# Repeated Games with Costly Imperfect Monitoring*

Eiichi Miyagawa† Kobe University
Yasuyuki Miyahara‡ Kobe University
Tadashi Sekiguchi§ Kyoto University

February 10, 2009

**Abstract**

We study repeated games where players can make costly investments in monitoring and improve quality of information about the other players' actions. We assume players can pay for arbitrarily precise information, but unlike our previous work, we allow the case where perfect information cannot be bought. For repeated prisoners' dilemma, we show that the standard folk theorem obtains in the model with arbitrarily large observation costs. The folk theorem is based on the belief-free approach, and extends existing approximate folk theorems to a setting with endogenous, costly monitoring.

*JEL Classification Numbers:* C72, C73.

*Key Words*: repeated games, imperfect information, costly monitoring, folk theorem.

*Very preliminary.

†Graduate School of Economics, Kobe University, 2-1 Rokkodai-cho, Nada-ku, Kobe 657-8501, Japan; miyagawa@econ.kobe-u.ac.jp.

‡Graduate School of Business Administration, Kobe University, 2-1 Rokkodai-cho, Nada-ku, Kobe 657-8501, Japan; miyahara@kobe-u.ac.jp.

§Institute of Economic Research, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan; sekiguchi@kier.kyoto-u.ac.jp.

## 1 Introduction

Some recent papers have raised the issue of information acquisition in repeated games (Ben-Porath and Kahneman (2003), Kandori and Obara (2004), Miyagawa, Miyahara and Sekiguchi (2008)). On the ground that the apparently exogenous monitoring structure reflects players' efforts to obtain information and that those efforts are never costless, they have formulated models where players have an option to pay costs to acquire perfect or almost perfect information about each other's past actions, in addition to private signals they can obtain costlessly.[1] All those papers show that the folk theorem or efficiency results extend to the case of costly information acquisition.

One common assumption of those papers is that the monitoring decision is binary; each player simply decides whether to obtain (almost) perfect information at a cost or not to obtain additional information at all. This paper relaxes that assumption, and allows players to have much more flexibility in choosing quality of information, possibly at varying cost levels. More concretely, we assume that players have costly options to control the amount of noise the additional information may contain, and that they can pay for options arbitrarily close to the zero noise (perfect monitoring). In other words, we assume costly, almost perfect monitoring. Our formulation includes existing ones as a special case, where players can pay for perfect information (Ben-Porath and Kahneman (2003), Miyagawa, Miyahara and Sekiguchi (2008)). However, ours covers the case where perfect monitoring per se is not purchasable.

We state our results only in the repeated prisoners' dilemma model like Kandori and Obara (2004). At the end of each period, after selecting an action and receiving a costless signal, each player independently and privately decides how much to invest in quality of information he additionally obtains. The additional information is an indicator of the opponent's action, and the more a player invests, the less noise it contains.[2] The information acquisition activities are never observable to the other players. Due to the assumption of costly almost perfect information, the investment levels arbitrarily close to the zero noise, possibly the zero noise itself, are feasible.

Under this setting, we prove that the standard folk theorem extends to the case of costly almost perfect monitoring. That is, given a stage game structure (including monitoring technology), any payoff pair which Pareto dominates the payoff pair of the static equilibrium can be approximated as an equilibrium if the players are arbitrarily patient. Thus the result shows that the folk theorem is robust to extensions where good information must come with costs, and the information may not be completely perfect.

Roughly speaking, existing results on repeated games with costly observations can be compared from three dimensions: (i) whether the costly observation is perfect or

[1] Among others, realized period-payoffs are a typical example of the free signals.

[2] We assume that the players anyway obtain the information even if they invest nothing in monitoring. In this case, the information will be uninformative at all.

almost perfect, (ii) whether observational decisions are (almost) observable or not, and (iii) whether the observational decision takes place simultaneously with the choice of stage-actions, or it takes place after the stage-action is selected (namely, simultaneous or sequential observational decision). The first paper in the literature is Ben-Porath and Kahneman (2003), and they assume costly perfect monitoring with simultaneous, unobservable decisions. Ben-Porath and Kahneman (2003) also assume that explicit communication is available. Miyagawa, Miyahara and Sekiguchi (2008) assume costly perfect monitoring with sequential, unobservable decisions, but they do not assume explicit communication. Both papers prove a standard, minmax folk theorem for general stage games. Kandori and Obara (2004) consider the case of costly imperfect monitoring, with simultaneous and almost observable decisions. They show that in the limit of vanishing noise in costly observations, cooperation is sustainable in the prisoners' dilemma or its generalization. This paper assumes costly almost perfect monitoring with sequential, unobservable decisions, but limits attention to the two-player prisoners' dilemma. Therefore all those papers report independent results.

In our model, due to lack of perfect observations, it is difficult to coordinate future play as in the construction by Ben-Porath and Kahneman (2003) which uses explicit communication, or the construction by Miyagawa, Miyahara and Sekiguchi (2008) which uses stage actions and mutual monitoring for implicit communication. Thus, like Kandori and Obara (2004), we invoke the belief-free approach in repeated games with imperfect monitoring (Piccione (2002), Ely and Valimaki (2002), Ely, Horner and Olszewski (2005), and Kandori and Obara (2006)). Furthermore, due to unobservability of observational decisions, it is difficult to provide incentives to invest in monitoring. That is why our construction has an examination state, as in Miyagawa, Miyahara and Sekiguchi (2008), whose details are explained later.

Here we briefly explain the main idea of our construction, assuming that the target payoff pair is Pareto inferior to the payoff pair of mutual cooperation. As in Ely and Valimaki (2002), the equilibrium strategies are state-based ones, but in our construction each strategy has two cooperation states and one *examination state*. Each player randomly starts play with one of the cooperation states, and its probability distribution determines the opponent's payoff. The cooperation states are either a *strong cooperation state*, where the player cooperates with a probability close to 1, or a *weak cooperation state*, where he rather defects with a probability close to 1. In either state, observational decisions are coordinated by a public randomization device. The players do not invest in monitoring at all and remain in the same cooperation state in the next period with some large probability. With the remaining, small probability, each player is prescribed to invest in monitoring for very precise information. After the mutual monitoring, the state shifts to the examination state in the next period. It is important that the public randomization device is available *in the middle of the period*. In order to reduce expected monitoring costs, our construction relies on random monitoring, and there must be uncertainty as to whether a player is monitored when he selects a stage-action.

The crux of our construction is the introduction of the examination state, which always follows after a cooperation state with mutual monitoring. In this state, a public randomization device selects one player as an *examiner*, the other as an *examinee*. The sunspot also decides the type of the examination; in one type, the examinee is expected to answer the examiner's action in the previous period, and in the other type, he is expected to answer what the examiner did not play in that period. Note that if the examinee had chosen the prescribed investment level in the previous period, he can choose a correct answer almost surely. The examiner decides whether the examinee answers correctly or not, based on her costless private signal.

In the examination state, no player invests in monitoring, and the state in the next period is one of the cooperation states. Which cooperation state is selected depends on the role of a player and on what happened in the previous two periods. If a player is an examinee, then he stochastically selects the state so as to keep the opponent indifferent over cooperation and defection in the previous cooperation period. This ensures belief-freeness of the play in the cooperation states. If he is an examiner, then he stochastically chooses the state so that the opponent apparently sending a wrong answer is punished, and the opponent apparently sending a correct answer is rewarded. The probabilities are selected so that the opponent has an incentive to invest in monitoring in the previous period.[3]

Though a primary contribution of our folk theorem is to establish a robustness of existing results on repeated games with observation activities, it also has some implications on repeated games with private monitoring, where monitoring is given exogenously. Starting from Sekiguchi (1997), much of the literature on private monitoring deals with *approximate* folk theorems (or efficiency results): the target payoff vectors can be approximated as an equilibrium if the players are sufficiently patient *and if the underlying monitoring structure is sufficiently close to perfect monitoring.*[4] In other words, this line of research first fixes expected payoffs, but not stage games. This is in sharp contrast with standard folk theorems under public monitoring (Fudenberg, Levine and Maskin (1994)), where the stage game is fixed together with its public monitoring structure. Notable exceptions are Matsushima (2004) and Yamamoto (2007), who study the case where private signals are conditionally independent or satisfy some similar conditions, and prove standard folk theorems when the stage game is a prisoners' dilemma or its generalization.[5] Therefore our model can be regarded as another class of stage games for which the standard folk theorem holds.

[3] Due to the introduction of examination states, our equilibrium is not a belief-free equilibrium in the literal sense (Ely, Horner and Olszewski (2005)). Since the correct answer depends on the opponent's history, the continuation play in an examination state does not satisfy belief-freeness.

[4] See Bhaskar and Obara (2002), Ely and Valimaki (2002), Piccione (2002) and Horner and Olszewski (2006), for instance.

[5] A recent paper by Fong, Gossner, Horner and Sannikov (2007) deals with a different monitoring structure.

Our folk theorem also has a more direct implication on the approximate folk theorems under private monitoring. To see that, suppose players can choose their monitoring investments *costlessly*. Then we can prove the same folk theorem by just applying the argument of Ely and Valimaki's (2002) approximate folk theorem by belief-free strategies. This fact implies that existing approximate folk theorems by belief-free strategies can be interpreted as a standard folk theorem in models with costless monitoring activities where almost perfect observations are available. Thus our theorem is also an extension of those results to settings with costly information acquisition, thereby establishing a robustness of that interpretation.

The rest of this paper is organized as follows. Section 2 introduces the model of repeated prisoners' dilemma with costly imperfect observations. Section 3 reports the main result, a folk theorem for the repeated prisoners' dilemma. Section 4 offers some discussions on the result.

# 2 Model

The stage game is a simple two-player prisoners' dilemma, and the only complication is asymmetry. Let $A_1 = A_2 = \{C, D\}$ be each player's stage action set. Let $\Omega_i$ be player $i$'s signal space, which is assumed to be finite only for simplicity. The (costless) monitoring structure is characterized by $P(\omega|a)$, the probability of a signal profile $\omega = (\omega_1, \omega_2)$ under an action profile $a = (a_1, a_2)$.

Player $i$'s realized stage-payoff (disregarding observation costs) depends only on his action and signal, and is described by a function $\pi_i : A_i \times \Omega_i \to \mathbb{R}$. It is standard to define each player $i$'s expected payoff function $u_i : A_i \times A_j \to \mathbb{R}$, which we assume is represented as the following prisoners' dilemma in Figure 1. We assume $g_i > 0$, $l_i > 0$

| | | Player 2: $C$ | Player 2: $D$ |
|---|---|---|---|
| Player 1 | $C$ | $1, 1$ | $-l_1, 1+g_2$ |
| | $D$ | $1+g_1, -l_2$ | $0, 0$ |

Figure 1: The stage game

and $g_i - l_j < 1$ for each $i$ and $j \neq i$. Hence $D$ is a dominant action for each player, and $(C, C)$ is efficient.

The marginal distribution of $P(\omega|a)$ for each $\omega_i$ is denoted by $P_i(\omega_i|a)$.

**Assumption 1 (individual full support)**

$$\underline{p} \equiv \min_i \min_a \min_{\omega_i \in \Omega_i} P_i(\omega_i \mid a) > 0.$$

Assumption 1 is standard under imperfect monitoring. No signal reveals that certain action pairs have not been played.

Players make observational decisions after they choose actions and receive private signals.[6] We assume that each player obtains additional information after his observational activity, and the observational decision affects quality of the information. We call this additional information an "observation," in order to distinguish it from the costless signal $\omega_i$. The observational decision of player $i$ consists of choosing from a set of *monitoring investments* $M_i$. We assume each $M_i$ is a subset of an interval $[1/2, 1]$. The set of observations for each player is $\{C, D\}$. We assume that if player $i$ chooses a certain monitoring investment $m_i \in M_i$, then the probability that his observation equals his opponent's action is always $m_i$. The observations are independent across players, and are independent of any other information.

The monitoring investments are costly, and $c_i : M_i \to \mathbb{R}_+$ denotes each player $i$'s cost function; it costs $c_i(m_i)$ to player $i$ choosing $m_i \in M_i$. The costly monitoring structure satisfies the following assumptions.

**Assumption 2** *For each $i$,*

*(i)* $1/2 \in M_i$ *and* $c_i(1/2) = 0$*,*

*(ii)* 1 *belongs to the closure of* $M_i$*, and*

*(iii)* *there exist an increasing sequence on* $M_i$*,* $(\rho_n)_{n=1}^{\infty}$*, and an increasing sequence of positive numbers,* $(\kappa_n)_{n=1}^{\infty}$*, such that:*

$$\lim_{n\to\infty} \rho_n = 1, \quad \lim_{n\to\infty} \kappa_n = \infty,$$
$$c_i(m_i) \geq \kappa_n(m_i - \rho_n) + c_i(\rho_n) \ \forall n \ \forall m_i \in M_i$$

Assumption 2(i) simply means that it is always possible to buy no information. Since the observation is uninformative at all if $m_i = 1/2$, this is equivalent to no information acquisition. Assumption 2(ii) is our main assumption of almost perfect monitoring. Each player can get as precise information as possible if he wishes. It is trivially satisfied if $1 \in M_i$; that is, perfect information is purchasable. Finally, Assumption 2(iii) plays an important role in proving our folk theorem: it states that we can always find an arbitrarily steep slope with which some point in the graph of the function $c_i$ is tangent. Assumption 2 is satisfied if either $M_i = [1/2, 1)$ or $M_i = [1/2, 1]$, and if $c_i$ is an increasing, convex function with $\lim_{m_i \to 1} c_i'(m_i) = \infty$.[7]

Let us compare our formulation with existing ones. Ben-Porath and Kahneman (2003) and Miyagawa, Miyahara and Sekiguchi (2008) consider the case of $M_i = \{1/2, 1\}$. That is, the choice of monitoring investments is binary in those models, a choice

[6] This assumption on timing follows that of Miyagawa, Miyahara and Sekiguchi (2008). Ben-Porath and Kahneman (2003) and Kandori and Obara (2004) rather assume that players choose a stage-action and make an observational decision simultaneously.

[7] A restriction of such a function to $M_i$ satisfying Assumption 2(i)(ii) also satisfies the assumption.

between no information acquisition and obtaining perfect information. Note that Assumption 2(iii) is satisfied, if we set $\rho_n = 1$ and $\kappa_n$ arbitrarily large for each $n$.[8] Kandori and Obara (2004) consider the case $M_i = \{1/2, 1-\varepsilon\}$, and study the limit case of making $\varepsilon \to 0$.[9] Clearly one advantage of our model is to allow more diversity in monitoring decisions and to allow the case where one cannot buy perfect information.

The infinitely repeated game with the above stage game and a common discount factor $\delta \in (0,1)$ is denoted by $G(\delta)$. We assume availability of two types of sunspots each period; the sunspot at the beginning of the period, and the sunspot in the middle of the period. The latter realizes just before observational decisions, so that the players can coordinate their decisions. The players' payoff criteria are average discounted sums, and the solution concept is sequential equilibrium.

## 3 Results

The purpose of this section is to prove a folk theorem for our repeated prisoners' dilemma with observational decisions. Namely, any interior feasible and individually rational payoff pair can be sustained by a sequential equilibrium if players are sufficiently patient. Note that the feasible and individually rational payoff pairs are:

$$V^* \equiv \text{convex hull of } \Big\{(1,1), \Big(\frac{g_1 + l_2 + 1}{l_2 + 1}, 0\Big), \Big(0, \frac{g_2 + l_1 + 1}{l_1 + 1}\Big), (0,0)\Big\}.$$

We prove the folk theorem, following the two-step argument in Ely and Valimaki (2002). In the first step, we prove that any interior point of the rectangle with the vertices $\{(1,1),(1,0),(0,1),(0,0)\}$ is sustained as an equilibrium if the players are sufficiently patient. This in particular implies that two vertices of $V^*$, $(1,1)$ and $(0,0)$, are approximately sustained by an equilibrium. In the second step, we prove that a payoff pair arbitrarily close to $\big(\frac{g_1+l_2+1}{l_2+1}, 0\big)$ is sustained by an equilibrium if the players are sufficiently patient. Since a symmetric argument works for $\big(0, \frac{g_2+l_1+1}{l_1+1}\big)$, we obtain the folk theorem due to availability of the public randomization device.

**Proposition 1** *Fix $\varepsilon > 0$. Then there exists $\underline{\delta} \in (0,1)$ such that any payoff pair $(v_1, v_2) \in [\varepsilon, 1-\varepsilon] \times [\varepsilon, 1-\varepsilon]$ is a sequential equilibrium payoff pair of $G(\delta)$ with any $\delta \geq \underline{\delta}$.*

**Proof.** First of all, for any fixed $a_i$, $u_i(a_i, a_j)$ nontrivially depends on $a_j$, because $g_i > 0$ and $l_i > 0$ for each $i$. Since a player's realized stage-payoff depends only on his action and signal, this particularly implies that

$$P_i(\,\cdot\mid a_i = D, a_j = C) \neq P_i(\,\cdot\mid a_i = a_j = D).$$

[8] More generally, in case of $1 \in M_i$, then Assumption 2 is satisfied if there exists *no* sequence $(\rho_n)_{n=1}^{\infty}$ such that $\lim_{n\to\infty} \rho_n = 1$ and $\lim_{n\to\infty} c_i(\rho_n) < c_i(1)$.

[9] Thus precisely speaking, their model does not satisfy Assumption 2.

Hence for each $i$, there exists a subset of $\Omega_i$, $\Omega_i^C$, such that

$$P_i\big(\Omega_i^C \mid a_i = D, a_j = C\big) > P_i\big(\Omega_i^C \mid a_i = a_j = D\big). \tag{1}$$

Let us define $\Omega_i^D \equiv \Omega_i \setminus \Omega_i^C$, and

$$P_i^C \equiv P_i\big(\Omega_i^C \mid a_i = D, a_j = C\big),$$
$$P_i^D \equiv P_i\big(\Omega_i^D \mid a_i = a_j = D\big).$$

Note that (1) is equivalent to $P_i^C + P_i^D > 1$.

Fix $\varepsilon > 0$. Choose a small $\eta > 0$ so that for each $i$ and $j \neq i$,

$$u_i(C, \alpha_j^S) > 1 - \frac{\varepsilon}{2}, \quad u_i(D, \alpha_j^W) < \frac{\varepsilon}{2}, \tag{2}$$

where

$$\alpha_i^S = (1-\eta)\cdot C + \eta \cdot D, \quad \alpha_i^W = \eta \cdot C + (1-\eta)\cdot D.$$

We then define

$$\hat{\eta} = \frac{\eta \underline{p}}{\eta \underline{p} + (1-\eta)(1-\underline{p})}.$$

By Assumption 1, we have $\hat{\eta} > 0$.

By Assumption 2(ii)(iii), there exist $\kappa_i^* > 0$ and $\rho_i^* \in M_i$ such that:

$$2\kappa_i^* > l_i, \tag{3}$$
$$\rho_i^* > 1 - \hat{\eta}, \tag{4}$$
$$\frac{1-\rho_i^*}{2\rho_i^* - 1}\max\{g_1, l_1, g_2, l_2\} < \frac{\varepsilon}{2}, \tag{5}$$
$$\kappa_i^* \rho_i^* - c_i(\rho_i^*) > \kappa_i^*(1-\hat{\eta}), \tag{6}$$
$$\kappa_i^* \rho_i^* - c_i(\rho_i^*) \geq \kappa_i^* m_i - c_i(m_i) \quad \forall m_i \in M_i \tag{7}$$

To see that (6) can be satisfied, note that Assumption 2(ii) implies existence of $\hat{m}_i \in M_i$ such that $\hat{m}_i > 1 - (\eta/2)$. Therefore, we can choose $\kappa_i^*$ so large as to satisfy

$$\kappa_i^* \hat{m}_i - c_i(\hat{m}_i) \geq \kappa_i^*(1-\hat{\eta}).$$

Then (7) applied for $\hat{m}_i$ implies (6).

Next we choose a small $\mu > 0$ and $\underline{\delta} \in (0,1)$ so that for any $i$, any $j \neq i$ and any

$\delta \geq \underline{\delta}$, all the following inequalities are satisfied.

$$
\begin{aligned}
V_i^S \equiv \frac{1}{1+\delta\mu}\Bigg[ & u_i(C,\alpha_j^S) - \mu c_i(\rho_i^*) - \frac{1-\rho_j^*}{2\rho_j^*-1}\{u_i(D,\alpha_j^S) - u_i(C,\alpha_j^S)\} \\
& + \frac{\delta\mu}{4}\left(1 + g_i - l_i + \frac{P_j^C - P_j^D}{P_j^C + P_j^D - 1} l_i\right) \\
& - \mu \frac{\kappa_i^*}{P_j^C + P_j^D - 1}\left\{\rho_i^* \frac{2 - P_j^C - P_j^D}{2} + (1-\rho_i^*)\frac{P_j^C + P_j^D}{2}\right\}\Bigg] \\
> 1 - \varepsilon, & \qquad (8)
\end{aligned}
$$

$$
\begin{aligned}
V_i^W \equiv \frac{1}{1+\delta\mu}\Bigg[ & u_i(D,\alpha_j^W) - \mu c_i(\rho_i^*) + \frac{1-\rho_j^*}{2\rho_j^*-1}\{u_i(D,\alpha_j^W) - u_i(C,\alpha_j^W)\} \\
& + \frac{\delta\mu}{4}\left(1 + g_i - l_i + \frac{P_j^C - P_j^D}{P_j^C + P_j^D - 1} l_i\right) \\
& + \mu \frac{\kappa_i^*}{P_j^C + P_j^D - 1}\left\{\rho_i^* \frac{P_j^C + P_j^D}{2} + (1-\rho_i^*)\frac{2 - P_j^C - P_j^D}{2}\right\}\Bigg] \\
< \varepsilon, & \qquad (9)
\end{aligned}
$$

$$
(1-\delta)\frac{2\kappa_j^* + \delta l_j}{\delta^2(P_i^C + P_i^D - 1)(V_j^S - V_j^W)} \in (0,1), \qquad (10)
$$

$$
2(1-\delta)\frac{\max\{g_j, l_j\}}{\mu\delta^2(2\rho_i^* - 1)(V_j^S - V_j^W)} \in (0,1). \qquad (11)
$$

To see that such $\mu$ and $\underline{\delta}$ indeed exist, note first that because of (2) and (5), (8) and (9) evaluated at $\delta = 1$ are satisfied if $\mu > 0$ is sufficiently small. Hence under this $\mu$, there exists $\underline{\delta}$ such that (8)–(11) are all satisfied if $\delta \geq \underline{\delta}$.

Fix a $\delta \geq \underline{\delta}$, and $(v_1, v_2) \in [\varepsilon, 1-\varepsilon] \times [\varepsilon, 1-\varepsilon]$. We prove the claim by first showing that there exists a *Nash equilibrium* of $G(\delta)$, denoted by $\hat{\sigma} = (\hat{\sigma}_1, \hat{\sigma}_2)$, which sustains the payoff pair $(v_1, v_2)$. We then show that $\hat{\sigma}$ has an outcome-equivalent sequential equilibrium. Each $\hat{\sigma}_i$ consists of the following three (private) states; Strong-Cooperation ($S$), Weak-Cooperation ($W$) and Examination ($E$). The play in each state and the transition rule is as follows.

*Cooperation States ($S$ or $W$).* If player $i$ is in the $Z$-Cooperation state ($Z \in \{S, W\}$), he plays $\alpha_i^Z$ in that period. If the realization of the middle-of-the-period sunspot is $y \geq \mu$, then he makes no monitoring investment; namely, he chooses $m_i = 1/2$. In this case, the next period remains to be in the same $Z$-Cooperation state. If $y < \mu$, then he chooses $\rho_i^*$, and the next period moves to the Examination state ($E$).

*Examination States ($E$).* This state always follows from a cooperation state in which the sunspot prescribed the monitoring investment of $\rho_i^*$ (i.e., $y < \mu$ in the previous period). In state $E$, the sunspot at the beginning of the period selects a number $e \in \{-2, -1, 1, 2\}$ equiprobably, on which each player $i$'s behavior in this period depends.

Let $a_j^{obs}$ be player $i$'s observation in the previous cooperation period. Then the action player $i$ should select, denoted by $b_i^*$, depends on $a_j^{obs}$ and $e$ in the following way.

$$b_i^* = \begin{cases} C & \text{if } (a_j^{obs}, e) \in \{(C, i), (D, -i)\}, \\ D & \text{otherwise.} \end{cases} \tag{12}$$

First, player $i$ plays what he observed in the previous period if the sunspot selects his number. Second, he plays what he did *not* observe in the previous period if the sunspot selects the minus of his number. Third, if $|e| = j$, then he plays $D$.

The state in the next period is either $S$ or $W$, which depends on the cooperation state player $i$ was in the previous period ($Z$), the pair of player $i$'s own action and observation $(a_i^{own}, a_j^{obs})$ in the previous period, the sunspot's selection in the current period ($e$), and $i$'s signal in the current period ($\omega_i$).

(i) Suppose player $i$ was in $Z$-Cooperation in the previous period, where he played $a_i^{own}$, $e \in \{-j, j\}$ was selected, and he observed $\omega_i$. Then he moves to $Z$-Cooperation with probability $1 - \zeta_i(Z, a_i^{own}, e, \omega_i)$, and moves to the other Cooperation state with probability $\zeta_i(Z, a_i^{own}, e, \omega_i)$, where:

$$\begin{aligned} \zeta_i(S, a_i^{own}, e, \omega_i) &= \begin{cases} q_i^C & \text{if } (a_i^{own}, e) \in \{(C, j), (D, -j)\} \text{ and } \omega_i \in \Omega_i^D, \\ q_i^D & \text{if } (a_i^{own}, e) \in \{(C, -j), (D, j)\} \text{ and } \omega_i \in \Omega_i^C, \\ 0 & \text{otherwise,} \end{cases} \\ \zeta_i(W, a_i^{own}, e, \omega_i) &= \begin{cases} q_i^C & \text{if } (a_i^{own}, e) \in \{(C, j), (D, -j)\} \text{ and } \omega_i \in \Omega_i^C, \\ q_i^D & \text{if } (a_i^{own}, e) \in \{(C, -j), (D, j)\} \text{ and } \omega_i \in \Omega_i^D, \\ 0 & \text{otherwise,} \end{cases} \end{aligned} \tag{13}$$

where $q_i^C$ and $q_i^D$ are defined as follows.

$$q_i^C = q_i + \Lambda_i, \quad q_i^D = q_i - \Lambda_i, \tag{14}$$

$$q_i \equiv \frac{2(1-\delta)\kappa_j^*}{\delta^2 (P_i^C + P_i^D - 1)(V_j^S - V_j^W)}, \tag{15}$$

$$\Lambda_i \equiv \frac{(1-\delta) l_j}{\delta (P_i^C + P_i^D - 1)(V_j^S - V_j^W)}. \tag{16}$$

Note that $q_i^C$ and $q_i^D$ are a probability due to (3) and (10).

(ii) Suppose player $i$ was in $Z$-Cooperation in the previous period, where he played $a_i^{own}$ and observed $a_j^{obs}$, and $e \in \{-i, i\}$ was selected. Then irrespective of the outcome of the current period, he moves to $Z$-Cooperation with probability

$1-\xi_i(Z, a_i^{own}, a_j^{obs})$, and moves to the other Cooperation state with probability $\xi_i(Z, a_i^{own}, a_j^{obs})$, where:

$$\begin{aligned} \xi_i(S,C,D) = \xi_i(W,C,C) &= \frac{2(1-\delta)g_j}{\mu\delta^2(2\rho_i^*-1)(V_j^S - V_j^W)}, \\ \xi_i(S,D,D) = \xi_i(W,D,C) &= \frac{2(1-\delta)l_j}{\mu\delta^2(2\rho_i^*-1)(V_j^S - V_j^W)}, \\ \xi_i(Z, a_i^{own}, a_j^{obs}) = 0 &\quad \text{if } (Z, a_j^{obs}) \in \{(S,C),(W,D)\}. \end{aligned} \tag{17}$$

Note that because of (11), we always have $\xi_i(Z, a_i^{own}, a_j^{obs}) \in [0,1)$.

*Initial Play.* For each $i$ and $j \neq i$, let $\lambda_i \in (0,1)$ be such that

$$\lambda_i V_j^S + (1-\lambda_i) V_j^W = v_i. \tag{18}$$

$\lambda_i$ exists because of (8) and (9). Then each player $i$ stochastically selects the initial state, and chooses $S$-Cooperation state with probability $\lambda_i$ and $W$-Cooperation state with probability $1-\lambda_i$.

We first compute the payoff of the strategy profile $\hat{\sigma} = (\hat{\sigma}_1, \hat{\sigma}_2)$, as well as its continuation payoffs. Note that whenever a player is in a Cooperation state, the other player is also in a Cooperation state, given the strategy profile.[10] Let $V_i^{\beta,C}$ be player $i$'s payoff when:

(i) player $i$ believes that player $j$ is in $S$-Cooperation with probability $\beta$ and in $W$-Cooperation with probability $1-\beta$,

(ii) player $j$ follows $\hat{\sigma}_j$, and

(iii) player $i$ follows the transition rule among states and follows $\hat{\sigma}_i$ in any Examination state and with respect to any observational decision, but he always plays $C$ in any Cooperation state.

Note that since player $i$ always plays $C$ in any cooperation state, it is irrelevant whether he is in $S$ or $W$ in the current period. Similarly, we define $V_i^{\beta,D}$, with the only difference being that the action in (iii) is replaced with $D$.

Suppose that player $i$ believes that player $j$ is in $S$-Cooperation with probability $\beta$ and in $W$-Cooperation with probability $1-\beta$, and player $i$ plays $a_i$. Depending on the sunspot in the middle of the period, player $j$'s state remains unchanged with probability $1-\mu$. So player $i$'s belief also remains to be $\beta$. With probability $\mu$, in contrast, the play moves to $E$. Given that, let $\hat{\beta}^{a_i}$ be the probability with which player $i$

[10] However, their cooperation states need not be the same.

believes that player $j$ is in $S$-cooperation in the next cooperation period. Using this notation, we obtain the following value equations.

$$\begin{aligned} V_i^{\beta,C} =&(1-\delta)\Big[u_i(C,\beta\alpha_j^S+(1-\beta)\alpha_j^W)-\mu c_i(\rho_i^*)\Big] \\ &+\delta\Big[(1-\mu)V_i^{\beta,C}+\mu\Big\{(1-\delta)\frac{1+g_i-l_i}{4}+\delta V_i^{\hat{\beta}^C,C}\Big\}\Big], \end{aligned} \tag{19}$$

$$\begin{aligned} V_i^{\beta,D} =&(1-\delta)\Big[u_i(D,\beta\alpha_j^S+(1-\beta)\alpha_j^W)-\mu c_i(\rho_i^*)\Big] \\ &+\delta\Big[(1-\mu)V_i^{\beta,D}+\mu\Big\{(1-\delta)\frac{1+g_i-l_i}{4}+\delta V_i^{\hat{\beta}^D,D}\Big\}\Big]. \end{aligned} \tag{20}$$

(19) and (20) are functional equations for $V_i^{\cdot,C}$ and $V_i^{\cdot,D}$, respectively.

If we define

$$\begin{aligned} \xi_j(S,\alpha_j^S,D) &= (1-\eta)\xi_j(S,C,D)+\eta\xi_j(S,D,D), \\ \xi_j(W,\alpha_j^W,C) &= \eta\xi_j(W,C,C)+(1-\eta)\xi_j(W,D,C), \end{aligned}$$

we can compute $\hat{\beta}^C$ and $\hat{\beta}^D$ as follows, by using (13) and (14):

$$\begin{aligned} \hat{\beta}^C =&\beta\left[1-\frac{1}{2}(1-\rho_j^*)\xi_j(S,\alpha_j^S,D)-\frac{1}{2}\left\{\rho_i^*\frac{2-P_j^C-P_j^D}{2}+(1-\rho_i^*)\frac{P_j^C+P_j^D}{2}\right\}q_j\right] \\ &+(1-\beta)\left[\frac{1}{2}\rho_j^*\xi_j(W,\alpha_j^W,C)+\frac{1}{2}\left\{\rho_i^*\frac{P_j^C+P_j^D}{2}+(1-\rho_i^*)\frac{2-P_j^C-P_j^D}{2}\right\}q_j\right] \\ &+\frac{P_j^C-P_j^D}{4}\Lambda_j, \\ \hat{\beta}^D =&\beta\left[1-\frac{1}{2}\rho_j^*\xi_j(S,\alpha_j^S,D)-\frac{1}{2}\left\{\rho_i^*\frac{2-P_j^C-P_j^D}{2}+(1-\rho_i^*)\frac{P_j^C+P_j^D}{2}\right\}q_j\right] \\ &+(1-\beta)\left[\frac{1}{2}(1-\rho_j^*)\xi_j(W,\alpha_j^W,C)+\frac{1}{2}\left\{\rho_i^*\frac{P_j^C+P_j^D}{2}+(1-\rho_i^*)\frac{2-P_j^C-P_j^D}{2}\right\}q_j\right] \\ &+\frac{P_j^C-P_j^D}{4}\Lambda_j. \end{aligned}$$

Let us substitute (15)–(17) into $\hat{\beta}^C$ and $\hat{\beta}^D$. Then we can solve (19) and (20), and the solutions are:

$$V_i^{\beta,C} = \beta V_i^S+(1-\beta)V_i^W, \tag{21}$$

$$V_i^{\beta,D} = \beta V_i^S+(1-\beta)V_i^W. \tag{22}$$

(21) and (22) imply that at any Cooperation state with belief $\beta$ about the other player's state, player $i$'s continuation payoff is independent of his action, as long as he will not deviate in any Examination state and in any observational activity. This

proves that, given the other player's strategy, player $i$ does not have an incentive to deviate in terms of action in any Cooperation state. The argument also proves that any player $i$'s continuation payoff given a Cooperation state depends solely on his belief about the other player's state, and has the form of

$$V_i = \beta V_i^S + (1-\beta) V_i^W. \tag{23}$$

This implies that the payoff pair of the whole strategy profile is $(v_1, v_2)$ by (18).

We have seen that each player $i$ does not have an incentive to deviate in action in any Cooperation state. Also player $i$ has no incentive to choose a different action in an Examination state with $|e| = j$, because he is prescribed to play a stage game dominant action $D$, and his action will not affect future play.

We next consider player $i$'s incentive in an Examination state with $|e| = i$. Let $\gamma^{Z,a_j}$ be the current belief with which player $i$ believes that player $j$ was in $Z$-Cooperation state and played $a_j$ in the previous period. We also define $\gamma^{a_j} = \gamma^{S,a_j} + \gamma^{W,a_j}$.

Suppose $e = i$. Let $\hat{\gamma}_{a_i}$ be the probability with which player $i$ believes that player $j$ will be in the $S$-Cooperation state in the next period, when player $i$ chooses $a_i$ in the current period. Then we have:

$$\begin{aligned} \hat{\gamma}_C =& \gamma^{S,C}\{1-(1-P_j^C)q_j^C\} + \gamma^{S,D}(1-P_j^C q_j^D) \\ &+ \gamma^{W,C} P_j^C q_j^C + \gamma^{W,D}(1-P_j^C) q_j^D, \end{aligned} \tag{24}$$

$$\begin{aligned} \hat{\gamma}_D =& \gamma^{S,C}(1-P_j^D q_j^C) + \gamma^{S,D}\{1-(1-P_j^D)q_j^D\} \\ &+ \gamma^{W,C}(1-P_j^D) q_j^C + \gamma^{W,D} P_j^D q_j^D. \end{aligned} \tag{25}$$

By (14)–(16) and (23), the necessary and sufficient condition that $C$ is optimal in the current period is:

$$-(1-\delta)l_i + \delta\big(\hat{\gamma}_C - \hat{\gamma}_D\big)\big(V_i^S - V_i^W\big) = \frac{1-\delta}{\delta} 2\kappa_i^*(2\gamma^C - 1) \geq 0. \tag{26}$$

Hence $C$ is optimal if and only if $\gamma^C \geq 1/2$.

Next, suppose $e = -i$. Again, let $\hat{\gamma}_{a_i}$ be the probability with which player $i$ believes that player $j$ will be in the $S$-Cooperation state in the next period, when player $i$ chooses $a_i$ in the current period. This time we have:

$$\begin{aligned} \hat{\gamma}_C =& \gamma^{S,C}(1-P_j^C q_j^D) + \gamma^{S,D}\{1-(1-P_j^C)q_j^C\} \\ &+ \gamma^{W,C}(1-P_j^C) q_j^D + \gamma^{W,D} P_j^C q_j^C, \end{aligned} \tag{27}$$

$$\begin{aligned} \hat{\gamma}_D =& \gamma^{S,C}\{1-(1-P_j^D)q_j^D\} + \gamma^{S,D}(1-P_j^D q_j^C) \\ &+ \gamma^{W,C} P_j^D q_j^D + \gamma^{W,D}(1-P_j^D) q_j^C. \end{aligned} \tag{28}$$

By (14)–(16) and (23), the necessary and sufficient condition that $C$ is optimal in the current period is:

$$-(1-\delta)l_i+\delta\big(\hat{\gamma}_C-\hat{\gamma}_D\big)\big(V_i^S-V_i^W\big)=\frac{1-\delta}{\delta}2\kappa_i^*(1-2\gamma^C)\geq 0. \tag{29}$$

Hence $C$ is optimal if and only if $\gamma^C\leq 1/2$.

Now we consider each player $i$'s incentive to choose $\rho_i^*$ and then play the subsequent examination period according to (12), if prescribed in a Cooperation state. Suppose that player $i$ is at some Cooperation state and has selected some action, and monitoring is prescribed (namely, the sunspot in the middle of that period is less than $\mu$). As before, let $\gamma^{Z,a_j}$ be the probability with which player $i$ believes that player $j$ is in $Z$-Cooperation state and played $a_j$ in the current period. We also let $\gamma^{a_j}=\gamma^{S,a_j}+\gamma^{W,a_j}$. Suppose player $i$ chooses $m_i\in M_i$, and let $\hat{\gamma}_{a_i}^{Z,a_j}$ be the posterior probability with which player $i$ believes that player $j$ is in $Z$-Cooperation state and played $a_j$ in the current period, after observing $a_i$. Then we have:

$$\hat{\gamma}_C^{Z,C}=\frac{\gamma^{Z,C}m_i}{\gamma^C m_i+(1-\gamma^C)(1-m_i)},\quad \hat{\gamma}_C^{Z,D}=\frac{\gamma^{Z,D}(1-m_i)}{\gamma^C m_i+(1-\gamma^C)(1-m_i)}, \tag{30}$$

$$\hat{\gamma}_D^{Z,C}=\frac{\gamma^{Z,C}(1-m_i)}{\gamma^C(1-m_i)+(1-\gamma^C)m_i},\quad \hat{\gamma}_D^{Z,D}=\frac{\gamma^{Z,D}m_i}{\gamma^C(1-m_i)+(1-\gamma^C)m_i}. \tag{31}$$

We also define $\hat{\gamma}_{a_i}^{a_j}=\hat{\gamma}_{a_i}^{S,a_j}+\hat{\gamma}_{a_i}^{W,a_j}$.

We consider a continuation strategy where player $i$ chooses $m_i$ and then follows $\hat{\sigma}_i$ in all subsequent periods, and examine how $m_i$ affects the continuation payoff. First, it costs $(1-\delta)c_i(m_i)$ in the current period to choose $m_i$. The monitoring decision does not affect future payoffs if the next examination period has $e\in\{j,-j\}$. If $e\in\{i,-i\}$, however, $m_i$ affects the probability with which player $j$ moves to $S$-Cooperation state in the subsequent cooperation period. With probability $\{\gamma^C m_i+(1-\gamma^C)(1-m_i)\}/4$, player $i$ observes $C$ and $e=i$. Thus $\hat{\sigma}_i$ prescribes $C$ in the examination period. Hence (24) applies, and player $i$ believes that player $j$ will move to the $S$-Cooperation state in the subsequent cooperation period with the probability

$$\hat{\gamma}_C^{S,C}\{1-(1-P_j^C)q_j^C\}+\hat{\gamma}_C^{S,D}(1-P_j^C q_j^D)+\hat{\gamma}_C^{W,C}P_j^C q_j^C+\hat{\gamma}_C^{W,D}(1-P_j^C)q_j^D.$$

With probability $\{\gamma^C m_i+(1-\gamma^C)(1-m_i)\}/4$, player $i$ observes $C$ and $e=-i$. Thus $\hat{\sigma}_i$ prescribes $D$ in the examination period. Hence (28) applies, and player $i$ believes that player $j$ will move to the $S$-Cooperation state in the subsequent cooperation period with the probability

$$\hat{\gamma}_C^{S,C}\{1-(1-P_j^D)q_j^D\}+\hat{\gamma}_C^{S,D}(1-P_j^D q_j^C)+\hat{\gamma}_C^{W,C}P_j^D q_j^D+\hat{\gamma}_C^{W,D}(1-P_j^D)q_j^C.$$

Similarly, player $i$ observes $D$ and has either $e=i$ or $e=-i$ with probability $\{\gamma^C(1-m_i)+(1-\gamma^C)m_i\}/4$, respectively. In either case, (25) or (27) applies. Therefore player $i$ believes that player $j$ will move to the $S$-Cooperation state in the subsequent cooperation period with the probability

$$\hat{\gamma}_D^{S,C}(1-P_j^D q_j^C)+\hat{\gamma}_D^{S,D}\{1-(1-P_j^D)q_j^D\}+\hat{\gamma}_D^{W,C}(1-P_j^D)q_j^C+\hat{\gamma}_D^{W,D}P_j^D q_j^D$$

or

$$\hat{\gamma}_D^{S,C}(1-P_j^C q_j^D)+\hat{\gamma}_D^{S,D}\{1-(1-P_j^C)q_j^C\}+\hat{\gamma}_D^{W,C}(1-P_j^C)q_j^D+\hat{\gamma}_D^{W,D}P_j^C q_j^C.$$

To sum up, by (30)–(31), the probability that the next Examination state has $e \in \{i, -i\}$ and player $j$ moves to $S$-Cooperation state in the subsequent cooperation period is

$$\frac{1}{2}q_j(P_j^C+P_j^D-1)m_i+\frac{1}{4}\big[2(\gamma^{W,C}+\gamma^{W,D})-(P_j^C+P_j^D)\big]q_j+\frac{1}{4}(P_j^C-P_j^D)\Lambda_j.$$

Using (23) and considering the term depending on $m_i$ only, we conclude that choosing $m_i$ has an effect on the continuation payoff by the amount

$$\delta^2\frac{1}{2}q_j(P_j^C+P_j^D-1)(V_i^S-V_i^W)m_i=(1-\delta)\kappa_i^* m_i,$$

where the equality follows from (15). Thus the continuation payoff of choosing $m_i$ and then conforming to $\hat{\sigma}_i$ is

$$(1-\delta)\big\{-c_i(m_i)+\kappa_i^* m_i\big\}$$

plus a constant, which is maximized at $m_i=\rho_i^*$ by (7).

Since $\hat{\eta} \le \gamma^C \le 1-\hat{\eta}$ for any belief $\gamma^{Z,a_j}$'s, it follows from (30) and (31) that $\hat{\gamma}_C^C \ge 1/2$ and $\hat{\gamma}_D^C \le 1/2$ for any $m_i \ge 1-\hat{\eta}$. Hence (26) and (29) imply that once player $i$ chooses $m_i \ge 1-\hat{\eta}$, then it is optimal to follow $\hat{\sigma}_i$ in all subsequent periods. By (4), this implies that choosing $\rho^*$ and then conforming to $\hat{\sigma}_i$ is optimal among all continuation strategies where player $i$ chooses $m_i \ge 1-\hat{\eta}$.

It remains to consider a continuation strategy where player $i$ chooses $m_i < 1-\hat{\eta}$. For such a strategy, it is possible that either $\hat{\gamma}_C^C < 1/2$ and $\hat{\gamma}_D^C > 1/2$ holds (however, both cannot hold simultaneously, because we always have $\hat{\gamma}_C^C \ge \hat{\gamma}_D^C$). If that happens, it is not optimal to follow $\hat{\sigma}_i$ in the next examination period.

Thus we have three cases to consider. First, consider $m_i$ such that both $\hat{\gamma}_C^C \ge 1/2$ and $\hat{\gamma}_D^C \le 1/2$ hold. Then it is optimal to follow $\hat{\sigma}_i$ in the subsequent examination period. Therefore, the above argument applies, and $m_i$ is inferior to $\rho_i^*$ by (7).

Second, consider $m_i$ such that $\hat{\gamma}_C^C < 1/2$. Then conforming to $\hat{\sigma}_i$ is not optimal if player $i$ observed $C$. With probability $\gamma^C m_i+(1-\gamma^C)(1-m_i)$, player $i$ observes $C$. Then the play reaches to an Examination state with $e \in \{i,-i\}$ with probability 1/2. If $e=i$, then his optimal action is $D$, while $\hat{\sigma}_i$ assigns $C$. The gain from optimally choosing $D$ is the minus of the value in (26), where $\gamma^C$ is replaced with $\hat{\gamma}_C^C$. If $e=-i$, then his optimal action is $C$, while $\hat{\sigma}_i$ assigns $D$. The gain from optimally choosing $C$ is the value in (29), where $\gamma^C$ is replaced with $\hat{\gamma}^C_C$. Thus the additional gain from optimally deviating from $\hat{\sigma}_i$ in the Examination state is:

$$\delta\frac{1}{2}\{\gamma^C m_i + (1-\gamma^C)(1-m_i)\}\frac{1-\delta}{\delta}2\kappa_i^*(1-2\hat{\gamma}^C_C) = (1-\delta)(1-\gamma^C-m_i)\kappa_i^*,$$

where the equality is due to (30). However, this additional gain is smaller than the difference in the continuation payoffs between $\rho_i^*$ and $m_i$, if player $i$ subsequently follows $\hat{\sigma}_i$. Indeed, by $\gamma^C \geq \hat{\eta}$ and (6), we obtain

$$(1-\delta)(1-\gamma^C-m_i)\kappa_i^* \leq (1-\delta)(1-\hat{\eta}-m_i)\kappa_i^* < (1-\delta)\big[\kappa_i^*\rho_i^* - c_i(\rho_i^*) - \kappa_i^* m_i + c_i(m_i)\big].$$

Hence it is not optimal to choose $m_i$.

Finally, consider $m_i$ such that $\hat{\gamma}^C_D > 1/2$. Then conforming to $\hat{\sigma}_i$ is not optimal if player $i$ observed $D$. With probability $\gamma^C(1-m_i)+(1-\gamma^C)m_i$, player $i$ observes $D$. Then the play reaches to an Examination state with $e \in \{i, -i\}$ with probability $1/2$. If $e = i$, then his optimal action is $C$, while $\hat{\sigma}_i$ assigns $D$. The gain from optimally choosing $C$ is the value in (26), where $\gamma^C$ is replaced with $\hat{\gamma}^C_D$. If $e = -i$, then his optimal action is $D$, while $\hat{\sigma}_i$ assigns $C$. The gain from optimally choosing $C$ is the minus of the value in (29), where $\gamma^C$ is replaced with $\hat{\gamma}^C_D$. Thus the additional gain from optimally deviating from $\hat{\sigma}_i$ in the Examination state is:

$$\delta\frac{1}{2}\{\gamma^C(1-m_i) + (1-\gamma^C)m_i\}\frac{1-\delta}{\delta}2\kappa_i^*(2\hat{\gamma}^C_D-1) = (1-\delta)(\gamma^C-m_i)\kappa_i^*,$$

where the equality is due to (31). Again by $\gamma^C \leq 1-\hat{\eta}$ and (6), we obtain

$$(1-\delta)(\gamma^C-m_i)\kappa_i^* \leq (1-\delta)(1-\hat{\eta}-m_i)\kappa_i^* < (1-\delta)\big[\kappa_i^*\rho_i^* - c_i(\rho_i^*) - \kappa_i^* m_i + c_i(m_i)\big].$$

For the same reason as above, it is not optimal to choose $m_i$. Hence player $i$ has no incentive to choose $m_i < 1-\hat{\eta}$, which completes the proof that $\hat{\sigma}$ is a Nash equilibrium.

$\hat{\sigma}$ is not a sequential equilibrium because the action in the Examination state after selecting $m_i < 1-\hat{\eta}$ in the previous period may not be optimal. Therefore, in order to obtain an outcome-equivalent sequential equilibrium, we modify each $\hat{\sigma}_i$ to the following new strategy $\sigma_i^*$. $\sigma_i^*$ coincides with $\hat{\sigma}_i$ at all histories except at the Examination states. If player $i$ is at an Examination state, then let $\gamma^C$ be the probability with which player $i$ believes that player $j$ played $C$ in the previous cooperation state. Then $\sigma_i^*$ assigns $C$ if and only if (i) $e = i$ and $\gamma^C \geq 1/2$, or (ii) $e = -i$ and $\gamma^C \leq 1/2$. By (26) and (29), this behavior is sequentially rational. Since $\sigma_i^*$ coincides with $\hat{\sigma}_i$ at all histories on the path, $\sigma^* = (\sigma_1^*, \sigma_2^*)$ has the same outcome as $\hat{\sigma}$, and satisfies sequential rationality. Q.E.D.

The equilibrium construction in the proof of Proposition 1 reveals that we have a family of *Nash* equilibria with simple structure, each of which sustains a payoff pair in $[\varepsilon, 1-\varepsilon] \times [\varepsilon, 1-\varepsilon]$ for a fixed $\varepsilon > 0$, if the players are sufficiently patient.

**Proposition 2** *Fix $\varepsilon > 0$. Then there exists $\underline{\delta} \in (0,1)$ such that for any $\delta \geq \underline{\delta}$, any $i$ and any $v_j \in [\varepsilon, 1-\varepsilon]$, there exists a strategy of player $i$, $\hat{\sigma}_i(\delta, v_j)$, with the following property: for any $\delta$ and any $(v_1', v_2') \in [\varepsilon, 1-\varepsilon] \times [\varepsilon, 1-\varepsilon]$, the strategy profile $(\hat{\sigma}_1(\delta, v_2'), \hat{\sigma}_2(\delta, v_1'))$ is a Nash equilibrium of $G(\delta)$, with a payoff pair $(v_1', v_2')$.*

**Proof.** Fix $\varepsilon > 0$, and fix $\underline{\delta}$ for which Proposition 1 holds. Recall that in the proof of Proposition 1, the Nash equilibrium $\hat{\sigma}$ sustaining some $(v_1, v_2) \in [\varepsilon, 1-\varepsilon] \times [\varepsilon, 1-\varepsilon]$ is such that each $\hat{\sigma}_i$ depends only on $v_j$. That is, each $v_i$ affects the construction only through $\lambda_i$ in (18), and $\lambda_i$ affects only $\hat{\sigma}_j$. Thus for $\delta \geq \underline{\delta}$ and $v_j \in [\varepsilon, 1-\varepsilon]$, if we set $\hat{\sigma}_i(\delta, v_j)$ as a Nash equilibrium strategy sustaining some $(v_i', v_j)$, we are done. Q.E.D.

We point out that this interchangeability result is only for Nash equilibrium, and does not extend to sequential equilibrium. The interchangeability requires that the equilibrium property is independent of what equilibrium the other player plays. This is true in our construction at all histories on the path. However, if we consider off-the-path histories at an Examination state where the player deviated in monitoring investment in the previous period, his sequentially rational behavior depends on his belief about the other player's action in the period, and the belief depends on the initial play.

Proposition 1 implies that two vertices of the set of feasible and individually rational payoff pairs, $(1,1)$ and $(0,0)$, can be approximated as an equilibrium. Therefore, the proof of the folk theorem is complete if the remaining two vertices can also be approximated as an equilibrium if players are sufficiently patient. It suffices to consider the approximation of $\left(\frac{g_1+l_2+1}{l_2+1}, 0\right)$, because the same line of argument also works for $\left(0, \frac{g_2+l_1+1}{l_1+1}\right)$.

**Proposition 3** *For any $\varepsilon > 0$, there exists $\underline{\delta} \in (0,1)$ such that any $G(\delta)$ with $\delta \geq \underline{\delta}$ has a sequential equilibrium whose payoff for player 1, $v_1$, satisfies*

$$v_1 > \frac{g_1 + l_2 + 1}{l_2 + 1} - \varepsilon. \tag{32}$$

**Proof.** Fix $\varepsilon > 0$. Note first that

$$\frac{1}{l_2+1} u_1(D,C) + \frac{l_2}{l_2+1} u_1(C,C) = \frac{g_1 + l_2 + 1}{l_2 + 1}.$$

Therefore there exists $r < 1/(l_2 + 1)$ such that

$$r u_1(D,C) + (1-r) u_1(C,C) > \frac{g_1 + l_2 + 1}{l_2 + 1} - \frac{\varepsilon}{2}. \tag{33}$$

Next, we choose $\nu > 0$ so that

$$u_1(D, (1-\nu)C + \nu D) > 1 + g_1 - \frac{\varepsilon}{4}, \tag{34}$$

$$\frac{(1-\nu)\underline{p}}{(1-\nu)\underline{p} + \nu(1-\underline{p})} > \frac{1}{2}. \tag{35}$$

We also define

$$\hat{\nu} \equiv \frac{\nu\underline{p}}{\nu\underline{p} + (1-\nu)(1-\underline{p})}.$$

By Assumption 2, there exist $\hat{\kappa}_1 > l_1$, $\hat{\rho}_1 \in M_1$ with $\hat{\rho}_1 > 1 - \hat{\nu}$, and $\Delta < 1$ such that

$$\hat{\kappa}_1\hat{\rho}_1 - c_1(\hat{\rho}_1) > \hat{\kappa}_1(1-\hat{\nu}), \tag{36}$$

$$\hat{\kappa}_1\hat{\rho}_1 - c_1(\hat{\rho}_1) \geq \hat{\kappa}_1 m_1 - c_1(m_1) \quad \forall m_1 \in M_1 \tag{37}$$

$$\frac{rl_2}{1-r}\frac{1}{(2\hat{\rho}_1 - 1)\Delta} < 1, \tag{38}$$

$$\frac{1+\Delta}{2} > 1 - \frac{\varepsilon}{2}. \tag{39}$$

Note that (38) can be satisfied because $r < 1/(l_2+1)$ implies $rl_2/(1-r) < 1$.

We then choose $\underline{\delta} \in (0,1)$ such that (i) it is greater than $\underline{\delta}$ appearing in Proposition 2 if we set $\varepsilon = (1-\Delta)/2$, and (ii) the following inequalities are all satisfied if $\delta \geq \underline{\delta}$.

$$\hat{\xi}_1 \equiv \frac{rl_2}{1-r}\frac{1}{\delta(2\hat{\rho}_1 - 1)\Delta} < 1, \tag{40}$$

$$\frac{(1-\delta)(\hat{\kappa}_1 + \delta l_1)}{\delta^2\left(P_2^C + P_2^D - 1\right)\Delta} < 1, \tag{41}$$

$$\frac{1-r}{r}\frac{1-\delta}{\delta}c_1(\hat{\rho}_1) < \frac{\varepsilon}{4}, \tag{42}$$

$$\begin{aligned} E_1 \equiv & -(1-\delta)\frac{l_1}{2} + \delta\frac{1+\Delta}{2} + \frac{1-\delta}{\delta}\hat{\kappa}_1\hat{\rho}_1 \\ & -(1-\delta)\frac{\left(P_2^C + P_2^D\right)\hat{\kappa}_1 + \delta\left(P_2^D - P_2^C\right)l_1}{2\delta\left(P_2^C + P_2^D - 1\right)} > 1 - \frac{\varepsilon}{2}. \end{aligned} \tag{43}$$

Note that (40) can be satisfied by (38), and (43) can be satisfied by (39).

Now we are ready to define the following strategy profile, which is described by the following two states, *initial states* and *unilateral examination states*.

*Initial States (I).* The play starts with this state. In this state, the players play a mixed action profile $(D, (1-\nu)C + \nu D)$. As for observational decisions, depending on the sunspot in the middle of the period, they choose $(m_1, m_2) = (1/2, 1/2)$ with probability $1 - \hat{\mu}$, where

$$\hat{\mu} = \frac{1-r}{r}\frac{1-\delta}{\delta}.$$

In this case, the next period is again in the initial state. With probability $\hat{\mu}$, player 1 chooses $\hat{\rho}_1$, while player 2 chooses 1/2. In this case, the play in the next period moves to the unilateral examination state, where only player 1 will take a test.

*Unilateral Examination States ($E_1$).* This state always follows an initial state where player 1 is prescribed to choose $\hat{\rho}_1$ as a monitoring investment. In this state, the sunspot at the beginning of the period selects a number $e \in \{-1, 1\}$ with equal probabilities. Let $a_2^{obs}$ be player 1's observation in the previous period. Then he is prescribed to play

$$\hat{b}_1 = \begin{cases} C & \text{if } (a_2^{obs}, e) \in \{(C, i), (D, -i)\}, \\ D & \text{otherwise.} \end{cases} \tag{44}$$

Player 2 is prescribed to play $D$. In this state, each player always chooses $m_i = 1/2$.

After the unilateral examination state, the play goes back to none of the two states $I$ or $E_1$, and players' continuation strategies are determined as follows. As for player 1, his continuation strategy depends entirely on $a_2^{obs}$ in the previous initial state. If $a_2^{obs} = C$, then his continuation strategy is $\hat{\sigma}_1(\delta, (1 + \Delta)/2)$, whose definition is in Proposition 2. If $a_2^{obs} = D$, then his continuation strategy is $\hat{\sigma}_1(\delta, (1 + \Delta)/2)$ with probability $1 - \hat{\xi}_1$, and $\hat{\sigma}_1(\delta, (1 - \Delta)/2)$ with probability $\hat{\xi}_1$, where $\hat{\xi}_1$ is defined by (40). Player 2's continuation strategy depends on his own action in the previous initial period $a_2^{own}$, the sunspot's selection $e$ and his private signal in the current period $\omega_2$. In any case, player 2 randomizes over $\hat{\sigma}_2(\delta, (1 + \Delta)/2)$ and $\hat{\sigma}_2(\delta, (1 - \Delta)/2)$. The probability that $\hat{\sigma}_2(\delta, (1 - \Delta)/2)$ is selected is

$$\begin{cases} \hat{q}_2^C & \text{if } (a_2^{own}, e) \in \{(C, 1), (D, -1)\} \text{ and } \omega_2 \in \Omega_2^D \\ \hat{q}_2^D & \text{if } (a_2^{own}, e) \in \{(C, -1), (D, 1)\} \text{ and } \omega_2 \in \Omega_2^C, \\ 0 & \text{otherwise,} \end{cases}$$

where

$$\hat{q}_2^C = \frac{(1 - \delta)(\hat{\kappa}_1 + \delta l_1)}{\delta^2 \big(P_2^C + P_2^D - 1\big)\Delta},$$
$$\hat{q}_2^D = \frac{(1 - \delta)(\hat{\kappa}_1 - \delta l_1)}{\delta^2 \big(P_2^C + P_2^D - 1\big)\Delta}.$$

Note that by (41) and $\hat{\kappa}_1 > l_1$, both $\hat{q}_2^C$ and $\hat{q}_2^D$ are a probability.

We show that this strategy profile is a Nash equilibrium of $G(\delta)$. First, note that if the current period is in $E_1$, then each player $i$'s continuation strategy from the next period on is either $\hat{\sigma}_i(\delta, (1 + \Delta)/2)$ or $\hat{\sigma}_i(\delta, (1 - \Delta)/2)$. By Proposition 2, the players have no incentive to deviate at any history on the path after a unilateral examination state.

Next we consider optimality of behavior at a unilateral examination state. Clearly player 2 is willing to conform to the strategy, because his action has no influence on the future play. Player 1's optimal action depends on his belief about the other player's action in the previous period. Let $\gamma^C$ be the probability with which player 1 believes that player 2 played $C$ in the previous period. Then the same line of argument showing (26) and (29) proves that under $e = 1$, it is optimal to play $C$ if and only if $\gamma^C \geq 1/2$, and under $e = -1$ it is optimal to play $C$ if and only if $\gamma^C \leq 1/2$. Since $\hat{\rho}_1 > 1 - \hat{\nu}$, it follows that at any history on the path in a unilateral examination period, $\gamma^C$ after observing $C$ in the previous period is no less than 1/2, and $\gamma^C$ after observing $D$ in the previous period is no greater than 1/2. Thus the prescribed action in any unilateral examination state is optimal on the path.

Finally we consider incentives in initial states. As for player 2, if he plays $D$ instead of $C$, his payoff increases by $(1-\delta)l_2$. It also increases the probability that player 1's observation is $D$ by $2\hat{\rho}_1 - 1$, if player 1 is prescribed to choose $\hat{\rho}_1$, which occurs with probability $\hat{\mu}$. If player 1 observes $D$ instead of $C$, the probability he chooses $\hat{\sigma}_1(\delta, (1-\Delta)/2)$ instead of $\hat{\sigma}_1(\delta, (1+\Delta)/2)$ after the subsequent unilateral examination period increases by $\hat{\xi}_1$. Hence player 2's future loss from choosing $D$ instead of $C$ is

$$\hat{\mu}(2\hat{\rho}_1 - 1)\hat{\xi}_1\delta^2\Delta,$$

which is equal to $(1-\delta)l_2$ by the definitions of $\hat{\xi}_1$ and $\hat{\mu}$. Therefore, player 2 is indifferent between $C$ and $D$, and is willing to randomize as specified.

Player 1's action in the initial state is optimal, because it does not change future play. Next, suppose that player 1 is prescribed to choose $\hat{\rho}_1$. Suppose player 1 chooses $m_1 \in M_1$ and then plays according to this strategy. By a similar argument to the one in the proof of Proposition 1, player 1's payoff of this continuation strategy is

$$(1-\delta)\big\{-c_1(m_1) + \hat{\kappa}_1 m_1\big\}$$

plus a constant, which is maximized at $m_1 = \hat{\kappa}_1$ by (37).

Playing according to the strategy after choosing $m_1$ is optimal if $m_1 \geq 1 - \hat{\nu}$. If player 1 chooses $m_1 < 1 - \hat{\nu}$, then it is possible that $\gamma^C$ after observing $D$ is greater than 1/2 (by (35), $\gamma^C$ after observing $C$ is always greater than 1/2). If that is the case, then the optimal action in the unilateral examination period is $C$ if $e = 1$ and $D$ if $e = -1$. In either case, the additional gain from playing optimally is $(1-\delta)(\gamma^C - m_1)\hat{\kappa}_1$. By $\gamma^C \leq 1 - \hat{\nu}$ and (36), we obtain

$$(1-\delta)(\gamma^C - m_1)\hat{\kappa}_1 \leq (1-\delta)(1-\hat{\nu}-m_1)\hat{\kappa}_1 < (1-\delta)\big[\hat{\kappa}_1\hat{\rho}_1 - c_1(\hat{\rho}_1) - \hat{\kappa}_1 m_1 + c_1(m_1)\big].$$

Hence it is optimal to choose $\hat{\rho}_1$, and the strategy profile is a Nash equilibrium.

We compute player 1's equilibrium payoff $v_1$. Note that if the play moves to the $E_1$ state, player 1's observation in the previous period is correct with probability $\hat{\rho}_1$. Hence by the definition of $\hat{q}_2^C$ and $\hat{q}_2^D$, the probability that player 2's continuation strategy from the next period on is $\hat{\sigma}_2(\delta, (1-\Delta)/2)$ is

$$\begin{aligned}&\frac{1}{2}\hat{\rho}_1\Big\{\big(1-P_2^C\big)\hat{q}_2^C+\big(1-P_2^D\big)\hat{q}_2^D\Big\}+\frac{1}{2}\big(1-\hat{\rho}_1\big)\big(P_2^C\hat{q}_2^D+P_2^D\hat{q}_2^C\big)\\&=-\frac{1-\delta}{\delta^2\Delta}\hat{\kappa}_1\hat{\rho}_1+(1-\delta)\frac{\big(P_2^C+P_2^D\big)\hat{\kappa}_1+\delta\big(P_2^D-P_2^C\big)l_1}{2\delta^2\big(P_2^C+P_2^D-1\big)\Delta}.\end{aligned}$$

Using this, we can show that the expected continuation payoff at $E_1$ state is equal to $E_1$ defined by (43). Consequently, it follows that

$$v_1=(1-\delta)\big\{u_1(D,(1-\nu)C+\nu D)-\hat{\mu}c_1(\hat{\rho}_1)\big\}+\delta(1-\hat{\mu})v_1+\delta\hat{\mu}E_1.$$

Substituting the definition of $\hat{\mu}$ and rearranging, we obtain

$$v_1=r\big\{u_1(D,(1-\nu)C+\nu D)-\hat{\mu}c_1(\hat{\rho}_1)\big\}+(1-r_1)E_1.$$

By (33), (34), (42) and (43), we have

$$v_1>\frac{g_1+l_2+1}{l_2+1}-\varepsilon.$$

Consequently, we have shown (32).

So far we have proved existence of a Nash equilibrium satisfying (32). The Nash equilibrium may not satisfy sequential rationality at histories in unilateral examination states, where player 1 deviated in terms of observational decision. However, the same argument as that in Proposition 1 demonstrates that there exists an outcome-equivalent sequential equilibrium. This completes the proof. Q.E.D.

We briefly explain basic ideas of our construction. For simplicity, we consider only equilibria approximating the efficient payoff pair $(1, 1)$.

In order to sustain efficient outcomes and provide incentives, the play must have both a cooperation phase and a punishment phase. Also players must reduce a probability that inefficient punishment occurs when they are actually cooperative. This requires that a shift from a cooperation phase to a punishment phase be based on a very precise information about the other player's action. That is why in our construction a player's action in a Cooperation state affects future payoffs only after mutual monitoring. Since the observation is very precise, inefficient punishment is avoided quite surely.

It remains to provide incentives to choose a large monitoring investment when prescribed to do so. This is done by a possibility that a player becomes an examinee in the subsequent examination state, and a wrong answer is punished by a greater probability that the other player is in the $W$-Cooperation state in the next period. A key fact is that if a player chooses a sufficiently large investment level, then the optimal action in the examination state as an examinee depends only on his observation. Namely, his private signal provides only coarse information about the other player's action, so it can never overturn what his observation suggests. Since the probability that his observation is wrong is linear in $m_i$, the effect on continuation payoffs when he chooses a large $m_i$ is also linear in $m_i$. If the players are sufficiently patient, the effect on the continuation payoff can be made arbitrarily large. As a result, by Assumption 2(iii), we can design future behavior so that it is optimal to choose a very large monitoring investment level if the sunspot prescribes monitoring.

Finally, we point out that the equilibrium strategy can be described by an automaton. For player $i$'s strategy, the automaton has two cooperation states and *eight* examination states, where each examination state is characterized by $(Z, a_i^{own}, a_j^{obs}) \in \{S, W\} \times A_i \times A_j$. If player $i$ is in $Z$-Cooperation state and then is prescribed to choose $\rho_i^*$, then in the next period he is in $(Z, a_i^{own}, a_j^{obs})$-Examination state, where $(a_i^{own}, a_j^{obs})$ is the pair of his action and his observation in that period.

# 4 Discussions

One contribution of our folk theorem is to extend existing folk theorems under costly perfect monitoring settings to the case with costly almost perfect monitoring. The theorem also has some implications on existing folk theorems in repeated games with imperfect private monitoring, where monitoring is given exogenously.

For understanding this implication, it is helpful to consider a variant of our model, where players can select monitoring investments *at no cost*. That is, while we maintain Assumption 2(i)(ii), we replace (iii) with $c_i(m_i) = 0$ for any $i$ and $m_i \in M_i$. In this alternative framework, there is no incentive problem on acquisition of as precise information as possible. Indeed, if perfect monitoring is available (namely, if $1 \in M_i$ for each $i$), then any equilibrium outcome in a repeated game with exogenously given perfect monitoring is sustained in this framework, too. Hence we would obtain a folk theorem if the players are sufficiently patient. What happens if $1 \notin M_i$? Then we can prove an analog of Proposition 1, by a direct application of the belief-free approach.

**Proposition 4** *Fix $\varepsilon > 0$. Then there exists $\underline{\delta} \in (0, 1)$ such that any payoff pair $(v_1, v_2) \in [\varepsilon, 1-\varepsilon] \times [\varepsilon, 1-\varepsilon]$ is a sequential equilibrium payoff pair of $G(\delta)$ with any $\delta \geq \underline{\delta}$. Moreover, the equilibrium strategy can be described by a two-state automaton.*

**Proof.** Fix $\varepsilon \in (0, 1/2)$. For each $i = 1, 2$, choose $\rho_i^* \in M_i$ so that

$$V_i^S \equiv 1 - \frac{1-\rho_j^*}{2\rho_j^* - 1} g_i \geq 1 - \varepsilon > \varepsilon \geq \frac{1-\rho_j^*}{2\rho_j^* - 1} l_i$$

for each $i$ and $j \neq i$. Also choose $\underline{\delta} \in (0,1)$ so that for any $\delta \geq \underline{\delta}$, both

$$q_i^S \equiv \frac{(1-\delta)g_j}{\delta(2\rho_i^* - 1)(V_j^S - V_j^W)} \in (0,1),$$
$$q_i^W \equiv \frac{(1-\delta)l_j}{\delta(2\rho_i^* - 1)(V_j^S - V_j^W)} \in (0,1)$$

hold.

Fix $(v_1, v_2) \in [\varepsilon, 1-\varepsilon] \times [\varepsilon, 1-\varepsilon]$. Now we define the following strategy of player $i$, $\sigma_i^*$. It consists of two states, which we again call Strong-Cooperation ($S$) and Weak-Cooperation ($W$) states. This time, player $i$'s action in each state is pure; player $i$ plays $C$ with probability 1 in state $S$, and plays $D$ with probability 1 in state $W$. Then he chooses $m_i = \rho_i^*$ in each state. The state transition rule is as follows. If player $i$ is in state $S$ and observed $C$, then he stays in the same state next period. If he observed $D$, then he moves to state $W$ with probability $q_i^S$, and stays in state $S$ with probability $1 - q_i^S$. If player $i$ is in state $W$ and observed $D$, then he stays in the same state next period. If he observed $C$, then he moves to state $S$ with probability $q_i^W$, and stays in state $W$ with probability $1 - q_i^W$. The initial state is stochastically selected, and player $i$ starts with state $S$ with probability $\lambda_i$ and starts with state $W$ with probability $1 - \lambda_i$, where $\lambda_i$ satisfies

$$\lambda_i V_j^S + (1-\lambda_i)V_j^W = v_j$$

for each $i$ and $j \neq i$.

The standard argument for belief-free strategies shows that (i) the payoff pair of the strategy profile $\sigma^* = (\sigma_1^*, \sigma_2^*)$ is $(v_1, v_2)$, and (ii) player $i$ is indifferent over all strategies if player $j$ plays $\sigma_j^*$. Indeed, player $i$ is indifferent over stage-actions because of the belief-freeness. He is also indifferent over monitoring investments, because all monitoring investment levels now have the same cost, and quality of his observation never affects quality of *the other player's observation*, which is the sole element determining $i$'s continuation payoff. Q.E.D.

Proposition 4 is concerned only with approximating the payoff pairs on the rectangle $[0,1] \times [0,1]$. We can also approximate other individually rational payoff pairs by strategies which also have an initial state (but not a unilateral examination state). The idea of the proof is similar to that of Proposition 3, we omit it.

Note that the line of argument in the proof of Proposition 4 is exactly the same as Ely and Valimaki's (2002) robust folk theorem. The only difference is that their folk theorem is an approximate result. Ely and Valimaki (2002) fix expected stage payoffs (namely, Figure 1) only, and then find a lower bound on discount factors and an upper bound on the noise of private monitoring structure, in order to sustain a given target payoff pair. Our argument therefore indicates that existing approximate folk theorems by belief-free strategies can be interpreted as a standard folk theorem in models with costless monitoring activities where almost perfect observations are available. From that perspective, our folk theorem can be regarded as an extension of those results to environments with costly information acquisition, thereby establishing a robustness of that interpretation.

There are some standard, non-approximate folk theorems in the literature on imperfect private monitoring. Matsushima (2004) and Yamamoto (2007) prove such folk theorems for prisoners' dilemma or its generalization, assuming that the players' private signals satisfy conditional independence or some similar property. Therefore our model can be regarded as another class of stage games for which the standard folk theorem obtains. Our model is similar to theirs in the sense that the observations are conditionally independent, and that assumption is crucial.[11] However, since free private signals need not be conditionally independent, our formulation cannot be interpreted as a special case or an extension of the models by Matsushima (2004) and Yamamoto (2007). In addition, our stage game is an extensive-form game, and we cannot replace it with its normal-form representation, because our construction depends on use of interim sunspots observed in the course of playing the extensive-form game.

# References


Ben-Porath, E., and M. Kahneman (2003): "Communication in Repeated Games with Costly Monitoring," *Games and Economic Behavior*, 44, 227–250.

Bhaskar, V., and I. Obara (2002): "Belief-based Equilibria in the Repeated Prisoners' Dilemma with Private Monitoring," *Journal of Economic Theory*, 102, 40–69.

Ely, J., J. Horner and W. Olszewski (2005): "Belief-Free Equilibria in Repeated Games," *Econometrica*, 73, 377–415.

Ely, J., and J. Valimaki (2002): "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory*, 102, 84–105.

Fong, K., O. Gossner, J. Horner and Y. Sannikov (2007): "Efficiency in a Repeated Prisoner's Dilemma with Imperfect Private Monitoring," mimeo.

Fudenberg, D., D. Levine, and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62, 997–1039.

Horner, J., and W. Olszewski (2006): "The Folk Theorem for Games with Private Almost-Perfect Monitoring," *Econometrica*, 74, 1499–1544.


[11] In fact, if the observations are correlated, our approach based on belief-freeness confronts difficulty, because then players may have an incentive to choose a different monitoring investment level in order to manipulate the noise in *the other player's observation.*

Kandori, M., I. Obara (2004): “Endogenous Monitoring,” mimeo.

Kandori, M., and I. Obara (2006): “Efficiency in Repeated Games Revisited: The Role of Private Strategies,” *Econometrica*, 74, 499–519.

Matsushima, H. (2004): “Repeated Games with Private Monitoring: Two Players,” *Econometrica*, 72, 823–852.

Miyagawa, E., Y. Miyahara, and T. Sekiguchi (2008): “The Folk Theorem for Repeated Games with Observation Costs,” *Journal of Economic Theory*, 139, 192–221.

Piccione, M. (2002): “The Repeated Prisoner’s Dilemma with Imperfect Private Monitoring,” *Journal of Economic Theory*, 102, 70–83.

Sekiguchi, T. (1997): “Efficiency in Repeated Prisoner’s Dilemma with Private Monitoring,” *Journal of Economic Theory*, 76, 345–361.

Yamamoto, Y. (2007): “Efficiency Results in N Player Games with Imperfect Private Monitoring,” *Journal of Economic Theory*, 135, 382–413.