# Communication in Repeated Games with Costly Monitoring

Elchanan Ben-Porath[1] and Michael Kahneman

January, 2002

[1]The department of Economics and the Center for Rationality, the Hebrew University of Jerusalem, and the Eitan Berglas School of Economics, Tel-Aviv University.  Correspondence: Elchanan Ben-Porath, the Center for Rationality, Feldman building, the Hebrew University, Givaat-Ram, Jerusalem, Israel. e-mail: benporat@math.huji.ac.il  Fax: 972-2-6513681.

**Abstract**

We study repeated games with discounting where perfect monitoring is possible, but costly. It is shown that if players can make public announcements, then every payoff vector which is an interior point in the set of feasible and individually rational payoffs can be implemented in a sequential equilibrium of the repeated game when the discount factor is high enough. Thus, efficiency can be approximated even when the cost of monitoring is high, provided that the discount factor is high enough.

**Key words: Repeated games, costly monitoring, communication.**

# 1   INTRODUCTION

The basic results in the theory of repeated games (Aumann and Shapley, 1976; Fudenburg and Maskin, 1986, and Rubinstein, 1979) establish that every individual rational payoff, (that is, a feasible payoff which is above the minmax point), is a perfect equilibrium payoff. In particular, repetition can enable players to obtain efficient outcomes in equilibrium. An important assumption that underlies these results is that each player can costlessly observe the past actions of the other players. Here we assume that perfect monitoring is possible but costly. It is shown that if players can make public announcements then, under a standard assumption of full dimensionality of the payoff set, the limit set of sequential equilibrium payoffs contains the set of individual rational payoffs. Thus, efficiency can be approximated even when the cost of monitoring is high, provided that the discount factor is high enough. In particular, as the discount factor tends to 1, monitoring can be more sparse so that the average cost of monitoring per period tends to zero.

The setup is as follows: The starting point is an $n$-person one-shot game $G$. We now consider an extension of $G, \bar{G}$, in which each player $i$ can observe the actions of some set of players $Q$ at a cost $C_i(Q)$. Formally, we associate with each action $a_i$ for player $i$ and any set of players different from $i$, $Q$, an action $(a_i, Q)$. When $(a_i, Q)$ is played, player $i$ observes the actions that were played by members of $Q$ and her payoff is her payoff in $G$ minus the cost of observation $C_i(Q)$.[1] The only assumption we make on $C_i(Q)$ is that it is increasing, so if $Q' \subseteq Q$ then $C_i(Q') \leq C_i(Q)$.[2] There are no restrictions on the costs of observing, so $C_i(Q)$ can be any finite non-negative number. The game $\bar{G}$ is played repeatedly and we assume that at the end of each period, players can make public announcements. The payoff in the repeated game is a discounted sum of the one-shot payoffs. We are interested in the behavior of the set

---

[1] To obtain our result we only need to assume that player $i$ can observe one other player at each period.

[2] This assumption is in fact not necessary but it makes the exposition simpler.

$V(\bar{G})$ which is the limit of the set of sequential equilibrium payoffs when the discount factor tends to 1. The result is that $V(\bar{G})$ contains the set of feasible and individually rational payoffs of the original game $G$ provided that the latter set is $n-$dimensional.

To keep the exposition simple, we assume that player $i$ observes only the action of the members of $Q$ and, of course, her own actions, but nothing else. However, our result can be immediately extended to the case where each player can observe some public, or even private, signal for free, provided that a standard assumption of common support is satisfied.[3] This extension is important because there are games of economic interest in the literature which have this structure. One example is partnership games which have been studied by Radner (1986) and Radner, Myerson and Maskin (1986). In these games, each worker/partner chooses a level of effort and the profile of efforts determines a distribution on the joint production. The joint production level is a public signal that is observed (for free) by everyone. When low production is observed it is not clear if one of the players has deviated or if it is just an unfortunate outcome. Radner, Myerson and Maskin provide an example which demonstrates that this problem may cause equilibrium payoffs to be bounded away from efficiency. Our result implies that if it is possible to observe the effort of a given worker for some cost, (which can be arbitrarily high), then efficiency can be approximated as the discount factor tends to 1.

Another celebrated example is repeated Cournot games with imperfect monitoring. In these games the quantity that each firm produces is not observed. The only observable parameter is the price of the product which is a stochastic function of the aggregate production. Green and Porter (1984) and Abreu, Pearce and Stachetti (1986) have shown that

---

[3]A signal function is a stochastic function that associates with each profile of actions of the players a probability distribution over a set of signals. The common support assumption is that all the probability distributions that are associated with the different action profiles have the same support.

in symmetric games optimal equilibria have a simple structure and that efficiency cannot be obtained. Again, our result implies that if monitoring is possible, then even if its cost is high, efficiency can be approximated.

Most of the literature on games with imperfect monitoring has focused on games with **public** signals.[4] For example, in the partnership game the signal that each player observes is the joint production level. In the Cournot game it is the price. In most real interactions players observe private signals. For example, a real situation which exhibits the problem modeled by the partnership game will almost surely have each worker observing some private signal of the effort invested by other workers. The study of repeated games with private monitoring has been relatively limited. Lehrer has studied two person games where payoffs are evaluated by the limit of the means (Lehrer 1990, 1992a, and 1992b). Fudenberg and Levine (1991) study N-person games with discounting but use an $\epsilon$-sequential equilibrium notion, which is a significantly weaker notion than sequential equilibrium. The closest studies to the current one are Compte (1998) and Kandori and Matsushima (1998). Kandori and Matsushima examine, as do we, the limit behavior of the set of sequential equilibrium payoffs in a repeated game with discounting. They assume, as it is done here, that players can make public announcements and that the signals that players observe have a common support. They then provide two conditions on the signals which ensure that every individual rational payoff belongs to the limit set of equilibrium payoffs. The first condition is that if either player $i$ or player $j$ deviates, then the other players, putting their information together, can statistically detect that. The second condition is that, given that player $i$ or player $j$ have deviated, the other players, (putting their information together), can statistically detect which one of them it was. Compte obtains a similar result using different

---

[4]General studies of such games can be found in Abreu, Pearce and Stachetti (1990), Fudenberg, Levine and Maskin (1994), and Lehrer (1990b).

assumptions on the signals. Thus, the relationship of these two papers to the current study is that the assumption of costly perfect monitoring is substituted by assumptions on the informativeness of the signals.

The assumptions that underlie the basic Falk Theorem can be decomposed as follows:

(1)    Each player can observe the action of every other player.

(2)    Each observation is a perfect observation.

(3)    Each observation is costless.

The current paper relaxes the third assumption.

In Ben-Porath and Kahneman (1996) we relaxed the first assumption and showed that if each player can be observed by at least **two** other players then the limit set of sequential equilibrium payoffs contains the set of individual rational payoffs[5].   In the equilibria that we construct in the current paper we do not need to assume that a given player is observed by more than two players. Since the construction here applies to the case where the cost of monitoring is zero as well, the result in the current paper extends the result in the previous paper for discounted games.  Furthermore, in the equilibria that we construct here each player is observed by just one other player in any given period.  (However, in case of a deviation the monitor of a given player $i$ is substituted, so we do need at least two players who have the ability to observe $i$.)

The equilibria that we will define are somewhat involved so it is worthwhile to list the main ideas in the construction. We consider first the case of three or more players and then comment on the case of two players. The first observation is that to obtain efficiency monitoring, which is costly, should be sparse. This immediately implies that the monitoring should be random and that the probability of being monitored in each particular period is

---

[5]In the first paper we also showed that if payoffs are evaluated by the $\liminf$ of the means then it is possible to cooperate even when each player can be monitored by only **one** other player.

low. Now, when the discount factor is close to one it is possible to inflict harsh punishments and therefore it is possible to deter a player from deviating even when the probability of detection is low. Next, we note that since the act of observing is costly and is itself unobserved there is a need to provide incentives to do it. So if player $j$ monitors player $i$ in period $t$ then player $j$ should be uncertain about the action of player $i$ in period $t$. To deal with these issues we propose an equilibrium which has the following general structure: Time is divided into segments of $K$ periods. Each player $i$ has a player which monitors him, $q(i)$. At the beginning of a segment player $q(i)$ chooses at random a period $t_1$ in the segment in which she will observe player $i$, and player $i$ chooses at random a period $t_2$[6] in which she 'switches' and plays an action which is different from the action that is specified by the main path. At the end of the segment player $i$ publicly announces the period $t_2$ and her action in that period and player $q(i)$ announces the period $t_1$ and her observation of the action of player $i$. The two announcements are made simultaneously. If the announcements are incompatible then one of the players has deviated. Conversely, if one of the players deviates - player $q(i)$ by guessing the action of player $i$ rather than observing it, or player $i$ by departing from the main path more than once - then with a probability of at least $\frac{1}{K}$, ($K$ is the number of periods in a segment), there will be a mismatch in the announcements. Thus, the way to deter players from deviating is to initiate a sufficiently severe punishment whenever there is a mismatch in the announcements.[7] Since it is impossible to tell which one of the two players,

---

[6] In fact player $i$ chooses two periods but the reason for that is not important.

[7] The idea of checking whether the player $q(i)$ has monitored player $i$ by having player $i$ switch from the main path is related to an idea which appears in Lehrer (1992a). Specifically, let $a_i$ and $\hat{a}_i$ be two actions for player $i$ such that $\hat{a}_i$ is less informative than $a_i$, that is, there exist actions $a_j$ and $\bar{a}_j$ for player $j$ such that if player $i$ plays $a_i$ she can distinguish between $a_j$ and $\bar{a}_j$, (i.e., she gets a different signal in each case), while if she plays $\hat{a}_i$ she gets the same signal. In the equilibrium which Lehrer constructs a deviation from $a_i$ to $\hat{a}_i$ is deterred in the following way: Player $j$ picks at random periods in which she randomizes between $a_j$ and $\bar{a}_j$. Player $i$ is then "asked" to reveal his signals in these periods. If he gives a wrong answer he is

6

$i$ or $q(i)$, has deviated both of them are punished. The problem in punishing two players is that in general it is impossible to punish the same pair of players $(i, q(i))$ repeatedly. So if $(i, q(i))$ continue to make contradicting statements repeatedly there will be a period in which it will not be possible to further decrease the payoffs of **both** of them. The reason for this difficulty is that in general the minmax profile of actions for player $i$ is different from the minmax profile for player $q(i)$ so it is impossible to minmax both players at the same time. The way to deal with this problem is to change the monitoring assignment $q(\ )$ each time that a mismatch of announcements occurs so that a new mismatch of announcements will involve a pair of players that is different from the last pair that was punished. If, for example, player $i$ and her monitor $k$ contradict each other, then the equilibrium moves to a phase in which $i$ and $k$ are punished and the new monitoring assignment is such that $i$ and $k$ do not monitor each other. So if there will be a new case of contradicting statements in this punishment phase it will involve a pair of players different from $(i, k)$. In this manner the same pair of players is not punished twice in a row.

In the case of two players each player must monitor the other all along the game so the construction has to be modified. The new feature here is that there are subgames in which the harshness of the punishment that can be implemented is limited and therefore there is a need to increase the density of the monitoring so that the probability of detection in case of a deviation will be high enough.

At this point the reader might wonder why we chose the $K-$periods segment structure rather than work with simpler strategies in which each player $i$ is monitored each period with a small probability. The problem with these simpler strategies is that to have player $j$ randomize between monitoring and not monitoring player $j$ has to be indifferent between these two options which means that she has to be compensated in case she monitors. Now

---

punished.

the point is that if player $j$ has to monitor player $i$ in a long sequence of consecutive periods, (such an event happens with a positive probability), then because of the discounting the compensation that player $j$ should receive in the subgame that follows is higher than what is feasible.[8] This problem of infeasible compensation does not arise in the equilibrium that we have outlined because in this equilibrium the density of the monitoring is fixed (one period every $K$ periods).

The paper is organized as follows. The model and the results are presented in Section 2. In Section 3 we illustrate the main feature of the equilibria that we propose in the context of a simple partnership game. Section 4 contains the proof of the result for the case where the number of players is at least three. In Section 5 the case of two players is addressed.A complete characterization of $V(\overline{G})$ seems to be difficult to obtain and is beyond the scope of this paper. However, in the concluding section, section 6, we make a few observations that are related to this problem.

## 2    The Model and the Result

The component game $G$ is defined as follows: Each player $i$, $i \in N \equiv \{1, ..., n\}$, chooses an action $a_i$ from a finite set of actions $A_i$. A profile of actions $a$, $a \in A \equiv \Pi_{i=1}^n A_i$, determines a payoff $g_i(a)$ for player $i$.

Let $\Sigma_i$ denote the set of mixed actions for player $i$ with a generic element $\alpha_i$. We let $g_i(\alpha)$, $\alpha \in \Sigma \equiv \Pi_{i=1}^n \Sigma_i$, denote the expected payoff when $\alpha$ is played.

We assume that player $i$ can perfectly observe the actions of a set of players $Q$, $Q \subseteq$

---

[8]Putting it a bit differently the problem is as follows: After a long sequence of consecutive periods in which player $j$ has monitored player $i$ her payoff in the continuation of the game is so close to the maximal payoff that is feasible that it will be impossible to compensate player $j$ for the cost of monitoring in the next period.

$N - \{i\}$ for some cost $C_i(Q)$, $0 \leq C_i(Q) < \infty$, where $C_i(\emptyset) = 0$ and $C_i(Q) \geq C_i(Q')$ if $Q \supseteq Q'$. Formally, we extend the game by associating with each action $a_i$ $2^{n-1}$ actions $(a_i, Q_i)$, $Q_i \subseteq N - \{i\}$, so that when $(a, Q)$ is played, $(a, Q) = (a_j, Q_j)_{j=1}^n$, player $i$ learns the action of each player in $Q_i$. Let $\bar{G}$ denote the extended game and let $\bar{A}_i$, $\bar{\Sigma}_i$, and $\bar{g}_i$ denote, respectively, the set of pure actions for player $i$, the set of mixed actions for player $i$, and the expected payoff for player $i$. So,

$$\bar{A}_i = A_i \times 2^{N-\{i\}}.$$

$$\bar{g}_i(a, Q) \equiv g_i(a) - C_i(Q_i).$$

The function $\bar{g}_i{}^9$ is extended to the space of mixed actions in the obvious way.

Let:

$$Z \equiv \max_i \max_{a \in A, a'_i \in A_i, Q_i \subseteq N - \{i\}} g_i(a'_i, a_{-i}) - \bar{g}_i(a, Q_i).$$

Thus, $Z$ is an upper bound on the benefit from a deviation in the extended game.

To state our results we need some additional definitions and notation. Let $y$ be some $n-$element vector. We let $y_{-i}$, $i \in N$, denote the corresponding $n - 1$ element vector where the i'th element is missing. For each $i \in N$ let $m^i \in \Sigma$ be a profile of mixed actions in $G$ such that:

$$m^i_{-i} \in \arg \min_{\alpha_{-i}} \max_{\alpha_i} g_i(\alpha_i, \ \alpha_{-i}).$$

$$m^i_i \in \arg \max_{\alpha_i} g_i(\alpha_i, m^i_{-i}).$$

---

[9]Since $\bar{g}_i$ depends only on $a$ and $Q_i$ we will sometimes write $\bar{g}_i(a, Q_i)$.

Let $\tilde{v}_i \equiv g_i(m^i)$. The profile $m^i_{-i}$ is a profile of minmax mixed actions against player $i$ and $\tilde{v}_i$ is the minmax payoff that players apart from player $i$ can force on player $i$. We can assume w.l.o.g. that:

$$\tilde{v} = (\tilde{v}_1, ..., \tilde{v}_n) = 0.$$

Let:

$$U \equiv \{(v_1, ..., v_n) \mid \exists \ a \in A \ s.t. \ g_i(a) = v_i\}.$$

$$V \equiv \text{ Convex hull of } U.$$

$$V^* \equiv \{v \in V \mid v_i > 0\}.$$

We can now turn to the repeated game. We assume that the players have at their disposal a public randomizing device. Such a device enables the players to implement in **each period** any payoff vector $v$ in the convex hull of the payoff vector set, $V^*$. This avoids uninteresting complications that come up when the convexification is done by playing different actions in different periods. The device selects an outcome from the interval $[0,1]$ according to the uniform distribution. Let $v = \sum_{i=1}^k \alpha_i v^i$ where $v^i \in U$ and $\sum_{i=1}^k \alpha_i = 1$. The payoff $v$ is implemented by partitioning the interval $[0,1]$ to $k$ sub-intervals where the length of the i'th subinterval, $P_i$, is $\alpha_i$. If the random device selects a point in $P_i$ then the players play a profile of actions with payoff $v^i$. Thus, the expected payoff is $v$. Formally, we introduce the notion of an **action function**, $f_i : [0, 1] \rightarrow A_i$, which specifies the action of a player $i$ as a function of the outcome of the randomizing device. Let $v \in V$. We let $f^v = (f_1^v, ..., f_n^v)$ denote a profile of action functions which induces a probability distribution on $A$ with an expected payoff vector $v$.

The repeated game is played as follows. Each period consists of three stages. In the first stage a public randomizing device selects an outcome from the interval [0,1] according to the uniform distribution. In the second stage the game $\bar{G}$ is played. In the third stage players simultaneously make public announcements. Formally each player $i$ selects a costless message $m_i$ from a set $M_i$ which is assumed to be countable[10]. Let $M \equiv \Pi_{i=1}^{n} M_i$. Let $\bar{\Omega}_i$ denote the set of possible observations of player $i$ in the extended game. That is:

$$\bar{\Omega}_i = \bigcup_{Q_i \in 2^{N-\{i\}}} A_{Q_i}$$

where

$$A_{Q_i} \equiv \Pi_{j \in Q_i} A_j.$$

A strategy for player $i$, $S_i$, consists of an action strategy $S_i^1$ and an announcement strategy $S_i^2$. These strategies specify the action and announcement of player $i$ at each period $t$ as a function of player $i's$ observations in the past. So, $S_i = (S_i^1, S_i^2)$ where:

$$S_i^1 = (S_i^1(t))_{t=0}^{\infty}, \ S_i^1(t) : [0,1]^t \times \bar{\Omega}_i^{t-1} \times M^{t-1} \to \bar{\Sigma}_i$$

and

$$S_i^2 = (S_i^2(t))_{t=0}^{\infty}, \ S_i^2(t) : \ [0,1]^t \times \bar{\Omega}_i^t \times M^{t-1} \to \Delta(M_i).$$

In the body of the paper we will describe the action strategy of a player $i$ in a period $t$ as a function from the history of the game (as perceived by player $i$), to her set of action functions. Specifically, a history of the game $h_{t-1} \in \bar{\Omega}^{t-1} \times M^{t-1}$, where $\bar{\Omega} = \Pi_{i=1}^{n} \bar{\Omega}_i$, determines an expected payoff vector $v$ for period $t$ which will be implemented by playing the profile of action functions $f^v$.

---

[10]Given a specific cost of monitoring and a payoff $v \in V^*$ to be implemented $M_i$ can be assumed to be finite. However, if we want $M_i$ to apply for every $v \in V^*$ we need to assume that it is countable.

Player $i's$ expected payoff when all the players play the profile $S$ and when the discount factor is $\delta$ is:

$$v_i(S,\delta) = (1-\delta)E[\Sigma_{t=0}^{\infty}\delta^t(\bar{g}_i(a(t),Q(t))]$$

where $(a(t),Q(t))$ is the profile of actions in the extended game $\bar{G}$ at time $t$, and where the expectation is taken with respect to the probability measure on the histories which is induced by the profile $S$.

Let $V(\bar{G},\delta)$ denote the set of sequential equilibrium payoff vectors when the discount factor is $\delta$. We are interested in:

$$V(\bar{G}) \equiv \lim_{\delta\to 1} V(\bar{G},\delta).^{11}$$

FM(1986) have shown that for the case of perfect monitoring $V(G) = V^*$ when $V^*$ is $n-$dimensional. Thus, with perfect monitoring the closure of $V(G)$ contains every feasible and individually rational payoff in $G$. The main result of this paper is that every feasible and individually rational payoff belongs to the limit set of sequential equilibrium payoffs, even when perfect monitoring is costly. Formally, we show:

**Theorem 1** *:If $V^*$ is $n-dimensional$, then:*

$$Int\ V^* \subseteq V(\bar{G}).$$

## 3   An Example

To demonstrate the main features of the equilibrium that we will present, we consider a simple example of a partnership game that is based on RMM(1986). There are two partners, each

---

[11]To be precise $v \in V(\bar{G})$ if there exists a $\bar{\delta} < 1$, $\bar{\delta} = \bar{\delta}(v)$, such that $v \in V(\bar{G},\delta)$ for every $\bar{\delta} < \delta < 1$.

of whom has two unobservable actions "work" ($W$) and "shirk" ($Sh$). There are two possible output levels: $y = 0$ and $y = 12$. The actions of the partners determine the probability of each outcome. If both partners work, then $y = 12$ with probability 5/6, if both shirk then $y = 12$ with probability 1/6, if one shirks and the other works then $y = 12$ with probability 1/2. Assume that the partners get an equal share of the output and that the disutilities from working and shirking are 4 and 1, respectively. The expected payoff function of each player is given in Figure 1.

|        | **w**   | **Sh**  |
|--------|---------|---------|
| **w**  | 1,1     | -1,2    |
| **Sh** | 2,-1    | 0,0     |

**Figure 1**

RMM show that if each partner observes only the output level then $V(\bar{G})$ does not contain a point in which the sum of the payoffs of the players exceeds 1. In particular, this implies that equilibrium payoffs are bounded away from efficiency.

Assume now that each partner can observe the action of the other partner for a certain cost, $C$. Let $\lambda$ be a number such that $\frac{1}{2} > \lambda > 0$. We will show that if $\delta$ is sufficiently close to 1, then there is an equilibrium in which the expected payoff for each player is $v^* = (1 - \lambda, 1 - \lambda)$.

We have pointed out in the introduction that to keep the exposition as simple as possible we assume that players do not observe anything for free beyond their own actions. For the sake of consistency we will make this assumption here as well, i.e., we assume that the players do not observe the output. At the end of the section we will show that the equilibrium that we construct for the case where the output is not observed is an equilibrium for the case where the outcome is observed as well.

Let $K$ be a number such that:

$$\frac{C}{K} + \frac{12}{K} < \lambda. \tag{1}$$

Consider the following profile of strategies, $S = (S_1, S_2)$. Time is divided into segments of $K$ periods. In each segment there is a target payoff $v$. (The way in which these target payoffs are related to $v^*$ will be explained later). The segment is played in the following way: player $i$ chooses at random a period $t_1^i$ in which she observes player $j$ and a pair of periods $(t_2^i, t_3^i)$ in which she switches from the action function $f_i^v$, that is, she plays the action that is different from the prescription of $f_i^v$ for that period. In all the periods different from $t_2^i$ and $t_3^i$ player $i$ plays according to $f_i^v$. The choices of $t_1^i$ and $(t_2^i, t_3^i)$ are independent and each one of them is made according to a uniform distribution. In the end of the segment each player $i$ announces the periods $t_k^i$, $k = 1, 2, 3$, that she selected and her observation of the action of player $j$ at $t_1^i$. This announcement is also interpreted as a statement that in the periods different from $t_2^i$ and $t_3^i$ player $i$ played according to $f_i^v$. We emphasize that each player $i$ is restricted to announcements of the type we have just described. So, for example, player $i$ cannot announce that she switched in more than two periods. (Formally, we consider equilibria in which an "illegitimate" announcement is a deviation.) If the announcements of the players are not consistent, that is, if the announcement of player $i$ on $j's$ action at period $t_1^i$ is different from what $j$ claims she did at $t_1^i$, then the players continue with $(Sh, Sh)$ forever. If the announcements are consistent then the next segment is played.

This part of the construction achieves the following goals:

1.      Each player is monitored in each period $t$ of the segment with probability $\frac{1}{K}$. So if a player deviates from the prescriptions of $f^v$ in more than two periods the probability for contradicting announcements is at least $\frac{1}{K}$.

2.      If player $i$ shirks from monitoring and guesses a period $t$ in which player $j$ did not switch, then the probability that the guess will be wrong is $\frac{2}{K}$.

3.     When $K$ is large the monitoring and switching are sparse so the average payoff in a segment is close to the target payoff. In particular, the average cost of monitoring per period, $\frac{C}{K}$, is small.

It follows from 1 and 2 that if a player deviates the probability of punishment is at least $\frac{1}{K}$. As we will see soon, even when $K$ is large, it is not worthwhile to risk a punishment when $\delta$ is high enough.

To complete the construction of the equilibrium $S$ we need to show how to tailor different segments together so that the given payoff $v^* \in V^*$ can be implemented. Second, we have to construct the equilibrium so that each player is indifferent to the periods in which she monitors and switches, (otherwise she will not pick these periods at random.) This is done as follows:

The target payoff of the first segment is $v^*$, the target payoff of the second segment is $v(2) = v^* + \varepsilon(2)$ where $\varepsilon(2)$ depends on the announcements of the players in the first segment and precisely offsets the net cost[12] of monitoring and switching in the first segment. Specifically, let $\bar{a}(t)$ denote the profile of actions that was played in period $t$ according to the announcements of the players at the end of the segment and let $a(t)$ denote the profile of actions that were prescribed by the action function $f^{v^*}$ in period $t$. Define:

$$x_i(t) \equiv g_i(a(t)) - g_i(\bar{a}(t)).$$

Thus, $x_i(t)$ is the cost (which might be negative) of switching at period $t$ for player $i$. (Note that in equilibrium $x_i(t)$ is zero if $t \neq t_2^i, t_3^i, i = 1, 2$.) It follows that the net cost of monitoring and switching for player $i$ is:

$$\delta^{t_1^i} C + \sum_{t=0}^{K-1} \delta^t x_i(t).$$

---

[12]This net cost might be negative.

We therefore define $\varepsilon(2)$ by the following equation:

$$\sum_{t=K}^{2K-1} \delta^t \varepsilon_i(2) = \delta^{t_1^i} C + \sum_{t=0}^{K-1} \delta^t x_i(t).$$

This definition of the term $\varepsilon_i(2)$ makes player $i$ indifferent to the periods in which she monitors and switches and thus ensures that choosing these periods randomly is consistent with the equilibrium. Note that inequality (1) and the fact that $\mid x_i(t) \mid = 0$ for $t \neq t_2^i,\ t_3^i,\ i = 1, 2,$ and $\mid x_i(t) \mid \leq 3$ for $t = t_2^i,\ t_3^i,\ i = 1, 2,$ imply that if $\delta$ is sufficiently close to 1, then $\mid \varepsilon(2) \mid < \lambda$ so that $v(2) = v^* + \varepsilon(2) \in V^*$. Thus, $v(2)$ is a feasible payoff.

The target payoffs of subsequent segments are defined in a similar way. Thus, the target payoff in a segment $k$ is $v^* + \varepsilon(k)$ where the $\varepsilon(k)$ term depends on the announcements of the players at the end of the $k-1$ segment and precisely offsets the net cost of monitoring and switching at that segment. Thus, the term $\varepsilon(k)$ makes player $i$ indifferent to the periods in which she monitors and switches in the segment $k-1$. Also, inequality (1) implies that all the target payoffs are in the set of feasible payoffs.

We now show that the expected payoff in $S = (S_1,\ S_2)$ is $v^*$. To see that note that the payoff of player $i$ in the $\ell' th$ segment, $seg(\ell)$, in units of the first period of the segment can be written as:

$$\sum_{t=0}^{K-1} \delta^t \cdot v_i^* + \sum_{t=0}^{K-1} \delta^t \cdot \varepsilon_i(\ell) + R_i(\ell)$$

where $R_i(\ell)$ is the net cost of switching and monitoring in $seg(\ell)$ and the $\varepsilon_i(\ell)$ term is the term that offsets the net cost of monitoring and switching in $seg(\ell-1)$. Thus $R_i(\ell)$ is offset by the $\varepsilon_i(\ell+1)$ term and the $\varepsilon_i(\ell)$ term is offset by $R_i(\ell-1)$. It follows that the expected payoff of player $i$ in $S$ is $(1-\delta)\sum_{t=0}^{\infty} \delta^t v_i^* = v_i^*$.

We have completed the definition of $S = (S_1, S_2)$ and have shown that it generates the payoff vector $v^*$. We now show that there exists $\bar{\delta} = \bar{\delta}(K)$ such that for every $\bar{\delta} \leq \delta < 1$

$S$ is a sequential equilibrium. We have to show that a player does not gain by deviating from $S$. A standard argument establishes that it is enough to demonstrate that a player will not gain by deviating at some segment $seg(\ell)$ and then play according to his equilibrium strategy in the following segments. The intuition as to why such a deviation is not worthwhile when $\delta$ is high is very simple. The reduction in the player's future payoffs in case his deviation is revealed outweighs any benefit he can make in the periods of $seg(\ell)$. Formally, if player $i$ does not deviate then his expected payoff calculated in units of the first period of $seg(\ell)$ is:

$$\sum_{t=0}^{K-1} \delta^t (v_i^* + \varepsilon_i(\ell)) + \delta^K \sum_{t=0}^{\infty} \delta^t \cdot v_i^*. \tag{2}$$

If he does deviate his expected payoffs are bounded from above by:

$$\sum_{t=0}^{K-1} \delta^t \cdot 2 + \frac{K-1}{K} \delta^K \sum_{t=0}^{\infty} \delta^t \cdot v_i^* \tag{3}$$

where (3) follows because the maximal payoff in the one-shot game is 2 and because the probability for punishment in case of a deviation is at least $\frac{1}{K}$. It is easy to see that if $\delta$ is high enough (2) is greater than (3). This completes the proof that $S$ is a sequential equilibrium.[13]

Consider now the original example of RMM where players observe the output. The equilibrium that we have defined is not upset by this signal. Specifically, if $\delta$ is close enough to 1 the strategies that we have constructed form an equilibrium even when players observe the output, (players simply ignore these observations). The reason why ignoring the signals

---

[13] A final clarification might be in place. The reader may wonder why the equilibrium strategy of a player prescribes switching twice rather than once. The point is that if player $i$ switches in only one period, then player $j$ can gain from the following deviation: If player $j$ observes a switch by player $i$ in period $t$ she announces that player $i$ did not switch in period 1! In this manner she increases her $\varepsilon-$term in the next segment because she reports a higher discounted cost of monitoring. If player $i$ switches twice, then this deviation will lead to contradicting statements with probability $\frac{1}{K-1}$.

is consistent with equilibrium behavior is that the signals are **imperfect**, player $i$ cannot tell from her observation of the output what player $j$ did and therefore she must monitor $j$ at some period. Specifically, if player $i$ deviates by shirking from monitoring and guesses the action of $j$ in some period, relying only on her signals, the probability that player $i$ will be wrong is bounded from below by some number $\rho > 0$. When $\delta$ is close enough to 1 the loss that player $i$ will incur in case she is wrong outweighs the cost of monitoring so that even when $\rho$ is small it is not worthwhile to shirk from monitoring.

The general results that we will obtain in sections 4 and 5 for the case where there are no signals can be extended to the case of public or private signals in the same manner provided that the signals have a common support. The significance of this assumption is that signals are imperfect so that a player cannot determine the action of another player with certainty. This imperfection ensures that player $i$ has an incentive to monitor. If, by contrast player $i$ had a perfect signal, *sig,* on the action of the player $j$ she monitors, so that when player $i$ observes *sig,* she knows with certainty the action of player $j$, then player $i$ would have had no incentive to monitor $j$. In such a case player $i$ prefers to wait until she observes *sig* and report the action of player $j$ at that period.

# 4   Three or More Players

In this section theorem 1 is proved for the case of three players or more. So let $v^* \in IntV^*$. We will construct a sequential equilibrium with payoff $v^*$.

To simplify the description of the equilibrium we will assume that the minmax profiles of actions $m^i$, $i \in N$ are pure. The extension to the general case is simple and is provided in the appendix.

The building blocks of the equilibrium are segments with a finite number of periods.

There are K-period and L-period segments. (The L-period segments are used to minmax one of the players). A segment, *seg*, is characterized by a target payoff $v$ and a monitoring assignment $q : N \to N$ such that $q(i) \neq i$. So each player $i$ has another player $q(i)$ which monitors him. The segment implements an average payoff that is 'close' to $v$ and ensures that if one of the players deviates from the equilibrium strategies then she is caught with a probability of at least $\frac{1}{M}$ where $M = K, L$ is the length of the segment. Specifically the segment is played as follows: player $i$ plays the action function $f_i^v$ in each period with the following exception: Player $i$ chooses at random a pair of periods $(t_i^1, t_i^2)$ in which she plays some action that is different from the action prescribed by $f_i^v$. In addition, if player $i$ monitors another player $k$, then player $i$ chooses at random a period $t_k$ in which she observes player $k$. The choices of the $t_k$ and $(t_i^1, t_i^2)$ are independent and each one of them is made according to the uniform distribution. At the end of the segment each player $i$ announces the periods $t_i^1$ and $t_i^2$ in which she switched from playing according to $f_i^v$, her pure actions at these periods, and her observations of the players that she monitored. (So, if she monitors player $k$ she will announce the period $t_k$ and the action of player $k$ at that period.) This announcement is also interpreted as a statement that, in all periods different from $t_i^1$ and $t_i^2$, player $i$ played according to $f_i^v$. We emphasize that player $i$ is restricted to such announcements. Specifically, she cannot announce that she switched in more than two periods or that she did not monitor a player $k$ which was assigned to her. (Formally, an "illegitimate" message will be considered a deviation).

Now it is easy to see that if player $i$ deviates from the equilibrium behavior, (either by lying about her action in a certain period, or by guessing the action of the player she monitors instead of actually observing him), then with a probability of at least $\frac{1}{M}$ her announcement will be inconsistent with the announcement of some other player. (The argument is similar to the one in the example in section 3). In such a case player $i$ and the player that she

contradicted will be punished in the segments that follow the current one.

Next, note that with the exception of at most $3N$ periods (three periods for each player), the expected payoff vector in every other period is $v$. It follows that if $K$ and $\delta$ are large, the expected average payoff on the segment is close to $v$.

We can now turn to a complete description of a profile of strategies $S = (S_1, ..., S_n)$ which is a sequential equilibrium with an expected payoff $v^*$. For this purpose it is convenient to introduce the notion of a 'phase'. A phase is a sequence of segments such that the announcements in each segment are consistent. A transition from one phase to another occurs whenever there is a mismatch between the announcements of some player and her monitor. If there is more than one pair of players that contradict each other then one pair of the set of contradicting pairs is selected at random. (Each pair with an equal probability).

There are two types of phases. The first type are phases $r(v)$ in which the expected payoff is $v$ and each phase is composed of a sequence of $K-$ period segments. This set of phases consists of the initial phase $r(v^*)$, and then, for every pair of players $i, j$, there is a phase $r(v_{i,j}^*)$ where $v_{i,j}^*$ is defined as follows: The payoffs of players $i$ and $j$ are $v_i^* - \eta$ and $v_j^* - \eta$ respectively and the payoff of a player $k \neq i, j$ is $v_k^*$. Thus, $r(v_{i,j}^*)$ is a phase in which players $i$ and $j$ are punished by reducing their payoffs by $\eta > 0$. Note that since $v^* \in IntV^*$ we can choose $\eta > 0$ so that $v_{i,j}^* \in IntV^*$ for every pair $i, j$. In fact we want to choose $\eta$ so that $v_i^* - 2\eta > 0$ for every $i \in N$. The second type of phases are phases $\hat{r}_{i,j}$ in which player $i$ is first minmaxed for $L$ periods and then in the subgame that follows the payoffs of players $i$ and $j$ are approximately $v_i - 2\eta$ and $v_j - \eta$ respectively. (By approximately we mean that there is a certain correction term which can be made negligible in comparison to $\eta$).

Before providing a detailed description of each phase it is worthwhile to see the general structure of the equilibrium as it is reflected by the transition rule. The basic idea is simple if the announcements of two players $i$ and $j$ contradict each other then one of them has

deviated and since it is unknown which one of them it was both are punished. Now, since it is, in general, impossible to punish the same pair of players repeatedly (due to the fact that it is, in general, impossible to jointly minmax a pair of players), the monitoring assignment changes from one phase to the other so that in a phase where players $i$ and $k$ are punished $(r(v_{i,k}^*), \; \hat{r}_{i,k}.$or$ \; \hat{r}_{k,i})$ they do not monitor each other and therefore if there will be contradicting statements in this phase they will involve a pair of players which is different from $(i,k)$.[14] The transition rule is defined as follows: If player $i$ and her monitor $j$ make contradicting statements, (which implies that one of them has deviated), then the next phase is $r(v_{i,j}^*)$ unless one of them, say, $i$, is already punished in the current phase (i.e., the current phase is $r(v_{i,k}^*), \; \hat{r}_{(i,k)}$ or $\hat{r}_{(k,i)}$ for some $k \neq i, j$) in this case the next phase is $\hat{r}_{(i,j)}$.(So player $i$ is, first, minmaxed and after that the payoffs of $i$ and $j$ are reduced by $\eta$).

We now turn to a full definition of the different phases. Consider first a phase $r(v)$ ($v \in Int \; V^*$). This phase is defined by a sequence of $K-$period segments $(seg(\ell)_{\ell=1}^{\infty}$. The target payoff of $seg(1), v(1),$ is $v$ and the target payoff of $seg(\ell), v(\ell),$ is $v + \varepsilon(\ell)$ where the term $\varepsilon(\ell)$ offsets the net costs of monitoring and switching in the previous segment, $seg(\ell-1)$. Specifically let $t$ denote the period $t$ in the segment $seg(\ell - 1)$. Let $C_i(t)$ denote the cost incurred by player $i$ for monitoring in period $t$ according to the announcement of player $i$. ($C_i(t)$ is zero unless player $i$ says that she monitored in $t$). Let $\bar{a}(t)$ denote the profile of actions that were played in period $t$ according to the announcements of the players, and let $a(t)$ denote the profile of actions that were prescribed by $f^{v(\ell-1)}$ in $t$. Define:

$$x_i(t) = g_i(a(t)) - g_i(\bar{a}(t))$$

The term $x_i(t)$ is the cost (which might be negative) of the switching at period $t$ for player $i$, (when no player switches $x_i(t) = 0$). The term $\varepsilon(\ell)$ is defined by the following

---

[14]So if $(\ell,$m$)$ are a pair of players that contradict each other in a phase where $i$ and $k$ are punished, then either $\ell \neq i, k$ or $m \neq i, k$.

equation:

$$\sum_{t=K}^{2K-1} \delta^t \varepsilon_i(\ell) = \sum_{t=0}^{K-1} \delta^t \left( C_i(t) + x_i(t) \right)$$

This definition of the term $\varepsilon(\ell)$ achieves two goals. First, it makes player $i$ indifferent to the periods in which she monitors and switches and in this way ensures that choosing randomly the periods of monitoring and switching is consistent with equilibrium. Second, the expected discounted payoff in the phase is $v$. (To see that recall the example).

Finally, note that since each player monitors in one period and switches only twice $C_i(t) + x_i(t)$ equals zero except for at most $3N$ periods. It follows that $\varepsilon_i(\ell)$ can be made as small as we wish by taking $K$ and $\delta$ to be large enough and therefore $v + \varepsilon(\ell)$ can be assumed to be a member of the set of feasible and individually rational payoffs, $V^*$.[15]

Consider, now, the phase $\hat{r}_{(i,j)}$. The phase is a sequence of segments $(seg(\ell))_{\ell=1}^{\infty}$. The segment $seg(1)$ is an $L$−period segment in which player $i$ is minmaxed. The target payoff in $seg(1)$ is $g(m^i)$, (recall that $m^i$ is the profile of action that minmaxes player $i$) and player $i$ does not monitor and is not monitored. Thus the monitoring assignment is restricted to the set of players different from $i$. (Note that there is no need to monitor player $i$ because his action in the profile $m^i$ is a best response). The segments that follow are $K$−period segments in which the monitoring assignment has the property that $i$ and $j$ do not monitor each other. Let $\hat{v}_{(i,j)}$ denote the payoff vector where the payoff of player $i$ is $v_i^* - 2\eta$, the payoff of player $j$ is $v_j^* - \eta$, and the payoff of a player $k$, $k \neq i,j$ is $v_k^*$. The target payoff of $seg(\ell)$, $\ell \geq 2$, is $\hat{v}_{(i,j)} + \varepsilon(\ell)$ where $\varepsilon(\ell)$ offsets the costs of monitoring and switching in the previous segment, $seg(\ell - 1)$.

We have completed the definition of the profile $S = (S_1, ..., S_n)$. We now show that we can choose the parameters $K, L,$ and $\eta$ so that there exists a number $\bar{\delta} < 1$ such that for

---

[15]The $K$ and $\bar{\delta}$ which ensure that the $\varepsilon - terms$ are below a given term $\bar{\varepsilon}$ are an increasing function of the cost $C$ and the number of players, $N$.

every discount factor $\delta$, $\bar{\delta} \leq \delta < 1$, $S$ is a sequential equilibrium in the repeated game. Start with $\eta$, the only restriction on $\eta$ is that the payoffs $v^*_{(i,j)}$ and $\hat{v}_{(i,j)}$ belong to $IntV^*$. Since $v^* \in IntV^*$ this can be done. Once $\eta$ is chosen we choose $K$ and $\bar{\delta}$ to be large enough so that the $\varepsilon-$terms which offset the costs and benefits of monitoring and switching are small in comparison to $\eta$. Such a choice will enable us to ignore these $\varepsilon - terms$ when we evaluate deviations from the equilibrium strategies.

We want to show that the expected gain from a deviation is negative. A standard argument implies that we only need to consider a deviation in one segment followed by confirming with the equilibrium in subsequent segments.[16]We will evaluate such a deviation by comparing its expected *benefit* with its expected *loss*. The *benefit* is the difference between the player's payoff in the current segment when she deviates and her payoff at this segment if her announcement is true, that is, if she actually did what she says she did. The *loss* from a deviation is the reduction in the payoff of the player in the subgame that starts at the end of the current segment.

Consider first a deviation of player $i$ in a phase $\hat{r}_{(i,j)}$. A deviation in the first segment, where $i$ is minmaxed, is not worthwhile because in this segment the equilibrium behavior in each period is a best response for $i$. So assume that player $i$ deviates in some subsequent segment. The benefit from this deviation in units of the first period of the segment is bounded from above by:

$$(1 - \delta) \sum_{t=0}^{K-1} \delta^t Z. \tag{4}$$

---

[16]We refer to the argument that establishes that, if there is a worthwhile deviation which involves deviations in an infinite number of periods, then it is possible to gain by deviating in only a finite number of periods as well. This, in turn, implies that if there is a worthwhile deviation, then there is also a worthwhile deviation of the type that we discuss.

(Recall that $Z$ is an upper bound on the benefit from a deviation in the one-shot game).

If player $i$ is caught (i.e., her statement contradicts the statement of another player), then the phase $\hat{r}_{(i,j)}$ will be played from the beginning, i.e., player $i$ will be minmaxed for $L$ periods and in the periods that follow her payoff will be $v_i^* - 2\eta$ per period. This means that the reduction in her payoff in the subgame that follows the current segment is $v_i^* - 2\eta$ in each one of the first $L$ periods. (When she is minmaxed she gets zero in each period while if she would have confirmed with the equilibrium she would have obtained a payoff $v_i^* - 2\eta$ in each period). The probability that player $i$ will be caught is at least $\frac{1}{K}$. Thus, the expected loss from the punishment in units of the first period of the current segment is bounded from below by:

$$\frac{1}{K}(1-\delta) \sum_{t=K}^{K+L-1} \delta^t (v_i^* - 2\eta). \tag{5}$$

It is easy to see that if $L$ is large enough and $\delta$ is sufficiently close to 1 then (5) is greater than (4) and therefore this deviation is not worthwhile.

Consider now other deviations of player $i$. Let $seg$ denote the segment in which player $i$ deviates and let $M$, $M \in \{K, L\}$, be the number of periods in $seg$. The benefit from the deviation in units of the first period of $seg$ is bounded from above by:

$$(1-\delta) \sum_{t=0}^{M-1} \delta^t Z. \tag{6}$$

To assess the expected loss from the deviation we distinguish between two cases: (a) Player $i$ is not punished in the current phase; (b) Player $i$ is punished. Consider, first, case (a). If player $i$ is caught, then her payoff in the subgame that follows $seg$ is $v_i^* - \eta$. If she would have stuck to the equilibrium her payoff would have been $v_i^*$. The probability that she

is caught is at least $\frac{1}{M}$. It follows that the expected loss from the deviation in units of the first period of $seg$ is bounded from below by:

$$\frac{1}{M}(1-\delta)\sum_{t=M}^{\infty}\delta^t\eta. \tag{7}$$

It is easy to see that if $\delta$ is close enough to 1 (7) is greater than (6) and therefore the deviation is not worthwhile.

Consider now Case (b). If the deviation of player $i$ is revealed, then player $i$ is punished by being minmaxed for $L$ periods and then getting an expected payoff $v_i^* - 2\eta$ in each period. If player $i$ would have confirmed with the equilibrium her expected payoff would have been $v_i^* - \eta$ in each period that follows $seg$. It follows that in the first $L$ periods the loss is $v_i^* - \eta$ per period and after that it is $\eta$ per period. Since $v_i^* - \eta > \eta$ we get that the expected loss in this case is bounded from below by (7) as well. Therefore we obtain again that if $\delta$ is close to 1 the deviation is not worthwhile.

We have shown that $S$ is a sequential equilibrium[17] and have thus proved that every payoff vector $v^*$ in which the payoff of each player is above her minmax payoff in **pure** actions can be implemented in a sequential equilibrium. In the appendix we show how the proof can be easily extended for payoff vectors in which the payoff of each player is above her minmax payoff in **mixed** actions.

---

[17]There is still room for some clarifying comments regarding: 1) Deviations where a player observes other players more than once. 2. The beliefs of players outside the equilibrium path. These comments are marginal to the main ideas in the construction and therefore they are delegated to the appendix.

# 5    Two Players

In this section we prove Theorem 1 for a two-player game. In the case of two players, there are histories in which players are *jointly minmaxed.* We will see that this requires a modification of the equilibrium that was described in Section 4.

We assume that the action by which player $i$ minmaxes player $j$ is a mixed action. This assumption is without loss of generality because since $v^* \in \text{Int } V^*$ we can find a mixed action (which is close to $m_i^j$), call it $\hat{m}_i^j$, such that $\max_{a_j \in A_j} g_j(a_j, \hat{m}_i^j) < v_j^*$. As will become clear from the proof this last inequality is all we need.

We can now turn to the description of an equilibrium $S$ with an expected payoff vector $v^* \in Int\ V$. The equilibrium $S$ consists of two phases, $r(v^*)$ and $\hat{r}$. The phase $r(v^*)$ is similar to the phase $r(v^*)$ that was described in Section 4. (Since there are only two players, the monitoring assignment is obvious – each player monitors the other). The phase $\hat{r}$ is defined as follows: Let $L$ and $\ell$ be integers such that $L = \ell^2$. In the first $L$ periods each player $i$ plays the action $m_i^j$ and in addition chooses a set of $\ell$ periods in which she observes player $j$. Each one of the sets of $\ell$ periods is selected with an equal probability. At the end of $L$ periods each player announces her sequence of actions and her observations of the other player's actions. If the announcements are inconsistent, then the players start the phase $\hat{r}$ from the beginning. If the announcements are consistent, then the continuation is identical to the phase $r(\bar{v})$ where $\bar{v}$ depends on the announcements of the players and is defined in the following way: Let $\sigma_i \in \bar{A}_i^L$, $\sigma_i = \sigma_i(t)$, $t = 1, ..., L$ be a sequence of $L$ pure actions for player $i$ in the extended game $\bar{G}$ (thus, an action $\sigma_i(t) \in \bar{A}_i$ determines an action in the original game $G$ and a decision whether to monitor player $j$ or not). Define $\bar{v}_i$, $\bar{v}_i = \bar{v}_i(\sigma_i)$, so that the following equation holds:

$$(1 - \delta) \left[ \sum_{t=0}^{L-1} \delta^t \bar{g}_i(\sigma_i(t), m_j^i) + \sum_{t=L}^{\infty} \delta^t \bar{v}_i \right]$$

$$= (1 - \delta) \left[ \sum_{t=0}^{L-1} \delta^t \cdot 0 + \sum_{t=L}^{\infty} \delta^t v_i^* \right] = \delta^L v_{i.}^*$$

Thus, for every sequence of pure actions in the extended game, $\sigma_i$, player $i$ is indifferent between playing $\sigma_i$ in the minmax segment and receiving afterwards a payoff $\bar{v}_i$ in each period, and the path of one-shot payoffs $\bar{P}$ in which she receives 0 in the first $L$ periods and $v_i^*$ in every subsequent period. In particular, player $i$ is indifferent among her pure actions in the original game $G$ and indifferent to the set of periods in which she observes player $j$. Note that: (i) since $g_i(a_i, m_j^i) \leq 0$ for every $a_i \in A_i$, $d_i(\sigma_i) = \bar{v}_i(\sigma_i) - v_i^* \geq 0$ for every $\sigma_i$; (ii) $d_i(\sigma_i)$ can be as small as we wish for every $\sigma_i$ if $\delta$ is sufficiently close to 1. So if $v^* \in Int\ V$ we can assume that $\bar{v} \in V$ and therefore $\bar{v}$ is a feasible payoff.

We now show that there is an $\bar{L}$ such that there exists a $\bar{\delta}$ so that, for every $\delta > \bar{\delta}$, $S$ is a sequential equilibrium. We need to consider two types of deviations:

($i$)  a deviation of a player in the minmax segment of the phase $\hat{r}$

   followed by confirming with the equilibrium after the end of the segment;

($ii$)  a deviation at any other segment followed, again, by confirming with

   the equilibrium

It will become apparent that in evaluating these deviations we need to consider only a finite horizon which consists of the segment in which the deviation occurs and the $L$ periods which follow it. Since $\bar{\delta}$ is as close to 1 as we wish the effect of the discounting in this finite horizon is negligible and can be ignored. Thus, we can evaluate a deviation in terms of one-shot payoffs. So, for example, we will evaluate a payoff $x$ in each of $h$ periods by $h \cdot x$.

Also, in evaluating a deviation we will ignore, (as we did in section 4), the $\varepsilon-$terms that offset the costs of monitoring and switching in $K-$period segments. These terms can

be made negligible by choosing $K$ and $L$ which are large enough.

Consider, first, a deviation of player $i$ in a segment which is different from the minmax segment. Such a segment consists of $K$ periods and therefore the benefit from the deviation is bounded by:

$$K \cdot Z. \tag{8}$$

Consider, now, the loss from the deviation. If the announcements are inconsistent, then player $i's$ expected payoff starting from the first period of the next segment is given by the path $\bar{P}$ (recall that $\bar{P}$ consists of $L$ periods with a zero payoff followed by a payoff $v^*$ in every subsequent period). Let $P_x$ denote the path that consists of the one-shot payoff $x$ in each period. If the announcements are consistent, then player $i's$ expected payoff is given by a path $P_x$ where $x = v_i^*$ or $x = \bar{v}_i(\sigma_i)$ for some $\sigma_i \in \bar{A}_i^L$. Now, since $\bar{v}_i(\sigma_i) \geq v_i^*$ it follows that the loss from the deviation in case the announcements are inconsistent is bounded from below by $L \cdot v_i^*$. (Recall that we are evaluating the loss only in the $L$ periods that follow the minmax segment. We can do this because it is an underestimation of the loss.) Since the probability that the announcements are inconsistent is at least $1/K$, the expected loss is at least:

$$\frac{1}{K} \cdot (L \cdot v_i^*). \tag{9}$$

It is easy to see that if $L$ is large enough, then (9) is larger than (8) and, therefore, the deviation is not worthwhile.

We turn now to evaluating a deviation of player $i$ in the minmax segment. Before demonstrating that the player does not gain from deviating, it is worthwhile explaining why a simpler construction where each player chooses only one period for monitoring would not work. So, consider a variant of our equilibrium where in the minmax segment a player

chooses only one period in which she monitors the other player. Let $a_i$ and $a'_i$ be two pure actions for player $i$ in $G$ such that:

$$y_i \equiv g_i(a_i, m^i_j) - g_i(a'_i, m^i_j) > v^*_i.$$

Consider a deviation where player $i$ lies and announces that at some period $t$ she played $a'_i$ while in fact she played the action $a_i$. Assume that this is the only deviation of player $i$. The gain from the deviation if player $i$ is not caught, that is, if the announcements at the end of the segment are consistent, is $y_i$. To evaluate the loss, note that the expected payoff of player $i$ starting with the first period of the next segment is $v^*_i$. If player $i$ is caught, then she will be minmaxed for $L$ periods and therefore her loss in this case is $L \cdot v^*_i$. The probability that player $i$ will be caught is $1/L$. It follows that expected loss is $v^*_i$ while the expected gain is $((L-1)/L) \cdot y_i$. Thus, if $L$ is large the deviation is worthwhile.

We now show that a choice of $\ell = \sqrt{L}$ periods of monitoring works and that $S$ is an equilibrium. We will evaluate a deviation of player $i$ by comparing her gain from the deviation in case she is not caught (that is, the case where the announcements at the end of the segment are consistent), with her loss in case she is caught.

Let $\bar{a}_i, \bar{a}'_i \in \bar{A}_i$ be pure actions of player $i$ in the extended game. Let $\bar{y}_i = \bar{g}_i(\bar{a}_i, m^i_j) - \bar{g}_i(\bar{a}'_i, m^i_j)$. Player $i$ gains $\bar{y}_i$ at period $t$ if she announces that she played $\bar{a}'_i$ at $t$ while in fact she played $\bar{a}_i$ at that period, and if she is not caught. Thus, the gain in a single period is bounded from above by $Z$ and therefore the gain from a deviation in $k$ periods is bounded by $k \cdot Z$. Consider now the loss of player $i$ in case she is caught. If player $i$ would have confirmed with the equilibrium, then her expected payoff would have been given by the path $\bar{P}$ in which she gets 0 for $L$ periods and $v^*_i$ in every subsequent period. If player $i$ deviates and is caught, then her payoff is bounded from above by the payoff that corresponds to the path $P_c$ in which player $i$ gets 0 for $2L$ periods and then $v^*_i$ in every subsequent period. This follows because the maximal one-shot payoff of player $i$ in each period in the minmax segment is

0 and if she is caught, then the game continues by playing the phase $\hat{r}$ from the beginning. It follows that the loss from being caught is bounded from below by $L \cdot v_i^*$. Let $p(k)$ denote the probability that player $i$ will be caught if she deviates in $k$ periods. It follows that the deviation is *not* worthwhile if the following inequality holds:

$$[1 - p(k)]kZ < p(k)Lv_i^*. \tag{10}$$

**Claim:** There exists $\bar{L}$ such that (10) holds for every $L \geq \bar{L}$.

**Lemma 2** *There exists $\bar{L}$ such that if $L \geq \bar{L}$ then:*

$$p(k) \geq 1 - \left[\left(1 - \frac{1}{\ell}\right)^k\right]. \tag{11}$$

**Proof.** It will be shown that:

$$1 - p(k) \leq \left(1 - \frac{1}{\ell}\right)^k. \tag{12}$$

There are two types of deviations: (i) shirking from monitoring and (ii) giving a false announcement regarding the pure action in the original game $G$, $a_i(t) \in A_i$, that was played in some period $t$.

Let $a_j$ and $a_j'$, $a_j, a_j' \in A_j$, denote respectively two pure actions in the support of $m_j^i$ which are assigned the highest and lowest probabilities, respectively. Let $p_j$ and $p_j'$ denote these probabilities, respectively. Now, assume that player $i$ monitors player $j$ in only $\ell - 1$ periods. Since player $i$ must report on the actions of player $j$ in $\ell$ periods her optimal continuation is to select some additional period $t$ and announce that she observed player $j$ in period $t$ and that player $j$ played the action $a_j$. The probability that player $i$ will be wrong is at least $p_j'$. If $L$ is large enough, $p_j' \geq 1/\ell$ and our claim follows for $k = 1$, (the case where player $i$ shirks in only one period). Suppose, now, that player $i$ shirks in $k$ periods where

$k > 1$. If player $i$ is not caught at the end of the segment, then player $i$ guessed correctly the actions of player $j$ in $k$ periods. The probability of that is at most $\left(1 - p'_j\right)^k$ since $p'_j \geq 1/\ell$, (12) follows.

Now suppose that player $i$ deviates by lying about his pure action $a_i(t) \in A_i$ in some period $t$. Player $i$ is caught if player $j$ monitors player $i$ at $t$. The probability of that is $\ell/L = 1/\ell$. So our claim follows for $k = 1$. If player $i$ lies about her actions in $k$ periods then the probability that she will not be caught is:

$$\prod_{h=0}^{k-1} \left(1 - \frac{\ell}{L - h}\right). \tag{13}$$

Since (13) is lower than the r.h.s. of (12) the lemma follows.

We can now turn to the proof of the claim.

Let

$$\lambda \equiv \frac{v_i^*}{Z}$$

and let $\theta$ be the smallest number such that $\theta \geq \lambda$

and

$$\frac{2e^{-\theta}}{1 - e^{-\theta}} < \frac{v_i^*}{Z} \tag{14}$$

where $e$ is the basis of the natural logarithm. Consider the following three cases:

1. $\quad 1 \leq k \leq \lambda\ell$.

2. $\quad \lambda\ell \leq k \leq \theta\ell$.

3. $\quad\quad \theta\ell < k$.

We have $p(k) \geq p(1) \geq 1/\ell$ for every $k > 1$. Thus, for $1 \leq k \leq \lambda\ell$ we have:

$$p(k)Lv_i^* \geq \ell v_i^* = \lambda \ell Z > (1 - p(k))kZ.$$

Next, consider case (2). By the lemma,

$$1 - p(k) \leq \left(1 - \frac{1}{\ell}\right)^k = \left[\left(1 - \frac{1}{\ell}\right)^\ell\right]^{\frac{k}{\ell}}. \tag{15}$$

By choosing $L$, and as a result $\ell$, large enough we can make the RHS of (15) as close as we wish to $e^{-(k/\ell)}$. It follows that if $L$ is large enough and, as a result, $L/\ell$ is large, then for $\lambda \ell \leq k \leq \theta \ell$ we have:

$$p(k)Lv_i^* \geq \left[1 - \left[(1 - \frac{1}{\ell})^\ell\right]^{\frac{k}{\ell}}\right] Lv_i^* \geq \frac{1}{2}(1 - e^{-(k/\ell)})Lv_i^*.$$

$$\geq \frac{1}{2}(1 - e^{-\lambda})Lv_i^* > \theta \ell Z > (1 - p(k))kZ$$

(The inequalities from left to right follow because of (15), $\left(1 - \frac{1}{\ell}\right)^\ell$ tends to $e^{-1}$, $\lambda < \frac{k}{\ell}$, $L/\ell$ is large, $\theta \ell > k$.)

Finally, (15) and (14) imply that when $k > \theta \ell$ :

$$p(k)Lv_i^* \geq (1 - e^{-\theta})Lv_i^* > e^{-\theta}LZ \geq (1 - p(k))kZ$$

where the last inequality follows because (15) and the fact that $1 - p(k)$ is decreasing in $k$ imply that $1 - p(k) \leq e^{-\theta}$ for $k > \theta \ell$.

This completes the proof of the claim. Thus, we have shown that with an appropriate choice of the parameters $K$ and $L$ there exists $\bar{\delta}$ such that $S$ is a sequential equilibrium whenever $\delta \geq \bar{\delta}$.

# 6 Conclusion

We have shown that $V(\overline{G})$ contains the interior of $V^*$. A complete characterization of $V(\overline{G})$ seems to be difficult to obtain and is beyond the scope of this paper. However, in this section we make a few observations that are related to this problem.

Let $\widehat{v}_i$ denote the minmax payoff for player $i$ in $G$ when the players different from $i$ can <u>correlate</u> their actions against $i$. That is,

$$\widehat{v}_i = \min_{\beta_{-i} \in \Delta(A_{-i})} \max_{a_i \in A_i} g_i(a_i, \beta_{-i})$$

Let $\overline{\beta_{-i}} \in \Delta(A_{-i})$ be a probability distribution such that $\max_{a_i} g_i(a_i, \overline{\beta_{-i}}) = \widehat{v}_i$. Define:

$$W = \{w \in R^n \mid \exists v \in V, \quad \widehat{v}_i \leq w_i \leq v_i, \ i = 1, .., n\}$$

We claim that $V(\overline{G}) \subseteq W$.

To see that we, first, observe that given a profile of strategies in the repeated game, $S_{-i}$, player $i$ can guarantee a payoff of at least $\widehat{v}_i$ as follows: Player $i$ never spends money on observing the actions of the other players. Let $\beta_{-i}^1 \in \Delta(A_{-i})$ denote the probability distribution on the profiles of actions of the players different from $i$ in the original game $G$ in the first period. The action of player 1 in the first period is $a_i^1$ where

$$a_i^1 = \arg \max_{a_i \in A_i} g_i(a_i, \beta_{-i}^1)$$

Assume that $a_i^1, ...., a_i^{t-1}$ have been defined. Let $\beta_{-i}^t$ denote the probability distribution on the profile of the actions of the players different from $i$ in period $t$ given that player $i$ played $a_i^1, ...., a_i^{t-1}$ in the first $t-1$ periods. Define:

$$a_i^t = \arg \max_{a_i \in A_i} g_i(a_i, \beta_{-i}^t)$$

It is easy to see that in each period $t$ player $i$ gets an expected payoff of at least $\widehat{v}_i$ and therefore he gets at least $\widehat{v}_i$ in the repeated game.

Next, we observe that a payoff, $w$, in the repeated game $\overline{G}(\delta)$ can be represented as $w = v - c$ where $v \in V$ is the corresponding payoff in the repeated game $G(\delta)$ and $c$ is the discounted sum of the costs of monitoring in the different periods.

These two observations establish the claim.

We note that the set $W$ is different from the set of individual rational payoffs, $V^*$, in two ways. Let $w \in W$ then,

1. $w_i$ may be smaller than $\overline{v_i}$, the individual rational payoff for player $i$.

2. The vector $w$ may not belong to $V$, the convex hull of the set of payoff vectors in $G$.

We now show that $V(\overline{G})$ may indeed differ from $V^*$ in these two ways. To see 1. consider a 4-person game in which player 1 does not observe anything for free, not even his payoff, while each one of the other players observes the actions of all the players for free. We will show that if the cost for player 1 of observing the actions of any other player is sufficiently high then it is possible to implement in equilibrium a payoff vector $w$ such that $w_i$ is "close" to $\widehat{v}_i$ .The idea is simple; If player 1 deviates in some period $\widehat{t}$ then the other players correlate their actions against him for some number of periods, $L$, that is large enough to offset the gain of player 1 in period $\widehat{t}$. This is done in the following way: Assume w.l.o.g. that $|A_2| \geq |A_j|$ for $j = 3, 4$. Player 2 conducts a private lottery according to the probability distribution $\overline{\beta_{-1}}$ on $A_{-1}$. (Recall that $\overline{\beta_{-1}}$ is a probability distribution such that $\max_{a_1} g_1(a_1, \overline{\beta_{-1}}) = \widehat{v}_i$. ) Let $(a_2, a_3, a_4)$ be the outcome of the lottery. Since $|A_2| \geq |A_j|$ for $j = 3, 4$ player 2 can communicate $a_j$ to player $j$ in one period. Specifically, let $f_j, f_j : A_j \rightarrow A_2$, be some bijection of $A_j$ into $A_2$. Player 2 communicates $a_j$ to player $j$ by playing $f_j(a_j)$ in period $\widehat{t} + (j - 2)$. So after two periods players 3 and 4 know $a_3$ and $a_4$ and starting with

period $\widehat{t}+3$ players 2,3 and 4 play $(a_2, a_3, a_4)$ for $L$ periods where $L$ is large enough to offset the gain of player 1 from the deviation in period $\widehat{t}$. Now the point is this, if the cost for player 1 of observing any action of any other player is high enough then it is not worthwhile for him to make any observation in any period $t, \widehat{t}+1 \leq t \leq \widehat{t}+L+2$. If player 1 doesn't observe any action he has no information on the actions of the other players and therefore his maximal expected payoff in each period $t$ in the punishment segment is $\widehat{v_1}$. Since along the equilibrium path player 1 gets an expected payoff $w_1$ in each period, where $w_1 \succ \widehat{v_1}$, then if $L$ is large enough the deviation is not worthwhile.

We have shown that $V(\overline{G})$ may contain points in which the payoffs of some of the players are below their individual rational payoffs. The set $V(\overline{G})$ may also contain points which do not belong to $V$. To see that we note, without proof, that there are equilibria, which are similar to the equilibria that were defined in section 4, in which some (or all) players observe other players many times in each segment so that the average cost of monitoring per period is not negligible. Thus, it is possible to diminish the payoff of players by forcing them to spend money on monitoring. In particular, given a $n$-person game $G$ a vector $v \in V^*$ and a vector $\alpha = (\alpha_1, ....., \alpha_n)$ such that $0 \prec \alpha_i \leq 1$, $i = 1, .., n$, there exist cost functions $C_i$, $i = 1, .., n$, such that the payoff vector $\alpha \cdot v$ belongs to $V(\overline{G})$.

## **APPENDIX**

We will show how to modify the equilibrium that was described in section 4 so that it can implement a payoff vector $v^*$ in which the payoff of each player is above her minmax payoff (in mixed actions). This will be done by demonstrating that the equilibrium can be modified so that in an L-period segment in which a player $i$ is minmaxed she is brought down to her minmax payoff in mixed actions.

So consider a phase $\hat{r}_{i,j}$ which is composed of a sequence of segments $(seg(\ell))_{\ell=1}^{\infty}$.

The segment $seg(1)$ is an $L$ period segment in which player $i$ is minmaxed. The segment is played as follows: A player $k$, $k \neq i$, for which $m_k^i$, (the action of player $k$ which minmaxes player $i$), is a mixed action, plays $m_k^i$ in each one of the $L$ periods. Player $k$ does not need to switch because as we will see mixing performs the function of switching in providing incentives to monitor. If the action $m_k^i$ happens to be a pure action then player $k$ chooses at random two periods in which she switches (just like the switching in the main proof). Each player $k$ is monitored by some other player $q(k)$ who is different from $i$. The monitoring is similar to the one that we have described in the main text, i.e., player $q(k)$ chooses at random a period $t$ in which she observes player $k$. Player $i$ plays $m_i^i$ (her best response to the minmaxing), in each one of the $L$ periods. She does not monitor and is not monitored by anyone.

At the end of the segment each player $k \neq i$ announces the sequence of $L$ pure actions that she played the periods in which she monitored other players and her observations in these periods. We need to make each player $k$ indifferent among the different sequences of pure actions that she can play and indifferent among the periods in which she monitors. The way to do that is to condition her payoff in the subgame that follows the minmax segment

on her announcements at the end of the segment so that her **total** discounted payoff is fixed. The precise way in which this is done is as follows. Let $(A_{-i})^L$ denote the sequence of profiles of pure actions of the players different from $i$ in the minmax segment. Let $d(A_{-i})^L \to R_+^{N-1}$ be a function such that for any two elements of $(A_{-i})^L$, $\sigma_{-i}(t)$ and $\tau_{-i}(t)$, $t = 1, ..., L$, and every $k \in N - \{i\}$, $k$ is indifferent between the path of payoffs that start with the sequence $g_k(\sigma_{-i}(t), m_i^i)$, $t = 1, ..., L$, and continues with a constant payoff $d(\sigma_{-i}(t))_k$, and the path of payoffs that starts with the sequence $g_k(\tau_{-i}(t), m_i^i)$ and continues with constant payoff $d(\tau_{-i}(t))_k$. Now it is easy to see that for any given $L$ and $\epsilon > 0$, if $\delta$ is close enough to 1, we can define $d$ so that for every $\sigma_{-i}(t) \in (A_{-i})^L$, $d(\sigma_{-i}(t))_k$ is a nonnegative number which is smaller than $\epsilon$. Let $C_t$ be the monitoring cost incurred by player $k$, according to her announcement, at period $t$, $1 \le t \le L$, in the minmax segment. Define $\epsilon_k'$ by the equation:

$$\sum_{t=L}^{\infty} \delta^t \epsilon_k' = \sum_{t=0}^{L-1} \delta^t \cdot C_t.$$

Thus, $\epsilon_k'$ is the increase per period in the payoff of $k$ which would precisely compensate $k$ for her cost of monitoring. Again, it is easy to see that for any $L$ and $\epsilon > 0$ if $\delta$ is close enough to 1 then $\epsilon_k' < \epsilon$.

Define the vector payoff $\bar{v}$ as follows:

$$\begin{aligned}
\bar{v}_k &= v_k^* + d(\sigma_{-i}(t))_k + \epsilon_k', \ k \neq i, j \\
\bar{v}_j &= v_j^* - \eta + d(\sigma_{-i}(t))_j + \epsilon_j' \\
\bar{v}_i &= v_i^* - 2\eta.
\end{aligned}$$

If, at the end of the minmax segment, the announcements of the players are consistent, then the continuation of the phase is identical to the phase $r(\bar{v})$ and the monitoring assignment has the property that players $i$ and $j$ do not monitor each other.

If the announcements are inconsistent then the next phase is determined by the transition rule that was defined in the main proof (that is, the pair of players that contradict each other are punished).

**Some additional clarifying comments to the proof in Section 4 regarding:**

      **1)     Deviations where a player observes other players more than once.**

      **2)     The beliefs of players outside the equilibrium path.**

1.     Suppose that player 1 deviates by observing the actions of other players in several, (or all), periods in a given segment. It is easy to see that player $i$ cannot gain from such a deviation because player $i$ remains uncertain about the period in which his action is monitored so he has no incentive to deviate by switching more than twice from the prescriptions of the action function $f^v$.

However, there is an additional potential difficulty in the construction of our equilibrium. Consider an out of equilibrium history, $h$, in which player $i$ has deviated by observing the actions of all, or some, of the other players in, say, all the periods except the last one in the current segment, $seg$. (That is, the period that follows $h$ is the last period in $seg$). Assume furthermore that some players different from $i$ have deviated by switching from the action function more than twice so that player $i$ knows that there is a positive probability that players different from himself will make contradicting statements at the end of the segment. (This probability might in fact be close to 1 if, for example, player $i$ observes that some player $j$ has deviated in all the $M-1$ periods, $M = K, L$, that have been observed.) Now this might make it worthwhile for player $i$ to deviate from $f^v$ in the last period because the probability that he will be punished is smaller than $\frac{1}{M}$. Specifically, suppose that player $i$ has switched twice in the first $M$ periods of, $seg$, and now upon observing that other players have deviated decides to switch for the third time in the last period of $seg$. To see that this deviation is in fact not worthwhile recall that in the case where there are several pairs

of players that contradict each other the transition function selects at random (from the set of contradicting pairs), a pair of players to be punished. Now this means that even in the case where player $i$ observed deviations by other players his decision to deviate at the last period of $seg$, (or at any other period), increases the probability that he will be selected for punishment by at least $\frac{1}{M} \cdot \frac{1}{N}$. This follows because the probability that the monitor of $i$, $q(i)$, will choose to observe $i$ in the last period of $seg$ is $\frac{1}{M}$ and because there are at most $N$ pairs of players which can contradict each other (a pair $(j, q(j))$ for every player $j$.)

Going back to the proof, we easily observe that for any positive probability $p$, (in particular for $p = \frac{1}{M} \cdot \frac{1}{N}$), there exists $\bar{L}(p)$ and $\bar{\delta}(p)$ such that if $L > \bar{L}(p)$ and $\delta > \bar{\delta}(p)$ then it is not worthwhile to deviate when the probability of punishment is greater than $p$. Specifically, the arguments on pages 22, 23 and 24 carry through if we substitute $\frac{1}{K}$ in expression (5) and $\frac{1}{M}$ in expression (7) by any positive number $p$.

2. We have not been explicit about the definition of the beliefs outside the equilibrium path. This deserves a comment. The point is that the beliefs of the players are not really important, or more precisely, the condition of sequential rationality is satisfied for a general class of beliefs including **all** the beliefs which satisfy the condition of consistency. To see this note that since the game is a repeated game with complete information, that is, there is no uncertainty about the payoffs in the one-shot game, the only significance of the imperfect monitoring to the satisfaction of the condition of sequential rationality is the uncertainty that it generates about the behavior of the players in the continuation of the game. In other words the beliefs of a player about what happened in the past do not change her optimal behavior in the continuation of the game as long as her beliefs about the behavior of the players in the continuation of the game remain the same. Now, the announcements of the players at the end of each segment[18] determine the strategies of the players in the continuation of the

---

[18]Plus the outcome of a lottery if several pairs of players contradicted each other.

game. So in a period $t, 1 \leq t \leq, M$ in a segment, *seg*, of length $M$, $M = K, L$, the beliefs of a player $i$ about what happened in the periods that preceded *seg* are completely irrelevant, any belief, and in particular beliefs that satisfy consistency, would work. What matters are the beliefs of player $i$ about the play in *seg*. The discussion in the previous paragraph highlights the fact that the only property of the beliefs of player $i$ which is required to make her strategy an optimal response is that for each period $\bar{t}$, $\bar{t} \geq t$, player $i$ assigns a positive probability which is bounded from below to being monitored at $\bar{t}$. Call this property $P$. It is easy to see that each system of beliefs which satisfies consistency satisfies $P$. Along the equilibrium path $P$ follows from Bayesian updating and outside the equilibrium path $P$ follows from belief revision which assumes a minimal number of deviations, (a condition which is implied by consistency).

# References

[1] D.Abreu, D.Pearce and E.Stachetti, "Optimal Cartel Equilibria with Imperfect Monitoring," **Journal of Economic Theory**, 39 (1986), 251-269.

[2] D.Abreu, D.Pearce and E.Stachetti, "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," **Econometrica**, 58 (1990), 1041-1064.

[3] R.J.Aumann and L.Shapley, "Long-term Competition: A Game Theoretic Analysis," mimeo, 1976.

[4] E.Ben-Porath and M.Kahneman, "Communication in Repeated Games with Private Monitoring," **Journal of Economic Theory**, 76 (1996), 281-297.

[5] O. Compte, "Communication in Repeated Games with Imperfect Private Monitoring," **Econometrica**, 66 (1998), 597-626.

[6] E. Green and R. Porter, "Non-Cooperative Collusion under Imperfect Price Information," **Econometrica**, 52 (1984), 87-100.

[7] D.Fudenberg and D.Levine, "An Approximate Folk Theorem with Imperfect Private Information," **Journal of Economic Theory**, 54 (1991), 26-47.

[8] D.Fudenberg, D.Levine and E.Maskin, "The Folk Theorem with Imperfect Public Information," **Econometrica**, 62 (1994), 997-1040.

[9] D.Fudenberg and E.Maskin, "Folk Theorems for Repeated Games with Discounting and Incomplete Information," **Econometrica**, 54 (1986), 533-554.

[10] M.Kandori and H.Matsushima, "Private Observation, Communication and Collusion," **Econometrica**, 66 (1998), 627-652.

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

[12] E.Lehrer, "Nash Equilibria of n-player Repeated Games with Semi-standard Information," **International Journal of Game Theory**, 19 (1990).

[13] _____ "Correlated Equilibria in Two-Player Repeated Games with Nonobservable Actions," **Mathematics of Operations Research**, 17 (1992), 175-199.

[14] _____ "Two-Player Repeated Games with Nonobservable Actions and Observable Payoffs," **Mathematics of Operation Research**, 17 (1992), 200-224.*

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

[16] R.Radner, R.Myerson and E.Maskin, "An Example of a Repeated Partnership Game with Discounting and with Uniformly Inefficient Equilibria," **Review of Economic Studies**, 53 (1986), 59-70.

[17] A.Rubinstein, "Equilibrium in Supergames with the Overtaking Criterion," **Journal of Economic Theory**, 27 (1979), 1-9.