CARESS Working Paper #99-13

The Repeated Prisoner's Dilemma with Private Monitoring: a N-player case*

Ichiro Obara
Department of Economics
University of Pennsylvania
obara@ssc.upenn.edu

September 99

Abstract

This paper studies the repeated prisoner's dilemma with private monitoring for arbitrary number of players. It is shown that a mixture of a grim trigger strategy and permanent defection can achieve an almost efficient outcome for some range of discount factors if private monitoring is almost perfect and the number of players is large. This result also holds when the number of players is two for any prisoner's dilemma as long as monitoring is almost perfect and symmetric. A detailed characterization of this sequential equilibrium is provided.

*I am grateful to my advisers George Mailath and Andrew Postlewaite for their encouragement and valuable suggestions, and Johannes Horner, Stephen Morris, Nicolas Persico, Tadashi Sekiguchi for their helpful comments. All remaining errors are mine.

1. Introduction

This paper examines the repeated prisoner's dilemma for arbitrary number of players, where players only observe private and imperfect signals about the other players' actions. This game belongs to the class of repeated games with private monitoring. While repeated games with public monitoring have been extensively analyzed in, for example, Abreu, Pearce and Stachetti [1] or Fudenberg, Levin, and Maskin [7], few things are known about repeated games with private monitoring. It is shown in Compte [4] and Kandori and Matsushima [8] that a Folk Theorem still holds in this class of game with communication between players, but it is di¢cult to analyze it without communication because the simple recursive structure is lost.

The two player prisoner's dilemma was already examined in Sekiguchi [13], which is the …rst paper to show that the e¢cient outcome can be achieved in some repeated prisoner's dilemmas with almost perfect private monitoring. This paper is an extension of Sekiguchi [13] in the sense that (1): a similar grim trigger strategy is employed, (2): the e¢cient outcome is obtained for any prisoner's dilemma with two players, (3): this e¢ciency result for the two player case is extended to the case of arbitrary number of players with some additional assumptions, and (4): the sequential equilibrium corresponding to this grim trigger Nash equilibrium is explicitly constructed.

In Sekiguchi [13], the critical step of the arguments is to obtain the unique optimal action with respect to a player's subjective belief about the other player's continuation strategy. Since players randomize between the grim trigger strategy and the permanent defection in the …rst period, a player's continuation strategy is always one of these two strategies after any history. This means that a player's subjective belief about the other player's strategy can be summarized in one parameter: a subjective probability of the permanent defection being played by the other player. In Sekiguchi [13], it is shown that a player should start defecting if she is very con…dent that the other player has started defecting, and a player should cooperate if she is really con…dent that the other player is still cooperating. However, it is not clear what a player should do if the belief is somewhere in the middle. In this paper, the clear cut characterization of the optimal action is provided, which makes it possible to extend the e¢ciency result to any prisoner's dilemma.

Although the same kind of clear characterization of the optimal action is possible with many players, it is not straightforward to extend this e¢ciency result to the n player case. The dynamics of belief is richer with more than two players. In particular, it is possible to have a belief that some player started defecting but other players are still cooperating. In such a case, a player might think that it is better to continue cooperating because it might keep cooperative players from starting defection. So, it is no longer clear when players should pull the trigger.

Under the assumption that the probability of any signal pro…le depends on the number of the total errors it contains, it is shown that the e¢ciency outcome can be supported with the mixture of the permanent defection and a certain kind of grim trigger strategy, where players start defecting if they observe even one signal of deviation by any other player. This strategy generates an extreme belief dynamics under the assumption on the signal distribution, which in turn rationalizes the use of this strategy. As soon as a player observes any bad signal from any other player, the player expects that some other players also got some bad signals with high probability. Then, she becomes pessimistic enough to start defecting for herself because defection should prevail among all players using the same strategy at least in the next period.

A sequence of papers have re…ned the result of Sekiguchi [13] for the two player case. Piccione [12] also achieves the e¢cient outcome for any prisoner's dilemma with two players and almost perfect private monitoring. Moreover, he establishes an almost Folk Theorem using a strategy which allows players to randomize between cooperation and defection after every history. The strategy used in his paper can be represented as an automaton with countably in…nite states. Ely and Välimäki [6] prove a Folk Theorem using a similar strategy, but their strategy is "simple" in the sense that it is a two states automaton. Bhaskar [2] is closest to this paper in terms of results and strategies employed in the two player case. He essentially shows (2) and (4), and also proves a Folk Theorem for a class of prisoner's dilemma through a di¤erent line of attack from Piccione [12] or Ely and Välimäki [6].

Mailath and Morris [9] is the …rst paper to deal with the n player case in the private monitoring framework. They show that a subgame perfect equilibrium with public monitoring is robust to the introduction of private monitoring if players' continuation strategies are approximately common knowledge after every history and information is almost public. A Folk theorem can be obtained when information is almost public and almost perfect. Although the stage game in this paper has a more speci…c structure, the information structure allowed in this paper is not nested in their information structure. Especially, private signals can be independent over players.

This paper is organized as follows. In section 2, the model is described. In Section 3, the assumptions on the information structure are presented. Section 4 discusses the optimal action with respect to player's beliefs and the belief dynamics generated by the equilibrium strategy proposed in this paper. A sequential equilibrium is constructed in Section 5. Section 6 gives a detailed characterization of the sequential equilibrium constructed in Section 5. Section 7 concludes.

2. The Model

Let $N = \{1, 2, ..., n\}$ be the set of players and g be the stage game played by those players. The stage game g is as follows. Player i chooses an action a_i from the action set $A_i = \{C, D\}$. Actions are not observable to the other players and taken simultaneously. A n-tuple action profile is denoted by $a \in A = \prod_{i=1}^{N} A_i$. A profile of all player's actions but player i's is $a_{-i} \in \prod_{i \neq j} A_i$.

Each player receives a private signal profile $\omega_i \in \{C, D\}^{n-1} = \Omega_i$; which is a n-tuple of signals about all the other players' actions within that period. Let $\omega_i = (\omega_{i,1}, ..., \omega_{i,i-1}, \omega_{i,i+1}, ..., \omega_{i,n})$ be a generic signal received by player i; where $\omega_{i,j}$ stands for the signal player i receives about the action taken by player j. A generic signal profile is denoted by $\omega = (\omega_1, ..., \omega_N) \in \Omega$. All players have the same payoff function u. Player i's payoff $u(a_i, \omega_i)$ depends on her own action a_i and private signal ω_i. Other players' actions affect a player i's payoff only through the distribution over the signal which player i receives. The distribution conditional on a is denoted by $p(\omega|a)$. It is assumed that $p(\omega|a)$ are full support, that is, $p(\omega|a) > 0 \ \forall a \forall \omega$. The space of a system of full support distributions $\{p(\omega|a)\}_{a \in A}$ is denoted by P.

I now introduce the perfectly informative signal distribution $P_0 = \{p_0(\omega|a)\}_{a \in A}$, where, for any $a \in A$, $p_0(\omega|a) = 1$ if $\omega_i = a_{-i}$ for all i. The whole space of the information structure $P \cup P_0$ is endowed with the Euclidean norm.

Since I am interested in the situation where information is almost perfect, I restrict my attention to a subset of P where information is almost perfect. Information is almost perfect when every person's signal profile is equal to the actual action profile taken by the other players in that period with probability more than $1 - \varepsilon$ for some small number ε.

To sum up, the space of the information structure I am mainly concerned with is a subset of P:

$$P_\varepsilon = \left\{ \{p(\omega|a)\}_{a\in A} \in \mathbb{R}_{++}^{n\times(n-1)\times 2^n} \;\middle|\; \begin{array}{l} p(\omega|a) \geq 1-\varepsilon \text{ if } \omega_i = a_{-i} \text{ for all } i; \\ \text{and } \forall a,\ \sum_\omega p(\omega|a) = 1 \end{array} \right\}$$

and p_ε is a generic element of P_ε.

The stage game payoff only depends on the number of signals "C" and "D" a player receives. Let $d(\omega_i)$ be the number of "D" contained in ω_i. Then, $u(a_i;\omega_i') = u(a_i;\omega_i'')$ if $d(\omega_i') = d(\omega_i'')$ for any a_i. I denote by $u\left(a_i;D^k\right)$ the payoff of player i when $d(\omega_i) = k$. The deviation gain when k defections are observed is $\Delta(k) = u\left(D;D^k\right) - u\left(C;D^k\right)$.

The stage game expected payoff is $U_i(a:p) = \sum_\omega u(a_i;\omega_i)\,p(\omega|a)$. It is assumed that D is a dominant action, that is, $U_i((D;a_{-i}):p) > U_i((C;a_{-i}):p)$ for any a_{-i}. The payoffs $U_i((C;\dots;C):p)$ and $U_i((D;\dots;D):p)$ are normalized to 1 and 0 respectively for all i. It is assumed that $(1;\dots;1)$ is an efficient stage game payoff.

The stage game g is repeated infinitely many times by N players, who discount their payoffs with a common discount factor $\delta \in (0;1)$. Time is discrete and denoted by $t = 1;2;\dots$. Player i's private history is $h_i^t = \left((a_i^1;\omega_i^1);\dots;\left(a_i^{t-1};\omega_i^{t-1}\right)\right)$ for $t \geq 2$ and $h_i^1 = \emptyset$. Let H_i^t be the set of all such history h_i^t and $H_i = \bigcup_{t=1}^{\infty} H_i^t$. Player i's strategy is a sequence of mappings $\sigma_i = (\sigma_{i;1};\sigma_{i;2};\dots)$; each $\sigma_{i;t}$ being a mapping from H_i^t to probability measures on A_i.

Since the equilibrium constructed later is based on a grim trigger strategy and permanent defection, it is convenient to introduce some notations for this specific construction.

First of all, the grim trigger and the permanent defection, denoted by σ_C and σ_D respectively, are:

$$\sigma_C(h_i^t) = \begin{cases} C & \text{if } h_i^t = ((C;C);\dots;(C;C)) \text{ or } t=1 \\ D & \text{otherwise} \end{cases}$$

$$\sigma_D(h_i^t) = D \text{ for all } h_i^t \in H_i$$

I also use σ_C or σ_D for any continuation strategy which is identical to σ_C or σ_D after some period t, that is, any continuation strategy at period t such that $\sigma\left(h_i^{t+k}\right) = \sigma_{a_i}\left(h_i^k\right)$ for $k = 1;2;\dots$ and $a_i = C$ or D. Moreover, any continuation strategy which is realization equivalent to σ_C or σ_D is also denoted by σ_C or

σ_D respectively[1]. This grim trigger strategy is the harshest one among all the variations of grim trigger strategies in the n player case. Players using σ_C switch to σ_D as soon as they observe any signal profile which is not full cooperation. When player i is mixing σ_C and σ_D with probability $(1 - q_i, q_i)$, that strategy is denoted by $(1 - q_i)\sigma_C + q_i\sigma_D$.

Suppose that σ_C or σ_D is chosen in the first period by all players. Let $\theta \in \Theta$ be the number of players using σ_D as a continuation strategy among n players. Then a probability measure $q_{-i}(h_i^t; p)$ on the space $\Theta = \{0, 1, ..., n-1\}$ is derived conditional on the realization of the private history h_i^t. Clearly, this measure also depends on the initial level of mixture between σ_C and σ_D by every player, but this dependence does not appear explicitly as it is obvious. Player i's conditional subjective probability that at least one player is using σ_D is denoted by $\phi(q_{-i}(h_i^t; p)) = 1 - q_{-i}(h_i^t; p)(0)$. The probability of this event is important because the number of players who are playing permanent defection does not make much difference to what happens in the future given everyone's strategy. As soon as someone starts playing σ_D, every other player starts playing σ_D with very high probability from the very next period on by the assumption of almost perfect monitoring. What is important is not how many players have switched to σ_D, but whether anyone has switched to σ_D or not.

Discounted average payoff is $V_i(\sigma : p; \delta) = (1 - \delta)\sum_{t=1}^{\infty} \delta^{t-1} E[u((a_i^t, \omega_i^t)) | \sigma; p]$, where the probability measure on H_i^t is generated by $(\sigma; p)$. Let $V_i(\sigma_i; k : p; \delta)$ be player i's discounted average payoff when k other players are playing σ_D and $n - k - 1$ other players are playing σ_C. This notation is justified under the assumption of the symmetry distribution, which is introduced in the next section. I also use the following notations:

$$U_i(a_i; q_{-i} : p) = \sum_{\theta=0}^{n-1} U_i\left(\left(a_i; D^\theta\right) : p\right) q_{-i}(\theta)$$

$$V_i(\sigma_i; q_{-i} : p; \delta) = \sum_{\theta=0}^{n-1} V_i(\sigma_i; \theta : p; \delta)\, q_{-i}(\theta)$$

$$\text{and } \mathbf{M}(q_{-i}; p) = \sum_{\theta=0}^{n-1} \left\{ U_i\left(\left(D; D^\theta\right) : p\right) - U_i\left(\left(C; D^\theta\right) : p\right)\right\} q_{-i}(\theta).$$

[1] A strategy is realization equivalent to another strategy if the former generates the same outcome distribution as the latter independent of the other players' strategies.

The largest deviation gain and the smallest deviation gain are $\overline{\mathbf{M}}$ and $\underline{\mathbf{M}}$ respectively, where $\overline{\mathbf{M}} = \max_{1 \leq k \leq n-1} \mathbf{M}(k)$ and $\underline{\mathbf{M}} = \min_{1 \leq k \leq n-1} \mathbf{M}(k)$: The least upper bound and largest lower bound of the discounted average payoff are denoted by $\overline{V}$ and $\underline{V}$ respectively.

3. Information Structure

In this section, various assumptions on the information structure are proposed and discussed. In the following sections, a sequential equilibrium is constructed with a mixture of grim trigger strategy and permanent defection, which achieves an approximately efficient outcome for some range of discount factors. As is the case with any equilibrium based on simple grim trigger strategies, this equilibrium satisfies the following property; players stick to the grim trigger strategy as long as they have an optimistic belief about the others, and they switch to the permanent defection once they become pessimistic and never come back. This property is satisfied in games with perfect monitoring, but not so easily satisfied in games with imperfect monitoring. In order to achieve a certain level of coordination, which is necessary for an equilibrium with trigger strategies, I impose some assumptions on $p(\omega|a)$ in addition to the assumption that it is almost perfect.

The first assumption, which is maintained throughout this paper, is

Assumption 1

$$p(\omega|a) = p\left(\left(\omega_{\tau(i)\tau(j)}\right) \middle| \left(a_{\tau(1)}, \ldots, a_{\tau(i)}, \ldots, a_{\tau(N)}\right)\right) \text{ for any permutation } \tau: N \to N:$$

This implies that the conditional distribution on signals received by the other players is the same over all the players given the same action, the same personal signal profile and the same belief about the action profile which is actually taken. This assumption makes it possible to treat agents symmetrically combined with the assumption of the common utility function.

Although Assumption 1 is strong enough to achieve an almost efficient outcome for two players, a stronger assumption is called upon to achieve similar results with more than two players. Let $\#(\omega|a)$ stand for the number of errors in ω: The following assumptions is strong enough for that purpose:

Assumption 2.

$$p(\omega'|a) = p(\omega''|a) \text{ if } \#(\omega'|a) = \#(\omega''|a) \text{ for any } \omega', \omega'' \in \Omega, a \in A$$

A couple of remarks on these assumptions are in order.

First, Assumption 1 is a relatively weak assumption about the symmetry of a signal distribution and satisfied in most of the papers in reference which analyze the repeated prisoner's dilemma with almost perfect private monitoring. Second, while Assumption 2 is much stronger than Assumption 1 in general, it is very close to Assumption 1 in the two player case. Consequently, this assumption is also satisfied in those papers as most of them concentrate on the two player case.

Assumption 2 means that the probability of some signal profile only depends on the number of errors contained in that profile. For example, given that everyone is playing C; the probability that a player receives two "D" signals while the other players get correct signals is equal to the probability that two players receive one "D" while the rest of the players gets correct signals.

For example, the following information structure satisfies Assumption 2 for general n:

- **Example: Totally Decomposable Case**

  $p(\omega|a) = \prod_j \prod_{i \neq j} p(\omega_{i,j}|a_j)$ for all $a \in A$ and $\omega \in \Omega$

  Given the action by player j; the probability that player $i \neq j$ receives the right signal or the wrong signal about player j's action is the same across $i \neq j$. Also note that players' signals are conditionally independent over players.

4. Belief Dynamics and Best Response

In Section 5, an approximately efficient sequential equilibrium is constructed. Since this game belongs to a class of games called a game of nonobservable deviation, corresponding to any Nash equilibrium, there exists a sequential equilibrium which generates the same outcome distribution as the Nash equilibrium[2]. So, finding a particular sequential equilibrium is essentially equivalent to finding the corresponding Nash equilibrium, which is an easier task in general.

Later, a strategy profile is proposed and shown to be an approximately efficient Nash equilibrium, henceforth sequential equilibrium. In order to verify that the

[2] See Lemma 2 in [8] for detail.

proposed profile is a Nash equilibrium, it is shown that one sufficient condition for Nash equilibrium is satisfied, that is, it is checked that any action which is not assigned by the proposed strategy cannot be optimal at every history which is realized with positive probability[3]. In other word, the proposed strategy always assigns the unique optimal action at any such history. This section provides a couple of preliminary results for this procedure.

The strategy used to construct a Nash equilibrium is a mixture of σ_C and σ_D. So, players continuation strategies are always either grim trigger or permanent defection after any private history including histories which are never reached. This implies that the only crucial information is whether some player has started defecting or not. This is why players can restrict attention to belief q_{-i} with respect to their own decision. Players' best response strategy is just a function of q_{-i}, or more precisely, a function of $\phi(q_{-i})$. This fact allows one to decompose the argument into two parts. In the first subsection, the optimal action is characterized as a function of q_{-i}. The next subsection analyzes the dynamics of q_{-i} for an initial level mixture of σ_C and σ_D. Finally, combining these pieces together, it is proved in Section 5 that some mixture of σ_C and σ_D assigns the unique optimal action at any history which is realized with positive probability.

4.1. Beliefs and Optimal Action

Take a grim trigger strategy equilibrium with perfect monitoring in the two player case to get some insight into the imperfect monitoring case. Figure 1 shows the payoff difference between σ_C and σ_D depending on q_{-i}, a probability to play σ_D. The payoff difference $V_i(\sigma_C; q_{-i} : p_0; \delta) - V_i(\sigma_D; q_{-i} : p_0; \delta)$ is linear and decreasing in q. The level of mixture which makes the other player indifferent between σ_C and σ_D is denoted by $q^*(\delta; p_0)$. With perfect monitoring, given any q_{-i}, C (resp. D) is the optimal action and σ_C (resp. σ_D) is actually the optimal continuation strategy when σ_C (resp. σ_D) is preferred to σ_D (resp. σ_C). So, the optimal action and continuation strategy are functions of q_{-i}.

[3] This is the path dominance argument in Sekiguchi [13].

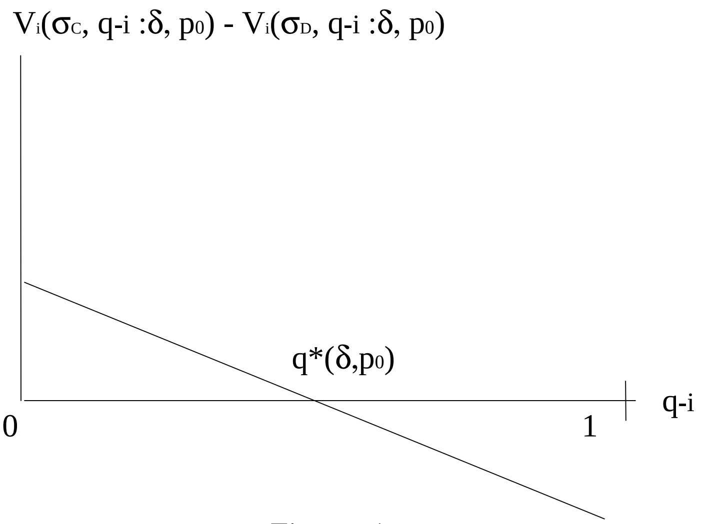

Figure 1

It is easy to check that $(\sigma_C; \sigma_C)$ is a subgame perfect equilibrium for a large enough δ. This is because the value of q_{-i} takes only 0 or 1 after any history and the unique optimal action is clearly C or D respectively. Similarly, $(\sigma_D; \sigma_D)$ and $((1-q^*)\sigma_C + q^*\sigma_D; (1-q^*)\sigma_C + q^*\sigma_D)$ are also subgame perfect equilibria. However, it turns out that $(\sigma_C; \sigma_C)$ is not robust with respect to the introduction of private noise. In order to achieve efficiency, I use the last equilibrium, which is robust to the introduction of private noise in the sense that there exists a sequential equilibrium which is close to the equilibrium.

When the number of players is more than two; $V_i(\sigma_C; q_{-i} : p_0; \delta) - V_i(\sigma_D; q_{-i} : p_0; \delta)$ is a slightly more complex object. Even when players randomize independently and symmetrically, that is, $q_{-i}(k) = \sum_{k=0}^{n-1} q^k (1-q)^{n-1-k} \binom{n-1}{k}$ for $k = 0; \ldots; n-1$, it is a $n-1$ degree polynomial in $q \in (0;1)$: Potentially, this equation may have $n-1$ solutions between 0 and 1 as shown in figure 2. In such a case, $q^*(\delta; p_0)$ is defined to be the solution which is closest to 0.

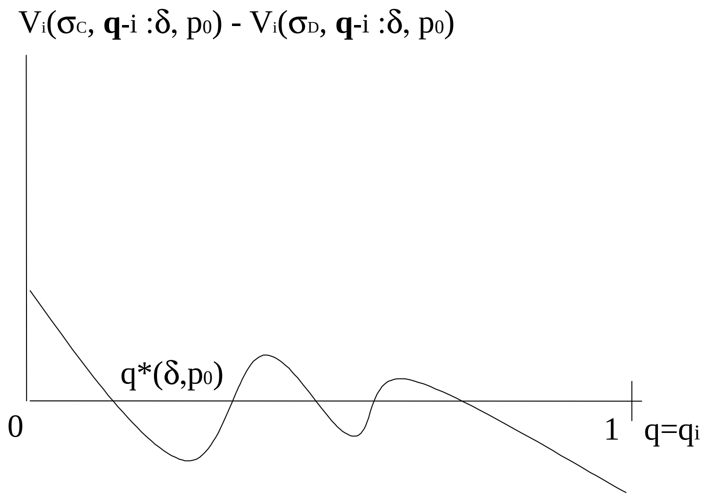

Figure 2

Now, let's move to the world of imperfect private monitoring. First, I need to find $q^*(\delta; p_\varepsilon)$ because I let players to randomize between σ_C and σ_D in the first period: When monitoring is almost perfect, $V_i(\sigma_C; \mathbf{q}_{-i} : p_\varepsilon; \delta) - V_i(\sigma_D; \mathbf{q}_{-i} : p_\varepsilon; \delta)$ is very close to $V_i(\sigma_C; \mathbf{q}_{-i} : p_0; \delta) - V_i(\sigma_D; \mathbf{q}_{-i} : p_0; \delta)$: Actually, it is easy to confirm that the former converges to the latter uniformly in q as $\varepsilon \to 0$.[4] So, $q^*(\delta; p_\varepsilon)$ is very close to $q^*(\delta; p_0)$ when ε is very small.

Whenever $\delta > \frac{\mathbf{M}(0)}{1+\mathbf{M}(0)}$; then $V_i(\sigma_C; 0 : p_0; \delta) - V_i(\sigma_D; 0 : p_0; \delta) > 0$; which implies that there exists $q^*(\delta; p_0)$ between 0 and 1. The following lemma is useful later to construct an almost efficient sequential equilibrium.

Lemma 1. $q^*(\delta; p_0) \to 0$ as $\delta \downarrow \frac{\mathbf{M}(0)}{1+\mathbf{M}(0)}$

Proof. See Appendix.

Next, I characterize the optimal action for each $\mathbf{q}_{-i}$ with private monitoring. When information is perfect, it is trivial to see what is the best response because

[4] Also note that convergence of $V_i(\sigma_i; q_j : p_\varepsilon)$ to $V_i(\sigma_i; q_j : p_0)$ is independent of the choice of associated sequence $\{p_\varepsilon\}$ because of the definition of P_ε.

the dynamics of q_{-i} is very simple. It is common knowledge whether everyone is playing σ_C or σ_D at any history after the initial period. On the other hand, when information is almost perfect, the dynamics of q_{-i} can be very complex because it cannot reach the absorbing states where a player is confident that everyone is playing the same strategy σ_C or σ_D.

However, it can be expected that the optimal action as a function of q_{-i} is similar to the one with perfect monitoring if the dynamics of q_{-i} is very close to the dynamics of q_{-i} with perfect monitoring. As a first step to show that, the following lemma shows that σ_D is still optimal if a player knows that someone has switched to the permanent defection and ε is small.

Lemma 2. There exists a $\bar{\varepsilon} > 0$ such that $V_i(\sigma_i; q_{-i} : p_\varepsilon; \delta)$ is maximized by σ_D for any $p_{\bar{\varepsilon}}$, if $q_{-i}(\theta) = 1$ for any $\theta \neq 0$.

Proof.

Take σ_D and any strategy which starts with C. The least deviation gain is $(1-\delta)\underline{\triangle}$. The largest loss caused by the difference in continuation payoffs with σ_D and the latter strategy is $\delta\varepsilon\overline{V}$. Setting $\bar{\varepsilon}$ small enough guarantees $(1-\delta)\underline{\triangle} > \delta\varepsilon\overline{V}$ for any $\varepsilon \in (0, \bar{\varepsilon})$. Then, D must be the optimal action for any such ε. Since players are using permanent defection, $q_{-i}(\theta) = 1$ for some $\theta \neq 0$ in the next period. This implies that D is the optimal action in all the following periods. ■

Using $p(\cdot|\cdot)$ and given the fact that players are playing either σ_C or σ_D, I can define a transition probability of the number of players who have switched to σ_D. Let $\pi(l|m)$ be a probability that l players will play σ_D from the next period when m players are playing σ_D now. In other words, this $\pi(l|m)$ is a probability that $l-m$ players playing C receive the signal D when $n-m$ players play C and m players play D. Of course, $\pi(l|m) > 0$ if $l \geq m$ and $\pi(l|m) = 0$ if $l < m$. The following lemma provides various informative and useful bounds on the variations of discounted average payoffs caused by introducing small imperfectness in private monitoring.

Lemma 3.

1. $\inf_{p_\varepsilon \in P_\varepsilon} V_i(\sigma_C; 0 : p_\varepsilon; \delta) \geq \frac{(1-\delta)+\delta\varepsilon\overline{V}}{1-\delta(1-\varepsilon)}$

2. Given $\delta \in \left(\frac{M(0)}{1+M(0)}, 1\right)$, There exists a $\bar{\bar{\varepsilon}} > 0$ such that for any $\varepsilon \in [0, \bar{\bar{\varepsilon}}]$,

$$\sup_{\sigma_i; p_\varepsilon \in P_\varepsilon} V_i(\sigma_i; 0 : p_\varepsilon; \delta) \leq \frac{1-\delta+\delta\varepsilon\overline{V}}{1-\delta(1-\varepsilon)}$$

Proof.

(1): For any $\varepsilon \in (0,1)$ and $p_\varepsilon \in P_\varepsilon$;

$$V_i(\sigma_C; 0 : p_\varepsilon; \delta) = (1-\delta) + \delta\pi(0|0)\, V_i(\sigma_C; 0 : p_\varepsilon; \delta) + \delta(1-\pi(0|0))\,\underline{V}$$

So,

$$V_i(\sigma_C; 0 : p_\varepsilon; \delta) = \frac{(1-\delta) + \delta(1-\pi(0|0))\underline{V}}{1-\delta\pi(0|0)} = \frac{(1-\delta) + \delta\varepsilon\underline{V}}{1-\delta(1-\varepsilon)}$$

(2): Given $\delta \in \left(\frac{\mathbf{M}(0)}{1+\mathbf{M}(0)}; 1\right)$; it is easy to check that $V_i(\sigma_C; 0 : p_0; \delta) > V_i(\sigma_D; 0 : p_0; \delta)$: Pick $\bar\varepsilon$ small enough such that (i) $V_i(\sigma_C; 0 : p_{\bar\varepsilon}; \delta) > V_i(\sigma_D; 0 : p_{\bar\varepsilon}; \delta)$ for any $p_{\bar\varepsilon}$ and (ii) $\bar\varepsilon < \bar{b}$: Let σ_0^* be the optimal strategy given that everyone is using σ_C.[5] Suppose that σ_0^* assigns D for the first period. Then for any $\varepsilon \in [0;\bar\varepsilon]$;

$$V_i(\sigma_0^*; 0 : p_\varepsilon) \le \left\{ (1-\delta)\, U((D; D^0) : p_\varepsilon) + \delta\left(\pi(1|1)\, V_i(\sigma_0^*; 0 : p_\varepsilon; \delta) + \sum_{k=2} \pi(k|1)\, V_i(\sigma_D; k-1 : p_\varepsilon; \delta)\right) \right\}$$

In this inequality, the second component represents what player i could get if she knew the true continuation strategies of her opponents at each possible state. To see that this additional information is valuable, suppose that the continuation strategy of σ_0^* leads to a higher expected payoff than $V_i(\sigma_0^*; 0 : p_\varepsilon; \delta)$ or $V_i(\sigma_D; k-1 : p_\varepsilon; \delta)$ at the corresponding states, then this contradicts the optimality of σ_0^* or σ_D by Lemma 2. So this inequality holds.

Then, for any $\varepsilon \in [0;\bar\varepsilon]$;

$$\begin{aligned} V_i(\sigma_0^*; 0 : p_\varepsilon; \delta) &\le \frac{(1-\delta)\, U(D; D_0 : p_\varepsilon) + \delta \sum_{k=2} \pi(k|1)\, V_i(\sigma_D; k-1 : p_\varepsilon; \delta)}{1-\delta\pi(1|1)} \\ &= V_i(\sigma_D; 0 : p_\varepsilon; \delta) \\ &< V_i(\sigma_C; 0 : p_\varepsilon; \delta) \end{aligned}$$

Since this contradicts the optimality of σ_0^*; σ_0^* has to assign C for the first period.

[5] This σ^* exists because the strategy space is a compact space in product topology, on which discounted average payoff functions are continuous. Of course, this σ^* depends on the choice of p_ε:

Now,

$$V_i(\sigma_0^*; 0 : p_\epsilon; \delta) \leq (1-\delta) + \delta\pi(0|0) V_i(\sigma_0^*; 0 : p_\epsilon; \delta) + \delta(1-\pi(0|0))\overline{V}$$

So,

$$V_i(\sigma_0^*; 0 : p_\epsilon; \delta) \leq \frac{(1-\delta) + \delta(1-\pi(0|0))\overline{V}}{1-\delta\pi(0|0)} \leq \frac{(1-\delta) + \delta\epsilon\overline{V}}{1-\delta(1-\epsilon)}$$

This implies that $\sup_{\sigma_i; p_\epsilon \in P_\epsilon} V_i(\sigma_i; 0 : p_\epsilon; \delta) \leq \frac{1-\delta+\delta\epsilon\overline{V}}{1-\delta(1-\epsilon)}$ for any $\epsilon \in [0; \overline{\epsilon}]$. ■

(1) means that a small departure from the perfect monitoring does not reduce the payoff of σ_C much when all the other players are using a grim trigger strategy. (2) means that there is not much to be exploited by using other strategies than σ_C with a small imperfection in the private signal as long as all of the other players are using a grim trigger strategy.

The main result in this section shows that the unique optimal action is almost completely characterized as a function of q_{-i} except for an arbitrary small neighborhood and equivalent to the optimal action with perfect monitoring

Proposition 1. Given δ; for any $\eta > 0$; there exists a $\overline{\epsilon} > 0$ such that for any p_ϵ;

- it is not optimal to play C for player i if q_{-i} satisfies $\phi(q_{-i}) \geq 1 - \frac{1-\delta}{\delta} M(q_{-i}; p_0) + \eta$
- it is not optimal to play D for player i if q_{-i} satisfies $\phi(q_{-i}) \leq 1 - \frac{1-\delta}{\delta} M(q_{-i}; p_0) - \eta$

Proof:
(1): It is not optimal to play C if

$$\begin{aligned}(1-\delta)\ M\ (q_{-i}; p_\epsilon) &\\ > \delta\ (1-\phi(q_{-i})) &\left\{(1-\epsilon)\sup_{\sigma_i} V_i(\sigma_i; 0 : p_\epsilon; \delta) + \epsilon\overline{V}\right\} + \phi(q_{-i})\epsilon\overline{V}\end{aligned}$$

By Lemma 3.2., this inequality is satisfied for any $\epsilon \in [0; \overline{\epsilon}]$ and any p_ϵ if

$$
\begin{aligned}
&(1-\delta)\,\mathbf{M}\,(q_{-i},p_\varepsilon)\\
&\quad > \delta\left[(1-\phi(q_{-i}))\left\{(1-\varepsilon)\frac{1-\delta+\delta\varepsilon\overline{V}}{1-\delta(1-\varepsilon)}+\varepsilon\overline{V}\right\}+\phi(q_{-i})\,\varepsilon\overline{V}\right]
\end{aligned}
$$

LHS converges to $(1-\delta)\,\mathbf{M}\,(q_{-i},p_0)$ and RHS converges to $\delta(1-\phi(q_{-i}))$ as $\varepsilon \to 0$. So, if q_{-i} satisfies $\phi(q_{-i}) \geqq 1-\frac{1-\delta}{\delta}\mathbf{M}\,(q_{-i},p_0)+\eta$ for any $\eta>0$, then there exists a $\varepsilon'(\delta,\eta,q_{-i}) \in (0,\bar{\varepsilon})$ and a neighborhood $B(q_{-i})$ of q_{-i} such that C is not optimal for any $p_{\varepsilon'(\delta,\eta,q_{-i})}$ and any $q'_{-i} \in B(q_{-i})$. This $\varepsilon'(\delta,\eta,q_{-i})>0$ can be set independent of q_{-i} by the standard arguments because q_{-i} is in a compact space.

(2). It is not optimal to play D if

$$
\begin{aligned}
&(1-\delta)\,\mathbf{M}\,(q_{-i},p_\varepsilon)\\
&\quad < \delta\left[(1-\phi(q_{-i}))\{(1-\varepsilon)V_i(\sigma_C,0:p_\varepsilon,\delta)+\varepsilon\underline{V}\}+\phi(q_{-i})\,\varepsilon\underline{V}-\varepsilon\overline{V}\right]
\end{aligned}
$$

this inequality is satisfied for $\varepsilon \in (0,1)$ and any p_ε if

$$
\begin{aligned}
&(1-\delta)\,\mathbf{M}\,(q_{-i},p_\varepsilon)\\
&< \delta(1-\phi(q_{-i}))\left\{(1-\varepsilon)\frac{1-\delta+\delta\varepsilon\underline{V}}{1-\delta(1-\varepsilon)}+\varepsilon\underline{V}\right\}+\delta\phi(q_{-i})\,\varepsilon\underline{V}-\delta\varepsilon\overline{V}
\end{aligned}
$$

This inequality converges to $\phi(q_{-i}) \leqq 1-\frac{1-\delta}{\delta}\mathbf{M}\,(q_{-i},p_0)$ as $\varepsilon \to 0$. So, if q_{-i} satisfies $\phi(q_{-i}) \leqq 1-\frac{1-\delta}{\delta}\mathbf{M}\,(q_{-i},p_0)-\eta$ for any $\eta>0$, there exists a $\varepsilon''(\delta,\eta,q_{-i})$ such that D is not optimal for any $p_\varepsilon \in P_{\varepsilon''(\delta,\eta,q_{-i})}$ and any q'_{-i} around q_{-i}. Again, $\varepsilon''(\delta,\eta,q_{-i})$ can be set independent of q_{-i}.

Finally, setting $\bar{\varepsilon}(\delta,\eta) = \min\{\varepsilon'(\delta,\eta),\varepsilon''(\delta,\eta)\}$ completes the proof. ■

This proposition implies that the optimal action can be completely characterized except for an arbitrary small neighborhood of the manifold satisfying $\phi(q_{-i}) = 1-\frac{1-\delta}{\delta}\mathbf{M}\,(q_{-i},p_0)$ in a $n-1$ dimensional simplex[6], where player i is indifferent between σ_C and σ_D with perfect monitoring.

[6] Abusing notation, q_{-i} is used for both a mapping and a point on a $n-1$ dimentional simplex.

Although a similar argument is given in Sekiguchi [13] for $n = 2$, this proposition for general n actually provides a weaker incentive constraint with respect to the optimality of D when $n = 2$. The lower bound of q_{-i} to make D optimal for player i, given a certain level of δ, is lower than the bound in Sekiguchi [13], which has restricted the class of prisoner's dilemma for which almost efficient outcome can be achieved. In fact, this proposition serves as a preliminary result to achieve the almost efficient outcome for any prisoner's dilemma in the two player case.

An immediate corollary of this proposition is that C is the unique optimal action given that ψ is close to 0, $\delta > \frac{\mathbf{M}(0)}{1+\mathbf{M}(0)}$ and ε is small.

Corollary 1. Given $\delta > \frac{\mathbf{M}(0)}{1+\mathbf{M}(0)}$, there exists $\underline{\psi} > 0$ and $\varepsilon > 0$ such that for any p_{ε}, it is not optimal for player i to play D if $\psi \in \left[0, \underline{\psi}\right]$.

4.2. Belief Dynamics

Since all players are playing either grim trigger strategy or permanent defection, the most important information is whether there is anyone who has switched to permanent defection or not. The number of players who have started defecting is not important with almost perfect monitoring. Players only need to keep track of $\psi_i^t = \psi(q_{-i}(h_i^t))$. In this subsection, the dynamics of ψ_i^t under the grim trigger strategy and permanent defection is analyzed.

Since the unique optimal action is almost characterized in the last subsection, all I have to make sure is that q_{-i} stays in the "C area" described by Proposition 1 as long as player i has observed full cooperation from the beginning and q_{-i} stays in the "D area" once player i received a bad signal or started playing defection for herself. Assumption 2 on the signal distribution is required here for the first time as the following arguments show.

First, consider the history where players have observed perfect cooperation. Suppose that every player i mixes σ_C and σ_D with $(1 - q^*, q^*)$ at the first period. It is not difficult to see that $\psi(q_{-i})$ moves into the "C area" in the second period if ε is set small enough. Since private signals are almost perfect, (C, C) clearly signals that every player has picked σ_C and observed correct signals. Then, q_{-i} moves into the "C area" in the second period and cannot move outside of the area as long as full cooperation continues to be observed.

Second, consider the history where player i observes some defection for the first time. If this is the first period, player i interprets this as a signal of σ_D

rather than as an error if ε is small[7]. Suppose next that this kind of history is reached after the first period. Also suppose that the number of players is three for simplicity and player 1 observes 1 defection by player 2. With Assumption 2, player 1 can interpret this as a 1-error event and still believe that everyone is cooperative. On the other hand, it is equally likely that player 2's observation contained 1 error in the last period and the current signal is correct. Note that there are two such events. The player for whom player 2 observed "D" last period can be player 1 or player 3. Since someone should have already defected after all other possible histories, the probability that someone has switched to σ_D is at least $\frac{2}{3}$. Obviously, this flexibility of interpretation increases as the number of players increases, which makes it easier to move ψ closer to 1 after this kind of history. Note that this lower bound of ψ does not depend on the level of ε.

Finally, consider the history where player i has already started defection. Suppose that all players but player i have been cooperative until the present. Also suppose again that the number of players is three and $i = 1$ for the sake of simple exposition. For everyone to be still cooperative after the current period, all players but player 1 should have observed the wrong signal "C" about player 1 and the correct signal "C" about the other players in the current period. Again, there are other events with the same probability, where some player switches to σ_D. For example, player 2 may observe the correct signal "D" about player 1 and the wrong signal "D" about player 3. This event contains the same number of errors. Since there are 5 such events, the probability that someone has switched to σ_D is at least $\frac{5}{6}$ even though it is assumed that all players but player i have been cooperative until the current period. With positive probability that someone has already started defection, the posterior ψ is strictly higher than $\frac{5}{6}$. This argument is again independent of the level of ε.[8]

The following proposition summarizes these arguments.

Proposition 2. Suppose that every player plays $(1 - q^*)\sigma_C + q^*\sigma_D$ with $q^* \in (0, 1)$ in the first period, and (i) : Assumption 1 is satisfied and $n = 2$, or (ii) : Assumption 2 is satisfied. Then for all i and $t = 2, 3, \ldots$

- For any $\psi' > 0$, there exists ε' such that for any $\varepsilon \in (0, \varepsilon')$

[7] This argument needs players to randomize between σ_C and σ_D in the initial period. If players start with, say, σ_C with probability 1, no learning occurs after the initial period. This is first observed by Matsushima [10]. Note that this is the only reason why the initial randomization is needed. The rest of arguments does not depend on this initial randomization.

[8] This last argument is specific to the $n = 3$ player case. The two player case has to be treated separately. See Sekiguchi [13] for that case.

$\chi(q_{-i}(h_i^t)) \leqq \chi'$ after $h_i^t = ((C,C), \ldots, (C,C))$.

- $\chi(q_{-i}(h_i^t)) \geqq \frac{n-1}{n}$ after histories such as

  – $\begin{cases} h_i^t = \left(h_i^{t-1} = ((C,C), \ldots, (C,C)), \left(C, \omega_i^{t-1}\right)\right) & \text{for } t \geqq 3 \\ h_i^t = \left(C, \omega_i^{t-1}\right) & \text{for } t = 2 \end{cases}$

  with $\omega_i^{t-1} \neq C$

  or

  – $\begin{cases} h_i^t = \left(h_i^{t-1}, \left(D, \omega_i^{t-1}\right)\right) & \text{for } t \geqq 3 \\ h_i^t = \left(D, \omega_i^{t-1}\right) & \text{for } t = 2 \end{cases}$

Proof. See appendix.

5. Sequential Equilibrium with the "Grim Trigger" Strategy

Let $Q^D = \left\{q_{-i} | \chi(q_{-i}) \geqq \frac{n-1}{n}\right\}$ and $Q_0^I = \left\{q_{-i} | \chi(q_{-i}) = 1 - \frac{1-\delta}{\delta}\Delta(q_{-i}, p_0)\right\}$ be a subset of a $n-1$ dimensional simplex on Σ. The former subset Q^D is a set containing the absorbing set of the dynamics of q_{-i} under the grim trigger. The latter subset Q_0^I is a manifold where player i is indifferent between σ_C and σ_D with $\varepsilon = 0$. In particular, $q^*_{-i}(\delta; p_0) \in Q_0^I$ by definition. The main proposition of this section shows that if these sets are disjoint, then there exists a sequential equilibrium which is realization equivalent to $(\ldots, (1 - q^*(\delta; p_\varepsilon))\sigma_C + q^*(\delta; p_\varepsilon)\sigma_D, \ldots)$[9].

Proposition 3. Suppose that (i): Assumption 1 is satisfied and $n = 2$, or (ii): Assumption 2 is satisfied. Given $\delta \in \left(\frac{M(0)}{1+M(0)}, 1\right)$, if $Q^D \cap Q^I = \emptyset$, then there is $\varepsilon > 0$ such that for any p_ε, there exists a sequential equilibrium which generates the same outcome distribution as $(\ldots, (1 - q^*(\delta; p_\varepsilon))\sigma_C + q^*(\delta; p_\varepsilon)\sigma_D, \ldots)$.

Proof. Since $\delta \in \left(\frac{M(0)}{1+M(0)}, 1\right)$, $q^*(\delta; p_\varepsilon)$ exists in $(0,1)$ if ε is small enough. First, we show that $(\ldots, (1 - q^*(\delta; p_\varepsilon))\sigma_C + q^*(\delta; p_\varepsilon)\sigma_D, \ldots)$ forms a Nash equilibrium.

Suppose that a player chooses σ_C as her strategy in the first period. Set $\chi' < \underline{\chi}$ and $\varepsilon > 0$ small enough for Corollary 1 and Proposition 2 to hold. If she observed

[9] In precise, any profile which is realization equivalent to $(\ldots, (1 - q^*(\delta; p_\varepsilon))\sigma_C + q^*(\delta; p_\varepsilon)\sigma_D, \ldots)$ is a Nash equilibrium for which there exists a corresponding payoff equivalent sequential equilibrium.

$\omega_i = C$ at the first period; then her belief goes down to $\psi_i^2 = \psi(q_{-i}(C,C)) \leq \psi^0 < \underline{\psi}$ and $\{\psi_i^t\}_{t=2}^{\infty}$ will never go above ψ^0 as long as she continues observing C by Proposition 2. Then, by Corollary 1, it is always optimal to play C after such a history as $h_i^t = ((C,C),\ldots,(C,C))$.

Consider a history where player i observed some D for the first time or a history where player i started playing D. After this sort of history, $\psi_i^t = \psi(q_{-i}(h_i^t))$ is going to stay in Q^D forever by Proposition 2 because this player is using σ_D. It is possible to take a small number $\eta' > 0$ such that $\forall q_{-i} \in Q^D, \psi(q_{-i}) \geq 1 - \frac{1-\delta}{\delta} M(q_{-i}, p_0) + \eta'$ because Q^D is compact, connected and $Q^D \cap Q_0^I = \emptyset$. Then D is the unique optimal action for any p_ε and $q_{-i} \in Q^D$ if ε is small enough by Proposition 1.

Taking ε small such that all the above arguments go through, we can confirm that $(\ldots,(1-q^*(\delta,p_\varepsilon))\sigma_C + q^*(\delta,p_\varepsilon)\sigma_D,\ldots)$ is a symmetric Nash equilibrium. Finally, there exists a sequential equilibrium which generates the same outcome distribution as this grim trigger strategy Nash equilibrium because this equilibrium has no observable deviation. ∎

Since the probability that everyone chooses σ_C in this sequential equilibrium, $(1-q^*(\delta,p_\varepsilon))^{n-1}$, gets closer to 1 as δ gets closer to $\frac{M(0)}{1+M(0)}$ by Lemma 1, an outcome arbitrary close to the efficient outcome can be achieved for δ arbitrary close to $\frac{M(0)}{1+M(0)}$. For high δ, Ellison's trick in [5] can be used to achieve an almost efficient outcome although the strategy is more complex and no longer a grim trigger. It is also possible to use a public randomization device to reduce δ effectively as in [2]. Here is the corollary of Proposition 3 with regard to an approximately efficient outcome.

Corollary 2. Suppose that (i): Assumption 1 is satisfied and $n = 2$, or (ii): Assumption 2 is satisfied. For any $k > 0$, if $Q^D \cap Q^I = \emptyset$, then there is a $\varepsilon > 0$ such that for any p_ε, there exists a sequential equilibrium whose symmetric equilibrium payoff is more than $1 - k$.

When is $Q^D \cap Q_0^I = \emptyset$ satisfied? First of all, this is always satisfied when $n = 2$ for a range of δ with $q^*(\delta,p_0) \in \left(0,\frac{1}{2}\right]$; that is, when δ is between $\frac{M(0)}{1+M(0)}$ and $\frac{M(0)+M(1)}{1+M(0)+M(1)}$. This is a special case of the following proposition which provides sufficient conditions for $Q^D \cap Q_0^I = \emptyset$ for general n.

Proposition 4. :

- If $\mathbf{M}(k) = \mathbf{M}$ for $k = 1, \ldots, n-1$, then $Q^D \cap Q^I_0 = \emptyset$ and $Q^I_0 \neq \emptyset$ for $\delta \in \left(\frac{\mathbf{M}}{1+\mathbf{M}}, \frac{n\mathbf{M}}{1+n\mathbf{M}}\right)$
- Regarding n as a parameter, take a sequence of the stage game with $n = 2, 3, \ldots.$ If there exists a lower bound $\underline{\mathbf{M}} > 0$ such that $\min_{1 \le k \le n-1} \mathbf{M}(k) \geq \underline{\mathbf{M}}$ independent of n, then there exists $\underline{n}$ such that for all $n \geq \underline{n}$, $Q^D(n) \cap Q^I_0(n) = \emptyset$ and $Q^I_0(n) \neq \emptyset$ for $\delta \in \left(\frac{\mathbf{M}(0)}{1+\mathbf{M}(0)}, \frac{n\underline{\mathbf{M}}}{1+n\underline{\mathbf{M}}}\right)$

Proof: If the deviation gain is constant, $Q^I_0 = \left\{ q_{-i} \mid \phi(q_{-i}) = 1 - \frac{1-\delta}{\delta}\Delta \right\}$ So, $\frac{n-1}{n} > 1 - \frac{1-\delta}{\delta}\Delta$ iff $Q^D \cap Q^I_0 = \emptyset$. Combining this inequality with $\frac{\mathbf{M}}{1+\mathbf{M}} < \delta$ for $Q^I_0 \neq \emptyset$, $\delta \in \left(\frac{\mathbf{M}}{1+\mathbf{M}}, \frac{n\mathbf{M}}{1+n\mathbf{M}}\right)$ is obtained.

If $\underline{\mathbf{M}}$ is independent of n, $Q^D(n) \cap Q^I_0(n) = \emptyset$ if $\frac{n-1}{n} > 1 - \frac{1-\delta}{\delta}\underline{\mathbf{M}}$. So, $Q^D(n) \cap Q^I_0(n) = \emptyset$ and $Q^I_0(n) \neq \emptyset$ for $\delta \in \left(\frac{\mathbf{M}(0)}{1+\mathbf{M}(0)}, \frac{n\underline{\mathbf{M}}}{1+n\underline{\mathbf{M}}}\right)$ for all $n \geq \underline{n}$ if $\underline{n}$ is chosen such that $\frac{\mathbf{M}(0)}{1+\mathbf{M}(0)} < \frac{\underline{n}\underline{\mathbf{M}}}{1+\underline{n}\underline{\mathbf{M}}}$ ■

6. Characterization of the Sequential Equilibrium

When the other players are playing either σ_C or σ_D, Proposition 1 almost characterizes the optimal action as a function of belief q_{-i}, the probability on the number of players playing σ_D, if the information is almost perfect. For a fixed information structure with $\varepsilon > 0$, the optimal action is not characterized yet when q_{-i} is in some small neighborhood containing Q^I_0 where player i is indifferent between σ_C and σ_D. Although this area can be made arbitrary small by setting ε small, there always remains an area for fixed ε, where the optimal action is unknown. This is not a problem to construct a Nash equilibrium using a grim trigger strategy because q_{-i} never falls in such an area on the equilibrium path by construction. On the other hand, it is still possible that q_{-i} falls in that area off the equilibrium path, that is, after one's own deviation. Although it is true that there exists a sequential equilibrium realization equivalent to the grim trigger Nash equilibrium and that equilibrium assigns some optimal action after any private history, the way this sequential equilibrium is constructed does not give any information about off the equilibrium path behavior.

In this section, the sequential equilibrium constructed in the last section is examined in detail. It turns out that this sequential equilibrium has a natural structure. Given the probability q_{-i}, the unique optimal action and the optimal repeated game strategy for player i has the same structure as in the perfect monitoring case. In the two player case, if $q_{-i} < q^*(\delta, p_\varepsilon)$, then C is the unique optimal action and the optimal strategy is realization equivalent to σ_C. On the other hand, D is the uniquely optimal action and the optimal strategy is realization equivalent to σ_D if $q_{-i} > q^*(\delta, p_\varepsilon)$.

To sum up, the best response strategy to any mixture of σ_C and σ_D is realization equivalent to σ_C or σ_D and unique for almost all level of mixture. Note that the continuation strategy after any history which is never reached is almost completely determined with almost perfect private monitoring while it is totally arbitrary with perfect monitoring.

The number of players is set to be two in the proof for simplicity, but the following proof directly carries over to the n player case as long as the assumptions for Proposition 4 are satisfied.

Proposition 5. Suppose that (i) : Assumption 1 is satisfied and $n = 2$, or (ii) : Assumption 2 is satisfied. Given $\delta \in \left(\frac{\mathbf{M}(0)}{1+\mathbf{M}(0)}, 1\right)$, if $Q^D \bigcap Q_0^I = \emptyset$, then there exists an $\varepsilon > 0$ such that for any p_ε,

- If q_{-i} satisfies $\phi(q_{-i}) > 1 - \frac{1-\delta}{\delta}\mathbf{M}(q_{-i}, p_0)$, then D is the unique optimal action and the best response strategy is realization equivalent to σ_D.

- If q_{-i} satisfies $\phi(q_{-i}) < 1 - \frac{1-\delta}{\delta}\mathbf{M}(q_{-i}, p_0)$, then C is the unique optimal action and the best response strategy is realization equivalent to σ_D.

Proof (n=2): Pick any $\eta > 0$ in Proposition 1 and set ε small enough for this proposition and Corollary1 to be true. Define $\overline{\phi} = 1 - \frac{1-\delta}{\delta}\Delta\left(q^*_{-i}, p_0\right) + \eta$. The unique optimal action for any $q_{-i} \in \left[\overline{\phi}, 1\right]$ is D and C is the unique optimal action for any $q_{-i} \in \left[0, \underline{\phi}\right]$. Pick any q_{-i} such that

$$q^*_{-i}(\delta, p_\varepsilon) < q_{-i} < \overline{\phi}$$

Note that $q^*_{-i}(\delta, p_\varepsilon) \in \left(\underline{\phi}, \overline{\phi}\right)$ because both C and D can be the optimal action with this belief $q^*_{-i}(\delta, p_\varepsilon)$.

Set ε very small such that $q^0_{-i} = \phi\left(f\left(\overline{\phi}, (C, C), p_\varepsilon\right)\right) \leqq \underline{\phi}$, where $f(q_{-i}, (C, C), p_\varepsilon)$ is a value of posterior q_{-i} given prior q_{-i} and the current action and signal (C, C). This function is strictly increasing function of q_{-i}. Now suppose that the optimal action for this q_{-i} is C. Since $q_{-i} < \overline{\phi}$, $\phi(f(q_{-i}, (C, C), p_\varepsilon)) < q^0_{-i} \leqq \underline{\phi}$ if she observes a "C": So the optimal continuation strategy is realization equivalent to σ_C with the dynamics of belief described in Proposition 2 if ε is small enough. On the other hand, if she observes D; the optimal continuation strategy is realization equivalent to σ_D because ψ is going to be more than $\frac{1}{2}$and stay in Q^D: This means that σ_C should be the optimal strategy given this q_{-i} if C is played now. This is a contradiction because σ_C is dominated by σ_D for this q_{-i} by definition. So, the unique optimal action for this q_{-i} is D. This in turn implies that the optimal continuation strategy for this q_{-i} is realization equivalent to σ_D given the dynamics of belief analyzed in Proposition 2.

The other case follows a similar logic. Pick any q_{-i} such that

$$\underline{\psi} < q_{-i} < q^*_{-i}(\delta; p_\varepsilon)$$

If the optimal continuation strategy assigns D now, the continuation strategy is going to be σ_D: This means that σ_D is actually one of the optimal continuation strategy for this q_{-i}; but this is a contradiction because σ_D is dominated by σ_C by definition. So, the optimal action is C and the optimal continuation strategy is realization equivalent to σ_C for this q_{-i} by Proposition 2 ■

7. Conclusion

In this paper, I clarify the incentive structure in a general repeated prisoner's dilemma with private monitoring when players are using a mixture of a grim trigger strategy and permanent defection, and provide the sufficient conditions under which the simple grim trigger strategy supports the efficient outcome as a sequential equilibrium for some range of discount factors. It is also shown that the best response to a mixture of grim trigger strategy and permanent defection can be characterized almost uniquely, which makes it possible to provide the clear representation of the sequential equilibrium supporting the efficient outcome .

There are two lines of research pursuing sustainability of the efficient outcomes or Folk Theorem in this class of game. One direction of research is based on grim trigger strategies. Such papers as Bhaskar [2], Sekiguchi [13] belong to this literature and so does this paper. The emphasis of these papers are on coordination of players' actions and beliefs. The assumption of almost public signal in Mailath and Morris [9] also works as a device to generate coordination.

The other direction of research is based on a complete mixed strategies which makes the other player indifferent over many strategies so that . Ely and Välimäki [6] and Piccione [12] are among papers in this direction.[10]

One advantage to the former approach is:

1. The equilibrium needs to use mixing only at the beginning of the game, while the latter approach uses the behavior strategy which let players to randomize at every period after every history.

   Another advantage, which is closely related to the first one, is as follows:

2. Since the strategy is an almost pure strategy, it is very easy to justify the use of a mixed strategy. Purification is straightforward for our strategy by introducing a small amounts of uncertainty into stage game payoffs. On the other hand, payoff uncertainty in stage game payoffs has to depend on a private history in a peculiar way to purify the completely mixed behavior strategy used in the latter approach.[11] It is also easy to adopt Nash's population interpretation to purify the former equilibrium. What I have in my mind is a pool of players who are matching with the other players to play a repeated game, where most of players use the grim trigger strategy and only a small portion of the players use permanent defection.

A relative disadvantage of the former approach is that monitoring is almost public or almost perfect to generate strong coordination among players. For example, ε has to be very small in this paper to make monitoring almost perfect. However, the latter approach might work in a more noisy environment. One such example can be found in Piccione [12], where an approximately efficient outcome is sustained with a monitoring technology which is neither almost perfect nor almost public.

[10] Obara [11] uses the same kind of strategy for repeated partenership games with public monitoring and constructs a sequential equilibrium which cannot be supported by public perfect equilibria.

[11] See [3] for the refinement of mixed strategies in repeated games along this line.

Appendix.

Proof of Lemma 1.

When $\delta = \frac{\Delta(0)}{1+\Delta(0)}$; $q^*(\delta; p_0) = 0$ is the solution of the equation in q:

$$V_i(\sigma_C; q_{-i} : p_0; \delta) - V_i(\sigma_D; q_{-i} : p_0; \delta) = 0$$

where $q_{-i}(k) = \sum_{k=0}^{n-1} q^k (1-q)^{n-1-k} \binom{n-1}{k}$ for $k = 0; ...; n-1$.

I just need to show that $\frac{\partial q^*(\delta;p_0)}{\partial \delta}\Big|_{\delta=\frac{\Delta(0;p_0)}{1+\Delta(0;p_0)}} > 0$ using the implicit function theorem. Since

$$\begin{aligned} & V_i(\sigma_C; q_{-i} : p_0; \delta) - V_i(\sigma_D; q_{-i} : p_0; \delta) \\ = \; & (1-\delta) \sum_{k=0}^{n-1} q^k (1-q)^{n-1-k} \binom{n-1}{k} \Delta(k; p_0) - \delta(1-q)^{n-1} \end{aligned}$$

$$\begin{aligned} \frac{\partial q^*(\delta; p_0)}{\partial \delta}\Big|_{\delta=\frac{\Delta(0)}{1+\Delta(0)}} &= -\frac{\frac{\partial V_i(\sigma_C;q^*:p_0;\delta) - V_i(\sigma_D;q^*:p_0;\delta)}{\partial \delta}}{\frac{\partial V_i(\sigma_C;q^*:p_0;\delta) - V_i(\sigma_D;q^*:p_0;\delta)}{\partial q}}\Bigg|_{\delta=\frac{\Delta(0)}{1+\Delta(0)}} \\ &= \frac{1+\Delta(0)}{(1-\delta)(n-1)\Delta(1) + \delta(n-1)}\Bigg|_{\delta=\frac{\Delta(0)}{1+\Delta(0)}} \\ &= \frac{1}{n-1}\frac{1}{\Delta(0)+\Delta(1)} > 0 \end{aligned}$$

■

Proof of Proposition2

case 1:

$h_i^t = ((C; C); ...; (C; C))$

Applying Bayes' Rule[12],

$$\begin{aligned} \phi_i^t &= \phi(q_{-i}(h_i^t)) \\ &= \frac{(1-\phi_i^{t-1})P(\exists j; \omega_j^{t-1} \neq C \text{ and } \omega_i^{t-1} = C \mid \theta_{t-1}=0) + \phi_i^{t-1}P(\omega_i^{t-1} = C \mid \theta_{t-1} \neq 0; h_i^{t-1})}{(1-\phi_i^{t-1})P(\omega_i^{t-1} = C \mid \theta_{t-1}=0) + \phi_i^{t-1}P(\omega_i^{t-1} = C \mid \theta_{t-1} \neq 0; h_i^{t-1})} \end{aligned}$$

[12] All the conditional distributions implicitly depend on the definition of σ_C; σ_D; and the level of initial mixture between σ_C; σ_D:

This function is increasing in ϕ_i^{t-1} and crosses 45° line once. Note that this function is bounded above by $\varphi\left(\phi_i^{t-1}\right) = \frac{\varepsilon}{\left(1-\phi_i^{t-1}\right)(1-\varepsilon)+\phi_i^{t-1}\varepsilon}$. Let $\hat{\phi}$ be the unique fixed point of this mapping. Given that players are mixing between σ_C and σ_D with $(1-q^*, q^*)$, it is easy to see that $\varphi\left(\phi_i^1\right)$ can be made smaller than any $\phi' > 0$ by choosing ε small enough. As long as players continue to observe C, $\varphi^n\left(\phi_i^1\right)$ is going to decrease monotonically to $\hat{\phi}$. On the other hand, since $\phi_i^2 \leq \varphi\left(\phi_i^1\right)$ and $\phi_i^n \leq \varphi\left(\phi_i^{n-1}\right)$, ϕ_i^n is less than $\varphi^n\left(\phi_i^1\right)$ for any n. This implies that $\left(\phi_i^t\right)_{t=2}$ is always below ϕ'.

$$\text{case 2: } \begin{cases} h_i^t = \left(h_i^{t-1} = ((C,C),\ldots,(C,C)), \left(C, \omega_i^{t-1}\right)\right) & \text{for } t \geq 3 \\ h_i^t = \left(C, \omega_i^{t-1}\right) & \text{for } t = 2 \end{cases}$$

with $\omega_i^{t-1} = \omega'_i \neq C$

Suppose that $t \geq 3$. By Bayes' Rule,

$$\phi_i^t = \phi\left(q_{-i}\left(h_i^t\right)\right)$$
$$= \frac{\left(1-\phi_i^{t-2}\right)P\left(\exists j;\ \omega_j^{t-2} \text{ or } \omega_j^{t-1} \neq C, \text{ and } \left(\omega_i^{t-2},\omega_i^{t-1}\right)=\left(C,\omega'_i\right) \mid \theta_{t-2}=0\right)+\phi_i^{t-2}P\left(\left(\omega_i^{t-2},\omega_i^{t-1}\right)=\left(C,\omega'_i\right) \mid \theta_{t-2}\neq 0;\ h_i^{t-2}\right)}{\left(1-\phi_i^{t-2}\right)P\left(\left(\omega_i^{t-2},\omega_i^{t-1}\right)=\left(C,\omega'_i\right) \mid \theta_{t-2}=0\right)+\phi_i^{t-2}P\left(\left(\omega_i^{t-2},\omega_i^{t-1}\right)=\left(C,\omega'_i\right) \mid \theta_{t-2}\neq 0;\ h_i^{t-2}\right)}$$

This is bounded below by

$$\frac{P\left(\exists j;\ \omega_j^{t-2} \text{ or } \omega_j^{t-2} \neq C, \text{ and } \left(\omega_i^{t-2}, \omega_i^{t-1}\right) = \left(C, \omega'_i\right) \mid \theta_{t-2} = 0\right)}{P\left(\left(\omega_i^{t-2}, \omega_i^{t-1}\right) = \left(C, \omega'_i\right) \mid \theta_{t-2} = 0\right)}$$
$$= \frac{(n-1)^{\#\left(\omega'_i \mid C\right)} P\left(\#\left(\omega'_i \mid C\right)\right)}{P\left(\#\left(\omega'_i \mid C\right)\right) + (n-1)^{\#\left(\omega'_i \mid C\right)} P\left(\#\left(\omega'_i \mid C\right)\right)}$$
$$= \frac{(n-1)^{\#\left(\omega'_i \mid C\right)}}{1 + (n-1)^{\#\left(\omega'_i \mid C\right)}}$$
$$\geq \frac{n-1}{n}$$

where $P\left(\#\left(\omega'_i \mid C\right)\right)$ is the probability of the event that $\#\left(\omega'_i \mid C\right)$ errors occur.

So, once players observed a bad signal for the first time, the posterior ϕ_i^t jumps up at least above $\frac{n-1}{n}$ independent of the prior ϕ_i^{t-1} or q_{-i}^{t-1} for $t \geq 3$. This argument is independent of the level of ε.

When $t = 2$, ψ_i^t is high enough to be more than $\frac{n-1}{n}$ if ε is very small. This is because players do not interpret it as an error but as a signal of σ_D at the first period.:

$$\text{case 3:} \quad \left\{ \begin{array}{ll} h_i^t = \left(h_i^{t-1}, \left(D, \omega_i^{t-1}\right)\right) & \text{for } t \geq 3 \\ h_i^t = \left(D, \omega_i^{t-1}\right) & \text{for } t = 2 \end{array} \right. \quad \text{with } \omega_i^{t-1} = \omega_i''$$

we have to treat (i) $n \geq 3$ and (ii) $n = 2$ separately again.

(i): $n \geq 3$

By Bayes' Rule,

$$\psi_i^t = \psi\left(q_{-i}\left(h_i^t\right)\right)$$
$$= \frac{\left(1-\psi_i^{t-1}\right)P\left(\exists j \neq i, \omega_j^{t-1} \neq C \text{ and } \omega_i^{t-1} = \omega_i'' | \theta_{t-1} = 1\right) + \psi_i^{t-1}P\left(\omega_i^{t-1} = \omega_i'' | \theta_{t-1} > 1, h_i^{t-1}\right)}{\left(1-\psi_i^{t-1}\right)P\left(\omega_i^{t-1} = \omega_i'' | \theta_{t-1} = 1\right) + \psi_i^{t-1}P\left(\omega_i^{t-1} = \omega_i'' | \theta_{t-1} > 1, h_i^{t-1}\right)}$$

This is bounded below by

$$\frac{P\left(\exists j \neq i, \omega_j^{t-1} \neq C \text{ and } \omega_i^{t-1} = \omega_i'' | \theta_{t-1} = 1\right)}{P\left(\omega_i^{t-1} = \omega_i'' | \theta_{t-1} = 1\right)}$$
$$= \frac{P\left(\#\left(\omega_i'' | C\right) + n - 1\right) \sum_{m=1}^{n-1} \binom{(n-2)(n-1)}{m} \binom{n-1}{m}}{P\left(\#\left(\omega_i'' | C\right) + n - 1\right) + P\left(\#\left(\omega_i'' | C\right) + n - 1\right) \sum_{m=1}^{n-1} \binom{(n-2)(n-1)}{m} \binom{n-1}{m}}$$
$$= \frac{\sum_{m=1}^{n-1} \binom{(n-2)(n-1)}{m} \binom{n-1}{m}}{1 + \sum_{m=1}^{n-1} \binom{(n-2)(n-1)}{m} \binom{n-1}{m}}$$
$$\geq \frac{5}{6} \text{ (This holds with equality when } n = 3\text{)}$$

This argument is independent of ε, too.

(ii): $n = 2$

$$\psi\left(q_{-i}\left(h_i^{t-1}, (D, C)\right)\right) = \frac{\left(1-\psi_i^{t-1}\right)(1-2P(1)-P(2)) + \psi_i^{t-1}(P(1)+P(2))}{\left(1-\psi_i^{t-1}\right)(1-P(1)-p(2)) + \psi_i^{t-1}(P(1)+P(2))}$$
$$\psi\left(q_{-i}\left(h_i^{t-1}, (D, D)\right)\right) = \frac{\left(1-\psi_i^{t-1}\right)P(1) + \psi_i^{t-1}(1-P(1)-P(2))}{\left(1-\psi_i^{t-1}\right)(P(1)+P(2)) + \psi_i^{t-1}(1-P(1)-P(2))}$$

where $P(k)$ is a probability that k errors occur.

It can be shown that $\psi\left(q^t_{-i}\left(h^{t-1}_i; (D; \omega_{-i})\right)\right) \geq \frac{1}{2}$ when ε is small. See Sekiguchi [13] for detail.

With case 2 and case 3, I can conclude that $\psi(q_{-i}(h^t_i)) \geq \frac{n-1}{n}$ after any history such as

$$\begin{cases} h^t_i = \left(h^{t-1}_i = \left((C;C);\ldots;(C;C)\right); \left(C; \omega^{t-1}_i\right)\right) & \text{for } t \geq 3 \\ h^t_i = \left(C; \omega^{t-1}_i\right) & \text{for } t = 2 \end{cases}$$

with $\omega^{t-1}_i \neq C$

or

$$\begin{cases} h^t_i = \left(h^{t-1}_i; \left(D; \omega^{t-1}_i\right)\right) & \text{for } t \geq 3 \\ h^t_i = \left(D; \omega^{t-1}_i\right) & \text{for } t = 2 \end{cases}$$

$\blacksquare$

References

[1] D. Abreu, D. Pearce, and E. Stachetti, Toward a theory of discounted repeated games with imperfect monitoring, Econometrica 58 (1990), 1041-1064.

[2] V. Bhaskar, Sequential equilibria in the repeated prisoner's dilemma with private monitoring, mimeo, 1999.

[3] V. Bhaskar, The robustness of repeated game equilibria to incomplete payo¤ information, mimeo, 1999.

[4] O. Compte, Communication in repeated games with private monitoring, Econometrica 66 (1998), 597-626.

[5] G. Ellison, Cooperation in the prisoner's dilemma with anonymous random matching, J. Econ. Theory 61 (1994), 567-588.

[6] J. C. Ely and J. Välimäki, A robust folk theorem for the prisoner's dilemma, mimeo, 1999.

[7] D. Fudenberg, D. Levine, and E. Maskin, The folk theorem with imperfect public information", Econometrica 62 (1994), 997-1040.

[8] M. Kandori and H. Matsushima, Private observation, communication, and collusion, Econometrica 66 (1998), 627-652.

[9] G. Mailath and S. Morris, Repeated games with imperfect private monitoring: notes on a coordination perspective, mimeo, 1997.

[10] H. Matsushima, On the theory of repeated game with non-observable actions, part I: Anti-Folk Theorem without communication. Econ. Letters 35 (1990), 253-256.

[11] I. Obara, Private strategy and e¢ciency: repeated partnership games revisited, mimeo, 1999.

[12] M. Piccione, The repeated prisoner's dilemma with imperfect private monitoring, mimeo, 1998.

[13] T. Sekiguchi, E¢ciency in repeated prisoner's dilemma with private monitoring, J. Econ. Theory 76 (1997), 345-361.