Belief-Based Equilibria in the Repeated Prisoners' Dilemma with Private Monitoring*

V. Bhaskar†
Department of Economics
University of Essex

Ichiro Obara‡
Department of Economics
University of Pennsylvania

May 2001

Abstract

We analyze infinitely repeated prisoners' dilemma games with imperfect private monitoring. We construct mixed trigger strategy equilibria, and show that such strategies have a simple representation, where a player's action only depends upon her belief that her opponent(s) are continuing to cooperate. When monitoring is almost perfect, the symmetric efficient outcome can be approximated in any prisoners' dilemma game, while every individually rational feasible payoff can be approximated in a class of prisoner dilemma games. We also extend the approximate efficiency result to n-player prisoners' dilemma games. Our results require that monitoring be sufficiently accurate but do not require very low discounting when a public randomization device is available.

1. Introduction

We analyze a class of infinitely repeated prisoners' dilemma games with imperfect private monitoring and discounting. We construct and build upon mixed trigger strategy equilibria, and show that these can be represented in the form of "belief-based" strategies, where a player's continuation strategy is a function only of her

*This paper combines independent papers by each of the authors. We are grateful to Jeff Ely, Johannes Horner, Michihiro Kandori, Stephen Morris, Nicolas Persico and Tadashi Sekiguchi for their comments on the earlier versions of this paper. The first author thanks the University of Tokyo for its hospitality while part of this work was done. The second author is grateful to George Mailath and Andrew Postlewaite for their advice and encouragement since the very early stage of this research.

†Email: vbhas@essex.ac.uk
‡Email: obara@ssc.upenn.edu

beliefs about her opponent's continuation strategy. This representation simplifies the analysis considerably — in the two-player case, we explicitly construct sequential equilibria, enabling us to invoke the one-step deviation principle of dynamic programming. By doing so, we prove that one can approximate the symmetric efficient payoff in any prisoners' dilemma game provided that the monitoring is sufficiently accurate. Furthermore, for a class of prisoners' dilemma games, one can approximate every individually rational feasible payoff. Our efficiency results also generalize to the n player case, where we show that the symmetric efficient payoff can similarly be approximated.

These results are related to an important paper by Sekiguchi [13], who uses trigger strategies to show that one can approximate the efficient payoff in two-player prisoners' dilemma games provided that the monitoring is sufficiently accurate. Sekiguchi's result applies for a class of prisoners' dilemma payoffs, and relied on the construction of a Nash equilibrium which achieves approximate efficiency. Our results can be viewed as an extension and generalization of the approach taken in Sekiguchi's paper.

Our substantive results are also related to those obtained in recent papers by Piccione [12] and Ely and Välimäki [6], which adopt a very different approach. The current paper (and Sekiguchi's) utilizes initial randomization to ensure that a player's beliefs adjust so that she has incentive to punish or reward her opponent(s) as is appropriate. The Piccione-Ely-Välimäki approach on the other hand relies on making each player indifferent between her different actions at most information sets, so that her beliefs do not matter.[1] While we defer a more detailed discussion of the two approaches to the concluding section of this paper, it may be noted that our approach relies on extending trigger strategy equilibria to the private monitoring context.[2] The advantage of trigger strategies is their simplicity. With two players and private monitoring, if a player adopts such a (possibly mixed) trigger strategy, the support of the beliefs of her opponent, at any stage of the game, consist only of two pure strategies — the grim trigger strategy and the strategy which always defects. More generally, the idea underlying our construction is to find a finite set of repeated game strategies that is *closed*, so that for all "relevant" beliefs, a best response is contained within this

[1] Obara [11] independently found a similar strategy which has the same property, but used it for repeated games with *imperfect public monitoring* to construct a sequential equilibrium which Pareto-dominates perfect public equilibria in simple repeated partnership games.

[2] While the idea underlying the construction is that of a trigger strategy equilibrium, the strategy may require that players return to cooperation depending, for example, upon a sunspot. This is essential in Sekiguchi's case as well as ours. As Compte [4] has shown, if one is restricted to equilibria in strategies where players do not cooperate once they have defected, then equilibrium payoffs must tend to the minmax payoff as discounting vanishes.

set. Initial randomization by a player provides a suitable "prior", so that the set of relevant beliefs for her opponents' is sufficiently restricted that closedness obtains.

The rest of this paper is as follows. Section 2 constructs sequential equilibria which approximate the efficient outcome in the two-player case, while section 3 approximates the set of individually rational feasible payoffs in this case. Section 4 shows that the efficiency result can be generalized to n player prisoners' dilemma games. We conclude in section 5.

2. Approximating the Efficient Payoff

We consider the prisoners' dilemma where each player chooses her action from the action set $A = \{C, D\}$. Each player only observes her own action and a private signal from the set $\Omega = \{c, d\}$, where c (resp. d) is more likely when the opponent plays C (resp. D). The signalling structure is assumed symmetric, in the sense that the probability of errors does not depend on the action profile played and the identity of players. Given any action profile $a = (a_1, a_2)$, $a_i \in A$, the probability that exactly one player receives a wrong signal is $\varepsilon > 0$, and the probability that both receive wrong signals is $\xi > 0$. The realized payoff of a player depends upon her own action and upon the signal she observes. Given this realized payoff function, one may write down the expected payoff of a player, as a function of the action profile chosen (see [13] for the details of this construction). This is the strategic form of the stage game, and is shown in Fig. 1, where the row indicates the player's own action and the column indicates her opponent's action. We assume that $g, l > 0$ and $g - l < 1$, so that this is the usual prisoners' dilemma, where playing (C, C) yields a higher payoff than an equi-probable randomization over (C, D) and (D, C).

	C	D
C	1	$-l$
D	$1+g$	0

Fig. 1

Players maximize the expected sum of stage game payoffs discounted at rate δ. We also assume that at the end of each period, players observe the realization of a public randomization device uniformly distributed on the unit interval.

Our approach is closely related to Sekiguchi's [13]: we show that one can construct a mixed trigger strategy sequential equilibrium which achieves partial cooperation. With public randomization or by "dividing up the game" as in Ellison [5], one can modify the strategy appropriately in order to approximate full cooperation as discounting vanishes. Our approach involves the construction of a "belief-based" strategy, i.e. a strategy which is a function of the player's beliefs about her opponent's continuation strategy. This results in a major simplification as compared to the more conventional notion of a strategy which is a function of the private information of the player.

We begin by defining *partial continuation strategies*. In any period t, define the partial continuation strategy σ_D as follows: play D at period t, and at period $t+1$ play σ_D if the realized outcomes in period t are (Dc) or (Dd). Define the partial continuation strategy σ_C as follows: in any period t play C; at period $t+1$ play σ_C if the realized outcomes in period t is (Cc), and play σ_D if the realized outcome at t is (Cd). We call σ_C and σ_D a partial continuation strategy since each of these fully specifies the player's actions in every subsequent period at every information set that arises given that she confirms to the strategy. In consequence, the (random) path and payoffs induced by any pair of partial continuation strategies is well defined. However, a partial continuation strategy does not specify the player's actions in the event that she deviates from the strategy at some information set. This is deliberate, since our purpose is to construct the full strategies that constitute a sequential equilibrium. Note also that for any player i, only the partial continuation strategy of player j is relevant when computing i's payoffs in any equilibrium.

The key parameters in our analysis are the informational parameters ε, ξ and the discount factor δ. Our focus is in particular upon the situation when ε and ξ are small. We shall assume that the strategic form of the stage game (i.e. the payoff parameters g and l) are fixed, and do not vary as we consider different values of ε, ξ. This assumption is slightly different from that adopted by Sekiguchi [13], where the payoff as a function of the action and signal is assumed fixed, so that the strategic form also varies with ε, ξ.

Let $V_{ab}(\delta, \varepsilon, \xi)$, $a, b \in \{C, D\}$ denote the repeated game payoff of σ_a against σ_b — these payoffs are well defined since the path induced by each pair is well defined. We have that $V_{DD} > V_{CD}$, for all parameter values. Furthermore, if $\delta > \frac{g}{1+g}$, and $(\varepsilon + \xi)$ is sufficiently small, then $V_{CC} > V_{DC}$. Suppose that player i believes that her opponent is playing either σ_C or σ_D, and is playing σ_C with probability μ. Then the difference between the payoff from playing σ_C and the payoff from playing σ_D is given by

$$\Delta V(\mu; \delta, \varepsilon, \xi) = \mu(V_{CC} - V_{DC}) - (1 - \mu)(V_{DD} - V_{CD}) \tag{2.1}$$

Hence $\Delta V(\mu)$ is increasing and linear in μ and there is a unique value, $\pi(\delta, \varepsilon, \xi)$, at which it is zero. Suppose now that at $t = 1$ both players are restricted to choosing between σ_C and σ_D. There is a mixed equilibrium of the restricted game, where each player plays the strategy σ which plays σ_D with probability $1-\pi$ and σ_C with probability π. Call this partial mixed strategy σ. Note that $\pi(\delta,\varepsilon,\xi)$ increases to 1 as we decrease δ towards its lower bound $\frac{g}{1+g}$. Let δ be such that $\pi > \frac{1}{2}$.

For future reference we emphasize that equation (2.1) applies to any period — if a player believes that her opponent's continuation strategy is σ_C with probability μ and σ_D with probability $1-\mu$, then she prefers σ_C to σ_D if $\mu > \pi$ and prefers σ_D to σ_C if $\mu < \pi$. Note also that if a player's opponent begins at $t = 1$ with a strategy in $\{\sigma_C, \sigma_D\}$, her continuation strategy also belongs to this set, since σ_D induces only σ_D, while σ_C may induce either σ_C or σ_D, depending upon the private history that the opponent has observed.

We define the following four belief revision operators. Starting with any initial belief μ, we can define a player's new beliefs when she takes action a and receives signal ω. Her new belief (i.e. the probability that j's continuation strategy is σ_C) will be given by $\chi_{a\omega}(\mu)$. We have four belief operators, χ_{Cc}, χ_{Cd}, χ_{Dc}, χ_{Dd}, where each $\chi_{a\omega} : [0,1] \to [0,1]$ is defined, using Bayes rule, as follows

$$\chi_{Cc}(\mu) = \frac{\mu(1-2\varepsilon-\xi)}{\mu(1-\varepsilon-\xi)+(\varepsilon+\xi)(1-\mu)} \tag{2.2}$$

$$\chi_{Cd}(\mu) = \frac{\mu\varepsilon}{\mu(\varepsilon+\xi)+(1-\varepsilon-\xi)(1-\mu)} \tag{2.3}$$

$$\chi_{Dc}(\mu) = \frac{\mu\varepsilon}{\mu(1-\varepsilon-\xi)+(\varepsilon+\xi)(1-\mu)} \tag{2.4}$$

$$\chi_{Dd}(\mu) = \frac{\mu\xi}{\mu(\varepsilon+\xi)+(1-\varepsilon-\xi)(1-\mu)} \tag{2.5}$$

Starting with any initial belief $\hat{\mu}$ at the beginning of the game, a player's belief at any private history, i.e. after an arbitrary sequence $\langle(a\omega)_r\rangle_{r=1}^{t}$, can be computed by iterated application of the appropriate belief operators. Let $\Xi(\hat{\mu})$ be the set of *possible beliefs,* i.e. $\mu \in \Xi(\hat{\mu}) \Leftrightarrow \exists <\mu_r>_{r=1}^{t}$: $\mu_1 = \hat{\mu}$, $\mu_t = \mu$ and $\mu_{r+1} = \chi_{(a\omega)_r}(\mu_r), (a\omega)_r \in A \times \Omega, 1 \le r \le t-1$. Let τ be a (full) strategy, which is defined at every information set, i.e. after arbitrary private histories. Clearly, τ is a best response to σ after every private history if and only if it is optimal to play τ at every belief $\mu \in \Xi(\hat{\mu})$, i.e. at all possible beliefs given the initial belief $\hat{\mu}$. In this section, the initial belief $\hat{\mu}$ will usually coincide with π, the mixing probability where a player is indifferent between σ_C and σ_D.

We now examine the properties of these belief operators. Each is a strictly increasing function. Let us also assume that $\max\{\varepsilon,\xi\} < \frac{1}{6}$. The operator χ_{Cc}

has an interior fixed point at $\theta(\varepsilon,\xi)$, with $\chi_{Cc}(\mu) \lesseqgtr \mu$ as $\mu \gtreqless \theta$, where

$$\theta(\varepsilon,\xi) = \frac{1-3\varepsilon-2\xi}{1-2\varepsilon-2\xi} > \frac{1}{2} \tag{2.6}$$

We now show $\chi_{a\omega}(\mu) < \mu$ for each of the other three belief operators, χ_{Cd}, χ_{Dc} and χ_{Dd}. To verify this, take any typical expression from (2.3)-(2.5), and divide by μ. This yields ε (or ξ) in the numerator, while the denominator is strictly larger since it is a convex combination of $(\varepsilon+\xi)$ and $(1-\varepsilon-\xi)$.

Since $\mu < \theta \Rightarrow \chi_{a\omega}(\mu) < \theta$ for any belief operator, this immediately implies that if the initial belief equals $\pi < \theta$, $\Xi(\pi) \subseteq [0,\theta)$, since we have demonstrated that no point $\mu' > \theta$ is the image of any $\mu \leq \theta$ under *any* belief operator. Hence, provided that initial beliefs are given by $\pi < \theta$, it suffices to define our belief based strategy for beliefs in the set $[0,\theta]$. Let $\rho : [0,\theta] \to \{C, D, \pi\}$ be defined as follows: $\rho(\mu) = C$ if $\mu \in (\pi,\theta]$ and $\rho(\mu) = D$ if $\mu \in [0,\pi)$. If $\mu = \pi, \rho(\mu) = \pi$, i.e. ρ plays C with probability π and D with probability $1-\pi$. Hence the pair (π,ρ), i.e. ρ in conjunction with an initial belief $\pi \leq \theta$, specifies an action at every possible belief, and hence a complete strategy.

The advantage of this specification is that a player's continuation strategy is specified even at information sets which arise due to a player's deviating from ρ in the past. The belief based strategy (π,ρ) is *realization equivalent* to the partial strategy σ if it induces the same probability distribution over actions at every private history where σ is defined. This reduces to the following condition:

Definition 2.1. *(π,ρ) is realization equivalent to σ if $\mu \in [\pi,\theta] \Rightarrow [\chi_{Cc}(\mu) > \pi$ and $\chi_{Cd}(\mu) < \pi]$ and $\mu \in [0,\pi] \Rightarrow [\chi_{Dc}(\mu) < \pi$ and $\chi_{Dd}(\mu) < \pi]$.*

Lemma 2.2. *If $\frac{1}{2} < \pi < \theta(\varepsilon,\xi)$, (ρ,π) is realization equivalent to σ.*

Proof. To verify that $\mu \in [\pi,\theta] \Rightarrow \chi_{Cc}(\mu) > \pi$, recall that $\chi_{Cc}(\mu) > \mu$ if $\mu < \theta$, so that $\chi^k_{Cc}(\pi) > \pi$ for any k. To verify $\mu \in [\pi,\theta] \Rightarrow \chi_{Cd}(\mu) < \pi$, it suffices to verify that $\chi_{Cd}(\theta) \leq \pi$, since χ_{Cd} is strictly increasing.

$$\chi_{Cd}(\theta) = \frac{\varepsilon(1-3\varepsilon-2\xi)}{(\varepsilon+\xi)(1-3\varepsilon-2\xi)+(1-\varepsilon-\xi)\varepsilon} < \frac{\varepsilon}{2\varepsilon} = \frac{1}{2} < \pi \tag{2.7}$$

$\mu \leq \pi \Rightarrow \chi_{Dc}(\mu) < \pi$ and $\mu \leq \pi \Rightarrow \chi_{Dd}(\mu) < \pi$ follow from the already established fact that χ_{Dc} and χ_{Dd} lie below the 45° line.■

Note that if ε and ξ are sufficiently small, we can select $\delta > \frac{g}{1+g}$ so that $\pi(\delta,\varepsilon,\xi) \in (\frac{1}{2},\theta)$ — this follows from the fact that $\pi(\delta,\varepsilon,\xi) \to 1$ as $\delta \to \frac{g}{1+g}$ and

$(\varepsilon + \xi) \to 0$, while $\pi(\delta, \varepsilon, \xi) \to 0$ if $\delta \to 1$ and $(\varepsilon + \xi) \to 0$. Henceforth we shall assume that δ is such that $\pi \in (\frac{1}{2}, \theta)$ so that (π, ρ) defines a full strategy which is realization equivalent to σ.

Proposition 2.3. *If $\frac{1}{2} < \pi < \theta(\varepsilon, \xi)$, the strategy profile where each player plays (π, ρ) is a sequential equilibrium.*

Proof. Note first that if $\mu = \pi$, a player is indifferent between playing σ_C and σ_D, and hence a one-step deviation from playing ρ is not profitable. Since the payoffs from playing σ, σ_C and σ_D are equal at belief π, one may also, for the purposes of computing payoffs, use σ_C or σ_D as is computationally convenient in the event of belief π.

Consider first the case when $\mu > \pi$. A one-step deviation from ρ is to play D, and to continue with ρ in the next period. The following sub-cases arise:

a) Suppose that $\chi_{Dc}(\mu) \leq \pi$ and $\chi_{Dd}(\mu) \leq \pi$. In this case, a one-step deviation from ρ is to play σ_D, whereas $\rho(\mu) = \sigma_C$. However, (2.1) establishes that in this case σ_C is preferable to σ_D, and hence a one-step deviation from ρ is unprofitable.

b) Suppose that $\chi_{Dc}(\mu) \leq \pi$ and $\chi_{Dd}(\mu) > \pi$, so that the one-step deviation is to play D today and continue with σ_D if Dc is reached, and to continue with σ_C if Dd is reached. Let $\Delta\tilde{V}(\mu)$ be payoff difference between the equilibrium strategy and the one-step deviation. Note that the one step deviation differs from σ_D only at the information set Dd; at this information it continues by playing σ_C whereas σ_D continues with σ_D. Hence we can write $\Delta\tilde{V}(\mu)$ as the payoff difference between σ_C and σ_D minus the payoff difference between σ_C and σ_D conditional on Dd being reached, as follows:

$$\Delta\tilde{V}(\mu) = \Delta V(\mu) - \delta[\mu(\varepsilon + \xi) + (1 - \mu)(1 - \varepsilon - \xi)][\Delta V(\chi_{Dd}(\mu))]$$

Note that $\chi_{Dd}(\mu) < \mu$. Equation (2.1) shows that this implies that $\Delta V(\mu) > \Delta V(\chi_{Dd}(\mu))$. Since the coefficient multiplying $\Delta V(\chi_{Dd}(\mu))$ is strictly less than one, this implies that $\Delta\tilde{V}(\mu) > 0$. Hence if $\mu > \pi$, a one-step deviation is unprofitable.

c) Finally, we establish that $\chi_{Dc}(\mu) < \pi \ \forall \ \mu \leq \theta$, so that no other sub-case need be considered. Evaluating χ_{Dc} at the upper bound θ, we have

$$\chi_{Dc}(\theta) = \frac{\varepsilon(1 - 3\varepsilon - 2\xi)}{(1 - \varepsilon - \xi)(1 - 3\varepsilon - 2\xi) + (\varepsilon + \xi)\varepsilon} < \frac{\varepsilon}{1 - \varepsilon - \xi} < \frac{1}{2}$$

where the last step follows from the assumption that $\max\{\varepsilon, \xi\} < \frac{1}{6}$.

Consider now the case when $\mu < \pi$. In this case, a one-step deviation from ρ is to play C today, and to continue with σ_C if $\chi_{Cc}(\mu) \geq \pi$, but to continue with σ_D if $\chi_{Cc}(\mu) < \pi$. (Note that $\mu < \pi \Rightarrow \chi_{Cd}(\mu) < \pi$, so the continuation strategies do not differ in this event.) In the first sub-case, the one-step deviation from ρ corresponds to playing σ_C, and (2.1) establishes that in this case σ_D is preferable to σ_C, and hence a one-step deviation from ρ is unprofitable. In the second sub-case, the one-step deviation differs from σ_C only at the information set Cc — it plays σ_D at this information set rather than σ_C. Let $\Delta\hat{V}(\mu)$ denote the payoff difference between the one-step deviation and the equilibrium strategy σ_D. We have

$$\Delta\hat{V}(\mu) = \Delta V(\mu) - \delta[\mu(1-\varepsilon-\xi) + (1-\mu)(\varepsilon+\xi)][\Delta V(\chi_{Cc}(\mu))]$$

Since $\pi > \chi_{Cc}(\mu) > \mu, \Delta V(\mu) < \Delta V(\chi_{Cc}(\mu)) < 0$. Also, the coefficient multiplying $\Delta V(\chi_{Cc}(\mu))$ is less than 1, which establishes that $\Delta\hat{V}(\mu) > 0$.

We have therefore established that if a player's opponent j plays the strategy σ (which randomizes between σ_C and σ_D), it is optimal for player i to play ρ, with initial belief π. However, (ρ, π) is consistent and realization equivalent to the strategy σ. Hence the profile where both players play (ρ, π) is a sequential equilibrium.■

Note that π plays a dual role in the construction of the mixed strategy equilibrium. On the one hand it is initial belief — the randomization probability in the first period. On the other hand, it is simply a number which defines the threshold at which behavior changes. These roles are obviously distinct, and this distinction is relevant when we discuss the folk theorem in the following section.

With the construction of the mixed equilibrium, one can approximate full cooperation by using one of two devices. If a public randomization device is available, then it is immediate that the equilibrium payoff set is monotonically increasing (in the sense of set inclusion) in δ — given any $\delta' > \delta$, players may simply re-start the game with probability $m = \frac{\delta(1-\delta')}{\delta'(1-\delta)}$ in every period. In the absence of such public randomization, one may use the construction introduced by Ellison [5] (see also, Sekiguchi [13]), of dividing the game into n separate repeated games, thereby reducing the discount factor.

Lemma 2.4. *Let $\delta_0 < \delta_1 < 1$, and let there be a strategy profile which is a sequential equilibrium of the repeated game for any $\delta \in (\delta_0, \delta_1)$, yielding payoff $v(\delta) \geq v$ for any $\delta \in (\delta_0, \delta_1)$. There exists $\hat{\delta} < 1$ such that the repeated game has a sequential equilibrium with payoff greater than v for any $\delta \geq \hat{\delta}$. If a public randomization device is available and (v_1, v_2) is a sequential equilibrium payoff for some $\delta \in (0, 1)$, it is also an equilibrium payoff for any $\delta' > \delta$.*

Proof. For the proof of the first part of this lemma, see Ellison [5]. To prove the second part, let τ be the strategy profile giving the required payoff given δ. Given δ', let $m = \frac{\delta(1-\delta')}{\delta'(1-\delta)}$. Players play a sequence of games: they begin with the strategy profile τ. If the sunspot in any period is larger than m, they play a new game and re-start with τ.■

Proposition 2.5. *Fix g and l. For any $x < 1$, there exists a symmetric sequential equilibrium with payoff greater than x if ε and ξ are sufficiently small, provided that either (i) δ is sufficiently close to 1 or (ii) $\delta > \frac{g}{1+g}$ and a public randomization device is available.*

Proof. Proposition 2.3 implies that if (ε, ξ) are sufficiently small, so that $\theta(\varepsilon, \xi)$ is close to 1, we have an open interval of values of π such that (π, ρ) is a sequential equilibrium. In this range, $\pi(\delta, \varepsilon, \xi)$ is a strictly decreasing function of δ, and hence if (ε, ξ) are sufficiently small, there is an open interval of values of δ such that $(\pi(\delta, \varepsilon, \xi), \rho)$ is a sequential equilibrium. Since (ε, ξ) are close to zero, we can select this interval of values of π close to 1, so that the payoff in any such equilibrium is greater than x. Part (i) of the proposition then follows from the first part of lemma 2.4. If a public randomization device is available, let $(\varepsilon, \xi) \to (0, 0)$ and $\delta \to \frac{g}{1+g}$, so that $\pi \to 1$. The equilibrium payoff tends to one. Lemma 2.4 ensures that this result holds for all $\delta > \frac{g}{1+g}$.■

The assumptions on the information structure can be significantly relaxed. The crucial assumption we need for our construction is that a player is almost sure that the other player played C and observed c when she played C and observed c. This means that the event that both players play σ_C is almost common knowledge after the fully cooperative private history. As long as this condition is satisfied, we can relax almost perfect monitoring or incorporate many other signals under a weak assumption on the information structure. The details of this more general result are available in the discussion paper version of our paper.

Although it is common to allow for vanishing discounting in proving folk theorems in repeated games, it is worth pointing out that in order to obtain approximate efficiency, such vanishing discounting is not required. Indeed, the construction relies on the effective discount factor tending to its lower bound, $\frac{g}{1+g}$. If a public randomization device is available, for any greater discount factor, players can effectively reduce it by re-starting the game with some probability. In the absence of such a randomization device, one does require vanishing discounting, essentially due to an "integer" problem.

3. Approximating Any Individually Rational Feasible Payoff

We now build on the construction of the previous section and show how to approximate any individually rational feasible payoff. We assume in this section that a public randomization device is available. The key step is to approximate the payoff $(\frac{1+g+l}{1+l}, 0)$, which is player 1's maximal payoff within the set of individually rational and feasible payoffs. Since the payoff $(1, 1)$ has already been approximated in the previous section, and $(0, 0)$ is a stage game equilibrium payoff, one can then use public randomization to approximate any individually rational feasible payoff.

It might be useful to outline the basic construction and to explain the complications that arise. The basic idea in approximating the extremal asymmetric payoff is that play begins in the *asymmetric phase* where player 1 plays D and player 2 randomizes, playing C with a high probability, υ. This asymmetric phase continues or ends, depending upon the realization of a public randomization device. Thus player 1's per-period payoff in the asymmetric phase is approximately $1 + g$ while player 2's per-period payoff is approximately $-l$. Since the latter is less than the individually rational payoff for player 2, he must be rewarded for playing C. To ensure this, when the asymmetric phase ends, both player's continuation strategies depend upon their private information. Player 1 continues with σ_C if she has observed the signal c in the last period (i.e. if her information is Dc)and continues with σ_D if she has observed d (i.e. if her information is Dd). This ensures that player 2 is rewarded for playing C in the asymmetric phase. Similarly, player 2 continues with σ_C if his private information is Cd, the information set which is most likely when he plays C,and continues with σ_D if his private information is Dd. Hence, if the noise is small, player 2's continuation payoff when the asymmetric phase ends is approximately 1 if he has played C in the previous period and approximately zero if he has played D. Hence if δ is large relative to l ($\delta > \frac{l}{1+l}$), we can, by choosing the value of the sunspot appropriately, make player 2 indifferent between C and D in the asymmetric phase. The payoffs in this equilibrium converge to $(\frac{1+g+l}{1+l}, 0)$ as the noise vanishes.

However, one must also verify that the players find it optimal to play σ_C and σ_D, as appropriate, at each information set after the end of the asymmetric phase. A complication arises here, as compared to the previous section, since player 1 does not randomize in the asymmetric phase, i.e. she plays D for sure. (Indeed, she cannot play C with positive probability, since in that case her payoff in the asymmetric phase is bounded above by 1 and hence cannot approximate $1 + g$).[3]

[3] This argument is more general and implies that one cannot have a folk theorem in completely mixed strategies for stage games with non-degenerate payoffs. Let $\hat{v}_1$ be the supremum of the

Hence when player 2 receives the signal c, he knows that there has been at least one error in signals, and his beliefs about player 1's continuation strategy depend upon the relative probability of one (ε) versus two errors (ξ). In other words, his continuation strategy at the information sets Cc and Dc depends upon the correlation structure of signals. Since player 2's continuation strategy depends upon this correlation structure, this implies that player 1's beliefs also depend upon the correlation structure.

We adopt two alternative approaches to handle this problem. First, we show that if signals are positively correlated, so that the probability of two errors is at least as large as the probability of one error, then one can approximate the asymmetric payoff, without any restriction upon payoffs. Second, we show that one does not need such positive correlation of signals provided that one can choose δ so that $\pi(\delta, \varepsilon, \xi)$ sufficiently close to one. This result applies to any prisoners' dilemma game where $g \geq l$ — in any such game one can approximate the asymmetric payoff arbitrarily closely. However, this second approach does not work if $l > g$, since in this case one cannot have $\pi(\delta, \varepsilon, \xi) \to 1$. The reason for this is that for π to be close to 1, we must have $\delta \to \frac{g}{1+g}$. However, in the asymmetric phase, player 2 incurs a loss of l by playing C, whereas his continuation payoff gain is no more than 1. Hence player 2 will be willing to play C in the asymmetric phase only if $\delta > \frac{l}{1+l}$. Hence if $l > g$, one cannot have π close to 1 since δ is bounded away from $\frac{g}{1+g}$.

We make the following assumption for this section:

Assumption A: Either A1: $\xi \geq \varepsilon$ or A2: $g \geq l$ and $\xi(1-\xi)(1-3\varepsilon-2\xi) > \varepsilon^3$.

Note that A1 is a relatively strong assumption that signals are positively correlated, but does not require any assumption on payoffs. On the other hand, A2 requires an assumption on payoffs but is a mild assumption about the relative probability of errors. It is always satisfied if signals are positively correlated, or independent. In the independent signal case, the left hand side is a term of order ε whereas the right hand side is a term of order ε^3. Hence A2 is satisfied even if signals are negatively correlated provided that they are not too highly so.

We now define the players' strategies more precisely. In any period $t-1$ in the asymmetric phase, player 1 plays D for sure, while player 2 randomizes between C and D, choosing C with a constant probability υ which is close to 1. At the end of period, players observe the realization of a sunspot which is uniformly distributed on $[0, 1]$. If it is larger than $1 - \lambda$, both players continue in the asymmetric phase

continuation payoffs of player 1 in any equilibrium where player 1 randomizes in every period at every information set. Since $\hat{v}_1 \leq (1-\delta)\min_{a_1}\{\max_{a_2} u_1(a_1, a_2)\} + \delta\hat{v}_1$, this implies $\hat{v}_1 \leq \min_{a_1}\{\max_{a_2} u_1(a_1, a_2)\}$.

for the next period. If it is smaller than λ, the asymmetric phase ends for both players, and is never reached again. In this case, players' continuation strategies (i.e. their states) depend upon the realization of their private information at date $t-1$ (i.e. players ignore their private information from previous dates). Let ν_{t-1} denote the player's private information realization at date $t-1$. Player 1 continues with σ_C if $\nu_{t-1} = Dc$; if $\nu_{t-1} = Dd$, she continues in period t with σ_D.[4] Player 2 continues with σ_C if $\nu_{t-1} = Cd$, and continues with σ_D if $\nu_{t-1} = Dd$. If $\nu_{t-1} \in \{Cc, Dc\}$, player 2 continues with σ_C if $\mu_2(\nu_{t-1}) > \pi(\delta, \varepsilon, \xi)$ and with σ_D if $\mu_2(\nu_{t-1}) \leq \pi(\delta, \varepsilon, \xi)$.

Our analysis proceeds as follows. First, we show that player 2 is willing to randomize in the asymmetric phase provided that λ is appropriately chosen, and that the payoffs associated with this class of equilibria tend to $(\frac{1+g+l}{1+l}, 0)$ as the noise vanishes. Subsequently, we shall demonstrate that all players are choosing optimally at every information set.

Write $W_2(D)$ for the payoff of player 2 in the asymmetric phase given that he plays D, and $W_2(C)$ for the payoff in the asymmetric phase from playing C. Since $W_2(D) = W_2(C) = W_2$, we have

$$W_2(D) = \delta(1-\lambda)W_2 + \delta\lambda V_2(D)$$

where $V_2(D)$ is the expected payoff to player 2 conditional on the fact that the asymmetric phase has ended and that he has played D. Similarly, letting $V_2(C)$ be the expected payoff to player 2 conditional on the fact that the asymmetric phase has ended and that he has played C, we have

$$W_2(C) = (1-\delta)(-l) + \delta(1-\lambda)W_2 + \delta\lambda V_2(C)$$

Clearly, $V_2(D) \to 0$ as $(\varepsilon, \xi) \to (0,0)$, while it is easy to see that $V_2(C) \to 1$ as $(\varepsilon, \xi) \to (0,0)$. Hence if $\varepsilon + \xi$ is sufficiently small and $\delta > \frac{l}{1+l}$, there exists a value of λ which equates $W_2(C)$ and $W_2(D)$. Further, as $(\varepsilon + \xi) \to 0$, $\lambda \to \frac{(1-\delta)l}{\delta}$, and player 2's payoff converges to zero.

If $\upsilon \to 1$, player 1's per-period payoff tends to $(1+g)$ in the asymmetric phase, and 1 in the cooperative phase. By substituting for the limiting value of λ which is$\frac{(1-\delta)l}{\delta}$, we see that player 1's payoff converges to $\frac{1+g+l}{1+l}$. (We shall establish later that $\upsilon \to 1$).

[4] We show that any strategy which plays C in the asymmetric phase is dominated, and hence we need not define precisely the optimal continuation strategy after playing C. The existence of an optimal continuation strategy follows from the same argument as in Sekiguchi [13]. Since player 1 never plays C in the asymmetric phase, his continuation after his own deviation does not affect player 2's incentives.

It is easy to verify that each player plays optimally in the asymmetric phase: for player 2 by construction, since he is indifferent between C and D. Player 1 also plays optimally, since she is choosing her one shot best response.[5]

We now show that players' behavior is optimal in cooperative phase. The critical part is in the transition, i.e. in the first period after the sunspot signals the end of the asymmetric phase. Since players only condition on their private information in the previous period, we may focus on this alone. Player 1 has two possible information sets, (Dc) and (Dd), whereas player 2 has four possible information sets. Let $\mu_i(\nu)$ denote the probability assigned by player i to her opponent's continuation strategy being σ_C, given that i is at information set ν. As in the previous section, we shall assume that $\max\{\varepsilon, \xi\} < \frac{1}{6}$. Furthermore, as in the previous section, we may also choose parameters so that $\pi(\delta, \varepsilon, \xi) \in (\frac{1}{2}, \theta)$. This is clearly feasible if we invoke the assumption $g \geq l$ in A2, and in this case we may also choose π to be arbitrarily close to its upper bound. If $l > g$, then $\frac{l}{1+l}$ is the lower bound for δ, and we find that $\pi(\delta, \varepsilon, \xi) \to \frac{l}{2l-g} > \frac{1}{2}$ as δ decreases to this lower bound and $(\varepsilon, \xi) \to (0, 0)$. Hence we may assume that $\pi(\delta, \varepsilon, \xi) \in (\frac{1}{2}, \theta)$ in this case as well. We shall also assume that $\upsilon \in (\pi, \theta(\varepsilon, \xi))$. Since $\theta \to 1$ as $(\varepsilon, \xi) \to (0, 0)$ we can also have $\upsilon \to 1$.

Sequential Rationality: Player 2

Consider first the beliefs of player 2. Let $\mu_2(.)$ denote the probability assigned by player 2 to the event that player 1's continuation strategy is σ_C. Since player 1 plays σ_C at Dc and σ_D at Dd, and since player 1 does not play C in the asymmetric phase, we have

$$\mu_2(Cd) = \frac{1 - 2\varepsilon - \xi}{1 - \varepsilon - \xi} > \theta$$

Since $\pi < \theta$, it is optimal to continue with σ_C today at information set Cd. Further, we have

$$\chi_{Cd}(\mu_2(Cd)) = \frac{(1 - 2\varepsilon - \xi)\varepsilon}{(1 - 2\varepsilon - \xi)(\varepsilon + \xi) + (1 - \varepsilon - \xi)\varepsilon} < \frac{1}{2}$$

Hence it is optimal for player 2 to switch to the defection phase if he receives the signal Cd at any date in the future.

[5] It is possible that playing C in the asymmetric phase increases player 1's continuation payoff in the cooperative phase. However, it is easy to see that such an increase can never offset the loss from playing C. A simple proof is as follows. If playing C in the asymmetric phase is optimal for 1, then playing C in every period in the asymmetric phase is also optimal. The overall payoff of this strategy is approximately 1 if the noise is small, whereas the payoff of player 1 in the equilibrium tends to $\frac{1+g+l}{1+l}$, which is strictly greater.

At Dd, we have

$$\mu_2(Dd) = \frac{\varepsilon}{1-\varepsilon-\xi}$$

This is clearly less than $\frac{1}{2}$ since $\max\{\varepsilon, \xi\} < \frac{1}{6}$, so that it is optimal to continue with σ_D.

The delicate part of the construction is at information sets (Cc) and (Dc), i.e. where player 2 knows that there has been at least one error in the signals.

$$\mu_2(Dc) = \frac{\xi}{\xi+\varepsilon}$$

$$\mu_2(Cc) = \frac{\varepsilon}{\xi+\varepsilon}$$

Recall that player 2 plays σ_D at least at one of these information sets, since the above probabilities cannot be both greater than π, which is greater than one-half. Hence there are three possibilities: either both $\mu_2(Dc)$ and $\mu_2(Cc)$ are less than π, or exactly one of these is greater than π. Now if $\mu_2(.) < \pi$ at any information set, it is optimal to continue with σ_D today, and at every future date. Hence it remains to verify the case when $\mu_2(.) \geq \pi$.

Suppose that $\frac{\xi}{\xi+\varepsilon} > \pi$, so that player 2 plays σ_C at Dc. Since $\pi < \theta$, it will be optimal to continue to play σ_C as long as good signals are received. On the other hand, $\frac{\xi}{\xi+\varepsilon} > \pi \Rightarrow \xi \geq \varepsilon$ and $\chi_{Cd}(1) \leq 1/2$. Since χ_{Cd} is increasing, it follows that it is optimal to continue with σ_D once signal d is observed.

Finally, we consider the case where that player 2 plays σ_C at Cc, i.e. when $\frac{\varepsilon}{\varepsilon+\xi} > \pi$. Note that in this case A1 is violated. Hence we assume A2, which ensures that we can make π arbitrarily close to its upper bound θ by selecting δ sufficiently close to $\frac{g}{1+g}$. We can find a value of π such that $\chi_{Cd}(\frac{\varepsilon}{\varepsilon+\xi}) < \pi$ provided that $\chi_{Cd}(\frac{\varepsilon}{\varepsilon+\xi})$ is less than the upper bound for π, i.e.

$$\chi_{Cd}(\frac{\varepsilon}{\varepsilon+\xi}) = \frac{\varepsilon^2}{\varepsilon^2+\xi-\xi^2} < \theta$$

It is easily verified that the inequality above is ensured by condition A2.

Sequential Rationality: Player 1

Consider now the beliefs of player 1. Her beliefs will depend upon player 2's strategy, which in turn depends upon the parameters of the signal distribution, and as we have seen, there are three possible cases.

i) Player 2 plays σ_C only at information set Cd.

In this case, if 1 observes Dc, her belief is given by

$$\mu_1(Dc) = \frac{\upsilon(1-2\varepsilon-\xi)}{\upsilon(1-\varepsilon-\xi)+(1-\upsilon)(\varepsilon+\xi)}$$

Note that the expression is such that $\mu_1(Dc) = \chi_{Cc}(\upsilon)$, where χ_{Cc} is the belief revision operator defined in the previous section. Hence it follows that if $\upsilon \in [\pi, \theta)$, it follows that $\chi^k_{Cc}(\upsilon) \in (\pi, \theta)$, $\forall k$, and hence it is optimal for player 1 to continue with σ_C at every information set.

On the other hand if 1 observes (Dd), it is easy to verify that $\mu_1(Dd) = \chi_{Cd}(\upsilon)$, and since $\upsilon < \theta$, it is optimal to continue with σ_D at this information set.

ii) Player 2 plays σ_C at Cd and Cc and σ_D at Dd and Dc. In this case, assumption A2 applies, so that we may choose π close to its upper bound. At information set Dc,

$$\mu_1(Dc) = \frac{\upsilon(1-\varepsilon-\xi)}{\upsilon(1-\varepsilon-\xi)+(1-\upsilon)(\varepsilon+\xi)}$$

If $\upsilon > \pi$, then $\mu_1(Dc) > \pi$ so that it is optimal for player 1 to start by playing σ_C at information set Dc. In any subsequent period, player 1 will find it optimal to switch to σ_D on receiving a bad signal, provided that

$$\chi_{Cd}(\mu_1(Dc)) = \frac{\upsilon\varepsilon}{\varepsilon+\xi} < \pi$$

Now if $\mu_1(Dc) \le \theta$, lemma 2.2 has verified that a player who begins with σ_C will switch to σ_D on receiving signal Cd in any subsequent period. If $\mu_1(Dc) > \theta$, it suffices to verify that $\chi_{Cd}(\mu_1(Dc)) < \theta$, which is the upper bound for π. This yields the condition

$$\upsilon < \frac{(\varepsilon+\xi)\theta}{\varepsilon} \tag{3.1}$$

Since $\upsilon < \theta$, this condition is also satisfied.

At information set Dd,

$$\mu_1(Dd) = \frac{\upsilon(\varepsilon+\xi)}{\upsilon(\varepsilon+\xi)+(1-\upsilon)(1-\varepsilon-\xi)}$$

which is less than $\frac{1}{2}$ if $\upsilon < 1-\varepsilon-\xi$.

iii) Finally, we consider the case where Player 2 plays σ_C at Cd and Dc and σ_D at Dd and Cc.

$$\mu_1(Dc) = \frac{\upsilon(1-2\varepsilon-\xi)+(1-\upsilon)\xi}{\upsilon(1-\varepsilon-\xi)+(1-\upsilon)(\varepsilon+\xi)} < \frac{1-2\varepsilon-\xi}{1-\varepsilon-\xi}$$

Hence it suffices to evaluate χ_{Cd} at the upper bound, which yields

$$\chi_{Cd}(\frac{1-2\varepsilon-\xi}{1-\varepsilon-\xi}) = \frac{(1-2\varepsilon-\xi)\varepsilon}{(1-2\varepsilon-\xi)(\varepsilon+\xi)+(1-\varepsilon-\xi)\varepsilon} < \frac{1}{2}$$

Hence $\chi_{Cd}(\mu_1(Dc)) < \frac{1}{2}$ for every value of υ. We have therefore shown that σ_C is optimal at Cc, and also at any subsequent period. At information set Dd,

$$\mu_1(Dd) = \frac{\varepsilon}{\upsilon(\varepsilon+\xi)+(1-\upsilon)(1-\varepsilon-\xi)}$$

So, $\mu_1(Dd) \leq \frac{\varepsilon}{\varepsilon+\zeta}$, which is less than $\frac{1}{2}$ in this case (since player 2 plays σ_C at Dc). Hence it is optimal for player 1 to play σ_D at Dd.

We have therefore proved that the payoff $(\frac{1+g+l}{1+l}, 0)$ (and obviously the payoff $(0, \frac{1+g+l}{1+l})$ can be approximated under assumption A provided that $\delta > \max\{\frac{g}{1+g}, \frac{l}{1+l}\}$ and provided that ε and ξ are sufficiently small. The payoff $(1,1)$ has been approximated under a weaker set of assumptions ($\delta > \frac{g}{1+g}$ and ε and ξ sufficiently small), and the payoff $(0,0)$ is a static Nash payoff. Since any payoff individually rational feasible payoff is a convex combination of these payoffs, and can be achieved via public randomization, we have proved the following theorem.

Theorem 3.1. *Fix g and l and assume that Assumption A is satisfied, and players observe a public randomization device, then for any individually rational feasible payoff vector $u = (u_1, u_2)$ and any number $\zeta > 0$, there exist $\varepsilon(\zeta) > 0, \xi(\zeta) > 0$ such that there exists a sequential equilibrium with payoffs within ζ distance of u provided that $\varepsilon < \varepsilon(\zeta)$ and $\xi < \xi(\zeta)$ and $\delta > \max\{\frac{g}{1+g}, \frac{l}{1+l}\}$.*

This result is most closely related to those obtained in a paper by Piccione [12], who also analyzes the prisoners' dilemma with imperfect private monitoring. Our results differ, both in terms of substance and in the techniques/strategies used. Piccione's substantive results are that full cooperation can always be approximated, and further, any individually rational feasible payoff can be approximated in a class of prisoners' dilemma games, i.e. for games where $l \geq g$. The "folk theorem" condition A in the present paper is, in a sense, the opposite of Piccione's condition. More recently, Ely and Välimäki [6] have considerably simplified the technique used in Piccione, and generalized the folk theorem obtained there. We shall discuss the differences between the approach of the present paper and the approach of Piccione and Ely-Välimäki in the concluding section.

4. The N-player Case

In this section, we extend the approximate efficiency result to the n-player prisoners' dilemma. Let $N = \{1, 2, ..., n\}$ be the set of players. Player i chooses an action a_i from the action set $A_i = \{C, D\}$. A n-tuple action profile is denoted by $a \in A = \Pi_{i=1}^{n} A_i$. An action profile of all players but player i is $a_{-i} \in \Pi_{j \neq i} A_j$. Each player receives a private signal $\omega_i = (\omega_{i,1}, ..., \omega_{i,i-1}, \omega_{i,i+1}, ..., \omega_{i,n}) \in \{c, d\}^{n-1}$, where $\omega_{i,j}$ stands for player i's signal about player $j's$ action. A generic signal profile is denoted by $\omega = (\omega_1, ..., \omega_n) \in \Omega$. Conditional distributions of private signals $p(\omega|a)$ are assumed to have full support, that is, $p(\omega|a) > 0 \ \forall a \forall \omega$. We denote by p_ε an information structure for which $p(\omega|a) \geq 1 - \varepsilon$ if $\omega_i = a_{-i}$ for all i.[6] We impose a following symmetry condition on the information structure. For any $a \in A, \omega \in \Omega$, the probability of $(\omega_1, ..., \omega_i, ..., \omega_j, ..., \omega_n)$ given $(a_1, ..., a_i, ..., a_j, ..., a_n)$ is equal to the probability of $(\omega_1, ..., \omega_j, ..., \omega_i, ..., \omega_n)$ given $(a_1, ..., a_j, ..., a_i, ..., a_n)$.[7] This implies that $p(.|.)$ is invariant with respect to any permutation about players' identities.

The following is an example of the information structure which satisfies these requirements.

$$p(\omega|a) = \Pi_j \Pi_{i \neq j} f(\omega_{i,j}|a_j) \text{ for all } a \in A \text{ and } \omega \in \Omega$$

$f(\omega|a)$ is a density function on $\{c, d\}$ such that $\omega = a$ with very high probability. For this information structure, the probability that player $i \neq j$ receives the right signal or the wrong signal about player j's action is the same across $i \neq j$ and all these private signals are conditionally independent.

The realized payoff of a player depends upon her own action and private signals. We assume that a player's expected payoff does not depend on the permutation of the other players' actions. This implies that every player's expected payoff is described as a function of her own action and the number of the other players playing D. We also assume that this function is common to all players, which we denote by $u(a_i, m_i)$. As in the last sections, we fix the expected payoff structure throughout this section, hence we choose suitable realized payoffs to keep the expected payoff structure fixed as we vary the information structure. Playing D is still strictly dominant in the stage game and $(C, ..., C)$ is the efficient action profile as before. We normalize $u(C, 0)$ and $u(D, n-1)$ to 1 and 0 respectively. The deviation gain when m other players are playing D is $g(m) > 0$. For

[6] p_0 corresponds to the perfect monitoring case.

[7] This class of symmetric information structure is slightly more general than the one in the previous sections. For example, the probability that one player receives a wrong signal can depend on the action profile chosen.

example, $g(0) = u(D,0) - u(C,0)$. The largest deviation gain and the smallest deviation gain are denoted by $\overline{g} = \max_{0 \le m \le n-1} g(m)$ and $\underline{g} = \min_{0 \le m \le n-1} g(m)$ respectively.

Player i's t period private history is $h_i^t = \left(\left(a_i^1, \omega_i^1\right), ..., \left(a_i^t, \omega_i^t\right)\right) \in H_i^t$, and $H_i = \cup_{t=0}^{\infty} H_i^t$ with $H_i^0 = \emptyset$. Player i's strategy σ_i is a mapping from H_i to $\triangle A_i$. Discounted average payoffs are defined in the standard way.

For any period t, we define σ_D to be the partial continuation strategy where a player plays D in the current period and continues with σ_D after $\left(D, \omega_i^{t+1}\right)$. We define the partial continuation strategy σ_C as follows; play C in the current period and continue with σ_C after $(C, \mathbf{c} = (c, ..., c))$ and continue with σ_D after $\left(C, \omega_i^{t+1} \neq \mathbf{c}\right)$. When player i is mixing σ_C and σ_D with probability $(\mu_i, 1 - \mu_i)$, such partial mixed strategy is denoted by $\mu_i \sigma_C + (1 - \mu_i) \sigma_D$. Our goal is to approximate the efficient payoff by constructing symmetric equilibria where each player is randomizing between σ_C and σ_D in the initial period as in the two player case. The symmetry assumptions, on the information structure and payoffs, allow us to focus on only one player without loss of generality because we analyze symmetric equilibria.

Let $m_i \in M$ be the number of players using σ_D as a continuation partial strategy among all players except for player i. Player i can derive a probability measure $\boldsymbol{\mu}_{-i}\left(h_i^t, p\right)$ on the space M conditional on the realization of the private history h_i^t and any partial mixed strategy profile $\{\mu_i \sigma_C + (1 - \mu_i) \sigma_D\}_{i=1,\ldots,n}$ in the initial period. We do not explicitly show this dependence on the initial level of mixture $\{\mu_i\}_{i=1,\ldots,n}$ as it is clear from the context. Let $V(\sigma_i, m_i : \delta, p)$ be player i's discounted average payoff when m_i other players are playing σ_D and $n - m_i - 1$ other players are playing σ_C. Player i's strategy σ_i in $V(\sigma_i, m_i : \delta, p)$ can be a partial continuation strategy or a complete strategy. We use the following notations to simplify the exposition.

$$\begin{aligned} V\left(\sigma_i, \boldsymbol{\mu}_{-i} : \delta, p\right) &= \sum_{m_i=0}^{n-1} V(\sigma_i, m_i : \delta, p) \, \boldsymbol{\mu}_{-i}(m_i) \\ \triangle V(m_i : \delta, p) &= V(\sigma_C, m_i : \delta, p) - V(\sigma_D, m_i : \delta, p). \end{aligned}$$

In the two player case, M was essentially $\{0, 1\}$ and the belief on M was 1-dimensional. Here $M = \{0, 1, ..., n-1\}$ and each player's belief is $n-1$ dimensional. However, there is a convenient way to summarize the relevant information. We classify M into two sets; $\{0\}$ and $\{1, ..., n-1\}$, that is, the state where no other player have started playing σ_D and the state where there is at least one player who has already switched to σ_D. Player i's subjective probability of the event $m_i = 0$ is denoted by $\phi_i = \phi\left(\boldsymbol{\mu}_{-i}\right) = \boldsymbol{\mu}_{-i}(0)$. It is clear that this index captures the effective level of cooperation among the other players well. The exact number of players who are playing σ_D does not make much difference to what will happen in the future given that everyone is playing the grim trigger strategies. As soon as someone starts playing σ_D, every other player starts playing σ_D with very high probability from the very next period if the monitoring is almost perfect. What is important is not how many players have switched to σ_D, but whether anyone has switched to σ_D or not.

In the two player case, the difference in payoffs by σ_C and σ_D is linear and there is a unique $\pi(\delta, \varepsilon, \xi)$ where a player is indifferent between σ_C and σ_D. When there are more than two players, the equation $V\left(\sigma_C, \boldsymbol{\mu}_{-i} : \delta, p\right) - V\left(\sigma_D, \boldsymbol{\mu}_{-i} : \delta, p\right) = 0$ becomes slightly more complicated. For example, when all players are playing σ_C and σ_D with probability $(\mu, 1-\mu)$, $\boldsymbol{\mu}_{-i}(m_i) = (1-\mu)^{m_i} \mu^{n-1-m_i} \binom{n-1}{m_i}$ for $m_i = 0, ..., n-1$. So,

$$
\begin{aligned}
& V\left(\sigma_C, \boldsymbol{\mu}_{-i} : \delta, p\right) - V\left(\sigma_D, \boldsymbol{\mu}_{-i} : \delta, p\right) \\
= & \sum_{m_i=0}^{n-1} (1-\mu)^{m_i} \mu^{n-1-m_i} \binom{n-1}{m_i} \triangle V(m_i : \delta, p) = 0
\end{aligned}
$$

is an $n-1$ degree polynomial equation in $\mu \in [0,1]$. However, the following lemma holds as in the two player case, whose proof is left in the appendix.

Lemma 4.1. *If $\delta > \frac{g(0)}{1+g(0)}$, there exists a unique $\pi(\delta, p_\varepsilon) \in (0,1)$ for a small ε such that $\sum_{m_i=0}^{n-1} (1-\pi)^{m_i} \pi^{n-1-m_i} \binom{n-1}{m_i} \triangle V(m_i : \delta, p_\varepsilon) = 0$, and $\pi(\delta, p_0) \to 1$ as $\delta \downarrow \frac{g(0)}{1+g(0)}$*

Let $\boldsymbol{\mu}_{-i}^{\pi}$ be player i's belief on M when each player is randomizing between σ_C and σ_D independently with $\pi(\delta, p_\varepsilon)$, and let $\phi_i^{\pi(\delta, p_\varepsilon)} = \phi\left(\boldsymbol{\mu}_{-i}^{\pi(\delta, p_\varepsilon)}\right)$.

We proved that the cut-off strategy ρ is optimal in the two player case by applying one shot deviation argument. The strategy which corresponds to ρ here is the following:

$$
(*) \begin{cases} \text{play } C \text{ if } V\left(\sigma_C, \boldsymbol{\mu}_{-i} : \delta, p_\varepsilon\right) > V\left(\sigma_D, \boldsymbol{\mu}_{-i} : \delta, p_\varepsilon\right) \\ \text{play } D \text{ if } V\left(\sigma_C, \boldsymbol{\mu}_{-i} : \delta, p_\varepsilon\right) < V\left(\sigma_D, \boldsymbol{\mu}_{-i} : \delta, p_\varepsilon\right) \end{cases}
$$

Unfortunately, the n-player case is a bit more complicated to apply one shot deviation argument. The belief $\boldsymbol{\mu}_{-i}$ is an $n-1$ dimensional object with n players while it was one dimensional in the two player case. Moreover, since private signals are very rich with many players, we need many belief operators to keep track of the belief after each distinct realization of private signals. These difficulties prevent us from simply applying the one shot deviation argument.

We instead apply the path dominance argument adopted in Sekiguchi [13] in the following. For some set of beliefs, we prove that C is the unique optimal action by showing that *any strategy which assigns D in the current period is dominated by a strategy which assigns C in the current period, namely σ_C*. Similarly, for another set of beliefs, we show that any strategy which starts with C is dominated by σ_D to prove that D is the unique optimal action for such beliefs. By applying this path dominance argument, we can show that the cut-off strategy described above indeed almost assigns the unique optimal action.

Proposition 4.2. *Given δ, for any $\eta > 0$, there exists $\varepsilon > 0$ such that for any p_ε,*

- *it is not optimal to play D for player i if $\boldsymbol{\mu}_{-i}$ satisfies*

$$V\left(\sigma_C, \boldsymbol{\mu}_{-i} : \delta, p_0\right) \geqq V\left(\sigma_D, \boldsymbol{\mu}_{-i} : \delta, p_0\right) + \eta$$

- *it is not optimal to play C for player i if $\boldsymbol{\mu}_{-i}$ satisfies*

$$V\left(\sigma_C, \boldsymbol{\mu}_{-i} : \delta, p_0\right) \leqq V\left(\sigma_D, \boldsymbol{\mu}_{-i} : \delta, p_0\right) - \eta$$

The proof roughly goes as follows. Suppose that

$$V\left(\sigma_C, \boldsymbol{\mu}_{-i} : \delta, p_0\right) \geqq V\left(\sigma_D, \boldsymbol{\mu}_{-i} : \delta, p_0\right) + \eta$$

and the monitoring is almost perfect. If player i plays D, then it will be observed by all the other players and trigger the collective punishments with very high probability. Hence, her total payoff is approximately $V\left(\sigma_D, \boldsymbol{\mu}_{-i} : \delta, p_0\right)$. On the other hand, similarly she can guarantee at least $V\left(\sigma_C, \boldsymbol{\mu}_{-i} : \delta, p_0\right)$ approximately by playing σ_C because $V\left(\sigma_C, \boldsymbol{\mu}_{-i} : \delta, p_\varepsilon\right) \to V\left(\sigma_C, \boldsymbol{\mu}_{-i} : \delta, p_0\right)$ as $\varepsilon \to 0$. Clearly, the latter payoff dominates the former for small ε.

We have a couple of remarks on this proposition. First, this argument extends the path dominance argument used in Sekiguchi [13] for $n = 2$ to the general $n-$player case. Moreover, it provides a sharper characterization of the unique optimal action even for $n = 2$. Second, we can show that we do not need η, that is, $(*)$ assigns the unique optimal action for the two player case. This implies that ρ indeed assigns not only an optimal action, but also the unique optimal action for all beliefs.

Although this proposition characterizes the unique optimal action in terms of $\boldsymbol{\mu}_{-i}$, it turns out that it is more useful to express the unique optimal action in terms of ϕ_i; the probability that everyone is playing σ_C. As we noted before, ϕ_i works as a good index to measure the degree of cooperation among the other players when they are using the grim trigger strategies. If δ is close to $\frac{g(0)}{1+g(0)}$, the unique optimal action can be expressed in terms of ϕ_i.

Proposition 4.3. *There exists* $\overline{\delta} \in \left(\frac{g(0)}{1+g(0)}, 1\right)$ *and for each* $\delta \in \left(\frac{g(0)}{1+g(0)}, \overline{\delta}\right)$, *there exists* $\overline{\phi}(\delta)$, $\underline{\phi}(\delta)$ *and* ε *for which the following holds; for any* p_ε,

- *it is not optimal to play* D *for player* i *if* $\phi_i \geqq \overline{\phi}(\delta)$
- *it is not optimal to play* C *for player* i *if* $\phi_i \leqq \underline{\phi}(\delta)$

  Moreover, $\overline{\phi}(\delta), \underline{\phi}(\delta) \to 1$ *as* $\delta \to \frac{g(0)}{1+g(0)}$.

Now we show that $\pi(\delta, p_\varepsilon)\sigma_C + (1 - \pi(\delta, p_\varepsilon))\sigma_D$ is a symmetric Nash equilibrium for a middle range of discount factors and it approximates the efficiency as $\delta \to \frac{g(0)}{1+g(0)}$. Note that, once we find a Nash equilibrium, the existence of a realization equivalent sequential equilibrium is guaranteed because any deviation is not observable to the other players.[8] Also note that we can use a public randomization device or divide the game into component games to approximate the efficiency for a large δ as before.

For $\delta \in \left(\frac{g(0)}{1+g(0)}, \overline{\delta}\right)$,we know that there are three numbers $\overline{\phi}(\delta)$, $\underline{\phi}(\delta)$, and $\phi_i^{\pi(\delta,p_\varepsilon)} \in \left(\underline{\phi}(\delta), \overline{\phi}(\delta)\right)$, all of which are close to 1. We also know that, if the monitoring is almost perfect, player i's unique optimal action is C when ϕ_i is higher than $\overline{\phi}(\delta)$, D when ϕ_i is lower than $\underline{\phi}(\delta)$, and player i is indifferent when $\phi_i = \phi^{\pi(\delta,p_\varepsilon)}$. So, in order to show that $\pi(\delta, p_\varepsilon)\sigma_C + (1 - \pi(\delta, \underline{p}_\varepsilon))\sigma_D$ is a Nash equilibrium, we only need to show that ϕ_i is always above $\overline{\phi}(\delta)$ after $((C, \mathbf{c}), ..., (C, \mathbf{c}))$ and falls below $\underline{\phi}(\delta)$ when d is first observed or D is being played by player i herself.

First, suppose that player i observed the cooperative profile.$\left(\omega_i^t = \mathbf{c}\right)$ in period t. With probability at least $\phi_i^{t-1}(1-\varepsilon)$, all players were cooperative at period $t-1$ and observed cooperative signals. With probability at most $\left(1 - \phi_i^{t-1}\right)\varepsilon$, some players were already playing σ_D and player i observed $\omega_i^t = \mathbf{c}$ by accident. Then, it is not difficult to see that ϕ_i^t is larger than $\frac{\phi_i^{t-1}(1-\varepsilon)}{\phi_i^t + \left(1-\phi_i^{t-1}\right)\varepsilon}$ by Bayes Rule. Let $f(\phi) = \frac{\phi(1-\varepsilon)}{\phi+(1-\phi)\varepsilon}$. If ε is small enough, then $f(\phi^\pi) > \overline{\phi}(\delta)$. This implies that ϕ_i^2 is above $\overline{\phi}(\delta)$ given that $(C, \mathbf{c})$ was observed in the first period. Moreover, if

[8] See Sekiguchi [13, Proposition 3]

$\phi_i^t > \overline{\phi}(\delta)$, we can obtain $\phi_i^{t+1} \geq f(\phi_i^t) > f(\overline{\phi}(\delta)) > f(\phi^\pi) > \overline{\phi}(\delta)$ by the strict monotonicity of f. Hence, ϕ_i^t is always strictly above $\overline{\phi}(\delta)$ by induction when the cooperative profile has been observed.

In order to show that ϕ_i falls below $\underline{\phi}(\delta)$ after the other relevant histories, we need an assumption. The following assumption is too strong for our purpose, but it is easy to describe and satisfied in both the two player case and the example in the beginning of this section.

Assumption B

The probability of a signal profile only depends on the number of errors contained in the signal profile.

For example, when everyone is playing C, the probability that a player observes two d signals while the other players observe correct signals is equal to the probability that two players observe one d while the rest of the players observe correct signals.[9]

Now consider the dynamics of player i's belief after the other relevant histories under Assumption B. First, consider the history where player i observes some defection for the first time. If this occurs in the first period, player i interprets it as a signal of σ_D being chosen in the first period rather than as an observation error if ε is small. Hence, ϕ_i^2 falls below $\underline{\phi}(\delta)$. Suppose next that this kind of history is reached in period t. Also suppose that the number of players is three for simplicity and player 1 observes 1 defection by player 2. Player 1 can interpret this as there being an error only in her private signals. This is the only case where everyone is still cooperative. On the other hand, it is equally likely that player 2's observation contained 1 error in the last period (observed d even if every player played C) and player 1's current signal is indeed 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 for any other possible histories involving more than one observation error, ϕ_i^{t+1} is at most $\frac{1}{3}$. In general, ϕ_i^{t+1} is at most $\frac{1}{(n-1)^m+1}$ when there are n players and m defections are observed. As the number of players increases, it becomes easier to interpret such history in a pessimistic way.

Second, consider the history where player i has already started defection. Suppose that all players but player i have been cooperative until period t. Also suppose again that the number of players is three and $i = 1$ for the sake of simplicity. For everyone to be still cooperative after the current period, all players

[9] With this assumption, it can be shown that $(*)$ indeed assigns the unique optimal action, hence constitutes a sequential equilbirum.

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. There are other events where some player switches to σ_D from the next period on. For example, player 2 may have observed the correct signal d about player 1 and the wrong signal d about player 3. This event contains the same number of errors, hence occurs with the same probability under Assumption B. Since there are five such events, ϕ_i^{t+1} is at most $\frac{1}{6}$ even though we assumed that $\phi_i^t = 1$. If $\phi_i^t < 1$, ϕ_i^{t+1} is smaller after such histories. When there are $n\,(> 3)$ players and player i played D and observed k defections, it is again straightforward to show that ϕ_i^{t+1} is even much smaller.[10]

The following lemma summarizes these arguments.

Lemma 4.4. *Suppose that every player plays $(1-\pi)\,\sigma_C + \pi\sigma_D$ with $\pi \in (0,1)$ in the first period and Assumption B is satisfied. Then, there exists $\varepsilon > 0$ such that for any p_ε,*

- $\phi_i^t > \overline{\phi}\,(\delta)$ *after* $h_i^t = ((C,\mathbf{c})\,, ..., (C,\mathbf{c}))$.
- $\phi_i^t < \frac{1}{n}$ *after histories such as*
  - $h_i^t = \left(h_i^{t-1} = ((C,\mathbf{c})\,, ..., (C,\mathbf{c}))\,, \left(C,\omega_i^t\right)\ \right)$ *with* $\omega_i^t \neq \mathbf{c}$
    or
  - $h_i^t = \left(D,\omega_i^t\right)$

for all i and $t = 1, 2, ...$

Combining Lemma 4.1, Proposition 4.3, and Lemma 4.4 (and a remark after Proposition 4.3) together, we obtain the following theorem.

Theorem 4.5. *Suppose that Assumption B is satisfied. For any $x < 1$, there exists $\varepsilon(\delta) > 0$ such that there exists a symmetric sequential equilibrium with payoff greater than x for any $p_{\varepsilon(\delta)}$, provided that either (i) δ is sufficiently close to 1 or (ii) $\delta > \frac{g(0)}{1+g(0)}$ and a public randomization device is available.*

[10] This last argument is particular to the case with three or more players. As shown before, $\chi_{D\omega_i}(\mu)$ is lower than μ for the two player case. This guarantees that μ is below $\underline{\phi}\,(\delta)$ once player i starts playing D.

5. Concluding Comments

This paper has developed "belief-based" strategies as a way of constructing sequential equilibria in repeated games with private monitoring. This affords a major simplification as compared to the traditional method of analysis. While our construction has been restricted to the prisoners' dilemma, and to a mixed trigger strategy profile which consists only of two continuation strategies, the idea underlying this simplification is generalizable. If player i starts with a finitely complex (mixed) strategy which induces k possible continuation strategies, then the state space or the set of possible beliefs for player j for the entire repeated game is a $k-1$ dimensional simplex.

The approach of the present paper is based on generalizing "trigger strategy" equilibria to the private monitoring context. Under perfect or imperfect public monitoring, such trigger strategy can be constructed so as to provide strict incentives for players to continue with their equilibrium actions at each information set. Mailath and Morris [9] show that one can construct equilibria which provide similar strict incentives under private monitoring which is "almost-public". However, if private signals are not sufficiently correlated, pure trigger strategy profiles fail to be equilibria.

The approach in the present paper, as in previous works such as Bhaskar and van Damme [3] and Sekiguchi [13], relies on approximating the grim trigger strategy with a mixed strategy. In the basic construction, a player is indifferent between cooperating and defecting in the initial period, but has strict incentives to play the equilibrium action at every subsequent information set. In particular, player i's behavior strategy is measurable with respect to her beliefs about player j's continuation strategy. As Bhaskar [2] shows, such mixed strategies are robust to a small amount of incomplete payoff information as in Harsanyi's [7]. In particular, Bhaskar [2] shows in the context of the repeated prisoners' dilemma, where stage game payoffs are random and private, there exists a strict equilibrium with behavior corresponding to that of the mixed equilibrium of the present paper.[11] In the initial period, a player plays C for some realizations of his private payoff information, and D for other realizations, and continues with a trigger strategy in subsequent periods, independent of their payoff information.

The alternative approach to constructing non-trivial repeated game equilibria with private monitoring is due to Piccione [12] and Ely and Välimäki [6].[12] This

[11] This result is relevant since Bhaskar [1] considers a overlapping generations game with private monitoring and shows that incomplete payoff information as in Harsanyi implies an anti-folk theorem — players must play Nash equilibrium of the stage game in every period.

[12] See also the work of Kandori [8] in the context of a finitely repeated game.

approach relies on using player j's mixed strategy to make a player i indifferent between playing C and D at every information set. Since player i is so indifferent, she is likewise willing to randomize so as to make j also indifferent between his actions at each information set. In this approach, beliefs are irrelevant, since a player's continuation payoff function does not depend upon her continuation strategies. Such equilibria seem to be less likely to survive if there is private payoff information, and indeed this question is the subject of current research.

Appendix.

Proof of Lemma 4.1

Proof. Let $f_\varepsilon(\mu,\delta) = \sum_{m_i=0}^{n-1} (1-\mu)^{m_i} \mu^{n-1-m_i} \binom{n-1}{m_i} \triangle V(m_i : \delta, p_\varepsilon)$. Since $f_\varepsilon(1,\delta) > 0$ (by $\delta > \frac{g(0)}{1+g(0)}$) for small ε, $f_\varepsilon(0,\delta) < 0$ and f is continuous, existence of such π is guaranteed. To prove uniqueness, we show $f_\varepsilon(\mu,\delta) = 0 \Rightarrow \frac{\partial f_\varepsilon(\mu,\delta)}{\partial\mu} > 0$. Suppose that $f_\varepsilon(\mu,\delta) = 0$. Then,

$$\begin{aligned}
&\frac{\partial f_\varepsilon(\mu,\delta)}{\partial\mu} \\
&= \sum_{m_i=0}^{n-1} \left\{ \begin{array}{c} (1-\mu)^{m_i}(n-1-m_i)\mu^{n-2-m_i} \\ -m_i(1-\mu)^{m_i-1}\mu^{n-1-m_i} \end{array} \right\} \binom{n-1}{m_i} \triangle V(m_i : \delta, p_\varepsilon) \\
&= -\sum_{m_i=0}^{n-1} m_i(1-\mu)^{m_i}\mu^{n-1-m_i}\left(\frac{1}{\mu}+\frac{1}{1-\mu}\right)\binom{n-1}{m_i} \triangle V(m_i : \delta, p_\varepsilon)
\end{aligned}$$

(by $f_\varepsilon(\mu,\delta) = 0$)

Since $\triangle V(m_i : \delta, p_0) \to -(1-\delta)g(m_i) < 0$ as $\varepsilon \to 0$ for $m_i = 1,..,n-1$, $\frac{\partial f_\varepsilon(\mu,\delta)}{\partial\mu} > 0$ if ε is small enough.

Next, we prove that $\pi(\delta, p_0) \to 1$ as $\delta \downarrow \frac{g(0)}{1+g(0)}$. When $\delta = \frac{g(0)}{1+g(0)}$, $\pi(\delta,p_0) = 1$ is the solution of the equation:

$$\sum_{m_i=0}^{n-1} (1-\mu)^{m_i}\mu^{n-1-m_i}\binom{n-1}{m_i} \triangle V(m_i : \delta, p_0) = 0$$

We just need to show that $\frac{\partial\pi(\delta,p_0)}{\partial\delta}|_{\delta=\frac{g(0)}{1+g(0)}} < 0$ using the implicit function theorem. Since

$$f_0(\mu,\delta) = \sum_{m_i=0}^{n-1} (1-\mu)^{m_i}\mu^{n-1-m_i}\binom{n-1}{m_i} \triangle V(m_i : \delta, p_0)$$

$$= -(1-\delta)\sum_{m_i=0}^{n-1}(1-\mu)^{m_i}\mu^{n-1-m_i}\binom{n-1}{m_i}g(m_i)+\delta\mu^{n-1}$$

a straightforward calculation gives the desired result as follows.

$$\begin{aligned}\frac{\partial\pi(\delta,p_0)}{\partial\delta}\Big|_{\delta=\frac{g(0)}{1+g(0)}} &= -\frac{\frac{\partial f_\varepsilon(\mu,\delta)}{\partial\delta}}{\frac{\partial f_\varepsilon(\mu,\delta)}{\partial\mu}}\Big|_{\delta=\frac{g(0)}{1+g(0)},\mu=1}\\ &= -\frac{1+g(0)}{(1-\delta)(n-1)(g(1)-g(0))+\delta(n-1)}\Big|_{\delta=\frac{g(0)}{1+g(0)}}\\ &= -\frac{1}{n-1}\frac{(1+g(0))^2}{g(1)}<0\end{aligned}$$

■

Proof of Proposition 4.2

We first prove a lemma. Let $\overline{V}=\max_{a,m}u(a,m)$ and $\underline{V}=\min_{a,m}u(a,m)$.

Lemma 5.1. *If $\boldsymbol{\mu}_{-i}(m_i)=1$ for any $m_i\neq 0$, then there exists $\varepsilon>0$ such that σ_D maximizes $V(\sigma_i,\boldsymbol{\mu}_{-i}:\delta,p_\varepsilon)$ for any p_ε.*[13]

Proof. Suppose that $\boldsymbol{\mu}_{-i}(m_i)=1$ for any $m_i\neq 0$. Take σ_D and any strategy which starts with C. The least deviation gain in the current period is $(1-\delta)\underline{g}$. The largest loss by playing σ_D in terms of the continuation payoffs is $\delta\varepsilon\overline{V}$. Setting ε small enough guarantees that $(1-\delta)\underline{g}>\delta\varepsilon\overline{V}$. Then, D must be the optimal action for any p_ε. Since $\boldsymbol{\mu}_{-i}(m_i)$ is always 1 after this period, D is the unique optimal action in all the following periods. Hence, σ_D is the optimal strategy.■

Let $q(m''|m')$ be a probability that m'' players will play σ_D from the next period when m' players are playing σ_D now. In other words, this $q(m''|m')$ is a probability that $m''-m'$ players playing C observe the signal d when $n-m'$ players play C and m' players play D. By definition, $q(m''|m')>0$ for $m''\geqq m'$ and $q(m''|m')=0$ for $m''<m'$. 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 5.2. *1.* $V(\sigma_C,0:\delta,p_\varepsilon)\geqq\frac{(1-\delta)+\delta\varepsilon\underline{V}}{1-\delta(1-\varepsilon)}$ *for any p_ε.*

[13] What this lemma means is that any strategy which is realization equivalent to σ_D is the maximizer.

2. *Given* $\delta \in \left(\frac{g(0)}{1+g(0)}, 1\right)$, *There exists* $\varepsilon > 0$ *such that for any* p_ε,

$\max_{\sigma_i} V(\sigma_i, 0 : \delta, p_\varepsilon) \leqq \frac{1-\delta+\delta\varepsilon\overline{V}}{1-\delta(1-\varepsilon)}$

Proof. (1): For any p_ε,

$$V(\sigma_C, 0 : \delta, p_\varepsilon) \geqq (1-\delta) + \delta q(0|0) V(\sigma_C, 0 : p_\varepsilon, \delta) + \delta(1 - q(0|0)) \underline{V}$$

So,

$$V(\sigma_C, 0 : \delta, p_\varepsilon) \geqq \frac{(1-\delta) + \delta(1-q(0|0))\underline{V}}{1-\delta q(0|0)} \geqq \frac{(1-\delta)+\delta\varepsilon\underline{V}}{1-\delta(1-\varepsilon)}$$

(2): Given $\delta \in \left(\frac{g(0)}{1+g(0)}, 1\right)$, it is easy to check that $V(\sigma_C, 0 : \delta, p_0) > V(\sigma_D, 0 : \delta, p_0)$. Pick ε small enough such that (i) $V(\sigma_C, 0 : \delta, p_\varepsilon) > V(\sigma_D, 0 : \delta, p_\varepsilon)$ for any p_ε and (ii) Lemma 5.1 holds. Let σ_i^* be the optimal strategy when everyone plays σ_C.[14] Suppose that σ_i^* assigns D for the first period. Then for any p_ε,

$$\begin{aligned} V(\sigma_i^*, 0 : \delta, p_\varepsilon) \quad & \leqq (1-\delta) u(D, 0) + \\ & \delta \left\{ q(1|1) V(\sigma_i^*, 0 : \delta, p_\varepsilon) + \textstyle\sum_{m=2}^n q(m|1) V(\sigma_D, m-1 : \delta, p_\varepsilon) \right\} \end{aligned}$$

In RHS of 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_i^* leads to a higher expected payoff than $V(\sigma_i^*, 0 : \delta, p_\varepsilon)$ or $V(\sigma_D, m-1 : \delta, p_\varepsilon)$ at the corresponding states, then this contradicts the optimality of σ_i^* or σ_D by Lemma 5.1. Hence this inequality should hold.

Then, for any p_ε,

$$\begin{aligned} V(\sigma_i^*, 0 : \delta, p_\varepsilon) \quad & \leqq \quad \frac{(1-\delta) u(D, 0) + \delta \sum_{m=2}^n q(m|1) V(\sigma_D, m-1 : \delta, p_\varepsilon)}{1 - \delta q(1|1)} \\ & = \quad V(\sigma_D, 0 : \delta, p_\varepsilon) \\ & < \quad V(\sigma_C, 0 : \delta, p_\varepsilon) \end{aligned}$$

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

Then,

$$V(\sigma_i^*, 0 : \delta, p_\varepsilon) \leqq (1-\delta) + \delta q(0|0) V(\sigma_i^*, 0 : \delta, p_\varepsilon) + \delta(1 - q(0|0)) \overline{V}$$

[14] Such σ_i^* exists because the strategy space is a compact space in product topology, on which discounted average payoff functions are continuous.

hence,

$$V\left(\sigma_i^*, 0:\delta, p_\varepsilon\right) \leqq \frac{(1-\delta)+\delta\left(1-q\left(0|0\right)\right)\overline{V}}{1-\delta q\left(0|0\right)} \leqq \frac{(1-\delta)+\delta\varepsilon\overline{V}}{1-\delta\left(1-\varepsilon\right)}$$

This implies that $\max_{\sigma_i} V\left(\sigma_i, 0: p_\varepsilon, \delta\right) \leqq \frac{1-\delta+\delta\varepsilon\overline{V}}{1-\delta(1-\varepsilon)}$ for any p_ε.■

(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 σ_C. (2) means that there is not much to be exploited by using other strategies than σ_C with almost perfect private monitoring as long as all the other players are using σ_C.

With this Lemma, we can prove Proposition 4.2.

Proof. (1): Given player i's belief $\boldsymbol{\mu}_{-i}$, the current period gain by playing D is $g\left(\boldsymbol{\mu}_{-i}\right)$. The continuation payoff by σ_D is at most $\delta\varepsilon\overline{V}$, while the continuation payoff from playing σ_C is at least

$$\delta\left[\phi\left(\boldsymbol{\mu}_{-i}\right)\left\{(1-\varepsilon)V\left(\sigma_C, 0:\delta, p_\varepsilon\right)+\varepsilon\underline{V}\right\}+\left(1-\phi\left(\boldsymbol{\mu}_{-i}\right)\right)\varepsilon\underline{V}\right]$$

Hence, it is not optimal to play D if

$$\begin{aligned}&(1-\delta)\,g\left(\boldsymbol{\mu}_{-i}\right)\\ <\;&\delta\left[\phi\left(\boldsymbol{\mu}_{-i}\right)\left\{(1-\varepsilon)V\left(\sigma_C, 0:\delta, p_\varepsilon\right)+\varepsilon\underline{V}\right\}+\left(1-\phi\left(\boldsymbol{\mu}_{-i}\right)\right)\varepsilon\underline{V}-\varepsilon\overline{V}\right]\end{aligned}$$

is satisfied because then any strategy which plays D now is dominated by σ_C.

By Lemma 5.2.1., this inequality is satisfied for any p_ε if

$$\begin{aligned}&(1-\delta)\,g\left(\boldsymbol{\mu}_{-i}\right)\\ <\;&\delta\left[\phi\left(\boldsymbol{\mu}_{-i}\right)\left\{(1-\varepsilon)\frac{1-\delta+\delta\varepsilon\underline{V}}{1-\delta\left(1-\varepsilon\right)}+\varepsilon\underline{V}\right\}+\left(1-\phi\left(\boldsymbol{\mu}_{-i}\right)\right)\varepsilon\underline{V}-\varepsilon\overline{V}\right]\end{aligned}$$

Since RHS converges to $\delta\phi\left(\boldsymbol{\mu}_{-i}\right)$ as $\varepsilon\to 0$, if $\boldsymbol{\mu}_{-i}$ satisfies $\delta\phi\left(\boldsymbol{\mu}_{-i}\right) \geqq (1-\delta)\,g\left(\boldsymbol{\mu}_{-i}\right)+\eta$ for some $\eta>0$, which is equivalent to $V\left(\sigma_C, \boldsymbol{\mu}_{-i}:\delta, p_0\right) \geqq V\left(\sigma_D, \boldsymbol{\mu}_{-i}:\delta, p_0\right)+\eta$, then there exists a $\varepsilon'\left(\delta,\eta,\boldsymbol{\mu}_{-i}\right)$ and a neighborhood $B\left(\boldsymbol{\mu}_{-i}\right)$ of $\boldsymbol{\mu}_{-i}$ such that D is not optimal for any $p_{\varepsilon'\left(\delta,\eta,\boldsymbol{\mu}_{-i}\right)}$ and $\boldsymbol{\mu}'_{-i}\in B\left(\boldsymbol{\mu}_{-i}\right)$. This $\varepsilon'\left(\delta,\eta,\boldsymbol{\mu}_{-i}\right)>0$ can be set independent of $\boldsymbol{\mu}_{-i}$ by the standard argument because $\left\{\boldsymbol{\mu}_{-i}|\delta\phi\left(\boldsymbol{\mu}_{-i}\right)\geqq(1-\delta)\,g\left(\boldsymbol{\mu}_{-i}\right)+\eta\right\}$ is a compact subset in a $n-1$ dimensional simplex.

(2): Similarly, it is not optimal to play C if

$$\begin{aligned}&(1-\delta)\,g\left(\boldsymbol{\mu}_{-i}\right)\\ >\;&\delta\left[\phi\left(\boldsymbol{\mu}_{-i}\right)\left\{(1-\varepsilon)\max_{\sigma_i}V\left(\sigma_i, 0:\delta, p_\varepsilon\right)+\varepsilon\overline{V}\right\}+\left(1-\phi\left(\boldsymbol{\mu}_{-i}\right)\right)\varepsilon\overline{V}\right]\end{aligned}$$

because then any strategy which plays C now is dominated by σ_D.

By Lemma 5.2.2., this inequality is satisfied for any p_ε if

$$\begin{aligned} & (1-\delta)\, g\left(\boldsymbol{\mu}_{-i}\right) \\ > \;& \delta\left[\phi\left(\boldsymbol{\mu}_{-i}\right)\left\{(1-\varepsilon)\frac{1-\delta+\delta\varepsilon\overline{V}}{1-\delta(1-\varepsilon)}+\varepsilon\overline{V}\right\}+\left(1-\phi\left(\boldsymbol{\mu}_{-i}\right)\right)\varepsilon\overline{V}\right] \end{aligned}$$

Since RHS converges to $\delta\phi\left(\boldsymbol{\mu}_{-i}\right)$ as $\varepsilon \to 0$, if $\boldsymbol{\mu}_{-i}$ satisfies $\delta\phi\left(\boldsymbol{\mu}_{-i}\right) \leqq (1-\delta)\, g\left(\boldsymbol{\mu}_{-i}, p_0\right) - \eta$ for some $\eta > 0$, then there exists a $\varepsilon''\left(\delta,\eta,\boldsymbol{\mu}_{-i}\right)$ such that C is not optimal for any $p_{\varepsilon''\left(\delta,\eta,\boldsymbol{\mu}_{-i}\right)}$ and any $\boldsymbol{\mu}'_{-i}$ close to $\boldsymbol{\mu}_{-i}$. Again, $\varepsilon''\left(\delta,\eta,\boldsymbol{\mu}_{-i}\right)$ can be set independent of $\boldsymbol{\mu}_{-i}$.

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

Proof of Proposition 4.3

Proof. Fix any $\delta > \frac{g(0)}{1+g(0)}$. Then, since $\delta > (1-\delta)\, g(0) + \eta'$ for some positive number η', there exists $\overline{\phi}(\delta)$ such that for any $\phi_i \geq \overline{\phi}(\delta)$, $\delta\phi_i > (1-\delta)\, g\left(\boldsymbol{\mu}_{-i}\right) + \eta'$, where $g\left(\boldsymbol{\mu}_{-i}\right) = \sum_{m_i=0}^{n-1} \boldsymbol{\mu}_{-i}(m_i)\, g(m_i)$. Note that if $\phi_i = \phi\left(\boldsymbol{\mu}_{-i}\right)$ is close to 1, $g\left(\boldsymbol{\mu}_{-i}\right)$ is close to $g(0)$. Proposition 4.2 implies that it is not optimal to play D for any such $\phi_i \geq \overline{\phi}(\delta)$ if ε is small enough.

Next, by definition

$$V\left(\sigma_C, \boldsymbol{\mu}_{-i}^{\pi(\delta,p_0)} : \delta, p_0\right) - V\left(\sigma_D, \boldsymbol{\mu}_{-i}^{\pi(\delta,p_0)} : \delta, p_0\right) = 0 \tag{5.1}$$

Consider any $\boldsymbol{\mu}_{-i}$ for which $\boldsymbol{\mu}_{-i}^{\pi(\delta,p_0)}(0) - \kappa \geq \phi_i\left(\boldsymbol{\mu}_{-i}\right)$ for some $\kappa > 0$. There are three effects which this change of $\boldsymbol{\mu}_{-i}$ has on the LHS of (5.1). First, it reduces LHS by at least $\kappa\triangle V(0 : \delta, p_0) < 0$ for $m_i = 0$. Second, this probability κ should be moved on the states $m_i = 1, .., n-1$. Since $\triangle V(m_i : \delta, p_0) \to -(1-\delta)\, g(m_i)$ as $\varepsilon \to 0$ for $\theta = 1, .., n$, it reduces LHS at least by $\kappa(1-\delta)\underline{g}$. Finally, the probability $\left(1 - \boldsymbol{\mu}_{-i}^{\pi(\delta,p_0)}(0)\right)$ can be redistributed among $m_i = 1, .., n-1$ in any way. This increases LHS at most by $\left(1 - \boldsymbol{\mu}_{-i}^{\pi(\delta,p_0)}(0)\right)(1-\delta)\left(\overline{g} - \underline{g}\right)$. Note that $\triangle V(0 : \delta, p_0) \to 0$ and $\boldsymbol{\mu}_{-i}^{\pi(\delta,p_0)}(0) \to 1$ as $\delta \to \frac{g(0)}{1+g(0)}$. Hence, we can choose $\overline{\delta}$ in such a way that for any $\delta \in \left(\frac{g(0)}{1+g(0)}, \overline{\delta}\right)$, the second effect dominates the other effects, that is, $V\left(\sigma_C, \boldsymbol{\mu}_{-i} : \delta, p_0\right) - V\left(\sigma_D, \boldsymbol{\mu}_{-i} : \delta, p_0\right) \leq \eta''$ for any such $\boldsymbol{\mu}_{-i}$ for some positive η''. Define $\underline{\phi}(\delta) = \boldsymbol{\mu}_{-i}^{\pi(\delta,p_0)}(0) - k$. Then, Proposition 4.2 implies that it is not optimal to play C for any $\phi_i \leq \underline{\phi}(\delta)$ if ε is small enough. Since κ can be arbitrarily small as δ is chosen arbitrarily close to $\frac{g(0)}{1+g(0)}$, the proof is complete.■

References

[1] V. Bhaskar, Informational Constraints and the Overlapping Generations Model: Folk and Anti-Folk Theorems, *Rev. Econ. Studies,* 65 (1998) , 135-149.

[2] V. Bhaskar, The Robustness of Repeated Game Equilibria to Incomplete Payoff Information, mimeo, 2000.

[3] Bhaskar, V., and E. van Damme, Moral Hazard and Private Monitoring, mimeo, 1997.

[4] O. Compte, On Failing to Cooperate When monitoring is Private, mimeo, 1999.

[5] G. Ellison, Cooperation in the Prisoner's Dilemma with Anonymous Random Matching, *Rev. Econ. Studies,* 61 (1994), 567-588.

[6] J. C. Ely and J. Välimäki, A Robust Folk Theorem for the Prisoner's Dilemma, mimeo, 1999.

[7] J. Harsanyi, Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points, *Int. J. Game Theory* 2 (1973), 1-23.

[8] M. Kandori,Cooperation in Finitely Repeated Games with Imperfect Private Information, mimeo, 1991.

[9] G.Mailath, and S. Morris, Repeated Games with Almost-Public Monitoring, CARESS Working Paper #99-09, 1999.

[10] I. Obara, Repeated Prisoners' Dilemma with Private Monitoring: an N-player Case, CARESS Working Paper #99-13, 1999.

[11] I. Obara, Private Strategy and Efficiency: Repeated Partnership Games Revisited, mimeo, 1999.

[12] M. Piccione, The Repeated Prisoners' Dilemma with Imperfect Private Monitoring, mimeo, 1998.

[13] T. Sekiguchi, Efficiency in the Prisoners' Dilemma with Private Monitoring, *J. Econ. Theory,*76 (1997), 345-361.