Efficiency in Repeated Games Revisited: The Role of Private Strategies*

Michihiro Kandori
Department of Economics
University of Tokyo

Ichiro Obara
Department of Economics
UCLA

January 26, 2004

Abstract

Most theoretical or applied research on repeated games with imperfect monitoring has restricted attention to *public strategies*; strategies that only depend on history of publicly observable signals, and *perfect public equilibria* (PPE); sequential equilibria in public strategies. The present paper sheds light on the role of *private strategies*; strategies that depend on players' own actions in the past as well as observed public signals. Our main finding is that players can sometimes make better use of information by using private strategies and *efficiency in repeated games can often be drastically improved*. We illustrate this for games with a small signal space, where the Folk Theorem fails, as well as for games with a large signal space, for which the Folk Theorem holds. Our private strategy consists of two states and has the property that the opponent's incentives are independnt of the state the player is in. We provide two different characterizations of our two-state equilibrium for general two-person repeated games with imperfect public monitoring.

Keywords: Efficiency, Imperfect Public Monitoring, Mixed Strategy, Partnership Game, Private Equilibrium, Private Strategy, Repeated Games, Two-State Machine.

JEL classification codes: C72, C73, D82

1   Introduction

The theory of repeated games under imperfect monitoring provides a formal framework to explore the possibility of cooperation in long term relationships,

where each agent's action is not directly observable. It has been successfully applied to a number of economic problems: cartel enforcement, internal labor market, and international policy coordination, to name a few. However, almost all existing works (including Abreu, Pearce and Stacchetti [2] and Fudenberg, Levine and Maskin [4]) focus on a simple class of strategies, known as public strategies. In the present paper, we illustrate that players can make better use of information by using non-public strategies, which we call *private strategies*, and show that *efficiency in repeated games can often be drastically improved*.

A public strategy is a strategy which only depends on past realizations of the public signal. In contrast, a private strategy can also depend on private history; one's own actions in the past. Our private strategy is a mixed[1] strategy which has the following feature: a player's continuation strategy depends on the realization of his or her mixed actions in the past. For example, take a model of dynamic Cournot competition with stochastic demand (Green and Porter [5]). In this context, our private strategy has the property that each firm randomly chooses its present output level and the realized output affects the future output level.

A rough intuition for the improved efficiency by private strategies may be obtained by the following observation. It is often the case that a player has a socially inefficient action that helps to monitor other players' actions more accurately. For example, in a team production (or partnership) problem with decreasing returns to scale, observable output can be more sensitive to the opponent's effort, when a player's effort is low. Hence, if such "monitoring" action is played with a small probability and the opponents are rewarded/punished *only after* such an action is taken, the opponents' moral hazard problem can be resolved in a more efficient way. Note that in this story it is vital that the players' future behavior (reward/punishment) depends on their past actions (hence the use of private strategies).

Let us explain our point in more detail with an explicit example. Consider a simple repeated partnership game with two actions $\{C, D\}$, two public signals $\{\text{"}good\text{"}, \text{"}bad\text{"}\}$, where the stage game (expected) payoff matrix has the structure of the standard prisoners' dilemma. Assume that one's effort (to play C) reduces the likelihood of "bad" by a larger margin when the opponent is making no effort (i.e., when he is taking D). This can be regarded as the decreasing returns to scale we discussed, and it is formulated as $\Pr(\text{"}bad\text{"}|C, C) + \varepsilon = \Pr(\text{"}bad\text{"}|C, D) = \Pr(\text{"}bad\text{"}|D, C) << \Pr(\text{"}bad\text{"}|D, D)$, where $\varepsilon > 0$ is a small number. First note that cooperation cannot be sustained with the standard trigger strategy when ε is very small because the public signal is insensitive to a deviation when (C, C) is played. One way to get around this problem within the class of *public* strategies is to use a mixed trigger strategy; mix D with some probability in the cooperative phase. Although playing D with positive probability causes inefficiency in the stage game payoffs, the public signal becomes more sensitive to deviations. This may allow players to use a

---

[1]It is well known that public strategies achieve the same equilibrium outcomes as private strategies, if we consider pure strategies. See footnote 7 for the precise statement.

mutual punishment after the public signal "*bad*" to sustain a certain level of co-operation.[2],[3] Now note that we can improve efficiency further by using *private* strategies. When a player is randomizing between C and D, the most informative action-signal pair with respect to a deviation is $(D, "bad")$. Therefore it is more efficient to start a punishment only after $(D, "bad")$. In contrast, the trigger strategies based only on public information are less efficient, because the punishment can occur even when C is played, that is, when the public signal is not informative at all. Our first contribution is to point out that such an advantage exists in private strategies.

Our second contribution is to find a method to construct a private equilibrium which manifests the above intuition. Note that, with private strategies, each player may not know the other player's continuation strategy because it depends on the realizations of past private actions. Hence, players have to compute their beliefs (by Bayes' rule) about what the opponents are going to do, and the computation generally becomes fairly complex over time.[4] Thus it is not an easy task to construct an equilibrium which manifests the above idea. To see this, note that the punishment by the private strategy in the above example cannot be a coordinated action. When you are supposed to punish the other player, you are not sure whether the other player is also going to punish you. Hence your incentive to punish generally depends on your belief about the opponent's future actions, which changes over time in a complex way. It is thus not obvious how we can provide the right incentive to implement punishment at the right moment.

Our private strategy can be regarded as a *machine* which consists of two states; *rewarding state* and *punishment state*. A player may play a mixed action at each state, and the transition probability between the states depend on a realization of action-signal pair. Note that there is always uncertainty regarding the other player's strategy as we pointed out above. At each point of time, each player may not know which state the other player is in, or which continuation strategy is used by the other player. The trick is to choose the right mixed actions and the right transition probabilities to make one's incentives *identical* (and adequate to support equilibrium actions), *no matter which state the other player is in*. Such construction makes one's belief about what the other player is going to do irrelevant, in the sense that one's action is optimal independent of her belief.

This idea is very powerful in dealing with private information. Indeed it can deal with not only private information about past actions but also private signals, if any. A similar idea was first applied by Piccione [16] in the framework of repeated games with private monitoring. Two-state machine strategies were

---

[2] The level of punishment can be adjusted so that players are actually indifferent between C and D.

[3] Note that the efficiency can be improved *even within the class of public strategies* by using a mixed strategy. To the best of our knowledge, this has not been pointed out in the existing literature.

[4] This is the same difficulty as the diffficuly in repeated games with private monitoring. See Kandori [7].

first independently found by Ely and Välimäki [3] in repeated games with private monitoring, and by Obara [14] on which this paper is based.

The paper is organized as follows. We describe our model and provide a brief review of previous results in Section 2. In Section 3, we show that private equilibria are more efficient than PPE under a certain conditions for games with a small signal space. We use a version of partnership games for which the Folk Theorem fails as shown by Radner, Myerson, and Maskin[17]. In Section 4, we illustrate that our result holds even for games with a large signal space. Again we use a partnership game, but with a rich signal space this time. Our model satisfies the sufficient condition of the Fudenberg-Levine-Maskin Folk theorem, thus the efficiency is approximately achieved by PPE as $\delta \to 1$. However, we show that our equilibrium approximates efficiency faster, and is always more efficient than any PPE as long as there exists any nontrivial PPE. Section 5 provides two different characterizations of our two-state machine equilibrium for general two-person repeated games with imperfect public monitoring. Section 6 discusses related literature and concludes.

2   The Model and Review of Previous Results

In this section we present a general model of repeated games with imperfect public monitoring and review some of the existing results. Throughout the paper we consider two-player games. In the stage game, each player $i = 1, 2$ chooses an action $a_i \in A_i$, and they publicly observe a signal $\omega \in \Omega$. We assume that A_i and Ω are finite sets and let $p(\omega|a)$ denote the probability of signal ω given action profile $a \in A = A_1 \times A_2$. The realized payoff to player i is $u_i(a_i, \omega)$ (so that it conveys no more information than a_i and ω do), and the expected payoff is given by $g_i(a) = \sum_{\omega \in \Omega} u_i(a_i, \omega) p(\omega|a)$. Denote i's mixed action by $\alpha_i \in \Delta_i$, and with an abuse of notation, let $g_i(\alpha)$ be player i's expected payoff associated with mixed action profile $\alpha \in \Delta_1 \times \Delta_2$. We also let $p(\cdot|\alpha)$ be probability distribution on public signals given $\alpha \in \Delta_1 \times \Delta_2$. The stage game is played over an infinite time horizon $t = 0, 1, 2, ...$, and player i's (average) payoff in the repeated game is given by $(1 - \delta) \sum_{t=0}^{\infty} \delta^t g_i(a(t))$, where $a(t)$ is the action profile at time t and $\delta \in (0, 1)$ is the discount factor.

The existing literature restricts attention to *public strategies*, where action at time t depends only on the history of publicly observable signal $(\omega(0), ..., \omega(t-1))$. A sequential equilibria in public strategies[5] is called a *perfect public equilibrium (PPE)*, and the structure of this class of equilibria is now well understood. Roughly speaking, efficiency can be achieved in this class as $\delta \to 1$, if the signal space is large enough (the Fudenberg-Levine-Maskin folk theorem [4]). More precisely, when $|\Omega| \geq |A_1| + |A_2| - 1$, for a generic choice of payoffs and signal distributions[6], any feasible and individually rational payoff profile can be asymptotically achieved by a PPE as $\delta \to 1$ (the Fudenberg-Levine-Maskin folk

---

[5]Deviations to non-public strategies are allowed.

[6]This is because Fudenberg-Levine-Maskin's full rank conditions and full dimensionality condition are satisfied generically.

theorem). When the signal space is small, in contrast, there are well-known examples where the public perfect equilibria do not achieve efficiency (Radner, Myerson, and Maskin [17]).

In the present paper, we consider a more general class of strategies, where action at time t depends on the history of publicly observable signal *and* his own action $(\omega(0), ..., \omega(t-1), a_i(0), ..., a_i(t-1))$. Such a strategy is called *private*, and a sequential equilibrium in private strategies is called a *private equilibrium(PE)*.[7] We demonstrate that private equilibria can outperform PPE in either case discussed above.[8] Namely, for the case where the folk theorem fails, we show that private equilibria can be fully or almost efficient when PPE are far away from the Pareto frontier for all $\delta \in (0,1)$. Even when the folk theorem holds, we show that private equilibria can be strictly more efficient than PPE *for each* sufficiently large $\delta < 1$, although in the limit $\delta \to 1$ they both achieve efficiency.

3 The Advantage of Private Strategies in Games with a Small Signal Space

We first examine the case where the signal space is small, so that efficiency cannot be achieved by PPE (the folk theorem fails). We present a series of examples where private equilibria outperform PPE, each of which is a special case of the following model.

The stage game is a symmetric game with $\Omega = \{X, Y\}$ and $A_1 = A_2 = \{C, D\}$. (Note that the folk theorem condition $|\Omega| \geq |A_1| + |A_2| - 1$ fails.) We assume the stage game payoffs have the prisoners' dilemma structure

	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, that is, it Pareto-dominates the equal (public) randomization between (C, D) and (D, C)). We also assume that the signalling structure is symmetric $(p(\omega|C, D) = p(\omega|D, C))$. This is a simplified version of the model examined by Radner, Myerson and Maskin [17] (RMM from now on).[9]

We derive an upper bound of the PPE payoffs as a benchmark. Our analysis is a generalization of RMM [17] and Abreu, Milgrom, and Pearce [1] (AMP from now on). The main difference is that we consider mixed strategies, while those

---

[7]These terms "private strategy" and "private equilibrium" are first coined in Obara [14].

[8]Note that it would be without loss of generality to restrict attention to public strategies if we were to consider only pure strategies (Abreu, Pearce, and Stacchetti [2]). More precicely, for any pure strategy, there exists a payoff equivalent public pure strategy. Since a pure strategy Nash equilibrium is a PPE when the signal distribution has a full support, we need to use *a mixed strategy (whose behavior strategy representation is private)* in order to outperform PPE.

[9]The action set is continuum in Radner, Myerson, and Maskin [17].

papers only consider pure strategies. Besides finding an upper bound, this analysis serve two other purposes. First, we demonstrate a reason why the best PPE is sometimes in mixed strategies, which has not been pointed out before. Second, this fact helps to understand a general reason why private strategies can outperform public strategies (see Remark in Sections 3.2).

We first consider the best trigger strategy equilibrium payoff to sustain a symmetric profile (α^q, α^q) where α^q is the mixed action playing D with (small) probability $q \in [0, 1)$. Let $\omega(q) \in \Omega = \{X, Y\}$, be the signal satisfying $p(\omega(q)|D, \alpha^q) > p(\omega(q)|C, \alpha^q)$. It is the signal to detect player i's profitable deviation (D) from the profile (α^q, α^q).

Consider the following trigger strategy:[10]

$$(\#) \begin{cases} (1): & \text{Play } (\alpha^q, \alpha^q) \text{ in the stage game.} \\ (2): & \text{If } \omega \neq \omega(q) \text{ is observed, go back to (1)} \\ & \text{If } \omega(q) \text{ is observed, start playing } (D, D) \text{ forever with} \\ & \text{probability } \rho \text{ and go back to (1) with probability } 1 - \rho. \end{cases}$$

The average symmetric payoff v^q satisfies the dynamic programming equations[11]

$$v^q = (1 - \delta)g(C, \alpha^q) + \delta(1 - \rho p(\omega(q)|C, \alpha^q))v^q \tag{1}$$

and

$$v^q \geq (1 - \delta)g(D, \alpha^q) + \delta(1 - \rho p(\omega(q)|D, \alpha^q))v^q. \tag{2}$$

If α^q mixes C and D, they must provide an equal payoff so that (2) is satisfied with equality. If α^q plays C with probability one ($q = 0$), (2) may not be binding. However, to obtain the best equilibrium, we need to choose the smallest probability of punishment ρ so that (2) is binding. From the two dynamic programming equalities, we obtain a closed form expression of the equilibrium payoff

$$v^q = g(C, \alpha^q) - \frac{d^q}{L^q - 1}, \tag{3}$$

where

$$d^q = g(D, \alpha^q) - g(C, \alpha^q) \tag{4}$$

is the gain from deviation and

$$L^q = \frac{p(\omega(q)|D, \alpha^q)}{p(\omega(q)|C, \alpha^q)} = \frac{p(\omega(q)|D, C)(1 - q) + p(\omega(q)|D, D)q}{p(\omega(q)|C, C)(1 - q) + p(\omega(q)|C, D)q} \tag{5}$$

is the likelihood ratio which measures the degree of observability of deviation. Plugging into (3) the values in the payoff table, we obtain

$$v^q = 1 - q - hq - \frac{(1 - q)d + qh}{L^q - 1}. \tag{6}$$

---

[10] We allow the use of public randomization device for PPE, but not for PE, and still show that PE can Pareto-domate PPE.

[11] Note that we used player 1's payoff without loss of generality, as the payoff is symmetric.

Let $q^* \in \arg\max_{q \in [0,1]} v^q$. Then v^{q^*} is the best trigger strategy equilibrium payoff (indeed the best strongly symmetric equilibrium payoff) if and only if it is positive (otherwise, $(0,0)$ is the best strongly symmetric equilibrium payoff).[12] Note that this formula is independent of δ, thus implies that any trigger strategy equilibrium payoff is inefficient (< 1) how patient players are.

Note that this formula reduces to APM [1]'s formula for the best (pure) trigger strategy equilibrium payoff when $q = 0$. We can interpret $1 - q - hq$ as the stage game payoff in the cooperative phase and the last term $\frac{(1-q)d+qh}{L^q-1}$ as the efficiency loss which stems from imperfect monitoring and unnecessary punishments. This formula clarifies why mixed (public) strategies may help to improve efficiency. *While taking the inefficient action D with a larger probability q reduces the stage game payoff $(g(C, \alpha^q))$, it may improve the quality of monitoring (increase L^q) and reduce the inefficiency* associated with unnecessary punishments (it may also reduce the deviation gain d^q). So q^* may not be 0 in general, that is, the mixed trigger strategy may achieve a better outcome. Indeed it is very easy to construct a such example.

In order to get the bound of all the PPE payoffs, we also need to take into account cases where the optimal strategy pair is asymmetric. Let us define $\omega_i(\alpha) \in \{X, Y\}$ by $p(\omega_i(\alpha)|D, \alpha_{-i}) > p(\omega_i(\alpha)|C, \alpha_{-i})$. This is the signal outcome to detect Player i's defection at mixed action profile α. When $\omega_1(\alpha) \neq \omega_2(\alpha)$, different players' deviations can be statistically distinguished, and asymmetric punishment can enforce the given action α without welfare loss. Namely, when $\omega_1(\alpha)$ is realized, we *transfer* future payoffs from player 1 to player 2 (symmetric argument applies to $\omega_2(\alpha)$). Such a movement of future payoffs can be made in a close vicinity of the Pareto frontier when δ is large, so that the associated welfare loss is negligible. This is the key observation to derive the folk theorem (Fudenberg, Levine and Maskin [?]). Now let Q be the set of α such that $\omega_1(\alpha) \neq \omega_2(\alpha) \in \{X, Y\}$. The actions in Q can be sustained by the efficient asymmetric punishment, so that the sum of equilibrium payoffs could potentially be equal to

$$g^* \equiv \sup_{\alpha \in Q} \{g_1(\alpha) + g_2(\alpha)\}.$$

It turns out that the sum of all PPE payoffs are bounded by one of the following four numbers; $2v^{q^*}$, the sum of asymmetric efficient points $1 + d - h$, g^*, or the Nash payoff 0, whichever is the largest in the following sense:

Proposition 1 *Any PPE payoff profile (v_1, v_2) satisfies $V^* \geqq v_1 + v_2$, where $V^* = \max\{2v^{q^*}, 1 + d - l, g^*, 0\}$.*[13]

---

[12] A public strategy is strongly symmetric if players play the same action after every public history.

[13] We can drop the symmetry assumption for signalling structure. This proposition is still valid if $2v^{q^*}$ is replaced by $v_1^{q_2^*} + v_2^{q_1^*}$, where $v_i^{q_j^*} = g_i(C, \alpha_j^{q_j^*}) - \frac{d_i^{q_j^*}}{L_i^{q_j^*} - 1}$ ($d_i^{q_j^*}$ and $L_i^{q_j^*}$ are defined in a similar way to (4) and (5))

7

Proof. See Appendix. ∎

It is easy to construct examples for which each of these becomes the effective upper bound of $v_1 + v_2$. Note that when $2v^{q^*} = V^*$, it can be exactly achieved by a mixed trigger strategy (#) using the symmetric profile $\left(\alpha^{q^*}, \alpha^{q^*}\right)$ and $\omega(q^*)$. Kandori and Obara [10] shows a couple of examples for which $V^* = g^*$, and g^* is not exactly achieved (Section 2) or exactly achieved (Section 4).

3.1   An Example of Efficient Private Equilibria

We first present a special case of the above model, where private equilibria asymptotically achieves full efficiency, while PPE payoffs are bounded away from the efficiency frontier.

Let us assume $0 < p(X|C,C) < 1$, $0 < p(X|D,D) < 1$, and $p(X|C,D) = p(X|D,C) = 0$. Under this information structure, no PPE can approximate the efficient point $(1,1)$. This is formally shown as follows. By Proposition 1, for $(1,1)$ to be achieved by a PPE, the best trigger payoff v^q should be equal to 1. Expression (6) shows that this requires $q = 0$ (C is played for sure in the first period) and $L^0 = \infty$ (deviation to D is perfectly detected), but this is impossible: As deviation to D makes $\omega = Y$ more likely when $q = 0$, we have $\omega(0) = Y$ and $L^0 = \frac{P(Y|D,C)}{p(Y|C,C)} = \frac{1}{p(Y|C,C)} < \infty$.

In contrast, the efficient payoff profile can be approximated by the following *private* strategy, which starts at (1) below.

$$(\#\#) \begin{cases} (1): & \text{Mix between } C \text{ and } D. \text{ Choose action } D \\ & \text{with a (small) probability } q \in (0,1). \\ (2): & \text{If the signal is } X \text{ and one's own action was } D, \\ & \text{then play } D \text{ forever. Otherwise, go back to (1).} \end{cases}$$

Note well that (i) the players' future actions partly depend on their current actions (so that it is a *private* strategy) and (ii) thanks to the assumption $p(X|C,D) = p(X|D,C) = 0$, when a player chose D and observes X, it is common knowledge that the other player also chose D (and of course observes X). The equilibrium conditions are

$$v = (1-\delta)(1 - q - qh) + \delta v \tag{7}$$

$$v = (1-\delta)(1-q)(1+d) + \delta\{1 - qp(X|D,D)\}v \tag{8}$$

Equation (7) represents the average payoff when a player plays C today (while the opponent is employing the above strategy). Note that punishment is surely avoided in this case, as defection is triggered if and only if both players play D and the signal is X. Equation (8) shows the average payoff when the player chooses D today. Punishment is triggered when the opponent also chooses D and the signal is X, which happens with probability $qp(X|D,D)$. Equation (7) and (8), taken together, imply that the player is just indifferent between choosing C and D (the condition for a mixed strategy equilibrium).

From (7), we have

$$v = 1 - q - qh \qquad (9)$$

Also, by (7) and (8) we get

$$(1 - \delta)\left\{(1 - q)d + qh\right\} = \delta q p(X|D, D)v. \qquad (10)$$

This and (9) result in a quadratic equation in q;

$$(1 - \delta)\left\{(h - d)q + d\right\} = \delta q p(X|D, D)(1 - q - qh) \qquad (11)$$

It is easy to show that there is a root of this equation in $(0, 1)$ which tends to 0 as $\delta \to 1$. Equation (9) then implies that, as q tends to 0, the average payoff tends to 1, the payoff from full cooperation. This leads us to the following result.

Proposition 2 *In the game defined above, there is a private equilibrium that approximately attains the fully efficient average payoff ($= 1$) as the discount factor tends to unity, while any perfect public equilibrium average payoff is bounded away from 1 independent of the discount factor.*

Proof. See Appendix ∎

Remark (The Essential Idea of Our Paper): Since it is much easier to detect the other player's defection when one defects herself, it is more efficient to trigger a punishment only after such a (private) history. More precisely, private strategies allow players to start a punishment only after a realization of *action-signal pair* for which the likelihood ratio with respect to a defection is maximized. This high likelihood ratio helps to reduce the inefficiency term in the formula we obtained before. For this particular example, indeed the inefficiency term vanished completely because the likelihood ratio is infinity. It is based on a simple familiar principle: the efficient statistical inference, which is one of the main underlying theme in any moral hazard model. Note that public strategies can also use this high likelihood ratio to some extent, but D needs to be played with a very high probability to do so at the cost of reducing efficiency in stage games. We can avoid this kind of trade-off between efficiency and detectability by utilizing private strategies. This is the essential idea why private strategies can outperform public strategies.

The reader may wonder why the kind of private strategies presented above, based on a simple intuition, have rarely appeared in the existing literature. The answer hinges on the assumption of the moving support in this particular example. It allows players to coordinate their punishments after playing D and observing X. In another word, a realization of the action-signal pair which triggers punishment is common knowledge. Suppose that X is observed with probability ε when (C, D) or (D, C) is played. If ε is very small, this information structure is very close to the information structure of our example. However, our simple private trigger strategy is not an equilibrium any more. The problem is that cooperation/punishment is not common knowledge once X is observed.

Each player cannot be sure whether the opponent is in the punishment mode or in the cooperation mode. Such an inference problem becomes increasingly more complex as time passes by, even though the opponent is using a fairly simple strategy as above. As a result, specifying the best action in each period is usually a formidable task (and the action specified by our trigger strategy is suboptimal after certain histories).[14] The complexity of inference is the most difficult problem associated with private strategies, and this is why this class of strategies has not been fully explored. We present a way to address this issue in the next sub-section.

3.2 Two-State Machine Private Equilibria

In this section, we demonstrate how to construct a private equilibrium when the signal has full support. We assume that

$$0 < p\left(X|CC\right) < p\left(X|DC\right) = p\left(X|CD\right) < p\left(X|DD\right) < 1$$

The idea is that "bad" signal X is more likely to realize as more players defect. This is similar to the information structure in RMM [**?**] or any other standard partnership game. We also assume that defection is easier to detect when one is playing D.

$$\frac{p\left(X|D,C\right)}{p\left(X|C,C\right)} < \frac{p\left(X|D,D\right)}{p\left(X|C,D\right)} \tag{12}$$

Note that this implies a form of decreasing returns to scale.[15] Let us denote $p\left(X|CC\right) = p_0$, $p\left(X|CD\right) = p\left(X|DC\right) = p_1$, and $p\left(X|DD\right) = p_2$ in this section. With this notation, the likelihood ratio defined by (5) is expressed as $L^q = \frac{(1-q)p_1+qp_2}{(1-q)p_0+qp_1}$

Now consider the following private strategy, which we call a "*two-state machine*". The strategy has two states, R and P, and it begins with state R. Furthermore, it has the following structure:

- State R (State to Reward the opponent):

  Choose D with probability q_R (a small number). Go to state P with probability $\rho_R \in (0,1)$ if D was taken and X was observed (otherwise, stay in State R).

- State P (State to punish the opponent):

  Choose D with probability q_P (a large number). Go to state R with probability $\rho_P \in (0,1)$ if D was taken and Y was observed (otherwise, stay in State P).

Figure 2 describes this machine graphically.

---

[14] For example, since a player is indifferent between playing C and playing D only if she believes that her opponent is cooperating with probability 1, she cannot stay in the cooperative phase even after she chose C and observed X.

[15] The probability of "success" Y is increasing more for the first imput of effort, that is, $p\left(Y|CD\right) - p\left(Y|DD\right) > p\left(Y|CC\right) - p\left(Y|DC\right)$.

(C,X)
(C,Y)
(D,Y)

(D,X)

(C,X)
(C,Y)
(D,X)

ρ_R

$1 - \rho_R$

R
q_R

P
q_P

$1 - \rho_P$

ρ_P

q_R is very small

q_P is very large

(D,Y)

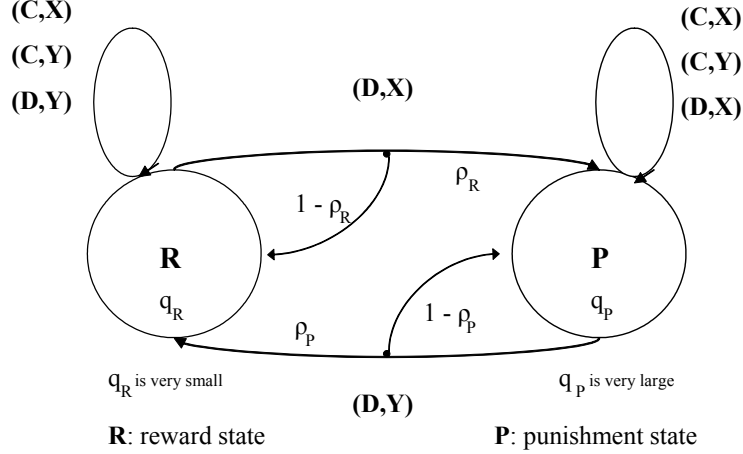**R**: reward state

P: punishment state

Figure 1:

First note that this private strategy has the same feature as the one in the previous section. Each player moves to state P only after (D, X); the most informative action-signal pair of defection, as (12) shows. Similarly, players move to state R only after (D, Y), which turns out to be the most informative action-signal pair of cooperation. Second, note that there is a strategic uncertainty we described before. A player is not sure whether the other player is in state R or state P from the second period (and never will). How can we check if this machine is a best response strategy at every history given such a ever-changing belief? To resolve this problem, we choose $(q_R, q_P, \rho_R, \rho_P)$ in such a way that *no matter which state player 2 is in, player 1 is always indifferent between choosing C and choosing D*. This means that any repeated game strategy is a best response to the machine, hence so is the machine itself.

Since the game and the strategy is symmetric, subscript i is omitted from here on as long as it does not cause any confusion. A set of parameters $(q_R, q_P, \rho_R, \rho_P)$ is chosen to satisfy the following four equations. When player 2 is in state R, the equilibrium conditions for player 1 are

- (player 1 plays C today)

$$V_R = (1 - \delta)(1 - q_R - q_R h) + \delta\{(1 - q_R p_1 \rho_R)V_R + q_R p_1 \rho_R V_P\} \quad (13)$$

and

- (player 1 plays D today).

$$V_R = (1-\delta)(1-q_R)(1+d) + \delta\{(1-q_R p_2 \rho_R)V_R + q_R p_2 \rho_R V_P\} \quad (14)$$

When player 2 is in state P, the equilibrium conditions for player 1 are

- (player 1 plays C today)

$$V_P = (1-\delta)(1-q_P - q_P h) + \delta\left[\begin{array}{c} q_P(1-p_1)\rho_P V_R \\ + \{1 - q_P(1-p_1)\rho_P\}V_P \end{array}\right] \quad (15)$$

and

- (player 1 plays D today)

$$V_P = (1-\delta)(1-q_P)(1+d) + \delta\left[\begin{array}{c} q_P(1-p_2)\rho_P V_R \\ + \{1 - q_p(1-p_2)\rho_P\}V_P \end{array}\right] \quad (16)$$

where V_Z can be interpreted as the discounted average payoff for player 1 when player 2 is in state $z = R, P$.

Equations (13) and (14) imply that player 1 is indifferent between C and D when player 2 is in state R and if her continuation payoff is completely determined by her opponent's state. Similarly, (15) and (16) imply that player 1 is indifferent between C and D when player 2 is in state P. A system of these equations implies that player 1 is completely indifferent among all the repeated game strategies, thus player 2's state indeed determines player 1's continuation payoff completely as we assumed. Any payoff difference one can make in the current period is exactly offset by the difference of the continuation payoffs caused by the change of the other player's transition probability. Let us emphasize again that a player never knows what is the opponent's continuation strategy or which state the opponent is in during the game. However, players do not have to know them because their expected payoffs cannot be affected by their own strategies. Note that this logic is somewhat similar to the one for a totally mixed strategy equilibrium in a finite normal form game. What is interesting here is that the same thing is done for an infinite game with only a finite number of equations and value functions.

The above system of equilibrium conditions consists of four equations (13)-(16) and six unknowns $(V_Z, q_Z^*, \rho_Z^*, Z = R, P)$. We can show that, when δ is close to 1, this system has a solution such that (i) the probabilities $(q_Z^*, \rho_Z^*, Z = R, P)$ are in $[0, 1]$, and (ii) $q_R^* \to 0$ as $\delta \to 1$.[16] By a manipulation similar to the one to derive the formula for the trigger equilibrium payoff (3), we can obtain $V_R^* = 1 - q_R^* - q_R^* h - \frac{(1-q_R^*)d + q_R^* h}{L^1 - 1}$. Hence the property (ii) means that the payoff arbitrary close to $1 - \frac{d}{L^1 - 1}$ is achieved as a PE as $\delta \to 1$.

_____

[16] It turns out that q_p^* can be set to 1.

Proposition 3 *Suppose that $p_2 - p_1 > p_1 d + (1 - p_2) h$.[17] Then for any $\eta > 0$, there exists a $\underline{\delta}$ such that for all $\delta \in (\underline{\delta}, 1)$, there exists a two-state machine private equilibrium whose payoff $V(\delta)$ is larger than $1 - \frac{d}{L^1 - 1} - \eta$.*

The direct proof is found in the Appendix. Later we show how to construct a two-state machine equilibrium for general two players stage game, and the above Proposition will be derived as a special case (see Section 5.2.1.).

Note that this formula uses the likelihood ratio $L^1 \left(> L^0\right)$ instead of L^0 or L^q for any $q \in (0, 1)$, but otherwise it looks exactly like the best trigger strategy payoff. The advantage of private equilibrium comes from this larger likelihood ratio. As a corollary to Propositions 1 and 3, we can show that the PE Pareto-dominates the best symmetric PPE payoff.

Corollary 4 *Suppose that $p_2 - p_1 > p_1 d + (1 - p_2) h$, $h > d$, and $1 - \frac{d}{L^1 - 1} > \frac{1 + d - h}{2}$, then there exists a $\underline{\delta}$ such that for all $\delta \in (\underline{\delta}, 1)$, there exists a two-state machine equilibrium which Parero-dominates the best symmetric PPE payoff.[18]*

Proof. See Appendix. ∎

Although certain restrictions are imposed on the structure of the stage game for this corollary, there still exists an open set of parameters which satisfies all these restrictions. As an example of Corollary 4, we construct an example where private equilibria are nearly efficient, while the only PPE is the repetition of the stage game Nash equilibrium.

Example : Assumed that $d = \kappa > 0$, $h = 1 + \kappa > 0$, and

$$
\begin{cases}
p(X|CC) = \frac{1}{2} \\
p(X|CD) = p(X|DC) = \frac{1}{2} + \epsilon \\
p(X|DD) = 1 - \epsilon
\end{cases}
$$

where ϵ is a small positive number.[19]

It is easy to show that the assumptions for Proposition 4 is satisfied for small ϵ if $\kappa < 1$. As ϵ becomes small, it becomes more difficult to detect the opponent's deviation when (C, C) is played. The formula (6) $1 - q - qh - \frac{(1-q)d + qh}{L^q - 1}$ becomes approximately $1 - q(2 + \kappa) - \frac{(1-q)\kappa + q(1+\kappa)}{q} = -q(2 + \kappa) - \frac{\kappa}{q}$ for small ϵ, which is a negative number. This means that the trigger strategy cannot sustain

---

[17] This assumption turns out to be equivalent to $V_R(\delta) > V_P(\delta)$, where $V_R(\delta)$ and $V_P(\delta)$ are derived from the equations (13) - (16). Proposition 12 in Section 5.2.1 derives expressions for $V_R(\delta)$ and $V_P(\delta)$ in a general setting, and the detailed explanation of the present example can be found there.

[18] Note that the best symmetric PPE is the most efficient PPE as we allow public randomization devices for PPE. Thus our PE (which does not use a device) is more efficient than any PPE.

[19] When a player is playing C, the distribution of the public signal is not so sensitive to the other player's action. This implies that the realized payoffs have to vary large to generate the fixed expected payoff matrix as ϵ becomes small. In particular, $u(C, X) \to -\infty$ and $u(C, Y) \to \infty$ as $\epsilon \to 0$.

anything other than the repetition of (D, D). Another candidate of the upper bound for symmetric PPE payoffs is simply $\frac{1+d-h}{2} = 0$ by Proposition 1. Hence there exists a $\bar{\epsilon}$ such that for $\epsilon \in (0, \bar{\epsilon})$ the only PPE is the repetition of the one shot Nash equilibrium independent of discount factor. On the other hand, the private equilibrium achieves $1 - \frac{d}{L^1 - 1}$ as $\delta \to 1$, which is approximately $1 - \kappa$ with small ϵ. Since κ can be an arbitrary small positive number, we can construct an example where the PE is approximately efficient and the only PPE is the repetition of the one-shot Nash equilibrium.

4 The Advantage of Private Equilibria with a Large Signal Space

The preceding analysis focuses on the case with signal space, where the folk theorem in public strategies fails. In this section we present an example to show that even when the folk theorem holds, so that efficiency is asymptotically achieved by PPE, PE may do better than any PPE for *each* sufficiently high[20] discount factor $\delta < 1$. It is a version of the prisoners' dilemma, whose expected stage game payoffs are given by the following table.

	C	D
C	$1, 1$	$-6, 2$
D	$2, -6$	$0, 0$

The public signal ω takes on three values, X, Y_1, and Y_2, and the probability distributions are given below.

	X	Y_1	Y_2
(C, C)	$1/3$	$1/3$	$1/3$
(D, C)	0	$1/2 + \epsilon$	$1/2 - \epsilon$
(C, D)	0	$1/2 - \epsilon$	$1/2 + \epsilon$
(D, D)	$1/3$	$1/3$	$1/3$

Note that, as long as $\epsilon > 0$, the pairwise full rank condition (PFR) is satisfied at (C, C), that is, the first three rows are linearly independent.[21] This means that each player's defection at (C, C) is statistically discriminated (player i's deviation makes signal Y_i more likely, $i = 1, 2$). So the Fudenberg-Levine-Maskin Folk Theorem applies, and the efficient payoff $(1, 1)$ can be approximately achieved by a PPE as $\delta \to 1$. Also note that this model is similar to the model in Section 2, where signal X arises only when both players take the same action. Therefore, it is easy to check that the efficient payoff $(1, 1)$ can also be approximately achieved by a PE as $\delta \to 1$ by the same strategy $(\#\#)$ used in

---

[20]If δ is small, the only equilibrium is a trivial one (the repetition of the stage game equilibrium), which is by definition a PPE. In our example, a PE does strictly better than any PPE whenever a non-trivial equilibrium exists.

[21]When $\epsilon = 0$, PFR fails at any (possibly mixed) action profile, because at most two rows in the above table are linearly independent.

Section 3.1. In summary, both PPE and PE asymptotically achieves efficiency as $\delta \to 1$ in this example. We can show, however, that the PE in Section 2 does better than *any* PPE for all sufficiently large $\delta < 1$, if ϵ is small enough.

Formally, we derive the following upper bound of the best symmetric PPE payoffs.

Proposition 5 *For any (large) $H > 0$, there is a (small enough) value of the signal distribution parameter $\epsilon > 0$ such that*

$$\max \left\{ 1 - (\frac{1 - \delta}{\delta})H, 0 \right\}$$

is an upper bound of the best symmetric PPE payoffs under δ.

Note that, when H is large, the upper bound is a steep (almost linear) curve for δ sufficiently close to 1 (and otherwise it is 0). The proof is given in Appendix B. Intuitively, this bound is derived by the following observation. It turns out that in our example positive payoffs cannot be sustained if we punish the players *simultaneously*. However, as long as $\epsilon > 0$, we can utilize an *asymmetric* punishment where we "transfer" player i's future payoff to player j, when player i's defection is suspected (i.e., when Y_i arises). Hence to support a payoff profile by a PPE, we must require the future payoffs to vary in the northwest/southeast directions around the payoff profile to be supported. As the players' defections become indistinguishable ($\epsilon \to 0$), however, we need huge payoff transfers to support cooperation, and for such transfers to be in the feasible payoff set, the discount factor should be sufficiently large. This observation provides a lower bound of δ to support the given payoff profile, which in turn provides the upper bound of the PPE payoffs for each δ in Proposition 5.

On the other hand, the private equilibrium in Section 3.1 relies only on the assumption $p(X|D, D) > 0 = p(X|D, C) = p(X|C, D)$, so that it also works in the present example, irrespective of the level of ϵ. As in Section 2, we can derive the equilibrium probability q_i of defection for each player i by solving the following quadratic equation in q;

$$(1 - \delta) \left\{ (h - d)q + d \right\} = \delta q p(X|D, D)(1 - q - qh) \qquad (17)$$

Note that, in the current example, we have $h = 6$, $d = 1$ and $p(X|DD) = 1/3$. Hence (17) becomes

$$f(q) \equiv 7\delta q^2 + (15 - 16\delta)q + 3(1 - \delta) = 0$$

. As we are interested in the most efficient equilibrium (hence the one with the smallest q), we choose the smaller root, denoted $q(\delta)$. Computation shows that this solution is real and lies in $[0, 1]$ when $\delta \geq 0.992$. The associated symmetric private equilibrium payoff for each player is $v(\delta) = 1 - 7q(\delta)$. Figure 2 plots this along with the upper bound of symmetric PPE payoffs in Proposition 5: $1 - (\frac{1-\delta}{\delta})H$, where H is set to be 500 by choosing a suitable small ϵ. The horizontal axis represents the discount factor δ. The solid curve represents the private equilibrium payoff, while the thin dotted straight line is an upper bound of all PPE payoffs.
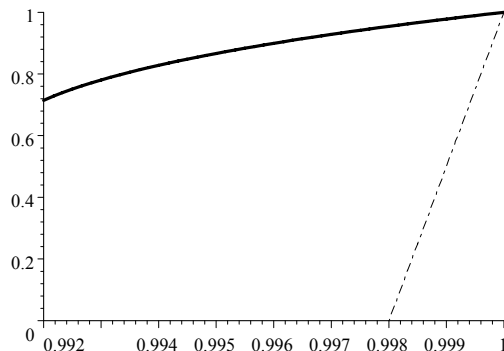
Figure 2

Hence, for each $\delta < 1$, a PE does better than any PPE (whenever something other than the repetition of the stage game equilibrium can be sustained).

5 Generalization

In this section, we demonstrate how to construct a two-state machine equilibrium in general two-person repeated games. As we generalize our construction of private strategy equilibria, it may help to generalize two state machine as well. An obvious way to generalize it is to incorporate more than two states. We first show that restricting attention to two-state machine entails no loss of generality. In the next subsection, we introduce general machines with many states, which share the same property with the simple two state machine in the previous sections, and show that they can be reduced to a two state machine.

5.1 Two State Is Enough

We formally define generalized machine with many states as follows. A machine M_i for player $i = 1, 2$ is $\left\{ \{\theta_i^n\}_{n=0}^{l_i}, \alpha_i, \mu_i \right\}$ (l_i can be ∞), where $\{\theta_i^n\}_{n=0}^{l_i}$ is the set of player $i's$ states with θ_i^0 being the initial state. Player $i's$ behavior strategy at the state θ_i^n is $\alpha_i^n \in \Delta_i$, and $\mu_i^{nm}(a_i, \omega)$ is the probability to transit from θ_i^n to θ_i^m when a_i is played and ω is observed.

Let $\text{supp}(\alpha_i^n)$ be the support of α_i^n. Suppose that (M_1, M_2) satisfies the following conditions for $i = 1, 2$ for some bounded sequence of real numbers $V = \left(\{V_1^n\}_{n=0}^{l_2}, \{V_2^n\}_{n=0}^{l_1} \right)$.

$$
\begin{aligned}
For \ n \ &= \ 0, 1, ..., l_j \\
\forall a_i \ &\in \ A_i^*, \ V_i^n = (1 - \delta) \, g_i(a_i, \alpha_j^n) + \\
&\quad \delta \sum_{a_2 \in A_2} \sum_{\omega \in \Omega} \sum_{m=1}^{l_j} \alpha_j^n(a_j) \, p(\omega | a_i, a_j) \, \mu_j^{nm}(a_j, \omega) \, V_i^m
\end{aligned}
\tag{18}
$$

16

$$\forall a_i \quad \notin \quad A_i^*, \ V_i^n \geq (1 - \delta) g_i(a_i, \alpha_j^n) +$$

$$\delta \sum_{a_2 \in A_2} \sum_{\omega \in \Omega} \sum_{m=1}^{l_j} \alpha_j^n(a_j) p(\omega | a_i, a_j) \mu_j^{nm}(a_j, \omega) V_i^m$$

$$A_i^* \quad = \quad \cup_{n=1}^{l_i} supp(\alpha_i^n)$$

This is clearly a straightforward generalization of the equations (13)-(16). This machine is basically a many-state analogue of the two state machine in previous sections.

Consider any sequential equilibrium which consists of a pair of machines with many states, which satisfies the above equations. We can show that there exists a sequential equilibrium with a two state machine which achieves the same equilibrium payoff.[22]

Proposition 6 *If a pair of machines (M_1, M_2) with many states $(2 \leq l_1, l_2 \leq \infty)$ satisfies (18), there exists a pair of two state machines which constitute a sequential equilibrium with the payoff profile*

$$\left(V_1'^0, V_2'^0\right) = \left(\sup_{n=0,\dots,l_2} \{V_1^n\}, \ \sup_{n=0,\dots,l_1} \{V_2^n\} \right).$$

Proof. See Appendix. ∎

The intuition of the proof is very simple. Player i's state θ_i^n determines player j's continuation payoff completely. If the number of player i's states is finite, then there exists player i's state $\overline{\theta}_i$ which maximizes player j's continuation payoff and $\underline{\theta}_i$ which minimizes player j's continuation payoff. Then player i can always generate player j's payoff at any other state θ_i^n by randomly moving to $\overline{\theta}_i$ and $\underline{\theta}_i$ when she is supposed to move to θ_i^n. Hence she needs only two states to generate any payoff of player j associated with all states. When the number of the states is not finite, we may not be able to find such $\overline{\theta}_i$ and $\underline{\theta}_i$. However, we can still find a sequence of the states (and mixed actions associated with them) to approximate $V_j'^0 = \sup_{n=0,\dots,l_i} \{V_j^n\}$ and $V_j'^1 = \inf_{n=0,\dots,l_i} \{V_j^n\}$, and we can construct a two state machine whose states correspond to $\sup_{n=0,\dots,l_i} \{V_j^n\}$ and $\inf_{n=0,\dots,l_i} \{V_j^n\}$ $i, j = 1, 2$, by choosing a convergent subsequence.

Remark 7 *It turns out that any payoff profile $(V_1, V_2) \in \left[V_1'^1, V_1'^0\right] \times \left[V_2'^1, V_2'^0\right]$ can be supported by using the two state machine we constructed. For example, if player i chooses $V_j'^0$ and $V_j'^1$ with probability $(1 - \lambda_i, \lambda_i)$ in the first period, this is still a sequential equilibrium and player j's expected average payoff is $(1 - \lambda_i) V_j'^0 + \lambda_i V_j'^1$, $j \neq i$.*

---

[22] Recall that Piccione [16] used such a machine with countable states in the context of repeated games with private monitoring, and Ely and Valimaki [3] succeeded to simplify it to a two state machine. The following result provides an algorithm to reduce the number of states to two in more general settings.

5.2 General Two State Machine

Now we can focus on two state machines. We use R and P to denote "reward" and "punishment" states as in Section 3.2. Let A_i^Z denote the support of the equilibrium mixed action α_i^Z for state $Z = R, P$. We are going to show that the equilibrium condition ((18) for the two-state case) can be simplified in two ways. The first characterization result shows that the equilibrium condition is equivalent to the following simpler system of linear inequalities:

(LI) For $i, j = 1, 2$ and $j \neq i$, there exist $x_i^R : \Omega \times A_j^R \to [0, \infty)$ and $x_i^P : \Omega \times A_j^P \to [0, \infty)$ such that

$$\forall a_i \in A_i^* \quad V_i^R = g_i(a_i, \alpha_j^R) - E[x_i^R(\omega, a_j)|a_i, \alpha_j^R] \tag{19}$$

$$\forall a_i \notin A_i^* \quad V_i^R \geqq g_i(a_i, \alpha_j^R) - E[x_i^R(\omega, a_j)|a_i, \alpha_j^R] \tag{20}$$

$$\forall a_i \in A_i^* \quad V_i^P = g_i(a_i, \alpha_j^P) + E[x_i^P(\omega, a_j)|a_i, \alpha_j^P], \tag{21}$$

$$\forall a_i \notin A_i^* \quad V_i^P \geqq g_i(a_i, \alpha_j^p) + E[x_i^P(\omega, a_j)|a_i, \alpha_j^P], \text{ and} \tag{22}$$

$$V_i^R > V_i^P. \tag{23}$$

$$A_i^* = A_i^R \cup A_i^P \tag{24}$$

Proposition 8 (Linear Inequalities Characterization) *If there is a two-state machine equilibrium which satisfies the equilibrium condition (18), then (LI) is satisfied. Conversely, if (LI) holds, then there is a two-state machine equilibrium which satisfies the equilibrium condition (18) for $(\alpha^Z, V^Z, Z = R, P)$, provided that discount factor δ is close enough to unity.*

The proof is given in the Appendix. Intuitively, x_i^R and x_i^P in (LI) represent the future payoff variations in each state of a two-state machine. Condition (LI) reveals that there is a certain restriction on the actions that can be used in a two-state machine:

Proposition 9 *The (potentially mixed) actions used in a two-state machine equilibrium α_i^R and α_i^P, and their support $A_i^* = supp(\alpha_i^R) \cup supp(\alpha_i^R)$ must satisfy the **separation condition***

$$\min_{a_i \in A_i^*} g_i(a_i, \alpha_j^R) > \max_{a_i \in A_i} g_i(a_i, \alpha_j^P) \tag{25}$$

Proof. Condition (19) and the non-negativity of x_i^R implies $g_i(a_i, \alpha_j^R) \geq V^R$ for all $a_i \in A_i^*$. In contrast, (21), (22), and the non-negativity of x_i^P shows $V^P \geq g_i(a_i, \alpha_j^P)$ for all $a_i \in A_i$. The corollary follows from $V_i^R > V_i^P$. ∎

The separation condition is necessary for a two-state machine equilibrium, but it is also sufficient under good observability. To state formally what "good

observability" means, let us introduce some more notations first. Let $g_i(\mathbf{a}, \alpha_j^R)$ be the vector of player's $i's$ expected payoffs given α_j^R. Let $P(a_j)$ be a positive $|A_i| \times |\Omega|$ matrix whose k, l element is $p\left(\omega^l | a_i^k, a_j\right)$ and define, for any $A_j' \subset A_j$, $P(A_j')$ by $P(A_j') = \left(P(a_j^1), ..., P\left(a_j^{|A_j'|}\right)\right)$ (a $|A_i| \times |\Omega| |A_j'|$ matrix). Finally, let $\mathbf{x}_i^Z(a_j) = \alpha_j^Z(a_j) \cdot \left(x_i^Z(\omega^1, a_j), ..., x_i^Z(\omega^{|\Omega|}, a_j)\right)$ and $\mathbf{x}_i^Z\left(A_j'\right) = \left(\mathbf{x}_i(a_j^1), ..., \mathbf{x}_i(a_j^{|A_j'|})\right)'$ for $Z = R, P$. Then, (19)-(22) can be compactly expressed as

$$
\begin{aligned}
g_i(\mathbf{a}, \alpha_j^R) - V_i^R \cdot \mathbf{I} + \mathbf{h}_i^R &= P\left(A_j^*\right) \cdot \mathbf{x}_i^R(A_j^*) \qquad (26) \\
g_i(\mathbf{a}, \alpha_j^P) - V_i^P \cdot \mathbf{I} + \mathbf{h}_i^P &= -P\left(A_j^*\right) \cdot \mathbf{x}_i^P(A_j^*)
\end{aligned}
$$

where $\mathbf{I} = (1, ..., 1)' \in \Re^{|A_i|}$ and $\mathbf{h}_i^Z \geq \mathbf{0}$ denotes non-negative slack variables, which correspond to the difference between the left and right hand sides of incentive constraints in (LI) (hence h_i^k is 0 if the corresponding action profile a_i^k is in A_i^*). Geometrically, this means that the left hand side is contained in the convex cone generated by the column vectors of $P\left(A_j^*\right)$ (or $-P\left(A_j^*\right)$), which we denote by $cone\left(P\left(A_j^*\right)\right)$ $\left(-cone\left(P\left(A_j^*\right)\right)\right)$.

These equations (26) imply that a more informative signalling structure (in the sense of Blackwell) leads to a better two state machine. Let $\widetilde{w}'$ and $\widetilde{w}$ be a random public signal and assume that $\widetilde{w}'$ is a garbling of $\widetilde{w}$. Then the following result is immediately obtained.

Proposition 10 *If $\widetilde{w}'$ is a (strict) garbling of $\widetilde{w}$ in the sense of Blackwell, then for any two state machine equilibrium with $\widetilde{w}'$ with the equilibrium payoff $\left(V_1'^R, V_2'^R\right)$, there exists a two state machine equilibrium with $\widetilde{w}$ with the equilibrium payoff $\left(V_1^R, V_2^R\right) \geq (>) \left(V_1'^R, V_2'^R\right)$.*

Proof. Let p' and p be density functions for $\widetilde{w}'$ and $\widetilde{w}$ respectively. Then there exist density functions $q(\cdot | \omega')$ on Ω for each $\omega' \in \Omega$ and $p'(\omega | a) = \sum_{\omega'} q(\omega | \omega') p(\omega' | a)$. This implies that $cone\left(P'\left(A_i^*\right)\right) \in cone\left(P\left(A_i^*\right)\right)$. Therefore, for any two state machine $\left(\alpha_j'^Z, V_i'^Z, \mathbf{h}_i'^Z, \mathbf{x}_i'^Z(A_j^*), z = R, P\right)$ which satisfies (26) with p', there exists $\mathbf{x}_i^Z(A_j^*)$ to satisfy (26) with p for the same $\left(\alpha_j'^Z, V_i'^Z, \mathbf{h}_i'^Z\right)$.

Next suppose that $\widetilde{w}'$ is a strict garbling of $\widetilde{w}$, that is, $q(\cdot | \omega') > 0$ for all $\omega' \in \Omega$. Then, $cone\left(P'\left(A_i^*\right)\right) / \{\mathbf{0}\}$ is in the relative interior of $cone\left(P\left(A_i^*\right)\right)$. Note that $\mathbf{I}$ is also in the relative interior of $cone\left(P\left(A_i^*\right)\right)$.[23] Thus, for any two state machine $\left(\alpha_j'^Z, V_i'^Z, \mathbf{h}_i'^Z, \mathbf{x}_i'^Z(A_j^*), z = R, P\right)$ which satisfies (26) with p', there exists $\eta_i \in \Re_{++}$ and $\mathbf{x}_i^Z(A_j^*)$ which satisfies

$$
g_i(\mathbf{a}, \alpha_j'^R) - \left(V_i'^R + \eta_i\right) \cdot \mathbf{I} + \mathbf{h}_i'^R = P\left(A_j^*\right) \cdot \mathbf{x}_i^R(A_j^*)
$$

---

[23] $\mathbf{I}$ is in the relative interior of $cone\left(P\left(A_j^*\right)\right)$ if and only if there exists some $a_j \in A_j^*$ and $\omega \in \Omega$ for which $p(\omega | a)$ is not constant for all $a_i \in A_i$. Note that this cannot happen when there exists a two state machine equilibrium for p' because then the separation condition is violated.

$$g_i(\mathbf{a}, \alpha_j'^P) - V_i'^P \cdot \mathbf{I} + \mathbf{h}_i'^P \;\; = \;\; -P\left(A_j^*\right) \cdot \mathbf{x}_i^P(A_j^*)$$

Thus we can obtain a two state machine equilibrium with an even higher equilibrium payoff $\left(V_1'^R, V_2'^R\right) = \left(V_1^R + \eta_1, V_2^R + \eta_1\right) > \left(V_1^R, V_2^R\right).$ $\blacksquare$

The proof of this proposition suggests that a two state machine can be constructed more easily if a cone associated with the signalling structure is larger.[24] Clearly, the upper bound of all such cones is $R_+^{|A_i|}$. We can show that if $P\left(A_j\right)$ is close enough to $R_+^{|A_i|}$ (this is what "under good observability" means), then the separation condition (25) is sufficient for the construction of a two state machine with $(\alpha_i^R, \alpha_i^P), i = 1, 2$.

Proposition 11 *Suppose that (α_i^R, α_i^P), $i = 1, 2$ satisfy the separation condition (25). Then a profile of the two-state machine equilibrium with those actions can be constructed if* $cone\left(P\left(A_i\right)\right)$ *is close enough to* $R_+^{|A_i|}$ *for* $Z = R, P$ *and* $i = 1, 2$.

Proof. See Appendix. $\blacksquare$

One example of monitoring structure to satisfy the above assumption is as follows. Suppose that, when player j mixes actions over $A_j' \subset A_j$, the opponent i's action is *perfectly detected* with a positive probability, that is,

$$\forall a_i^0 \in A_i \quad \exists(\omega^0, a_j^0) \in (\Omega \times A_j') \;\; \text{such that} \;\; p(\omega^0 | a_i, a_j^0) \begin{cases} > 0 & \text{if } a_i = a_i^0 \\ = 0 & \text{otherwise} \end{cases}.$$

This is equivalent to

$$cone(P(A_j')) = R_+^{|A_i|}.$$

Note that this means that there is a *combination* of ω^0 *and* a_j^0 to detect given action a_i^0. The advantage of private equilibrium is its ability to mix such detecting action (a_j^0) and to punish the opponent only when the detecting action is taken. Also note that, for each action of the opponent we may have a different detecting action.

We now present a second characterization of the two-state machine equilibria, which is a generalization of the closed form formula of an equilibrium payoff with relevant likelihood ratio (equation (3) in Section 3). To this end, let us first define a benchmark action and payoff for each state $Z = R, P$ as

$$a_i^Z \in A_i^*, \text{ and} \tag{27}$$

$$g_i^Z = g_i(a_i^Z, \alpha_j^Z), \tag{28}$$

---

[24] Indeed what we need for Proposition 10 is simply that the cone associated with $\widetilde{w}$ (strictly) contains the cone associated with $\widetilde{w}'$. That $\widetilde{w}'$ is a garbling of $\widetilde{w}$ is just a simple sufficient condition for this more general condition.

and define the deviation gain from this benchmark by

$$d_i^Z(a_i) = g_i(a_i, \alpha_j^Z) - g_i(a_i^Z, \alpha_j^Z). \tag{29}$$

The choice of benchmark is somewhat arbitrary, and for any choice satisfying (27) we have the following equivalence (Theorem 12). The theorem holds under the following "full support" assumption:

The full support assumption: $\quad \forall (\omega, a) \quad p(\omega|a) > 0.$

Proposition 12 *(Likelihood Ratio Representation) Under the full support assumption, (LI) is satisfied with x_i^R and x_i^P which are not identically equal to zero, if and only if the following set of conditions holds:*

(LR) For each player $i = 1, 2$ and each state for $Z = R, P$, there exist a weight function $\beta_i^Z : \Omega \times A_j^Z \to [0, 1]$ (such that $\sum_{\omega, a_j} \beta_i^Z(\omega, a_j) = 1$) and slack variables $h_i^Z : A_i \to [0, \infty)$ which satisfy

$$L_i^Z(a_i) \equiv \sum_{\omega, a_j} \beta_i^Z(\omega, a_j) \frac{p(\omega|a_i, a_j)}{p(\omega|a_i^Z, a_j)},$$

$$L_i^Z(a_i) = 1 \quad \text{if and only if} \quad d_i^Z(a_i) + h_i^Z(a_i) = 0, \tag{30}$$

$$V_i^Z = g_i^Z - \frac{d_i^Z(a_i) + h_i^Z(a_i)}{L_i^Z(a_i) - 1} \quad \text{if } d_i^Z(a_i) + h_i^Z(a_i) \neq 0, \tag{31}$$

$$\forall a_i \in A_i^* \quad h_i^Z(a_i) = 0, \tag{32}$$

and also

$$g_i^R > V_i^R > V_i^P > g_i^P,$$

is satisfied.

Remark 13 *In the above statement, the slack variable $h_i^Z(a_i)$ is equal to zero if and only if the corresponding incentive constraint in (LI) is binding, and we have*

$$L_i^Z(a_i) = \frac{E[x_i^Z(\omega, a_j)|a_i, \alpha_j^Z]}{E[x_i^Z(\omega, a_j)|a_i^Z, \alpha_j^Z]}.$$

21

The proof is found in the Appendix, and here we offer interpretations of those conditions. First, we consider state $S = R$. The remark shows that deviation to a_i increases the expected penalty by the factor of $L_i^R(a_i)$, and Proposition 12 shows that $L_i^R(a_i)$ is expressed as a convex combination of likelihood ratios[25].

The number $\frac{d_i^R(a_i)+h_i^R(a_i)}{L_i^R(a_i)-1}$ represents the welfare loss associated with the penalty scheme $x_i^R(\omega, a_j)$. Note that condition $g_i^R > V_i^R$ ensures that this welfare loss is positive. The merit of (LR) is to provide an explicit expression for the welfare loss in a simple form: It is proportional to the gain from deviation (plus the slack variable) and decreasing in the degree of observability (the likelihood ratio $L_i^R(a_i)$ in the denominator).

The conditions for state P admit similar interpretations. In this case, condition $V_i^P > g_i^P$ implies $-\frac{d_i^P(a_i)+h_i^P(a_i)}{L_i^P(a_i)-1} > 0$, so that the condition (31) may be easier to interpret if we rewrite it as

$$V_i^P = g_i^P + \frac{d_i^P(a_i) + h_i^P(a_i)}{1 - L_i^P(a_i)} \text{ if } d_i^P(a_i) + h_i^P(a_i) \neq 0.$$

The positive number $\frac{d_i^P(a_i)+h_i^P(a_i)}{1-L_i^P(a_i)}$ represents the expected bonus to satisfy incentive constraints. Note that we have bonus instead of fine, as incentives in state P are controlled by changing the probability of going to the better state R. Note also that $L_i^P(a_i)$ can be interpreted as the likelihood ratio associated with the bonus scheme $x_i^P(\omega, a_j)$.

We can state the analysis of example in Section 3.2 as a special case of this general theorem. In the example, we have, for $i = 1, 2$,

$$A_i^* = A_j^R = \{C, D\}, \text{ and } A_j^P = \{D\}.^{26}$$

When we choose

$$a_i^R = C, \text{ and } a_i^P = D,$$

[25] To obtain more concrete interpretation, let us define

$$\overline{x}_i^R \equiv \max_{(\omega, a_j)} x_i^R(\omega, a_j).$$

We could interpret that the penalty scheme $x_i^R(\omega, a_j)$ effectively imposes a *fixed* fine $\overline{x}_i^R > 0$ with probability $\frac{x_i^R(\omega, a_j)}{\overline{x}_i^R}$ when (ω, a_j) is observed. As $E[x_i^R(\omega, a_j)|a_i, \alpha_j^R] = \overline{x}_i^R \times \text{Prob}(\overline{x}_i^R$ is imposed $|a_i, \alpha_j^R)$, we have

$$L_i^R(a_i) = \frac{\text{Prob}(\overline{x}_i^R \text{ is imposed } |a_i, \alpha_j^R)}{\text{Prob}(\overline{x}_i^R \text{ is imposed}|a_i^R, \alpha_j^R)}$$

Hence, $L_i^R(a_i)$ *is the likelihood ratio of getting the fixed penalty* $\overline{x}_i^R$. *In other words, the probability of getting the fixed fine* $\overline{x}_i^R$ *is increased by this factor if player i deviates from the benchmark action a_i^R to any other action a_i.*

we have

$$L_i^R(D) = \frac{p(X|D,D)}{p(X|C,D)} = \frac{p_2}{p_1}, \text{ and}$$

$$L_i^P(C) = \frac{p(Y|C,D)}{p(Y|D,D)} = \frac{1-p_1}{1-p_2}.$$

Hence, the two-state machine can be constructed if and only if

$$V_i^R = g_i^R - \frac{d_i^R(D)}{L_i^R(D) - 1} > V_i^P = g_i^P + \frac{d_i^R(C)}{1 - L_i^P(C)} \tag{33}$$

Note that we have (i) $(g_i^R, d_i^R(D)) \to (1, d)$ as $\alpha_j^R(C) \to 1$ and (ii) $(g_i^P, d_i^R(C)) = (0, -h)$ when $\alpha_j^P(D) = 1$. Therefore, the above condition (33) is satisfied for $\alpha_j^R(C) \simeq 1$ and $\alpha_j^P(D) = 1$ if

$$1 - \frac{d}{\frac{p_2}{p_1} - 1} > \frac{-h}{1 - \frac{1-p_1}{1-p_2}}$$

which is equivalent to our condition $p_2 - p_1 > p_1 d + (1 - p_2)h$ in Proposition 3.

6 Related Literature and Comments

Private Monitoring

The private monitoring model is obtained by replacing the public signal ω with privately observed signals ω_i, $i = 1, 2$, whose joint distribution is given by $p(\omega_1, \omega_2|a)$. We claim that the private strategies we constructed also work under private monitoring. Consider the general many-state machine equilibrium in Section 5.1. If we replace ω in the dynamic programming condition (18) with ω_i, it provides an equilibrium in the private monitoring case. This condition (18) shows that each player has an incentive to follow equilibrium actions, no matter which state the opponent is in. Call this *many-state machine equilibrium in the private monitoring case*. Then, the following results are obtained in the private monitoring case, just by replacing ω with ω_i in our proofs.

1. The equilibrium payoffs achieved by a many-state machine equilibrium can also be achieved with a two-state machine (Proposition 6, where ω is replaced with ω_i).

2. Any two-state machine can be characterized by a system of linear inequalities (Proposition 8, where ω is replaced with ω_i).

3. Any two-state machine equilibrium payoff admits the likelihood representation (Proposition 12, where ω is replaced with ω_i).

23

One interesting point to note is that our equilibrium has certain continuity properties on the boundary between the private and public monitoring structures[30]. Suppose, for example, that players observe a public signal perturbed by independent private noise. This game falls in the class of repeated games with private monitoring. Since each player does not need to know the other player's state, it is not important whether a player can observe her opponent's signals. Thus our private equilibrium can also be an equilibrium for repeated games with private monitoring.

Ely and Välimäki [3] independently[31] found a similar two-state machine strategy in the framework of repeated games with private monitoring. As in this paper, a player is indifferent among all the repeated game strategies regardless of the state the opponent is in. The idea behind these strategies goes back to Piccione [16], where the equilibrium strategy is basically a machine with a countably infinite number of states.

However, there is a critical difference between our paper and Ely and Välimäki [3]. In Ely and Välimäki, a player plays a pure action at each state. In contrast, we consider mixed actions because it is crucial for our result that a player does not know which action the opponent is choosing. If a player knows the opponent's action, she is more tempted to defect when C is being played and more likely to cooperate when the "monitoring" action D is being played. Since players need to use the action-signal pair without being noticed for the efficient punishment, they need to play a mixed action at the reward state in our paper. Indeed, Ely and Välimäki's two-state machine, which uses a pure action in each state, can sometimes be strictly improved by using a mixed action at each state. The efficient use of the signaling structure is the key to our efficient private equilibria. This idea of efficient monitoring is not new. It is an old and simple idea which lies at the heart of any moral hazard model. One contribution of this paper is to find a way to use this idea to its full extent in the context of repeated games/dynamic moral hazard models.

Private Strategy

Recently, Mailath, Matthews, and Sekiguchi [12] found examples of finitely repeated games with public monitoring for which there exists a PE which is better than any PPE. Lehrer [11] used a private strategy as an endogenous correlation device in repeated games without discounting. Kandori [8] shows that FLM's sufficient condition for the folk theorem can be relaxed when players can communicate. It is based on a certain type of private strategies where players randomize over their actions and announce their realizations. Obara

---

[30]Mailath and Morris [13] provided some conditions under which a particular PPE remains a sequential equilibrium with almost public monitoring when a public signal structure is perturbed slightly with private noise. Their conditions require players to have almost common knowledge about the other players' continuation strategies at all times. Note that our PE does not satisfy this sufficient condition. On the contrary, its property is rather orthogonal to this requirement; players do not have to have any knowledge about the opponent's continuation strategy at any point.

[31]Simultaneously with the original version of this paper (Obara [14]) .

[15] applies a similar idea in the context of mechanism design. Kandori and Obara [9] considers repeated games where players may choose to pay some costs to obtain additional (private) information about the other players' actions. We examine the scope and limitations of the equilibria where (i) each player randomizes between monitoring and nonmonitoring, and (ii) each player's continuation strategy depends on whether she monitored or not. This is similar to the private strategy in this paper in the sense that players' continuation strategies depend on realizations of their mixed actions, although the signals are private, not public.

Robustness

There are a couple of comments on the robustness of the private equilibria. First, when the parameters such as $(d, h, p(X|CC), p(X|CD), p(X|DD))$ change slightly, there exists a PE close to the original PE. This is due to the regularity of equations characterizing the parameters of two-state machines. Secondly, suppose that each player can observe additional signals which are informative about the other player's state. Our PE still continues to be a sequential equilibrium in this setting because a player does not have to know the other player's state.

Open Issue

Finally, there is one important open question left unanswered. Although we were able to show that a PE can be far more efficient than any PPE, we have not characterized the best private equilibrium payoff yet. This is due to the lack of recursive structure of private monitoring equilibria, which makes the characterization of all private equilibria quite difficult (see Kandori [7]). In general, when PPE payoffs are inefficient, is there also an efficiency bound for private equilibria? Or, do private equilibria achieve full efficiency? This is left as an important topic for future research.

Appendix A: proofs

Proof of Proposition 1.

Proof. Let us first prove a useful lemma which generally holds for the best symmetric PPE payoff (which may be based on asymmetric strategies and public correlation device).

Lemma 14 *Let (v^*, v^*) be the best symmetric PPE payoff for a repeated game with symmetric stage game payoffs. Then there exists a PPE which achieves the same total payoff $2v^*$ and do not use any public correlation device in the initial period. Furthermore, the sum of the expected stage payoffs in the initial period is no less than $2v^*$.[32]*

Proof. When the best symmetric PPE payoff is achieved by public randomization over some PPE, each of them must obtain the same *total* payoff $2v^*$ (otherwise, we can just pick up (v_1, v_2) with the highest total payoff and achieve a higher symmetric payoff by equally randomizing over (v_1, v_2) and (v_2, v_1), a contradiction). Pick up any one of those PPE. By definition, it does not use any public randomization in the first period, and therefore it is achieved by a current (possibly mixed) action profile α and continuation payoffs $(V_1(\omega), V_2(\omega))$ such that

$$2v^* = (1 - \delta)(g_1(\alpha) + g_2(\alpha)) + \delta E\left[V_1(\omega) + V_2(\omega)|\alpha\right],$$

where g_i is player i's payoff function and $E\left[\cdot\,|\alpha\right]$ is the expectation under α.

For the second part, note that, if $g_1(\alpha) + g_2(\alpha) < 2v^*$, the sum of the expected continuation payoffs would be

$$E\left[V_1(\omega) + V_2(\omega)|\alpha\right] > 2v^*$$

This contradicts our assumption that (v^*, v^*) is the best symmetric PPE payoff profile. Hence $g_1(\alpha) + g_2(\alpha) \geq 2v^*$.∎

We continue the proof of Proposition 1. Let $\overline{v}_s\ (> 0)$ be the best symmetric PPE payoff in the repeated partnership game. Lemma 14 implies that (i) there exists a PPE payoff profile $(\overline{v}_1, \overline{v}_2)$ such that $2\overline{v}_s = \overline{v}_1 + \overline{v}_2$, (ii) players do not use a public correlation device in the initial period. Let $\alpha^q = (\alpha_1^q, \alpha_2^q)$ be the mixed action profile in the first period of such PPE.

First, note that Lemma 14 provides an obvious upper bound on $2\overline{v}_s$ (i) if $\alpha^q = (D, D)$, then $g_1(D, D) + g_2(D, D) = 0$ is an upper bound, (ii) if $q \in Q$, then $g^* \geq g_1(\alpha^q) + g_2(\alpha^q)$ is an upper bound, and (iii) if one of α_1^q or α_2^q is D, then $1 + g - l \geq g_1(\alpha^q) + g_2(\alpha^q)$ is an upper bound.

Suppose otherwise, that is, both players are playing C with positive probability and $q \notin Q$. Then the following two inequalities hold for $i = 1, 2$.:

$$\overline{v}_i \ =\ (1 - \delta)\, g_i(C, \alpha_j^q) + \delta\left\{\left(1 - p(\omega(q)|C, \alpha_j^q)\right) V_i\left(\omega'(q)\right)) + p(\omega(q)|C, \alpha_j^q)V_i\left(\omega(q)\right)\right\}$$

$$\overline{v}_i \ \geq\ (1 - \delta)\, g_i(D, \alpha_j^q) + \delta\left\{\left(1 - p(\omega(q)|D, \alpha_j^q)\right) V_i\left(\omega'(q)\right)) + p(\omega(q)|D, \alpha_j^q)V_i\left(\omega(q)\right)\right\}$$

where $\omega'(q) = \Omega/\{\omega(q)\}$. Note that $V_i(\omega'(q)) \geq V_i(\omega(q))$. Replace $V_i(\omega'(q)) - V_i(\omega(q))$ by $\rho_i V_i(\omega'(q))$ $(\rho_i \in [0,1])$ so that the second inequality holds as an equality. Then we obtain

$$
\begin{aligned}
\overline{v}_i &\leq (1-\delta)\,g_i(C,\alpha_j^q) + \delta(V_i(\omega'(q)) - p(\omega(q)|C,\alpha^q)\rho_i V_i(\omega'(q))) \quad (34)\\
\overline{v}_i &= (1-\delta)\,g_i(D,\alpha_j^q) + \delta(V_i(\omega'(q)) - p(\omega(q)|D,\alpha^q)\rho_i V_i(\omega'(q)))
\end{aligned}
$$

which implies

$$
(1-\delta)\,d_i^q \leq \delta\,(p(\omega(q)|D,\alpha^q) - p(\omega(q)|C,\alpha^q))\,\rho_i V_i(\omega'(q))
$$

where $d_i^q = g_i(D,\alpha_j^q) - g_i(C,\alpha_j^q)$. Combining this with (34), we obtain

$$
\frac{\overline{v}_i - \delta V_i(\omega'(q)))}{1-\delta} \leq g_i(C,\alpha_j^q) - \frac{d_i^q}{L_i^q - 1}
$$

where L_i^q is L^q with $(\omega(q),\alpha^q,q)$ being replaced by $(\omega_j(q),\alpha_j^q,q_j)$.

Since $V_1'(\omega(q)) + V_2'(\omega(q)) \leqq 2\overline{v}_s$, the following inequality is obtained by adding the above two inequalities;

$$
\begin{aligned}
2\overline{v}_s &\leq \sum_{i=1,2} g_i(C,\alpha_j^q) - \frac{d_i^q}{L_i^q - 1}\\
&\leq 2v^{q^*}
\end{aligned}
$$

∎

Proof of Proposition 2.

Proof. To show the efficiency of the private equilibrium given above, we need to prove that a root of equation (11) lies in $(0,1)$ and tends to unity as δ tends to 1. At $q = 0$, the left hand side of (11) is strictly positive but the right hand side is equal to zero. Now let q be any number $q' \in (0, \frac{1}{1+h})$ and let δ tends to 1. The left hand side of (11) tends to zero, while the right hand side tends to

$$
q'p(X|D,D)\{1 - q'(1+h)\} > 0
$$

Thus equation (11) has a solution in $(0,q')$ as δ tends to 1, where q' is any number close to 0. ∎

Proof of Proposition 3.

Proof. From (13) and (14), we can obtain

$$
(1-\delta)\{(1-q_R)\,d + q_R h\} = \delta q_R \rho_R\,(p_2 - p_1)\,(V_R - V_P) \quad (35)
$$

As before, we can use this equation to derive the following equation from (13)

$$
V_R = 1 - q_R - q_R h - \frac{(1-q_R)\,d + q_R h}{L^1 - 1} \quad (36)
$$

27

Similarly, we can derive the following two equations from (15) - (16).

$$V_P = 1 - q_P - q_P h + \frac{(1 - q_P) d + q_P h}{L^1 - 1} \frac{1 - p_1}{p_1} \qquad (37)$$

$$(1 - \delta) \{(1 - q_P) d + q_P h\} = \delta q_P \rho_P (p_2 - p_1) (V_R - V_P) \qquad (38)$$

This system of equations is equivalent to (13)-(16).

First note that ρ_R should be set equal to 1. If there exists a solution of these equations with $\rho_R < 1$, then you can reduce q_R and raise ρ_R to increase V_R via (36) while (35) is maintained, and reduce ρ_P so that (38) is still satisfied. In this way, we can obtain another solution with higher equilibrium payoff V_R. Second, q_P can be also set equal to 1. If not, you can increase q_P to reduce V_P via (37), while lowering ρ_R and ρ_P so that (35) and (38) is satisfied. This leads to $V_P = \frac{(1 - p_2) h}{p_2 - p_1}$ from (37).

Now we are left with three equations (35), (36), (38) and three unknowns (q_R, ρ_P, V_R). Once q_R is obtained, V_R is also obtained from (36) and $\rho_P = \frac{q_R h}{(1 - q_R) d + q_R h} \in [0, 1]$ is obtained from (35) and (38). Thus we only need to find q_R in $[0, 1]$.

These three equations reduce to a quadratic equation for q_R;

$$c_2 (\delta) q_R^2 + c_1 (\delta) q_R + c_0 (\delta) = 0$$

where

$$
\begin{aligned}
c_2 (\delta) &= \delta \{p_2 (1 + h) - p_1 (1 + d)\} \\
c_1 (\delta) &= (1 - \delta) (h - d) + \delta \{p_1 d + (1 - p_2) h - (p_2 - p_1)\} \\
c_0 (\delta) &= (1 - \delta) d
\end{aligned}
$$

One root of this quadratic equation is clearly $q_R = 0$ when $\delta = 1$. Since $\frac{\partial F}{\partial q_R}|_{(q_R, \delta) = (0, 1)} \neq 0$ by the assumption $p_2 - p_1 > p_1 d + (1 - p_2) h$, the implicit function theorem can be applied to obtain a C^1 function $q_R (\delta)$ around $\delta = 1$ such that $\frac{dq_R (1)}{d\delta} = -\frac{\frac{\partial F}{\partial \delta}|_{(q_R, \delta) = (0, 1)}}{\frac{\partial F}{\partial q_R}|_{(q_R, \delta) = (0, 1)}} = \frac{d}{p_1 d + (1 - p_2) h - (p_2 - p_1)}$, which is negative by assumption. Thus there exists a $q_R (\delta) \in (0, 1)$ for large enough δ such that $q_R (\delta) \to 0$ as $\delta \to 1$. Hence we get a solution for (35) - (38) parameterized by δ around $\delta = 1$.

Clearly this two state machine generates a sequential equilibrium combined with the belief obtained via Bayes' rule.[33] Since the equilibrium payoff $V_R (\delta)$ converges to $1 - \frac{d}{L^1 - 1}$ as $\delta \to 1$, for any $\eta > 0$ we can find $\underline{\delta}$ such that the equilibrium payoff exceeds $1 - \frac{\alpha}{L^1 - 1} - \eta$ for any $\delta \in (\underline{\delta}, 1)$. ∎

Proof of Proposition 4

---

[33] Belief can be simply derived by Bayes rule at any history. Since any deviation is not observable to the opponent, a player always updates her belief assuming that the opponent has never deviated.

Proof. Since $Q = \emptyset$, we just need to show that $1 - \frac{d}{L^1 - 1} > \max\left\{\max_q v_q, \frac{1+d-h}{2}, 0\right\}$ by Proposition 1. First, $1 - \frac{d}{L^1 - 1} > 0$ follows from $p_2 - p_1 > p_1 d + (1 - p_2) h$ because

$$1 - \frac{d}{L^1 - 1} > \frac{(1 - p_2) h}{p_2 - p_1} > 0$$

Secondly, $1 - \frac{d}{L^1 - 1} > \frac{1+d-h}{2}$ is just assumed. Finally, it is easy to see that $v_q = 1 - q - qh - \frac{(1-q)d + qh}{L^q - 1}$ is decreasing in $q \in [0, 1]$ by $L^1 > L^0$ and $h > d$. Thus

$$
\begin{aligned}
1 - \frac{d}{L^1 - 1} \quad &> \quad 1 - \frac{d}{L^0 - 1} \\
&\geqq \quad 1 - q - qh - \frac{(1 - q) d + qh}{L^q - 1}
\end{aligned}
$$

∎

Proof of Proposition 6

Proof. Suppose first that both M_1 and M_2 has only a finite number of states. Then there exists player 2's state which corresponds to the largest V_1^n. Suppose without loss of generality that $n = 0$. Similarly, let $n = 1$ be the state which minimizes the value function of player 1. We modify player $2's$ machine in the following way. When player 2 is supposed to move to θ_2^n from θ_2^0 after some action and signal is observed, he instead move to θ_2^0 and θ_2^1 with probability $1 - \lambda_2^n$ and λ_2^n where λ_2^n is defined by $V_1^n = (1 - \lambda_2^n) V_1^0 + \lambda_2^n V_1^1$. Then, we obtain the following system of (in)equalities;

$$
\begin{aligned}
For\ n \quad &= \quad 0, 1, \\
\forall a_1 \quad \in \quad & A_1^*, \ V_1^n = (1 - \delta) g_1(a_1, \alpha_2^n) + \\
& \delta \sum_{a_2 \in A_2} \sum_{\omega \in \Omega} \sum_{k=0}^{1} \alpha_2^n(a_2) p(\omega | a_1, a_2) \mu_2'^{nk}(a_2, \omega) V_1^k \\
\forall a_1 \quad \notin \quad & A_1^*, \ V_1^n \geq (1 - \delta) g_1(a_1, \alpha_2^n) + \\
& \delta \sum_{a_2 \in A_2} \sum_{\omega \in \Omega} \sum_{k=0}^{1} \alpha_2^n(a_2) p(\omega | a_1, a_2) \mu_2'^{nk}(a_2, \omega) V_1^k
\end{aligned}
$$

where for $n = 0, 1$

$$
\begin{aligned}
\mu_2'^{n0}(a_2, \omega) \quad &= \quad \sum_{m=0}^{l_2} \mu_2^{nm}(a_2, \omega)(1 - \lambda_2^m) \\
\mu_2'^{n1}(a_2, \omega) \quad &= \quad \sum_{m=0}^{l_2} \mu_2^{nm}(a_2, \omega) \lambda_2^m
\end{aligned}
$$

This defines a two state machine for player 2. We can repeat the same procedure to obtain a two state machine for player 1. This new pair of machines $M_i' = \left\{ \{\theta_i^n\}_{n=0}^1, \ \alpha_i, \ \mu_i' \right\}, i = 1, 2$ clearly satisfies (18), hence constitute a sequential equilibrium which supports the payoff profile $\left(V_1^0, V_2^0\right)$ with the initial state $\left(\theta_1^0, \theta_2^0\right)$.

If the number of the states is countable, we might not able to find the best state and the worst state. In such a case, we construct them in the following way. Suppose that M_2 has a countable number of the states. Since $\{V_1^n\}_{n=0}^{l_2}$ is bounded by assumption, there exists $V_1'^0 = \sup_{n=0,\dots,l_2} \{V_1^n\}$ and $V_1'^1 = \inf_{n=0,\dots,l_2} \{V_1^n\}$ such that $-\infty < V_1'^1 \le V_1'^0 < \infty$. Since V_1^n (hence, $\sum_{m=0}^{l_2} \mu_2^{nm}(a_2, \omega) V_1^m$) and α_2^n are in the compact sets ($\left[V_1'^1, V_1'^0\right]$ and $\triangle_2$ respectively), for $n = 0, 1$, we can find a subsequence $\theta_{2,}^{n(k)}, k = 1, 2, \dots$ such that $V_1^{n(k)} \to V_1'^n$, $\alpha_2^{n(k)} \to \alpha_2'^n$, and $\sum_{m=0}^{l_2} \mu_2^{n(k)m}(a_2, \omega) V_1^m \to \widetilde{V}_1^n(a_2, \omega)$ as $k \to \infty$. Then, $V_1'^n, \alpha_2'^n$ and $\widetilde{V}_1^n(a_2, \omega)$ satisfy

$$
\begin{aligned}
For \ n \ &= \ 0, 1, \\
\forall a_1 \ &\in \ A_1^*, \ V_1'^n = (1 - \delta) \, g_1(a_1, \alpha_2'^n) + \\
&\quad \delta \sum_{a_2 \in A_2} \sum_{\omega \in \Omega} \alpha_2'^n(a_2) \, p(\omega|a_1, a_2) \, \widetilde{V}_1^n(a_2, \omega) \\
\forall a_1 \ &\notin \ A_1^*, \ V_1'^n \ge (1 - \delta) \, g_1(a_1, \alpha_2'^n) + \\
&\quad \delta \sum_{a_2 \in A_2} \sum_{\omega \in \Omega} \alpha_2'^n(a_2) \, p(\omega|a_1, a_2) \, \widetilde{V}_1^n(a_2, \omega)
\end{aligned}
$$

Now we can define the new transition probability by

$$
\widetilde{V}_1^n(a_2, \omega) = \mu_2'^{n0}(a_2, \omega) V_1'^0 + \mu_2'^1(a_2, \omega) V_1'^1
$$

Then we obtain a two state machine $M_2 = \left\{ \{\theta_2^n\}_{n=0}^1, \alpha_2', \mu_2' \right\}$ to satisfy (18) for $V_1'^0$ and $V_1'^1$. We can construct a two state machine M_1 in a similar way and (M_1, M_2) constitutes a sequential equilibrium with the payoff profile $\left(V_1'^0, V_2'^0\right)$.
∎

Proof of Proposition 8

Proof. Consider the following transition rule for player j in the two-state machine (or Markov) strategy: go to state P with probability $\rho_j^z(\omega, a_j)$ when the current state (for j) is $z = R, P$ and the current signal and j's action are ω and a_j (otherwise, go to state R). Consider the dynamic programming equation for the average payoff for player i when j is in state $z = R.P$,

$$
V_i^z \ge (1 - \delta) g_i(a_i, \alpha_j^z) + \delta E[(1 - \rho_j^z(\omega, a_j)) V_i^R + \rho_j^z(\omega, a_j) V_i^P | a_i, \alpha_j^z], \quad (39)
$$

where the equality should be satisfied for $a_i \in \text{supp}\alpha_i^R \cup \text{supp}\alpha_i^P$. Consider first the case $z = R$. Subtracting δV_i^R from both sides and dividing through by

$(1 - \delta)$, we obtain

$$V_i^R \geq g_i(a_i, \alpha_j^R) - E[\frac{\delta}{1 - \delta}\rho_j^R(\omega, a_j)(V_i^R - V_i^P)|a_i, \alpha_j^R],$$

where equality holds for $a_i \in \text{supp}\alpha_i^R \cup \text{supp}\alpha_i^P$. A similar manipulation for state $z = P$ shows

$$V_i^P \geq g_i(a_i, \alpha_j^P) + E[\frac{\delta}{1 - \delta}(1 - \rho_j^P(\omega, a_j))(V_i^R - V_i^P)|a_i, \alpha_j^P],$$

where equality holds for $a_i \in \text{supp}\alpha_i^R \cup \text{supp}\alpha_i^P$. Hence, if we have an equilibrium in the two-state machine strategy, conditions $(19)-(23)$ are satisfied with

$$x_i^R(\omega, a_j) = \frac{\delta}{1 - \delta}\rho_j^R(\omega, a_j)(V_i^R - V_i^P) \text{ and} \tag{40}$$

$$x_i^P(\omega, a_j) = \frac{\delta}{1 - \delta}(1 - \rho_j^P(\omega, a_j))(V_i^R - V_i^P). \tag{41}$$

Conversely, suppose that conditions $(19)-(23)$ are satisfied. Then, (40) and (41) can be satisfied for $\rho_j^z(\omega, a_j) \in [0,1]$, $z = R, P$, for sufficiently high δ. Hence we obtain the equilibrium condition (39) and the two-state machine equilibrium to support payoffs (V_i^R, V_i^P) for $i = 1, 2$. ∎

Proof of Proposition 11

Proof. First we show that $\mathbf{x}_i^Z(A_j^*)$ which satisfies $(19)-(22)$ can be easily found. Take (19) and (20). If information is almost perfect, $P(A_j^*)$ is clearly full row rank. Thus there exists $\mathbf{x}_i^R(A_j^*)$ to satisfy (26) for any V_i^R and $\mathbf{h}_i^R$. To make sure that $\mathbf{x}_i^R(A_j^*)$ is positive, we can choose a large enough numbers $K_i \geq 0$ so that

$$\widehat{\mathbf{x}}_i^R(A_j^*) \equiv \mathbf{x}_i^R(A_j^*) + \begin{pmatrix} K_i \\ \vdots \\ K_i \end{pmatrix} \geq 0 \text{ and}$$

for $i = 1, 2$, and define

$$\widehat{V}_i^R \equiv V_i^R - K_i$$

for $i = 1, 2$. Then we have $\widehat{\mathbf{x}}_i^R(A_j^*) \geq 0$ and $\widehat{V}_i^R$ for $i = 1, 2$ to satisfy (26). The same argument applies to conditions (21) and (22).

Next we need to show that V_i^R and V_i^P can be constructed so that the feasible condition (23) is satisfied . If information is almost perfect with A^* $(cone(P(A_j^*)))$ is almost $\Re_+^{|A_i|}$ for $i \neq j)$[34], V_i^R is roughly given by $\arg\max_{\varepsilon \in \Re, \mathbf{h}_i^R} g_i(\cdot, \alpha_j^R)+$

---

[34] Although what we mean by "close" is obvious, formally we need a metric to define closeness of two cones, say. One example of such distances between two cones E and $F \in \Re_+^n$ would be

$$\|E \cap B_1(\mathbf{0}) - F \cap B_1(\mathbf{0})\|$$

where $\|\cdot\|$ is Hausdorff distance and $B_1(\mathbf{0})$ is an open ball around $\mathbf{0}$ with radius 1.

$\mathbf{h}_i^R - \varepsilon \mathbf{I}$ subject to $g_i\left(\cdot, \alpha_j^R\right) + \mathbf{h}_i^R - \varepsilon \mathbf{I} \subset \Re_+^{|A_i|}$ as it should be maximized to satisfy (26)(see Figure 3 below). Since we can make the slack positive variables $\mathbf{h}_i^R$ as large as possible for any $a_i \notin A_i^*$, this is equal to $\min_{a_i \in A_i^*} g_i\left(a_i, \alpha_j^R\right)$ (which can be negative). On the other hand, we need to make V_i^P as small as possible so that (23) is satisfied. Hence V_i^P is roughly given by $\arg\min_{\varepsilon \in \Re} g_i\left(\cdot, \alpha_j^P\right) + \mathbf{h}_i^P - \varepsilon \mathbf{I}$ subject to $g_i\left(\cdot, \alpha_j^P\right) + \mathbf{h}_i^P - \varepsilon \mathbf{I} \subset -\Re_+^{|A_i|}$. Since $\mathbf{h}_i^P$ can be set to 0 without loss of generality to minimize V_i^P, this is equal to $\max_{a_i \in A_i} g_i\left(a_i, \alpha_j^P\right)$. So, the feasibility condition (25) can be satisfied for some V_i^R and V_i^P if

$$\min_{a_i \in A_i^*} g_i\left(a_i, \alpha_j^R\right) > \max_{a_i \in A_i} g_i\left(a_i, \alpha_j^P\right)$$

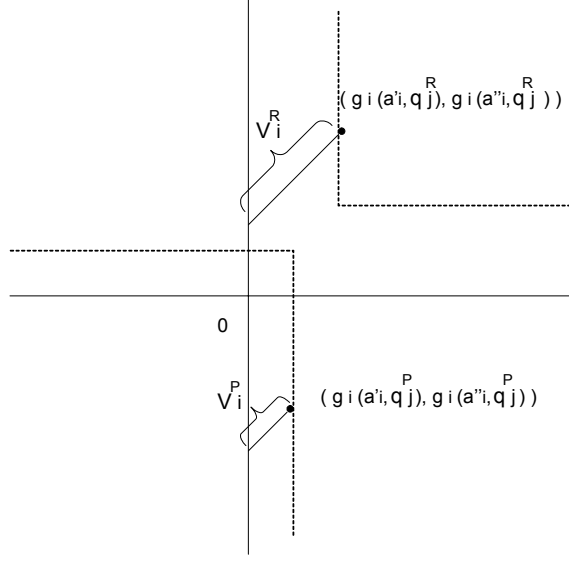when information is almost perfect. This is exactly the condition (25). ∎

Figure 3

Proof of Proposition 12

Proof. The proof is given in two steps. First, we show that (LI) is equivalent to the following condition:

(LR') Conditions in (LR) are satisfied, except that $L_i^Z(a_i)$ is given by

$$L_i^Z(a_i) \equiv \frac{\sum_{\omega, a_j} \lambda_i^Z(\omega, a_j) p(\omega | a_i, a_j) \alpha_j^Z(a_j)}{\sum_{\omega', a_j'} \lambda_i^Z(\omega', a_j') p(\omega' | a_i^Z, a_j') \alpha_j^Z(a_j')}. \tag{42}$$

32

for some $\lambda_i^Z : \Omega \times A_j^Z \to [0,1]$ that is not identically equal to zero.

Theorem 12 is then obtained by this equivalence and the following lemma:

Lemma 15 *Condition (42) is satisfied for some* $\lambda_i^Z : \Omega \times A_j^Z \to [0,1]$ *that is not identically equal to zero if and only if*

$$L_i^Z(a_i) = \sum_{\omega,a_j} \beta_i^Z(\omega,a_j) \frac{p(\omega|a_i,a_j)}{p(\omega|a_i^Z,a_j)}, \tag{43}$$

for some $\beta_i^Z : \Omega \times A_j^Z \to [0,1]$, *such that* $\sum_{\omega,a_j} \beta_i^Z(\omega,a_j) = 1$.

Proof of (LI)$\Longleftrightarrow$(LR'):
(Part I) We show that (LI) implies (LR'). Let us assume that (LI) is satisfied, and define the transfer to player i by

$$t^Z(\omega,a_j) = \begin{cases} -x_i^R(\omega,a_j) & \text{for } Z = R \\ x_i^P(\omega,a_j) & \text{for } Z = P \end{cases}$$

and let

$$T_i^Z \equiv \begin{cases} -\max_{(\omega,a_j)} x_i^R(\omega,a_j) & \text{for } Z = R \\ \max_{(\omega,a_j)} x_i^P(\omega,a_j) & \text{for } Z = P \end{cases}.$$

and also define $\lambda_i^Z : \Omega \times A_j^Z \to [0,1]$ that is not identically equal to zero by

$$\lambda_i^Z(\omega,a_j) \equiv \frac{t_i^Z(\omega,a_j)}{T_i^Z}. \tag{44}$$

We can interpret that $x_i^Z(\omega,a_j)$ effectively imposes fixed fine $T_i^R < 0$ (for $Z = R$) or bonus $T_i^P > 0$ (for $Z = P$) with probability $\lambda_i^Z(\omega,a_j)$, when (ω,a_j) is realized. The probability of getting the fixed fine or bonus T_i^Z given action a_i is given by

$$p_i^Z(a_i) \equiv \sum_{\omega,a_j} \lambda_i^Z(\omega,a_j)p(\omega|a_i,a_j)\alpha_j^Z(a_j). \tag{45}$$

This is always strictly positive, because of the full support assumption and the fact that $\lambda_i^Z(\omega,a_j)$ is not identically equal to zero. Note, by definition, we have

$$-E[x_i^R(\omega,a_j)|a_i,\alpha_j^R] = p_i^R(a_i)T_i^R, \text{ and}$$

$$E[x_i^P(\omega,a_j)|a_i,\alpha_j^P] = p_i^P(a_i)T_i^P.$$

Let $h_i^Z(a_i) \geq 0$ be the slack variables to make the incentive constraints in (LI) binding:

$$\forall a_i \quad V_i^Z = g_i(a_i,\alpha_j^Z) + h_i^Z(a_i) + p_i^Z(a_i)T_i^Z. \tag{46}$$

33

Note that (LI) requires
$$\forall a_i \in A_i^* \quad h_i^Z(a_i) = 0.$$

Let us recall that the following four crucial conditions are satisfied

$$\forall a_i \quad p_i^Z(a_i) > 0, \tag{47}$$

$$V_i^Z = g_i^Z + p_i^Z(a_i^Z)T_i^Z, \tag{48}$$

$$T_i^R < 0, \text{ and} \tag{49}$$

$$T_i^P > 0. \tag{50}$$

Those conditions directly shows $V_i^R < g_i^R$ and $V_i^P > g_i^P$. Since we have $g_i(a_i, \alpha_j^Z) = g_i^Z + d_i^Z(a_i)$ by definition, we have the following equivalence whenever the above four conditions are satisfied:

$$(46) \Longleftrightarrow d_i^Z(a_i) + h_i^Z(a_i) = (p_i^Z(a_i^Z) - p_i^Z(a_i))T_i^Z.$$

$$\Longleftrightarrow^{35} \begin{cases} p_i^Z(a_i^Z) = p_i^R(a_i) & \text{if and only if } d_i^Z(a_i) + h_i^Z(a_i) = 0 \\ \\ T_i^Z = \frac{d_i^Z(a_i) + h_i^Z(a_i)}{p_i^Z(a_i^Z) - p_i^Z(a_i)} & \text{if } d_i^Z(a_i) + h_i^Z(a_i) \neq 0 \end{cases}$$

$$\Longleftrightarrow^{36} \begin{cases} L_i^Z(a_i) = 1 & \text{if and only if } d_i^Z(a_i) + h_i^Z(a_i) = 0 \\ \\ V_i^Z = g_i^Z - \frac{d_i^Z(a_i) + h_i^Z(a_i)}{L_i^R(a_i) - 1} & \text{if } d_i^Z(a_i) + h_i^Z(a_i) \neq 0 \end{cases},$$

where

$$L_i^Z(a_i) \equiv \frac{p_i^Z(a_i)}{p_i^Z(a_i^Z)},$$

so that (42) is satisfied. As (LI) requires $V_i^R > V_i^P$, all conditions in (LR') are satisfied.

(Part II) We now show (LR')$\Rightarrow$(LI). Suppose (LR') holds. First define $p_i^Z(a_i)$ by (45). As λ_i^Z is not identically equal to zero, $p_i^Z(a_i) > 0$ for all a_i (condition (47)) holds. Let us now choose a number T_i^Z to satisfy $V_i^Z =$

---

[35]Because $T_i^Z \neq 0$ by (49) and (50).

[36]By (47), $L_i^Z(a_i)$ is well-defined. The formula for V_i^Z is obtained by (48) as

$$\begin{aligned} V_i^Z &= g_i^Z + p^Z(a_i^Z)T_i^Z \\ &= g_i^Z - p^Z(a_i^Z)\frac{d_i^Z(a_i) + h_i^Z(a_i)}{p_i^Z(a_i) - p_i^Z(a_i^Z)} \\ &= g_i^Z - \frac{d_i^Z(a_i) + h_i^Z(a_i)}{\frac{p_i^Z(a_i)}{p^Z(a_i^Z)} - 1}. \end{aligned}$$

$g_i^Z + p_i^Z(a_i^Z)T_i^Z$ (condition (48)). As we have $V_i^R < g_i^R$ and $V_i^P > g_i^P$, we have $T_i^R < 0$ and $T_i^P > 0$. Hence, conditions (49) and (50) are satisfied. Since (i) the four conditions in Part I, (47), (48), (49), and (50), are satisfied and (ii) $L_i^Z(a_i) \equiv \frac{p_i^Z(a_i)}{p_i^Z(a_i^Z)}$ by (42), we can follow the equivalence relations in Part I to reach equality (46). Now define $x_i^Z(\omega, a_j)$ by $x_i^R(\omega, a_j) = -T_i^R \times \lambda_i^P(\omega, a_j)$ and $x_i^P(\omega, a_j) = T_i^P \times \lambda_i^P(\omega, a_j)$ so that we have $-E[x_i^R(\omega, a_j)|a_i, \alpha_j^R] = p_i^R(a_i)T_i^R$, and $E[x_i^P(\omega, a_j)|a_i, \alpha_j^P] = p_i^P(a_i)T_i^P$. Plugging this into the obtained equality (46), we can show that the incentive constraints in (LI) are satisfied. ∎.

Note that, the slack variables h_i^Z satisfy the requirement in the "Furthermore" part of Theorem 12 by construction and we have

$$L_i^Z(a_i) \equiv \frac{T_i^Z \times p_i^Z(a_i)}{T_i^Z \times p_i^Z(a_i^Z)} = \frac{E[x_i^Z(\omega, a_j)|a_i, \alpha_j^Z]}{E[x_i^Z(\omega, a_j)|a_i^Z, \alpha_j^Z]}.$$

Now we turn to Lemma 15.

Proof of Lemma 15:
Note that

$$L_i^Z(a_i) = \frac{\sum_{\omega, a_j} \lambda_i^Z(\omega, a_j)p(\omega|a_i, a_j)\alpha_j^Z(a_j)}{\sum_{\omega', a_j'} \lambda_i^Z(\omega', a_j')p(\omega'|a_i^Z, a_j')\alpha_j^Z(a_j')}$$

can be rearranged as

$$= \sum_{\omega, a_j} \left(\frac{\lambda_i^Z(\omega, a_j)p(\omega|a_i^Z, a_j)\alpha_j^Z(a_j)}{\sum_{\omega', a_j'} \lambda_i^Z(\omega', a_j')p(\omega'|a_i^Z, a_j')\alpha_j^Z(a_j')} \right) \frac{p(\omega|a_i, a_j)}{p(\omega|a_i^Z, a_j)},$$

and let $\beta_i^Z(\omega, a_j)$ be the number in the parenthesis. Note that $\sum_{\omega, a_j} \beta_i^Z(\omega, a_j) = 1$ by definition and we obtain (43).

We now examine if the converse is true. Now suppose

$$L_i^Z(a_i) = \sum_{\omega, a_j} \beta_i^Z(\omega, a_j)\frac{p(\omega|a_i, a_j)}{p(\omega|a_i^Z, a_j)},$$

for some $\beta_i^Z : \Omega \times A_j^Z \to [0, 1]$, such that $\sum_{\omega, a_j} \beta_i^Z(\omega, a_j) = 1$. Define $\lambda_i^Z(\omega, a_j)$ by

$$K\beta_i^Z(\omega, a_j) = \lambda_i^Z(\omega, a_j)p(\omega|a_i^Z, a_j)\alpha_j^Z(a_j),$$

where $K > 0$ is a sufficiently small number to ensure $\lambda_i^Z(\omega, a_j) \in [0, 1]$. Then, we have

$$L_i^Z(a_i) = \sum_{\omega, a_j} \frac{1}{K}\lambda_i^Z(\omega, a_j)p(\omega|a_i, a_j)\alpha_j^Z(a_j)$$

$$= \frac{\sum_{\omega, a_j} \lambda_i^Z(\omega, a_j) p(\omega|a_i, a_j) \alpha_j^Z(a_j)}{K \sum_{\omega', a_j'} \beta_i^Z(\omega', a_j')} = \frac{\sum_{\omega, a_j} \lambda_i^Z(\omega, a_j) p(\omega|a_i, a_j) \alpha_j^Z(a_j)}{\sum_{\omega', a_j'} \lambda_i^Z(\omega', a_j') p(\omega'|a_i^Z, a_j') \alpha_j^Z(a_j')}.$$

∎

Appendix B: Example 2

Let q_i be the probability that player i chooses action D. Given q_j, the probability of X when player i chooses C and D are respectively $\frac{1}{3}(1 - q_j)$ and $\frac{1}{3}q_j$, as X arises only when both players take the same action. Hence we have the following simple but useful observation.

Lemma 16 *When player i deviates from C to D while the opponent chooses D with probability q_j, outcome X becomes less likely iff $q_j < 1/2$.*

Let F be the sum of the expected stage payoffs under (q_1, q_2),

$$\begin{aligned} F &= (1+1)(1-q_1)(1-q_2) + (2-6)q_1(1-q_2) + (2-6)(1-q_1)q_2 \\ &= 2 - 6q_2 - 6q_1 + 10q_1q_2. \end{aligned}$$

We note that this is positive only if both players choose D with sufficiently low probability.

Lemma 17 *The sum of the stage payoffs is positive only if $q_1, q_2 < 1/3$.*

Proof. Note that $F(q_1, q_2)$ is linear in q_1 and that both $F(0, q_2) = 2 - 6q_2$ and $F(1, q_2) = 4(q_2 - 1)$ are non-positive if $q_2 \geq 1/3$. Hence $F(q_1, q_2)$, which is a convex combination of those values, is non-positive if $q_2 \geq 1/3$. Symmetric argument shows that F is non-positive if $q_1 \geq 1/3$. Hence F is positive only if $q_1, q_2 < 1/3$. ∎

The following is a immediate corollary from the above two lemmata.

Corollary 18 *When the sum of the stage game payoffs is positive, outcome X becomes less likely if player i defects given player j's mixed action.*

Combining Lemma 17 and Corollary 18 with Lemma 14, we have:

Proposition 19 *For any parameter of information structure $\epsilon \in [0, 1/2)$, if the best symmetric PPE payoff v^* is not 0, then there is a (possibly asymmetric) PPE with the same total payoff $2v^*$, where in the first period (i) no public correlation device is used,(ii) each player chooses D with probability less than $1/3$, and (iii) unilateral defection of each player makes outcome X less likely.*

Now we use this fact to show the following.

Proposition 20 *The best symmetric PPE payoff is 0 for all $\delta \in [0, 1)$ when the parameter of the information structure ϵ is equal to 0.*

Proof. Suppose $v^* > 0$ and choose the equilibrium stated in the above Proposition. When $\epsilon = 0$, we can regard Y_1 and Y_2 as a single outcome Y. Note that as D is dominant in the stage game, a player always has a short-term incentive to defect, irrespective of the opponent's mixing probability q_j. Then the above Proposition shows that both payers must be punished when Y realizes. The associated likelihood ratio for player i given player j's mixed action is

$$L_i^{q_j} \equiv \frac{(1-q_j)\Pr(Y|D,C) + q_j\Pr(Y|D,D)}{(1-q_j)\Pr(Y|C,C) + q_j\Pr(Y|C,D)} = \frac{(1-q_j) + \frac{2}{3}q_j}{\frac{2}{3}(1-q_j) + q_j} = \frac{3-q_j}{2+q_j}$$

, and by a similar argument to the proof of Proposition 3, we have

$$2v^* \leq \left(1 - 7q_2 - \frac{1+5q_2}{L_1^{q_2}-1}\right) + \left(1 - 7q_1 - \frac{1+5q_1}{L_2^{q_1}-1}\right). \tag{51}$$

Note that $1 - 7q_j$ is the stage payoff when player i plays C and player j is choosing D with probability q_j, and $(1-q_j) \times 1 + q_j \times 6 = 1 + 5q_j$ is player i's current gain from defection in the same situation. As $L_i^{q_j} \leq 2/3$ for $q_j \leq 1/2$, we have

$$1 - 7q_j - \frac{1+5q_j}{L_i^{q_j}-1} < 1 - \frac{1}{\frac{3}{2}-1} = 1 - 2 < 0 \text{ for } i,j = 1,2 \text{ and } j \neq i.$$

which, together with (51), contradicts our presumption $v^* > 0$. Hence we conclude that best symmetric equilibrium payoff is 0 when $\epsilon = 0$. ∎

Next we derive an upper bound the symmetric PPE payoffs. Let $v^*(\delta)$ be the best symmetric PPE payoff under δ. We suppress δ when no confusion ensues. If v^* is positive, the Proposition 19 shows that there is a PPE achieving the same total payoff $2v^*$, where a possibly mixed action is chosen (but no public correlation device is used) in the first period. Let q_i be the probability that player i chooses action D in the first period $(i = 1, 2)$. The average payoff profile of such an equilibrium, denoted (v_1^0, v_2^0), must satisfy the following "dynamic programming" conditions.

$$v_1^0 + v_2^0 = 2v^* \tag{52}$$

$$v_i^0 = (1-\delta)(1 - 7q_j) + \delta \sum_\omega v_i(\omega)p(\omega|C, q_j), \text{ for } i,j = 1,2 \text{ and } j \neq i \tag{53}$$

$$v_i^0 = (1-\delta)(2 - 2q_j) + \delta \sum_\omega v_i(\omega)p(\omega|D, q_j), \text{ for } i,j = 1,2 \text{ and } j \neq i \tag{54}$$

In the above expression $p(\omega|a, q)$ denotes the probability of ω when a player chooses action a $(a = C, D)$ and the opponent chooses D with probability q (note the symmetry of $p(\omega|\cdot, \cdot)$). The continuation payoff profile is represented by $(v_1(\omega), v_2(\omega))$. Equations (53) and (54) respectively represent player i's payoff when she plays C or D in the first period. Together they imply that player i is indifferent between C and D.

By summing up $(1 - q_i) \times (53) + q_i \times (54)$ for $i = 1, 2$ and using (52), we can calculate the total payoff associated with the mixed strategy profile as

$$2v^* = (1 - \delta)(2 - 6q_2 - 6q_1 + 10q_1q_2) + \delta \sum_\omega (v_1(\omega) + v_2(\omega))p(\omega|q_1, q_2).$$

Note that the first term is $(1 - \delta)$ times the sum of expected stage payoffs, which we formerly defined as F. Also note that $p(\omega|q_1, q_2)$ is the probability of ω when players mix D with probabilities q_1 and q_2. Subtract $2\delta v^* = 2\delta(v_1^0 + v_2^0)$ from both sides and divide by $(1 - \delta)$ to obtain

$$2v^* = (2 - 6q_2 - 6q_1 + 10q_1q_2) + \sum_\omega (\Delta_1(\omega) + \Delta_2(\omega))p(\omega|q_1, q_2), \qquad (55)$$

where $(\Delta_1(\omega), \Delta_2(\omega))$ represents total (as opposed to average) future payoff variations (around the "best" PPE payoff profile (v_1^0, v_2^0)):

$$\Delta_i(\omega) = \frac{\delta}{1 - \delta}(v_i(\omega) - v_i^0), \text{ for } i = 1, 2. \qquad (56)$$

Note that the future payoff variations $(\Delta_1(\omega), \Delta_2(\omega))$ have to satisfy some conditions. First, it must provide right incentive for each player. Subtracting (53) from (54) and dividing through by $(1 - \delta)$, we have (binding) incentive constraints

$$1 + 5q_j = \sum_\omega \Delta_i(\omega)[p(\omega|C, q_j) - p(\omega|D, q_j)], \text{ for } i, j = 1, 2 \text{ and } j \neq i. \qquad (57)$$

Note that the left hand side is the short term gain from defection, while the right hand side shows the reduction of the future payoffs. Secondly, the future payoffs $(v_1(\omega), v_2(\omega))$ should be chosen from the set of PPE payoffs $V^{PPE}(\delta)$. By the definition (56), this condition is represented as

$$\forall \omega \ \frac{1 - \delta}{\delta}(\Delta_1(\omega), \Delta_2(\omega)) + (v_1^0, v_2^0) \in V^{PPE}(\delta) \qquad (58)$$

Let us now summarize what we have found.

Lemma 21 *Let v^* be the best symmetric PPE payoff under discount factor δ. Then, there exist $q_1, q_2 \in [0, 1/2)$ and $(\Delta_1(\omega), \Delta_2(\omega))$ that satisfy the dynamic programming value equation (55), the incentive constraint (57) and the PPE condition (58) for some feasible payoff profile (v_1^0, v_2^0) such that $v_1^0 + v_2^0 = 2v^*$.*

To get an upper bound for v^*, we will relax condition (58). First, let V^F be the feasible payoff set, that is, the convex hull of stage payoffs

$$V^F = Co\{(1, 1), (2, -6), (-6, 2), (0, 0)\}.$$

Note that $V^{PPE}(\delta) \subset V^F$. As $2v^*$ is the maximized sum of the two players' payoffs over $V^{PPE}(\delta)$, we also have $V^{PPE}(\delta) \subset \{v \mid v_1 + v_2 \leq 2v^*\}$. Hence (58) implies

$$\frac{1 - \delta}{\delta}(\Delta_1(\omega), \Delta_2(\omega)) + (v_1^0, v_2^0) \subset V^F \cap \{v \mid v_1 + v_2 \leq 2v^*\}. \qquad (59)$$

The part of the efficient frontier connecting two payoff profiles $(1, 1)$ and $(2, -6)$ is given by

$$7v_1 + v_2 = 8,$$

and by symmetry

$$v_1 + 7v_2 = 8$$

is the other part of the efficient frontier connecting $(1, 1)$ and $(-6, -2)$. Hence any feasible payoff profile v in V^F must satisfy $7v_1 + v_2 \leq 8$ and $v_1 + 7v_2 \leq 8$. Therefore, (59) implies

$$\forall \omega \ \frac{1 - \delta}{\delta}(\Delta_1(\omega), \Delta_2(\omega)) + (v_1^0, v_2^0) \subset \{v \mid v_1 + v_2 \leq 2v^*, \ 7v_1 + v_2 \leq 8, \ v_1 + 7v_2 \leq 8\}. \tag{60}$$

Let us now derive an upper bound of symmetric PPE payoffs. To this end, we first find a lower bound of discount factor to support a symmetric payoff $v^* \in (0, 1)$. Fix any $v^* \in (0, 1)$. Lemma 21 shows that there is a feasible payoff profile (v_1^0, v_2^0) such that $v_1^0 + v_2^0 = 2v^*$. Then, condition (60) implies (by the first inequality on the right hand side) $\frac{1-\delta}{\delta}(\Delta_1(\omega) + \Delta_2(\omega)) + v_1^0 + v_2^0 \leq 2v^*$, which is equivalent to

$$\forall \omega \ \Delta_1(\omega) + \Delta_2(\omega) \leq 0. \tag{61}$$

Also the value equation (55) and Lemma 21 show

$$2v^* - (2 - 6q_2 - 6q_1 + 10q_1q_2) = \sum_\omega (\Delta_1(\omega) + \Delta_2(\omega))p(\omega|q_1, q_2).$$

As $(2 - 6q_2 - 6q_1 + 10q_1q_2)$ is the sum of stage payoffs, it is less than or equal to 2. This and $0 \leq v^*$ imply

$$-2 \leq \sum_\omega (\Delta_1(\omega) + \Delta_2(\omega))p(\omega|q_1, q_2). \tag{62}$$

Let r be the minimum probability of outcome X when players choose D with probabilities $q_1, q_2 \in [0, 1/2]$: $r = \min_{q_1, q_2} p(X|q_1, q_2)$ subject to $q_1, q_2 \in [0, 1/2]$. Note that $p(X|q_1, q_2) < p(Y_i|q_1, q_2)$, $i = 1, 2$ independent of $\epsilon > 0$. Clearly, $r > 0$, and (61) and the definition of r implies $\sum_\omega (\Delta_1(\omega) + \Delta_2(\omega))p(\omega|q_1, q_2) \leq r \min_\omega (\Delta_1(\omega) + \Delta_2(\omega))$. Hence the condition (62) implies $-2 \leq r \min_\omega (\Delta_1(\omega) + \Delta_2(\omega))$. Thus we have another condition for $(\Delta_1(\omega), \Delta_2(\omega))$;

$$\forall \omega \ -2/r \leq \Delta_1(\omega) + \Delta_2(\omega). \tag{63}$$

Now we present a crucial observation that we need large payoff variations of $(\Delta_1(\omega), \Delta_2(\omega))$ in the northwest/southeast directions as $\epsilon \to 0$. That is, as we approach the information structure where the pairwise full rank condition fails, we need large payoff transfers between the players to support a positive payoff v^*.

Lemma 22 *For any (large) $K > 0$, there is (small) $\epsilon > 0$ such that for each $q_1, q_2 \in [0, 1/2]$, if $(\Delta_1(\cdot), \Delta_2(\cdot))$ satisfies conditions (57), (61) and (62), then $\forall \omega \ \Delta_1(\omega), \Delta_2(\omega) \leq K$ cannot hold..*

Proof. Suppose that the assertion is not true. Then, there is a sequence $\{\epsilon^n, \Delta_1^n, \Delta_2^n, q_1^n, q_2^n\}$ such that $\epsilon^n \to 0$, as $n \to \infty$, which satisfies (57), (61), (62), and $\forall \omega \ \Delta_1(\omega), \Delta_2(\omega) \leq K$. The condition (61), (63) implied by (62) and $\forall \omega \ \Delta_1(\omega), \Delta_2(\omega) \leq K$ imply that the sequence lies in a compact set, and we can choose a converging subsequence. Let $(\Delta_1^0, \Delta_2^0 q_1^0, q_2^0)$ be its limit, where (Δ_1^0, Δ_2^0) supports C with probability more than $1/2$ for each player when $\epsilon = 0$. However, since we can regard Y_1 and Y_2 as a single outcome Y when $\epsilon = 0$, the following inequality holds as in Proposition 20.

$$2 - 6q_2 - 6q_1 + 10q_1q_2 + \sum_\omega (\Delta_1^0(\omega) + \Delta_2^0(\omega))p(\omega|q_1, q_2)$$

$$\leq \quad \left(1 - 7q_2 - \frac{1 + 5q_2}{L_1^{q_2} - 1}\right) + \left(1 - 7q_1 - \frac{1 + 5q_1}{L_2^{q_1} - 1}\right)$$

which implies

$$\sum_\omega (\Delta_1^0(\omega) + \Delta_2^0(\omega))p(\omega|q_1, q_2)$$

$$\leq \quad -6q_1 - 6q_2 - 10q_1q_2 - \frac{1 + 5q_2}{L_1 - 1} - \frac{1 + 5q_1}{L_2 - 1}$$

$$< \quad -4,$$

This contradicts the fact that the limit (Δ_1^0, Δ_2^0) also satisfies (62). ∎

Note that given $K > 0$, the choice of ϵ is independent of the initial choice of (v_1^0, v_2^0) and v^* in the above proof. If ϵ chosen is small enough, then $\forall \omega \ \Delta_1(\omega), \Delta_2(\omega) \leq K$ cannot hold for any (v_1^0, v_2^0) and v^*.

Now define

$$A = \{(\Delta_1, \Delta_2) \,|\, \Delta_1 + \Delta_2 \leq 0 \text{ and } -2/r \leq \Delta_1 + \Delta_2\}, \text{ and}$$

$$B(K) = A \cap \{(\Delta_1, \Delta_2) \,|\, \Delta_1, \Delta_2 \leq K\}.$$

Conditions (61), (63) and Lemma 22 implies that we can always choose (small enough) ϵ in such a way that for some ω, $(\Delta_1(\omega), \Delta_2(\omega))$ lies in the region $A \setminus B$. Let us now summarize what we have found as follows.

Proposition 23 *For any (large) $K > 0$, we can find a value of the signal distribution parameter $\epsilon > 0$ for which the following holds: Let $v^* \in (0, 1)$ be the best symmetric PPE payoff under discount factor δ. Then, there exists a feasible payoff profile (v_1^0, v_2^0) such that $v_1^0 + v_2^0 = 2v^*$, where we have*

$$\forall \omega \ \left\{\frac{1 - \delta}{\delta}(A \setminus B(K)) + (v_1^0, v_2^0)\right\} \cap \{v \mid v_1 + v_2 \leq 2v^*, \ 7v_1 + v_2 \leq 8, \ v_1 + 7v_2 \leq 8\} \neq \emptyset.$$

$$(64)$$

As this condition (64) becomes more stringent as $K \to \infty$, if we choose (small) ϵ that corresponds to a large K, we need a fairly large discount factor δ

to support v^*. Note that condition (64) is satisfied if δ is sufficiently large, as in Figure A. Hence, when we have the situation depicted in Figure B with small δ, condition (64) fails for any feasible payoff profile (v_1^0, v_2^0) such that $v_1^0 + v_2^0 = 2v^*$. Therefore, the value of δ given by Figure B is a lower bound of the discount factor that supports the symmetric PPE payoff v^*.
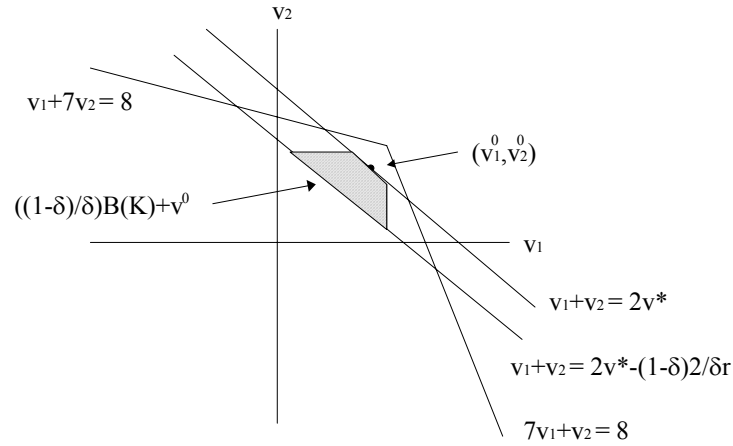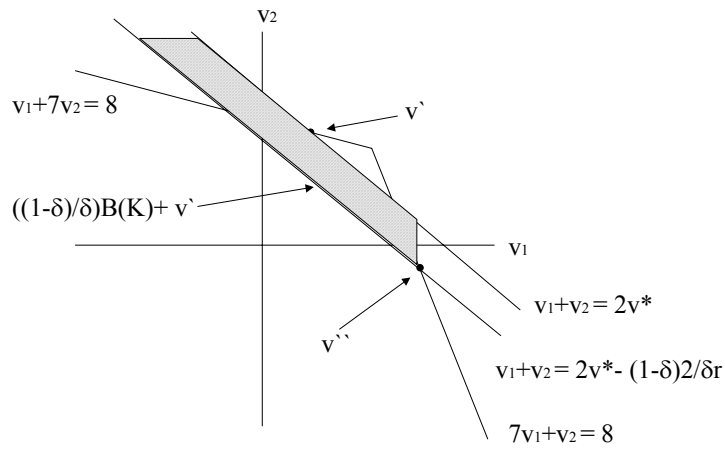
Figure A

Figure B

By the definition of $B(K)$, points v' and v'' in Figure B must satisfy

$$v_1'' - v_1' = \frac{1-\delta}{\delta}K. \tag{65}$$

The value of v_1' is obtained by solving $v_1 + v_2 = 2v^*$ and $v_1 + 7v_2 = 8$, and we find $v_1' = \frac{7v^* - 4}{3}$. Similarly, v_1'' is determined by $v_1 + v_2 = 2v^* - (\frac{1-\delta}{\delta})\frac{2}{r}$ and $7v_1 + v_2 = 8$, and we find $v_1'' = \frac{8 - 2v^* + (\frac{1-\delta}{\delta})\frac{2}{r}}{6}$. By plugging those in equation (65), we obtain a lower bound of the discount factor to support v^*;

$$\delta(v^*) = \frac{3K - \frac{1}{r}}{3K - \frac{1}{r} + 8(1 - v^*)}. \tag{66}$$

Note that this is an increasing function with $\delta(1) = 1$ and $\delta(0) \to 1$ as $K \to \infty$. This means that to support any positive value, we need a fairly large discount factor when the signal distribution parameter ϵ is small (hence K is large). The inverse function of $\delta(\cdot)$,

$$\overline{v}(\delta) = 1 - (\frac{1-\delta}{\delta})\frac{3K - \frac{1}{r}}{8} \tag{67}$$

is concave and depicted in Figure C.
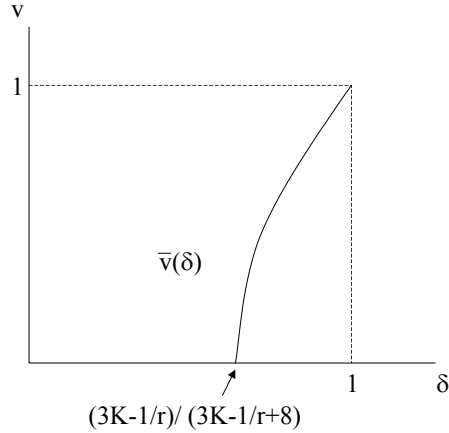

v

1

$\overline{v}(\delta)$

1

δ

(3K-1/r)/ (3K-1/r+8)

Figure C

By the definition of this function, the maximum symmetric PPE payoff under δ must be located to the right of the graph of $\overline{v}(\delta)$, and hence $\overline{v}(\delta)$ is an upper bound of the maximum symmetric PPE payoff under δ, whenever it is positive. Proposition 5 is then given by defining H by $H = \frac{3K - \frac{1}{r}}{8}$.

42

References

[1] D. Abreu, P. Milgrom, and D. Pearce. Information and timing in repeated partnerships. *Econometrica*, 59(6):1713–1733, November 1991.

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

[3] J. C. Ely and J. Välimäki. A robust folk theorem for the prisoner's dilemma. *Journal of Economic Theory*, 102(1):84–105, 2002.

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

[5] E. Green and R. Porter. Noncooperative collusion under imperfect price formation. *Econometrica*, 52:87–100, 1984.

[6] M. Kandori. Check your partners' behavior by randomization: New efficiency results on repeated games with imperfect monitoring. Technical Report CIRJE-F-49, University of Tokyo, 1999.

[7] M. Kandori. Introduction to repeated games with private monitoring. *Journal of Economic Theory*, 102(1):1–15, 2002.

[8] M. Kandori. Randomization, communication, and efficiency in repeated games with imperfect public monitoring. *Econometrica*, 71:345–353, 2003.

[9] M. Kandori and I. Obara. Endogenous monitoring. 2003. mimeo.

[10] M. Kandori and I. Obara. Less is more: An observability paradox in repeated games. 2003. mimeo.

[11] E. Lehrer. Internal correlation in rapeated games. *International Journal of Game Theory*, 19:431–456, 1991.

[12] G. J. Mailath, S. A. Matthews, and T. Sekiguchi. Private strategies in finitely repeated games with imperfect public monitoring. *Contributions to Theoretical Economics*, 2(1), 2002. http://www.bepress.com/bejte/contributions/vol2/iss1/art2.

[13] G. J. Mailath and S. Morris. Repeated games with almost-public monitoring. *Journal of Economic Theory*, 102(1):189–228, 2002.

[14] I. Obara. Private strategy and efficiency: Repeated partnership game revisited. University of Pennsylvania, 1999.

[15] I. Obara. The full surplus extraction theorem with hidden actions. 2003. mimeo.

[16] M. Piccione. The repeated prisoner's dilemma with imperfect private monitoring. *Journal of Economic Theory*, 102(1):70–83, 2002.

[17] R. Radner, R. Myerson, and E. Maskin. An example of a repeated partnership game with discounting and with uniformly inefficient equilibria. *Review of Economic Studies*, 53:59–69, 1986.