Confession and Pardon in Repeated Games with Private Monitoring and Communication

Galit Ashkenazi-Golan

December 21, 2006

Abstract

We investigate multi-player discounted repeated games with private monitoring and communication. After each period, every player observes a random private signal whose distribution depends on the common action taken.

Principal-agent relationships, those between business-partners and between investors and entrepreneurs, these are all examples of economical environments that involve a small number of agents, and relationships that are long-lasting and enable frequent communication. We focus on observation assumptions and an equilibrium construction that are most likely to occur in such cases.

We assume that deviating from certain actions might reduce the contents of information received by the deviator, or may even be observed with some probability. Under this assumption we obtain a Nash-threat folk theorem.

In equilibrium players are provided with incentives to report a deviation when they detect one. Moreover, in equilibrium, the deviating player has an incentive to confess his deviation. This is done by making a punishment that follows a confession lighter than a punishment that does not follow a confession. Thus, a confession induces a pardon, and a choice between a variety of sanctions enables efficient outcomes.

Key Words: repeated games; private monitoring; moral hazard.

1 Introduction

Long-lasting economical relationships between a small number of agents are often characterized by frequent communication and private monitoring. Such interaction enables flexibility to choose from a variety of continuations in case problems occur. One example of such relationships are principle-agents games. Another example are games of partnership. The theory of infinitely-repeated private-monitoring games with communication attempts to model such interactions.

Consider, as a general example, principal-agent relationships under subjective evaluation (which is equivalent to private monitoring). In case where a deviation is not necessarily detected, there could be room for shirking. To discourage shirking, the principal might resort to a harsh punishment (such as firing) upon detection. Firing, however, would eliminate the possibility of future gains from the continuation of the relationships. Thus, harsh punishments are a non-credible threat in a situation where there is a positive continuation value to the relationship for the principal, and, in addition, the agent cannot know whether the shirking was detected or not. We offer a construction that restores contract-enforcing and, in turn, efficiency. This can be facilitated by the incorporation of a range of sanctions to be applied once a deviation is detected. With regard to the literature of optimal contracting, this suggests that termination of the contract is not sufficient for efficiency, a result in the spirit of McLeod (2003). From within the range of sanctions, lighter punishments will be employed if shirking was followed by a "confession". Hence, a post-deviation discussion between the employer and the employee, followed by a potential amicable arrangement is a necessary part of our construction.

A second example one could imagine is of two partners who are working together on a project from a distance (programming a software, designing a campaign, writing a paper etc.). Partner A decides to take a few days off. This action is un-observable by Partner B. Partner B calls and demands to have answers regarding e-mails he sent. Partner A can either "confess" not reading them or he can try to guess their content and try and relate to them. If Partner A attempts to "fake" reading the e-mails he risks exposing his ignorance which in turn reveals his shirking.

Yet a third example is that of a dinner in a restaurant. Occasionally, the restaurant manager figures out that the dinner's favorite dish falls short of standards. On such occasions, the manager faces a dilemma. On the one hand he could keep silent, hoping that the dinner would not notice the difference, thus risking disappointment and dissatisfaction, which may lead the dinner to refrain from returning to the restaurant for a long period of time (a "harsh" punishment). On the other hand, he could "confess", maybe loose the current meal of the dinner (a "soft" punishment), and retain the customer's trust.

In the principal-agent example, a deviation from an agreement results in a positive probability for the deviation being detected. In the partnership example, the deviation results in a loss of information, the content of the e-mails, and when communication between the agents takes place, this loss, which indicates a deviation, may be exposed.

We are interested in long-term interactions that involve communication, where the agents try to establish cooperation and trust under possibly profitable deviations. If both players are interested in preserving cooperation, and if there is only one sanction available, then the private monitoring could make the implementation of such a sanction un-credible. If the deviator cannot know whether his deviation is observed, then the opponent would choose to overlook the deviation in order to sustain the cooperation. However, if there are a few levels of sanctions, then a more truthful communication can be established. Assume that if a deviator "confesses" a deviation, a lighter sanction is employed, and in case a deviation is detected and no confession is made, a heavier sanction is executed. In such a case both players are motivated to communicate truthfully about their observations and past actions. This is in line with the notion that confessing a "mistake" might reduce the long-lasting damage to the trust in a relationship.

Trust is essential when players cannot fully monitor each other, and it is easier to establish when the number of players is relatively small and they communicate with each other frequently.

Theoretical results that apply to the model of two-player repeated discounted games with private monitoring are relatively scarce. There are some results concerning prisoner's dilemma and a recent working paper by Obara (2005) which assumes an observation mechanism different from ours (see details in the literature survey).

The main result of the paper is based on an assumption that refers only to payoffs, called extreme payoffs, which are extreme points of the set of all possible payoffs (the feasible set). We assume that every deviation from a common action, whose payoff is extreme, is detectable in one of two ways. The first is by observing the deviation. This might occur when a deviation induces a positive probability for at least one profile of private signals (of the opponents) assigned a zero probability under the distribution corresponding to the agreed upon joint action.

The second way to detect a deviation is the indirect one: the deviation may result in a loss of information that the deviator would otherwise receive. On equilibrium paths, the players are supposed to publicly report their private signal after each period. A report that is not consistent with the equilibrium path, signifies that at least one player has deviated. While deviating a player might loose information necessary to report in a manner that is consistent with others' report. Thus, a possible loss of information upon deviation makes the deviation detectable.

A signal of a deviator will be called sufficiently informative if it has two properties. First, it lets her know that her opponents' signals are consistent with the equilibrium path; and second, it allows her to complete these signals with a signal of her own, so that all together they are consistent with the equilibrium path. A sufficiently informative signal enables a player to get away with a deviation. We assume that, having deviated from an action profile whose payoff is extreme, a deviator observes a sufficiently informative signal with a probability strictly less than 1.

A detection of a deviation is followed by a punishment. The following punishment mechanism enforces truth-telling when the players are communicating. There are three types of punishment phases, differing according to their durations: a short-term, a medium-term, and a long-term. During any punishment phase, a one-period Nash-Equilibrium is played. A short-term punishment phase will take place when first, all the players, but the deviator, announce a combination of signals which means that a deviation took place and, second, the deviator confesses a deviation. A medium-term punishment phase takes place when only the deviator confesses, and the combination of signals of the other players is consistent with the agreed upon action. And a long-term punishment phase takes place when the deviator does not confess, while the reports are inconsistent with the equilibrium path.

According to the construction, when a player "confesses" a deviation, the harshest punishment she might get is the medium-term punishment. When a player does not confess, she risks the long-term punishment. The duration of the various punishment phases are designed to provide a deviator a motivation to "confess" her deviation even if the deviation has a small probability of being detected. The opponents, in turn, know that whenever they observe a signal-profile indicating a deviation, it will be followed by a confession. Such a confession results in a short-term punishment if the opponents' reports will indicate a deviation and in a medium-term punishment otherwise. Since, a short term punishment is preferred to a medium-term punishment, the players have incentives to truthfully report their signals. After the punishment, the players restart the equilibrium path.

A punishment phase is triggered by both the deviator's confession and her opponents' reports. Thus, for a player observing the signal which indicates a deviation, overlooking the deviation is unprofitable since it is accompanied by a confession of the deviator. For a deviator who receives a signal that is not sufficiently informative, trying to avoid punishment by not confessing is un-profitable due to the positive probability of the deviation being detected.

The sequel of the paper is organized as follows: Section 2 presents examples of games for which our model and assumptions apply naturally. In Section 3 introduces the formal model. Section 4 contains the main result - a Nash-threat folk theorem. This result is discussed in Section 5. A literature survey is conducted in Section 6. The appendix integrates our results with those of Kandori and Matsushima (1998) and Compte (1998), in order to obtain a richer set of sequential equilibria payoffs.

2 Detecting Deviations with Private Monitoring and Communication - Examples

2.1 Example 1: a partnership game

Consider the following partnership game. In this game there are two partners, each one of them can "work" (w) or "shirk" (s)[1]. The expected payoffs from the actions are given in table 1:

	w	s
w	$(1,1)$	$(-1,2)$
s	$(2,-1)$	$(0,0)$

Since the profitable deviations are only from "work" to "shirk", we shall concentrate on deterring only such deviations.

[1]This example appears in Fudenberg et al(1994)

One example of an observation mechanism that will enable the player to detect deviations is a variation of the standard-trivial model, introduced by Lehrer (1990). In the standard-trivial model the signals of the players are either the action taken by the opponent or a null signal. If there is always a positive probability that the opponent observes a signal indicating the action, then deviations can be detected. We don't need to make any assumption about the correlation of the signals, so they can be perfectly correlated (which gives a public-monitoring game), independent, or any other form.

Note that in the independent signals case (as well as in many other cases), when a player observes a signal which indicates a deviation, he knows that the deviator does not know whether the deviation has been detected. We shall use the one-period equilibrium payoff $(0, 0)$ as a punishment, so the punishment will be costly for the punisher as well. In this case, we will need to motivate a player observing a deviation to report it, even though the punishment is costly, and even though he knows that the deviator does not know that the deviation has been detected. Motivating such a truthful communication will be done using the pardon and the variety of sanctions available in this game.

2.2 Example 2: A Principal-Agent Game

A second example of an observation mechanism that satisfies our assumption can be found in principal-agent games.

The actions of the principal are perfectly observed; the agent exerts effort, and (only) the principal observes a stochastic result. The imperfectness of the monitoring gives rise to moral-hazard situations, and the one-period equilibrium is often strictly Pareto-dominated by the efficient outcome.

In some games, when the agent chooses not to exert enough effort (not to study a subject thoroughly enough, not to be present in certain situations or places) the agent risks losing information. In the following communication between the principal and the agent, if there is a chance that this loss of information will be exposed, then we will say that the deviation (of not exerting enough effort) is detectable.

3 Preliminaries

The model outlined below features n-player repeated games with stochastic private signals. After each period, the players observe private signals, and after observing their private signals, the players simultaneously send public messages.

3.1 The Stage Game

In the stage game, players move simultaneously and each player $i \in N$ chooses an action a_i from a finite set of actions A_i. After actions are played, each player i observes a signal, y^i which is not observed by the opponents. Let n be the number of players, $|N|$. Let Y_i be the finite set of possible private signals for player i. A *signal profile* is an n-tuple $y = (y^1, ..., y^n) \in Y = \Pi_{i \in N} Y_i$. Each *action profile* $a = (a_1, ..., a_n) \in A \equiv \times_{i \in N} A_i$ induces a probability distribution over signal profiles. Let $p(\cdot|a)$ be the common distribution of the private signals, conditioned on the common action a. Let $(a_{-i}, a'_i) \in \times_{j \neq i} A_j \times A_i$ be the action profile where all the players but i follow the action profile a, and player i is playing a'_i. Let $y_{-i} \in Y_{-i} = \times_{j \neq i} Y_j$ be the signal profile of the opponents of player i. Let $p_{-i}(\cdot|a)$ be the common distribution of the signals of the opponents of player i when the common action taken is a. Let $q_i(y_{-i}|a, y_i)$ be the probability that the opponents of player i received the signal profile y_{-i} when the action profile was a and the signal player i received was y_i.

Each player i's mean payoff $g_i(a)$ depends on the action profile played. The realized payoff can be dependent on the signals, and in general is not known to the player.

We allow players to *communicate* with each other. After choosing actions and observing their private signals, the players simultaneously and publicly announce messages. Player i announces a message taken from the finite set M_i. Thus, a profile of messages is $(m_1, ..., m_n)$, where $m_i \in M_i$ for $i = 1, ..., n$.

3.2 The Repeated Game

At each date $t = 1, 2, ...$ the stage game is played and private signals are observed. At the end of period t, the private history of player i consists of player i's past actions, past private signals, and the public messages: $h_i^t = (a_i(1), y^i(1), w(1), ..., a_i(t), y^i(t), w(t))$. We denote by $h_i(0)$ the null private history of player i. We denote by h the sequence of actions, signals and messages taken so far (the history of the game), and by $\mathcal{H}$ the set of all possible histories. A pure strategy (σ_i, τ_i) for player i is a pair of sequences of maps , $\{\sigma_i^t\}_{t=1}^{\infty}, \{\tau_i^t\}_{t=1}^{\infty}$, where σ_i^t maps each history that ends with the public messages, to an action in A_i to be taken in the next period, and τ_i^t maps each history that ends with a private signal, to the public message the player should announce.

Formally, a strategy is (σ_i^t, τ_i^t),

$\sigma_i^t : \times_{t'=1,..,t} A_i \times_{t'=1,..,t} Y_i \times_{t'=1,..,t} (W(1), .., W(n)) \to A_i$

$\tau_i^t : \times_{t'=1,..,t} A_i \times_{t'=1,..,t-1} Y_i \times_{t'=1,..,t-1} (W(1), .., W(n)) \to Y_i$

Each strategy profile $(\sigma, \tau) = \times_{i \in N}(\sigma_i, \tau_i)$ generates a probability distribution over future streams of actions, payoffs and messages, which in turn induces a distribution over future payoffs. Players are assumed to discount future payoffs with a common discount factor δ.

Player i's average discounted expected payoff from σ is

$v_i(\sigma, \tau) = (1 - \delta)E[\sum_{t \geq 1} \delta^{t-1} g_i(a(t))]$.

3.3 Sequential Equilibria

A strategy profile (σ, τ) is a Nash equilibrium if and only if for any player i and for any strategy (σ_i', τ_i'), $v_i(\sigma_i, \tau_i) \geq v_i((\sigma_i', \tau_i')(\sigma_{-i}, \tau_{-i}))$.

An *assessment* is a pair (σ, μ), where σ is a profile of behavioral strategies and μ is a function that assigns to every information set a probability measure on the set of histories in the information set. We shall refer to $\mu(h_i, h)$ as the beliefs of the players, the probabilities a player assigns to the history $h \in \mathcal{H}$ conditional on the private history h_i being observed.

The assessment (σ, τ, μ) is *sequentially rational* if for every player $i \in N$ and for every information set of player i, the strategy of player i is a best response to the strategies of the other players, given the information set.

An assessment is *consistent* if there is a sequence $((\sigma^n, \tau^n \mu^n))_{n=1}^{\infty}$ of assessments that converges to (σ, τ, μ) in Euclidian space and has the properties that each strategy profile (σ^n, τ^n) is completely mixed (meaning that it assigns positive probability to every action at every information set) and that each belief system μ^n is derived from (σ^n, τ^n) using Bayes' rule.

An assessment is a *sequential equilibrium* if it is sequentially rational and consistent.

3.4 Observation Assumption

We shall call a private signal observed by a deviating player *sufficiently informative* if it enables the deviator to choose a private signal that will complete the signals of her opponents to a signal profile that is consistent with the agreed upon action. The signal will need to indicate that the opponents' common signal does not indicate a deviation and to allow choosing a private signal that will complete any signal-profiles that the opponents might have (assuming the opponents did not deviate and given the private signal) to a common signal that is consistent with the equilibrium path.

We assume that every profitable deviation from a pure action profile whose payoff is an extreme point of the set of feasible payoffs induces a probability strictly less than one for the deviator to observe a sufficiently informative signal.

Formally, let $EX \in A$ be the set of action profiles whose payoff is an extreme point of the set of feasible payoffs.

Definition 1 *A message of player i, announcing that he observed a signal y_i following a period when player i played a'_i and observed y'_i is* consistent with action-profile a *if for every signal profile of player i's opponents, $y_{-i} \in Y_{-i}$ such that $p(y_{-i}y'_i|a_{-i}, a'_i) > 0$ it holds that $p(y_{-i}, y_i|a) > 0$.*

Definition 2 *A signal y_i of player i, following a deviation a'_i from the common action profile a is* sufficiently informative *if there exists a signal y'_i that is consistent with action-profile a.*

We make the following assumption:

Assumption 1: *Every profitable deviation from a pure action profile $a \in EX$ induces a probability strictly less than one for the deviator to observe a sufficiently informative signal.*

Assumption 1 is equivalent to assuming that every profitable deviation from a pure action profile $a \in EX$ is detectable.

4 Communication, Deviations and Confessions - Nash Threat Folk Theorem

In this section we shall prove the Nash-threat folk theorem.

The main idea of our equilibrium construction is to provide a player with an incentive to "confess" a deviation if she indeed deviated and to "report" a deviation if she observed it.

Denote by V the set of feasible payoffs, by V^* the set sequential equilibria payoffs and by V^{**} the set of feasible payoffs Pareto dominating one-period equilibrium of the repeated game.

Theorem 1 *When the players are allowed to communicate, $V^{**} \subseteq V^*$.*[2]

Proof.

Let $v = (v_1, v_2, ..., v_n) \in V^{**}$. We shall follow the pure-action equilibrium path, as constructed by Fudenberg and Maskin. In their equilibrium path the continuation payoff is always within an ε distance of v. The exact ε that we shall use is dependent on the payoffs and the information structure, as will be shown in the following.

The equilibria strategy we construct is to follow the equilibrium path as in Fudenberg and Maskin, and after each period convey a public message that informs the opponents of the private signal that was observed, until a deviation from this path has occurred. If a player deviated and observes a message that is not sufficiently informative, he conveys a special message, "confessing" his deviation, a message that informs the opponents that he deviated in the last period.

To create the proper incentives to convey these messages (that could trigger a punishment phase), three different punishments are constructed: a short-term punishment in case both the deviator confessed his deviation and his opponents conveyed a private messages profile that indicates that a deviation took place; a medium-term punishment in case only the deviator conveyed his confessing message; and an "eternal" punishment in case the signal profile reported is inconsistent with the equilibrium path, but no player confessed a deviation.[3]

[2]In fact, Let W be the convex-hull of all the payoffs of pure-actions such that every profitable deviation induces a probability strictly less than one for the deviator to observe a sufficiently informative signal. Then the following proof can be used to show that any payoff inside W can be supported as sequential equilibrium payoff for patient enough players.

[3]In fact, the definition of sequential equilibrium demands that we shall also deal with cases of simultaneous deviations of two players or more. It is easy to see that it can be solved in the same fashion with the assumption of more messages, for example, a message which means :"I deviated and I observe the signal x". The length of the punishment will be determined, again, by how much all messages are consistent.

The three possible punishments will be three durations of punishment phases when all players play the one-period equilibrium, followed by restarting the equilibrium path. The length (number of periods) of the short-term punishment will be L_1, of the medium-term L_2 and the long-term punishment will last forever. Without loss of generality, we shall assume that the dominated one-period equilibrium payoffs are 0 for all players. The lengths of punishments will be the same for the different deviations of the different players, and therefore there is no need to specify who was the deviator in case that there was no confession. The punishments are the same.

Let $\overline{G}$ be the maximal one-period payoff over all players. Let $1-p$ be the maximal probability, over all players and all profitable deviations from all common actions whose payoffs are in EX, that the deviator will observe a signal that is sufficiently informative (that he will "get away" with the deviation). We shall induce punishment whenever the signal observed by the deviator is not sufficiently informative.

Now, with probability at least p the deviator will have a signal that is not sufficiently informative. When the signal is not sufficiently informative, then every choice of the deviator of a message to convey leads with a positive probability to detecting the deviation (because the signal profile reported will be inconsistent with the equilibrium path). Let that (positive) probability be r. Let r' be the minimum of the r's, over all players, and all their combinations of deviation and not sufficiently informative signals.

We shall describe the three punishment phases - short-term in case both the deviator's opponents report a signal combination that is inconsistent with the agreed upon common action and a signal of confession is observed, medium-term in case there was only a confession and long-term, in case the signal profile announced is inconsistent with the equilibrium path instructions, but no confession was announced.

From the deviator's point of view, when his signal is not sufficiently informative, confessing will be followed, at worst, with a medium-term punishment, (a "pardon"). Sufficiently large difference between the long-term and the medium-term punishments will induce the deviator to confess. From the deviator's opponents point of view, if before sharing the private signal, the opponent does not know that his private signal will help detecting a deviation - then there is no harm in announcing it. If he does know, then the following argument holds: if his signal is a part of a signal profile which indicates a deviation, then the deviator cannot have a signal that is sufficiently informative, hence he will confess. So the choice is between reporting and continuing to the short-term punishment and not reporting, which will result in the medium-term punishment. Any difference between the short-term and the medium-term punishment will suffice to induce reporting a deviation.

For the description above to be an equilibrium, it should be that for all players:

When the deviator observes a signal that is not sufficiently informative, confessing is more profitable than not confessing:

(1) $v_i^M > (1 - r')(v_i + \varepsilon)$

Staying in the equilibrium path is more profitable than deviating (when deviating, with probability at most $1 - p$ there is no punishment, and with probability at least p there is at least the small punishment):

(2) $v_i - \varepsilon > \overline{G}(1 - \underline{\delta}) + \underline{\delta}(1 - p)(v_i + \varepsilon) + \underline{\delta} p v_i^S$

Reporting is more profitable than not reporting:

(3) $v_i^S > v_i^M$

where,

(4) $v_i^S = \underline{\delta}^{L_1} v_i$

(5) $v_i^M = \underline{\delta}^{L_2} v_i$

We need to show that for δ close enough to 1, there are values for v_i^S and v_i^M which solve (1), (2) and (3) for all the players.

First, we note that since $r' > 0$, we can find v_i^S and v_i^M such that

$(1-r')v_i + \frac{1}{2}r'v_i < v_i^S < (1-r')v_i + \frac{3}{4}r'v_i$

$(1-r')v_i + \frac{1}{4}r'v_i < v_i^M < (1-r')v_i + \frac{1}{2}r'v_i$

We get the following inequality:

(*) $v_i^M > (1-r')v_i$

In addition, we have:
(**) $v_i^S < (1-r')v_i + \frac{3}{4}r'v_i < v_i$

Inequality (*) is inequality (1) for $\varepsilon = 0$; inequality (**) is inequality (2) for $\varepsilon = 0$ and $\delta = 1$; and inequality (3) is also satisfied. Since the inequalities are satisfied strictly and since they are continuous in δ and ϵ, then for δ close enough to 1 and ε close enough to 0 they will be satisfied as well. We might need to further increase δ in order to have enough flexibility for choosing proper L_1 and L_2 such that:

$(1-r')v_i + \frac{1}{2}r'v_i < \delta^{L_1}v_i < (1-r')v_i + \frac{3}{4}r'v_i$

$(1-r')v_i + \frac{1}{4}r'v_i < \delta^{L_2}v_i < (1-r')v_i + \frac{1}{2}r'v_i$

■

5 Discussion

5.1 Variety of Sanctions as a Tool to Enable Efficiency, Transparency and Trust

When monitoring is private, the transparency is lost, hence there is a need to establish some form of "trust" between the players. Communication enables us to restore this transparency. However, when punishments are too harsh, achieving truthful communication when shirking takes place becomes problematic. Restoring transparency and trust is due to the variety of sanctions that is available to the players. The possibility to "Pardon" and to reduce the punishment in case of truthful revelation of past "mistakes" is crucial.

This notion is in line with some basic intuitions. First, that flexibility is an important tool in long-lasting relationships; and second, it is plausible that truthful communication regarding past "mistakes" enhances the "trust" in the relationship, i.e., a deviation that is "confessed" is less destructive for future cooperation than a deviation that is not accompanied by a confession.

5.2 More Than Two Players and the Identity of the Deviator

If the game that is played is a game of more than two players, for example, when a principal is facing a team of agents, then often it is a challenge is to establish the identity of the deviator. Since the punishment in our construction is a punishment for all the players - this problem is simply solved. If a deviation was detected, and none of the players "confessed" then all the players are being simultaneously punished with the harsh punishment.

5.3 The Connection to Fudenberg and Maskin (1991)

We use the equilibrium-path description of Fudenberg and Maskin, however, we obtain a weaker result - they obtain all the feasible individually rational payoffs as perfect equilibrium payoffs while we obtain only those Pareto dominating one-period equilibrium payoffs. The reason is that we use a different punishment system because of the imperfect monitoring. Their punishments are to minimax the deviator for a number of periods. Since in our construction we rely on the players reporting the deviations of their opponents, we cannot trivially use their method since in general the player reporting a deviation can profit from minimaxing the alleged deviator, which would trigger false reports.

5.4 Mutual Minimaxing and Other Threats

It is natural to consider other mutual threat-points instead of the one-period Nash equilibrium. However, attempting to implement the confession-and-pardon construction in order to deter deviations *during* the punishment phase will fail. It will fail because it requires three lengths of punishment. "Restarting" a punishment phase in case a deviation during the punishment phase is detected will make deviations at the beginning of a long-term punishment profitable - it may replace the long punishment with a shorter one. Adding periods of punishments to the existing punishment phase will fail as well, since after large enough number of deviations, the weight of the additional punishment will become insignificant, and a deviation will become profitable.

Hence, stronger observation assumptions are required for the mutual-threat point that is not a one-period Nash equilibrium . We can instruct the players to add one period to the punishment in which a deviation is detected. That will make a deviation un-profitable if the detection probability is large enough compared to the possible profit. It will, on the other hand, lead the players to try and "avoid" detecting deviations.

For example, if there are two signals for player 1 that are consistent with the equilibrium path, but such that signal a "detects" certain deviations but signal b "detects" only a subset of those deviations, then player 1 will choose to report observing the "less informative" signal, signal b.

Exploring the possibility of enforcing a detection of deviation from the punishment phase calls for an extensive discussion regarding the "informativeness" of signals given the set of possible deviations and the observation mechanism. We feel that this is an interesting direction for further research.

6 Literature Review

The literature on discounted repeated games can be divided into two branches: games with perfect monitoring and games with imperfect monitoring. In the model of perfect monitoring each player observes after each period the actions played by all the others. Aumann and Shapley (1994) and Rubinstein (1994) characterize the equilibrium payoffs in such games with no discount. They state that every feasible and individually rational payoff vector is an equilibrium payoff. This result is known as folk theorem. Fudenberg and Maskin (1986, 1991) analyzed discounted games and obtained a folk theorem for perfect equilibrium.

Games with imperfect monitoring are further divided into games with public monitoring and games with private monitoring. Games with public monitoring are games where the players observe after each period a commonly known random signal $-$ a public signal. Players are assumed to have a perfect recall and they are allowed to condition their actions on previous data, including their own payoffs and signals. A player's strategy specifies how she should choose an action at any time and after every eventuality. A public strategy is such that actions are conditioned only on the history of public signals and not on the player's own previous actions.

In a perfect public equilibrium players are restricted to employ only public strategies. The set of perfect public equilibria payoffs of those games has been thoroughly investigated (Green and Porter, 1984, Abreu et al., 1990, Fudenberg et al., 1994). When the strategies of the players are not restricted to public strategies, and the players are allowed to condition their choices on the private histories of their own actions, the set of equilibria payoffs can be strictly larger than with public strategies. This phenomenon might occur when private histories can serve as a correlation means between the players' actions (see for instance, Mailath et al., 2002).

Games with imperfect private monitoring are games where each player observes a random private signal after each period. Such games present new difficulties. One of the prominent ones is to precisely characterize the ability to correlate between players who try to punish a deviator by using their private signals.

One way to overcome these difficulties is to allow the players to publicly communicate, as done by Matsushima (1990). After every stage the players are allowed to convey public messages that may depend on the history of their private signals. Ben-Porath and Kahneman (1996) proved a folk theorem for a model where each player can be perfectly observed by at least two others. In this model the players use a communication channel to report a deviation when they detect one. Compte (1998) and Kandori and Matsushima (1998) obtained folk theorems for games with three players or more, when the players are allowed to communicate. The latter two papers assume a full support on the set of signals. That is, each signal profile is observed with a positive probability after every history of actions. In addition, they assumed that any deviation induced a distribution over signal-profiles that allowed the non-deviating players to statistically detect the deviation and, moreover, the identity of the deviator.

Two recent developments in this area are those of Tomala (2005) and Obara (2005). Tomala employs a communication device and uses the concept of communication equilibrium (see Myerson, 1982 and Forges, 1986). He obtains conditions for folk theorem for the case three players or more. Obara generalizes Compte (1998) so that the results hold also for the case of two-players, and the assumptions of full-support of the signal-profiles, and the independence of the private signals are relaxed. He still requires full-support of the signals of each player (i.e., every signal is observed by each player with a positive probability after any common action played), so his informational assumptions are different, maybe complementary, to ours.

A model in which every profitable deviation is detected with a positive probability, is a version of the model of standard-trivial observation, introduced by Lehrer (1990). In this information structure for any two players, i and j, either player i fully observes the action of player j, or obtains no information of it. Lehrer characterized the set of equilibria payoffs for the two-players undiscounted infinitely repeated game, when the information is deterministic and symmetric - either both players are fully informed of the action played or both receive the null signal. This model is a public monitoring model. In contrast with Lehrer's model, here the signals can be either independent or correlated, as long as there is a positive probability for the deviations to be detected. Another paper which has an information structure that resembles ours, is that of Renault and Tomala (1998). They analyzed undiscounted repeated games where each player observed the actions of a subset of players.

Our construction of the equilibrium path is similar to that of Fudenberg and Maskin (1991) for the perfect observation case. In this equilibrium path the players are instructed to play a sequence of pure actions whose discounted average payoff approaches the desired one, and for which the continuation payoff is always within a predetermined ε-distance from that desired payoff. This way, since our desired payoff is strictly Pareto-dominating some one-period Nash equilibrium payoff, for a sufficiently small ε , the continuation payoff will also Pareto-dominate the same one-period Nash equilibrium payoff. Hence punishments using the one-period Nash equilibrium are always effective punishments.

In Ben Porath and Kahneman (1996) and Renault and Tomala (1998) the continuation payoff of each player depends on the reports of at least two of his opponents. Therefore, "overlooking" a deviation might lead to contradicting reports, and the players reporting would then be punished. In our paper we also rely on comparisons of two sources of information about a deviation. We reward reports of deviations by reducing the length of the punishment when they are announced simultaneously with the deviator's confession.

In our construction, since the players punish only in case a deviation indeed took place, punishments that do not preserve efficiency can be used without decreasing the equilibrium path payoff. To keep efficiency in the full-support of signal-profiles case, Compte (1998) and Kandori and Matsushima (1998) assume three players or more - they demonstrated that one can create incentives for players to report their signals by making each player's continuation payoff independent of his own message. In case there is a need to reduce another player's payoff as a result of the message, efficiency can be kept by giving the "surplus" to a third player. Observe that the kind of construction found in Kandori and Matsushima and Compte is applicable only in cases of three players or more. The difference in the monitoring assumptions allows our results to be applicable also to two-player games.

7 References

Abreu, D., D. Pearce and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring", *Econometrica*, **58**, 1041-1063.

Aumann, R. and L. Shapley (1994): "Long-Term Competition: A Game Theoretic Analysis," in Essays in Game Theory, edited by N. Megiddo, Springer-Verlag, pp. 1-15.

Ben-Porath, E. and M. Kahneman (1996): "Communication in Repeated Games with Private Monitoring", *Journal of Economic Theory*, **70**, 281-297.

Compte, O. (1998): "Communication in Repeated Games with Imperfect Private Monitoring ", *Econometica*, **66**, 597-626.

Forges, F. (1986): "An Approach to Communication Equilibria", *Econometrica*, **54**, 1375-1385.

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

Fudenberg, D., D. Levine and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information", *Econometrica*, **58**, 1041-1063.

Fudenberg, D. and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting or with Incomplete information", *Econometrica*, **54**, 533-554.

Fudenberg, D. and E. Maskin (1991): "On the Dispensability of Public Randomization in Discounted Repeated Games", *Econometrica*, **53**, 428-438.

Green, E. J. and R. H. Porter (1984): "Moral Hazard and Private Monitoring", *Econometrica*, **52**, 87-100.

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

Lehrer, E.(1990): "Nash Equilibria of n-Player Repeated Games with Semi-Standard Information", *International Journal of Game Theory*, **19**, 191-217.

— (1992a): "Two-Player Repeated Games with Nonobservable Actions and Observable Payoffs", *Mathematics of Operation Research*, **17**, 200-224.

— (1992b): "On the Equilibrium Payoffs Set of Two Player Repeated Games with Imperfect Monitoring", *International Journal of Game Theory*, **20**, 211-226.

MacLeod, W. B. (2003): "Optimal Contracting with Subjective Evaluation", *The American Economic Review*, **93**, 216-240.

Mailath, J. and S. Morris (2002): "Repeated Games with Almost-Public Monitoring", *Journal of Economic Theory*, **102**, 189-228.

Mailath J., A. Matthews, and T. Sekiguchi (2002): "Private Strategies in Finitely Repeated Games with Imperfect Public Monitoring", *Contributions to Theoretical Economics* Vol. 2: No. 1, Article 2.

Matsushima, H. (1990): "On the Theory of Repeated Games with Private Information: part II", *Economics Letters*, **35**, 257-261

Myerson, R.B. (1982): "Optimal Coordination Mechanisms in Generalized Principal - Agent Problems", *Journal of Mathematical Economics*, **10**, 67-81.

Obara, I. (2006): "Note on Folk Theorem with Communication", mimeo

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

Renault, J. and T. Tomala (1998): "Repeated Proximity Games", *International Journal of Game Theory*, **27**, 539-559.

Rubinstein, A. (1994): "Equilibrium in Supergames," in Essays in Game Theory, edited by N. Megiddo, Springer-Verlag, pp. 17-27.

Tomala, T (1998): "Pure Equilibria of Repeated Games with Public Observation", *International Journal of Game Theory*, **27**, 93-109.

— (2005): "Perfect Communication Equilibria in Repeated Games with Imperfect Monitoring", mimeo.

Sekiguchi, T. (1997): "Efficiency in Repeated Prisoners' Dilemma with Private Monitoring", *Journal of Economic Theory*, **76**, 345-361.

8 Appendix: Combined Theorem - Constant and Moving Support

In the papers of Kandori and Matsushima(1998) and Compte(1998), the authors prove several folk theorem for games with full support of the signals (all signal-profiles are observed with a positive probability after every action profile) and communication, when the number of players is at least 3. These results can be combined with ours in several ways, to enlarge the set of payoffs that can be supported as sequential equilibria payoffs. We first present those papers' results and main ideas, then two examples to demonstrate the synergy between their methods and ours, and then the combined theorem.

8.1 The Results of Kandori and Matsushima (1998), and Compte (1998)

Both the paper of Kandori and Matsushima (1998) and the paper of Compte (1998) prove folk theorem for games with private monitoring, when communication is allowed and with full support of the private signals profiles. Both papers use dynamic programming techniques and the assumption of at least three players. The papers use delay of the communication (meaningful communication is carried on only every k periods) to achieve efficiency.

Here are sufficient conditions under which there exists a folk theorem:

First assumption: Every deviation of a player i from the common action minimaxing player j, $j \neq i$, is either not profitable or statistically detectable by player j's opponents.

Officially: Let μ^i be the minimax profile for player i, when μ^i_j is the (possibly mixed) strategy of player j when player i is to be minimaxed.

(A1) - For all i and $j \neq i$, if there is a mixed strategy $\alpha_j \in \Delta A_j$ such that $p_{-j}(\cdot|\mu^i) = p_{-j}(\cdot|\mu^i_{-j}, \alpha_j)$ then $g_j(\mu^i) \geq g_j(\mu^i_{-j}, \alpha_j)$.

Second assumption: All mixed strategy deviations of every player i, are statistically detected by the i, j opponents, for every $j \neq i$. Define, for each pair $i \neq j$ and each action profile $a \in A$, $Q_{ij}(a) = \{p_{-ij}(a_{-i}, a'_i)|a'_i \in A_i\{a_i\}\}$. This is a collection of distributions of ij-opponents' signals, generated by player i's deviations from the profile a.

(A2) - For each player $i \neq j$ and each $a \in EX$,
$p_{-ij} \notin conv(Q_{ij}(a) \cup Q_{ji}(a))\}$.

Third assumption: For every two players i and $j \neq i$, the opponents of i and j can statistically discriminate player i's (possibly mixed) deviations from player j's. The deviations of the different players create different distributions of the signals of their opponents.

(A3) - For each pair $i \neq j$ and each $a \in Ex(A)$,
$conv(Q_{ij}(a) \cup \{p_{-ij}(a)\}) \cap conv(Q_{ji}(a) \cup \{p_{-ij}(a)\}) = \{p_{-ij}(a)\}$

Let v_i^* be the minimax value of player i and define the feasible and individually rational payoff set by

$W = \{v \in co(g(A)) | v \geq v^*\}$. Assume perfect support of the private signals profiles.

The main theorem is:

Theorem (Kandori and Matsushima): Suppose that there are more than two players ($n > 2$) and the information structure satisfies condition (A1), (A2) and (A3). Also suppose that the dimension of W is equal to the number of players. Then, any interior point in W can be achieved as a sequential equilibrium average payoff profile of the repeated game with communication, if the discount factor δ is close enough to 1.

8.2 Using Confessions and Reports Method to Support Dynamic Programming Methods

Consider the following game:

L	l	r	R	l	r
t	$(1,0,0)$	$(0,1,0)$	t	$(0,0,1)$	$(0,0,0)$
b	$(0,0,0)$	$(0,0,1)$	b	$(0,1,0)$	$(1,0,0)$

Assume that the signals to the three players are according to assumptions (A1) (A2) and (A3).

The one-period equilibrium is when each player randomizes with probability half for each action, and the payoff is $(1/4, 1/4, 1/4)$.

Now consider the following addition to the above game:

L	l	r	R	l	r
t	$(1,0,0)$	$(0,1,0)$	t	$(0,0,1)$	$(0,0,0)$
b	$(0,0,0)$	$(0,0,1)$	b	$(0,1,0)$	$(1,0,0)$
bb	$(5,-7,-7)$	$(5,-7,-7)$	bb	$(0,0,0)$	$(0,0,0)$

Note that now there is an additional equilibrium (bb, l, R) with the payoff $(0,0,0)$.

Assume that we add now another private signal for player 2 . Assume that when player 3 plays L and player 1 plays bb, this additional private signal is observed by player 2 and that the signal is observed with probability 0 when player 1 does not play the additional action, bb. Since the convex-hull of the original game is Pareto dominating the one-period equilibrium, we can still have the entire set of feasible individually rational payoffs as sequential equilibria payoffs. The set of feasible individually rational payoffs are all in the convex-hull of the original game (without the additional action). We can get all the payoffs of the original game through the method of Kandori and Matsushima, and in case player 1 deviates to bb when player 3 plays L, we can use our method of confession and reports - player 3 will convey a signal whose meaning is that a deviation took place, and player 1 will confess (the one period equilibrium that will be used as a punishment can be $(0,0,0)$). Under that construction, when player 3 plays R, there is no reason for player 1 to play his additional action, bb.

8.3 Supporting Confession and Reports Method with Dynamic Programming Methods

Consider the following version of the prisoners' dilemma. The signals can take the values 1 or 0:

	c	d
C	$(2,2)$	$(1-L, 2+H)$
D	$(2+H, 1-L)$	$(1,1)$

Kandori and Matsushimam (1998) proved folk theorem for this specific game, under the following monitoring-technology conditions:

- The signals of the players are independent given any pure action profile.

- The marginal distributions of the private signal of the players are symmetric, and $p_1(1|D,d) > p_1(1|D,c)$ and $p_1(1|D,c) > p_1(1|C,c)$.

Now consider the game with additional actions to the players:

	a	b	c	d
C	$(3, 1-0.5L)$	$(0,-1)$	$(2,2)$	$(1-L, 2+H)$
D	$(0,-1)$	$(0,-1)$	$(2+H, 1-L)$	$(1,1)$
E	$(0,-1)$	$(-1,-1)$	$(-1,0)$	$(-1,0)$

The minimax payoff is $(0,0)$

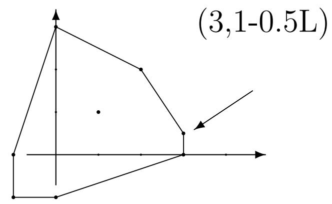

Figure 1. - Prisoner's Dilemma with additional actions

In order to support the entire efficient frontier as sequential equilibria payoffs, the common action (C,a) should be supported (see figure 1) . The payoff $(3, 1-0.5L)$ does not dominate the one-period equilibrium payoff $(1,1)$, however, if we assume that a deviation of player 2 from a to c or d when player 1 plays C induces a positive probability player 2 to observe a signal that is not sufficiently informative, we can still support this common action.

To see how, first note that now the entire convex-hull of $\{(2,2),(1-L,2+H),(2+H,1-L),(1,1)\}$ is individually rational. Let this convex-hull be U, Looking carefully at the construction of Kandori and Matsushima, one can verify that it still holds for the entire U.

Second, note that the payoff $(3, 1 - 0.5L)$ is Pareto dominating a two-dimensional non-empty subset of U. We can now replace the three lengths of punishments with three different continuation payoffs. There is a $\gamma > 0.5$ such that the payoff $(2 + 0.5H, 1 - \gamma L)$ is in the interior of U. We shall pick our three possible continuation payoffs, which are analog to the three lengths of punishment on the line connecting $(2 + 0.5H, 1 - \gamma L)$ and $(3, 1 - 0.5L)$. The continuation payoff $(2 + 0.5H, 1 - \gamma L)$ will be played in case player 1 announces observing the signal that is observed when player 1 deviates to playing a and player 2 is not confessing (long-term punishment) and two other points, closer to $(3, 1 - 0.5L)$ on this line can be chosen to supply the incentives for player 1 to confess and for player 2 to report a deviation (short-term and medium-term punishments analogs) if the players are patient enough. The logic of the proof is the same.

In this case, we support the confession-and-report construction not by the one-period equilibrium punishments, but rather by a set of payoffs that is itself achieved via dynamic programming construction. Note that this set has to be of dimension n.

8.4 The General Construction

In general, it is easy to see that one can use the following algorithm to find out the set of payoffs that can be supported as sequential equilibria payoffs when communication is allowed (denote it E), by combining dynamic-programming and confessions-and-reports methods:

1. Let E be the convex-hull of the one-period Nash Equilibrium payoff.

2. Add to E the convex-hull of the payoffs of all sub-matrices which:

a. Follow the conditions of Kandori and Matsushima,

b. Any deviation from the sub-matrix is either unprofitable or detectable.

c. Pareto dominating either one-period equilibrium payoff or three payoffs in the existing E such that one Pareto dominates the second which in turn Pareto dominates the third.

3. Go back to 2.

The proof follows the same logic as the two examples above.