

The Repeated Prisoners' Dilemma with Private Monitoring: a N-player case[□]

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

First version: December, 1998
This version: April, 2000

Abstract

This paper studies the repeated prisoners' 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 symmetric, and if the number of players is large. This approximate efficiency result also holds when the number of players is two for any prisoners' dilemma as long as monitoring is almost perfect. A detailed characterization of these sequential equilibria is provided.

1. Introduction

This paper examines the repeated prisoners' dilemma with 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 [8], few things are known about repeated games with private monitoring. It is shown in Compte [5] and Kandori and Matsushima [10] that a folk theorem still holds in this class of game with communication between players, but it is difficult to analyze it without communication because the simple recursive structure is lost.

The two player prisoner's dilemma was already examined in Sekiguchi [16], which is the first paper to show that the efficient outcome can be achieved in some repeated prisoner's dilemmas with almost perfect private monitoring. This paper is an extension of Sekiguchi [16] in the sense that (1): a similar grim trigger strategy is employed,

[□]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.

(2): the efficient outcome is obtained for any prisoner's dilemma with two players, (3): this efficiency result for the two player case is extended to the case of arbitrary number of players with some symmetric condition on the monitoring structure, and (4): the sequential equilibrium corresponding to this grim trigger Nash equilibrium is explicitly constructed.

In Sekiguchi [16], 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 a grim trigger strategy and a permanent defection in the first period, each 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 this paper, the clear cut characterization of the optimal action is provided as a function of this belief. In Sekiguchi [16], it is shown that a player should start defecting if she is very confident that the other player has started defecting, and a player should cooperate if she is really confident 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. This paper fills in that gap.

Through a detailed examination of players' incentive, we also find that the payoff restriction imposed in Sekiguchi [16] is not really necessary, that is, the approximate efficiency result is valid for any prisoner's dilemma game.

Although the same kind of clear characterization of the optimal action is possible with many players, it is not straightforward to extend this efficiency 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 the 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 profile only depends on the number of the total errors it contains, it is shown that the efficiency 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 refined the result of Sekiguchi [16] for the two player case. Piccione [15] also achieves the efficient outcome for any prisoners' dilemma with two players and almost perfect private monitoring. Moreover, he establishes a folk theorem for a class of prisoners' dilemma 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 infinite states. Ely and Välimäki [7] prove a folk theorem using a similar strategy, but their strategy

is "simple" in the sense that it is a two states automaton. Bhaskar [3] 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 different line of attack from Piccione [15] or Ely and Välimäki [7] using a public randomization device.

Mailath and Morris [11] is the first 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. A folk theorem can be obtained when information is almost public and almost perfect. Although the stage game in this paper has a more specific structure, the information structure allowed in this paper is not nested in their information structure. Especially, private signals can be independent over players in this paper.

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 6 concludes with a comparison to other related papers.

2. The Model

Let $N = \{1, 2, \dots, n\}$ 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 $A = \prod_{i=1}^n A_i$; A profile of all player's actions but player i 's is $a_{-i} = \prod_{j \in N, j \neq i} A_j$.

Each player receives a private signal about all the other players' actions within that period. Let $\theta_i = (\theta_{i,1}, \dots, \theta_{i,i-1}, \theta_{i,i+1}, \dots, \theta_{i,n}) \in \{C, D\}^{n-1}$ be a generic signal received by player i ; where $\theta_{i,j}$ stands for the signal which player i receives about player j 's action: A generic signal profile is denoted by $\theta = (\theta_1, \dots, \theta_n) \in \{C, D\}^n$. All players have the same payoff function u . Player i 's payoff $u(a_i, \theta_i)$ depends on her own action a_i and private signal θ_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 $p(\theta | a)$. It is assumed that $p(\theta | a)$ are full support distributions, that is, $p(\theta | a) > 0 \forall \theta$. The space of a system of full support distributions $\{p(\theta | a)\}_{a \in A}$ is denoted by P .

Since we are interested in the situation where information is almost perfect, we restrict 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 in that period with probability larger than $1 - \epsilon$ for some small number ϵ :

To sum up, the space of the information structure we mainly deal with in this

paper is the following subset of P :

$$P_{\omega} = \left\{ p(\cdot | j_a) g_{a_2 A} \mid \begin{array}{l} p(\cdot | j_a) > \frac{1}{2} \text{ if } j_i = a_i \text{ for all } i; \\ \text{and } \sum_i p(\cdot | j_a) = 1 \end{array} \right\}$$

and p_{ω} is a generic element of P_{ω} :

We also introduce the perfectly informative signal distribution $P_0 = \{p_0(\cdot | j_a) g_{a_2 A}\}$, where, for any $a \in A$; $p_0(\cdot | j_a) = 1$ if $j_i = a_i$ for all i . The whole space of the information structure P . P_0 is endowed with the Euclidean norm.

The stage game payoff only depends on the number of signals C and D a player receives. Let $d(i, j)$ be the number of "D"s contained in j ; Then, $u(a_i; j) = u(a_i; j')$ if $d(i, j) = d(i, j')$ for any a_i . Let $u^i(a_i; D^k; C^l)$ be the payoff of player i when $d(i, j) = k$. The deviation gain when k defections are observed is $M(k) = u^i(D; D^k; C^l) - u^i(C; D^k; C^l) > 0$. The payoffs $u^i(C; j(C; \dots; C))$ and $u^i(D; j(D; \dots; D))$ are normalized to 1 and 0 respectively for all i . It is assumed that $(1; \dots; 1)$ is an efficient stage game payoff. The largest deviation gain and the smallest deviation gain are \bar{M} and \underline{M} respectively, where $\bar{M} = \max_{k, n_i} M(k)$ and $\underline{M} = \min_{k, n_i} M(k)$.

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 $h_i^t = (a_i^1; j_1^1; \dots; a_i^{t-1}; j_{t-1}^{t-1})$ for $t = 2$ 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 ; Discounted average payoff is $V_i(\sigma; p; \delta) = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} E[u((a_i^t; j^t)) | \sigma_i; p]$; where the probability measure on H_i^t is generated by $(\sigma; p)$. The least upper bound and largest lower bound of the discounted average payoff are denoted by \bar{V} and \underline{V} respectively.

For this n_i player repeated prisoner's dilemma, the grim trigger σ_C and the permanent defection σ_D are defined as follows:

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

We also use σ_C or σ_D for any continuation strategy which is identical to σ_C or σ_D that is, any continuation strategy from the period $t+1$ such that $\sigma_{a_i}^i(h_i^{t+k}) = \sigma_{a_i}^i(h_i^k)$ 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 respectively¹. This grim trigger strategy is the harshest one among all the variations of the 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 fully cooperative. When player i is mixing σ_C and σ_D with probability $(q_i; 1 - q_i)$; this strategy is denoted by $q_i \sigma_C + (1 - q_i) \sigma_D$.

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

Suppose that $\frac{3}{4}_C$ or $\frac{3}{4}_D$ is chosen by all players. Let $\mu \in \{0, 1, \dots, n-1\}$ be the number of players using $\frac{3}{4}_D$ as a continuation strategy among n players. Then a probability measure $q_{i-1}(h_i^t; p)$ ² on the space $\mathcal{E} = \{0, 1, \dots, n-1\}$ 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 $\frac{3}{4}_C$ and $\frac{3}{4}_D$ by every player, but this dependence is not shown explicitly as it is obvious.

In the two player case, a player's strategy can be represented as a function of belief, using the fact that the other player is always playing either $\frac{3}{4}_C$ or $\frac{3}{4}_D$ on and off the equilibrium path. Here the space of the other player's "types" is much larger. However, there is a convenient way of summarizing information further. We classify \mathcal{E} into the two sets: $\{0\}$ and $\{1, \dots, n-1\}$; that is, the state no one has ever switched to $\frac{3}{4}_D$ and the state where there is at least one player who has already switched to $\frac{3}{4}_D$. Player i 's conditional subjective probability that no player has started using $\frac{3}{4}_D$ is denoted by $\hat{A}(q_{i-1}) = q_{i-1}(h_i^t; p)(0)$: The reason why we just focus on this number is because the exact number of players who are playing permanent defection does not make much difference in what will happen in the future given everyone's strategy. As soon as someone starts playing $\frac{3}{4}_D$, every other player starts playing $\frac{3}{4}_D$ 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 $\frac{3}{4}_D$; but whether anyone has switched to $\frac{3}{4}_D$ or not.

Let $V_i(\frac{3}{4}_i; \mu; p; \pm)$ be player i 's discounted average payoff when μ other players are playing $\frac{3}{4}_D$ and $n-1-\mu$ other players are playing $\frac{3}{4}_C$: This notation is justified under the assumption of the symmetry distribution, which is introduced in the next section. Finally, we need the following notations:

$$V_i(\frac{3}{4}_i; q_{i-1}; p; \pm) = \sum_{\mu=0}^{n-1} V_i(\frac{3}{4}_i; \mu; p; \pm) q_{i-1}(\mu)$$

$$M(q_{i-1}; p) = \sum_{\mu=0}^{n-1} [U_i(D; D^\mu; p) - U_i(C; D^\mu; p)] q_{i-1}(\mu)$$

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

²Here, q_{i-1} is a mapping from the space of private history (and the monitoring technology p) to a space of probability measures on \mathcal{E} : However, we also use q_{i-1} as a point in a $(n-1)$ dimensional simplex $\{p_i \geq 0, \sum_{i=0}^{n-1} p_i = 1\}$; where p_i means the probability of the event $\mu = i$:

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 private monitoring. In order to achieve a certain level of coordination, which is necessary for an equilibrium with trigger strategies, we impose some assumptions on $p(\omega, j)$ in addition to the assumption that it is almost perfect.

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

Assumption 1

$$p(\omega, j) = \prod_{i \in N} p_i(a_{\omega(i)} | a_{\omega(-i)}) \text{ for any permutation } \omega : N \rightarrow N$$

This assumption implies the followings. First, each player's expected payoff is the same under the same situation combined with the assumption on u : Second, only the number of C and D played matters in terms of expected payoff. Finally, the dynamics of belief for each player is the same with the same sort of private history. This assumption allows us to treat agents symmetrically and justifies all the notations introduced in Section 2.

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 $\#(\omega, j)$ stand for the number of errors in ω : The following assumption is strong enough for that purpose:

Assumption 2.

$$p(\omega^0, j) = p(\omega^{00}, j) \text{ if } \#(\omega^0, j) = \#(\omega^{00}, j) \text{ for any } \omega^0, \omega^{00} \in \Omega; a \in A$$

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 :

2 Example: Totally Decomposable Case

$$p(i|ja) = \prod_{i \in j} p(i|_i j a_j) \text{ for all } a \in A \text{ and } i \in j; \text{ where } p(i|_i j a_j) = 1 - a_j \text{ if } i = j \text{ and } a_j = 1 - a_j \text{ if } i \neq j$$

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

4. Belief Dynamics and Best Response

We represent our equilibrium strategy as a mapping from belief to actions $\sigma: \Delta \rightarrow \{C, D\}$. More explicitly, our strategy takes the following form:

$$\sigma(q_i) = \begin{cases} C & \text{if } q_i \in Q^C \\ q & \text{if } q_i \in Q^I \\ D & \text{if } q_i \in Q^D \end{cases}$$

where $Q^C; Q^I; Q^D$ are mutually exclusive and exhaustive sets in the space of the probability measures on \mathcal{E} ; and q means playing C with probability q and playing D with probability $1 - q$;

The steps to achieve the approximate efficient payoff is as follows. First, we give a almost complete characterization of the unique optimal action as a function of belief. As a next step, we analyze the dynamics of belief when players are playing either $\frac{3}{4}C$ or $\frac{3}{4}D$: Since the dynamics of belief in the n_i player case has some particular feature which does not appear in the two player case, we analyze it in detail. The third step is to construct a sequential equilibrium. We first have to find $\sigma(q_i)$ which assigns the optimal action for any possible belief on and off the equilibrium path. For this purpose, we strengthen the characterization result in the first step to get this function $\sigma(q_i)$: In order to prove that $\sigma(q_i)$ is a sequential equilibrium, we also have to check consistency, that is, to check if players are actually playing either $\frac{3}{4}C$ or $\frac{3}{4}D$ on the equilibrium path by following $\sigma(q_i)$. Since this sequential equilibrium is constructed for a discount factor in the middle range, as the final step, we divide the original repeated game to component repeated games or use a public randomization device to implement the same payoff by reducing high δ effectively.

4.1. Belief and the Best Response

Before analyzing the unique optimal action, we first extend one property which obviously holds in the two player case to the n_i player case. In the two player case, the difference in payoffs by $\frac{3}{4}C$ and $\frac{3}{4}D$ is linear and there is a unique $q^*(\delta; p_0)$ where a player is indifferent between $\frac{3}{4}C$ and $\frac{3}{4}D$ with perfect monitoring if δ is high enough

as shown in Figure 1. When private monitoring is almost perfect, there is $q^*(\pm; p^*)$ where a player is indifferent between $\frac{3}{4}_C$ and $\frac{3}{4}_D$ and $q^*(\pm; p^*) \rightarrow q^*(\pm; p_0)$ as $\epsilon \rightarrow 0$.

Put Figure 1 here.

When the number of players is more than two; the corresponding object $V_i(\frac{3}{4}_C; q_{-i}; p_0; \pm)$ $V_i(\frac{3}{4}_D; q_{-i}; p_0; \pm)$ for the n_i player case is a slightly more complex object. Even when players randomize independently and symmetrically playing C with probability q and D with probability $(1 - q)$, that is, $q_{-i}(\mu) = \sum_{\mu=0}^{n_i-1} (1 - q)^\mu q^{n_i-1-\mu} \mu^{i_{n_i}-1} \binom{n_i-1}{\mu}$ for $\mu = 0, \dots, n_i - 1$, it is a $n_i - 1$ degree polynomial in $q \in [0, 1]$. Potentially, this equation may have $n_i - 1$ solutions between 0 and 1 as shown in Figure 2. We define $q^*(\pm; p^*)$ to be the solution which is closest to 0 for $V_i(\frac{3}{4}_C; q_{-i}; p_0; \pm) - V_i(\frac{3}{4}_D; q_{-i}; p_0; \pm) = 0$ where $q_{-i}(\mu) = \sum_{\mu=0}^{n_i-1} (1 - q)^\mu q^{n_i-1-\mu} \mu^{i_{n_i}-1} \binom{n_i-1}{\mu}$.

Put Figure 2 here.

When monitoring is almost perfect, $V_i(\frac{3}{4}_C; q_{-i}; p^*; \pm) - V_i(\frac{3}{4}_D; q_{-i}; p^*; \pm)$ is very close to $V_i(\frac{3}{4}_C; q_{-i}; p_0; \pm) - V_i(\frac{3}{4}_D; q_{-i}; p_0; \pm)$. Actually, it is easy to confirm that the former converges to the latter uniformly in q as $\epsilon \rightarrow 0$.³ If $\pm > \frac{M(0)}{1+M(0)}$; then $V_i(\frac{3}{4}_C; 0; p_0; \pm) - V_i(\frac{3}{4}_D; 0; p_0; \pm) > 0$; which implies that there exists $q^*(\pm; p_0)$ between 0 and 1. The following lemma extends a useful property for the two player case to the n_i player case.

Lemma 1. $q^*(\pm; p_0) \rightarrow 1$ as $\pm \rightarrow \frac{M(0)}{1+M(0)}$

Proof. See Appendix

Now we show that the optimal action as a function of q_{-i} is still 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 $\frac{3}{4}_D$ is still optimal if a player knows that someone has switched to the permanent defection and ϵ is small.

Lemma 2. There exists a $\delta > 0$ such that $V_i(\frac{3}{4}_i; q_{-i}; p^*; \pm)$ is maximized by $\frac{3}{4}_D$ for any p^* ; if $q_{-i}(\mu) = 1$ for any $\mu \in \mathbb{N}$.

³Also note that convergence of $V_i(\frac{3}{4}_i; q_j; p^*)$ to $V_i(\frac{3}{4}_i; q_j; p_0)$ is independent of the choice of associated sequence $\{p_n\}$ because of the definition of P^* .

Proof.

Take $\frac{3}{4}_D$ and any strategy which starts with C. The least deviation gain is $(1 - \beta) \frac{4}{3}$. The largest loss caused by the difference in continuation payoffs with $\frac{3}{4}_D$ and the latter strategy is $\beta \bar{V}$. Setting β small enough guarantees that $(1 - \beta) \frac{4}{3} > \beta \bar{V}$ for any $\beta \in (0, \beta)$. Then, D must be the optimal action for any such β . Since players are using permanent defection, $q_i(\mu) = 1$ for some $\mu \in [0, 1]$ in the next period. This implies that D is the optimal action in all the following periods. \square

Using $p(\theta)$ and given the fact that all players are playing either $\frac{3}{4}_C$ or $\frac{3}{4}_D$; we introduce a notation for the transition probability of the number of players who have switched to $\frac{3}{4}_D$: Let $\frac{1}{4}(l, m)$ be a probability that l players will play $\frac{3}{4}_D$ from the next period when m players are playing $\frac{3}{4}_D$ now. In other words, this $\frac{1}{4}(l, m)$ 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, $\frac{1}{4}(l, m) > 0$ if $l = m$ and $\frac{1}{4}(l, m) = 0$ if $l < m$. The following lemma, which is a sort of Maximum theorem, provides various informative and useful bounds on the variations of discounted average payoffs caused by introducing small imperfectness in private monitoring.

Lemma 3.

- $\inf_{p \in P} V_i(\frac{3}{4}_C; 0; p; \beta) = \frac{(1 - \beta) \frac{4}{3} + \beta \bar{V}}{1 - \beta(1 - \mu)}$
- Given $\beta \in (0, \frac{M(0)}{1 + M(0)}; 1$; There exists a $\epsilon > 0$ such that for any $\beta \in [0, \epsilon]$:

$$\sup_{\frac{3}{4}_i; p \in P} V_i(\frac{3}{4}_i; 0; p; \beta) \leq \frac{1 - \beta + \beta \bar{V}}{1 - \beta(1 - \mu)}$$

Proof.

(1): For any $\beta \in (0, 1)$ and $p \in P$;

$$V_i(\frac{3}{4}_C; 0; p; \beta) = (1 - \beta) \frac{4}{3} + \beta \sum_{j=0}^m \frac{1}{4}(j, 0) V_i(\frac{3}{4}_C; 0; p; \beta) + \beta \sum_{j=1}^m \frac{1}{4}(j, 0) \bar{V}$$

So,

$$V_i(\frac{3}{4}_C; 0; p; \beta) = \frac{(1 - \beta) \frac{4}{3} + \beta \sum_{j=1}^m \frac{1}{4}(j, 0) \bar{V}}{1 - \beta \sum_{j=0}^m \frac{1}{4}(j, 0)} = \frac{(1 - \beta) \frac{4}{3} + \beta \bar{V}}{1 - \beta(1 - \mu)}$$

(2): Given $\beta \in (0, \frac{M(0)}{1 + M(0)}; 1$; it is easy to check that $V_i(\frac{3}{4}_C; 0; p_0; \beta) > V_i(\frac{3}{4}_D; 0; p_0; \beta)$:

Pick ϵ small enough such that (i) $V_i(\frac{3}{4}_C; 0; p; \beta) > V_i(\frac{3}{4}_D; 0; p; \beta)$ for any p and (ii) $\epsilon < \beta$: Let $\frac{3}{4}_0$ be the optimal strategy given that everyone is using $\frac{3}{4}_C$:⁴ Suppose that $\frac{3}{4}_0$ assigns D for the first period. Then for any $\beta \in [0, \epsilon]$;

$$V_i(\frac{3}{4}_0; 0; p) \leq \frac{1}{2} (1 - \beta) U_i(D; D^0; p) + \frac{1}{2} \sum_{\mu=2}^m \frac{1}{4} (1 - \beta) V_i(\frac{3}{4}_0; 0; p; \beta) + \frac{1}{2} \sum_{\mu=2}^m \frac{1}{4} (\mu - 1) V_i(\frac{3}{4}_D; \mu - 1; p; \beta)$$

⁴This $\frac{3}{4}_0$ exists because the strategy space is a compact space in product topology, on which discounted average payoff functions are continuous. Of course, this $\frac{3}{4}_0$ depends on the choice of p :

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 $\frac{3}{4}_0^B$ leads to a higher expected payoff than $V_i(\frac{3}{4}_0^B; 0; p^{**}; \pm)$ or $V_i(\frac{3}{4}_D; \mu_i - 1; p^{**}; \pm)$ at the corresponding states, then this contradicts the optimality of $\frac{3}{4}_0^B$ or $\frac{3}{4}_D$ by Lemma 2. So this inequality holds.

Then, for any $\epsilon \in [0; \epsilon^*]$:

$$\begin{aligned} V_i(\frac{3}{4}_0^B; 0; p^{**}; \pm) &\leq \frac{(1 - \epsilon) U(D; D_0; p^{**}) + \sum_{\mu=2}^P \frac{1}{4} (k_j - 1) V_i(\frac{3}{4}_D; \mu_i - 1; p^{**}; \pm)}{1 - \epsilon \frac{1}{4} (1 - j)} \\ &= V_i(\frac{3}{4}_D; 0; p^{**}; \pm) \\ &< V_i(\frac{3}{4}_C; 0; p^{**}; \pm) \end{aligned}$$

Since this contradicts the optimality of $\frac{3}{4}_0^B$, $\frac{3}{4}_0^B$ has to assign C for the first period. Now,

$$V_i(\frac{3}{4}_0^B; 0; p^{**}; \pm) \leq (1 - \epsilon) + \epsilon \frac{1}{4} (0 - j) V_i(\frac{3}{4}_0^B; 0; p^{**}; \pm) + \epsilon (1 - \frac{1}{4} (0 - j)) \bar{V}$$

So,

$$V_i(\frac{3}{4}_0^B; 0; p^{**}; \pm) \leq \frac{(1 - \epsilon) + \epsilon (1 - \frac{1}{4} (0 - j)) \bar{V}}{1 - \epsilon \frac{1}{4} (0 - j)} \leq \frac{(1 - \epsilon) + \epsilon \bar{V}}{1 - \epsilon (1 - \epsilon)}$$

This implies that $\sup_{\frac{3}{4}_i; p^{**} \in P^{**}} V_i(\frac{3}{4}_i; 0; p^{**}; \pm) \leq \frac{1 - \epsilon + \epsilon \bar{V}}{1 - \epsilon (1 - \epsilon)}$ for any $\epsilon \in [0; \epsilon^*]$: \forall

(1) means that a small departure from the perfect monitoring does not reduce the payoff of $\frac{3}{4}_C$ 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 $\frac{3}{4}_C$ with a small imperfection in the private signal as long as all the other players are using a grim trigger strategy.

The following result 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 ϵ ; for any $\delta > 0$; there exists a $\epsilon^* > 0$ such that for any p^{**} :

2 it is not optimal to play C for player i if q_i satisfies $\hat{A}(q_i) \leq \frac{1 - \epsilon}{\epsilon} M(q_i; p_0)_i$

2 it is not optimal to play D for player i if q_i satisfies $\hat{A}(q_i) = \frac{1 - \epsilon}{\epsilon} M(q_i; p_0) +$

Proof:

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

$$(1 - \epsilon) M(q_i; p) \geq \frac{1}{2} \sup_{\frac{3}{4}} V_i(\frac{3}{4}; 0; p; \epsilon) + \frac{3}{4} \bar{V} + (1 - \epsilon) \hat{A}(q_i) \bar{V}^\epsilon$$

By Lemma 3.2., this inequality is satisfied for any $\epsilon \in [0, \epsilon^*]$ and any p if

$$(1 - \epsilon) M(q_i; p) \geq \frac{1}{2} \frac{1 - \epsilon + \epsilon \bar{V}}{1 - \epsilon(1 - \epsilon)} + \frac{3}{4} \bar{V} + (1 - \epsilon) \hat{A}(q_i) \bar{V}^\epsilon$$

LHS converges to $(1 - \epsilon) M(q_i; p_0)$ and RHS converges to $\hat{A}(q_i)$ as $\epsilon \rightarrow 0$: So, if q_i satisfies $\hat{A}(q_i) \leq \frac{1 - \epsilon}{\epsilon} M(q_i; p_0) + \epsilon$ for any $\epsilon > 0$; then there exists a $\epsilon^0(\epsilon; q_i) \in (0, \epsilon^*)$ and a neighborhood $B(q_i)$ of q_i such that C is not optimal for any $p \in P^{\epsilon^0(\epsilon; q_i)}$ and any $q_i^0 \in B(q_i)$: This $\epsilon^0(\epsilon; q_i) > 0$ 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

$$(1 - \epsilon) M(q_i; p) < \frac{1}{2} \hat{A}(q_i) f(1 - \epsilon) V_i(\frac{3}{4}; 0; p; \epsilon) + \frac{3}{4} \underline{V} + (1 - \epsilon) \hat{A}(q_i) \underline{V} \bar{V}^\epsilon$$

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

$$(1 - \epsilon) M(q_i; p) < \frac{1}{2} \hat{A}(q_i) \frac{1 - \epsilon + \epsilon \underline{V}}{1 - \epsilon(1 - \epsilon)} + \frac{3}{4} \underline{V} + (1 - \epsilon) \hat{A}(q_i) \underline{V} \bar{V}^\epsilon$$

This inequality converges to $\hat{A}(q_i) = \frac{1 - \epsilon}{\epsilon} M(q_i; p_0)$ as $\epsilon \rightarrow 0$: So, if q_i satisfies $\hat{A}(q_i) = \frac{1 - \epsilon}{\epsilon} M(q_i; p_0) + \epsilon$ for any $\epsilon > 0$; there exists a $\epsilon^0(\epsilon; q_i)$ such that D is not optimal for any $p \in P^{\epsilon^0(\epsilon; q_i)}$ and any q_i^0 around q_i : Again, $\epsilon^0(\epsilon; q_i)$ can be set independent of q_i :

Finally, setting $\epsilon^*(\epsilon) = \min \{ \epsilon^0(\epsilon); \epsilon^0(\epsilon) \}$ completes the proof. \square

This proposition implies that the optimal action can be completely characterized except for an arbitrary small neighborhood of the manifold satisfying $\hat{A}(q_i) = \frac{1 - \epsilon}{\epsilon} M(q_i; p_0)$ in a $n_i - 1$ dimensional simplex; where player i is indifferent between $\frac{3}{4}_C$ and $\frac{3}{4}_D$ with perfect monitoring: Note that this unique optimal action is equivalent to the one with perfect monitoring. Since the continuation payoffs with private monitoring converge to the continuation payoffs with perfect monitoring as monitoring gets accurate, a player prefers C(D) to D(C) in private monitoring if and only if she prefers

C (D) to D (C) in perfect monitoring except for the region where she is almost indifferent between C and D with perfect monitoring. A slight noise still matters in this region, but we also characterize the best response for this region later.

Although this argument is essentially the path dominant argument used in Sekiguchi [16] for $n = 2$, it extends that argument to the n_i player case and provides a sharper characterization even for $n = 2$: In fact, this proposition is the reason why an almost efficient outcome is achieved 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 \bar{A} is sufficiently close to 1, $\epsilon > \frac{M(0)}{1+M(0)}$; and μ is small:

Corollary 1. Given $\epsilon > \frac{M(0)}{1+M(0)}$; there exists $\bar{A} > 0$ and $\mu > 0$ such that for any p ; it is not optimal for player i to play D if $\bar{A} \geq \frac{\bar{A}}{1 - \mu}$:

4.2. Belief Dynamics

Since the optimal action is almost characterized as a function of $q_{i,i}$; we have to analyze the dynamics of $q_{i,i}$ associated with grim trigger strategies. Given the optimal action shown above, all we have to make sure for the consistency of grim trigger strategies is that $q_{i,i}$ stays in the "C area" described by Proposition 1 as long as player i has observed fully cooperative signals from the beginning and $q_{i,i}$ 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 a 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 $\frac{3}{4}_D$ rather than as an error if μ is small⁵. 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 no one has switched to $\frac{3}{4}_D$ is at most $\frac{1}{3}$. Obviously, this flexibility of interpretation increases as the number of players increases, which makes it easier to move \bar{A} closer to 0 after this sort of history. Note that this upper bound of \bar{A} does not depend on the level of μ :

Second, consider a 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:

⁵We need to let players to randomize between $\frac{3}{4}_C$ and $\frac{3}{4}_D$ in the initial period for this argument to work in the initial period. If players start with, say, $\frac{3}{4}_C$ with probability 1, no learning occurs after the initial period. This is first observed by Matsushima [12]. Note that this is the only reason why the initial randomization is needed. The rest of arguments does not depend on this initial randomization.

For everyone to be still cooperative after the current period, all players but player 1 should observe 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 $\frac{3}{4}D$: 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 at least one player has switched to $\frac{3}{4}D$ 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 \hat{A} is strictly less than $\frac{1}{6}$: This argument is again independent of the level of ϵ .⁶

The above arguments strongly suggest that once players start playing $\frac{3}{4}D$; then they continue to do so. The following proposition summarizes the above arguments and also provide another useful property to support $\frac{3}{4}C$: Let $W_k = \{q_i, j | \hat{A}(q_i, j) = \frac{1-\epsilon}{\epsilon} M(q_i, j; p_0) \geq k\}$; where $k > 0$ and $0 < \frac{1-\epsilon}{\epsilon} M(n-1, j) \leq k$: When k is small, W_k just covers the region where C is the unique optimal action and the unique optimal action is indeterminate. We show that $\hat{A}_t > \bar{A}$ if a player plays C , observes C ; and $\hat{A}_{t-1} \geq W_k$.

Proposition 2. Suppose that every player plays $q\frac{3}{4}C + (1-q)\frac{3}{4}D$ 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 $t = 2, 3, \dots$:

2 For any $\hat{A}^0 > 0; k > 0$; there exists ϵ such that for any $\epsilon \in (0, \epsilon)$
 $\hat{A}^i(q_i, j | h_i^{t+n}) = \hat{A}^0$ for $h_i^{t+n} = (h_i^{t-1}, (C; C); \dots; (C; C))$ when $\hat{A}^i(q_i, j | h_i^{t-1}) \geq W_k$

2 $\hat{A}(q_i, j | h_i^t) \geq \frac{1}{n}$ after histories such as

$$\begin{aligned} & \frac{1}{2} h_i^t = (h_i^{t-1}, C; !_{i^{t-1}}) \text{ for } t = 3 \\ & - h_i^t = (C; !_{i^{t-1}}) \text{ for } t = 2 \\ & \text{with } !_{i^{t-1}} \in C \end{aligned}$$

$$\begin{aligned} & \text{or} \\ & \frac{1}{2} h_i^t = (h_i^{t-1}, D; !_{i^{t-1}}) \text{ for } t = 3 \\ & - h_i^t = (D; !_{i^{t-1}}) \text{ for } t = 2 \end{aligned}$$

Proof. See appendix.

⁶This last argument is specific to the $n = 3$ player case. The two player case has to be treated separately. See Sekiguchi [16] for that case.

5. Sequential Equilibrium with the "Grim Trigger" Strategy

As in section 2, the equilibrium strategy is represented as a mapping from the space of belief Q to $\{C, D\}$. The following notations are useful:

$$\begin{aligned} Q^C &= \{q_i \mid V(\frac{3}{4}C; q_i; p) > V(\frac{3}{4}D; q_i; p)\} \\ Q^I &= \{q_i \mid V(\frac{3}{4}C; q_i; p) = V(\frac{3}{4}D; q_i; p)\} \\ Q^D &= \{q_i \mid V(\frac{3}{4}C; q_i; p) < V(\frac{3}{4}D; q_i; p)\} \\ Q^n &= \{q_i \mid \Delta(q_i) \leq \frac{1}{n}\} \end{aligned}$$

These are subsets of the $n-1$ dimensional simplex on E . The subset Q^n is a set containing the absorbing set of the dynamics of q_i under the grim trigger. Q^I is a manifold where player i is indifferent between $\frac{3}{4}C$ and $\frac{3}{4}D$. In particular, $q_i^*(p) \in Q^I$ by definition. Note that Q^C converges to $Q_0^C = \{q_i \mid \Delta(q_i) > \frac{1-\epsilon}{\epsilon} \Delta(q_i; p_0)\}$ and Q^D converges to $Q_0^D = \{q_i \mid \Delta(q_i) < \frac{1-\epsilon}{\epsilon} \Delta(q_i; p_0)\}$ as $\epsilon \rightarrow 0$.

Now define $\frac{1}{2}^\epsilon$ as a mapping from $q_i \in Q$ to $\{C, D\}$ in the following way:

$$\frac{1}{2}^\epsilon(q_i) = \begin{cases} C & \text{if } q_i \in Q^C \\ q^*(p) & \text{if } q_i \in Q^I \\ D & \text{if } q_i \in Q^D \end{cases}$$

We know from Proposition 1 that this function assigns the best response action almost everywhere except for a neighborhood of Q^I when ϵ is small. Thanks to the last proposition about the dynamics of belief, now we can verify that $\frac{1}{2}^\epsilon(q_i)$ actually assigns the unique optimal action for any belief $q_i \notin Q^I$ with one weak assumption.

Proposition 3. Given ϵ ; if $Q^n \subset Q_0^D$ and ϵ is small enough, then

$\frac{1}{2}^\epsilon$ is the unique optimal action if and only if $q_i \in Q^C \setminus Q^I$

$\frac{1}{2}^\epsilon$ is the unique optimal action if and only if $q_i \in Q^D \setminus Q^I$

Proof:

Fix $\epsilon > 0$ in Proposition 1 and set $k > 0$ slightly larger than $\epsilon > 0$: Now take any q_i such that $\Delta(q_i) > \frac{1-\epsilon}{\epsilon} \Delta(q_i; p_0) + k$ and $q_i \in Q^D$: We prove that D is the unique optimal action in this region. If a player plays C , then Proposition 2 implies that the continuation strategy is $\frac{3}{4}C$ if ϵ is sufficiently small: Since $\frac{3}{4}C$ is dominated by $\frac{3}{4}D$ in this region, the unique optimal action should be D : Similarly, take any q_i such that $\frac{1-\epsilon}{\epsilon} \Delta(q_i; p_0) + \epsilon > \Delta(q_i)$ and $q_i \in Q^C$: If D is played, then Proposition 2 and the assumption $Q^n \subset Q_0^D$ implies that the continuation strategy is $\frac{3}{4}D$ as long as ϵ is small enough. Since $\frac{3}{4}D$ is dominated by $\frac{3}{4}C$ in this region, the unique optimal action should be C :

Since any other $q_i \in Q^D$ satisfies $\Delta(q_i) \leq \frac{1-\epsilon}{\epsilon} \Delta(q_i; p_0) + k$ and any other $q_i \in Q^C$ satisfies $\Delta(q_i) = \frac{1-\epsilon}{\epsilon} \Delta(q_i; p_0) + \epsilon$ satisfies $q_i \in Q^C$ for small $\epsilon > 0$; the proof is complete. \square

The last thing we have to make sure is that players are actually playing either $\frac{3}{4}C$ or $\frac{3}{4}D$ on the equilibrium path after they initially randomize between these strategies. It is almost obvious from the arguments we used to prove the propositions.

Proposition 4. Suppose that (i) : Assumption 1 is satisfied and $n = 2$, or (ii) : Assumption 2 is satisfied. Given $\pm \geq \frac{M(0)}{1+M(0)}; 1$; if $Q^n \frac{1}{2} Q_0^D$; then there is a $\epsilon > 0$ such that for any p^n ; The following proposition shows that $\frac{1}{2} (q_{i,i})$ is a symmetric sequential equilibrium, which generates the same outcome distribution as $(\dots; q^n(\pm; p^n) \frac{3}{4}C + (1 - q^n(\pm; p^n)) \frac{3}{4}D; \dots)$:

Proof. $\exists \epsilon$

Since $\pm \geq \frac{M(0)}{1+M(0)}; 1$; $q^n(\pm; p^n)$ exists in $(0, 1)$ if ϵ is small enough. $Q^n \frac{1}{2} Q_0^D$ is also satisfied for small ϵ :

Suppose that player i chooses $\frac{3}{4}C$ as her strategy in the first period. Set $\epsilon > 0$ small enough for Corollary 1 and Proposition 2 to hold. As long as $(C; C)$ is the outcome, $A > \bar{A}$ and C is the unique optimal action by Corollary 1 and Proposition 2.

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, $\hat{A}_i^t = \hat{A}(q_{i,i}(h_i^t))$ is going to stay in Q^n forever by Proposition 2 as long as she continues playing D : Since $Q^n \frac{1}{2} Q_0^D$, this continuation strategy is clearly $\frac{3}{4}D$:

We also know that $\frac{1}{2} (q_{i,i})$ assigns the best response after any of the equilibrium history. Moreover, consistency is obvious.

Taking ϵ small such that all the above arguments go through, we conclude that $\frac{1}{2} (q_{i,i})$ is a symmetric sequential equilibrium, which generates the same outcome distribution as

$(\dots; q^n(\pm; p^n) \frac{3}{4}C + (1 - q^n(\pm; p^n)) \frac{3}{4}D; \dots)$ by construction. \forall

Since the probability that everyone chooses $\frac{3}{4}C$ in this sequential equilibrium; $q^n(\pm; p^n)^{n-1}$ gets closer to 1 as \pm gets closer to $\frac{M(0)}{1+M(0)}$ by Lemma 1, an outcome arbitrary close to the efficient outcome can be achieved for \pm arbitrary close to $\frac{M(0)}{1+M(0)}$. For high \pm ; we can use a public randomization device again to reduce \pm effectively. It is also possible to use Ellison's trick as in Ellison [6] to achieve an almost efficient outcome although the strategy is more complex and no longer a grim trigger. Here is the corollary of Proposition 3 regarding 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 $Q^n \frac{1}{2} Q_0^D$; then there is a $\epsilon > 0$ such that for any p^n ; there exists a sequential equilibrium whose symmetric equilibrium payoff is more than $1 - k$:

Finally, we have to check if the assumption $Q^n \frac{1}{2} Q_0^D$ used above to prove propositions is not so strong. First of all, this is always satisfied when $n = 2$ for a range of \pm with $q^n(\pm; p_0) \geq \frac{1}{2}; 1$; that is, when \pm is between $\frac{M(0)}{1+M(0)}$ and $\frac{M(0)+M(1)}{1+M(0)+M(1)}$. The following proposition provides sufficient conditions for $Q^n \frac{1}{2} Q_0^D$ to hold for general n .

Proposition 5. :

- 2 If $M(k) = M$ for $k = 1, \dots, n_i - 1$, then $Q^n \geq Q_0^D$ and $Q_0^I \leq \frac{M}{1+M} + \frac{nM}{1+nM}$; for ± 2
- 2 Regarding n as a parameter, take a sequence of the stage game with $n = 2; 3; \dots$:
 If there exists a lower bound $\underline{M} > 0$ such that $\min_{k=1, \dots, n_i-1} M(k) = \underline{M}$ independent of n ; then there exists \underline{n} such that for all $n = \underline{n}$: $Q^n \geq Q_0^D(n)$ and $Q_0^I(n) \leq \frac{M(0)}{1+M(0)} + \frac{nM}{1+nM}$ for ± 2

Proof: If the deviation gain is constant, $Q_0^I = q_i \Delta(q_i) = \frac{1-i}{\pm} 4$ So, $\frac{1}{n} < \frac{1-i}{\pm} 4$ is $Q^n \geq Q_0^D(n)$: Combining this inequality with $\frac{M}{1+M} < \pm$ for $Q_0^I \leq \frac{M}{1+M} + \frac{nM}{1+nM}$ is obtained:

If \underline{M} is independent of n ; $Q^n \geq Q_0^D(n)$ if $\frac{1}{n} < \frac{1-i}{\pm} \underline{M}$: So, $Q^n \geq Q_0^D(n)$ and $Q_0^I(n) \leq \frac{M(0)}{1+M(0)} + \frac{nM}{1+nM}$ for all $n = \underline{n}$ if \underline{n} is chosen such that $\frac{M(0)}{1+M(0)} < \frac{nM}{1+nM}$

The assumption of constant deviation gain is natural in the team production, where a player's cost comes from disutility associated with the effort level.

6. Conclusion

In this paper, we 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 have been two lines of research exploring the sustainability of the efficient outcome or the possibility of any folk theorem in repeated prisoners' dilemma with private monitoring. One direction of research is based on grim trigger strategies. Bhaskar [3] and Sekiguchi [16] belong to this literature as well as this paper. The emphasis of these papers are on coordination of players' actions and beliefs. As shown in Mailath and Morris [11], the almost common knowledge of continuation strategies is very important for subgame perfect equilibria or perfect public equilibria to be robust to the perturbation with respect to private noise because it enables players to keep coordinating with others without public signals. In this paper, in fact a player strongly believes that the other players are playing $\frac{3}{4}C$ when she continues playing C; and believes that the other players are playing $\frac{3}{4}D$ once she started playing D: However, note that our strategy does not exhibit such a strong coordination as required in Mailath and Morris [11]. While Mailath and Morris [11] requires that players' continuation strategies are almost common knowledge after any history, a discoordination of belief or continuation strategy can occur almost surely for our

mixing grim trigger strategy. When a player observes D for the first time, she is still uncertain about the other player's continuation strategies.

The other direction of research is based on a complete mixed strategies which makes the other player indifferent over many strategies. Ely and Välimäki [7] and Piccione [15] are among papers in this direction.⁷

One advantage of the first approach compared to the second approach is:

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

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

2. Since our strategy is almost pure, it is very easy to justify the use of a mixed strategy. Especially, our strategy satisfies the refinement criterion for repeated game equilibria with respect to incomplete information on players' payoffs, which is proposed by Bhaskar[4]. This criterion prohibits players to play different continuation strategies after different histories if players are indifferent between these continuation strategies after all these different histories. Strategies in Ely and Välimäki [7] or Piccione [15] violate this criterion at the second period when private signals are independent. While purification is straightforward for our strategy by introducing a small amount of uncertainty into stage game payoffs, payoff uncertainty has to depend on a private history in a particular way to purify equilibria in Ely and Välimäki [7] or Piccione [15]. It is also easy to adopt Nash's population interpretation to purify our equilibrium. What we have in our mind is a pool of players who are matching with the other players to play repeated games, where most of the players use grim trigger and only a small portion of the players use permanent defection.

The case where private monitoring is not almost perfect has not been systematically investigated and it remains as a important topic for future research.

⁷Obara [14] uses the same kind of strategy for repeated partnership games with imperfect public monitoring and constructs a sequential equilibrium which cannot be supported by public perfect equilibria.

Appendix.

Proof of Lemma 1.

When $\pm = \frac{4(0)}{1+4(0)}$; $q^\pm(\pm; p_0) = 1$ is the solution of the equation in q :

$$V_i(\frac{3}{4}C; q_i; \pm; p_0; \pm) - V_i(\frac{3}{4}D; q_i; \pm; p_0; \pm) = 0$$

where $q_i(\mu) = \prod_{\mu=0}^{n_i-1} (1 - q)^\mu q^{n_i-1-\mu} i^{n_i-1-\mu}$ for $k\mu = 0; \dots; n_i - 1$:

We just need to show that $\frac{\partial q^\pm(\pm; p_0)}{\partial \pm} \Big|_{\pm = \frac{4(0)}{1+4(0)}} < 0$ using the implicit function theorem. Since

$$\begin{aligned} & V_i(\frac{3}{4}C; q_i; \pm; p_0; \pm) - V_i(\frac{3}{4}D; q_i; \pm; p_0; \pm) \\ &= \prod_{\mu=0}^{n_i-1} (1 - q)^\mu q^{n_i-1-\mu} i^{n_i-1-\mu} [4(\mu; p_0)_i \pm q^{n_i-1} \\ & \frac{\partial q^\pm(\pm; p_0)}{\partial \pm} \Big|_{\pm = \frac{4(0)}{1+4(0)}} = i \frac{\frac{\partial V_i(\frac{3}{4}C; q_i^\pm; p_0; \pm)}{\partial \pm} - \frac{\partial V_i(\frac{3}{4}D; q_i^\pm; p_0; \pm)}{\partial \pm}}{\frac{\partial V_i(\frac{3}{4}C; q_i^\pm; p_0; \pm)}{\partial q} - \frac{\partial V_i(\frac{3}{4}D; q_i^\pm; p_0; \pm)}{\partial q}} \Big|_{\pm = \frac{4(0)}{1+4(0)}} \\ &= i \frac{1 + 4(0)}{(1 - \pm)(n_i - 1)4(1) + \pm(n_i - 1)} \Big|_{\pm = \frac{4(0)}{1+4(0)}} \\ &= i \frac{1}{n_i - 1} \frac{1}{4(0) + 4(1)} > 0 \end{aligned}$$

¥

Proof of Proposition 2

case 1:

Let $h_i^\dagger = i h_i^{t_i-1}; (C; C)^\dagger$ such that $\hat{A}_i^{t_i-1} = \hat{A}_i q_i i h_i^{t_i-1} \in W_k$:

Applying Bayes' rule⁸,

$$\begin{aligned} \hat{A}_i^\dagger &= \hat{A}(q_i; (h_i^\dagger)) \\ &= \frac{\hat{A}_i^{t_i-1} P(I_i^{t_i-1} = C_j | h_i^{t_i-1})}{(1 - \hat{A}_i^{t_i-1}) P(I_i^{t_i-1} = C_j | \mu_{t_i-1} \neq 0; h_i^{t_i-1}) + \hat{A}_i^{t_i-1} P(I_i^{t_i-1} = C_j | h_i^{t_i-1})} \end{aligned}$$

This function is increasing in $\hat{A}_i^{t_i-1}$ and crosses 45° line once. Note that this function is bounded below by $\hat{A}_i^{t_i-1} = \frac{\hat{A}_i^{t_i-1}(1 - \hat{A}_i^{t_i-1})}{(1 - \hat{A}_i^{t_i-1}) + \hat{A}_i^{t_i-1}}$: Let \hat{A} be the unique fixed point of this mapping. Given that $\hat{A}_i^\dagger = \hat{A}_i q_i i h_i^{t_i-1} \in W_k$; it is easy to see that

⁸All the conditional distributions implicitly depend on the level of initial mixture between $\frac{3}{4}C$; $\frac{3}{4}D$:

$\hat{A}_i^{t_i-1}$ and \hat{A} can be made larger than $\hat{A}^0 > 0$ for any q_i uniformly by choosing " ϵ " small enough. If $\hat{A}_i^{t_i-1} < \hat{A}$; then, as long as players continue to observe C; $\hat{A}_i^{t_i-1}$ is going to increase monotonically to \hat{A} . On the other hand, since $\hat{A}_i^t = \hat{A}_i^{t_i-1}$ and $\hat{A}_i^{t+n} = \hat{A}_i^{t+n_i-1}$; $\hat{A}_i^{t+n_i-1}$ is larger than $\hat{A}_i^{t_i-1}$ for any $n = 0; 1; \dots$. If $\hat{A}_i^{t_i-1} = \hat{A}$; then $\hat{A}_i^{t+n_i-1} = \hat{A} > \hat{A}^0$ for any n . This implies that $\hat{A}_i^{t+n_i-1}$ is always above \hat{A}^0 :

$$\begin{aligned}
 \text{case 2: } \frac{1}{2} \quad h_i^t &= \frac{1}{2} h_i^{t_i-1} + \frac{1}{2} C; \hat{A}_i^{t_i-1} & \text{for } t = 3 \\
 h_i^t &= C; \hat{A}_i^{t_i-1} & \text{for } t = 2 \\
 \text{with } \hat{A}_i^{t_i-1} &= \hat{A} \notin C
 \end{aligned}$$

Suppose that $t = 3$: By Bayes' Rule,

$$\begin{aligned}
 \hat{A}_i^t &= \hat{A}(q_i | (h_i^t)) \\
 &= \frac{\hat{A}_i^{t_i-2} P((\hat{A}_i^{t_i-2}; \hat{A}_i^{t_i-1}; \hat{A}_i^{t_i-1}) = (C; C; \hat{A}_i^0) | \mu_{t_i-2} = 0)}{(\hat{A}_i^{t_i-2}) P((\hat{A}_i^{t_i-2}; \hat{A}_i^{t_i-1}) = (C; \hat{A}_i^0) | \mu_{t_i-2} = 0; h_i^{t_i-2}) + \hat{A}_i^{t_i-2} P((\hat{A}_i^{t_i-2}; \hat{A}_i^{t_i-1}) = (C; \hat{A}_i^0) | \mu_{t_i-2} = 0)}
 \end{aligned}$$

This is bounded above by

$$\begin{aligned}
 &\frac{P(\hat{A}_i^{t_i-2}; \hat{A}_i^{t_i-1}; \hat{A}_i^{t_i-1}) = (C; C; \hat{A}_i^0) | \mu_{t_i-2} = 0}{P(\hat{A}_i^{t_i-2}; \hat{A}_i^{t_i-1}) = (C; \hat{A}_i^0) | \mu_{t_i-2} = 0} \\
 &\leq \frac{P(\#(\hat{A}_i^0 | C))}{P(\#(\hat{A}_i^0 | C)) + (n_i - 1) \#(\hat{A}_i^0 | C) P(\#(\hat{A}_i^0 | C))} \\
 &= \frac{1}{1 + (n_i - 1) \#(\hat{A}_i^0 | C)} \\
 &\leq \frac{1}{n}
 \end{aligned}$$

where $P(\#(\hat{A}_i^0 | C))$ is the probability of the event that $\#(\hat{A}_i^0 | C)$ errors occur.

So, once players observed a bad signal for the first time, the posterior \hat{A}_i^t jumps down at least below $\frac{1}{n}$ independent of the prior $\hat{A}_i^{t_i-1}$ or $q_i^{t_i-1}$ for $t = 3$: This argument is independent of the level of " ϵ ":

When $t = 2$; \hat{A}_i^t decreases enough to be less than $\frac{1}{n}$ if " ϵ " is very small. This is because players do not interpret it as an error but as a signal of $\frac{3}{4}D$ at the first period.

$$\begin{aligned}
 \text{case 3: } \frac{1}{2} \quad h_i^t &= \frac{1}{2} h_i^{t_i-1} + \frac{1}{2} D; \hat{A}_i^{t_i-1} & \text{for } t = 3 \\
 h_i^t &= D; \hat{A}_i^{t_i-1} & \text{for } t = 2 \quad \text{with } \hat{A}_i^{t_i-1} = \hat{A}^0
 \end{aligned}$$

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

(i): $n = 3$

By Bayes' Rule,

$$\hat{A}_i^t = \hat{A}(q_{i-1}(h_i^t)) = \frac{A_i^{t-1} P((!_{i-1}^{t-1}; !_{i-1}^{t-1}) = (C; !_{i-1}^{00}) j \mu_{t_i-1} = 0)}{(1_i A_i^{t-1}) P(!_{i-1}^{t-1} = !_{i-1}^{00} j \mu_{t_i-1} \in 0) + A_i^{t-1} P(!_{i-1}^{t-1} = !_{i-1}^{00} j \mu_{t_i-1} = 0)}$$

This is bounded above by

$$\begin{aligned} & \frac{P(!_{i-1}^{t-1}; !_{i-1}^{t-1}) = (C; !_{i-1}^{00}) j \mu_{t_i-1} = 0}{P(!_{i-1}^{t-1} = !_{i-1}^{00} j \mu_{t_i-1} = 0)} \\ & \leq \frac{P(\#(!_{i-1}^{00} j C) + n_{i-1})}{P(\#(!_{i-1}^{00} j C) + n_{i-1}) + P(\#(!_{i-1}^{00} j C) + n_{i-1}) \prod_{m=1}^{n_{i-1}} \frac{1}{1 + \frac{P(!_{i-1}^{00} j C)}{m}}} \\ & = \frac{1}{1 + \prod_{m=1}^{n_{i-1}} \frac{1}{1 + \frac{P(!_{i-1}^{00} j C)}{m}}} \\ & \leq \frac{1}{6} \text{ (This holds with equality when } n = 3 \text{)} \end{aligned}$$

This argument is independent of "i"; too.

(ii): n = 2

$$\begin{aligned} \hat{A}_i^t(q_{i-1}(h_i^t); (D; C)) &= \frac{A_i^{t-1} (1_i P(1_i) P(2)) + (1_i A_i^{t-1}) (P(1) + P(2))}{A_i^{t-1} (1_i P(1_i) P(2)) + (1_i A_i^{t-1}) (P(1) + P(2))} \\ \hat{A}_i^t(q_{i-1}(h_i^t); (D; D)) &= \frac{A_i^{t-1} P(1) + (1_i A_i^{t-1}) (1_i P(1) P(2))}{A_i^{t-1} (P(1) + P(2)) + (1_i A_i^{t-1}) (1_i P(1) P(2))} \end{aligned}$$

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

It can be shown that $\hat{A}_i^t(q_{i-1}(h_i^t); (D; !_i)) \leq \frac{1}{2}$ when "i" is small. See Sekiguchi [16] for detail.

With case 2 and case 3, I can conclude that $\hat{A}_i^t(q_{i-1}(h_i^t)) \leq \frac{1}{n}$ after any history such as

$$\begin{aligned} & \frac{1}{2} h_i^t = !_{i-1} h_{i-1}^{t-1}; !_{i-1} C; !_{i-1}^{t-1} \text{ for } t = 3 \\ & h_i^t = !_{i-1} C; !_{i-1}^{t-1} \text{ for } t = 2 \\ & \text{with } !_{i-1}^{t-1} \in C \end{aligned}$$

or

$$\begin{aligned} & \frac{1}{2} h_i^t = !_{i-1} h_{i-1}^{t-1}; !_{i-1} D; !_{i-1}^{t-1} \text{ for } t = 3 \\ & h_i^t = !_{i-1} D; !_{i-1}^{t-1} \text{ for } t = 2 \end{aligned}$$

¥

References

- [1] Abreu, D. Pearce, and E. Stachetti, Toward a theory of discounted repeated games with imperfect monitoring, *Econometrica* 58 (1990), 1041-1064.
- [2] V. Bhaskar, Informational constraints and overlapping generations model: folk and anti-folk theorem, *Rev. of Econ. Stud* 65, (1998), 135-149.
- [3] V. Bhaskar, Sequential equilibria in the repeated prisoner's dilemma with private monitoring, mimeo, 1999.
- [4] V. Bhaskar, The robustness of repeated game equilibria to incomplete payoff information, mimeo, 1999.
- [5] O. Compte, Communication in repeated games with imperfect private monitoring, *Econometrica* 66 (1998), 597-626.
- [6] G. Ellison, Cooperation in the prisoner's dilemma with anonymous random matching, *J. Econ. Theory* 61 (1994), 567-588.
- [7] J. C. Ely and J. Välimäki, A robust folk theorem for the prisoner's dilemma, mimeo, 1999.
- [8] D. Fudenberg, D. Levine, and E. Maskin, The folk theorem with imperfect public information, *Econometrica* 62 (1994), 997-1040.
- [9] J. Harsanyi, Games with randomly distributed payoffs: a new rational for mixed-strategy equilibrium points, *Int. J. Game Theory* 2, (1973), 1-23.
- [10] M. Kandori and H. Matsushima, Private observation, communication, and collusion, *Econometrica* 66 (1998), 627-652.
- [11] G. Mailath and S. Morris, Repeated games with almost-public monitoring, mimeo, 1999.
- [12] 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.
- [13] I. Obara, Repeated prisoners' dilemma with private monitoring: a N-player case, mimeo, 1999
- [14] I. Obara, Private strategy and efficiency: repeated partnership games revisited, mimeo, 1999.
- [15] M. Piccione, The repeated prisoner's dilemma with imperfect private monitoring, mimeo, 1998.
- [16] T. Sekiguchi, Efficiency in repeated prisoner's dilemma with private monitoring, *J. Econ. Theory* 76 (1997), 345-361.