The Repeated Prisoner's Dilemma with Imperfect Private Monitoring

Michele Piccione*
University of Southampton

May 31, 1999

Abstract

In this paper, I study a repeated Prisoner's Dilemma game in which monitoring is private and imperfect.

*I am grateful to Eddie Dekel, Stephen Morris, Sujoy Mukerji, Ariel Rubinstein, and two anonymous referees for their comments.

1 Introduction

The objective of this paper is to analyze a repeated Prisoner's Dilemma game in which the actions of the players are monitored *privately* and *imperfectly*.

Repeated games are generally studied under the assumption that players observe public (noisy) signals of the action played. Such signals can be used by the players to generate dynamic incentives capable of sustaining equilibrium outcomes which differ from static Nash behavior (see Green and Porter (1984), Abreu, Pearce, and Stacchetti (1986, 1990), Fudenberg, Levine, and Maskin (1994)). When public signals of the actions played are not available, the construction of dynamic incentives is less manageable and few results have thus far been obtained. Folk Theorems have been shown in Ben-Porath and Kahneman (1996), Compte (1998) and Kandori and Matsushima (1998) by allowing the players to communicate publicly. Mailath and Morris (1998) show that a Folk Theorem can also be obtained when signals are "almost" public in that they are highly correlated regardless of the actions chosen in the previous period.

In the model of this paper, no communication is allowed and signals can be either correlated or independent. In a similar framework, Sekiguchi (1997) shows that, when errors in observations are small and the discount rate is sufficiently high, a nearly efficient equilibrium can be achieved for some Prisoner's Dilemma games. I extend Sekiguchi's result in two directions. I show that, when errors are small and the discount rate is high, almost (in a sense to be explained later) the *entire* individually rational convex hull can be supported as the average payoff of Sequential equilibria in *any* Prisoner's Dilemma game.

The paper is organized as follows. In section 2, the model is presented. In section 3, an example is provided and, in section 4, the main result is proved. In section 5 the issue of non vanishing observation errors is discussed briefly and in section 6 some conclusions are given.

2 The Model

Consider an infinitely repeated (two player) Prisoner's Dilemma game with private monitoring. In each period, a player chooses an action from the set $\{C, D\}$. A player always correctly observes his own stage-game action but can only observe a (possibly imperfect) private signal of his opponent's action drawn from the set $\{C^{Ob}, D^{Ob}\}$. Specifically, I shall assume that if player i plays action $s_i \in \{C, D\}$ and player j plays action $s_j \in \{C, D\}$, there is a probability $\varepsilon_i(s_i, s_j)$ that player i will observe the opponent's action incorrectly, that is, player i will observe the signal $\tilde{s}_j^{Ob}$ different from s_j^{Ob}, $\tilde{s}_j^{Ob}, s_j^{Ob} \in \{C^{Ob}, D^{Ob}\}$. Let $\bar{\varepsilon}_i$ denote the 4-tuple consisting of the probabilities $\varepsilon_i(s_i, s_j)$ (the order is inessential).[1]

The stage-game payoffs of player i, $U_i : \{C, D\} \times \{C^{Ob}, D^{Ob}\} \to \Re$, depend on the action played by i and the action of the opponent observed by i. They are normalized as follows:

$U_i(C, C^{Ob}) = 1$
$U_i(C, D^{Ob}) = -L$
$U_i(D, C^{Ob}) = 1 + G$
$U_i(D, D^{Ob}) = 0$

for positive L and G and $G - L - 1 \geq 0$. The ex-ante stage game payoffs of player i from a stage-game action profile (s_i, s_j) are defined as $U_i^{ex} : \{C, D\} \times \{C, D\} \to \Re$ and can be normalized as follows:

$U_i^{ex}(C, C) = 1$
$U_i^{ex}(C, D) = -l(\bar{\varepsilon}_i)$
$U_i^{ex}(D, C) = 1 + g(\bar{\varepsilon}_i)$
$U_i^{ex}(D, D) = 0$

Let H^i be the set of (pure) histories observed by player i. A strategy for player i is a function $f_i : H^i \to [0, 1]$, with the convention that $f_i(h)$ is the probability with which player i plays C. For convenience, a history $h \in H^i$ is sometimes written as a pair (h_i, h_j), with the obvious interpretation. Given two histories, h and h', define their concatenation by hh'. Given a strategy profile f, player i's repeated game payoffs are

$$\Upsilon_i(f) = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} E(u_i(t) \mid f)$$

where $E(u_i(t) \mid f)$ denotes i's expected stage game payoffs at time t under f. Denote this game by Γ^{∞}.

[1] The primitive structure of the model would have to specify a probability distribution over the observations conditional on the each action profile played. For the purpose of this paper, however, the vector $\bar{\varepsilon}_i$ suffices for the analysis.

3 An Example

In this section, I provide an example which conveys the intuition behind the proof of the main result. Suppose that $L = 3$, $G = 2$,[2] and that, for any player i and pair of action (s_i, s_j), $\varepsilon_i(s_i, s_j) = \lambda$, $0 < \lambda < 1$. The argument consists of two steps. First, I consider the "imaginary" game in which both played and observed actions are common knowledge. By this step, the recursive structure is recovered. Second, I construct a strategy for player 2 for this new game which depends only the actions observed by player 2 in the *original* game and makes player 1 indifferent between playing C and D after any possible history. This strategy is constructed by generating the appropriate value function recursively. By constructing the identical strategy for player 1, one then obtains an equilibrium of the imaginary game. Since these strategies depend only on the (observed) histories in the original game, they are an equilibrium for this game as well.

Suppose that in the first period player 2 plays C and consider the system

$$V^0 = (1-\delta)(1-4\lambda) + \delta((1-\lambda)V^0 + \lambda V^1)$$

$$V^0 = (1-\delta)3(1-\lambda) + \delta((1-\lambda)V^1 + \lambda V^0).$$

Solving, $V^0 = \dfrac{1-8\lambda+7\lambda^2}{1-2\lambda}$ and $V^1 = \dfrac{-2-\lambda+3\delta-7\delta\lambda+7\delta\lambda^2}{\delta(1-2\lambda)}$. Hence, if player 1 gets a continuation value V^1 when player 2 observes D and V^0 when player 2 observes C, player 1 is indifferent between C and D.

Now consider the following system of difference equations:

$$V^t = (1-\delta)((1-4\lambda)m^t + (1-m^t)(-3+4\lambda)) + \delta((1-\lambda)V^0 + \lambda V^{t+1})$$

$$V^t = (1-\delta)(3(1-\lambda)m^t + (1-m^t)3\lambda) + \delta((1-\lambda)V^{t+1} + \lambda V^0)$$

If player 2 plays C with probability m^t and player 1 gets a continuation value V^{t+1} when player 2 sees D and V^0 when player 2 sees C, player 1 is indifferent between C and D. This system can be rewritten as,

$$m^t = \frac{14\delta\lambda^2 - 7\lambda^2 + 7\lambda - 15\delta\lambda + 2\lambda V^t - 3 - V^t + 4\delta}{(1-2\lambda)(7\lambda-4)(1-\delta)}$$

$$V^{t+1} = \frac{1}{\delta(4-7\lambda)}V^t + \frac{49\delta\lambda^3 - 77\delta\lambda^2 + 49\delta\lambda - 12\delta - 18\lambda + 9}{(1-2\lambda)\delta(7\lambda-4)}$$

[2] Sekiguchi assumes that $L > G^2$ which is of course violated in this example.

The limit of the coefficient of V^t as λ goes to zero is $\frac{1}{4\delta}$. The limit of the steady state value of V^t as λ goes to zero is $3\frac{4\delta-3}{4\delta-1}$. For small λ and sufficiently high δ, m^t declines towards $\frac{4\delta-3}{4\delta-1} \approx 0.33$. Suppose that player 2 adopts the strategy to play C with probability equal to m^t when observing t consecutive plays of D by player 1 and to return to C with probability equal to one whenever player 1 plays C in the last period (note that this strategy is independent of player 2's own play). By construction, player 1 is always indifferent between C and D after any possible history. It follows that playing the same strategy as player 2 is a best reply for player 1. Thus, for small λ's and sufficiently high δ's, there is sequence of Nash equilibria with payoffs close to 1. Furthermore, these Nash equilibria are observationally equivalent to Sequential equilibria since all unreached information sets of a player can be reached by and only by her own deviations (Kandori and Matsushima (1998), Lemma 2).

4 Theorems

In this section, I will prove the main result. The proof is an adaptation of the example in the previous section to the general case. The crucial step is the construction of a value function for a player and a mixed strategy for the opponent that make the player, at some histories, indifferent between playing C or D. It should be noted that a player's value function will map the histories observed by the *opponent* to continuation values. However, as illustrated in the above example, this dependence will be rendered harmless by making a player's best reply independent of such histories.

The main argument relies on equilibrium strategies that are similar to *tit-for-tat* strategies. Specifically, each player repeats an n-tuple of stage-game actions until a deviation of the opponent is observed. Such a deviation triggers a punishment phase which is relinquished by switching back to the initial n-tuple whenever the initial n-tuple of the opponent is observed again.

Given two n-tuples of stage game actions, a and b, define $\pi(a,b)$ to be player 1's discounted ex-ante payoff of n periods when player 1 chooses a and player 2 chooses b. Consider a pair of n-tuples (a^0, b^0) consisting of (C,C)'s in the first R periods, followed by (C,D) in the next S periods, by (D,C) in the next T periods, and by (D, D) in the remaining periods. Thus,

$$\pi(a^0, b^0) = \sum_{t=1}^{R} \delta^{t-1} - \delta^R \sum_{t=1}^{S} \delta^{t-1} l(\bar{\varepsilon}_1) + \delta^{R+S} \sum_{t=1}^{T} \delta^{t-1}(1 + g(\bar{\varepsilon}_1)).$$

For reasons that will become clear later, I will assume that (a^0, b^0) satisfies the following restriction:

$$R - SL + T(1 + G) > R(G - L). \qquad (A)$$

Let Θ^z denote the z-dimensional null vector. The next objective is to show that if $\lim_{\delta \to 1, \bar{\varepsilon}_1 \to \Theta^4} \pi(a^0, b^0)$ is positive and $\bar{\varepsilon}_1$ and $\bar{\varepsilon}_2$ are sufficiently close to Θ^4, then there exists a strategy of player 2 such that player 1 is indifferent between playing C or D at all histories of length $tn + z$, $t = 1, 2, ..., \infty$, $0 \leq z < R + S$ and strictly prefers D at all other histories. To this end, I will initially proceed as if the stage-game consists of playing n-tuples of actions.

Consider all K possible n-tuples of actions a^k obtained from a^0 by replacing C's with D's (not vice-versa) and let $a^K = (D, D, ..., D)$. An n-tuple of stage game actions a is said to be a **bad defection** of an n-tuple a^k, $k = 0, 1, ..., K$, if it is the obtained from a^k by replacing D's with C's in the last $n - R - S$ periods, that is, in the periods in which a^0 specifies D.

Let Q^n be the set of r-tuples of stage game actions such that r is a multiple of n and given q, q' in Q^n define qq' to be their concatenation. Define b^0/m to be the n-tuple of *mixed* stage game actions obtained from b^0 by replacing each C with the mixture "C with probability m". Let $p^k(m_q)$ be a $(K + 1)$ dimensional vector of probabilities, $k = 0, 1, ..., K$. Each component $p^k_\ell(m_q)$ is to be interpreted as the probability that after n periods player 2 observes a^ℓ or any bad defection of a^ℓ, when player 1 plays a^k and player 2 uses the mixed strategy b^0/m_q. Note that $p^k_\ell(m_q)$ is also the probability that after n periods player 2 observes a^ℓ or any bad defection of a^ℓ, when player 1 plays a bad defection of a^k. Consider the value function $V : Q^n \to R$ defined recursively as follows from the system of equations below.

$$
\begin{aligned}
V(q) &= (1-\delta^n)\pi(a^0, b^0/m_q) + \delta^n \sum_{\ell=0}^{K} p_\ell^0(m_q) V(qa^\ell) \\
V(q) &= (1-\delta^n)\pi(a^1, b^0/m_q) + \delta^n \sum_{\ell=0}^{K} p_\ell^1(m_q) V(qa^\ell) \\
V(q) &= (1-\delta^n)\pi(a^2, b^0/m_q) + \delta^n \sum_{\ell=0}^{K} p_\ell^2(m_q) V(qa^\ell) \qquad (1) \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
V(q) &= (1-\delta^n)\pi(a^K, b^0/m_q) + \delta^n \sum_{\ell=0}^{K} p_\ell^K(m_q) V(qa^\ell)
\end{aligned}
$$

Let $\emptyset$ denote the empty r-tuple. At $q = \emptyset$, set $m_q = 1$, $V(a^0) = V(\emptyset)$, and solve for $V(\emptyset)$, and $V(a^k)$, $k = 1, 2, ..., K$. Set $V(a) = V(a^k)$ for any bad defection a of a^k, $k = 1, 2, ..., K$, and $V(a) = V(\emptyset)$, for any bad defection a of a^0. In step t, given $V(\emptyset)$ and each $V(q)$ obtained in step $t-1$, set $V(qa^0) = V(\emptyset)$ and solve for m_q and $V(qa^k)$, $k = 1, 2, ..., K$. Then set $V(qa) = V(qa^k)$ for any bad defection a of a^k, $k = 1, 2, ..., K$, and $V(qa) = V(\emptyset)$, for any bad defection a of a^0.

Remark: The value function is such that player 1 is indifferent among all n-tuples a^k, $k = 0, 1, , ..., K$. Moreover, since the value function following a bad deviation of a^k is identical to the value function following a^k, player 1 strictly prefers a^k to any of its bad deviations.

By plugging the 2^n values for $V(\cdot)$ obtained in the first iteration, one obtains 2^{2n} values for $V(\cdot)$ and 2^n for m_q in the second iteration. At the t iteration, one obtains 2^{nt} values for $V(\cdot)$ and $2^{n(t-1)}$ values for m_q.[3] If at every iteration (1) admits a solution such that each m_q is between zero and one, (1) is said to be **solvable.** Given a point x in $\Re^z$, let $\mathcal{N}^z(x, \omega)$ denote the intersection between the open ball centered in x with radius ω and $\Re^z_{++}$.

Theorem 1: Suppose that $\lim_{\delta \to 1, \bar{\varepsilon}_1 \to \Theta^4} \pi(a^0, b^0) > 0$. Then, there exist $\hat{\omega}$ and $\hat{\delta}$, $\hat{\delta} < 1$, such that, for $\bar{\varepsilon}_1, \bar{\varepsilon}_2 \in \mathcal{N}^4(\Theta^4, \hat{\omega})$ and $\hat{\delta} < \delta < 1$, (1) is solvable.

Proof: First, write the system of equations (1) as

[3] Of course, only 2^{Kt} values for $V(\cdot)$ and $2^{K(t-1)}$ values for m_q are (possibly) distinct.

$$\begin{bmatrix} V(q) \\ V(q) \\ V(q) \\ \ldots\ldots \\ V(q) \end{bmatrix} = (1-\delta^n)\mathcal{P}(m_q)^{-1} \begin{bmatrix} \pi(a^0, b^0/m_q) \\ \pi(a^1, b^0/m_q) \\ \pi(a^2, b^0/m_q) \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \pi(a^K, b^0/m_q) \end{bmatrix} + \delta^n \begin{bmatrix} V(\emptyset) \\ V(qa^1) \\ V(qa^2) \\ \ldots\ldots \\ V(qa^K) \end{bmatrix} \tag{2}$$

where $\mathcal{P}(m_q)$ is the $(K+1)$ dimensional square matrix whose k^{th} row is $p^k(m_q)$. Define the element of $\mathcal{P}(m_q)^{-1}$ in the k^{th} row and ℓ^{th} column by $r^k_\ell(m_q)$ and $\hat{\pi}^k(m_q) = \sum_{\ell=0}^{K} \pi(a^\ell, b^0/m_q) r^k_\ell(m_q)$. It should be noted that $\mathcal{P}(m_q)$ converges to the Identity matrix when $\bar{\varepsilon}_2$ approaches Θ^4 and hence has an inverse in a neighborhood of Θ^4. In what follows, it will be assumed that such inverse exists. Also, $\hat{\pi}^k(m_q)$ converges to $\pi(a^k, b^0/m_q)$ when $\bar{\varepsilon}_1$ and $\bar{\varepsilon}_2$ approach Θ^4.

After making the appropriate substitutions, one obtains from (2) the equation

$$\delta^n \hat{\pi}^0(m_{qa^k}) = \hat{\pi}^0(m_q) - \hat{\pi}^k(m_q) + \delta^n \hat{\pi}^0(1) \tag{3}$$

which describes the dynamics of m_q. If both R and S are equal to zero, then (3) is trivially solved by $m_q = 1$ for any q since all deviations are bad deviations. Hence assume that either R or S is positive and rewrite $\hat{\pi}^k(m_q)$ as $\psi^k(m_q, \delta, \bar{\varepsilon}_1, \bar{\varepsilon}_2)$ so that all the relevant arguments are explicit. Also, restrict the domain of ψ^k to $[0,1] \times [0,2] \times \mathcal{B} \times \mathcal{B}$, where $\mathcal{B}$ is an open neighborhood of Θ^4 such that ψ^k is continuous in all its arguments. The existence of this neighborhood is easy to verify.

We first analyze existence of a solution for (3) when $m_q = 1$ and $m_q = 0$. Note that

$$\psi^k(m_q, 1, \Theta^4, \Theta^4) = m_q(R - \hat{R}^k) + m_q(1+G)\hat{R}^k - (1-m_q)(R - \hat{R}^k)L + m_q T(1+G) - (S - \hat{S}^k)L$$

where $\hat{R}^k$ and $\hat{S}^k$ are the number of deviations from a^0 to a^k in the first R periods and the subsequent S periods respectively. For the case $m_q = 1$, it is easily shown that

$$\psi^0(1,1,\Theta^4,\Theta^4) > \psi^0(1,1,\Theta^4,\Theta^4) - \psi^k(1,1,\Theta^4,\Theta^4) + \psi^0(1,1,\Theta^4,\Theta^4)$$

and

$$\psi^0(0,1,\Theta^4,\Theta^4) < \psi^0(1,1,\Theta^4,\Theta^4) - \psi^k(1,1,\Theta^4,\Theta^4) + \psi^0(1,1,\Theta^4,\Theta^4)$$

The first inequality is straightforward. To see that the second inequality also holds, one can verify that the expression

$$\psi^0(0,1,\Theta^4,\Theta^4) - \big(\psi^0(1,1,\Theta^4,\Theta^4) - \psi^k(1,1,\Theta^4,\Theta^4) + \psi^0(1,1,\Theta^4,\Theta^4)\big)$$

simplifies to $-LR - R - T - TG + R^k G + L\hat{S}^k$, which is bounded by $-LR - R - T - TG + RG + SL$ and is negative by restriction (A). By continuity, there exist μ_1^*, δ_1^*, and ω_1^* such that

$$\delta^n \hat{\pi}^0(1) > \hat{\pi}^0(m_q) - \hat{\pi}^k(m_q) + \delta^n \hat{\pi}^0(1)$$

and

$$\delta^n \hat{\pi}^0(0) < \hat{\pi}^0(m_q) - \hat{\pi}^k(m_q) + \delta^n \hat{\pi}^0(1)$$

for $m_q \in (\mu_1^*, 1]$, $\delta \in (\delta_1^*, 1]$, $\bar{\varepsilon}_1, \bar{\varepsilon}_2 \in \mathcal{N}^4(\Theta^4, \omega_1^*)$. Hence, for $m_q \in (\mu_1^*, 1]$, $\delta \in (\delta_1^*, 1]$, $\bar{\varepsilon}_1, \bar{\varepsilon}_2 \in \mathcal{N}^4(\Theta^4, \omega_1^*)$, (3) admits a solution m_{qa^k} in $(0,1)$. Repeating this argument for the case $m_q = 0$, one finds that

$$\psi^0(1,1,\Theta^4,\Theta^4) > \psi^0(0,1,\Theta^4,\Theta^4) - \psi^k(0,1,\Theta^4,\Theta^4) + \psi^0(1,1,\Theta^4,\Theta^4)$$

is again straightforward and that

$$\psi^0(0,1,\Theta^4,\Theta^4) < \psi^0(0,1,\Theta^4,\Theta^4) - \psi^k(0,1,\Theta^4,\Theta^4) + \psi^0(1,1,\Theta^4,\Theta^4)$$

simplifies to $-LR + LR^k + L\hat{S}^k - R - T - TG < 0$. The left-hand side is bounded by $LS - R - T - TG$ which is negative by hypothesis. Hence, there exist μ_0^*, δ_0^*, and ω_0^* such that, for $m_q \in [0, \mu_0^*)$, $\delta \in (\delta_0^*, 1]$, $\bar{\varepsilon}_1, \bar{\varepsilon}_2 \in \mathcal{N}^4(\Theta^4, \omega_0^*)$, (3) also admits a solution m_{qa^k} in $(0,1)$.

Since $\psi^k(m_q, 1, \Theta^4, \Theta^4)$ is linear in m_q,

$$\psi^0(1,1,\Theta^4,\Theta^4) > \psi^0(m,1,\Theta^4,\Theta^4) - \psi^k(m,1,\Theta^4,\Theta^4) + \psi^0(1,1,\Theta^4,\Theta^4)$$

and

$$\psi^0(0,1,\Theta^4,\Theta^4) < \psi^0(m,1,\Theta^4,\Theta^4) - \psi^k(m,1,\Theta^4,\Theta^4) + \psi^0(1,1,\Theta^4,\Theta^4)$$

hold for an arbitrary $m \in (0,1)$. Hence, for any $m \in (0,1)$, there exist an open interval $\mathcal{I}(m)$ containing m, $\delta^*(m)$, and $\omega^*(m)$ such that, for $m_q \in \mathcal{I}(m)$, $\delta \in (\delta^*(m), 1]$, $\bar{\varepsilon}_1, \bar{\varepsilon}_2 \in \mathcal{N}^4(\Theta^4, \omega^*(m))$, (3) admits a solution m_{qa^k} in $(0,1)$. Since the union of $[0, m_0^*)$, $(m_1^*, 1]$, and $\mathcal{I}(m)$, $m \in (0,1)$, is an open cover of $[0,1]$, it has a finite subcover. Hence, by standard arguments, there exist $\hat{\delta}$ and $\hat{\omega}$ such that, for $m_q \in [0,1]$, $\delta \in (\hat{\delta}, 1]$, $\bar{\varepsilon}_1, \bar{\varepsilon}_2 \in \mathcal{N}^4(\Theta^4, \hat{\omega})$, (3) admits a solution m_{qa^k} in $(0,1)$. QED

The result stated below is an application of Theorem 1. Consider the "imaginary" modification of Γ^∞ such that both played histories, (h_1^{Pl}, h_2^{Pl}), and observed histories, (h_1^{Ob}, h_2^{Ob}), are common knowledge. Denote this game by $\tilde{\Gamma}^{\infty}$. A strategy for this game maps histories $(h_1^{Pl}, h_2^{Pl}, h_1^{Ob}, h_2^{Ob})$ into the probability of playing C. Define the strategy $\tilde{f}_2^*$ for player 2 in $\tilde{\Gamma}^{\infty}$ such that

$\tilde{f}_2^*(h_1^{Pl}, h_2^{Pl}, h_1^{Ob}, h_2^{Ob}) = m_q$

whenever there exists $q \in Q^n$ such that $h_1^{Ob} = qz$ for an r-tuple z where $0 \leq r < R$ and $R + S \leq r < R + S + T$, and

$\tilde{f}_2^*(h_1^{Pl}, h_2^{Pl}, h_1^{Ob}, h_2^{Ob}) = 0$

otherwise.

Theorem 2: Consider the game $\tilde{\Gamma}^{\infty}$ and suppose $\lim_{\delta \to 1, \bar{\varepsilon}_1 \to \Theta^4} \pi(a^0, b^0) > 0$. Then, there exist $\hat{\omega}$ and $\hat{\delta}$, $\hat{\delta} < 1$, such that, for $\bar{\varepsilon}_1, \bar{\varepsilon}_2 \in \mathcal{N}^4(\Theta^4, \hat{\omega})$ and $\hat{\delta} < \delta < 1$, if player 2 plays $\tilde{f}_2^*$, player 1 is indifferent between playing C or D at all histories of length $tn + z$, $t = 1, 2, ..., \infty$, $0 \leq z < R + S$ and strictly prefers D at all other histories.

Proof: See Appendix

Now define the **modified individually rational convex hull**, $\mathcal{M}$, as the individually rational convex hull satisfying an additional restriction:

$$\begin{aligned}\mathcal{M} = \{(v_1, v_2) \in \Re_+ \mid (v_1, v_2) = \alpha(1,1) + \beta(1 + G, -L) + \gamma(-L, 1 + G),\\ \alpha, \beta, \gamma \geq 0,\ \alpha + \beta + \gamma \leq 1, \text{ and } \alpha + \beta(1 + G) - \gamma L > \alpha(G - L),\\ \text{and } \alpha + \gamma(1 + G) - \beta L > \alpha(G - L)\}\end{aligned}$$

Note that $\mathcal{M}$ differs (slightly) from the individually rational convex hull when $0 < G - L < 1$ and is consistent with the n-periods payoffs obtainable under restriction (A). The reader can verify that the area excluded in the upper left corner of the convex hull is bounded (strictly) from below by the line segment joining the point $(0, 1 + G - L)$ and $(1, 1)$. Hence, the square with vertices in $(0, 0)$ and $(1, 1)$ is always in $\mathcal{M}$.

The exclusion of a portion of the individually rational convex hull is caused by the use of strategies that return to *cooperation* (say, b^0) when *cooperation (say,* a^0) is observed in the preceding n periods. Consider the limit case of no errors in observations. If player 1 defects for n periods and cooperates for the following n periods she gets, for δ large and under the assumption that player 2 punishes player 1 by playing D for n periods, $((R + T)(1 + G) + (R + S)(-L))$. This needs to be smaller than $2(R - SL + T(1 + G))$, yielding condition (A).

Theorem 2 leads the main Theorem:

Theorem 3: Take any pair $(v_1, v_2) \in \mathcal{M}$. Then, there exist $\hat{\omega}$ and $\hat{\delta}$, $\hat{\delta} < 1$, such that, for $\bar{\varepsilon}_1, \bar{\varepsilon}_2 \in \mathcal{N}^4(\Theta^4, \hat{\omega})$ and $\hat{\delta} < \delta < 1$, the game Γ^∞ has a Sequential equilibrium with payoffs arbitrarily close to (v_1, v_2).

Proof: Find the n-tuples (a^0, b^0) whose n-periods payoffs approximate (v_1, v_2) as closely as required and consider first the game $\tilde{\Gamma}^\infty$ and the strategy $\tilde{f}_2^*$. Then, by construction, player 1 is indifferent between playing C or D at all histories of length $tn + z$, $t = 1, 2, ..., \infty$, $0 \leq z < R + S$ and strictly prefers D at all other histories. Hence, having constructed the corresponding strategy $\tilde{f}_1^*$ for player 1, $(\tilde{f}_1^*, \tilde{f}_2^*)$ is a Nash equilibrium for this game. Now consider the original game Γ^∞. Since $\tilde{f}_i^*$ depends only on the realization of j's actions, one can construct an identical strategy f_i^* for this game. Furthermore, the strategy sets of Γ^∞ are subsets of the strategy sets of $\tilde{\Gamma}^\infty$. Hence, (f_1^*, f_2^*) is a Nash equilibrium. By Kandori and Matsushima (1998), Lemma 2, this Nash equilibrium is observationally equivalent to a Sequential equilibrium since all unreached information sets of a player can be reached by and only by her own deviations. QED

Remark: One referee has pointed out that the equilibria in Theorem 3 are not robust to payoff perturbations. The basic feature of these equilibria is that a player is indifferent between C and D independently of whether a deviation has been detected in the previous period. Following a small shock to the payoffs, a player would then choose the same (pure) action regardless of the previous observations of the opponent's action.

5 Larger Errors

The above results are derived under the assumption that errors in observations vanish in the limit. However, some equilibria can be computed also for the case when errors are not infinitesimal. Consider the example of section 3, that is $(a^0, b^0) = (C, C)$, for arbitrary G, L and λ, and suppose, for simplicity, that $L > G$ and $\lambda < 1/2$. To determine the equilibrium payoff, solve the following equations:

$$V^0 = (1 - \delta)\,(1 - \lambda - L\lambda) + \delta((1 - \lambda)V^0 + \lambda V^1$$

$$V^0 = (1 - \delta)(1 + G)(1 - \lambda) + \delta((1 - \lambda)V^1 + \lambda V^0)$$

The solution is

$V^0 = (1-\lambda)\,\dfrac{1-2\lambda-L\lambda-G\lambda}{1-2\lambda}$ and

$$V^1 = \frac{2\delta\lambda^2 - 3\delta\lambda + \delta + \delta G + \delta G\lambda^2 + \delta L\lambda^2 - 2\delta G\lambda - G + G\lambda - L\lambda}{\delta\,(1-2\lambda)}$$

The next step is to solve the system

$$V^t = (1-\delta)((1-\lambda-L\lambda)\,m^t + (1-m^t)(-L(1-\lambda)+\lambda)) + \delta((1-\lambda)V^0 + \lambda V^{t+1})$$

$$V^t = (1-\delta)\,((1+G)\,(1-\lambda)m^t + (1-m^t)\,(1+G)\,\lambda) + \delta((1-\lambda)V^{t+1} + \lambda V^0)$$

The system can be written as

$$m^t = \frac{1}{(1+L-2\lambda-L\lambda-G\lambda)\,(1-\delta)} V^t + const$$

$$V^{t+1} = \frac{L-G}{(1+L-2\lambda-L\lambda-G\lambda)\,\delta} V^t + const$$

By the parameter restrictions, m^t is increasing in V^t and, if V^0 is positive, the coefficient of V^t is the second difference equation is between zero and one. Hence, m^t declines to its steady state value

$m^\infty = \dfrac{2\lambda^2\delta\,(L+2+G) - \lambda\,(3\delta L + \delta G + 4\delta + G - L) - L + \delta + \delta L}{(-1+2\lambda)\,(\delta L\lambda + \delta G\lambda + 2\delta\lambda - \delta + L - G - \delta L)}$. Furthermore,

$$\lim_{\delta\to1} V^\infty = (1-\lambda)\,\frac{1-L\lambda-G\lambda-2\lambda}{1-2\lambda}$$

$$\lim_{\delta\to1} m^\infty = \frac{2L\lambda^2 + 4\lambda^2 - 2G\lambda + 2G\lambda^2 - 2L\lambda - 4\lambda + 1}{(1-2\lambda)\,1+G-L\lambda-G\lambda-2\lambda)}$$

Hence, when $m^\infty \geq 0$, there will exist equilibria with payoffs V^0. Note that the expected payoff pair $(1,1)$ is not obtained when errors are not small even if δ is very large.

6 Conclusion

This paper establishes a partial (limit) Folk Theorem for a repeated Prisoner's Dilemma when the actions are imperfectly observable. The strategies employed in the proof of this result resemble *tit-for-tat* strategies in that they revert to cooperative behavior whenever cooperative behavior is observed in the preceding unit of time. While these strategies make the proof fairly simple and go a considerable way towards establishing a standard Folk Theorem, they fail to go the entire distance. Stronger results can be obtained by using less "forgiving" strategies which return to cooperation after longer "redeeming" phases.

More importantly, the approach introduced in this paper can be extended to a larger class of games with private monitoring. I hope to explore this issue in further research.

References

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

[2] Abreu D, Pearce D, Stacchetti E, Toward a Theory of Discounted Repeated Gamed with Imperfect Monitoring, *Econometrica*, 1990, 58, 1041-1063.

[3] Ben-Porath E, Kahneman M, Communication in Repeated Games with Private Monitoring, *Journal of Economic Theory*, 1996, 70, 281-297.

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

[5] Fudenberg D, Levine D, Maskin E, The Folk Theorem with Imperfect Public Information, *Econometrica*, 1994, 62, 997-1040.

[6] Green E, Porter R, Noncooperative Collusion under Imperfect Price Formation, *Econometrica*, 1984, 52, 87-100.

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

[8] Mailath G, Morris S, Repeated Games with Imperfect Private Monitoring: Notes on a Coordination Perspective, CARESS Working Paper 98-07, 1998.

[9] Sekiguchi T, Efficiency in Repeated Prisoner's Dilemma with Private Monitoring, *Journal of Economic Theory*, 1997, 76, 345-361.

7 Appendix

Proof of Theorem 2: Take x and y such that $x + y = n$ and consider the set of x-tuples $A(x) = \{a^i(x)\}_{i=1,2,\dots,N(x)}$ and of y-tuples $A(y) = \{a^j(y)\}_{j=1,2,\dots,N(y)}$ such that each a^k can be written as a the concatenation $a^{i(k)}(x)a^{j(k)}(y)$, where $i(k)$ and $j(k)$ select the appropriate x tuple and y-tuple in $A(x)$ and $A(y)$. As before, an x-tuple (y-tuple) is a bad deviation of $a^i(x)$ ($a^j(y)$) if it is obtained from $a^i(x)$ ($a^j(y)$) by replacing D with C in periods where $a^{i(0)}(x)$ ($a^{j(0)}(y)$) plays D. Since errors in observation are independent over time, we can write $p^k_\ell(m_q) = p^{i(k)}_{i(\ell)}(x)p^{j(k)}_{j(\ell)}(y)$,[4] where $p^{i(k)}_{i(\ell)}(x)$ is the probability of observing the x-tuple $a^{i(\ell)}(x)$ or its bad deviations when $a^{i(k)}(x)$ is played and $p^{j(k)}_{j(\ell)}(y)$ is defined analogously. Now rewrite $\pi(a^k, b^0/m_q)$ as

$$\pi(a^k, b^0/m_q) = \pi^x(a^k(x), b^0/m_q) + \delta^x \pi^y(a^k(y), b^0/m_q)$$

where $\pi^x(a^k(x), b^0/m_q)$ is the first x periods payoff and $\pi^y(a^k(y), b^0/m_q)$ is the payoff in the last y periods with discounting beginning in period $x+1$. Define $\mathcal{V}(q) = (1-\delta^n)^{-1}V(q)$. Equation k in (1) can then be written as

$$\mathcal{V}(q) = \pi^x(a^{i(k)}(x), b^0/m_q) + \delta^x \pi^y(a^{j(k)}(y), b^0/m_q) \\ +\delta^n \sum_{h=0}^{N(x)} p^{i(k)}_h(x) \sum_{g=0}^{N(y)} p^{j(k)}_g(y)\mathcal{V}(qa^h(x)a^g(y)).$$

Define

$$E(a^i(x), a^j(y)) = \pi^y(a^j(y), b^0/m_q) + \delta^y \sum_{g=0}^{N(y)} p^j_g(y)\mathcal{V}(qa^i(x)a^g(y))].$$

$\mathcal{V}(q)$ can be rewritten as

$$\mathcal{V}(q) = \pi^x(a^{i(k)}(x), b^0/m_q) + \delta^x \sum_{h=0}^{N(x)} p^{i(k)}_h(x)E(a^h(x), a^j(y)).$$

Fix an $a^j(y)$ and consider the sub-system of (1) with $N(x)$ equations

$$\begin{bmatrix} \mathcal{V}(q) \\ \mathcal{V}(q) \\ \dots\dots \\ \mathcal{V}(q) \end{bmatrix} = \begin{bmatrix} \pi^x(a^1(x), b^0/m_q) \\ \pi^x(a^2(x), b^0/m_q) \\ \dots\dots\dots\dots\dots\dots \\ \pi^x(a^{N(x)}(x), b^0/m_q) \end{bmatrix} + \delta^x \mathcal{P}_x \begin{bmatrix} E(a^1(x), a^j(y)) \\ E(a^2(x), a^j(y)) \\ \dots\dots\dots\dots\dots \\ E(a^{N(x)}(x), a^j(y)) \end{bmatrix}.$$

[4] m_q is omitted to simplify the notation.

where $\mathcal{P}_x$ is a probability matrix whose element in the i^{th} row and z^{th} column is the probability of observing $a^z(x)$ or its bad deviations when playing $a^i(x)$. Hence,

$$\begin{bmatrix} \mathcal{V}(q) \\ \mathcal{V}(q) \\ \ldots\ldots \\ \mathcal{V}(q) \end{bmatrix} - \mathcal{P}_x^{-1} \begin{bmatrix} \pi^x(a^1(x), b^0/m_q) \\ \pi^x(a^2(x), b^0/m_q) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \pi^x(a^{N(x)}(x), b^0/m_q) \end{bmatrix} = \delta^x \begin{bmatrix} E(a^1(x), a^j(y)) \\ E(a^2(x), a^j(y)) \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ E(a^{N(x)}(x), a^j(y)) \end{bmatrix}$$

Since the left-hand side is independent of $a^j(y)$, so is $E(a^i(x), a^j(y))$. Thus, the value function $\mathcal{V}$ can be extended to the set of all r-tuples of actions by setting

$$\mathcal{V}(qa) = E(a^i(x), a^j(y))$$

for $a = a^i(x)$ and any bad deviation a of $a^i(x)$. Now extend the definition of ε_i to

$$\varepsilon_i(m, s_j) = \varepsilon_i(C, s_j)m + \varepsilon_i(D, s_j)\,(1-m)$$

Hence, for any r-tuple a, $r < R$,

$$\mathcal{V}(qa) = m_q - (1-m_q)l\,(\varepsilon_1) + \delta((1-\varepsilon_2(m_q, C))\mathcal{V}(qaC) + \varepsilon_2(m_q, C)\mathcal{V}(qaD))$$

$$\mathcal{V}(qa) = m_q(1 + g\,(\varepsilon_1)) + \delta((1-\varepsilon_2(m_q, D))\mathcal{V}(qaD) + \varepsilon_2(m_q, D)\mathcal{V}(qaC))$$

and, for any r-tuple a, $R \leq r < R+S$

$$\mathcal{V}(qa) = -l\,(\varepsilon_1) + \delta((1-\varepsilon_2(m_q, C))\mathcal{V}(qaC) + \varepsilon_2(m_q, C)\mathcal{V}(qaD))$$

$$\mathcal{V}(qa) = \delta((1-\varepsilon_2(m_q, D))\mathcal{V}(qaD) + \varepsilon_2(m_q, D)\mathcal{V}(qaC)).$$

Consider the game $\tilde{\Gamma}$ and the value function

$$\tilde{\mathcal{V}}(h_1^{Pl}, h_2^{Pl}, h_1^{Ob}, h_2^{Ob}) = \mathcal{V}(h_1^{Ob})$$

for any history $(h_1^{Pl}, h_2^{Pl}, h_1^{Ob}, h_2^{Ob})$. Then, player 1 is indifferent between playing C or D at all histories of length $tn+z$, $t = 1, 2, ..., \infty$, $0 \leq z < R+S$. At any other h_1^{Ob}, $\mathcal{V}(h_1^{Ob}C) = \mathcal{V}(h_1^{Ob}D)$. Hence, player 1 strictly prefers D. QED