MATHEMATICS OF OPERATIONS RESEARCH
Vol. 24, No. 1, February 1999
*Printed in U.S.A.*

# REPEATED GAMES WITH INCOMPLETE INFORMATION ON ONE SIDE: THE CASE OF DIFFERENT DISCOUNT FACTORS

EHUD LEHRER AND LEEAT YARIV

Two players engage in a repeated game with incomplete information on one side, where the underlying stage-games are zero-sum. In the case where players evaluate their stage-payoffs by using different discount factors, the payoffs of the infinitely repeated game are typically non zero-sum. However, if players grow infinitely patient, then the equilibrium payoffs will sometimes approach the zero-sum result, depending on the asymptotic relative patience of the players. We provide sufficient conditions that ensure a zero-sum limit. Moreover, we provide examples of games violating these conditions that possess "cooperative" equilibria whose payoffs are bounded away from the zero-sum payoffs set.

**1. Introduction.** Two players engage in a repeated game with incomplete information. The state of nature is chosen according to a known prior and the selected state is informed only to one player. Then, the zero-sum game that corresponds to the selected state is played infinitely many times. The uninformed player partially learns during the course of the game about the realized state by receiving signals, which depend also on the informed player's actions. For an extensive discussion on repeated games with incomplete information the reader is referred to Mertens, Sorin, and Zamir (1994). As opposed to most of the existing literature, we assume different time preferences: the players discount their future payoffs with respect to different discount factors.

The issue of different discounting was treated by Lehrer and Pauzner (1995), who looked at complete information repeated games. They showed that in the case of a zero-sum game, the only Nash equilibrium payoff of the repeated game is the value of the stage-game. Thus, although the repeated game is non zero-sum, the only equilibrium payoff is zero-sum. This, however, is not the case in incomplete information games.

Consider, for instance, a game where the informed player is very impatient or almost myopic. He can exploit, right at the outset of the game, his informational advantage, receive relatively high payoffs at the beginning and reveal what he knows to the relatively patient player. The latter, who cares more about long-run stage-payoffs, receives high payoffs at most stages due to the information he acquired. Thus, the players receive high payoffs in different stages. Since the players have different time preferences, this plan (a) can be sustained by an equilibrium, and (b) induces a non zero-sum pair of payoffs in the repeated game (for an explicit example, see Example 1 in §2).

This scenario attests to the fact that not only is the game itself non zero-sum, but also that some of the equilibrium payoffs may be non zero-sum. In other words, despite the fact that the stage-games are purely competitive, different discounting gives rise to some cooperation. The high payoff received by the informed impatient player is a result of a complete utilization of his information in the beginning of the game. However, in case the informed player is not so impatient, in order to take advantage of his extra information, he should split the revealed information over a greater number of periods. This, in turn, entails that at each period his

Received November 28, 1995; revised November 8, 1996; September 3, 1998; April 22, 1998; July 15, 1998.
*AMS 1991 subject classification.* Primary: 90D05, 90D20.
*OR/MS Index 1978 subject classification.* Primary: Games/Repeated.
*Key words.* Repeated games, incomplete information, different discount factors.



0364-765X/99/2401/0204/$05.00

payoff must be reduced (compared to the payoff received at a period in which all the information is revealed at once).

These arguments suggest that the informed player cannot exploit his information for a long period of time. Moreover, if he becomes more and more patient, then his informational advantage becomes negligible: after a short while, when almost all his information is being revealed, the game played is like a complete information game, where the equilibrium payoff is zero-sum. Therefore, one may expect that when players become patient, the equilibrium payoffs will become zero-sum.

It turns out that this intuition is misleading. The equilibrium payoff may or may not become, in the limit, zero-sum, depending on the converging path of both discount factors (henceforth, discounting path) to 1. Finding out when equilibrium payoffs are necessarily zero-sum in the limit when players become increasingly patient, is the goal of this paper.

We divide the discussion into two cases: the first is when the informed player (player 1) is more patient than the uninformed one and the second is when player 1 is less patient. Due to the asymmetric nature of information, the results in these two cases are qualitatively different. It turns out that the ratio $(1 - \lambda_2)/(1 - \lambda_1)$ is the appropriate measurement for the players' relative patience, where $\lambda_i$ is player $i$'s discount factor. This ratio will determine whether the equilibrium payoffs are necessarily zero-sum in the limit or not.

When the informed player is more patient, he will get, in any Nash equilibrium, a payoff very close to the value of the infinitely repeated zero-sum undiscounted game. If, in addition, the uninformed player is not significantly less patient than the informed player, in the sense that the ratio $(1 - \lambda_2)/(1 - \lambda_1)$ is bounded from above, then his equilibrium payoffs will also converge to the same value. Thus, asymptotically the equilibrium payoffs are zero-sum. The precise statement of these results will be provided in Theorem 1 in §3.

In §4 we link together the discounting paths that ensure a zero-sum limit payoff and the convergence speed of the discounted repeated games' values, with identical discount factors, as the discount factor converges to 1. Theorem 2 states that if the speed of convergence is high relative to the ratio $(1 - \lambda_2)/(1 - \lambda_1)$, then the limit payoff is zero-sum.

In the other case, when the uninformed player is more patient, if both players become more and more patient in a similar manner (in the sense that the ratio $(1 - \lambda_2)/(1 - \lambda_1)$ converges to 1), then the equilibrium payoffs will converge to the zero-sum value. An explicit formalization will be given in Theorem 3 in §5.

Regardless of whether the informed player is more patient or less, when the discount factors become singular (in the sense that they assign most of the weight to disjoint sets of stages), the limit of the equilibrium payoffs may be non zero-sum. We provide two examples to illustrate this phenomenon: one for the case where the informed player is more patient (in §3) and one for the other case (in §4).

Players with different discount factors have also been recently dealt with in the context of reputation effects. In the reputation literature, a very patient player, the informed one, plays against one or a sequence of impatient (frequently myopic) players (see for instance, Fudenberg and Levine (1989, 1992), Schmidt (1993), Celentani, Fudenberg, Levine and Pesendorfer (1996), Evans and Thomas (1997) and Sorin (1997) for a comprehensive overview of the reputation literature). The message of the reputation literature is that the informed player, having the advantages of both information and patience, obtains at least his true type Stackelberg payoff. He can do so by behaving as if he was of a certain type, his "commitment type," for a long enough time, building up a reputation of being "committed" and by exploiting this reputation.

Our paper differs from the reputation literature in spirit and in technique. First, having two long-lived players, as we have, gives rise to agreements. The relatively patient player might agree to act benevolently towards his opponent in the beginning of the game, with the promise to get reimbursed in the future, which is less significant to the impatient player. In this way both players can achieve strictly more than their individually rational levels (see

Example 3 in §4). In other words, despite the zero-sum underlying games, there is room for mutually beneficial intertemporal trading of payoffs. This cooperative behavior stands in sharp contrast with the spirit of the reputation literature results.

Second, and related to the first, in the reputation literature, myopic players (or those who play stationary strategies over finite blocks) can use only the information they have gathered up to the point of play and, more importantly, have no punishing power. Hence, they cannot enforce a specific path of play. Our work, in contrast, deals with a less patient uninformed player who has an infinite horizon of moves at every stage of the game. Thus, like the informed player who can punish his opponent (by withholding information and/or by using an undesirable action), the uninformed player can also punish his opponent by reducing his payoff to the individually rational level of the corresponding infinitely repeated zero-sum game. Cripps and Thomas (1995) are an exception in that they treat the case of two players that have different discount factors. They show that the more patient informed player can ensure what he could get in the undiscounted game. However, a major difference between their model and ours is that there the uninformed, usually the impatient player, knows his own payoffs, while here the uninformed player typically does not know them. Third, since we deal with underlying zero-sum games, the Stackelberg payoff of each type is the equilibrium payoff of the corresponding zero-sum game and the lower bound on the payoff of the long-lived player is the individually rational level in our framework.

**2. The model and an example.** A two-person repeated game with incomplete information on one side consists of a finite state space $K$, a set of zero-sum games $\{A^k\}$, $k \in K$, a probability distribution $p$ over $K$, and an information structure $\varphi = (\varphi^k)_{k\in K}$. $\varphi^k = (\varphi_1^k, \varphi_2^k)$ is a function from all possible combinations of the players' actions to pairs of signals. $\varphi_i^k$ is the signal given to player $i$. We assume that the signal a player receives contains his last action. Typically, the players' signals are private, but in most cases treated in the literature the signals are the actions played. In other prevalent cases the signals contain, on top of players' own actions, their payoffs.

The repeated game proceeds as follows. At stage 0, a state $k$ from $K$ is chosen according to $p$ and is told only to player 1. At any stage of the game, starting at 1, each player takes an action, which may depend on his previous signals and acquired information. He then receives a signal according to $\varphi_i^k$ and a payoff, corresponding to $A^k$. Notice that since $A^k$ is a zero-sum game, if $a_t$ is the payoff of player 1, then $-a_t$ is the payoff of player 2.

Formally, let $H_i$ be the set of all player $i$'s finite histories of signals. A (behavioral) *strategy* of player 1 (resp. 2) is a function from $K \times H_1$ (resp. $H_2$) to the set of his mixed actions $\Delta(\Sigma_1)$ (resp. $\Delta(\Sigma_2)$). Typically a strategy of player 1 is denoted by $\sigma$ and of player 2 by $\tau$. A pair $(\sigma, \tau)$ and the prior $p$ induce a probability distribution over the infinite streams of stage-payoffs. Thus, the stage-payoffs are random variables.

Here we deal with infinitely repeated games in which the two players evaluate their infinite stream of payoffs differently. More precisely, we associate with player $i$ the discount factor $\lambda_i$. If $\{a_t\}$ is the sequence of the stage-payoffs received by player $i$, his repeated game payoff is

$$(1-\lambda_i)\sum_{t=1}^{\infty}\lambda_i^{t-1}a_t.$$

We denote the game defined above by $G(p; \lambda_1, \lambda_2)$. Although the prevailing matrices are zero-sum, $G(p; \lambda_1, \lambda_2)$ for $\lambda_1 \neq \lambda_2$, is typically not a zero-sum game.

Let $a_t$ be player 1's payoff at stage $t$. The pair $(\sigma, \tau)$ is a *Nash equilibrium* of $G(p; \lambda_1, \lambda_2)$ if $E_{\sigma',\tau'}\,[(1-\lambda_i)\sum_{t=1}^{\infty}\lambda_i^{t-1}(-1)^{i+1}a_t]$ is maximized with $\sigma$ when $\tau'$ is fixed at $\tau$ and $i$

$= 1$ and, moreover, if it is maximized with $\tau$ when $\sigma'$ is fixed at $\sigma$ and $i = 2$. We denote the set of equilibrium payoffs in the game $G(p; \lambda_1, \lambda_2)$ by $V(p; \lambda_1, \lambda_2)$. Note that since both players' payoffs over the compact sets of behavioral strategies are continuous, existence of an equilibrium pair follows from standard fixed point arguments. The projection of $V(p; \lambda_1, \lambda_2)$ to the $i$th coordinate is denoted by $V^i(p; \lambda_1, \lambda_2)$. In this set-up both players are payoff maximizers.

Let $v(p; \lambda)$ be the value of the zero-sum game $G(p; \lambda, \lambda)$. In the case where $\lambda_1 = \lambda_2 = \lambda$, the set $V(p; \lambda_1, \lambda_2)$ contains only one pair, $(v(p; \lambda), -v(p; \lambda))$.

It is known (see Mertens, Sorin, and Zamir 1994) that $v(p; \lambda)$ converges uniformly on $P$, as $\lambda$ goes to 1. We denote the limit by $v(p; 1)$.

EXAMPLE 1. Consider the following game where the signals consist of the players' own action. There are two states of nature (i.e., $K = \{1, 2\}$), and the distribution over $K$ is $p = (\frac{1}{2}, \frac{1}{2})$. The payoffs are given by the following matrices.

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

Consider any Nash equilibrium of the game $G(p; \lambda_1, \lambda_2)$, where the discount factor of player 1, $\lambda_1$, is very low and that of player 2, $\lambda_2$, is high.

Note that when $\lambda_1$ is low $v(p; \lambda_1)$ is close to $\frac{1}{2}$. It implies, in particular, that player 1's equilibrium payoff is greater than $\frac{1}{4}$. On the other hand, since player 1 is very impatient he reveals his information at the beginning of the game. Hence, in most of the stages the expected payoff is close to zero. Player 2 is patient, and therefore his equilibrium payoff is close to 0. Thus, the players' payoffs are not zero-sum.

The intuition behind this example is that player 1, having a low discount factor, is eager to receive high payoffs at the first stages of the game. In order to achieve this goal, player 1 reveals all his private information in the initial stages. Player 2 benefits from player 1's impatience by learning the true state of nature quite early.

This example suggests that the advantage of player 1 is in his impatience. In other words, the high payoff player 1 receives is merely a result of a complete utilization of his private information in the beginning of the game. One may, therefore, expect that since player 1 cannot exploit his information for a long period of time, if he grows infinitely patient, his advantage will become negligible. As a result, the limit payoffs will necessarily be zero-sum. We will see that this intuition is incorrect.

A natural question to ask then is, under what conditions does $V(p; \lambda_1, \lambda_2)$ tend to the singleton $\{(v(p; 1), -v(p; 1))\}$, as one of the discount factors or both converge to 1? It turns out that whether or not the set of Nash equilibrium payoffs tends to a singleton depends on the asymptotic relative patience of the players.

DEFINITION 1. A discounting path is a pair of continuous functions $\lambda_1, \lambda_2 : [0, 1] \to [0, 1]$ satisfying $\lambda_1(\omega), \lambda_2(\omega) < 1$ for every $\omega \in [0, 1)$ and either $\lambda_1(1) = 1$ or $\lambda_2(1) = 1$ (or both).

Recall that $V^i(p; \lambda_1, \lambda_2)$ is the set of all player $i$'s payoffs sustainable by Nash equilibria. We say that $\lim_{\omega\to 1} V^i(p; \lambda_1(\omega), \lambda_2(\omega)) = (-1)^{i+1}v(p; 1)$ if for every $\nu(\omega) \in V^i(p; \lambda_1(\omega), \lambda_2(\omega))$, $\nu(\omega)$ converges to $(-1)^{i+1}v(p; 1)$ as $\omega$ tends to 1. For the sake of convenience, we omit $\omega$ and write $V^i(p; \lambda_1, \lambda_2)$ for $V^i(p; \lambda_1(\omega), \lambda_2(\omega))$.

Player $i$ can always play his optimal strategy of $G(p; \lambda_i)$ and guarantee himself the value $v(p; \lambda_i)$. Therefore,

$$\min V^i(p; \lambda_1, \lambda_2) \geq (-1)^{i+1} v(p; \lambda_i), \qquad i = 1, 2.$$

We refer to these two inequalities as the individual rationality of the players.

## 3. A more patient informed player.

### 3.1. The limit payoffs are zero-sum.

The main result of this section is:

THEOREM 1. *Let* $(\lambda_1, \lambda_2)$ *be a discounting path with* $\lambda_1(1) = 1$. *If* $\lambda_1(\omega) \geq \lambda_2(\omega)$ *for every* $\omega \in [0, 1]$, *then*

(a) $\limsup_{\omega\to 1} V^1(p; \lambda_1, \lambda_2) \leq v(p; \lambda_2(1))$;

(b) *If* $\lambda_2(1) = 1$ *then* $\lim_{\omega\to 1} V^1(p; \lambda_1, \lambda_2) = v(p; 1)$;

(c) *If* $\lambda_2(1) = 1$ *and* $(1 - \lambda_1(\omega))/(1 - \lambda_2(\omega))$ *is bounded away from zero, then* $\lim_{\omega\to 1} V^2(p; \lambda_1, \lambda_2) = -v(p; 1)$.

Theorem 1 states that if the informed player is relatively more patient, then he will receive, in the limit of any Nash equilibrium, payoffs not more than the value of the repeated game with the identical discount factor $\lambda_2$. If, in addition, player 2 also becomes extremely patient, then the informed player receives equilibrium payoffs which are very close to the zero-sum value of the infinitely repeated **undiscounted** game. The same will be true for the uninformed player, if he is not significantly less patient than the informed one.

The proof of Theorem 1 requires a notation and a lemma. For any $\lambda \in [0, 1)$, $p \in P$, and a period $s \geq 1$, if $\{a_t\}$ is the sequence of stage-payoffs, we denote by $N_s(\lambda; \{a_t\})$ the normalized future-payoff at stage $s$. Formally,

$$N_s(\lambda; \{a_t\}) = (1 - \lambda) \sum_{t=s}^{\infty} \lambda^{t-s} a_t.$$

It turns out that $N_1(\lambda_1; \{a_t\})$ can be expressed as an affine combination of $N_s(\lambda_2; \{a_t\})$, as stated without a proof in the following lemma.

LEMMA 1.

$$N_1(\lambda_1; \{a_t\}) = (\lambda_1 - \lambda_2)\left(\frac{1 - \lambda_1}{1 - \lambda_2}\right) \sum_{s=2}^{\infty} \lambda_1^{s-2} N_s(\lambda_2; \{a_t\}) + \left(\frac{1 - \lambda_1}{1 - \lambda_2}\right) N_1(\lambda_2; \{a_t\}).$$

Therefore, in case player 1 is more patient than player 2, one obtains,

COROLLARY 1. *For* $\lambda_1 \geq \lambda_2$, $N_1(\lambda_1; \{a_t\})$ *is a convex combination of* $N_s(\lambda_2; \{a_t\})$, $s = 1, 2, \ldots$.

PROOF OF THEOREM 1. Let $(\sigma, \tau)$ be a Nash equilibrium of $G(p; \lambda_1, \lambda_2)$. Let $a_t$ be player 1's payoff at stage $t$. Each $a_t$ is a random variable with a distribution induced by $\sigma$ and $\tau$. The pair of payoffs corresponding to $(\sigma, \tau)$ is $(E(N_1(\lambda_1; \{a_t\})), -E(N_1(\lambda_2; \{a_t\})))$, where $E(\cdot)$ is the expectation operator. For the sake of convenience, we omit $\{a_t\}$ and write $N_s(\cdot)$ for $N_s(\cdot, \{a_t\})$.

(a) By Lemma 1 and the bounded convergence theorem,

$$E(N_1(\lambda_1)) = (\lambda_1 - \lambda_2)\left(\frac{1 - \lambda_1}{1 - \lambda_2}\right) \sum_{s=2}^{\infty} \lambda_1^{s-2} E(N_s(\lambda_2)) + \left(\frac{1 - \lambda_1}{1 - \lambda_2}\right) E(N_1(\lambda_2)). \tag{1}$$

Recall that player 2's stage payoffs are $-a_t$. Since player 2 can ignore his acquired information, he can ensure $-v(p;\ \lambda_2)$ at any stage. That is, $-E(N_s(\lambda_2)) \geq -v(p;\ \lambda_2)$ at any stage $s$. Therefore, $E(N_1(\lambda_1)) \leq v(p;\ \lambda_2)$ and (a) is proved.

(b) Since $v(p;\ \lambda_2)$ converges to $v(p;\ 1)$,

$$\limsup_{\omega\to 1} \max V^1(p;\ \lambda_1,\ \lambda_2) \leq v(p;\ 1).$$

On the other hand, from individual rationality, $\min V^1(p;\ \lambda_1,\ \lambda_2) \geq v(p;\ \lambda_1)$, and $\liminf_{\omega\to 1} \min V^1(p;\ \lambda_1,\ \lambda_2) \geq \lim_{\omega\to 1} v(p;\ \lambda_1) = v(p;\ 1)$. Hence, $\lim_{\omega\to 1} V^1(p;\ \lambda_1,\ \lambda_2) = v(p;\ 1)$.

(c) Let $K = \liminf_{\omega\to 1} ((1-\lambda_1)/(1-\lambda_2))$. By assumption $K > 0$. As in (a), $E(N_s(\lambda_2)) \leq v(p;\ \lambda_2)$ for stages $s \geq 2$. Plugging in (1),

$$\begin{aligned} E(N_1(\lambda_1)) &\leq (\lambda_1 - \lambda_2)\left(\frac{1-\lambda_1}{1-\lambda_2}\right)\sum_{s=2}^{\infty} \lambda_1^{s-2} v(p;\ \lambda_2) + \left(\frac{1-\lambda_1}{1-\lambda_2}\right) E(N_1(\lambda_2)) \\ &= \left(1 - \frac{1-\lambda_1}{1-\lambda_2}\right) v(p;\ \lambda_2) + \left(\frac{1-\lambda_1}{1-\lambda_2}\right) E(N_1(\lambda_2)). \end{aligned} \tag{2}$$

Let $\{\omega_n\}$ be a sequence between 0 and 1 that converges to 1 and satisfies $\lim_{\omega_n}((1-\lambda_1)/(1-\lambda_2))\ E(N_1(\lambda_2)) = \liminf_{\omega\to 1} ((1-\lambda_1)/(1-\lambda_2))\ E(N_1(\lambda_2))$. Taking lim inf of both sides of (2) and using individual rationality of player 1 one obtains,

$$\begin{aligned} v(p;\ 1) &\leq \liminf_{\omega\to 1} E(N_1(\lambda_1)) \leq \liminf_{\omega_n} E(N_1(\lambda_1)) \\ &= \liminf_{\omega_n}\left(1 - \frac{1-\lambda_1}{1-\lambda_2}\right) v(p;\ \lambda_2) + \lim_{\omega_n}\left(\frac{1-\lambda_1}{1-\lambda_2}\right) E(N_1(\lambda_2)) \\ &= \liminf_{\omega_n}\left(1 - \frac{1-\lambda_1}{1-\lambda_2}\right) v(p;\ \lambda_2) + \liminf_{\omega\to 1}\left(\frac{1-\lambda_1}{1-\lambda_2}\right) E(N_1(\lambda_2)) \\ &\leq \limsup_{\omega\to 1}\left(1 - \frac{1-\lambda_1}{1-\lambda_2}\right) v(p;\ \lambda_2) + \liminf_{\omega\to 1}\left(\frac{1-\lambda_1}{1-\lambda_2}\right) E(N_1(\lambda_2)) \\ &\leq (1-K) v(p;\ 1) + K \limsup_{\omega\to 1} E(N_1(\lambda_2)). \end{aligned}$$

Thus, $\limsup_{\omega\to 1} E(N_1(\lambda_2)) \geq v(p;\ 1)$. The individual rationality of player 2 now implies the desired. □

Theorem 1 has an application to the comparison of repeated games where players' discount factors remain fixed, but the time lap between consecutive stages converges to zero. An equivalent way to say that the frequency of interactions grows to infinity is to say that the discount factors converge to 1 along a discounting path where $(\ln(\lambda_1))/(\ln(\lambda_2))$ stays constant. Thus, if player 1 is more patient and the interactions become infinitely frequent, then $(\ln(\lambda_1))/(\ln(\lambda_2))$ is a constant greater than 1 and by (b) and (c) of Theorem 1, the equilibrium payoffs converge to the value of the undiscounted game.

**3.2 The impact of the speed of convergence.** The proof of Theorem 1 suggests that the speed of convergence of $V^1(p; \lambda_1, \lambda_2)$ to the value $v(p; 1)$, depends on the speed of convergence of $v(p; \lambda)$ to $v(p; 1)$, as $\lambda$ tends to 1. If this speed is quick enough relative to $(1 - \lambda_2)/(1 - \lambda_1)$, then $V^2(p; \lambda_1, \lambda_2)$ converges to $-v(p; 1)$, because the lower and upper bounds of the elements in $V^2(p; \lambda_1, \lambda_2)$ become close.

Denote $\Phi(p; \lambda) = v(p; \lambda) - v(p; 1)$. Thus, $\lim_{\lambda\to 1} \Phi(p; \lambda) = 0$ and, moreover, for any $(p, \lambda) \in P \times [0, 1]$, $\Phi(p; \lambda) \geq 0$ (see Mertens, Sorin, and Zamir 1994).

THEOREM 2. *Let* $(\lambda_1, \lambda_2)$ *be a discounting path with* $\lambda_1(\omega) \geq \lambda_2(\omega)$ *for every* $\omega \in [0, 1]$, *where* $\lambda_2(\omega) \to 1$. *Suppose that* $\lim_{\omega\to 1}((1 - \lambda_2)/(1 - \lambda_1))\, \Phi(p; \lambda_2) = 0$. *Then*, $\lim_{\omega\to 1} V^2(p; \lambda_1, \lambda_2) = -v(p; 1)$.

PROOF. Let $(\sigma, \tau)$ be an arbitrary Nash equilibrium in the game $G(p; \lambda_1, \lambda_2)$. By individual rationality, Theorem 1(a), and the definition of $\Phi(p; \lambda)$, one obtains,

(a) $v(p; 1) \leq v(p; \lambda_1) \leq E(N_1(\lambda_1))$;

(b) $E(N_s(\lambda_1)) \leq v(p; \lambda_2) = v(p; 1) + \Phi(p; \lambda_2)$, for any $s \geq 1$.

From Lemma 1 and Corollary 1,

$$E(N_1(\lambda_1)) - E(N_1(\lambda_2)) = (\lambda_2 - \lambda_1)\left(\frac{1-\lambda_2}{1-\lambda_1}\right)\sum_{s=2}^{\infty} \lambda_2^{s-2}[E(N_1(\lambda_1)) - E(N_s(\lambda_1))]$$

(by (a) and (b))

$$\leq (\lambda_2 - \lambda_1)\left(\frac{1-\lambda_2}{1-\lambda_1}\right)\sum_{s=2}^{\infty} \lambda_2^{s-2}[v(p; 1) - (v(p; 1) + \Phi(p; \lambda_2))]$$

$$= (\lambda_1 - \lambda_2)\left(\frac{1-\lambda_2}{1-\lambda_1}\right)\sum_{s=2}^{\infty} \lambda_2^{s-2}\Phi(p; \lambda_2) = \frac{\lambda_1 - \lambda_2}{1-\lambda_1}\,\Phi(p; \lambda_2)$$

$$= \left(\frac{1-\lambda_2}{1-\lambda_1} - 1\right)\Phi(p; \lambda_2) \xrightarrow[\omega\to 1]{} 0.$$

Therefore, $\limsup_{\omega\to 1} \max V^2(p; \lambda_1, \lambda_2) \leq -\lim_{\omega\to 1} V^1(p; \lambda_1, \lambda_2) = -v(p; 1)$. On the other hand, from individual rationality $\limsup_{\omega\to 1} \max V^2(p; \lambda_1, \lambda_2) \geq -\lim_{\omega\to 1} v(p; \lambda_2) = -v(p; 1)$. We conclude that $\lim_{\omega\to 1} V^2(p; \lambda_1, \lambda_2) = v(p; 1)$. □

We will show, in Example 2, that the condition of Theorem 2 cannot be improved upon.

REMARK 1. If the information structure is standard (after each stage the signals are the players' actions), then the speed of convergence is bounded as follows. $\Phi(p; \lambda) \leq C \sqrt{(1-\lambda)/(1+\lambda)}\, \Sigma \sqrt{p^k(1-p^k)}$, where $p^k$ is the initial probability of state $k$ and $C$ is a constant (see Mertens, Sorin, and Zamir 1994).

Denoting $C^*(\lambda) = (C/\sqrt{1+\lambda})\, \Sigma \sqrt{p^k(1-p^k)}$, we see that $(C/\sqrt{2})\, \Sigma \sqrt{p^k(1-p^k)} \leq C^*(\lambda) \leq C\, \Sigma \sqrt{p^k(1-p^k)}$ and $((1-\lambda_2)/(1-\lambda_1))\, \Phi(p; \lambda_2) \leq C^*(\lambda_2)(1-\lambda_2)^{3/2}/(1-\lambda_1)$. Hence, for standard information structure, Theorem 2 states that if $(1-\lambda_2)^{3/2}/(1-\lambda_1)$ converges to zero, then the asymptotic equilibrium payoffs are necessarily zero-sum. This implies, for instance, that for every $c < 1$, if $\lambda_1 = \lambda_2^c$, then all equilibrium payoffs converge to zero-sum payoffs.

**3.3 An example with non zero-sum equilibrium payoffs.** In the preceding section sufficient conditions were given for the convergence of $V^1(p; \lambda_1, \lambda_2)$ and $V^2(p; \lambda_1, \lambda_2)$ to the zero-sum value. In this section we give an example of discounting paths which do not satisfy the conditions of Theorems 1 and 2 and of Nash equilibrium payoffs that do not converge to the zero-sum value.

The idea of the following example is that the impatient player, here player 2, can force his opponent to play for a relatively long duration one of his "favorite" actions, starting at the beginning of the game. In case players' discount factors are sufficiently singular (in the sense that they violate the conditions of Theorems 1 and 2), this duration may be chosen so that the corresponding payoffs are significant to player 2 while they are practically negligible for player 1. Player 1 is forced to give up payoffs at the beginning of the game but is compensated later. Thus, he plays a player 2's "favorite" action in early stages without violating his own individual rationality: all his continuation payoffs are slightly higher than his individually rational level.

Player 2 can force player 1 to follow such a plan by introducing a punishing scheme, aimed at reducing the payoffs of player 1 to his individual rational level in case of defection. Thus, player 2 will be able to ensure himself a payoff strictly greater than the zero-sum value in games that have the following two properties: (a) player 1 can play an action that guarantees player 2 a payoff strictly higher than his individual rational level; and (b) player 1 can be ensured a payoff slightly higher than his individual rational level so player 2 can effectively punish him.

EXAMPLE 2. Suppose that there are two states of nature (i.e., $K = \{1, 2\}$), and that the distribution over $K$ is $p = (\frac{1}{2}, \frac{1}{2})$. The game matrices corresponding to $K$ are:

$$A^1 = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & \frac{1}{8} \\ 1 & 0 & 0 \end{pmatrix} \qquad A^2 = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & \frac{1}{8} \\ 0 & 1 & 1 \end{pmatrix}.$$

Assume that the information structure is such that the players observe, beyond their own action, only the payoffs received after each stage.

Denote player 1's actions by $\{T, I, B\}$ and player 2's actions by $\{L, M, R\}$. Note that the non-revealing mixed actions of player 1 are all the probability distributions over $\{T, I\}$ and so the value of the non-revealing game is 0. This implies that $v(p; 1) = 0$ (see Aumann and Maschler 1995, Sorin 1980, Mertens, Sorin, and Zamir 1994).

When both players have an identical discount factor, $\lambda$, an analogous equilibrium to that introduced in Example 1 can be constructed. Namely,

**Player 1:** Always play $T$.

**Player 2:** Play $L$ until $-1$ is detected. From then on, play $M$.

A similar calculation to that carried out in Example 1 shows that the value of the identically discounted game is: $v(p; \lambda) = \frac{1}{2}(1 - \lambda)$.

We will now deal with a more patient informed player. The play path of the strategies we are about to present consists of an initial block of $n$ stages followed by an infinite number of identical blocks containing $k + 2$ stages each. In each stage of the first block, player 2 will collect the payoff 1 (minus $-1$). Since $n$ is chosen large enough relative to $\lambda_2$, player 2's overall payoff is close to 1. In the subsequent $(k + 2)$-blocks, player 1 will receive non-negative payoffs so as to make his overall payoff positive and, therefore, individually rational.

During the construction we make sure that the continuation payoffs of player 1 are at least his individually rational level plus his potential one-time gain from any deviation. Likewise, player 2's continuation payoffs should satisfy similar conditions.

Consider the following strategies, which depend on the parameters $n$ and $k$ to be determined later.

**The Master Plan.** At the *first n stages* player 1 plays $T$ and player 2 plays $L$ and $M$ with probability of $\frac{1}{2}$ each. *From stage* $n + 1$ *and on* both play repeatedly the following block of length $k + 2$. Player 1 plays $I$ all the time, while player 2 plays $R$ at the first two stages and at the last $k$ stages he plays $L$ and $M$ with probability of $\frac{1}{2}$ each.

Thus, the play path of these strategies is:

$$\underbrace{\left(T,\frac{1}{2}*L + \frac{1}{2}*M\right),\ldots,\left(T,\frac{1}{2}*L + \frac{1}{2}*M\right),}_{n} \qquad (I, R), (I, R),$$

$$\underbrace{\left(I,\frac{1}{2}*L + \frac{1}{2}*M\right),\ldots,\left(I,\frac{1}{2}*L + \frac{1}{2}*M\right),}_{k} \qquad (I, R), (I, R),$$

$$\underbrace{\left(I,\frac{1}{2}*L + \frac{1}{2}*M\right),\ldots,\left(I,\frac{1}{2}*L + \frac{1}{2}*M\right),}_{k} \ldots$$

and the corresponding sequence of player 1's payoffs is:

$$\underbrace{-1,\ldots,-1,}_{n}\quad \frac{1}{8},\frac{1}{8},\quad \underbrace{0,\ldots,0,}_{k}\quad \frac{1}{8},\frac{1}{8},\quad \underbrace{0,\ldots,0}_{k}\quad \frac{1}{8},\frac{1}{8},\quad \underbrace{0,\ldots,0,}_{k}\quad \ldots.$$

If a defection is detected at stage $t$ (i.e., the payoff observed at stage $t$ does not equal to the $t$th element of the above sequence), the players start playing the following punishment plan.

**The punishment plan.** Player 1 plays $B$ forever. Player 2 plays $M$ until the payoff 1 is received and from then on he plays $L$.

Note that in case of no defection, player 1 plays a completely non-revealing strategy. Moreover, he receives an overall payoff slightly above zero, while player 2 receives almost 1.

The best opportunity of player 2 to deviate is immediately after the initial block. We will choose the parameter $k$ in a way that will render such a deviation unprofitable. Formally, $k$ should satisfy the following inequality.

$$\frac{1-\lambda_2}{1-\lambda_2^{k+2}}\left[-\frac{1}{8}+\lambda_2\left(-\frac{1}{8}\right)+\lambda_2^2\frac{1-\lambda_2^k}{1-\lambda_2}0\right] \geq (1-\lambda_2)0 + (1-\lambda_2)\lambda_2\left(-\frac{1}{2}\right)+\lambda_2^2 0.$$

The left side is the continuation payoff according to the prescribed master plan. On the right side, the $(1 - \lambda_2)0$ term is the evaluation of the payoff 0 (due to deviation) at period $n + 1$. The term $(1 - \lambda_2)\lambda_2(-\frac{1}{2})$ is the evaluation of 0 and $-1$ with probability $\frac{1}{2}$ each received at the post deviation stage (due to the punishment plan). Finally, the term $\lambda_2^2 0$ is the evaluation of the stream of zeros, determined by the punishment plan.

The above inequality holds if and only if $((1 + \lambda_2)/(1 - \lambda_2^{k+2}))\ \frac{1}{8} \leq \lambda_2\ \frac{1}{2}$, which is equivalent to

$$k \geq \frac{\ln(3\lambda_2 - 1) - \ln(4)}{\ln(\lambda_2)} - 3. \tag{3}$$

The parameter $n$ will be chosen large enough to make the first block substantial with respect to $\lambda_2$. For instance, we can choose $n$ so that the weight of the tail after the $n$th stage is less than $1 - \lambda_2$. In other words, we can choose $n$ satisfying $1 - \lambda_2^n > \lambda_2$, which is equivalent to

$$n \geq \frac{\ln(1 - \lambda_2)}{\ln(\lambda_2)}. \tag{4}$$

To summarize, condition (3) guarantees that any deviation of player 2 is unprofitable. Condition (4) makes player 2's payoffs close to $-1$.

We can now turn to the selection of the other parameters that ensure that any deviation of player 1 is also unprofitable. The parameters $n$ and $k$ impose two conditions that $\lambda_1$ must satisfy:

(a) In order to assure that player 1 has no incentive to defect right at the beginning of the game, $\lambda_1$ should be large enough, relative to the length of the first block, $n$, so as to make this block negligible for player 1's payoffs. More precisely, this means,

$$(1 - \lambda_1)\frac{1}{2} + \lambda_1 0 \leq -(1 - \lambda_1^n) + \lambda_1^n\left[\frac{1 - \lambda_1}{1 - \lambda_1^{k+2}}\left(\frac{1}{8} + \lambda_1\frac{1}{8} + \lambda_1^2\frac{1 - \lambda_1^k}{1 - \lambda_1}0\right)\right]$$

The left side is the evaluation of the best available deviation for player 1. The right side is the payoff generated by the master plan ($n$ times $-1$, 2 times $\frac{1}{8}$, $k$ times 0, . . .).

This inequality reduces to:

$$\lambda_1^n \geq \frac{((3 - \lambda_1)/2)}{(1 + ((1 - \lambda_1^2)/(1 - \lambda_1^{k+2}))\ \frac{1}{8})}. \tag{5}$$

(b) In order to assure that player 1 has no incentive to deviate during a $(k + 2)$-block, $\lambda_1$ must be large enough. We deal with two possible deviations separately.

(b1) Player 1 will not benefit from defecting during the first or second stages of each $(k + 2)$-block if and only if

$$\frac{1 - \lambda_1}{1 - \lambda_1^{k+2}}\left[\frac{1}{8} + \lambda_1\frac{1}{8} + \lambda_1^2\frac{1 - \lambda_1^k}{1 - \lambda_1}0\right] \geq (1 - \lambda_1)\frac{1}{2} + \lambda_1 0.$$

The left side is the continuation payoff, according to the prescribed strategies, at the beginning of any $(k + 2)$ block. The right side is the sum of the one-shot gain from deviation (i.e., $\frac{1}{2}$) and the evaluation of the zeros thereafter. This reduces to:

$$\lambda_1^{k+2} \geq 1 - \frac{1 + \lambda_1}{4}. \tag{6}$$

(b2) Player 1 will not benefit from defecting after the second stage of each $(k + 2)$-block if and only if

$$\lambda_1^k\left[\frac{1 - \lambda_1}{1 - \lambda_1^{k+2}}\left(\frac{1}{8} + \lambda_1^2\frac{1 - \lambda_1^k}{1 - \lambda_1}0\right)\right] \geq (1 - \lambda_1)\frac{1}{2} + \lambda_1 0.$$

This reduces to

$$\lambda_1^k \geq \frac{4(1-\lambda_1^{k+2})}{1+\lambda_1}. \tag{7}$$

To summarize, if the parameters are chosen in a way that satisfies inequalities (3)–(7), then the strategies defined form an equilibrium. It turns out that for every $\epsilon > 0$, one can choose $n$, $k$ and $\lambda_1 \geq 1 - (1-\lambda_2)^{2+\epsilon}$ so that (3)–(7) are all satisfied, even when $\lambda_2$ is fairly close to 1. In case $\lambda_2$ is bounded away from 1 and only $\lambda_1$ is converging to 1, choosing $n$ and $k$ that satisfy (3)–(7) is trivial.

REMARK 2. In Example 2, $\Phi(p; \lambda) = v(p; \lambda) - v(p; 1) = \frac{1}{2}(1-\lambda)$. Thus, the condition of Theorem 2 can be written as

$$\frac{(1-\lambda_2)^2}{(1-\lambda_1)} \to 0. \tag{8}$$

Example 2 attests to the fact that the condition of Theorem 2, which ensures asymptotic zero-sum payoff, cannot be replaced by any weaker condition of the form $\lim_{\omega\to 1}((1-\lambda_2)/(1-\lambda_1))\,\psi(p; \lambda_2) = 0$, where $\psi(p; \lambda_2)/\Phi(p; \lambda_2)(1-\lambda_2)^{\epsilon} \to 1$ for some fixed $\epsilon > 0$. This is so because the discounting path $\lambda_1 = 1 - (1-\lambda_2)^{2+\epsilon}$ does not satisfy (8) for every $\epsilon > 0$ and with an appropriate choice of $n$ and $k$, it satisfies inequalities (3)–(7). Therefore, it entails a **non** zero-sum asymptotic equilibrium payoffs.

By construction, $\limsup \max V^2(\frac{1}{2}; \lambda_1, \lambda_2) = 1 > -v(\frac{1}{2}; 1) = 0$. In particular, $V^2(p; \lambda_1, \lambda_2)$ does not converge to the set $-\{v(p; 1)\}$. Since $\lambda_1 > \lambda_2$, by Theorem 1, $\lim V^1(p; \lambda_1, \lambda_2) = v(p; 1) = 0$. We therefore obtain

PROPOSITION 1. *There exists a game in which*

$$\limsup_{\omega\to 1}\{\nu^1 + \nu^2 \mid \nu^i \in V^i(p; \lambda_1, \lambda_2)\} = v(p; 1) + m,$$

*where $m$ is the maximal player 2's payoff in the game.*

Notice that $\limsup V^1(p; \lambda_1, \lambda_2) = v(p; 1)$ and that $\limsup V^2(p; \lambda_1, \lambda_2) \leq m$. Therefore, the summation of the players' payoffs, as indicated in Proposition 1, cannot be greater.

**4. A more patient uninformed player.** We turn now to the case in which the discounting path is such that the uninformed player, player 2, is more patient. The first result is valid for any information structure.

PROPOSITION 2. *Let $(\lambda_1, \lambda_2)$ be a discounting path. If $\lambda_1(\omega) \leq \lambda_2(\omega)$ for every $\omega \in [0, 1]$, then $\lim_{\omega\to 1}((1-\lambda_1(\omega))/(1-\lambda_2(\omega))) = 1$ implies*

$$\lim_{\omega\to 1} V^1(p; \lambda_1, \lambda_2) = -\lim_{\omega\to 1} V^2(p; \lambda_1, \lambda_2) = v(p; 1).$$

In other words, if the players grow infinitely patient in a similar manner, then the game becomes almost completely competitive. In particular, the payoff of any Nash equilibrium converges to the zero-sum value.

PROOF OF PROPOSITION 2. Using the notation introduced in the proof of Theorem 1 and the representation established in Lemma 1, we obtain that the coefficient of $N_1(\lambda_2)$ in the representation of $N_1(\lambda_1)$, which is $(1-\lambda_1)/(1-\lambda_2)$ (see Lemma 1), converges to 1.

Proposition 2 follows now from the individual rationality of both players. □

In order to get more symmetric results than those achieved when the informed player is more patient, we use the following observation made in Megiddo (1980) (see also Mertens, Sorin, and Zamir 1994).

PROPOSITION 3. *In games with incomplete information on one side where the uninformed player, player* 2, *is told his payoff at each stage,* $v(p; 1)$ *is linear over* $P$.

When player 2 observes his own payoffs we obtain,

THEOREM 3. *In games where at each stage player* 1 *is told the actions taken by both players* (*standard information*) *and player* 2 *knows his own action and payoff, if* $(\lambda_1, \lambda_2)$ *is a discounting path, where* $\lambda_1(\omega) \leq \lambda_2(\omega)$ *for every* $\omega = [0, 1]$, *and* $\lambda_1(1) = 1$, *then*

(a) $\lim_{\omega \to 1} V^2(p; \lambda_1, \lambda_2) = -v(p; 1)$.

*(b) If* $(1 - \lambda_2(\omega))/(1 - \lambda_1(\omega))$ *is bounded away from zero, then* $\lim_{\omega \to 1} V^1(p; \lambda_1, \lambda_2) = v(p; 1)$.

PROOF. The proof resembles the proof of Theorem 1. (a) Let $(\sigma, \tau)$ be a Nash equilibrium of $G(p; \lambda_1, \lambda_2)$. At any stage $s \geq 1$, the induced posterior probabilities $q_1^s, q_2^s, \ldots, q_{ns}^s$, corresponding to histories $h_1^s, \ldots, h_{ns}^s$, can be calculated by player 1 since he observes the actions taken by player 2 at each stage and can thus deduce player 2's signals. From individual rationality of player 1, $E(N_s(\lambda_1)) \geq \sum_{j=1}^{ns} \text{Prob}\,(h_j^s) v(q_j^s; \lambda_1)$, for all $s \geq 2$, where Prob $(\cdot)$ is the probability induced by $(\sigma, \tau)$. Using Lemma 1, we get an analogue of (1) for player 2. Plugging in the above set of inequalities we get,

$$E(N_1(\lambda_2)) \geq (\lambda_1 - \lambda_2)\left(\frac{1 - \lambda_2}{1 - \lambda_1}\right) \sum_{s=2}^{\infty} \lambda_1^{s-2} \sum_{j=1}^{n_s} \text{Prob}(h_j^s) v(q_j^s; \lambda_1) + \left(\frac{1 - \lambda_2}{1 - \lambda_1}\right) v(p; \lambda_1).$$

As $v(p; \lambda_1)$ converges to $v(p; 1)$ uniformly on $P$, for every $\epsilon > 0$ and sufficiently large $\lambda_1$

$$E(N_1(\lambda_2)) \geq (\lambda_1 - \lambda_2)\left(\frac{1 - \lambda_2}{1 - \lambda_1}\right) \sum_{s=2}^{\infty} \lambda_1^{s-2} \sum_{j=1}^{n_s} \text{Prob}(h_j^s)(v(q_j^s; 1) - \epsilon)$$
$$+ \left(\frac{1 - \lambda_2}{1 - \lambda_1}\right)(v(p; 1) - \epsilon).$$

By Corollary 1 and Proposition 3, the right hand side equals $v(p; 1) - \epsilon$. Combined with individual rationality of player 2, this implies (a).

(b) Identical to the proof of Theorem 1(c). □

**4.2 An example with non zero-sum equilibrium payoffs.** In cases where player 2, the uninformed player, is more patient than the informed player, he can define an initial short learning phase. In this phase player 2 will use actions that maximize his learning. When players' discount factors are singular enough, this phase may be chosen to be significant for player 1 and insignificant for player 2. As a result, player 1 will be forced to reveal information. If the learning phase is long enough for player 1, the game played after the learning phase is similar to a complete information game. This means that player 2's payoff after the learning phase is very close to his individually rational level in the complete information game, which is typically higher than that of the incomplete information game. In short, player 1 will use his information at the beginning of the game and player 2 will use the revealed information afterwards. In so doing, both players may receive strictly more than the value of the undiscounted game. This intuition is the basis for the following example.

EXAMPLE 3. Consider the game in which the state space is $K = \{1, 2\}$, the distribution over $K$ is $p = (\frac{1}{2}, \frac{1}{2})$, and the payoffs are given by the following matrices.

$$A^1 = \begin{pmatrix} 1 & 0 & 4 \\ 0 & 0 & 4 \end{pmatrix} \qquad A^2 = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 1 & 4 \end{pmatrix}$$

and the information structure is defined by the common signaling matrix

$$\begin{pmatrix} a & a & c \\ b & b & c \end{pmatrix}.$$

Let player 1's set of actions be $\{T, B\}$ and player 2's set of actions be $\{L, M, R\}$. Suppose that the players are informed after every stage of their own action and of the common signal. The information structure is such that player 1 learns, after each stage whether player 2 has played $R$ or one out of $\{L, M\}$ and can never differentiate between $L$ and $M$. Player 2, on the other hand, can perfectly monitor player 1's actions only if he plays $L$ or $M$.

As the right column is dominated in both game matrices, it is easy to see that the value of the non-revealing game, for every distribution $p = (p, 1 - p)$ over $K$, is $p(1 - p)$ and hence $v(p; 1) = p(1 - p)$. In our particular case (i.e., $p = \frac{1}{2}$), $v(p; 1) = \frac{1}{4}$ (see Mertens, Sorin and Zamir 1994).

REMARK 3. (a) Any Nash equilibrium in which player 2 uses with positive probability only actions in $\{L, M\}$, depends only on $\lambda_1$. That is, $G(p; \lambda_1, \lambda_2)$ depends only on $\lambda_1$. The reason is that as long as player 2 plays only $L$ or $M$, player 1 collects no information about the actions of player 2. Thus, the best response of player 2, regardless of his discount factor, is to maximize the expected payoff of the next period. It follows that, $V^1(p; \lambda_1, \lambda_2) = v(p; \lambda_1)$ for every $p$, $\lambda_1$ and $\lambda_2$.

(b) It is easily verified that for every $\epsilon > 0$, $\lambda_1 \in [0, 1)$, and $p$ if player 1 uses an optimal strategy in $G(p; \lambda_1)$, then there exists $S \in \mathbb{N}$ such that for every stage $s \geq S$ and every state $k$, $E(|p(s) - \mathbb{1}_{\{k=1\}}|) < \epsilon$, where $p(s)$ is the posterior probability of the state 1 at time $s$ and $\mathbb{1}$ is the characteristic function. That is, an optimal strategy of player 1 is asymptotically revealing the true state. In fact, this follows from the representation of $v(p; \lambda)$ introduced in Mayberry (1967). His general formula implies that in our case when $p \geq \frac{1}{2}$,

$$v(p; \lambda) = \max_{0 \leq x \leq 1-p} \left( \lambda x + (1 - \lambda)\left( (1 - p)v\left( \frac{1 - p - x}{1 - p}; \lambda \right) + pv\left( \frac{x}{p}; \lambda \right) \right) \right) \tag{9}$$

(see Mertens, Sorin and Zamir 1994, p. 308).

The best completely non-revealing strategy at $p$ is to play with probability $p$ the top action (in both $A^1$ and $A^2$). This action corresponds to $x = p(1 - p)$. By differentiating the argument in the right side of (9), $\lambda x + (1 - \lambda)((1 - p)v((1 - p - x)/(1 - p); \lambda) + pv(x/p; \lambda))$, with respect to $x$, one can see that there are better choices of $x$ than the one corresponding to the non-revealing strategy, $p(1 - p)$. It means that the best action of player 1 at the first stage of the game with any $p < 1$ is partially revealing. It follows that the posterior probabilities cannot converge to any $0 < p < 1$. They converge to $p = 1$ in case the realized game is $A^1$ and to $p = 0$ when the realized the game is $A^2$.

For a fixed $\omega \in [0, 1)$, consider the following strategies for both players. The parameter $n$ will be determined later.

**The Master Plan.** *At the first n stages* player 1 plays $T$ and player 2 plays $R$. *From stage*

$n + 1$ *and on* each player plays an optimal strategy in $G(\frac{1}{2}; \lambda_1)$. Note that from stage $n + 1$ and on player 2 uses only the actions $L$ and $M$.

**The Punishment Phase.** If a defection is detected (by player 1) during the first $n$ stages of the game, player 1 will play $T$ from then on (completely non-revealing) and the average payments from that stage on will be $\frac{1}{4}$.

Fix an $\epsilon > 0$. We choose $n$ in a way that ensures that the first $n$ stages of the game, in which the payoff is 4, have a significant effect on the total payoff of player 1. Precisely, we choose $n$ so that $\lambda_1^n < \epsilon$.

Using part (b) of Remark 3, we may choose $S \in \mathbb{N}$ such that for every stage $s \geq S$ and every state $k$, $E(|p(s) - \mathbb{1}_{\{k=1\}}|) < \epsilon$, where $p(s)$ is the posterior probability of state 1 after stage $s$. We can now choose $\lambda_2$ large enough to satisfy $4(1 - \lambda_2^{n+S}) < \epsilon$. Thus, the initial block of $n$ stages affects player 2's overall payoffs by less than an $\epsilon$. Clearly, for $\lambda_2$ large enough, player 2 will have no incentive to defect. We conclude that the above strategies constitute a Nash equilibrium in $G(\frac{1}{2}; \lambda_1, \lambda_2)$ for $\epsilon$ small enough.

Player 1's payoff corresponding to this Nash equilibrium is greater than $(1 - \epsilon)4$, while player 2's payoff is greater than $-4\epsilon - \epsilon = -5\epsilon$. Thus, **both** players do better than the zero-sum value.

We complete this section by stating:

PROPOSITION 4. *There exists a game in which*

$$\limsup_{\omega \to 1} \{\nu^1 + \nu^2 | \nu^i \in V^i(p; \lambda_1, \lambda_2), i = 1, 2\} = M + m,$$

*where $M$ and $m$ are the maximal payoffs of players* 1 *and* 2, *respectively.*

## 6. Concluding comments.

**6.1 A non zero-sum limit with standard information.** Example 2 may be modified to the case of standard information (instead of observable payoffs) by replacing the payoff matrices with the following:

$$A^1 = \begin{pmatrix} -1 & -1 & -1 & -1 & -1 \\ 1 & 0 & 1/2 & 1/4 & 1/4 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

$$A^2 = \begin{pmatrix} -1 & -1 & -1 & -1 & -1 \\ 0 & 1 & 1/2 & 1/4 & 1/4 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

The three left most columns are dominated in the non-revealing game (i.e., the mixture of the two matrices). Using similar calculations to those introduced in the beginning of Example 3, it is easily seen that $v(p; 1) = \frac{1}{4}$.

The sequence of payoffs corresponding to an equilibrium, similar to the completely non-revealing one introduced in Example 2, is

$$\underbrace{-1, \ldots, -1}_{n}, \frac{1}{2}, \underbrace{\frac{1}{4}, \ldots, \frac{1}{4}}_{k}, \frac{1}{2}, \underbrace{\frac{1}{4}, \ldots, \frac{1}{4}}_{k}, \frac{1}{2}, \underbrace{\frac{1}{4}, \ldots, \frac{1}{4}}_{k}, \frac{1}{2}, \underbrace{\frac{1}{4}, \ldots, \frac{1}{4}}_{k}, \ldots.$$

Defection leads player 1 to use his two bottom actions, and thereby reducing player 2's payoffs to at most $-\frac{1}{4}$. Player 2 punishes player 1 in a case of deviation by using his two rightmost actions. This guarantees that player 1's payoff is close to $\frac{1}{4}$ for discount factors large enough.

Just like in Example 2, an appropriate choice of $n$, $k$, and a discounting path $(\lambda_1, \lambda_2)$ results in $\limsup_{\omega\to 1} \max V^2(\frac{1}{2}; \lambda_1, \lambda_2) = 1 > -v(\frac{1}{2}; 1) = -\frac{1}{4}$. Thus, $V^2(p; \lambda_1, \lambda_2)$ does not necessarily converge to the singleton $\{-v(p; 1)\}$.

**6.2 More on observable payoffs.** In Example 2 the values of both matrices were identical and the strategies used were non-revealing. However, when $\text{val}(A^k)$ are not all identical, a completely non-revealing Nash equilibrium, i.e., one in which the posterior at each stage is equal to the initial prior with probability 1, cannot be constructed. This is so because a completely non-revealing equilibrium must satisfy $E(N_s(\lambda_1)) \geq \max \text{val}(A^k)$ for every $s$ (otherwise, player 1 would have an incentive to defect at some stage $s$ and in some state $k$). On the other hand, for any equilibrium $E(N_s(\lambda_2)) \leq v(p; \lambda_2) \to_{\omega\to 1} v(p; 1) = \sum_k p^k \text{ val } (A^k) < \max \text{val}(A^k)$. The last strict inequality holds because in the case of observable payoffs, $v(p; 1) = \sum_k p^k \text{ val}(A^k)$ (see Proposition 3). Using the representation given in Lemma 1 for $\lambda_1 \geq \lambda_2$ and $\lambda_2$ large enough, we get a contradiction.

The authors do not know whether it is possible, in the case where all the values (of the state matrices) are distinct and when the payoffs are observable, to construct an example with $\limsup_{\omega\to 1} \max V^2(p; \lambda_1, \lambda_2) > -v(p; 1)$.

**6.3 On the sequence $\{v(p; \lambda)\}$.** Suppose $\lambda_1 \geq \lambda_2$ are two discount factors. If $(\sigma, \tau)$ is an arbitrary Nash equilibrium in $G(p; \lambda_1, \lambda_2)$, then $v(p; \lambda_1) \leq E(N_1(\lambda_1)) \leq v(p; \lambda_2)$. In particular, for all $p$, the sequence $\{v(p; \lambda)\}$ is monotonically non-increasing with respect to $\lambda$.

**Acknowledgments.** The authors are grateful to the two anonymous referees and the associate editor of *Mathematics of Operations Research.*

## References

Aumann, R.J., M. Maschler 1995. *Repeated Games with Incomplete Information.* The MIT Press, Cambridge, MA.

Celentani, M., D. Fudenberg, D. Levine, W. Pesendorfer 1996. Maintaining a reputation against a long-lived opponent. *Econometrica* **64** 691–704.

Cripps, M., J.P. Thomas 1995. The folk theorem in repeated games of incomplete information, D.P.

Evans, R., J.P. Thomas 1997. Reputation and experimentation in repeated games with two long-run players. *Econometrica* **65** 1153–1174.

Fudenberg, D., D. Levine 1989. Reputation and equilibrium selection in games with a patient player. *Econometrica* **57** 759–778.

———, ——— 1992. Maintaining a reputation when strategies are imperfectly observed. *Rev. Econ. Stud.* **59** 561–589.

Lehrer, E., A. Pauzner 1995. Repeated games with differential time preferences. Tel Aviv University Discussion Paper.

Mayberry, J.-P. 1967. Discounted repeated games with incomplete information. Report of the U.S. Arms Control and Disarmament Agency ST-116, Washington D.C., Chapter 5. 435–461.

Megiddo, N. 1980. On repeated games with incomplete information played by non-Bayesian players. *Internat. J. Game Theory* **9** 157–167.

Mertens, J.-F., S. Sorin, S. Zamir 1994. Repeated games. CORE Discussion Papers 9420–9422, Universite Catholique de Louvain, Louvain-la-Neuve, Belgium.

Schmidt, K. 1993. Reputation and equilibrium characterization in repeated games with conflicting interests. *Econometrica* **61** 325–351.

Sorin, S. 1980. An introduction to two-person zero-sum repeated games with incomplete information. Technical Report No. 312, IMSSS-Economics, Stanford University, Stanford, CA.

——— 1997. Merging, reputation and repeated games with incomplete information. mimeo.

E. Lehrer: School of Mathematical Sciences, Sackler Faculty of Exact Sciences, Tel Aviv University, Ramat Aviv 69978, Israel; e-mail: lehrer@math.tau.ac.il

L. Yariv: Department of Economics, Harvard University, Cambridge, Massachusetts 02138, USA; e-mail: yariv@fas.harvard.edu