Games and Economic Behavior **29,** 274–308 (1999)
Article ID game.1999.0722, available online at http://www.idealibrary.com on IDEAL®

Merging, Reputation, and Repeated Games with Incomplete Information

Sylvain Sorin*

Laboratoire d'Econométrie, Ecole Polytechnique, 1 rue Descartes, 75005 Paris, France, and MODAL'X and THEMA, U.F.R. SEGMI, Université Paris X Nanterre, 200 Avenue de la République, 92001 Nanterre Cedex, France
E-mail: sorin@poly.polytechnique.fr

Received September 8, 1997

We relate and unify several results that appeared in the following domains: merging of probabilities, perturbed games and reputation phenomena, and repeated games with incomplete information. *Journal of Economic Literature* Classification Numbers: C72, D83.

1. PRESENTATION

Consider a discrete stochastic process on a space X that follows a probability distribution Q. The paths of the form $x = x_1, \ldots, x_n, \ldots$ are announced stage after stage to an observer who ignores Q but holds an *a priori* probability or belief, P, on the process. On each path x, at each stage n, the probability of the observer on the future behavior of the process is revised, giving rise to a new belief which is the conditional distribution of P given $(x_1, \ldots, x_n)$. On the other hand, Q also determines conditional probabilities.

We first consider conditions for merging, namely, convergence of these beliefs (deduced from P) to the corresponding conditional probabilities (given Q). The initial concept of merging, due to Blackwell and Dubins, requires the two probabilities to be close to each other for all future events. This notion and the corresponding Blackwell and Dubins theorem (Result I) deal with asymptotic properties. They will be used in the framework of multistage games with undiscounted payoffs. On the other hand, in the discounted case only events in the near future matter. The corresponding notion of proximity for probabilities considers only these events. This leads

*I thank Abraham Neyman for extensive comments.

0899-8256/99 $30.00

to weak merging introduced by Kalai and Lehrer. However for the applications to discounted games a stronger property involving uniform rate of weak merging is needed. This corresponds to Result II. These two results (I and II) are not comparable: the first one says that two sequences of probabilities are asymptotically closed on a large set of events, the second one that they are closed on a smaller set of events but at a uniform rate.

The above results on merging directly apply to the study of multistage two-person games of incomplete information on one side, where uncertainty concerns either the payoffs (like in repeated games with incomplete information) or the strategies (like in perturbed games and reputation phenomena). In both models the informed Player 1 is one of several types k and her opponent holds an initial probability p on this set of types K. We study the set of equilibrium payoffs in such games and also compute bounds on the advantage the informed player can achieve due to her opponent's uncertainty. We show that these properties crucially depend on the relative patience of the players and on the specification of the signalling mechanism used during the play.

The argument goes as follows: consider equilibrium strategies for the players. In the standard signalling case the stochastic process $(x_1, \ldots, x_n, \ldots)$ is simply the play $((i_1, j_1), \ldots, (i_n, j_n), \ldots)$, namely, the sequence of moves used and observed stage after stage by the players. The belief P of the uninformed Player 2 on the plays of the game is generated by her own strategy, the collection of type-dependent strategies of her opponent and the initial distribution p on types. On the other hand, for each type k of Player 1, the (true) distribution Q^k of the process is generated by her strategy and the strategy of the uninformed agent. This framework provides specific relations between distributions P and Q^k which are sufficient for merging to occur. It follows that from the point of view of the uninformed player, P and Q^k induce conditional probabilities on the future plays that are close to each other from some stage on. By the equilibrium condition, Player 2 is playing a best reply. We thus derive a bound on her payoff under the distribution P, hence by merging similar results under Q^k. This provides conditions satified by the probability Q^k, thus finally a bound on the informed player's payoff while using her strategy generating Q^k (under the given strategy of Player 2).

This general methodology is first applied to two-person repeated games with incomplete information on one side and known own payoffs. It gives a short proof of the characterization of the set of equilibrium payoffs in the undiscounted case, using Result I. The same characterization is obtained, using the parallel Result II, in the discounted framework where the informed Player 1 is infinitely more patient than her opponent. This condition on the time preferences of the agents plays a crucial role and is actually necessary.

The characterization shows moreover that the payoff set for the informed player depends only on the support of the distribution p on types. Hence it allows for an interpretation in terms of perturbation when most of the probability is on the "true" type. A quantity $m(A, B)$, is explicitly computed. It corresponds to the best lower bound on equilibrium payoffs for Player 1 in the following class of games: the perturbation is in terms of Player 1's payoffs, the true payoffs for Players 1 and 2 being defined by the matrices A and B, respectively .

We then consider reputation phenomena or incomplete information in terms of strategies. Here the strategy the uninformed player is facing is a mixture of the actual strategy of the informed Player and of some perturbation. Each such reputation model is identified by three parameters: one concerns Player 1 and is the nature of the perturbations, the other is the type of payoffs of the players (basically their relative duration), the third one is the information transmitted during the play.

When playing against a sequence of Players 2, each one living for one stage, Player 1 can build a quite strong reputation. In fact, for the uninformed Player 2, the knowledge of the probability induced by P (hence by Q^k) on the next stage implies the knowledge of the one-stage strategy (of the informed player) she is facing. Since Player 2 is playing a best reply it is simple for Player 1 to give intructions to her and a bound $w(A, B)$, better than the one obtained in the previous framework $m(A, B)$, is actually reached. On the other hand when both players live forever with undiscounted or discounted payoffs, but again in the latter case the informed player being much more patient than her opponent, the bound obtained with the same merging tools, is lower and coincides with $m(A, B)$.

Facing longer lived opponents may be worse for the informed player. Even after obtaining, by merging, a good approximation of Q^k from P, Player 2 may be afraid of playing a best reply to the true and unknown strategy of Player 1: by experimenting she could be punished and then obtain a worse payoff. This phenomena, reminiscent of the concept of conjectural equilibria, explains why reputation effects are not monotonic with respect to the length of the interaction.

However if the uninformed player is using a completely mixed strategy, she generates against any perturbation of Player 1 a "revealing distribution." In this case, for Player 2, playing a best reply under P essentially implies playing a true best reply. It follows that by adding some noise to the model, so that the distribution on the signals to Player 1 is completely mixed, one can extend the precise monitoring results of the short lived situation to the case where Player 2 lives for n periods. The stochastic process is now on Player 2's signals and the weak merging refers to the next n stages. The game looks like a repeated version of the normalized n stage games with standard signalling.

In addition one can, in this situation, take advantage of the length of Player 2's life to build more elaborate perturbations in terms of monitoring. This leads to the best achievable payoff from the point of view of Player 1, denoted by $s(A, B)$.

Another approach to avoid the "conjectural effect" is to introduce more complex perturbations, such that there is an infinite hierarchy of punishments on plays compatible with the same perturbation. Rather than playing an active role, the uninformed player, trying to avoid punishment, is finally confronted with the perturbation and the merging property implies that she eventually identifies it.

This paper is basically an original presentation of a large number of results that appeared in different articles. It shows that a few basic properties in the theory of merging underline most of the proofs. The advantages of this approach are in several directions.

Concerning the proofs of the results, their basic structure is described explicitly: in order to get properties on the payoff of Player 1 under Q^k, one studies the payoff of Player 2 under P and then one uses the merging property. This allows also to shorten drastically the proofs (especially Proposition 2.5, and those in Sections 3 and 4). Finally it enables to extend the results: for example, to a sequence of finite lived Player 2 (in Sections 3, 4, and 6) or to perturbation in mixed strategies (Section 7).

Also this methodology allows to replace a series of specific martingale convergence results—that seem to be adapted to each particular case—by a general property independent on the probability space and much easier to use: see the change of filtrations, deterministic in 3.4 and random in Section 7. It also helps in understanding the importance of some hypotheses: for example, the need for the informed player to be much more patient than the other or the impact of the signalling structure on merging.

Finally this unified presentation reveals hidden connections: on one hand between incomplete information games with known payoffs and reputation phenomena, see the reasons for getting the same bound $m(A, B)$ in Sections 3 and 4; on the other hand between the undiscounted model and the one with a patient informed player, see 3.3 versus 3.4 and 4.4 versus 4.5. It also facilitates the comparison of reputation properties when facing a sequence of short or long lived players and clarifies the relations between the different lower bounds on the equilibrium payoffs of the informed player: $m(A, B)$, $w(A, B)$, and $s(A, B)$.

This methodology has the large potential of applications since it relies on a general bound on the number of bad observations in a merging framework. This property is clearly independent of the length of the players, of the type of signals, and of the nature of the perturbations.

The content of the paper is as follows:

Part 2 recalls first (Result I) the initial and fundamental result on merging of Blackwell and Dubins (1962), the alternative formulation (weak merging) and refinements of Kalai and Lehrer (1994). Then (Result II) a related uniform bound on the speed of convergence is given with a new proof of the property of Fudenberg and Levine (1992) in Proposition 2.5. A new result analog to theirs is also obtained as Proposition 2.4.

Part 3 builds on the above properties to characterize the set of equilibrium payoffs in two-person repeated games with incomplete information and known own payoffs. Result I is used for the undiscounted case (3.2) in the spirit of Shalev (1994) and Koren (1992) and one relies on Result II for the discounted case with a more patient informed Player in 3.4, following Cripps and Thomas (1995b). The corresponding lower bound on the payoff of the informed player for payoff-perturbed games, due to Shalev (1994) and Israeli (1996) is presented in 3.3.

The remainder of the paper deals with reputation phenomena. The lower bounds on the informed player's payoff at equilibrium depend upon the relative sizes of the players, the observability conditions, and the nature of the perturbation.

Part 4 is devoted to the standard signalling case with different sizes of players, starting with the seminal works of Fudenberg and Levine (1989, 1992), then following Cripps and Thomas (1995a) and Cripps *et al.* (1996) using Result I or II.

Part 5 deals with the case of a repeated game with signals, according to Fudenberg and Levine (1992) and studies the relation with conjectural equilibrium.

Part 6 considers games with noise where the whole perturbed strategy can be identified by the uninformed player hence allowing more precise monitoring, see Aoyagi (1996), Celentani *et al.* (1996).

Part 7 shows, following Evans and Thomas (1997), that with more complex perturbations (with unbounded memory) the informed player can force her opponent to follow any prespecified individually rational path, thus achieving the best bound $s(A, B)$.

2. NOTIONS OF MERGING

Let $(\Omega, \mathscr{F}, Q)$ be a probability space equipped with a filtration $\{\mathscr{F}_n\}$ (increasing sequence of sub σ-fields of $\mathscr{F}$) that generates $\mathscr{F} : \mathscr{F} = \sigma(\bigcup_n \mathscr{F}_n)$. We assume that each σ-field $\mathscr{F}_n$ is generated by a countable partition $\mathscr{F}_n^*$ and we denote by $F_n(w)$ the atom of $\mathscr{F}_n^*$ containing ω.

Let P be another probability distribution on $(\Omega, \mathscr{F})$. Q defines the law of the process $(x_1, \ldots, x_n, \ldots)$ with $x_n = F_n(\omega)$ (note that here x_n determines $(x_1, \ldots, x_n)$) and P corresponds to the beliefs of an observer. $\mathscr{F}_n$ describes the information available at stage n: given ω, the future behavior is governed by $Q(\cdot|F_n(\omega)) \equiv Q(.|\mathscr{F}_n)(\omega)$ and the beliefs are given by $P(\cdot|F_n(\omega)) \equiv P(.|\mathscr{F}_n)(\omega)$.

2.1. *Merging*

A notion of merging of P to Q refers to a type of convergence, under the true distribution Q, of the beliefs $P(\cdot|F_n(\omega))$ to the conditional probability distributions $Q(\cdot|F_n(\omega))$. It corresponds to the following situation: observing stage after stage the realization of the process, the observer eventually is able to make good predictions. The first approach and basic properties are due to Blackwell and Dubins (1962).

Let us first define:

$$f_n(P, Q)(\omega) = \sup_{A \in \mathcal{F}} |P(A|F_n(\omega)) - Q(A|F_n(\omega))|.$$

DEFINITION 2.1. *P merges* to Q if:

$$f_n(P, Q)(\omega) \longrightarrow 0 \quad \text{as } n \longrightarrow \infty \ Q \text{ a.s.}$$

The main result is

THEOREM 2.2 (Blackwell and Dubins, 1962). *Assume that Q is absolutely continuous with respect to P ($P \gg Q$). Then P merges to Q.*

Comments. **1.** Kalai and Lehrer (1994) have shown conversely that if P merges to Q for all filtrations $\mathcal{F}_n$ that generate $\mathcal{F}$, then $P \gg Q$. It follows that the notion of merging is actually independent of the filtration. It depends only on the data $(\Omega, \mathcal{F}, P, Q)$: P merges to Q iff $P \gg Q$.

2. A simple case where $P \gg Q$ holds is obtained when $P = \rho\, Q + (1 - \rho)\, Q'$, for some $0 < \rho \leq 1$, where Q' is another probability distribution on $(\Omega, \mathcal{F})$. In this case one says that the belief P contains a *grain of truth of size* ρ about Q (Kalai and Lehrer, 1994).

2.2. *Weak Merging*

Kalai and Lehrer (1994) also introduced a weaker notion of convergence based on the following function:

$$e_n(P, Q)(\omega) = \sup_{A \in \mathcal{F}_{n+1}} |P(A|F_n(\omega)) - Q(A|F_n(\omega))|.$$

DEFINITION 2.3. *P weakly merges* to Q along $\{\mathcal{F}_n\}$ if:

$$e_n(P, Q)(\omega) \longrightarrow 0 \quad \text{as } n \longrightarrow \infty \ Q \text{ a.s.}$$

This requires, with Q probability 1, the conditional distributions of P and Q at stage n to be close, for n large enough and for all events that occur at the next stage. Obviously this property extends to all events in a bounded future but not to all events in $\mathcal{F}$ like with $f_n(P, Q)$ in the definition of merging.

Kalai and Lehrer (1994) have shown that weak merging depends on the filtration $\{\mathcal{F}_n\}$. Necessary or sufficient conditions for weak merging (in particular weaker than absolute continuity) and refinements can be found in Kalai and Lehrer (1994), Lehrer and Smorodinsky (1996a, 1996b, 1997).

2.3. *Uniform Weak Merging*

The following notion is not involving tail events, like in the case of merging, but is stronger than weak merging. It is needed for the study of the discounted framework.

Consider the space of the process as an oriented graph with nodes (ω, n) in $\Omega \times \mathbb{N}$ corresponding to $F_n(\omega)$. Given δ positive, let A^δ denote the nodes where P and Q are δ weakly far away. Explicitly:

$$A^\delta = \{(w, n) \in \Omega \times \mathbb{N}; e_n(P, Q)(\omega) \geq \delta\}.$$

The following two lemmas control, under a grain of truth hypothesis, the size of the set A^δ.

The first lemma considers A_n^δ the section of A^δ for a fixed n; i.e., the set of paths ω for which the event: "the one stage prediction at stage n is inaccurate by more than δ" holds.

The second one deals, on each trajectory ω, with $A^\delta(\omega)$ which is the set of stages n where this event occurs.

LEMMA 2.4. *Given any positive constants δ, ε, and ρ^*, there exists $M = M(\delta, \varepsilon, \rho^*)$ such that for any P, Q with $P = \rho Q + (1-\rho)Q'$ and $\rho \geq \rho^*$, one has for all stages n except at most M of them*:

$$Q(A_n^\delta) \leq \varepsilon.$$

Proof. We consider P as a probability distribution on an extended probability space, where first a lottery chooses between Q and Q' and then the process follows the selected probability. Let $\mathscr{B}$ be the event: "the process follows the distibution Q." Under P, $\mathscr{B}$ has an initial probability $\rho = q_1$. We introduce $q_m = P(\mathscr{B}|\mathscr{F}_{m-1})$ which defines, under P, a martingale $q = \{q_m\}$ with values in $[0, 1]$, hence satisfies:

$$E_P\left(\sum_{m=1}^{\infty} E_P((q_{m+1} - q_m)^2|\mathscr{F}_{m-1})\right) = E_P\left(\sum_{m=1}^{\infty}(q_{m+1} - q_m)^2\right) \leq 1. \qquad (1)$$

We obtain a bound of the probability of A_n^δ by studying the variation of q.

Let $\mathscr{M}$ denote the set of stages where $E_P((q_{m+1} - q_m)^2) \geq 1/M$. By (1), its cardinality satisfies $\#\mathscr{M} \leq M$.

Now for a stage $m \notin \mathscr{M}$, the bound $E_P(E_P((q_{m+1} - q_m)^2|\mathscr{F}_{m-1})) \leq 1/M$ implies that $P(\omega; E_P((q_{m+1} - q_m)^2|\ F_{m-1}(\omega)) \geq 1/\sqrt{M}) \leq 1/\sqrt{M}$.

Define the conditional "one-stage variation" of the martingale q by:

$$V_m(q)(\omega) = E_P(|q_{m+1} - q_m|\ |F_{m-1}(\omega)).$$

Hence for $m \notin \mathscr{M}$, one obtains from the previous inequality and Cauchy Schwarz inequality: $P(\omega; V_m(q)(\omega) \geq 1/\sqrt[4]{M}) \leq 1/\sqrt{M}$ as well. Also, denoting by $\mathscr{D}_m(\omega)$ the family of atoms of the partition of $F_{m-1}(\omega)$, one has an explicit formula for the variation:

$$V_m(q)(\omega) = \sum_{a \in \mathcal{D}_m(\omega)} P(a|F_{m-1}(\omega))|q_{m+1}(a) - q_m(\omega)|.$$

Now for each $a \in \mathcal{D}_m(\omega)$ one has

$$q_{m+1}(a) = \frac{q_m(\omega)Q(a|F_{m-1}(\omega))}{P(a|F_{m-1}(\omega))},$$

so that one obtains:

$$V_m(q)(\omega) = q_m(\omega)e_m(P, Q)(\omega). \tag{2}$$

The distance between the conditional probabilities $e_m(P, Q)$ is thus proportional to the one-stage variation $V_m(q)$. It remains to bound the coefficient q_m which is the conditional probability of $\mathcal{B}$.

Let $\Omega_0^m(t) = \{\omega \in \Omega;\ q_m(\omega) \leq t\rho\}$ be the set of paths where this posterior probability of $\mathcal{B}$ has decreased by a ratio at least t. For $\omega \notin \Omega_0^m(t)$ the bound on e_m follows from (2):

$$Q(\omega \notin \Omega_0^m(t);\ e_m(P, Q)(\omega) \geq 1/t\rho\sqrt[4]{M}) \leq 1/\rho\sqrt{M}.$$

Finally on $\Omega_0^m(t)$, $\rho Q(F_{m-1}(\omega))/P(F_{m-1}(\omega)) = q_m(\omega) \leq t\rho$. Thus $Q(F_{m-1}(\omega)) \leq t\, P(F_{m-1}(\omega))$. By summing one obtains the following bound on the probability under Q of this set of "low posteriors": $Q(\Omega_0^m(t)) \leq t$.

Choose $t = \varepsilon/2$ and M such that both $\sqrt{M} \geq 2/\varepsilon\rho^*$ and $\sqrt[4]{M} \geq 1/t\delta\rho^*$ to get the result. ∎

The previous lemma was computing for a given stage the probability of the set of paths going through A^δ at that stage. This probability can be reduced as small as we want except on a uniformly finite number of stages. By uniform we mean that it does not depend on the space Ω and not on the probabilities P and Q under consideration but only on the bound ρ^* on the size of the grain of truth.

The next lemma checks, on each path, the number of nodes in A^δ. Except on a set of paths of probability as small as wanted this number is uniformly bounded.

LEMMA 2.5 (Fudenberg and Levine, 1992). *Given any positive constants δ, ε, and ρ^*, there exists $M = M(\delta, \varepsilon, \rho^*)$ such that for any P, Q with $P = \rho Q + (1-\rho)Q'$ and $\rho \geq \rho^*$, one has*:

$$Q(\omega;\ \# A^\delta(\omega) \geq M) \leq \varepsilon.$$

Proof. We use the notations and results from the proof of the previous lemma.

Let

$$\Omega_1(M, \alpha) = \{w \in \Omega;\ \#\{m;\ V_m(q)(\omega) \geq \alpha\} \geq M\}$$

be the set of paths ω where the martingale has a conditional one-stage variation greater than α on more than M stages. From (1) one obtains by Cauchy–Schwarz inequality: $\alpha^2 M\ \ P(\Omega_1(M, \alpha)) \leq 1$. Using (2), it follows that, except on $\Omega_1(M, \alpha)$, there are at most M stages where $e_m(P, Q)(\omega) \geq \alpha / q_m(\omega)$.

Let us consider the set of paths exhibiting eventually low posterior distributions:

$$\Omega_2(t) = \{\omega \in \Omega;\ q_n(\omega) \leq t\rho,\ \text{ for some } n\}.$$

One introduces the partition: $\Omega_2(t) = \bigcup_m \Omega_2^m(t)$ with $\Omega_2^m(t) = \{\omega \in \Omega;\ q_n(\omega) > t\rho,\ \text{ for } n < m, \text{ and } q_m(\omega) \leq t\rho\}$. For ω in $\Omega_2^m(t)$, $q_m(\omega) \leq t\rho$, hence as in the previous proof $Q(F_{m-1}(\omega)) \leq tP(F_{m-1}(\omega))$. So that $Q(\Omega_2^m(t)) \leq tP(\Omega_2^m(t))$ and also $Q(\Omega_2(t)) \leq tP(\Omega_2(t)) \leq t$.

Given a point $\omega \notin \Omega_2(t)$, $q_m(\omega) \geq t\rho$. If in addition $\omega \notin \Omega_1(M, \alpha)$, there are at most M stages n where:

$$e_m(P, Q)(\ \ )\omega \geq \frac{\alpha}{t\rho}.$$

Therefore, except on the set $\Omega_1(M, \alpha) \cup W_2(t)$, $\omega \in A_n^{\alpha/t\rho}$ for at most M stages. Choose $t = \varepsilon/2$, then $\alpha = t\rho^* \delta$ and finally $M \geq 2/(\alpha^2 \varepsilon \rho^*)$ to get the result. ■

Comments. **1.** Note that the previous results are independent of the filtration $\mathcal{F}_n$. They give, under a strong form of absolute continuity (P with grain of Q at least ρ^*), a uniform bound on the set of nodes (ω, n) where $e_n(P, Q)(\omega)$ is large. In other words, the number of observations that may lead to bad predictions is uniformly bounded.

2. To have a control on this number of stages is crucial in a discounted framework, by opposition to the undiscounted case where asymptotic results (like upper density) suffice.

3. REPEATED GAMES WITH INCOMPLETE INFORMATION

3.1. *Presentation*

Consider repeated two-person games with incomplete information on one side introduced by Aumann and Maschler (1995) and described as follows: Player 1's (resp., Player 2's) stage payoff is defined by a finite $I \times J$ matrix A^k (resp., B^k), where k belongs to a finite set K. The game is played in stages: at stage 0, k is selected according to a probability p, $p \in \Delta(K)$, the simplex on K. Player 1 is told the k chosen while Player 2 knows only p. At stage $n+1$, given the previous history of length n, $h_n \in H_n$, which is the sequence of moves $h_n = (i_1, j_1, \ldots, i_n, j_n)$ up to that stage, both players choose a move, say i_{n+1} and j_{n+1}, this couple is announced to both and the stage payoff is $a_{n+1} = A^k_{i_{n+1}, j_{n+1}}$ for Player 1 and $b_{n+1} = B^k_{i_{n+1}, j_{n+1}}$ for Player 2. Note that this payoff is not announced. As usual, strategies, $\sigma = \{\sigma^k\}$ for Player 1 and τ for Player 2, can be represented as mappings from the set of histories $H = \bigcup_{n \geq 0} H_n$ to mixed moves, namely, $\Delta(I)$ and $\Delta(J)$. Σ and $\mathcal{T}$ denote the corresponding sets of strategies. On the set of plays $H_\infty = (I \times J)^\infty$, the σ-fields $\mathcal{H}_n$ are generated by H_n and generate $\mathcal{H}_\infty$. Together with p, σ, and τ define a probability distribution on $(K \times H_\infty, 2^K \otimes \mathcal{H}_\infty)$, hence on the stream of payoffs.

Let $\overline{a}_n(\sigma, \tau) = E_{p,\sigma,\tau}(\frac{1}{n}\sum_{m=1}^n a_m)$ (and similarly $\overline{b}_n(\sigma, \tau)$) denote the average expected payoff of Player 1 (resp., Player 2) and let $\overline{a}^k_n(\sigma^k, \tau) = E_{k,\sigma^k,\tau}(\frac{1}{n}\sum_{m=1}^n a_m)$ specify this payoff given k so that $\overline{a}_n(\sigma, \tau) = \sum_k p^k\ \overline{a}^k_n\ (\sigma^k, \tau)$. Similarly, for $0 < \lambda \leq 1$, $\overline{a}_\lambda(\sigma, \tau) = E_{p,\sigma,\tau}(\sum_{m=1}^\infty \lambda(1-\lambda)^{m-1} a_m)$ is the average discounted payoff of Player 1 and so on. $\Gamma_n(p)$ (resp., $\Gamma_\lambda(p)$, $\Gamma_\infty(p)$) denotes the n stage (resp., λ discounted, infinitely repeated) game. In this last case, the payoffs are not well defined, however one introduces a set of equilibrium payoffs using the following procedure:

DEFINITION 3.1 (Hart, 1985). Given a Banach limit $\mathcal{L}$, an $\mathcal{L}$ equilibrium (σ, τ) is an equilibrium in the game with payoff $\gamma_{\mathcal{L}}(\sigma, \tau) = \mathcal{L}(\overline{a}_n(\sigma, \tau), \overline{b}_n(\sigma, \tau))$ (or equivalently with vector payoff $\{\mathcal{L}(\overline{a}^k_n(\sigma^k, \tau))\}$ for the informed Player 1). $E(p)$ is the set of all $\mathcal{L}$ equilibrium payoffs obtained as $\mathcal{L}$ varies.

DEFINITION 3.2 (Sorin, 1990, 1992). A pair (α, β) in $\mathbb{R}^2$ is a *uniform equilibrium payoff* if for any positive ε, there exists a couple of strategies (σ, τ) inducing an ε-equilibrium in any sufficiently large game $\Gamma_n(p)$ with a payoff within ε of (α, β). Formally:

$$\forall \varepsilon > 0 \qquad \exists(\sigma, \tau), \exists N : \forall n \geq N \qquad \forall(\sigma', \tau')$$
$$\overline{a}_n(\sigma, \tau) + \varepsilon \geq \alpha \geq \overline{a}_n(\sigma', \tau) - \varepsilon$$
$$\overline{b}_n(\sigma, \tau) + \varepsilon \geq \beta \geq \overline{b}_n(\sigma, \tau') - \varepsilon$$

or equivalently, for $(\alpha, \beta) \in \mathbb{R}^{K+1}$

$$\overline{a}^k_n(\sigma^k, \tau) + \varepsilon \geq \alpha^k \geq \overline{a}^k_n(\sigma'^k, \tau) - \varepsilon \qquad \forall k$$
$$\overline{b}_n(\sigma, \tau) + \varepsilon \geq \beta \geq \overline{b}_n(\sigma, \tau') - \varepsilon.$$

Note that (σ, τ) is also a 2ε-equilibrium in any discounted game $\Gamma_\lambda(p)$ with small discount factor λ.

Hart (1985) gave a characterization of $E(p)$ proving in addition that any point in it is a uniform equilibrium payoff.

There are two important subclasses of these games with lack of information on one side:

— zero-sum games correspond to the case where $A^k = -B^k$, for all k;

— "known own payoffs" games are obtained when Player 2's payoff is independent of k, $B^k = B$.

Note that in the first case the uninformed Player 2 does not know her own payoff.

3.2. *Known Own Payoffs Games*

We consider here the second subclass for which an explicit characterization of $E(p)$ is much easier (Shalev, 1994). One can deduce it from the general characterization of Hart (1985), using properties of bimartingales due to Aumann and Hart (1986), or directly, even for the case of lack of information on both sides (but with known own payoffs), like in Koren (1992) (see also Forges, 1992).

We first introduce few notations. For M an $I \times J$ matrix, x in $\Delta(I)$, and y in $\Delta(J)$, xMy denotes $\sum_{i,j} x_i M_{i,j} y_j$, $\operatorname{val}_1 M = \max_x \min_y \; xMy$ and $\operatorname{val}_2 M = \max_y \min_x \; xMy$. For π in $\Delta(I \times J)$, one defines $\langle \pi, M \rangle = \sum_{i,j} \pi_{i,j} M_{i,j}$. We also use $\langle \; , \; \rangle$ for the scalar product in $\mathbb{R}^K$. $A(p)$ is the average matrix $\sum_k p^k A^k$. $p \gg 0$ means that p has full support. For all x in $\Delta(I)$, let x^* denote the strategy in the repeated game that plays x i.i.d.

We assume $\# I \geq 2$. To make computations simpler we assume that all matrices are of norm less than 1: $\|A^k\| \leq 1$, $\|B\| \leq 1$.

We recall now basic results from *approachability theory*, due to Blackwell (1956).

DEFINITION 3.3. Let $c \in \mathbb{R}^K$. An orthant $\mathcal{O}(c) = \{c\} - \mathbb{R}^K_+$ is *approachable* by Player 2 in the game with vector payoffs $\{A^k\}$ if: $\forall \varepsilon > 0, \exists \tau$ and $\exists N$ such that for $n \geq N$, $\overline{a}^k_n(\sigma^k, \tau) \leq c^k + \varepsilon$, for all σ^k and all k.

Blackwell's theorem (1956) implies

PROPOSITION 3.4. *A necessary and sufficient condition for $\mathcal{O}(c)$ to be approachable by Player 2 is*:

(1) $\operatorname{val}_1 A(q) \leq \langle c, q \rangle, \forall q \in \Delta(K)$

or equivalently $\mathcal{O}(c)$ being convex:

(2) $\forall x \in \Delta(I), \exists y \in \Delta(J)$ *with* $xA^k y \leq c^k, \forall k \in K$.

Let us finally define a set of correlated distributions on $I \times J$ by: $\Pi(K) = \{\pi = \{\pi^k\}; \pi^k \in \Delta(I \times J), k \in K$ with:

(i) $\sum_k \langle q^k A^k, \pi^k \rangle \geq \text{val}_1 A(q), \ \forall q \in \Delta(K)$

(ii) $\langle B, \pi^k \rangle \geq \text{val}_2 B, \ \forall k$

(iii) $\langle A^k, \pi^k \rangle \geq \langle A^k, \pi^{k'} \rangle, \ \forall k, \ k' \in K\}$

The set of payoffs induced by these distributions is: $\mathscr{E}(p) = \{(\alpha, \beta) \in \mathbb{R}^{K+1}; \ \exists \pi \in \Pi(K)$ with:

(iv) $\langle A^k, \pi^k \rangle = \alpha^k, \ \forall k \in K$

(v) $\langle B, \sum_k p^k \pi^k \rangle = \beta\}$.

The following characterization means that these distributions generate all equilibrium payoffs.

Proposition 3.5 (Shalev, 1994; Koren, 1992). *Assume* $p \gg 0$. *Hence,*

$$E(p) = \mathscr{E}(p).$$

Proof. (1) To prove that any payoff in $\mathscr{E}(p)$ is a uniform equilibrium payoff is easy. The equilibrium strategies are as follows: Player 1 announces her type k and then both players follow a prespecified play w^k in H_∞ on which the correlated frequency of the moves converges to π^k. The induced payoffs satisfy then (iv) and (v). Any deviation during w^k is detectable and can be punished by condition (i) (using Proposition 3.4) for Player 1 and (ii) for Player 2. Finally condition (iii) prevents any undetectable deviation of Player 1 (pretending being of type k' while being of type k) from being profitable.

(2) For the converse we now show that any $\mathscr{L}$ equilibrium payoff is in $\mathscr{E}(p)$. Hence let (σ, τ) be an $\mathscr{L}$ equilibrium with payoff $(\alpha, \beta) \in \mathbb{R}^{K+1}$.

Denote by $\mathbf{P}$ (resp., $\mathbf{P}^k$) the probability induced by (p, σ, τ) (resp., p, σ^k, τ) on the set of plays $(H_\infty, \mathscr{H}_\infty)$. $\mathbf{E}$ (resp., $\mathbf{E}^k$) is the corresponding expectation. Obviously one has $\mathbf{P} = \sum_k p^k \mathbf{P}^k$ so that $\mathbf{P} \gg \mathbf{P}^k$ for any k and p^k is the size of the grain of truth of $\mathbf{P}^k$ in $\mathbf{P}$.

Define $\theta_n(i, j) = \frac{1}{n} \sum_{m=1}^n \mathbf{1}_{\{i, j\}}(i_m, j_m)$ and let $\pi^k_{i, j} = \mathscr{L}(\mathbf{E}^k\{\theta_n(i, j)\})$ be the asymptotic expected empirical frequency (a.e.e.f.) of the moves under σ^k and τ. Clearly, (iv) and (v) are satisfied by (α, β) and π.

We claim that π belongs to $\Pi(K)$. To prove (i), note that otherwise (by Proposition 2.1) there exists $q \in \Delta(K)$ and $x \in \Delta(I)$ such that $\sum_k q^k x A^k y > \langle \alpha, q \rangle$, $\forall y \in \Delta(J)$. If π^* denotes the a.e.e.f. corresponding to (x^*, τ), one obtains $\sum_k q^k \langle \pi^*, A^k \rangle > \langle \alpha, q \rangle$ so that for some $k^* : \langle \pi^*, A^{k^*} \rangle > \alpha^{k^*}$. To follow σ unless in state k^* where x^* is used would then be a profitable deviation for Player 1.

The equilibrium property for Player 1 implies obviously that the "no cheating condition" (iii) is satisfied.

It thus remains to prove (ii) and this relies on merging. The equilibrium condition for Player 2 implies obviously that his payoff is individually rational at the start of the game: $\langle B, \sum_k p^k \pi^k \rangle \geq \mathrm{val}_2 B$ and also along the play. Hence, for any m, $\mathbf{P}$ almost surely:

$$\langle B, \mathscr{L}(\mathbf{E}\{\theta_n(i,j)|\mathscr{H}_m\})\rangle \geq \mathrm{val}_2 B. \tag{3}$$

Theorem 2.2 implies that $\mathbf{P}$ merges to $\mathbf{P}^k$, in particular one has, taking expectation and limit, $\mathbf{P}^k$ a.s.,

$$\mathscr{L}(\mathbf{E}^k\{\theta_n(i,j)|\mathscr{H}_m\}) - \mathscr{L}(\mathbf{E}\{\theta_n(i,j)|\mathscr{H}_m\}) \longrightarrow 0 \quad \text{as } m \longrightarrow \infty, \tag{4}$$

which says that the asymptotic distributions under $\mathbf{P}$ or $\mathbf{P}^k$ on the moves in the future, computed at any stage m large enough, are close one to the other. Thus, (3) and (4) together imply that, $\mathbf{P}^k$ a.s.:

$$\lim_{m\longrightarrow\infty} \langle B, \mathscr{L}(\mathbf{E}^k\{\theta_n(i,j)|\mathscr{H}_m\})\rangle \geq \mathrm{val}_2 B. \tag{5}$$

Since

$$\begin{aligned}\mathbf{E}^k(\mathscr{L}(\mathbf{E}^k\{\theta_n(i,j)|\mathscr{H}_m\})) &= \mathscr{L}(\mathbf{E}^k(\mathbf{E}^k\{\theta_n(i,j)|\mathscr{H}_m\}))\\ &= \mathscr{L}(\mathbf{E}^k\{\theta_n(i,j)\})\\ &= \pi^k_{i,j},\end{aligned}$$

one obtains from (5) condition (ii).

Basically the merging property allows to deduce from the individual rationality condition $\langle B, \pi\rangle \geq \mathrm{val}_2 B$, the more precise property: $\langle B, \pi^k\rangle \geq \mathrm{val}_2 B$, $\forall k$. ∎

COROLLARY 3.6. *$E(p)$ is nonempty.*

Proof. We prove that $\Pi(K)$ is nonempty. Choose y optimal for Player 2 in game B, let x^k be a best reply of Player 1 to y in game A^k and define $\pi^k = x^k \otimes y$, i.e., $\pi^k_{i,j} = x^k_i y_j$. Conditions (ii) and (iii) are clearly satisfied. Finally if x' is a best reply of Player 1 to y in $A(q)$, one has $x' A^k y \leq x^k A^k y$ for all k, thus $\mathrm{val}_1 A(q) \leq x' A(q) y \leq \sum_k q^k x^k A^k y = \langle \alpha, q\rangle$. ∎

Comments. **1.** Part 1 of the previous proof (use of Blackwell's approachability criteria and joint plans) is now standard, see Aumann and Maschler (1995), Sorin (1983), Hart (1985).

2. The proof of Corollary 3.6 for the class of games with incomplete information on one side is much more difficult, see Sorin (1983) for $\#K = 2$ and Simon *et al.* (1995) for the general case. Note that Koren (1992) has an example showing that the existence result does not extend to the class of known own payoff games with lack of information on both sides.

3. The set of vector payoffs equilibria of Player 1 is the projection of $E(p)$ on $\mathbb{R}^K$ and depends only on the payoff matrices in the support of p.

For $p \gg 0$, denote it by $F(A^1, \ldots, A^K; B)$.

3.3. *Games with Perturbed Payoffs*

We consider here a two-person nonzero-sum repeated game defined by a pair of $I \times J$ matrices A and B and perturbed in the following way. With a small positive probability Player 2 believes that Player 1's payoff is given by one of finitely many matrices $A^2, \dots, A^K$. Assuming this uncertainty known by Player 1, the situation is then analogous to the incomplete information game described in 3.2, with $A = A^1$, in particular the set of vector equilibrium payoffs for Player 1 is independent of the precise specification of the perturbation probability as long as it has full support. We are interested in the payoff of the true type, more precisely we define her smallest equilibrium payoff:

$$\underline{m}(A, A^2, \dots, A^K; B) = \min\{\alpha^1; \alpha \in F(A^1, \dots, A^K; B)\}.$$

Let also:

$$m(A, B) = \sup_{x \in \Delta(I)} \inf_{y \in \Delta(J)} \{xAy; xBy \geq \operatorname{val}_2(B)\}. \tag{6}$$

Then one has the following bounds:

PROPOSITION 3.7 (Shalev, 1994; Israeli, 1996). (1) $\underline{m}(A, A^2, \dots, A^K; B) \leq m(A, B)$, *for any family* $A^2, \dots, A^K$

(2) $\underline{m}(A, -B; B) \geq m(A, B)$.

Proof. (1) We adapt Forges's idea (personal communication). By definition of $m(A, B)$, $\Phi = \{(x, y); xAy \leq m(A, B), xBy \geq \operatorname{val}_2 B\}$ is nonempty. For each k choose x^k and y^k that achieve $\alpha^k = \max\{xA^k y; (x, y) \in \Phi\}$. Finally let $\pi^k = x^k \otimes y^k$. We now prove that $\pi \in \Pi(K)$. Conditions (ii) and (iii) are clearly satisfied. Also $\langle \alpha, q \rangle = \sum_k q^k \max\{xA^k y; (x, y) \in \Phi\} \geq \max\{x(\sum_k q^k A^k)y; (x, y) \in \Phi\} \geq \operatorname{val}_1 A(q)$.

(2) For the converse, let α be an equilibrium payoff for the game $(A, -B; B)$. Conditions (i) and (ii) together imply $\alpha^2 = \operatorname{val}_1(-B) = -\operatorname{val}_2 B$. Again (i) and Proposition 3.4 lead to: $\forall x$, $\exists y$, $xAy \leq \alpha^1$ and $x(-B)y \leq \alpha^2$. Thus, $\alpha^1 \geq \min_y\{xAy : xBy \geq \operatorname{val}_2 B\}$ for all x, which gives point 2). ∎

Comments. **1.** The interpretation is that the most advantageous perturbation for Player 1 (in terms of lower bound on her equilibrium payoffs) is the one that induces maximal constraints on Player 2: this is the case when facing the perturbed type, Player 2 plays a two-person zero-sum game, i.e., $A^2 = -B$. Compared to the Folk theorem, the individually rational level of the informed player increases from $\operatorname{val}_1 A$ to $m(A, B)$.

2. This result indicates a strong discontinuity w.r.t. p of the set of equilibrium payoffs for the undiscounted case. As soon as the slightest amount of uncertainty is present the informed player can build a reputation, see Shalev (1994), Israeli (1996).

3. Note that the bound $m(A, B)$ may not be reached. Choose, for example:

$$A = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}.$$

Here $m(A, B) = \frac{1}{2}$, and for $x = \frac{1}{2}$ there is no constraint on y and $\min_y xAy = -1$.

3.4. *Discounted Case*

We consider now the case where both players use personal discount factors to evaluate their payoffs. For Player 1 the payoff is thus $\overline{a}_{\lambda_1}(\sigma, \tau) = E\ (\sum_{m=1}^{\infty} \lambda_1\ (1-\lambda_1)^{m-1} a_m)$ with $0 < \lambda_1 \le 1$ and similarly $\overline{b}_{\lambda_2}(\sigma, \tau) = E\ (\sum_{m=1}^{\infty} \lambda_2 (1-\lambda_2)^{m-1} b_m)$ for Player 2. We denote by $\Gamma_{(\lambda_1,\, \lambda_2)}(p)$ the corresponding game with incomplete information.

The results of this section are due to Cripps and Thomas (1995b). They first show that if the informed player is "infinitely patient" with respect to the other one, she can do as well as in the undiscounted case. This means in particular that in perturbed games like in Section 3.3, she can drastically reduce to her advantage her set of equilibrium payoffs, compared to the complete information case.

PROPOSITION 3.8 (Cripps and Thomas, 1995b). *For any $p \gg 0$, any discount factor λ_2 and any positive ε^*, there exists $\lambda_1^* = \lambda_1^*(p, \lambda_2, \varepsilon^*)$ such that, if $\lambda_1 \le \lambda_1^*$, any equilibrium vector payoff α of Player 1 in $\Gamma_{(\lambda_1,\lambda_2)}(p)$ is within ε^* of $F(A^1, \ldots, A^K; B)$.*

Proof. We use the notations introduced in the proof of Proposition 3.5. Let σ, τ be an equilibrium in $\Gamma_{\lambda_1,\lambda_2}(p)$. Given λ_2 and $\eta > 0$, let N such that the weight on stages $1, \ldots, N$, given λ_2 is at least $1 - \eta$ (N is the approximate horizon of Player 2). We define the λ_2-average frequency between stages m and n by

$$\theta_{\lambda_2}(m, n)(i, j) = \sum_{r=m}^{n} \lambda_2 (1-\lambda_2)^{r-m} \mathbf{1}_{\{i,\, j\}}(i_r, j_r),$$

and the corresponding payoff for Player 2 by:

$$b_{\lambda_2}(m, n) = \sum_{r=m}^{n} \lambda_2 (1-\lambda_2)^{r-m} b_r = \langle B, \theta_{\lambda_2}(m, n)\rangle.$$

The equilibrium condition for Player 2 implies as in section 3.2 that she is individually rational on each play. Hence, for all m, $\mathbf{P}$ a.s.,

$$\mathbf{E}(b_{\lambda_2}(m+1, \infty)|\mathscr{H}_m) \ge \operatorname{val}_2 B,$$

which implies, by the choice of N:

$$\mathbf{E}(b_{\lambda_2}(m+1, m+N)|\mathscr{H}_m) \geq \operatorname{val}_2 B - \eta.$$

In particular, for all m, $\mathbf{P}$ a.s., the payoff on the mth block of size N satisfies:

$$\mathbf{E}(b_{\lambda_2}(mN+1, (m+1)N)|\mathscr{H}_{mN}) \geq \operatorname{val}_2 B - \eta. \tag{7}$$

We now use Lemma 2.4 for the filtration $\mathscr{F}_n = \mathscr{H}_{nN}$ correponding to the sequence of blocks of size N and δ, ε to be chosen later. One obtains that, for all indices $m \notin \mathscr{M}$, with $\#\mathscr{M} \leq M$:

$$\mathbf{P}^k\Big\{h_{mN};\ \sup_{A\in\mathscr{H}_{(m+1)N}} |\mathbf{P}(A|h_{mN}) - \mathbf{P}^k(A|h_{mN})| \geq \delta\Big\} \leq \varepsilon. \tag{8}$$

Hence for $m \notin \mathscr{M}$, (7) and (8) imply that, with probability at least $1 - \varepsilon$ under $\mathbf{P}^k$, a similar bound holds for the payoffs of Player 2 evaluated under $\mathbf{P}^k$:

$$\mathbf{E}^k(b_{\lambda_2}(mN+1, (m+1)N)|\mathscr{F}_m) \geq \operatorname{val}_2 B - \eta - 2\delta.$$

So that for $m \notin \mathscr{M}$, taking expectation

$$\mathbf{E}^k(b_{\lambda_2}(mN+1, (m+1)N)) \geq \operatorname{val}_2 B - \eta - 2\delta - 2\varepsilon,$$

and finally, in this case, by the choice of N:

$$\mathbf{E}^k(b_{\lambda_2}(mN+1, \infty)) \geq \operatorname{val}_2 B - 2\eta - 2\delta - 2\varepsilon. \tag{9}$$

This last inequality means that except for finitely many values of m, the payoff of Player 2, at the beginning of the mth block is still (almost) individually rational under $\mathbf{P}^k$ (and not only under $\mathbf{P}$).

We want to deduce properties concerning the frequencies evaluated according to Player 1's criteria. Recall that if $\lambda_1 \leq \lambda_2$, $\theta_{\lambda_1}(m, \infty)$ is in the convex hull of the family $\theta_{\lambda_2}(n, \infty)$, $n \geq m$. Define now $\lambda_1^* \leq \lambda_2$ so that the weight of the first MN stages given λ_1^* is less than η. From (9) one obtains, for $\lambda_1 \leq \lambda_1^*$,

$$\mathbf{E}^k(b_{\lambda_1}(mN+1, \infty)) + 2\eta \geq \operatorname{val}_2 B - 2\eta - 2\delta - 2\varepsilon,$$

hence in particular, letting $\pi^k = \mathbf{E}^k(\theta_{\lambda_1}(1, \infty))$, one deduces

$$\langle B, \pi^k\rangle \geq \operatorname{val}_2 B - 4\eta - 2\delta - 2\varepsilon, \tag{10}$$

which means that also under Player 1's evaluation, Player 2's payoff is almost individually rational, given $\mathbf{P}^k$.

Note that the equilibrium vector payoff of Player 1 is $\alpha^k = \langle A^k, \pi^k\rangle$. The individual rationality condition for Player 1 implies, as in the proof of Proposition 3.5, that $\langle \alpha, q\rangle \geq \operatorname{val}_1 A(q)$, for all q in $\Delta(K)$, using Proposition 3.4. Finally the equilibrium condition obviously gives: $\langle \pi^k, A^k\rangle \geq \langle \pi^{k'}, A^k\rangle$, for all k, k'. To finish the proof of Proposition 3.8, we need

LEMMA 3.9. *Define, for* ζ *a real number and* $p \gg 0$: $\Pi_\zeta(K) = \{\pi = \{\pi^k\}; \pi^k \in \Delta(I \times J), k \in K$ *with*:

(i) $\langle \alpha, q \rangle \geq \mathrm{val}_1 A(q) - \zeta, \ \forall q \in \Delta(K)$

(ii) $\langle B, \pi^k \rangle \geq \mathrm{val}_2 B - \zeta, \ \forall k$

(iii) $\langle A^k, \pi^k \rangle \geq \langle A^k, \pi^{k'} \rangle - \zeta, \ \forall k, \ k' \in K\}$.

Then, for any ε^* *positive there exists a positive* ζ *such that* $\Pi_\zeta(K)$ *is included in the* ε^**-neighborhood of* $\Pi(K)$.

Proof of the Lemma. The lemma says that the correspondance $\zeta \mapsto \Pi_\zeta(K)$ is u.h.c., with $\Pi(K) = \Pi_0(K)$ and this is clear since all the constraints are continuous in ζ. ■

Let now β be defined by $\langle B, \sum_k p^k \pi^k \rangle$ and recall that the map from Π to E is nonexpansive. Note that (i)–(iii) are satisfied with $\zeta = 4\eta + 2\delta + 2\varepsilon$. Given ε^*, choose ζ according to the lemma, then $\eta = \delta = \varepsilon = \zeta/8$ to get Proposition 3.8. ■

Comment. **1.** The result holds a fortiori if Player 1's payoff is undiscounted, and also if Player 1 is facing a sequence of Player 2 that lives for a finite number of stages (see Section 4.1).

2. The previous proof uses in a crucial way the fact that the bound λ_1^* is a function of λ_2: λ_2 determines the horizon N. This allows to use weak merging through a new normalization of the process, but Player 1 has to be "infinitely" more patient than Player 2.

3. In fact a new and important result of Cripps and Thomas (1995b) shows that for p^1 near 1 and $\lambda_1 = \lambda_2$ near 0 the range of the component α^1 of the equilibrium payoff of Player 1 is like in the complete information case (Folk theorem) with payoffs A^1 and B.

The next result shows that for small discount factors the set of equilibrium payoffs almost contains $E(p)$.

PROPOSITION 3.10 (Cripps and Thomas, 1995b). *Assume* $p \gg 0$ *and* ε *positive. Then there exist* $\lambda^* = \lambda^*(\varepsilon)$, *such that any* (α, β) *in* $E_{-\varepsilon}(p) = \{(\alpha, \beta);\ \exists \pi \in \Pi_{-\varepsilon}(K), \langle A^k, \pi^k \rangle = \alpha^k, \ \forall k \in K, \ \langle B, \sum_k p^k \pi^k \rangle = \beta\}$ *is an equilibrium payoff of* $\Gamma_{(\lambda_1, \lambda_2)}(p)$ *for any* $\lambda_1 \leq \lambda^*$, $\lambda_2 \leq \lambda^*$.

Sketch of Proof. The proof follows Part 1 of the proof of Proposition 3.5. Choose a smooth play w^k (cf., Sorin, 1992, p. 82) then discounted factors small enough to generate the payoffs and to respect the individual rational constraints, taking into account the stages needed to code the different types k. (Note that the Definition 3.3 of approachability implies continuity with respect to the discount factor.) ■

Comments. Altogether the previous results show that the limit behavior, for discount factors near 0, of the set of equilibrium payoffs of $\Gamma_{(\lambda_1, \lambda_2)}(p)$ depends crucially on the relative size of the discount factors. There is a path $(\lambda_2, \lambda_1(\lambda_2))$ going to $(0, 0)$ on which continuity to $E(p)$ holds. The lim inf of these families of sets always contains the interior of $E(p)$ but might be much larger.

4. REPUTATION: BASIC RESULTS

4.1. *Presentation*

A two-person game defined by two $I \times J$ payoff matrices A and B is played repeatedly. We consider first the case of *standard signalling* where, after each stage n, the chosen actions (i_n, j_n) are announced to both players. Let $a_n = A_{i_n, j_n}$, $b_n = B_{i_n, j_n}$ be the corresponding payoffs for Players 1 and 2, respectively. We assume that Player 1 is a long run player, hence plays at every stage. Her opponent might be of several types: either there is a sequence of short lived Players 2 where each one lives for finitely many periods and is then replaced, or a single long Player 2 is present. In all cases, the agents playing at stage n are aware of the whole past history h_{n-1} of moves until that stage.

$G_{(\mu_1, \mu_2)}$ is the game with payoffs parameters μ_1 and μ_2. $\mu_1 = \lambda_1$ corresponds to the case where Player 1's payoff is discounted with factor λ_1 and $\mu_1 = \infty$ stands for the undiscounted case (see 3.1 for the evaluation of the payoffs). As for Player 2, both values $\mu_2 = \lambda_2$ or $\mu_2 = \infty$ are feasible but in addition we consider the case $\mu_2 = n$ where n is the duration of each Player 2's life.

We add now a perturbation to this game: there are some strategies (in the repeated game) to which, with positive probability, Player 1 is committed. Moreover this fact is known to Player 2.

The basic idea of the literature on reputation, see Kreps and Wilson (1982), Milgrom and Roberts (1982), Kreps *et al.* (1982), is that Player 1 can use this uncertainty, namely, play some specific strategy, say φ, in the repeated game to build a reputation effect for being of this type, i.e., being in fact committed to φ. Once convinced of facing this type Player 2 should adapt and this could benefit Player 1. Clearly this possibility depends on both players' characteristics μ_1 and μ_2. (The influence of the *signalling structure* is considered in the next sections).

We exhibit a lower bound on Player 1's payoff (normal type) in any equilibrium of a perturbation of the game $G_{(\mu_1, \mu_2)}$ where with positive probability ρ she is of "type φ." (This does not prevent other perturbations to be present as well). The result is obtained by computing the payoff of Player 1 mimicking her type φ and relies on merging properties and on the behavior of Player 2 at equilibrium. Denote this (class of) perturbed game by $G_{(\mu_1,\mu_2)}(\rho,\varphi)$.

4.2. *General Properties*

Let (σ,τ) be an equilibrium in the perturbed game $G_{(\mu_1,\mu_2)}(\rho,\varphi)$. $\widetilde{\sigma}$ denotes the strategy of Player 1 defined by σ, the commitment types and the perturbation probability. Formally $\tilde{\sigma}$ can be written as $\tilde{\sigma}=\theta\sigma+(1-\theta)\sigma'$ with $\theta\geq\rho$, and σ' being of the form $\sigma'=\frac{\rho}{\theta}\varphi+(1-\frac{\rho}{\theta})\varphi'$. P is the probability on $(H_\infty,\mathscr{H}_\infty)$ induced by $(\widetilde{\sigma},\tau)$ while Q is the one generated by the type φ and τ. Obviously one has $P=\rho Q+(1-\rho)Q'$, for some distribution Q'. It follows that P merges to Q and that the merging properties of Section 2.3 hold. (ρ is the size of the grain of truth.)

The general procedure to obtain the lower bound w on player is as follows:

Step (1) uses the fact that Player 2 is playing a best reply to $\widetilde{\sigma}$ in $G_{(\mu_1,\mu_2)}$ to get at each stage a lower bound under P on her conditional expected payoff for the future. This gives a corresponding property on the conditional probabilities on histories.

Step (2) deduces from the merging results of Section 2 a similar property on the conditional probabilities on histories when Player 1 is of type φ (i.e., under Q).

Step (3) computes then the lowest payoff of Player 1 under type φ, given this property. Since Player 1 can always choose to mimic any perturbation this gives a bound on her payoff.

Note that the two first steps of this procedure are actually similar to those used in the proofs of Propositions 3.5 and 3.8.

Let us also underline the following: If Player 1 is less patient than Player 2, a large part of Player 2's payoff is achieved while Player 1's payoff is essentially over, hence Player 1 cannot efficiently monitor Player 2. We thus consider here only the case where Player 1 is more patient than Player 2.

4.3. *The Case* $\mu_2=1$; *Fudenberg and Levine* (1989, 1992)

The results in this framework correspond to the first systematic treatment of reputation bounds in the literature, after the study of a collection of specific cases.

Given x in $\Delta(I)$, we assume that the perturbation φ is x^*, i.e., playing x i.i.d. We first deal with Step 1.

Let $h_n\in H_n$. Since Player 2 is playing an equilibrium in a one-stage game, $\tau(h_n)$ is a best reply to $\widetilde{\sigma}(h_n)$, with P probability 1.

Step 2 indicates that, under Q, Player 2 is almost always playing an almost best reply to x:

LEMMA 4.1. *Given $\eta' > 0$, there exists M such that, except at most on M stages m, one has*:

$$E_Q(b_m) \geq \max_y \ xBy - \eta'.$$

Proof. Denote by BR^2 the best reply correspondence of Player 2 and more generally, for $\varepsilon \geq 0$, $BR^2_\varepsilon(z) = \{y \in \Delta(J);\ zBy \geq zBy' - \varepsilon$, for all $y'\}$, so that $BR^2 = BR^2_0$. Given $x \in \Delta(I)$, choose $\delta > 0$ such that $\|x' - x\| \leq \delta$ implies $BR^2(x') \subset BR^2(x)$. We use Lemma 2.4 with this δ and $\varepsilon = \eta'$ to define an exceptional set of stages $\mathscr{M}$ of cardinality M. For $m \notin \mathscr{M}$ one has, with Q probability at least η', $e_m(P, Q)(w) \leq \delta$, hence in particular $\|\widetilde{\sigma}(h_m) - x\| \leq \delta$. So that the strategy that Player 2 is facing after h_m differs from x by less than δ. Since she is playing a best reply $\tau(h_m)$ belongs to $BR^2(x)$. ■

We are now ready for Step 3. Since Player 1 is patient, the finite set of exceptional stages defined in the previous result does not affect her payoff.

PROPOSITION 4.2 (Fudenberg and Levine, 1992). *For all x in $\Delta(I)$, $\rho > 0$, $\eta > 0$, there exists λ_1^* such that for any $\lambda_1 \leq \lambda_1^*$, any equilibrium payoff a of Player 1 in the game $G_{(\lambda_1, 1)}(\rho, x^*)$ satisfies*:

$$a \geq \min_y \{xAy;\ y \in BR^2(x)\} - \eta.$$

Proof. Let $g(x) = \min_y\{xAy;\ y \in BR^2(x)\}$ and $\varphi(x) = xBy$ for $y \in BR^2(x)$. Let $\eta' > 0$ be such that $xBy \geq \varphi(x) - \eta'$ implies $xAy \geq g(x) - \eta/2$. Apply then the previous lemma with η', and let λ_1^* be such that the weight of M stages is less than $\eta/2$. The bound on Player 1's payoff then follows. ■

Define

$$w(A, B) = \sup_{x \in \Delta(I)} \inf_{y \in \Delta(J)} \{xAy;\ xBy \geq xBy', \forall y' \in \Delta(J)\}.$$

This corresponds to the best previous lower bound of Propositin 4.2 if we let the perturbation x^* vary.

COROLLARY 4.3 (Fudenberg and Levine, 1992). *If the perturbation of the game $G_{(\lambda_1, 1)}$ has full support on $\Delta(I)$ (interpreted as i.i.d. strategies), then for any $\eta > 0$, there exists λ_1^* such that for any $\lambda_1 \leq \lambda_1^*$, any equilibrium payoff a of Player 1 satisfies*:

$$a \geq w(A, B) - \eta.$$

It follows easily that any equilibrium payoff of Player 1 *in the corresponding perturbation of the game* $G_{(\infty,1)}$, *where her payoff is undiscounted is also greater than* $w(A, B)$.

Finally one can show, see Fudenberg and Levine (1992), that this bound is tight.

Note that in the present case with $\mu_2 = 1$ and standard signalling, Player 2 can anticipate the mixed move of her opponent and play a stage after stage best reply. The bound we obtained has thus a Stackelberg flavour.

In the general case Player 2 is only playing a best reply to some strategy that coincides with Player 1's strategy on the set of histories compatible with both players' strategies. In particular the merging of P to Q does not imply the merging of the probability induced by $\widetilde{\sigma}$ and τ' to the one induced by (φ, τ'), for an alternative strategy $\tau' \neq \tau$.

4.4. $G_{(\infty,\infty)}$; *Cripps and Thomas* (1995*a*)

This case is the extreme opposite of the previous one since both players have undiscounted payoffs and, as expected, the results are quite similar to those of Section 3.3.: equilibrium payoffs of incomplete information games in the undiscounted case.

Step 1 can be written, with the notation of Section 3.2 as:

For all m, P almost surely:

$$\langle B, \mathscr{L}(E_P\{\theta_n(i,j)|\mathscr{H}_m\})\rangle \geq \text{val}_2\, B. \tag{11}$$

For Step 2, Theorem 2.2 implies that P merges to Q, in particular one has, Q a.s.:

$$\mathscr{L}(E_Q\{\theta_n(i,j)|\mathscr{H}_m\}) - \mathscr{L}(E_P\{\theta_n(i,j)|\mathscr{H}_m\}) \longrightarrow 0 \quad \text{as } m \longrightarrow \infty. \tag{12}$$

Equations (11) and (12) imply that, Q a.s.:

$$\lim_{m\longrightarrow\infty} \langle B, \mathscr{L}(E_Q\{\theta_n(i,j)|\mathscr{H}_m\})\rangle \geq \text{val}_2\, B. \tag{13}$$

Hence, taking expectation, the payoff of Player 2 given Q satisfies:

$$\langle B, \mathscr{L}(E_Q\{\theta_n(i,j)\})\rangle \geq \text{val}_2\, B.$$

For Step 3, just write:

$$a \geq \langle A, \mathscr{L}(E_Q\{\theta_n(i,j)\})\rangle.$$

Assume that the perturbation φ is equal to some x^*. Then the asymptotic frequency $\mathscr{L}(E_Q\{\theta_n(i,j)\})$ will be of the form $x \otimes y$. Thus we obtain:

PROPOSITION 4.4 (Cripps and Thomas, 1995a). *For all x in $\Delta(I)$, $\rho > 0$, any equilibrium payoff a of Player* 1 *in the game $G_{(\infty,\infty)}(\rho, x^*)$ satisfies*:

$$a \geq \inf_y \{xAy; xBy \geq \mathrm{val}_2(B)\}.$$

Letting x vary, we obtain:

COROLLARY 4.5 (Cripps and Thomas, 1995a). *If the perturbation has full support on $\Delta(I)$, any equilibrium payoff a of Player* 1 *in the perturbed version of the game $G_{(\infty,\infty)}$ satisfies*:

$$a \geq m(A, B) = \sup_x \inf_y \{xAy; xBy \geq \mathrm{val}_2(B)\}.$$

Comments. **1.** Note that one obtains the same bound as in the case of perturbed payoffs, cf. Proposition 3.7. While the result would have been obvious in the case of pure stationary perturbations (that can be obtained through a payoff matrix with the corresponding line having strictly higher payoffs than any other entry), it is not clear how to define a repeated game where x^* is a strictly dominant strategy.

2. This bound is the best one: see an explicit construction in Cripps and Thomas (1995a).

4.5. $G_{(\lambda_1,\lambda_2)}$, $\lambda_1 \gg \lambda_2$; *Cripps et al.* (1996)

The analysis of this discounted game, with a more patient perturbed Player 1 is very similar to the one done in Section 3.4: equilibrium payoffs of incomplete information games in the discounted case. It follows previous results obtained by Schmidt (1993) in a special case.

Steps 1 and 2 are exactly as in the proof of Proposition 3.8 and one obtains inequality (10). Assuming that the perturbation of Player 1 is x^* implies that $E_Q(\theta_{\lambda_1}(1,\infty))$ can be written as $x \otimes y$. Hence by continuity one obtains:

PROPOSITION 4.6 (Cripps *et al.* 1996). *For any x in $\Delta(I)$, any $\rho > 0$, any $\varepsilon > 0$, and any discount factor λ_2, there exists $\lambda_1^* = \lambda_1^*(\rho, \lambda_2, \varepsilon)$ such that, if $\lambda_1 \leq \lambda_1^*$, any equilibrium vector payoff a of Player* 1 *in $G_{(\lambda_1,\lambda_2)}(\rho, x^*)$ satisfies*:

$$a \geq \inf_y \{xAy; xBy \geq \mathrm{val}_2(B)\} - \varepsilon.$$

Similarly one has:

COROLLARY 4.7 (Cripps *et al.* 1996). *If the perturbation has full support on $\Delta(I)$, for any discount factor λ_2 and any positive ε, there exists λ_1^* such* *that, if* $\lambda_1 \leq \lambda_1^*$, *any equilibrium payoff* a *of Player* 1 *in the perturbed version of the game* $G_{(\lambda_1, \lambda_2)}$ *satisfies*:

$$a \geq m(A, B) - \varepsilon.$$

Comments. **1.** As in 3.4, it is crucial here to allow Player 1 to be infinitely patient compared to Player 2, see Cripps and Thomas (1997) for a kind of Folk theorem result when $\lambda_1 = \lambda_2$.

2. The results immediately extend to the case $\mu_2 = n$, where Player 1 is facing a sequence of finite lived Player 2.

3. Like in Section 4.4 the bound obtained for games with perturbed strategies is the same as the one obtained for games with perturbed payoffs, cf. Proposition 3.8.

4. Also here the bound is tight, see Cripps *et al.* (1996).

5. Both results in 4.4 and 4.5 indicate that the situation is less favourable for the perturbed player when she is facing longer opponents. (This still holds when allowing for a larger set of perturbations than playing i.i.d., for example, those implementable by a finite automaton, see Cripps *et al.* 1996). The reason is that Player 2 does know the strategy used by her opponent, even when merging occurs. She might then be afraid of a punishment inducing a worse payoff in case she will experiment. This phenomena is studied in the next section.

5. SIGNALS AND CONJECTURAL EQUILIBRIUM

5.1. *Reputation and Signals*

Assume that Player 1's pertubation is a strategy consisting of playing an i.i.d. sequence of $x \in \Sigma_N$, a strategy in the N-stage game. Let again $\widetilde{\sigma}$ denote the strategy induced by σ and the perturbation. Denote by $\widetilde{\sigma}_N[h_{nN}] \in \Sigma_N$ the strategy for the N next stages defined by $\widetilde{\sigma}$, after some history h_{nN}. The merging of the probability on histories induced by $\widetilde{\sigma}$ and a strategy τ of Player 2 to the one induced by x^* and τ does not imply the convergence of the strategy $\widetilde{\sigma}_N[h_{nN}]$ to x. But the distributions $\widetilde{\sigma}_N[h_{nN}](h)$ and $x(h)$ on I will be closed to each other on histories h having positive probability under x and τ. Since the perturbation of Player 1 depends upon the moves of Player 2, Player 2's strategy now plays an active role in the revelation process.

In fact standard signalling in a repeated game implies, when normalizing the strategies, say every N stages, that one deals with a game with signals: the information after each stage does not reveal entirely the normalized strategy used at that stage.

To completely describe such a game one needs to introduce signalling functions, l^1 from $I \times J$ to some signal space A for Player 1 and similarly l^2 from $I \times J$ to B for Player 2. The interpretation is as follows: at each stage n, given the moves (i_n, j_n), the signal to Player r is $l^r(i_n, j_n)$, $r = 1, 2$. As usual these functions are extended from $\Delta(I) \times \Delta(J)$ to $\Delta(A)$ and $\Delta(B)$ (and they could be random to start with).

5.2. *Conjectural Equilibrium*

In the framework of a game with signals, a notion of stable behavior is captured by the definition of *conjectural equilibrium*, see Hahn (1973, 1977): each player plays a best reply to some conjecture and receives signals consistent with her conjecture. Formally one has:

DEFINITION. $(x, y) \in \Delta(I) \times \Delta(J)$ is a *conjectural equilibrium* if there exists $(x', y') \in \Delta(I) \times \Delta(J)$ such that:

(1) $x \in BR^1(y')$ and $y \in BR^2(x')$

(2) $l^1(x, y) = l^1(x, y')$ and $l^2(x, y) = l^2(x', y)$.

Obviously this definition reduces to the definition of Nash equilibrium when *standard signalling* holds: $l^1(i, j) = l^2(i, j) = \{i, j\}$.

Related definitions and properties have been introduced by Battigalli *et al.* (1992), Fudenberg and Levine (1989): generalized best response, (1992): self-confirming response, (1993): self-confirming equilibria, Kalai and Lehrer (1993b): subjective equilibrium.

5.3. *Bound for Games with Signals*; *Fudenberg and Levine* (1992)

In the framework of Section 4.3 (Player 2 living for one stage), the analysis of reputation effects in a game with signals follows the same lines. If Player 1 plays x^*, then for almost all stages, Player 2 with a high probability uses a best reply to a strategy of Player 1 that induces almost the same signal than x.

Formally the set of histories of length n is now $H_n = (I \times A \times J \times B)^n$ on which strategies of both players induce a probability distribution. We work on the corresponding set of histories for Player 2, $H_n^2 = (J \times B)^n$, with the corresponding marginal probabilities.

With the notations of Section 4, Step 1 is as follows. With P probability 1, at $h \in H_n^2$, Player 2 plays a best reply to the marginal distribution on I_{n+1} of the conditional probability P given h, say $\pi[h]$.

Step 2 uses the uniform weak merging of P to Q to obtain that: except on M stages, with Q probability at least $1 - \varepsilon$, the distributions on signals for Player 2 are almost the same under her updated beliefs and the true perturbation x: $|l^2(\pi[h], \tau(h)) - l^2(x, \tau(h))| \le \delta$.

Step 3 shows then that Player 1 can achieve the lower bound $\inf_{x',y}\{xAy;\ y \in BR^2(x'),\ l^2(x,y) = l^2(x',y)\}$, for λ_1 small enough. Finally define the analog of $w(A,B)$ for this framework:

$$wl^2(A,B) = \sup_{x\in\Delta(I)} \inf_{\substack{x'\in\Delta(I)\\ y\in\Delta(J)}} \{xAy;\ y \in BR^2(x'),\ l^2(x,y) = l^2(x',y)\}.$$

Then one obtains the counterpart of Corollary 4.3:

PROPOSITION 5.1 (Fudenberg and Levine, 1992). *If the perturbation has full support on* $\Delta(I)$, *then for any* $\eta > 0$, *there exists* λ_1^* *such that for any* $\lambda_1 \le \lambda_1^*$, *any equilibrium payoff a of Player* 1 *in the perturbation of the game* $G_{(\lambda_1,1)}$ *with signals* (l^1, l^2) *satisfies*:

$$a \ge wl^2(A,B) - \eta.$$

Comments. **1.** Note that since Player 1 does not have to monitor Player 2, the signalling function l^1 can be arbitrary.

2. For the other cases, corresponding to Sections 4.4 and 4.5, the bound is unchanged since we only use an individually rational (and not a best reply) condition on Player 2's payoff, hence l^2 is irrelevant.

3. Similarly if $\mu_2 = n$ and the perturbation has full support on Player 1's strategies in the n-stage game, the bound on her equilibrium payoffs is $wl_n^2(A,B)$, which corresponds to the function wl^2 applied to the normalized n-stage repetition of (A,B). (Note that now l^1 matters through the definition of Player 1's strategy set).

6. SIGNALS AND REVELATION

6.1. *Revelation and Full Support*

Consider a repeated two-person game where again standard signalling for Player 2 holds. Given a pair of strategies (σ,τ), denote by $P_{(\sigma,\tau)}$ the induced probability on $(H_\infty, \mathscr{H}_\infty)$.

A straightforward but fundamental remark is that: as soon as τ is completely mixed (given any history h, $\tau(h)$ has full support on J), for any strategies σ, σ' of Player 1, $P_{(\sigma,\tau)}(C) = P_{(\sigma',\tau)}(C)$, for all $C \subset H_N$ implies that σ and σ' have the same reduced form (Recall that, following Shapley, σ and σ' have the same reduced form if they induce the same play given any strategy τ.) in the N-stage game.

The same property holds when dealing with the probability induced on the set of Player 2's histories, $H_n^2 = (B \times J)^n$ as soon as the following condition **R** (for revelation) holds:

(1) l^2 is *revealing* in the sense that: $l^2(x, y) = l^2(x', y)$ implies $x = x'$, for all y, and:

(2) for any $h \in H^2$ and $i \in I$, $\tau(h)$ induces a distribution $l^1(i, \tau(h))$ with full support on the set A of signals of Player 1.

By continuity one obtains then the following crucial result for strategies $\sigma, \sigma' \in \Sigma_N$ of Player 1 and $\tau \in \mathcal{T}_N$ of Player 2 in the N-stage game:

Property 6.1 (Celentani *et al.* 1996). Assume l^2 revealing and $l^1(i, \tau(h))(a) \geq \eta > 0$, for all $a \in A$, $i \in I$, and $h \in H^2$. Then, for any positive ρ, there exists a positive ε such that:

(1) $\sup_{C \subset H_N^2} |P_{(\sigma, \tau)}(C) - P_{(\sigma', \tau)}(C)| \leq \varepsilon$,

and

(2) τ belongs to $BR_\varepsilon^2(\sigma)$

imply that τ belongs to $BR_\rho^2(\sigma')$.

Comment. This property precisely prevents Player 2 to play an almost best reply to a conjecture while being far from a true best reply.

Two models exhibiting these features have been proposed: Aoyagi (1996) considers a version of the game with "trembling hand" on the side of Player 2 so that $\tau(h)(j) \geq \eta > 0$ for all $h \in H^2$ and all $j \in J$. Celentani *et al.* (1996) assume that $l^1(i, j)$ has full support on A for all $(i, j) \in I \times J$.

Consider now the repeated perturbed game and assume $\mu_2 = N$. The structure of the proof to get lower bounds on Player 1's equilibrium payoffs is as usual. Let (σ, τ) be an equilibrium of the perturbed game $G_{(\lambda_1, N)}(\rho, \varphi)$. $\widetilde{\sigma}$ is the perturbed strategy induced by σ and φ, P (resp., Q) is the probability defined by $(\widetilde{\sigma}, \tau)$ (resp., (φ, τ)) on $H_\infty = (I \times A \times J \times B)^\infty$.

Step 1. Let $\pi_N[h_{nN}]$ (resp., $\varphi_N[h_{nN}]$) be the strategy of Player 1 for the next N stages, defined by P (resp., Q) conditionally to $h_{nN} \in H_{nN}^2$. Then $\tau_N(h_{nN})$ is a best reply to $\pi_N[h_{nN}]$ under P.

Step 2. We apply the merging results for P and Q on $(H_{nN}^2, \mathcal{H}_{nN}^2)$. Thus for almost all stages n, with a high probability under Q one has: $|P(C|h_{nN}) - Q(C|h_{nN})|$ small, for all $C \subset H_N^2$. In particular, using Property 6.1, $\tau_N(h_n N)$ is an almost best reply to $\varphi_N[h_{nN}]$.

Step 3. The aim here is to define a perturbation φ independent of N that would induce a payoff for Player 1 near the bound:

$$s_N(A, B) = \sup_{x \in \Sigma_N} \inf_{y \in \mathcal{T}_N} \{\overline{a}_N(x, y);\ y \in BR^2(x)\}.$$

6.2. *$s(A, B)$ and an Adapted Stackelberg Strategy*

Define $s(A, B)$ to be the largest payoff of Player 1 feasible and compatible with Player 2 's individually rational level. Formally one has

DEFINITION. Let F be the convex hull of the one stage payoffs $\{(A_{ij}, B_{ij});\ i \in I, j \in J\}$. Hence,

$$s(A, B) = \max\{a|(a, b) \in F,\ b \geq \text{val}_2(B)\}.$$

To avoid technicalities we assume from now on the existence of a point $(a, b) \in F$ with $a > \text{val}_1(A)$ and $b > \text{val}_2(B)$. We aim at constructing first for any positive ε, a strategy x of Player 1 in some finitely repetition G_N such that for some positive η and any $y \in \mathcal{T}_N \bigcap BR^2_\eta(x)$ one has $\overline{a}_N(x, y) \geq s(A, B) - \varepsilon$. This shows in particular that $s_N(A, B)$ converges to $s(A, B)$.

In a game with standard signalling the strategy is essentially of the form: follows a cyclical play h^* achieving a payoff (a, b) with $a > s(A, B) - \varepsilon/2$ and $b > \text{val}_2(B)$ and punish Player 2 in case of deviation. (One can, for example, take an increasing sequence of payoffs for Player 2 along the cycle to avoid late deviations).

In the current framework of Celentani *et al.* (1996), the payoffs could be random (A_{ij} is then defined as the expectation) and the moves are not observable. We assume nevertheless (in addition to the previous conditions on l^1 and l^2, see 6.1) that the signal to Player 1 contains his payoff. Hence she can monitor Player 2 by constructing a strategy based on a sequence of tests comparing his empirical payoff on blocks to the theoretical one (see Celentani *et al.* 1996). Moreover the constructed strategy s (s for Stackelberg) will have the following robustness property:

Property 6.2 (Celentani *et al.* 1996). Given any positive ε, there exists N_0 and a strategy $s \in \Sigma_{N_0}$, N_1 and a positive η, such that given any history h and any $\tau \in \mathcal{T}_N \bigcap BR^2_\eta(s^*_N[h])$ with $N \geq N_1$ the payoff of Player 1 satisfies

$$\overline{a}_N(s^*_N[h], \tau) \geq s(A, B) - \varepsilon,$$

where $s^*_N[h]$ denotes the strategy for the next N stages, after history h, induced by the play of s^* which is s i.i.d.

In words this means that if Player 1 is using s^*, then after any history h, an approximate best reply of Player 2 to s^* in a long enough game G_N gives to Player 1 a payoff of almost $s(A, B)$.

6.3. *The Result*

Define Σ^*_M to be the discretisation of width $1/M$ of the simplex of normalized mixed strategies of Player 1 in the M-stage game. The perturbation is *complete* if it has full support on the countable set $\Sigma^* = \bigcup_M \Sigma^*_M$. As usual if x is selected, then x i.i.d. is played.

Both results by Aoyagi (1996) and Celentani *et al.* (1996) show that under complete perturbation, if N is large and λ_1 small the lower bound on Player 1's payoff approaches $s(A, B)$. In fact in Step 3 above one can choose φ in Σ^* arbitrarily near s satisfying Property 6.2. So that for most blocks of length N and with high probability the best reply τ to $\widetilde{\sigma}$ will be an approximate best reply to φ, hence an approximate best reply to s, so that the average payoff of Player 1 will be near $s(A, B)$. For λ_1 small, the weight of the bad blocks vanishes hence the result. However the precise conditions in both cases are quite different and described below.

In Aoyagi's approach, given ε, choose first a tremble η then a length of Player 2's life (to have Properties 6.1 and 6.2) and finally λ_1 small enough to get the result. In particular, assuming Player 1's payoff undiscounted one obtains:

PROPOSITION 6.3 (Aoyagi, 1996). *The payoff a of Player* 1 *in any trembling hand perfect equilibria of $G_{(\infty, N)}$ with complete perturbation satisfies*

$$a \geq c(N),$$

with

$$\lim_{N \longrightarrow \infty} c(N) = s(A, B).$$

On the other hand, Celentani *et al.* (1996) deal with a fixed noise in the signals of Player 1 and they prove:

PROPOSITION 6.4 (Celentani *et al.* 1996). *For any positive ε and any N, there exist $\lambda_1^*(\varepsilon, N)$ such that for any $\lambda_1 \leq \lambda_1^*$ any equilibrium payoff a of Player* 1 *in the game $G_{(\lambda_1, N)}$ with complete perturbation satisfies*

$$a \geq d(N) - \varepsilon,$$

with:

$$\lim_{N \longrightarrow \infty} d(N) = s(A, B).$$

Comments. **1.** A similar result has been obtained by Celentani (1996) in the framework $\mu_2 = 1$ with application to extensive games.

2. The same structure of proof works for $\mu_2 = \lambda_2$ small enough. In fact this is the original framework of both Aoyagi (1996) and Celentani *et al.* (1996). One chooses then N large enough so that the weight on $[1, N]$ given λ_2 is almost one and that the best reply (for the discounted payoff of Player 2) still satisfies Property 6.2.

3. Recall that to get the bound $s(A, B)$ a condition on the length of Player 2 is needed. If λ_2 is too large, Player 1 has not enough time to monitor Player 2 through her payoffs.

4. The three bounds obtained up to now obviously satisfy:

$$m(A,B) \le w(A,B) \le s(A,B).$$

This shows that reputation effect could benefit from richer signals if their meaning is clear and explicit. On the other hand to increase the set of perturbations (and to allow enough stages for Player 2 to adapt) may be worse for Player 1 if it is possible for Player 2 to hold wrong and durable conjectures. A simple example where the bounds differ is given by

	α	β	γ
a	$(4,1)$	$(0,0)$	$(1,4)$
b	$(0,0)$	$(1,4)$	$(0,0)$

where $m(A,B) = \frac{1}{2}$, $w(A,B) = 1$, and $s(A,B) = 3$.

7. REPUTATION AND COMPLEXITY

To summarize the results up to now:

— a lower bound, $m(A,B)$, on Player 1's equilibrium payoffs has been obtained, which shows that reputation effects occur under very weak conditions

— to make the reputation effect more profitable for her (in the sense of increasing this lower bound), Player 1 has to instruct and monitor precisely Player 2. This is done by a process of transmitting and receiving information. When it is possible for Player 1 to reveal completely her strategy (one case is when facing a Player 2 living for one stage like in Section 4.3., another way is to add noise which is the model in Section 6) then it is better for him to face a patient Player 2 (either with a small discounted rate λ_2 or living for a large number of periods). In fact the situation is like a Stackelberg game where Player 1 is playing first hence she benefits from having a larger strategy set. This way the best feasible and individually rational payoff $s(A,B)$ can be achieved.

A similar result has been obtained by Evans and Thomas (1997) without noise in the signals but it requires nonstationary perturbations, more precisely an unbounded hierarchy of punishments. Denote by $\mathscr{P}$ the set of couples (h,x) where h is a finite history in H and x a mixed move in $\Delta(I)$. To each $(h,x) \in \mathscr{P}$ is associated a strategy $\varphi = \varphi(h,x)$ of Player 1 in the repeated game as follows. Write $h = (h^1, h^2)$, where h^i corresponds to the sequence of moves of Player i. Player 1 starts by obeying the main path:play according to h^1 in cycles as long as Player 2 plays according to h^2 in cycles. After a first deviation of Player 2, Player 1 finishes the cycle, plays once x and starts the main path again. Inductively after the nth deviation of Player 2, Player 1 finishes the current cycle of h^1, plays n times the mixed move x i.i.d. and reverts to the main path. (Note that the behavior of Player 2 between a deviation and the end of the following "punishing" phase is irrelevant).

Choose h^* like in the construction of Section 6.2 (achieving an individually rational payoff (a, b) with a near $s(A, B)$) and x minmaxing Player 2. When facing $\varphi = \varphi(h^*, x)$ Player 2 suffers longer and longer punishments if she is not following h^*. Since her payoff is discounted, if the number of deviations is not bounded, the future expected payoff after some stage would be basically 0. By merging, the payoff would be the same facing the perturbed strategy while it would be better for Player 2 to use the alternative strategy: play a best reply to φ, and get at least otherwise (if φ is not played) her minmax. This property bounds the number of deviations of Player 2 and allows Player 1 to force her to follow h^* forever after finitely many stages.

PROPOSITION 7.1 (Evans and Thomas, 1997). *For any positive ε and ρ, there exists $(h, x) \in \mathcal{P}$ and λ_2^* such that given any $\lambda_2 \leq \lambda_2^*$, there exists λ_1^* such that for any $\lambda_1 \leq \lambda_1^*$, any equilibrium payoff a of Player* 1 *in the perturbed game $G_{(\lambda_1, \lambda_2)}(\rho, \varphi(h, x))$ satisfies*:

$$a \geq s(A, B) - \varepsilon.$$

Proof. Given the game (A, B), let φ be a strategy of Player 1 in the repeated game induced by $(h, x) \in \mathcal{P}$ and such that:

— the average payoffs of Players 1 and 2 on a cycle of h satisfy $a(h) \geq s(A, B) - \varepsilon/2$ and $b(h) > 0 = \mathrm{val}_2(B)$

— $b(x, y) \leq 0$ for all $y \in \Delta(J)$.

Define on the set of plays H_∞, the stopping time χ_1 corresponding to the first deviation of Player 2 with respect to φ. F_1 is the corresponding set of histories ($h \in F_1 \cap H_n$ if for any play h_∞, $\chi_1(h, h_\infty) = n$) and $\mathcal{F}_1$ is the σ-algebra generated by F_1. We introduce also χ_2, the stopping time corresponding to the second deviation of Player 2 w.r.t. φ, $\mathcal{F}_2$ the associated σ-algebra, and similarly χ_n and $\mathcal{F}_n$, for any n.

Given an equilibrium pair (σ, τ), P and Q are defined as usual as the probabilities induced by $(\widetilde{\sigma}, \tau)$ and (φ, τ), respectively, on $(H_\infty, \mathcal{H}_\infty)$. Recall that the initial probability of φ is ρ. We consider the merging properties with respect to the filtration $\{\mathcal{F}_n\}$ and we use Lemma 2.5. Thus except on a set of probability ε_1 (the other histories are called *regular*) the number of stages n where $e_n(P, Q)(f_n)$ is greater than δ is less than M (the other stages are *good*).

Assume h of length N_0. Let λ_2^* be small enough so that the weight on N_0 stages is at most δ and $\lambda_2 \leq \lambda_2^*$ implies $b_{\lambda_2}(h_\infty) > b(h) - \delta$ for all cyclical plays h_∞ compatible with h. Given such a λ_2, choose N so that the λ_2 discounted payoff of Player 2 after stage N is uniformly less than δ. This implies that for $n \geq N$, the expected future payoff of Player 2 on an history after the nth deviation, $b_{\lambda_2}(\varphi, \tau)(f_n)$ is at most 2δ.

Under Q, on the regular histories and on the good stages, merging gives the following relation on payoffs: $|b_{\lambda_2}(\widetilde{\sigma}, \tau)(f_n) - b_{\lambda_2}(\varphi, \tau)(f_n)| \leq \delta + 2\delta$. In particular, except on the set of nonregular histories, for all good stages $n \geq N$, this implies $b_{\lambda_2}(\varphi, \tau)(f_n) \leq 5\delta$. Call these stages *losing*.

Recall (see the proof of Lemma 2.5) that the probability under Q of the histories on which the posterior probability of φ is at least $t\rho$, is greater than $(1 - t)$. Call such histories *nice*. Note that Player 2 can always expect at least $b(h) - \delta$ if facing φ and playing according to h and at least her minmax, up to δ, otherwise. Let thus $t = \varepsilon_1$ and choose δ such that:

$$(b(h) - \delta)\rho\varepsilon_1 - (1 - \rho\varepsilon_1)\delta > 5\delta.$$

The equilibrium condition for Player 2 shows that a losing stage does not exist on regular and nice histories. So that except on a set of probability at most $2\varepsilon_1$, the number of deviations of Player 2 under Q is uniformly bounded by $N + M$. This implies that the cyclic path induced by h is played in bounded time and for λ_1 small enough $a_{\lambda_1}(\sigma, \tau) \geq s(A, B) - 2/3\varepsilon$. One finally gets the result by choosing $\varepsilon_1 \leq \varepsilon/12$. ■

Comments. **1.** The original proof of Evans and Thomas (1997) assumes that Player 1 can punish Player 2 with a pure strategy (namely, the above x is pure). Dealing with this class of games one can define a countable set of perturbations (replacing $\Delta(I)$ by I in the definition of $\mathcal{P}$), such that Proposition 7.1 holds for any probability with full support on it.

Note that the proof only requires that both players identify a strategy φ adapted to $s(A, B)$; as remarked by Evans and Thomas (1997) it is sufficient that the signal contains her own payoff for each player.

The result extends to a sequence of finitely lived Player 2 if they live much longer than the length of h and φ starts again cooperatively when facing a new player, but keeps counting the deviations and punishes accordingly.

Finally the proof is inadequate for $G_{(\infty, \infty)}$ since an undiscounted Player 2 can still expect a good payoff when facing $\varphi(h)$, whatever being the history h.

2. The class of perturbations used in the previous proof is necessarily complex. A counterexample involving strategies generated by finite automata can be found in Cripps *et al.* (1996). However not only the class of strategies is more complex (one could consider all x^*, x being a strategy in some finite repetition of the game), but each of the perturbations is so, since it needs an unbounded counter to recall the number of deviations. This aspect is crucial in order to avoid a finite number of future feasible payoffs given $\varphi(h)$, as h varies.

On the other hand, the strategies φ induced by $\mathcal{P}$ have the important property that after any history h, Player 2 can generate a finite history h' such that $\varphi(hh')$ is the best for her among the strategies she could generate given φ—which extends the property of the bounded recall strategies of Aumann and Sorin (1989), see also the forgiving strategies of Watson (1996).

8. FINAL COMMENTS

8.1. *Reputation and Rationality*

As observed by Watson (1993) and refined in Battigalli and Watson (1997), see also Watson (1996), much less than the equilibrium notion is needed for reputation effects to hold. One basically uses one level of rationalizability on the side of Player 2: the fact that she is playing a best reply to some perturbed strategy $\widetilde{\sigma}$ of Player 1. In particular one does not use the fact that Player 1 is playing a best reply.

Moreover the results are very robust: it is enough that the required perturbation φ be present with a positive probability ρ to get the bounds on payoffs.

Finally if one assumes, say in the discounted case, that with a small probability Player 2 may not be rational, all the results hold qualitatively provided one works with ε-consistency (Lehrer and Sorin, 1998): given any history compatible with her strategy, Player 2 is still behaving rationally, i.e., maximizing the future payoff, with a high probability. On the other hand it is easy to see that reputation may fail when facing an ε-maximizing Player 2 with a large discount factor. (Note that from this point of view the situation is qualitatively different if Player 1 is facing a sequence of finite lived Player 2).

8.2. *Related Results*

Celentani and Pesendorfer (1996) use the same logic in the framework of dynamic games where the stage game may vary. In addition they introduced a continuum of small players whose actions can neither be observed nor influence the other payoffs—but they care about the state variable.

Relation between merging and equilibrium, but not in a reputation framework, have been investigated by Kalai and Lehrer (1993a) and Jackson and Kalai (1997).

The main question in reputation effects is how Player 1 can transmit some information to Player 2, especially on the strategy she is actually playing: namely, how to reveal. A dual aproach is to investigate the behavior of a Player 2 who wants to learn Player 1's strategy. In term of merging this corresponds to the informational content of the strategy. A measure through the maximal variation of the martingale of posterior probabilities has been introduced by Stearns, see Aumann and Maschler (1995) and Mertens *et al.* (Chapters V and VI, 1994).

8.3. *Reputation on Both Sides*

The extension of the current results to situations where uncertainty is on both sides seems in general difficult.

In the discounted case, the order in which the discount factors were chosen was crucial for the properties to hold: the informed player has to adapt herself to the speed at which the uninformed player learns. Basically Player 2 learns first (and the time she needs to learn is a function of his own discount factor) and then, due to the established reputation, Player 1 gets good payoffs. In terms of interpretation, to emphasize this disymmetry, it might be better to present the model as an infinite Player 1 facing a sequence of Players 2 with finite life. In the undiscounted case one may face nonexistence results, see Koren (1992) and Watson (1996). The hope is to get advances in specific classes where revelation seems more natural, like common interest games in the spirit of Aumann and Sorin (1989).

REFERENCES

Aoyagi, M. (1996). "Reputation and Dynamic Stackelberg Leadership in Infinitely Repeated Games," *J. Econ. Theory* **71**, 378–393.

Aumann, R. J., and Hart, S. (1986). "Bi-Convexity and Bi-Martingales," *Israel J. Math.* **54**, 159–180.

Aumann, R. J., and Maschler, M. (1995). *Repeated Games with Incomplete Information*. Cambridge, MA: MIT Press, (with the collaboration of R. Stearns).

Aumann, R. J., and Sorin, S. (1989). "Cooperation and Bounded Recall," *Games Econ. Behavior* **1**, 5–39.

Battigalli, P., Gilli, M., and Molinari, M. C. (1992). "Learning and Convergence to Equilibrium in Repeated Strategic Interactions: An Introductory Survey," *Ricerche Econ.* **XLVI**, 335–378.

Battigalli, P., and Watson, J. (1997). "On "Reputation" Refinements with Heterogeneous Beliefs," *Econometrica* **65**, 369–374.

Blackwell, D. (1956). "An Analog of the Minmax Theorem for Vector Payoffs," *Pacific J. Math.* **6**, 1–8.

Blackwell, D., and Dubins, L. (1962). "Merging of Opinions with Increasing Information," *Ann. Math. Statist.* **33**, 882–886.

Celentani, M. (1996). “Reputation with Observed Actions,” *Econ. Theory* **7**, 407–419.

Celentani, M., Fudenberg, D., Levine, D., and Pesendorfer, W. (1996). “Maintaining a Reputation against a Long Lived Opponent,” *Econometrica* **64**, 691–704.

Celentani, M., and Pesendorfer, W. (1996). “Reputation in Dynamic Games,” *J. Econ. Theory* **70**, 109–132.

Cripps, M., Schmidt, K., and Thomas, J. (1996). “Reputation in Perturbed Repeated Games,” *J. Econ. Theory* **69**, 387–410.

Cripps, M., and Thomas, J. (1995a). “Reputation and Commitment in Two-Person Repeated Games without Discounting,” *Econometrica* **63**, 1401–1419.

Cripps, M., and Thomas, J. (1995b). “The Folk Theorem in Repeated Games of Incomplete Information,” preprint.

Cripps, M., and Thomas, J. (1997). “Reputation and Perfection in Repeated Common Interest Games,” *Games Econ. Behavior* **18**, 141–158.

Evans, R., and Thomas, J. P. (1997). “Reputation and Experimentation in Repeated Games with Two Long-Run Payers,” *Econometrica* **65**, 1153–1173.

Forges, F. (1992). “Repeated Games of Incomplete Information: Non-zero-Sum,” in *Handbook of Game Theory, I* (R. J. Aumann and S. Hart, Eds.), pp. 155–177, New York: Elsevier.

Fudenberg, D., and Levine, D. (1989). “Reputation and Equilibrium Selection in Games with a Patient Player,” *Ecomometrica* **57**, 759–778.

Fudenberg, D., and Levine, D. (1992). “Maintaining a Reputation when Strategies are Imperfectly Observed,” *Rev. Econ. Stud.* **59**, 561–579.

Fudenberg, D., and Levine, D. (1993). “Self-Confirming Equilibrium,” *Econometrica* **61**, 523–545.

Hahn, F. (1973). *On the Notion of Equilibrium in Economics: An Inaugural Lecture*. Cambridge, U.K.: Cambridge Univ. Press.

Hahn, F. (1977). “Exercises in Conjectural Equilibria,” *Scand. J. Econ.* **79**, 310–226.

Hart, S. (1985). “Nonzero-Sum Two-Person Repeated Games with Incomplete Information,” *Math. Oper. Res.* **10**, 117–153.

Israeli, E. (1996). “Sowing Doubt Optimally in Two-Person Repeated Games,” Center for Rationality, D.P. 112.

Jackson, M., and Kalai, E. (1997). “Social Learning in Recurring Games,” *Games Econ. Behavior* **21**, 102–134.

Kalai, E., and Lehrer, E. (1993a). “Rational Learning Leads to Nash Equilibrium,” *Econometrica* **61**, 1019–1045.

Kalai, E., and Lehrer, E. (1993b). “Subjective Equilibrium in Repeated Games,” *Econometrica* **61**, 1231–1240.

Kalai, E., and Lehrer, E. (1994). “Weak and Strong Merging of Opinions,” *J. Math. Econ.* **23**, 73–86.

Koren, G. (1992). “Two-Person Repeated Games where Players Know Their Own Payoffs,” Courant Institute, preprint.

Kreps, D., Milgrom, P., Roberts, J., and Wilson, R. (1982). “Rational Cooperation in the Finitely Repeated Prisoner’s Dilemma,” *J. Econ. Theory* **27**, 245–252.

Kreps, D., and Wilson, R. (1982). “Reputation and Imperfect Information,” *J. Econ. Theory* **27**, 253–279.

Lehrer, E., and Smorodinsky, R. (1996a). “Compatible Measures and Merging,” *Math. Oper. Res.* **21**, 697–706.

Lehrer, E., and Smorodinsky, R. (1996b). “Merging and Learning,” in *Statistics, Probability and Game Theory* (T. S. Ferguson, L. S. Shapley and J. B. MacQueen, Eds.), Vol. 30, pp. 147–168, Haywood, CA: Inst. Math. Statist.

Lehrer, E., and Smorodinsky, R. (1997). “Repeated Large Games with Incomplete Information,” *Games Econ. Behavior* **18**, 116–134.

Lehrer, E., and Sorin, S. (1998). “ε-Consistency in Repeated Games,” *Int. J. Game Theory* **27**, 231–244.

Mertens, J.-F., Sorin, S., and Zamir, S. (1994). “Repeated Games,” Core DP 9420-21-22.

Milgrom, P., and Roberts, J. (1982). “Predation, Reputation and Entry Deterrence,” *J. Econ. Theory* **27**, 280–312.

Schmidt, K., (1993). “Reputation and Equilibrium Characterization in Repeated Games with Conflicting Interests,” *Econometrica* **61**, 325–351.

Shalev, J. (1994). “Nonzero-Sum Two-Person Repeated Games with Incomplete Information and Known-Own Payoffs,” *Games Econ. Behavior* **7**, 246–259.

Simon, R., Spiez, S., and Torunczyk, H. (1995). “The Existence of Equilibria in Certain Games, Separation for Families of Convex Functions and a Theorem of Borsuk–Ulam type,” *Israel J. Math.* **92**, 1–21.

Sorin, S. (1983). “Some Results on the Existence of Nash Equilibrium for Non-Zero-Sum Games with Incomplete Information,” *Int. J. Game Theory* **12**, 193–205.

Sorin, S. (1990). “Supergames,” in *Game Theory and Applications* (T. Ichiishi, A. Neyman, and Y. Taumann, Eds.), pp. 46–63, San Diego: Academic Press.

Sorin, S. (1992). “Repeated Games with Complete Information,” in *Handbook of Game Theory, I* (R. J. Aumann and S. Hart, Eds.), pp. 71–107, New York: Elsevier.

Sorin, S.(1995). “Bluff et réputation,” *Revue d’Economie Politique* **105**, 583–600.

Watson, J. (1993). “A ‘Reputation’ Refinement without Equilibrium,” *Econometrica* **61**, 199–205.

Watson, J. (1994). “Cooperation in the Infinitely Repeated Prisoner’s Dilemma with Perturbations,” *Games Econ. Behavior* **7**, 260–285.

Watson, J. (1996). “Reputation in Repeated Games with No Discounting,” *Games Econ. Behavior* **15**, 82–109.