CIRJE-F-246

Less is More: An Observability Paradox in Repeated Games

Michihiro Kandori
The University of Tokyo

Ichiro Obara
UCLA

November 2003

Less is More: An Observability Paradox in Repeated Games*

KANDORI, Michihiro†and OBARA, Ichiro‡
Faculty of Economics, University of Tokyo
and Department of Economics, UCLA

November 6, 2003

Abstract

We present a repeated prisoners' dilemma game with imperfect public monitoring, which exhibits the following paradoxical feature: the equilibrium payoff set expands and asymptotically achieves full efficiency as the public signal becomes less sensitive to the hiden actions of the players.

*We would like to thank Drew Fudenberg and David Levine for helpful comments and discussion.

†e-mail: kandori@e.u-tokyo.ac.jp
‡e-mail: iobara@econ.ucla.edu

1 Introduction

Repeated games with imperfect public monitoring provide an analytical framework to study long term relationships, where participants have imperfect public information about each other's hidden action. Now well-developed theory (including Abreu-Pearce-Stacchetti [1] and Fudenberg-Levine-Maskin [5], FLM hereafter) shows how the participants can utilize the imperfect information to achieve cooperation in such a setting. Our intuition suggests that, as the observability of actions improves, it should become easier to sustain cooperation. In fact, Kandori [6] has formalized this idea, by showing that the equilibrium payoff set expands, when observability improves according to the standard definition in the statistical decision theory (Blackwell and Girshick [2]). In this paper we point out subtlety involved in this issue and show that the above intuitive idea about the observability and cooperation should be taken with a pinch of salt. In particular, we present a repeated game with imperfect public monitoring, which has the following seemingly paradoxical feature: the equilibrium payoff set expands and asymptotically achieves full efficiency as the public information becomes *less* sensitive to the hidden actions of the players.

Our example is quite simple. We consider the standard prisoner's dilemma stage game with the following payoff table.

	C	D
C	$1,1$	$-h,1+d$
D	$1+d,-h$	$0,0$

where $d,h>0$ (D is dominant) and $d-h<1$ ((C,C) is efficient[1]). Actions are not observable[2], but the players publicly observe a signal $\omega \in \Omega \equiv \{X,Y\}$. We suppose that the probability of signal ω given action profile $a \in \{C,D\} \times \{C,D\}$, denoted $p(\omega|a)$, is symmetric (i.e., $p(\omega|a_1,a_2) = p(\omega|a_2,a_1)$) and satisfies

$$p(X \mid C,C) < p(X \mid C,D), \text{ and}$$

[1] This is the condition that (C,C) Pareto-dominates the public randomization between (D,C) and (C,D) with an equal probability.

[2] We assume that the realized payoff of a player is a function of his own action and the signal (so that a player cannot detect the opponent's action simply by looking at his realized payoff), and we denote it by $u_i(a_i,\omega)$. The payoff table represents the *expected* payoff $g_i(a) = \sum_{\omega\in\Omega} u_i(a_i,\omega)p(\omega|a)$, where $p(\omega|a)$ is the probability of ω given action profile a. Throughout the paper we hold the expected payoff fixed when we change the information structure $p(\omega|a)$. One can check that such an operation is possible with a suitable choice of $u_i(a_i,\omega)$.

$$p(Y \mid D, C) < p(Y \mid D, D).$$

That is, one defection makes X more likely, while an additional defection makes Y more likely.

We show that this familiar stage game exhibits the following paradoxical feature: As the signal becomes less sensitive to a deviation at the efficient point (C, C), (i.e., $p(X \mid C, C) - p(X \mid D, C) \to 0$), the asymptotic payoff set $\lim_{\delta \to \infty} E(\delta)$ expands and converges to the set of all feasible and individually rational payoffs. (Here we denote the (public perfect) equilibrium payoff set in the repeated game under discount factor δ by $E(\delta)$). Moreover, this can be true even if observability is reduced everywhere (not just at the efficient point): As $p(\omega|a)$ for all ω and a converges to one point (so that a deviation at any point becomes hard to detect), the asymptotic payoff set expands and converges to the set of all feasible and individually rational payoffs.

The mechanism operating behind our example is a familiar one (at least to the specialists in repeated games). Our point is to show that the familiar mechanism can operate in a subtle and disguised way. For this reason (and just for fun) we pause here for a moment and ask the reader if s/he can see how our example works. The answer is given in the next section and we show how the trick is done.

2 Intuitive Explanation

The key for our result is the efficiency of asymmetric punishment. When monitoring is imperfect, a "bad" outcome arises with a positive probability even if no player deviates. To provide incentives to cooperate, a punishment should be triggered in such an eventuality. If all players are simultaneously punished (as in the trigger strategy), welfare loss is inevitable. However, when each player's action affects the signal asymmetrically, we can use an asymmetric punishment scheme to avoid welfare loss; if player 1's deviation is suspected, we transfer 1's future payoff to player 2 (and *vice versa*). When the discount factor δ is close to 1, such a transfer can be made in a close vicinity of the Pareto frontier, and therefore the welfare loss vanishes as $\delta \to 1$. Furthermore, this conclusion is independent of how informative the signal is (as long as each player's deviation affects the signal asymmetrically). This is the basic driving force in the FLM folk theorem [5].

In our example, each player's action affects the signal symmetrically at the efficient pint (C, C) (each player's deviation makes X more likely), so that we are *un*able to use the efficient asymmetric punishment at this point. Hence, the above argument seemingly implies that efficiency cannot be achieved. However, consider a nearby point (D_ε, C), where D_ε denotes the mixed strategy which plays C and D with probabilities $(1-\varepsilon)$ and $\varepsilon > 0$ respectively. Player 1's defection makes X more likely at this point, as at (C, C). What about player 2's defection? The distribution of signal when player 2 takes action a_2 is given by

$$(1-\varepsilon)p(\omega|C, a_2) + \varepsilon p(\omega|D, a_2).$$

Now suppose that the first term is sufficiently insensitive to a_2. Then, player 2's action affects the signal mostly through the second term. As we have $p(Y|D, C) < p(Y|D, D)$, we conclude that player 2's defection makes Y more likely at this point. Hence, each player's action affects the signal asymmetrically at (D_ε, C), so that the efficient asymmetric punishment is feasible. Note that ε can be made arbitrarily small as the signal becomes insensitive at the efficient point (C, C). Hence, as the observability decreases, the asymmetric punishment becomes feasible at almost efficient point (D_ε, C) (and symmetrically at (C, D_ε)). This is the source of efficiency in our examples. Hence, the crux of the matter is the feasibility of asymmetric punishment and *not* the amount of information *per se*. The twist of our example is that the former is attained only when the observability of actions is decreased (in a broad sense).

Now let us make the above argument more precisely. At (D_ε, C), the change in the probability of Y when player 2 deviates is equal to

$$(1-\varepsilon)[p(Y|C, D) - p(Y|C, C)] + \varepsilon[p(Y|D, D) - p(Y|D, C)]. \tag{1}$$

If this is positive, player 2's defection at (D_ε, C) makes Y more likely, so that each player's action affects the signal asymmetrically, as we desire. The above expression (1) can be rearranged as

$$-(1-\varepsilon)\xi + \varepsilon\Delta$$

where $\xi \equiv p(X|C, D) - p(X|C, C) > 0$ and $\Delta \equiv p(Y|D, D) - p(Y|D, C) > 0$. Hence, (1) is positive if and only if

$$\varepsilon > \frac{\xi}{\xi + \Delta}.$$

Since (D_ε, C) is not an efficient profile, this lower bound for ε prevents us from obtaining full efficiency for a given $\xi > 0$. However, this lower bound vanishes when the observability of deviation at (C, C) is reduced (i.e., $\xi \to 0$). Furthermore, the same is true even when the observability of action is reduced everywhere (i.e., $\xi \to 0$ and $\Delta \to 0$) as long as Δ converges to 0 at a slower rate than ξ. Thus we can employ an asymmetric punishment while approximating the efficient action profile (C, C) arbitrarily closely as $\xi \to 0$ for such Δ.

Although this is the essence of our argument, we should address other non-trivial issues to obtain the formal proof. First, when supporting $((1-\varepsilon)C + \varepsilon D, C)$, player 1 must be indifferent between C and D. This indifference condition determines the level of transfer between 1 and 2 in the efficient asymmetric punishment. In general, this transfer may not be sufficient to deter player 2's deviation. We will show[3] that this does not happen when $\xi \to 0$ and $\xi/\Delta \to 0$. Second, we need to show that the continuation payoffs in the asymmetric punishment are also sustained by the repeated game equilibria. The detailed analysis is given in the next section.

Lastly, let us comment on the relationship to the folk theorem under imperfect public monitoring (FLM [5]). Our example does *not* satisfy the FLM conditions for the folk theorem for each $\xi > 0$ and neither in the limit $\xi \to 0$. Their pairwise full rank condition requires that there is at least one (potentially mixed) action profile α for which matrix

$$\begin{pmatrix} p(X|C,\alpha_2) & p(Y|C,\alpha_2) \\ p(X|D,\alpha_2) & p(Y|D,\alpha_2) \\ p(X|\alpha_1,C) & p(Y|\alpha_1,C) \\ p(X|\alpha_1,D) & p(Y|\alpha_1,D) \end{pmatrix}$$

has rank 3. Obviously, this is impossible, because this matrix has only two columns. However, our example *is* based on the fundamental idea of FLM, the efficiency of asymmetric punishment. The point here is that the pairwise full rank condition is a *sufficient* condition to facilitate the efficient asymmetric punishment, but it is not necessary. The essence of our example is that (i) the efficient asymmetric punishment is feasible at some points, even though the pairwise full rank condition fails, and (ii) those points can be arbitrarily close to the efficient point, as the observability is reduced. This suggests that we may obtain the folk theorem without the pairwise full rank condition. In fact, we show in Section 4 that we can modify this model to obtain a fairly simple example which satisfies the folk theorem without the pairwise full rank condition.

We present the formal statement of our result and its proof in the next section, and then offer some discussion (on the relationship to the folk theorem in Section 4, and on the relationship to Kandori [6] in the last section).

[3] A detailed verbal explanation is provided in Case 3 of the proof of Lemma 2.

3 Analysis

We denote player i's expected payoff given $a \in A = A_1 \times A_2$ by $g_i(a)$, and, by abusing notation, $g_i(\alpha)$ for player i's expected payoff given a mixed action profile $\alpha \in \triangle A_1 \times \triangle A_2$. Let V^* be the individually rational and feasible set, i.e. $V^* = \{v \in \mathrm{Co}g(A) | v \geq 0\}$, where $\mathrm{Co}g(A)$ denotes the convex hull of set $g(A) = \{v | v = g(a), a \in A\}$. Note that in the prisoners' dilemma game we consider, 0 is the minimax payoff for each player.

We are going to use Fudenberg and Levine's algorithm to compute the asymptotic equilibrium payoff set ([4]). Let us briefly summarize their method for readers' convenience. For a given welfare weight vector $\lambda \in \Re^2 \diagdown \{\mathbf{0}\}$ and a mixed action profile α, let $k^*(\alpha, \lambda)$ be the maximized value for the following optimization problem[4].

$$\begin{array}{ll} \max_{v_i, x_i(\cdot), i=1,2} & \lambda \cdot v \text{ subject to} \\ & v_i = g_i(a_i, \alpha_{-i}) + E[x_i(\omega) | a_i, \alpha_{-i}] \\ & \qquad \text{for } a_i \in \mathrm{supp}\alpha_i \\ & v_i \geq g_i(a_i, \alpha_{-i}) + E[x_i(\omega) | a_i, \alpha_{-i}] \\ & \qquad \text{for } a_i \notin \mathrm{supp}\alpha_i \\ & \lambda \cdot x(\omega) \leq 0 \text{ for each } \omega \end{array}$$

where $\mathrm{supp}\alpha_i$ is the support of α_i. Roughly speaking, the term x in the above optimization problem represents the variation in continuation payoffs, and the constraint $\lambda \cdot x(\omega) \leq 0$ ensures that the payoff variations lie in the equilibrium payoff set. Let $k^*(\lambda) = \sup_\alpha k^*(\alpha, \lambda)$ and $H(\lambda)$ be a half-space given by $\{v \in \Re^2 | \lambda \cdot v \leq k^*(\lambda)\}$. Let $Q = \cap_{\lambda \in \Re^2 \diagdown \{\mathbf{0}\}} H(\lambda)$, and denote the set of public perfect equilibrium[5] payoffs for discount factor δ by $E(\delta)$. Then, it is known that the following is true.

The Fudenberg-Levine Algorithm ([4])**:** $Q = \lim_{\delta \to 1} E(\delta)$ when Q has an interior point.

Let us parametrize the information structure by

$$\begin{aligned} \xi &\equiv p(X|C,D) - p(X|C,C) > 0, \text{ and} \\ \Delta(\xi) &\equiv p(Y|D,D) - p(Y|D,C) > 0, \end{aligned}$$

[4]This is a bit different form their exposition. We used the fact that $v = (1-\delta)g(\alpha) + \delta E[u(\omega)|\alpha]$ is equivalent to $v = g(\alpha) + E[x(\omega)|\alpha]$ for $x(\omega) = \frac{\delta}{1-\delta}(u(\omega) - v)$ (the former is obtained by the latter multiplied by $(1-\delta)$).

[5]A public perfect equilibirum is a sequential equilibrium, in which each player's action in any period depends only on the history of the publicly observable signal.

while $p(\omega|C,C)$ is fixed. To indicate explicitly the dependence of Q in the above result on the parameter ξ, we denote it by $Q(\xi)$. We examine how $Q(\xi)$ is affected by ξ. For each ξ, we can show that $Q(\xi)$ is a strict subset of V^*. In particular, it is bounded away from the efficient payoff profile $(1,1)$. However, we show that $Q(\xi)$ expands and supports almost all individually rational payoff profile *as* ξ *converges to* 0 *as long as* $\Delta(\xi)$ *is bounded away from 0 or converges to 0 at much slower rate.* Therefore the full efficiency is achieved only in the limit as $\xi \to 0$. One example of such information structure would be $p(X \mid C,C) = 1/2$ and $\Delta(\xi) = \xi + 1/4$ (so that $p(Y \mid D,D) = 3/4$ is constant). Another example would be $p(X \mid C,C) = 1/2$ and $\Delta(\xi) = \sqrt{\xi}$ (so that $p(\omega|a) \to 1/2$ for all ω and a, as $\xi \to 0$). Now we prove the following result. Recall that $E(\delta)$ is the set of public perfect equilibrium payoffs under discount factor δ. Note also that[6], for any $\delta' \in (0,1)$, $E(\delta')$ is contained in $\lim_{\delta\to 1} E(\delta)$.

Theorem 1 *For each* $\xi > 0$*, the equilibrium payoffs are bounded away from the efficient point* $(1,1)$*; there is a neighborhood of* $(1,1)$ *which lies outside of* $\lim_{\delta\to 1} E(\delta)$*. However, as the signal becomes less sensitive to actions* ($\xi \to 0$)*, the inefficiency vanishes and we can sustain all feasible and individually rational payoffs* ($\lim_{\delta\to 1} E(\delta) \to V^*$ *as* $\xi \to 0$)*, when* $\lim_{\xi\to 0} \frac{\xi}{\Delta(\xi)} = 0$.

The proof of this theorem is given by the Fudenberg-Levine Algorithm and the following Lemmata.

Lemma 1 $(1,1) \notin Q(\xi)$ *for any* $\xi > 0$.

Lemma 2 $\lim_{\xi\to 0} Q(\xi) = V^*$ *if* $\lim_{\xi\to 0} \frac{\xi}{\Delta(\xi)} = 0$.

Note that Lemma 1 immediately implies the first part of Theorem 1, because $Q(\xi)(= \lim_{\delta\to 1} E(\delta))$ is closed (as it is an intersection of closed half-spaces).

Proof of Lemma 1: To show this, it is sufficient to demonstrate $(1,1) \notin H(\lambda)$ for such λ that $\forall \alpha\ \lambda \cdot (1,1) \geq \lambda \cdot g(\alpha)$. Conceptually, this is quite simple. To sustain a point near $(1,1)$, the players need to play C with a sufficiently large probability, and this implies that each player's defection makes X more likely. This means that both players should be punished when X arises. As this entails welfare loss, it is impossible to sustain any

[6] See Fudenberg and Levine ([4]), Theorem 3.1.

point near $(1,1)$. This intuitive argument can be formulated as follows. Recall that $(1,1) \notin H(\lambda)$ is equivalent to

$$\lambda \cdot (1,1) > \sup \lambda \cdot (g(\alpha) + E[x(\omega)|\alpha]), \tag{2}$$

where the supremum is taken over α and x such that α is enforced by x which satisfies $\lambda \cdot x(\omega) \leq 0$ for all ω. Recall that D_ε denotes the mixed strategy which plays D with probability ε. Let $\varepsilon''(\xi) \equiv \frac{\xi}{\xi + \Delta(\xi)}$ be the minimum ε such that $P(X|D_\varepsilon, D) \leq P(X|D_\varepsilon, C)$, and let $M \equiv \{\alpha | \alpha_i(D) \geq \varepsilon''(\xi),\ i = 1,2\}$. When $\alpha \in M$, we have[7]

$$\lambda \cdot (1,1) > \max_{\alpha \in M} \lambda \cdot g(\alpha) \geq \lambda \cdot (g(\alpha) + E[x(\omega)|\alpha]).$$

Hence, (2) is proved if we show that there is also a constant K such that for all $\alpha \notin M$

$$\lambda \cdot (1,1) > K \geq \lambda \cdot (g(\alpha) + E[x(\omega)|\alpha]). \tag{3}$$

Note that, when $\alpha \notin M$, each player plays D with probability less than $\varepsilon''(\xi)$. This means that, by the definition of $\varepsilon''(\xi)$, any player's deviation makes X more likely, so that both players should be punished when X is realized. This entails a welfare loss, which determines the above bound K. Player 1's incentive constraint is

$$d \leq \{p(X|D,\alpha_2) - p(X|C,\alpha_2)\}(x_1(Y) - x_1(X)),$$

and by using this we can obtain an upper bound for $v_1 = g_1(\alpha) + E[x_1(\omega)|\alpha]$;

$$\begin{aligned} v_1 &= g_1(C,\alpha_2) + E[x_i(\omega)|C,\alpha_2] \\ &= g_1(C,\alpha_2) + x_1(Y) - p(X|C,\alpha_2)(x_1(Y) - x_1(X)) \\ &\leq g_1(C,\alpha_2) + x_1(Y) - \frac{d}{L^{\alpha_2} - 1} \end{aligned}$$

where $L^{\alpha_2} = \frac{p(X|D,\alpha_2)}{p(X|C,\alpha_2)} > 1$ is the likelihood ratio given α_2. Similarly, we can obtain a bound for $v_2 = g_2(\alpha) + E[x_2(\omega)|\alpha]$ as $v_2 \leq g_2(\alpha_1, C) + x_2(Y) - \frac{d}{L^{\alpha_1} - 1}$. These inequalities, together with the constraint $\lambda \cdot x(Y) \leq 0$ imply

$$\lambda \cdot v \leq \lambda_1 \left(g_1(C,\alpha_2) - \frac{d}{L^{\alpha_2} - 1} \right) + \lambda_2 \left(g_2(\alpha_1, C) - \frac{d}{L^{\alpha_1} - 1} \right)$$

$$\leq \lambda \cdot (1,1) - (\lambda_1 + \lambda_2) \frac{d}{L - 1}$$

[7] This is because we have $\max_\alpha \lambda \cdot g(\alpha) = \lambda \cdot (1,1)$ (by our choice of λ) and $E[x(\omega)|\alpha] \leq 0$.

where $L \equiv \frac{p(X|D,C)}{p(X|C,C)}$. The inequalities are implied by (i) $\lambda_i > 0$, $i = 1, 2$ and $1 \geq \max\{g_1(C, \alpha_2), g_2(\alpha_1, C)\}$ and (ii) $L \geq L^{\alpha_i} > 1$, $i = 1, 2$. Thus we have shown the required condition (3) with $K = \lambda \cdot (1, 1) - (\lambda_1 + \lambda_2)\frac{d}{L-1}$. **Q.E.D.**

Proof of Lemma 2: To establish the claim we will identify certain point(s) contained in the half-space $H(\lambda)$ for each direction $\lambda \neq 0$. This provides a subset of $Q(\xi)$, and we will show that this subset tends to V^* as $\xi \to 0$.

Case 1, $\lambda_1, \lambda_2 \leq 0$ and $\lambda \neq 0$: $(0, 0) \in H(\lambda)$.

This is true because Nash equilibrium $(0, 0)$ is sustained by $x_i \equiv 0$, $i = 1, 2$. Note that the above claim implies $\cap_{\lambda_1,\lambda_2 \leq 0, \lambda \neq 0} H(\lambda) \supset \Re^2_+ \equiv \{v \in \Re^2 | v_i \geq 0\}$.

Case 2, $\lambda_1 > 0, \lambda_2 \leq 0$: $g(D, C) \in H(\lambda)$.

This is established by showing that (D, C) is enforceable while $\lambda \cdot x(\omega) = 0$ is satisfied. This is an easy case, because any player's deviation at this point makes Y more likely, and $\lambda_1 > 0, \lambda_2 \leq 0$ implies that the sidepayment scheme x can punish both players when Y arises without any welfare loss[8] (i.e., $\lambda \cdot x(\omega) = 0$). The incentive constraints are

$$\begin{aligned} g_1(D, C) + E[x_1(\omega)|D, C] &\geq g_1(C, C) + E[x_1(\omega)|C, C] \\ g_2(D, C) + E[x_2(\omega)|D, C] &\geq g_2(D, D) + E[x_2(\omega)|D, D], \end{aligned}$$

which reduce to

$$\begin{aligned} d &\geq \xi(x_1(Y) - x_1(X)) \\ h &\leq \Delta(\xi)(x_2(X) - x_2(Y)). \end{aligned}$$

First, $\lambda \cdot x(\omega) = 0$ can be satisfied by setting $x_1(Y) = x_2(Y) = 0$, $x_1(X) = -\frac{\lambda_2}{\lambda_1} x_2(X)$. The incentive constraints are satisfied by taking $x_2(X)$ large enough, because $-\frac{\lambda_2}{\lambda_1}$ is non-negative.

[8] Precisely speaking, when $\lambda_2 = 0$, the efficient punishment condition $\lambda \cdot x \equiv 0$ requies that player 1 receives no punishment ($x_1 \equiv 0$). However, this is not a problem, as player 1 is taking a myopic best reply at (D, C).

Recall that D_ε denotes a mixed action profile where D is played with probability ε. With this notation, we show the following.

Case 3, $\lambda_1 \geq \lambda_2 > 0$: When $\frac{\xi}{\Delta(\xi)}$ is sufficiently small, we have

$$g\left(D_{\varepsilon'(\xi)}, C\right), g(D, C) \in H(\lambda)$$

for some $\varepsilon'(\xi) \in (0,1)$ such that $\varepsilon'(\xi) \to 0$ as $\frac{\xi}{\Delta(\xi)} \to 0$.

It is sufficient to show that, for all $\varepsilon \in [\varepsilon'(\xi), 1]$, (D_ε, C) is enforceable while satisfying $\lambda \cdot x(\omega) = 0$ (note that $(D_\varepsilon, C) = (D, C)$ for $\varepsilon = 1$). Before going into the formal proof, let us provide some intuition. We first discuss the case $\varepsilon \neq 1$ so that player 1 is mixing C and D. Calculation shows that, if $\varepsilon > \frac{\xi}{\xi+\Delta(\xi)}$, the profitable deviation (to D) by each player affects the signal *a*symmetrically. That is, defection by 1 makes X more likely, while defection by 2 makes Y more likely. Hence, to deter the profitable deviations, we may transfer payoff from 2 to 1 when Y arises, and this causes no welfare loss ($\lambda \cdot x(\omega) = 0$).

However, this is not the end of the story, because 1 should be *indifferent* between C and D. This indifference condition determines the magnitude of the transfer x, and it should be sufficiently large to deter player 2's defection. There are two factors which help to satisfy player 2's incentive constraint.

First, the signal must not be too insensitive to 2's defection. Recall that 2's action makes Y more likely if player 1's mixing probability ε is more than $\frac{\xi}{\xi+\Delta(\xi)}$. Because the signal becomes completely insensitive to 2's action when ε is exactly equal to $\frac{\xi}{\xi+\Delta(\xi)}$, the mixing probability must be bounded away from $\frac{\xi}{\xi+\Delta(\xi)}$. This bound is equal to $\varepsilon'(\xi)$ (to be determined in what follows).

Second, the magnitude of transfer should be large enough. Our condition $\xi \to 0$ achieves this goal[9]. When ξ is small, the signal becomes insensitive to player 1's defection, so that we need a large payoff transfer to keep him indifferent between C and D.

Finally, the case $\varepsilon = 1$ can be treated as a special case of the above construction (although a much simpler argument works, because player 1 need not be indifferent between C and D in this case). This completes the

[9]Technically, this is seen as follows; a sufficiently small ξ guarantees that the lower bound of the mixing probability $\varepsilon'(\xi)$ obtained above is actually less than 1.

intuitive explanation, and now let us provide the formal proof for Case 3. The incentive constraints are

$$\begin{aligned} d &= \xi\left(x_1(Y) - x_1(X)\right) \text{ (1 is indifferent), and} \\ \varepsilon h + (1-\varepsilon)d &\leq \left\{p\left(Y|D_\varepsilon, D\right) - p\left(Y|D_\varepsilon, C\right)\right\}\left(x_2(X) - x_2(Y)\right) \\ &= \left\{-(1-\varepsilon)\xi + \varepsilon\Delta(\xi)\right\}\left(x_2(X) - x_2(Y)\right) \text{ (2 does not want to play } D\text{).} \end{aligned}$$

The first constraint and $\lambda \cdot x(\omega) = 0$ for $\omega \in (X, Y)$ are satisfied, if we choose $x_1(X) = 0$, $x_1(Y) = \frac{d}{\xi}$, $x_2(X) = 0$, and $x_2(Y) = -\frac{\lambda_1}{\lambda_2}\frac{d}{\xi}$. As for the second constraint, the RHS

$$\begin{aligned} &\left\{-(1-\varepsilon)\xi + \varepsilon\Delta(\xi)\right\}\left(x_2(X) - x_2(Y)\right) \\ = &\left\{\varepsilon\left(\xi + \Delta(\xi)\right) - \xi\right\}\frac{\lambda_1}{\lambda_2}\frac{d}{\xi} \\ \geq &\; d\left\{\varepsilon\frac{\xi + \Delta(\xi)}{\xi} - 1\right\} \text{ (because } \lambda_1 \geq \lambda_2 > 0\text{)} \end{aligned}$$

Calculation shows that this is larger than $\varepsilon h + (1-\varepsilon)d$ if $\varepsilon \geq \varepsilon'(\xi)$, where

$$\varepsilon'(\xi) \equiv \frac{2d}{d(\frac{\xi+\Delta}{\xi} + 1) - h}.$$

Symmetric arguments apply to the remaining cases. Hence, when $\frac{\xi}{\Delta(\xi)}$ is sufficiently small, $Q(\xi)$ contains

$$\underline{Q}(\xi) \equiv \Re^2_+ \cap \mathrm{Co}\{g\left(D_{\varepsilon'(\xi)}, C\right), g\left(C, D_{\varepsilon'(\xi)}\right), g(D, C), g(C, D), (0, 0)\},$$

where Co denotes the convex hull (see Figure A). The reader can verify from the figure that $\underline{Q}(\xi) \subset H(\lambda)$ for each case discussed above (Case 1 is depicted in the figure).

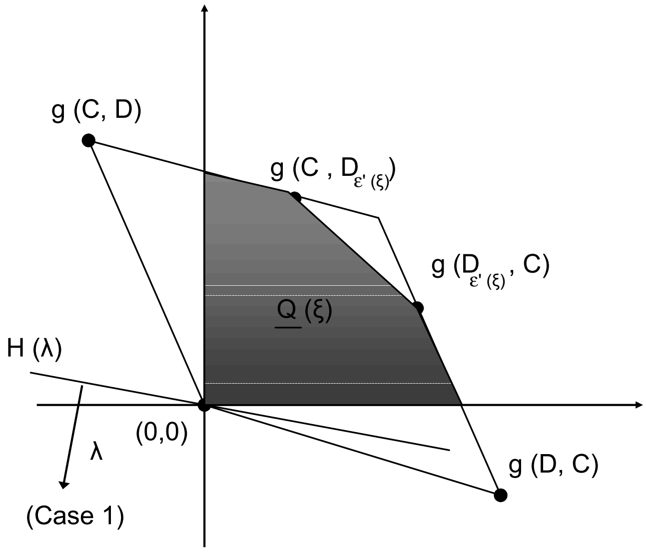

Figure A

Note that $\underline{Q}\,(\xi) \rightarrow V^*$as $\xi \rightarrow 0$, because $\varepsilon'(\xi) \rightarrow 0$. This and the fact $Q(\xi) = \lim_{\delta \rightarrow 1} \overline{E}\,(\delta) \subset V^*$ prove Lemma 2. **Q.E.D.**

4 A Simple Example of the Folk Theorem without the Pairwise Full Rank Condition

The feasibility of efficient asymmetric punishment is the basic driving force of the FLM folk theorem, and one of the sufficient conditions to facilitate asymmetric punishment is the pairwise full rank condition. As we have seen in Section 2, however, the preceding model has the property that the pairwise full rank condition is not satisfied but asymmetric punishment is feasible. This suggests that we may obtain the folk theorem under a weaker set of assumptions. In fact, a minor modification of the preceding model provides a fairly simple example where the folk theorem holds without the pairwise full rank condition.

Suppose that we have the same stage game payoff (the prisoners' dilemma) and assume that the signal $\omega = X, Y$ is distributed as follows.

$$
\begin{aligned}
P(X|CC) &= P(X|DD) = \frac{1}{2} \\
P(X|DC) &= \frac{1}{2} + \xi \\
P(X|CD) &= \frac{1}{2} - \xi
\end{aligned}
$$

The same argument as in Section 2 shows that the FLM pairwise full rank condition is violated at all the pure action profiles, basically because the number of outcomes is small relative to the number of actions.[10] Nonetheless, the basic idea behind FLM is still valid. For example, consider (C, C). It is possible to distinguish different players' deviations here; X is more likely if player 1 deviates, and Y is more likely if player 2 deviates (assuming that ξ is positive). Therefore efficient punishments based on transfer of utility (á la FLM) can still be employed to support the efficient action profile without any efficiency loss.

The more tricky part is to support the asymmetric profiles with respect to a variety of hyperplanes. For example, consider another efficient point (D, C). At this point, either player's deviation makes Y more likely, so that asymmetric punishment is *not* feasible. However, since player 1's deviation (to C) is *un*profitable, we may be able to transfer, without violating the incentive constraints, player 2's continuation payoff to player 1, when Y is realized. We have to check if this transfer, which makes player 1's deviation to C more attractive, is not too large to wipe out the loss associated with the unprofitable deviation to C. It turns out that we can overcome this potential problem by the particular payoff structure of the Prisoners' dilemma (see Case 2 in the proof).

The enforceability of a given action profile on various hyperplanes is the core to achieve the folk theorem. FLM's pairwise full rank condition ensures this by requiring that the *linear* combinations of relevant signal distributions are distinct. Kandori and Matsushima [7] pointed out that this can be weakened by requiring that the *convex* combinations of relevant signal distributions are distinct. An essential condition of theirs is satisfied at (C, C) in our example[11]. This fails at (D, C) and (C, D), however, as

[10] On the other hand, the individual full rank condition of FLM is satisfied at every action profile.

[11] At this point, Player 1, by deviating with a certain probablitiy, can create a convex combination of $(\frac{1}{2}, \frac{1}{2})$ and $(\frac{1}{2} + \xi, \frac{1}{2} - \xi)$ (those vectors represent (probability of X, probabiity of Y)). This is distinct from what Player 2's deviations create (convex combinations

either player's deviation creates the same distribution. Still, those points can be enforced on hyperplanes as we argued above, because of the special payoff structure. This suggests that we can obtain even weaker set of conditions for the enforceability on hyperplanes, by imposing restrictions *jointly* on the information structure *and* the deviation payoffs. This issue is being addressed in work in preparation by Harrison Cheng [3].

Now we are ready to prove the following result.

Theorem 2 *The model in this section violates the pairwise full rank condition but the folk theorem holds; for any* $\xi > 0$, $\lim_{\delta \to 1} E(\delta) = V^*$

Proof: We continue to use the Fudengerg-Levine Algorithm and demonstrate $Q(\xi)(= \lim_{\delta \to 1} E(\delta)) = V^*$.

Case 1, $\lambda_1, \lambda_2 \leq 0$ **and** $\lambda \neq 0$**:** $(0,0) \in H(\lambda)$.

This case is obvious.

Case 2, $\lambda_1 > 0$, $\frac{\lambda_2}{\lambda_1} \leq \frac{d}{1+h}$**:** $g(D,C) \in H(\lambda)$ (For such λ, $\lambda \cdot g(a)$ is maximized by (D,C))

This is established by showing that (D,C) is enforceable while $\lambda \cdot x(\omega) = 0$ is satisfied. The incentive constraints are

$$\begin{aligned} g_1(D,C) + E[x_1(\omega)|D,C] &\geq g_1(C,C) + E[x_1(\omega)|C,C] \\ g_2(D,C) + E[x_2(\omega)|D,C] &\geq g_2(D,D) + E[x_2(\omega)|D,D], \end{aligned}$$

which reduce to

$$\begin{aligned} d &\geq \xi(x_1(Y) - x_1(X)) \\ h &\leq \xi(x_2(X) - x_2(Y)). \end{aligned}$$

$\lambda \cdot x(\omega) = 0$ can be satisfied by setting $x_1(Y) = x_2(Y) = 0$, $x_1(X) = -\frac{\lambda_2}{\lambda_1} x_2(X)$. Then these two inequality constraints become

$$\begin{aligned} d &\geq \xi \frac{\lambda_2}{\lambda_1} x_2(X) \\ h &\leq \xi x_2(X). \end{aligned}$$

of $(\frac{1}{2}, \frac{1}{2})$ and $(\frac{1}{2} - \xi, \frac{1}{2} + \xi)$), even though $(\frac{1}{2}, \frac{1}{2})$, $(\frac{1}{2} + \xi, \frac{1}{2} - \xi)$, and $(\frac{1}{2} - \xi, \frac{1}{2} + \xi)$ are linearly dependent (so that the pairwise full rank condition fails).

If $\lambda_2 \leq 0$, these constraints can be satisfied by taking $x_2(X)$ large enough. So suppose not. Then we can find a $x_2(X)$ to satisfy these inequalities if and only if $\frac{\lambda_2}{\lambda_1} h \leq d$, which is always satisfied because $\frac{\lambda_2}{\lambda_1} \leq \frac{d}{1+h}$ by assumption.

Case 3, $\lambda_1, \lambda_2 > 0$**:** $g(C, C) \in H(\lambda)$.

The incentive constraints are

$$\begin{aligned} d &\leq \xi(x_1(Y) - x_1(X)) \\ d &\leq \xi(x_2(X) - x_2(Y)). \end{aligned}$$

These can be satisfied by, for example, $x_1(Y) = x_2(Y) = 0$, $x_1(X) = -\frac{\lambda_2}{\lambda_1} x_2(X)$ by taking $x_2(X)$ large enough while $\lambda \cdot x(\omega) = 0$ is satisfied.

Note that the symmetric argument of case 2 applies to the remaining case; $\lambda_2 > 0, \frac{\lambda_1}{\lambda_2} \leq \frac{d}{1+h}$. For this case, the incentive constraints are

$$\begin{aligned} h &\leq \xi(x_1(Y) - x_1(X)) \\ d &\geq \xi(x_2(X) - x_2(Y)). \end{aligned}$$

Then we can set $x_1(X) = x_2(X) = 0$ and $x_2(Y) = -\frac{\lambda_1}{\lambda_2} x_1(Y)$ and show that $g(C, D) \in H(\lambda)$ for such λ. Combining all these cases, it is clear that we have $Q(\xi) = V^*$. **Q.E.D.**

5 Relationship to the Blackwell-Monotonicity

The example in Section 3 seemingly contradicts Kandori [6], which shows that the equilibrium payoff set becomes smaller when observability is reduced in Blackwell's sense. In this section we explain the precise relationship between our example and Kandori [6].

Recall that a signal ω' is less informative than ω in Blackwell's sense, if there exists a function $q(\cdot|\cdot) \geq 0$ such that $\sum_{\omega' \in \Omega} q(\omega'|\omega) = 1$ and $p'(\omega'|a) = \sum_{\omega \in \Omega} q(\omega'|\omega)\, p(\omega|a)$ for each $\omega \in \Omega$ (where $p'(\omega'|a)$ and $p(\omega|a)$ denote the associated distribution functions). When this condition holds, we say that ω' is a *garbling* of ω. It is easy to see that, when public randomization device is available, any equilibrium strategy under the less informative signal ω' can be mimicked under the better signal ω. This is because the players can garble the signal by themselves via public randomization device. Inspection of the equilibrium conditions immediately shows that the strategy profile thus constructed is also an equilibrium under the better signal ω. Hence, when ω' is a garbling of ω, we have $E_{\omega'}(\delta) \subseteqq E_{\omega}(\delta)$ for any $\delta \in [0,1)$, where $E_{\omega'}(\delta)$ and $E_{\omega}(\delta)$ denote the public perfect equilibrium payoff sets under signal ω' and ω. Kandori's results [6] are built on this observation. Kandori pointed out that the same is true without public randomization when signal is a continuous variable, and he went on to show $E_{\omega'}(\delta) \subsetneqq E_{\omega}(\delta)$ under certain regularity conditions.

To demonstrate the precise relationship between those observations and our example, we now show that a similar result is obtained (without public randomization) in the current setting (i.e., with the finite signal space) *in the limit* ($\delta \to 1$). Let $Q_{\omega'}$ and Q_{ω} be the set Q $(= \lim_{\delta \to 1} E(\delta))$ associated with ω' and ω respectively.

Proposition 1 *If ω' is a garbling of ω, $Q_{\omega'} \subseteqq Q_{\omega}$.*

Proof: Take any feasible (α', x') for the optimization problem with p' for direction λ. Consider (α', x) where x is defined by $x(\omega) = \sum_{\omega' \in \Omega} q(\omega'|\omega)\, x'(\omega')$. Then (α', x) is feasible for the optimization problem with p for the direction λ because the feasible set of x is convex and the incentive constraints are automatically satisfied. Moreover, it is also clear that (α', x) achieves the same value as (α', x'). This implies that $H_{p'}(\lambda) \subset H_p(\lambda)$ for any λ, therefore $Q_{\omega'} \subseteqq Q_{\omega}$.**Q.E.D.**

This logically implies that, although our signalling structure gets "less informative" as $\xi \to 0$, it does so not in the sense of Blackwell. In fact, it is not so difficult to see that, in the limit, the assumption for Theorem 1 precludes a signalling structure which is less informative by Blackwell's criterion.

Suppose that p_{n+1} is a garbling of p_n for $n = 0, 1, ...$ and $\|p_n(\cdot|CD) - p_n(\cdot|CC)\|$ converges to 0 as $n \to \infty$, where $p_0(\omega|a) = p(\omega|a)$. Since $p_{n+1}(\cdot|a) = A_n p_n(\cdot|a)$ where A_n is a matrix defined by

$$A_n = \begin{pmatrix} q_n(X|X) & q_n(X|Y) \\ q_n(Y|X) & q_n(Y|Y) \end{pmatrix}$$

for some $q_n(\cdot|\cdot)$, we have

$$p_{n+1}(\cdot|a) = A_n ... A_1 A_0 p(\cdot|a)$$

Let $A^n = A_n ... A_1 A_0$. Since $\lim_{n\to\infty} \|A^n (p_1(\cdot|CD) - p_1(\cdot|CC))\| = 0$, you can show that

$$\lim_{n\to\infty} A^n = \begin{pmatrix} x & x \\ 1-x & 1-x \end{pmatrix}$$

for some $x \in [0,1]$. This implies that $\|p_n(\cdot|DD) - p_n(\cdot|DC)\|$ also converges to 0, and, moreover, $\|p_n(\cdot|CD) - p_n(\cdot|CC)\|$ and $\|p_n(\cdot|DD) - p_n(\cdot|DC)\|$ converges to 0 at the same speed.[12] This is in conflict with our assumption $\lim_{\xi \to 0} \frac{\xi}{\Delta(\xi)} = 0$.

---

[12] Formally, let $x = (x_1, x_2)$ and $y = (y_1, y_2)$ be any two distinct signal distributions on $\{X, Y\}$. Then

$$\|A^n(x-y)\| = \left\| \begin{pmatrix} \mu_n' & \mu_n'' \\ 1-\mu_n' & 1-\mu_n'' \end{pmatrix} (x-y) \right\|$$

This is 0 for any x, y if $\mu_n' = \mu_n''$. So suppose not. Since $x_1 - y_1 = -(x_2 - y_2)$,

$$\|A^n(x-y)\| = |\mu_n' - \mu_n''| \, \|x-y\|$$

Therefore,

$$\frac{\|A^n(p(\cdot|CC) - p(\cdot|CD))\|}{\|A^n(p(\cdot|DC) - p(\cdot|DD))\|} = \frac{\|p(\cdot|CC) - p(\cdot|CD)\|}{\|p(\cdot|DC) - p(\cdot|DD)\|}$$

References

[1] D. Abreu, D. Pearce, and E. Stacchetti. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica*, 58:1041–1063, 1990.

[2] D. Blackwell and M. A. Girshick. *Theory of Games and Statistical Decisions*. John Wiley and Sons, New York, 1954.

[3] H. Cheng. Efficiency and the no surplus condition in repeated games with imperfect monitoring. Mimeo., 2001.

[4] D. Fudenberg and D. Levine. Efficiency and observability with long -run and short-run players. *Journal of Economic Theory*, 62:103–135, 1994.

[5] D. Fudenberg, D. K. Levine, and E. Maskin. The folk theorem with imperfect public information. *Econometrica*, 62:997–1040, 1994.

[6] M. Kandori. The use of information in repeated games with imperfect monitoring. *Review of Economic Studies*, 59:581–594, 1992.

[7] M. Kandori and H. Matsushima. Private observation, communication and collusion. *Econometrica*, 66:627–652, 1998.