X_{i \neq j} \Sigma_i$ and $t \in B_k^m$, $R_t(w) = 1$ if each player i $(i \neq j)$ plays w_i at stage t and 0 otherwise. $Y_t(x) = 1$ if player j plays x at stage t and 0 otherwise. Define also $u(x) = (1/|B_k^m|) \Sigma_{t \in B_k^m} Y_t(x)$, $u(\bar{x}) = \Sigma_{x \in \bar{x}} u(x)$, $\hat{q}^m = (q_1^m,...,q_{j-1}^m, q_{j+1}^m,...,q_n^m)$, $Q(s) = \Pi_{i \neq j} q_i^m(w_i)$ = the probability that w will be acted. Let $a_j^t(w,x) = R_t(w)Y_t(x)h_j(w,x_j)$. Notice that $a_j^t = \Sigma_{w,x}\, a_j^t(w,x)$. By Lemma 5.6, with probability of at least $1 - (k^2/|B_k^m|)\, |X_{i \neq j} \Sigma_i|\, |\Sigma_j|$ the following holds,

$$(1/|B_k^m)\, \Sigma_{t \in B_k^m}\, a_j^t = 1/|B_k^m|\, \Sigma_{t \in B_k^m}\, \Sigma_x \Sigma_w\, a_j^t(w,x)$$

$$= \Sigma_x \Sigma_w\, (1/|B_k^m|)\, \Sigma_{t \in B_k^m}\, a_j^t(w,x) \leq \Sigma_{\bar{x}} \Sigma_{x \in \bar{x}} \Sigma_w [h_j(w,x)u(x)Q(w) + (1/k)W]$$

$$= \Sigma_{\bar{x}} \Sigma_{x \in \bar{x}} [h_j(\hat{q}^m,x)u(x) + |X_{i \neq j} \Sigma_i|\, (1/k)W]$$

$$\leq \Sigma_{\bar{x}} [(\max_{x \in \bar{x}} h_j(\hat{q}^m,x))u(\bar{x}) + |\bar{x}|\, |X_{i \neq j} \Sigma_i|\, (1/k)W] = (**).$$

In a case where player j will not be punished after B_k, namely, when $|u(\bar{x}) - q_j^m(\bar{x})| < 1/2k$, we get,

$$(**) \leq h_j(q^m) + 2|X_i \Sigma_i|\, (1/k)W = h_j(q^m) + c/k,$$

where c is a constant. The inequality holds because $q^m \in D_j$.

Denote by A_k^1 the event that $1/|B_k| \ \Sigma_{t \in B_k} a_j^t > \alpha_j + 2c/k$ will occur without player j being punished at $B_{k+1}, \ldots, B_{k^2}$. Because of (19c), whenever k is big enough, we get

$$pr(A_k^1) \leq \Sigma_{m=1}^{n+1} k^2/|B_k^m| \ |X_i \Sigma_i| \leq k^2 \bar{c}/|B_k^m|, \tag{26}$$

for a certain constant $\bar{c}$. By (19b), $\Sigma_{k=1}^{\infty} k^2 \bar{c}/|B_k^m| < \infty$. By the Borel-Cantelli lemma ([*B*], p. 412), the probability that A_k^1 will occur for infinitely many k's is zero.

Fix an $\eta > 0$. We claim that the probability that $\limsup_T 1/T \, \Sigma_{t=1}^T a_j^t > \alpha_j + \eta$ is included in the event $\{A^k$ occurs infinitely often$\}$, where A_k is the event $1/|B_k| \ \Sigma_{t \in B_k} a_j^t > \alpha_j + \eta/2$, and the average payoff of player j at $B_k \cup \ldots \cup B_{k^2}$ is also greater than $\eta/2$.

Define A_k^2 to be the event where player j is punished after B_k and the average of his payoffs at $B_{k+1} \cup \ldots \cup B_{k^2}$ is greater than $\alpha_j + \eta/2$. The event $\{A_k$ infinitely often$\}$ is included in the union of $\{A_k^1$ infinitely often$\}$ and $\{A_k^2$ infinitely often$\}$. As was shown before, the first event has probability zero and the second one, by similar arguments, has also probability zero.

To recapitulate, $\limsup_T 1/T \, \Sigma_{t=1}^T a_j^t > \alpha_j + \eta$ will occur with probability zero for any $\eta > 0$. This finishes the proof of Step 3. //

Remark 5.6: We actually proved more than what is required by the definitions; we proved a pointwise version. For the joint strategy f in Step 3, the payoff for player i is *almost surely* α_i. Moreover, any deviation will lead *almost surely* to an average payoff that is less than α_i.

Remark 5.7: The method of Step 3 can be generalized to any information structure. Define D to be the set of all the joint mixed actions, p, which satisfy:

(i) if player i has a profitable deviation p_i', then *all* other players can detect it. Player j can detect the deviation if he has an action a (not necessarily in the support of p_j) s.t. $((p_{-i}, p_i')_{-j}, a)$ and (p_{-j}, a) induce two different distributions over the signals of player j.

(ii) Two players do not have deviations (from p) that affect in the same way the signals' distributions.

6 Concluding Remarks

6.1 The Banach Equilibrium

The liminf and the limsup are two ways of evaluating a sequence of average payoffs that does not converge. Using these evaluations we have two notions of equilibria. We could take another approach of evaluating nonconverging sequences. A Banach limit, L, is a linear functional defined on bounded sequences and gives any converging sequence its limit as a value. In other words, the Banach limit is an extension of the usual limit to all the bounded sequences. There are many Banach limits and for each Banach limit L we can define the notion of L-equilibrium.

Definition 6.1: $(f_1,\dots,f_n)$ is an L-equilibrium if for any player i and a strategy $\overline{f}_i$

$$L(H_i^t(f_1,\dots,f_n))_t \geq L(H_i^t(f_1,\dots,\overline{f}_i,\dots,f_n))_t.$$

For any L we get the set of payoffs that are associated with L-equilibria. Denote this set by EP_L. By going along the outline of the proof of the previous section we get as a result that

$$\mathrm{UEP} = \mathrm{EP}_L \text{ for all } L.$$

(We parenthetically remark that the previous statement does not imply UEP = LEP.) From this point of view the upper equilibrium is more appealing than the lower equilibrium.

6.2 Uniform Equilibrium

We will introduce here another equilibrium notion which connects the long run game with the finitely repeated games, that is, the uniform equilibrium (see [S2]).

Definition 6.2: $f = (f_1,\dots,f_n)$ is a uniform equilibrium if

(i) $H^*(f)$ is well defined and,

(ii) for every $\epsilon > 0$ there exists N such that for all i and $g_i \in \Delta(\Sigma_i^*)$ if $n > N$ then $H_i^n(f) \geq H_i^n(f_{-i}, g_i) - \epsilon$.

The connection to finite games is suggested by the following (Proposition 3.2 in [S2]): f is a uniform equilibrium if $H^*(f)$ is well defined and if there exist a sequence ϵ_m decreasing to 0, and a sequence N_m such that f induces an ϵ_m-equilibrium in the N_m-fold repeated game.

We claim that the set of all the uniform equilibrium payoffs coincides with UEP. Step 2 in Section 5 proves that the set of all uniform equilibrium payoffs is contained in conv $h(\cap D_j)$.

In order to show that the strategy f, defined in Step 3 of that section, is also a uniform equilibrium, notice the two following facts:

(i) The length of any block relative to the length of its predecessors goes to zero;

(ii) $pr(A_k^1)$, the probability of earning something tangible, $(2c/k)$, at block B_k without being punished afterwards is less than $k^2\bar{c}/|B_k|$ (see (26)).

6.3 Stochastic Payoff

One can define a modification of our model in which the signal is deterministic but the payoffs are random. Providing that the signal is the same, the stochastic payoff can be replaced by its expectation and then all the results go through.

6.4 Other Related Results

The semi-standard information case is the only n-player case in which there is a full characterization of UEP. The semi-standard information has two characteristics which make the characterization possible:

(i) There is a common signal which is a function of the joint action.

(ii) By knowing his own action, a player cannot extract any further information about other players' actions. Formally,

if $\lambda_i(a_{-i}, a_i) = \lambda_i(a'_{-i}, a_i) = \lambda_i(a_{-i}, a'_i)$ then

$\lambda_i(a'_{-i}, a'_i) = \lambda_i(a_{-i}, a'_i)$.

These two features enable a player to compute the expected strategies of his opponent and, furthermore, no correlation can emerge from the histories. Under a general information case it could be the case where even LEP is not contained in IR. It happens because some players can use their private information as a correlation device and push the payoff of one of the others down below his minmax level.

Another point that differentiates between semi-standard information and a general information structure is that in the former, if one player discovers a deviation, all other players discover it as well, while in the latter case, this is not so. This means that a player is not required to transmit his private information about the alleged deviation to others in order to cooperate with them in the punishment. If such a transmission is necessary in a general information structure, it may give room for some players to transmit false messages about deviations and thereby to gain by other players' punishments.

In two player games with observable payoffs it is known (see [L2]) that UEP = LEP and it is characterized completely by using a similar formula of the payoffs set to that which appeared above, exchanging D_i with another set.

The characterization of LEP is easier. It is characterized in all the two-player repeated games with nonobservable actions (see [L1]). A recent result is a characterization of the correlated equilibrium (lower and upper) payoffs in all the two-player games [L3].

6.5 Open Problems

1. Find a characterization of UEP in a general two-player repeated game with nonobservable actions.
2. Give another nontrivial condition on the information functions that enable to characterize completely UEP in n-player games.

Acknowledgement: I am grateful to R. J. Aumann for his supervision, and to A. Neyman for helpful comments. I want also to acknowledge F. Forges, E. Kalai, J. F. Mertens, and S. Sorin.

References

[APS] Abreu D, Pearce D, and Stachetti E (1986) Optimal Cartel Equilibria with Imperfect Monitoring, Journal of Economic Theory (forthcoming)

[A1] Aumann RJ (1964) Mixed and Behavior Strategies in Infinite Extensive Games, in Advances in Game Theory, Ann. Math. Studies, 52, Dresher M et al. (eds.). Princeton, N.J.: Princeton University Press, 627-650

[A2] Aumann RJ (1976) Lecture Notes in Game Theory, Stanford University

[A3] Aumann RJ (1981) Survey of Repeated Games, in Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern, Bibliographishes Institute, Mannheim/Wein/Zurich, 11-92

[B] Breiman L (1968) Probability, Addison-Wesley Publishing Co.

[FM1] Fudenberg D, Maskin E (1986) Discounted Repeated Games with Unobservable Actions, I: One-Sided Moral Hazard, mimeo

[FM2] Fudenberg D, Maskin E (1986) Discounted Repeated Games with Unobservable Actions, II: Two-Sided Moral Hazard, mimeo

[GP] Green E, Porter R (1984) Noncooperative Collusion Under Imperfect Price Information, Econometrica, 52, 975-994

[H] Hart S (1985) Nonzero-Sum Two-Person Repeated Games with Incomplete Information, Mathematics of Operations Research, 10, 117-153

[K] Kuhn HW (1953) Extensive Games and the Problem of Information, in Contributions to the Theory of Games II, Ann. Math. Studies, 28, Kuhn HW, Tucker AW (eds.), Princeton, N.J.: Princeton University Press, 193-216

[L1] Lehrer E (1986) Lower Equilibrium Payoffs in Repeated Games with Non-Observable Actions, International Journal of Game Theory, 18, 57-89

[L2] Lehrer E (1986) Two Player Repeated Games with Non-Observable Actions and Observable Payoffs, Center for Research in Mathematical Economics and Game Theory (R. M. 76), The Hebrew University, Jerusalem

[L3] Lehrer E (1989) Correlated Equilibria in Two Player Repeated Games with Non-observable Actions, revised version, Northwestern University Working Paper

[R1] Radner R (1981) Monitoring Cooperative Agreements in Repeated Principal-Agent Relationship, Econometrica, 49, 1127-1148

[R2] Radner R (1986) Repeated Partnership Games with Imperfect Monitoring and No Discounting, Review of Economic Studies, 53 (1), 43-58

[RY] Rubinstein A, Yaari M (1983) Repeated Insurance Contracts and Moral Hazard, Journal of Economic Theory, 30, 74-97

[S1] Sorin S (1983) Some Results on the Existence of Nash Equilibria for Non-Zero Sum Games with Incomplete Information, International Journal of Game Theory, 12 (4), 193-205

[S2] Sorin S (1988) Repeated Games with Complete Information, CORE Discussion Paper No. 8822

Received July 1988
Revised version September 1989
Final version February 1990