A folk theorem in infinitely repeated prisoner's dilemma with small observation cost

Hino, Yoshifumi

Vietnam-Japan University

10 December 2018

A folk theorem in infinitely repeated prisoner's dilemma with small observation cost*

Yoshifumi Hino† [1]

[1]Business Administration Program, Vietnam–Japan University

This version: December 5, 2018

Abstract

We consider an infinitely repeated prisoner's dilemma under costly observation. If a player observes his opponent, then he pays an observation cost and knows the action chosen by his opponent. If a player does not observe his opponent, he cannot obtain any information about his opponent's action. Furthermore, no player can statistically identify the observational decision of his opponent. We prove an efficiency without any signals. Next, we consider a kind of delayed observations. Players decide their actions and observation decisions in the same period, but they choose observation decisions after they choose their actions. We introduce an interim public randomization instead of public randomization just before observation decision. We present a folk theorem with an interim public randomization device for a sufficiently small observation cost when players are sufficiently patient.

Keywords B to B business · Costly observation · Efficiency · Folk theorem · Prisoner's dilemma
JEL Classification: C72; C73; D82

1 Introduction

It is well known that prisoner's dilemma is a primitive model to represent the form of team production. In prisoner's dilemma, each player has two choices; exert a high effort for the team (cooperation) or do not exert a high effort

*This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

†Business Administration Program, Vietnam–Japan University, Luu Huu Phuoc Street, My Dinh 1 Ward, Nam Tu Liem District, Hanoi, 129160, Vietnam
Tel: +84 24 7306 6001
Email: y.hino@vju.ac.vn

for the team (noncooperation). To understand cooperative behavior in team production, the long-run relationship is also crucial.

One of the important factors in the long-run relationship is the monitoring structure. If a player wants to receive cooperation from the other player, he needs to monitor the other player and choose a cooperative action as a reward in the case where the other player cooperates with him. In reality, information is not free. We need time to collect information. We analyze what happens if monitoring is costly. More specifically, we consider whether efficiency or folk theorem holds or not.

In our model, we consider costly observation as a monitoring structure. Each player chooses his action and observational decision. If a player chooses to observe his opponent, then he can observe the action chosen by the opponent. The observational decision itself is unobservable. The player cannot obtain any information about his opponent in that period if he chooses not to observe that player. This means that the marginal distribution of private signals does not satisfy the full support condition.

Furthermore, no player can statistically identify the observational decision of his opponent. That is, our monitoring structure is neither almost-public private monitoring (Hörner and Olszewski (2009); Mailath and Morris (2002, 2006); Mailath and Olszewski (2011)) nor almost perfect private monitoring (Bhaskar and Obara (2002); Chen (2010); Ely and Välimäki (2002); Ely et al. (2005); Hörner and Olszewski (2006); Sekiguchi (1997); Piccione (2002); Yamamoto (2007, 2009))[1].

One of the application of this game is B to B business. Let us consider a price competition in in B to B business. In reality, the price of each products tends to be private. If the price of some company is public, the competitor will choose the price slightly lower than the competitor and obtain the client. If the company wants to know the price strategy of the competitor, the company needs to investigate with time and financial cost. This situation is costly observation.

We present two results. First, we show that a symmetric Pareto efficient payoff vector can be approximated by a sequential equilibrium without any signals under some assumptions regarding the payoff matrix when players are patient and the observation cost is small (efficiency). The second result is a type of folk theorem. We introduce an interim public randomization device. The public randomization device is realized just after the players choose their actions, and players can see the public randomization device before their observational decision. We present a folk theorem with an interim public randomization device under some assumptions regarding the payoff matrix when players are patient and the observation cost is small. The first result shows that a cartel is possible without any signal and communication in B to B business. The second result implies that companies need coordination device to archive asymmetric cartel

[1] Yamamoto (2012) shows some tractable subset of Nash equilibria under conditional independence and Sugaya (2011) modifies the equilibrium construction of Yamamoto (2012) and show the tractable subset of Nash equilibria without conditional independence. It is difficult to compare our result to their result because they assume the full support condition in their analysis, whereas costly monitoring does not satisfy the full support condition.

in B to B business.

The nature of our strategy is close to the *keep-them-guessing strategies* in Chen (2010). In our strategy, each player chooses C_i with certainty at the *cooperative* state, but randomizes the observational decision. Depending on the observation result, players change their actions from the next period on. If the player plays C_i and observes C_j, he remains in a cooperation state. However, in other cases (for example, the player does not observe his opponent), the player moves out of the cooperation state. From the perspective of his opponent, the player plays the game as if he randomizes C_i and D_i although he chooses pure actions in each state. Such randomized observations create uncertainty about the opponents' state in each period and gives the incentive to observe.

As with Chen (2010), our analysis is tractable. By construction, the concern of each player at each period is only whether his opponent is in a cooperation state or not. It is sufficient to keep track of this belief, which is the probability that the opponent is in a cooperation state.

Our main contribution is the efficiency result and folk theorem in infinitely repeated primitive prisoner's dilemma. Some previous literature shows that efficiency results hold if some tools to share information are available. For example, some literature assumes that communication is available. Another literature assumes that some information is available even if players do not observe the opponent. We will show these tools and discuss previous literature in Section 2. Our result shows that players can construct a cooperative relationship without any tools.

Another contribution is showing another approach to construct a sequential equilibrium. We consider the randomization of monitoring, whereas previous literature confines its attention to the randomization of actions. In costly monitoring model, the observational decision is supposed to be unobservable. Therefore, even if a player observes the opponent, he cannot know whether the opponent observes him or not. If the continuation strategy of the opponent depends on the observational decision in the previous period, the opponent randomizes actions from the perspective of the player although the opponent chooses pure actions in each history. This new approach enables us to construct a sequential equilibrium.

The rest of this paper is organized as follows. Section 3 introduces a model of repeated prisoner's dilemma with costly observation. We present our results in Section 5. We show an efficiency result with a small observation cost. We show a folk theorem with an interim public randomization device in Section 6. We will discuss asymmetric prisoner's dilemma in Section 7. Section 8 provides concluding remarks.

2 Literature Review

The previous literature shows efficiency results or folk theorems with some tools or assumptions. In this section, we explain these related literature on costly monitoring.

One of the biggest difficulty in costly monitoring is monitoring the monitoring activity of the opponent because the observational behaviors in the costly monitoring are assumed to be unobservable. Each player has to check these unobservable activity to motivate the other player to observe. One of the solution to this difficulty is assuming that the observational decision is observable. Kandori and Obara (2004); Lehrer and Solan (2018) assume that players can observe the other players' observational decision themselves.

Another approach is communication. Ben-Porath and Kahneman (2003) analyze an information acquisition model with communication. They show that players can share their information through explicit communication and present a folk theorem for any level of observation cost. Ben-Porath and Kahneman (2003) consider a strategy given which players randomize actions on the path. In their strategy, players report their observations each other. Then, each player can distinguish whether the other player observes him or not by the reports. Therefore, players can check observation activities of the other players.

An implicit communication has been shown in Miyagawa et al. (2008). Miyagawa et al. (2008) assume that communication is not allowed however players can obtain imperfect private signals about the other player's action even when they do not observe their opponent. They show that players can communicate with each other implicitly through the information and a folk theorem holds for any level of observation cost.

If these assumptions do not hold, that is, no costless information is available, then cooperation is difficult. There are two results which show folk theorems without costless information. Miyagawa et al. (2003) considers the same monitoring structure as used in this paper and presents a folk theorem with a small observation cost. Flesch and Perea (2009) also consider similar monitoring structures to our structure. In their model, players can obtain information about the other player if and only if they observe the other player. Furthermore, they assume that players can observe the actions chosen in the past if the players pay an additional cost. Flesch and Perea (2009) show a folk theorem for an arbitrary observation cost when each player can choose at least three actions. The above two studies consider an implicit communication using mixed actions. However, to use implicit communication by mixed action, the above two result needs more than two actions for each player. It means that their result does not hold in the infinitely repeated primitive prisoner's dilemma. We will discuss the implicit communications in Miyagawa et al. (2003); Flesch and Perea (2009) in Section 4 after we define our model in Section 3.

3 Model

The base game is a symmetric prisoner's dilemma. Each player i ($i = 1, 2$) chooses an action, C_i or D_i. Let $A_i \equiv \{C_i, D_i\}$ be the set of actions for player i. Given an action profile (a_1, a_2), the base game payoff for player i, $u_i(a_1, a_2)$, is displayed in Table 1.

We make the usual assumptions about the above payoff matrix.

<table>
<tr><td></td><td></td><td colspan="2">Player 2</td></tr>
<tr><td></td><td></td><td>C_2</td><td>D_2</td></tr>
<tr><td rowspan="2">Player 1</td><td>C_1</td><td>$1 \ , \ 1$</td><td>$-\ell \ , \ 1+g$</td></tr>
<tr><td>D_1</td><td>$1+g, \ -\ell$</td><td>$0 \ , \ 0$</td></tr>
</table>

Table 1: Prisoner's dilemma

Assumption 1. (i) $g > 0$ and $\ell > 0$; (ii) $g - \ell < 1$.

The first condition implies that action C_i is dominated by action D_i for each player i, and the second condition ensures that the payoff vector of action profile (C_1, C_2) is Pareto efficient. We impose an additional assumption.

Assumption 2. $g - \ell > 0$.

Assumption 2 is the same as Assumption 1 in Chen (2010).

The stage game is simultaneous form. Each player i chooses an action a_i and the observational decision simultaneously. Let m_i represent the observational decision for player i. Let $M_i \equiv \{0, 1\}$ be the set of observational decisions for player i, where $m_i = 1$ represents "to observe the opponent," and $m_i = 0$ represents "not to observe the opponent." If player i observes the opponent, he incurs an observation cost $\lambda > 0$, and he receives complete information about the action chosen by the opponent at the end of the stage game. If player i does not observe the opponent, he does not incur any observation cost and obtains no information about his opponent's action. We assume that the observational decision for a player is unobservable.

A stage behavior for player i is the pair of base game actions a_i for player i and observational decision m_i for player i and denoted by $b_i = (a_i, m_i)$. An outcome of the stage game is the pair of b_1 and b_2. Let $B_i \equiv A_i \times M_i$ be the set of stage behaviors for player i, and let $B \equiv B_1 \times B_2$ be the set of outcomes of the stage game. Given an outcome $b \in B$, the stage game payoff $\pi_i(b)$ for player i is given by

$$\pi_i(b) \equiv u_i(a_1, a_2) - m_i \lambda.$$

For any observation cost $\lambda > 0$, the stage game has a unique stage game Nash equilibrium outcome, $b^* = ((D_1, 0), (D_2, 0))$.

Let $\delta \in (0, 1)$ be a common discount factor. Players maximize their expected average discounted stage game payoffs. Given a sequence of outcomes of the stage games $(b^t)_{t=1}^{\infty}$, player i's average discounted stage game payoff is given by

$$(1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} \pi_i(b^t).$$

By the assumption of no free signals regarding player actions, a player receives no information about the action chosen by his opponent when he does not observe the opponent. This implies that each player does not receive the base game payoffs in the course of play. As in Miyagawa et al. (2003), we interpret the discount factor as the probability with which the repeated game continues, and it is assumed that each player receives the sum of the payoffs when the repeated game ends. Then, the assumption of no free signal regarding the actions is less problematic.

Let $o_i \in A_j \cup \{\phi_i\}$ be an observation result for player i. Observation result $o_i = a_j \in A_j$ implies that player i chose observational decision $m_i = 1$, and observed a_j. Observation result $o_i = \phi_i$ implies that player i chose $m_i = 0$, that is, he obtained no information about the action chosen by the opponent.

Let h_i^t be a (private) history of player i at the beginning of period $t \geq 2$: $h_i^t = (a_i^k, o_i^k)_{k=1}^{t-1}$. It is a sequence of his own actions and his observation results up to period $t-1$. We omit the observational decisions from h_i^t because observation result o_i^k implies the observational decision m_i^k for any k. Let H_i^t denote the set of all the histories for player i at the beginning of period $t \geq 1$, where H_i^1 is an arbitrary singleton set.

A (behavior) strategy for player i of the repeated game is a function of a history of player i to a probability distribution over the set $\Delta(B_i)$ of his stage behavior; $\sigma_i : \bigcup_{t=1}^{\infty} H_i^t \to \Delta(B_i)$.

A belief ψ_i^t of player i in period t is a function of the history h_i^t of player i in period t obtained from a probability distribution over the set of histories for player j in period t. Let $\psi_i \equiv (\psi_i^t)_{t=1}^{\infty}$ be a belief for player i, and $\psi = (\psi_1, \psi_2)$ denote a system of beliefs.

A strategy profile σ is a pair of strategies σ_1 and σ_2. Given a strategy profile σ, a sequence of completely mixed behavior strategy profiles $(\sigma^n)_{n=1}^{\infty}$ that converges to σ is called a *tremble*. Each completely mixed behavior strategy profile σ^n induces a unique system of beliefs ψ^n.

The solution concept is a sequential equilibrium. We say that a system of beliefs ψ is consistent with σ if there exists a tremble $(\sigma^n)_{n=1}^{\infty}$ such that the corresponding sequence of system of beliefs $(\psi^n)_{n=1}^{\infty}$ converges to ψ. Given the system of beliefs ψ, strategy profile σ is sequentially rational if, for each player i, the continuation strategy from each history is optimal given his belief of the history, and the opponent's strategy. It is defined that a strategy profile σ is a *sequential equilibrium* if there exists a consistent system of beliefs ψ for which σ is sequentially rational.

4 Cooperation failure in prisoner's dilemma (Miyagawa et al. (2003))

Let us explain some constraints in prisoner's dilemma. Table 2 below is the bilateral trade game with moral hazard in Bhaskar and van Damme (2002) simplified by Miyagawa et al. (2003).

		Player 2		
		C_2	D_2	E_2
Player 1	C_1	$1\ ,\ 1$	$-1\ ,\ 2$	$-1\ ,\ -1$
	D_1	$2\ ,\ -1$	$0\ ,\ 0$	$-1\ ,\ -1$
	E_1	$-1\ ,\ -1$	$-1\ ,\ -1$	$0\ ,\ 0$

Table 2: Extended Prisoner's Dilemma

Miyagawa et al. (2003) consider the following keep keep-them-guessing strategies to approximate payoff vector $(1, 1)$. There are three state: cooperation state, punishment state, and defection state. In the defection state, both player i choose E_i and the state remains the same. In the punishment state, both player i choose E_i for some periods and the state moves back to cooperation state. In both punishment state and defection state, players do not observe the opponent. In the cooperation state, each player chooses C_i with sufficiently high probability and chooses D_i with the remaining probability. Players observe the opponent in the cooperation state. If players observe (C_1, C_2) or (D_1, D_2), the state remains the same. The state moves to defection state if player i chooses E_i or observes E_j. When (C_1, D_2) or (D_1, C_2) is realized, the state moves to punishment state.

Players have an incentive to observe the opponent because the opponent randomizes actions C_j and D_j in the cooperation state. If player does not observe the opponent, player cannot know the state of the opponent in the next period. If the state of the opponent is cooperation state, then action E_i is suboptimal action because the opponent never chooses action E_j. That is, action E_i has some opportunity cost because the state of the opponent is cooperation state with a high probability. However, if the state of the opponent is defection state, then E_i is unique optimal action. Action C_i and D_i also have some opportunity cost because the state of the opponent is defection state with a positive probability. Therefore, players have an inventive to observe in order to avoid these opportunity costs.

These ideas do not hold in two-action game. Let us consider primitive prisoner's dilemma as an example. If players randomize C_i and D_i in the cooperation state, it means that players best response action always includes action D_i irrespective of the state of the opponent. For example, player can save observation cost in the cooperation state if he does not observe in the current period and chooses D_i and observe in the next period.

In addition, players can distinguish cooperation state and other states by the observation in extended prisoner's dilemma. Actions C_j and D_j mean that the state of the opponent is cooperation state. Action E_j is defection state. That is, player can convey some information by actions. This communication is also limited in a two-action game. [2]

In reality, players sometimes can choose only two types of actions (cooperation and non-cooperation). It means that there is no additional action for communication (e.g., action E_i in the extended prisoner's dilemma). Our results

[2] For further sophisticated application, see Flesch and Perea (2009)

give some understandings of cooperation in these primitive model to describe a reality.

5 No public randomization

In this section, we show our efficiency result without any randomization device. The following proposition shows that the symmetric efficient outcome is approximated by a sequential equilibrium if the observation cost λ is small and the discount factor δ is moderately low.

Proposition 1. *Suppose that Assumptions 1 and 2 are satisfied. For any $\varepsilon > 0$, there exist $\underline{\delta} \in \left(\frac{g}{1+g}, 1\right)$, $\overline{\delta} \in (\underline{\delta}, 1)$, and $\overline{\lambda} > 0$ such that for any discount factor $\delta \in [\underline{\delta}, \overline{\delta}]$ and for any observation cost $\lambda \in (0, \overline{\lambda})$, there exists a symmetric sequential equilibrium whose payoff vector (v_1^*, v_2^*) satisfies $v_i^* \geq 1 - \varepsilon$ for each $i = 1, 2$.*

Proof. See Appendix A. □

An illustration

While the proof in Appendix A provides the detailed construction of an equilibrium that approximates Pareto-efficient payoff vector, we here give its main idea.

Let us consider the following three four automaton: Initial state ω_i^1, cooperation states $(\omega_i^t)_{t=2}^{\infty}$, transition state ω_i^E, defection state ω_i^D. In initial state ω_i^D, player i randomizes three stage behavior: $(C_i, 1)$, $(C_i, 0)$, and $(D_i, 0)$. Player i chooses $(C_i, 1)$ with sufficiently high probability. In cooperation state $\omega_i^t (t \geq 2)$, player i chooses C_i and randomizes observation decision. Player i chooses $(C_i, 1)$ with sufficiently high probability. In transition state and defection state ω_i^D, player i chooses $(D_i, 0)$.

The state transition is described in Figure 1.

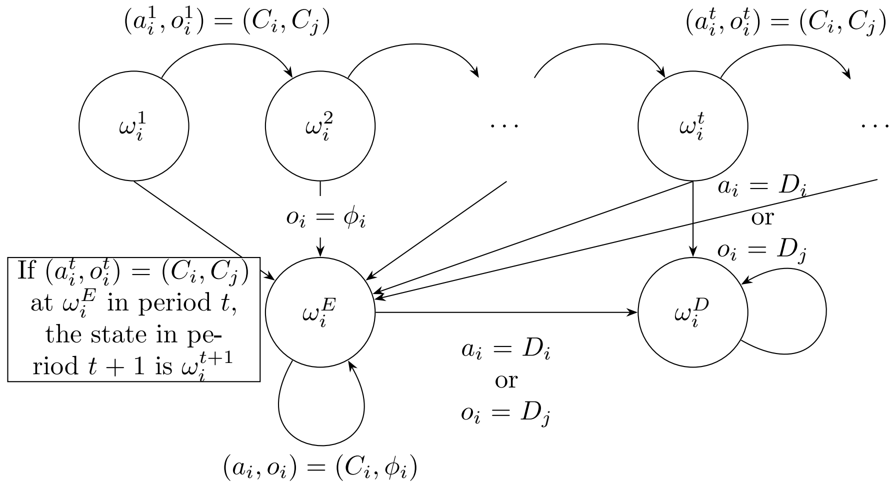

Figure 1: The state-transition rule

That is, player remains cooperation state only when he chooses C_i and observes C_j. Player i moves defection state if he chooses D_i or observes D_j. If player i does not observe the opponent in the cooperation state, he moves to transition state. Although, the stage-behavior in the transition state is the same with that in the defection state, the transition function differs from defection state. Player i moves back to cooperation state from the transition state if he observes (C_i, C_j), which is the event off the equilibrium path.

Another property of this strategy is that players never randomizes actions in cooperation state, whereas players randomizes action in cooperation state to induce the incentive to observe tn the previous literature. Furthermore, we will show in the Appendix that player i strictly prefers action C_i in cooperation state. However, from the perspective of the opponent, player i plays the game as if he randomizes C_i and D_i although he chooses pure actions in each state. It induces an incentive for the other player to observe.

Let us consider the sequential rationality in each state. The sequential rationality in the defection state is obvious. The state is defection state only when player i chose D_i or observed D_j. It implies that both player are sure that the opponent is also in the defection state. Hence player i does not have incentive to choose C_i nor $m_i = 1$ on the equilibrium path.

Next, let us consider off the equilibrium path. The defection state is the unique state off the path. Hence, a sufficient condition for the sequential rationality off the equilibrium path is that player i is certain that the state of the opponent is defection state. To this end, we consider the same belief with one in Miyagawa et al. (2008). That is, we consider a sequence of behavioral strategy profile $(\hat{\sigma}^n)_{n=1}^{\infty}$ such that each strategy profile puts a positive probability to every move but puts far smaller weights on the trembles with respect to the observational decisions than those with respect to actions[3]. This trembles induce a consistent system of beliefs that player i at any defection state is sure that the state of the opponent is defection state.

Let us discuss the sequential rationality in the cooperation state. We choose the randomization probability of observation decisions in the cooperation state so that player i is indifferent between $m_i = 1$ and $m_i = 0$ in the initial state and the cooperation state. Furthermore, this definition ensures that player i strictly prefers action C_i. Suppose that player i weakly prefers action D_i in the next cooperation state. Then, one of the best response action is D_i irrespective of his observation. It means that player i strictly prefers $m_i = 0$ because he can save the observation cost by choosing $(C_i, 0)$ in the current period and $(D_i, 1)$ in the next period. Therefore, the sequential rationality is satisfied in the cooperation state.

Next, let us consider the transition state. We show that why payer i prefers action D_i in the transition state. In transition state, there are two kinds of situations. The situation A is situations where player i is in cooperation state if he observed the opponent in the previous period. The other situation B is situations where player i is in defection state if he observed the opponent in the previous period. Of course, player i cannot distinguish these two situations because he did not observe the opponent. To understand the sequential rationality in the transition state, let us assume that the monitoring cost is almost zero. It means that the deviation payoff to $(D_i, 0)$ in the cooperation state is sufficiently close to the continuation payoff from the cooperation state. Otherwise, player i strictly prefers to observe in the cooperation state to avoid choosing action D_i in the situation A. Therefore, player i is almost indifferent between choosing C_i and D_i in the situation A, whereas player i strictly prefers action D_i in situation B. Hence, player i strictly prefers action D_i when observation cost is sufficiently small because both situations are realized with a positive probability.

Third, let us consider initial state. The indifference condition between C_i and D_i is ensured by the randomization probability between $(C_i, 1)$ and $(C_i, 0)$ in the initial state. If the monitoring probability is high enough, then player i is willing to choose action C_i. The indifference condition between $(C_i, 1)$ and $(C_i, 0)$ in the initial state is ensured by the randomization probability between $(C_i, 1)$ and $(C_i, 0)$ in the initial state in period 2. There is no incentive to choose $(D_i, 1)$ because the state of the opponent in the next period is not cooperation state for sure irrespective of the observation.

Lastly, let us consider the payoff. It is obvious that the equilibrium payoff vector is close to 1 if the probabilities of $(C_i, 1)$ in the initial state and cooperation state are close to 1 and the observation cost is close to 1. In Appendix A, we will show that the equilibrium payoff vector is close to 1 when discount factor is close to $\frac{g}{1+g}$. Another remaining issue is whether our strategy is well-defined or not. It will be proved by solving difference equation when Assumption 2 is satisfied.

We extend Proposition 1 by using Lemma 1.

[3]See Miyagawa et al. (2008) for further discussion.

Lemma 1. *Fix any payoff vector* v *and any* $\varepsilon > 0$*. Suppose that there exist* $\underline{\delta} \in \left(\frac{g}{1+g}, 1\right)$*,* $\overline{\delta} \in (\underline{\delta}, 1)$ *such that for any discount factor* $\delta \in [\underline{\delta}, \overline{\delta}]$*, there exists a sequential equilibrium whose payoff vector* (v_1^*, v_2^*) *satisfies* $|v_i^* - v_i| \geq \varepsilon$ *for each* $i = 1, 2$*. Then, there exist* $\underline{\delta}^* \in (g/1+g, 1)$ *such that for any discount factor* $\delta \in [\underline{\delta}^*, 1)$*, there exists a sequential equilibrium whose payoff vector* (v_1^*, v_2^*) *satisfies* $|v_i^* - v_i| \geq \varepsilon$ *for each* $i = 1, 2$*.*

Proof of Lemma 1 . We define $\underline{\delta}^* \equiv \underline{\delta}/\overline{\delta}$. Choose any discount factor $\delta \in (\underline{\delta}^*, 1)$. Then, we choose some integer n^* that satisfies $\delta^{n^*} \in [\underline{\delta}, \overline{\delta}]$. We divide the repeated game into n^* distinct repeated games. The first repeated game is played in period 1, $n^* + 1$, $2n^* + 1 \dots$, the second repeated game is played in period 2, $n^* + 1$, $2n^* + 2 \dots$, and so on. As each repeated game can be regarded as a repeated game with discount factor δ^{n^*}, strategy σ^* is a sequential rational in each game. Thus, this strategy is a sequential equilibrium. As the equilibrium payoff vector of the original game satisfies $|v_i^* - v_i| \geq \varepsilon$ for each $i = 1, 2$, the equilibrium payoff of this strategy also satisfies $|v_i^* - v_i| \geq \varepsilon$ for each $i = 1, 2$. □

We obtain an efficiency result for a sufficiently high discount factor.

Proposition 2. *Suppose that the base game satisfies Assumptions 1 and 2. For any* $\varepsilon > 0$*, there exist* $\underline{\delta}^* \in (0, 1)$ *and* $\overline{\lambda} > 0$ *such that for any discount factor* $\delta \in (\underline{\delta}^*, 1)$ *and any* $\lambda \in (0, \overline{\lambda})$*, there exists a sequential equilibrium whose payoff vector* (v_1^*, v_2^*) *satisfies* $v_i^* \geq 1 - \varepsilon$ *for each* $i = 1, 2$*.*

Proof of Proposition 2 . Apply Lemma 1 to Proposition 1. □

Remark 1. Proposition 2 shows a kind of monotonicity of the efficiency result on the discount factor. If an efficiency result holds given ε, observation cost λ and discount factor δ, then an efficiency result holds given a sufficiently large discount factor $\delta' > \delta$.

Theorem 1 (Necessary and sufficient condition)**.** *Suppose that Assumption 1 is satisfied. Then, the strategy* σ^* *is a sequential equilibrium for a sufficiently large discount factor and a sufficiently small observation cost if and only if Assumption 2 is satisfied.*

Proof of Theorem 1 . See Corollary 1.3 in Appendix A. □

6 Public randomization

In this section, we assume that an interim public randomization device is available. We assume that player i chooses an observational decision after he chooses his action and an interim public randomization device (sunspot) is realized. The distribution of the public signal is independent of the action profile chosen. Public signal x is uniformly distributed over $[0, 1)$ and each player observes the public signal without any cost.

The purpose of this section is to prove a folk theorem. To prove Theorem 2 (folk theorem), we present the proposition below first.

Proposition 3. *Suppose that an interim public randomization device is available, and the base game satisfies Assumptions 1 and 2. For any* $\varepsilon > 0$*, there exist* $\underline{\delta} \in \left(\frac{g}{1+g}, 1\right)$*,* $\overline{\delta} \in (\underline{\delta}, 1)$*, and* $\overline{\lambda} > 0$ *such that for any discount factor* $\delta \in [\underline{\delta}, \overline{\delta}]$ *and for any observation cost* $\lambda \in (0, \overline{\lambda})$*, there exists a sequential equilibrium whose payoff vector* (v_1^*, v_2^*) *satisfies* $v_1^* = 0$ *and* $v_2^* \geq \frac{1+g+\ell}{1+\ell} - \varepsilon$*.*

Proof of Proposition 3 . See Appendix B. □

An illustration

We will give the proof in Appendix B and provides the detailed construction of an equilibrium that approximates asymmetric Pareto-efficient payoff vector $(0, \frac{1+\ell+g}{1+\ell})$. We show its main idea in this section.

A rough idea of our idea is that players play (C_1, D_2) in the first period, and then players play the strategy in the proof of Proposition 1 from period 2 on. Applying the strategy in Section 5, let us consider the following strategy. In period 1, players play (C_1, D_2). If players did not play (C_1, D_2) in period 1, then players are in a defection state in period 2 onwards. If players did play (C_1, D_2) in period 1, players play a sequential equilibrium whose payoff vector is sufficiently close to $(1, 1)$, which is similar to the one in Section 5. We show that a similar strategy to the above strategy is a sequential equilibrium.

Let us describe the strategy in detailed. In the first period, player 1 randomizes C_1 and D_1, and does not observe the opponent. The state remains the same if the realized sunspot x is greater than $\hat{x}$. Player 1 moves to cooperation state if he chooses C_1 and $x \leq \hat{x}$. He moves to defection state if he chooses D_1 and $x \leq \hat{x}$. Player 2 randomizes the observational decision when the realized sunspot is not greater than $\hat{x}$. Otherwise, player 2 does not observe. Player 2 moves to cooperation state if he observes C_1, he moves to defection state if he observes D_1, and he moves to transition state if he does not observe the opponent.

The behavior and the transition function of player 1 in cooperation state in period 2 differ from the one in the proof of Proposition 1. Player 1 in cooperation state in period 2 plays the game as if he is in the "initial state" in the proof of Proposition 1. That is, player 1 randomizes $(C_1, 1)$, $(C_1, 0)$, and $(D_1, 0)$. If player 1 observes (C_1, C_2), the state remains the same. The state moves to defection state if player 1 chooses D_1 or observes D_2. The state is transition state if player 1 chooses $(C_1, 0)$.

The other construction of the strategy (e.g., defection state in period 2, strategy of player 2 in period 2, and the strategy from period 3 on ward) is the same with the one in the proof of Proposition 1.

Let us consider sequential rationalities of players. The sequential rationality in defection state both on and off the equilibrium path holds in the same manner in the Section 5. The sequential rationality in the cooperation state from period 3 on holds as well.

Let us consider the sequential rationality of player 1 in the cooperation state in period 2. Player 1 cannot distinguish whether the state of the opponent is cooperation state or not because the observational decision is unobservable. If player 2 observes in the previous period, he chooses C_2. Otherwise, player 2 chooses D_2. Therefore, from the viewpoint of player 1, player 2 randomizes three stage-behavior: $(C_2, 1)$, $(C_2, 0)$, and $(D_2, 0)$ like the initial state in the proof of Proposition 1. Hence, if player 2 chooses appropriate randomization probability of $(C_2, 1)$, $(C_2, 0)$, and $(D_2, 0)$, then player 1 is indifferent between $(C_1, 1)$, $(C_1, 0)$, and $(D_1, 0)$. Next, let us consider the sequential rationality of player 2 in the cooperation state in period 2. As player 1 randomizes $(C_1, 1)$, $(C_1, 0)$, and $(D_1, 0)$, it is easily satisfied when player 1 chooses appropriate randomization probability.

Let us consider the sequential rationality in period 1. As Assumption 2 is satisfied, the deviation to action D_i is more profitable in terms of the stage game payoff when the opponent chooses D_j than when the opponent chooses C_j. The incentive for player 1 to choose C_1 is higher than the one in the proof of Proposition 1. Therefore, we use an interim public randomization device to decrease the incentive to choose action C_1. The sequential rationality of player 2 holds as well because player 1 randomizes C_1 and D_1 with moderate probability. Therefore, the strategy will be a sequential equilibrium.

The last issue is the equilibrium payoff. Given this strategy, we have to consider the effect of interim public randomization device to the equilibrium payoff. Let V_i be the payoff for player i for each $i = 1, 2$. In the proof of Proposition 1, we have shown that Pareto efficient payoff vector $(1, 1)$ can be approximated by a sequential equilibrium when the discount factor is close to $\frac{g}{1+g}$. Therefore, the continuation payoff when player 1 moves to cooperation state in period 2 is close to 1. The value of $\hat{x}$ is given as the solution of the following equation.

$$-(1-\delta)\ell + \delta\hat{x} \cdot 1 + \delta(1-\hat{x})V_1 = (1-\delta) \cdot 0 + \delta\hat{x} \cdot 0 + \delta(1-\hat{x})V_1$$

The left-hand side is the payoff when player 1 chooses C_1, and the right-hand side is the one when he chooses D_1. Therefore, we have $\hat{x} = \frac{1-\delta}{\delta}\ell$. Then, the payoff V_2 of player 2 can be approximated by the following equation.

$$\begin{aligned} V_2 =& (1-\delta)(1+g) + \delta\hat{x} \cdot 1 + \delta(1-\hat{x})V_2 \\ =& \frac{(1-\delta)(1+g) + \delta\hat{x} \cdot 1}{1-\delta+\delta\hat{x}} \\ =& \frac{1+g+\ell}{1+\ell} \end{aligned}$$

We have obtained the desired result.

Corollary 3.1. *Suppose that an interim public randomization device is available, and the base game satisfies Assumptions 1 and 2. For any* $\varepsilon > 0$, *there exist* $\underline{\delta} \in \left(\frac{g}{1+g}, 1\right)$ *and* $\overline{\lambda} > 0$ *such that for any discount factor* $\delta \in [\underline{\delta}, 1)$ *and for any observation cost* $\lambda \in (0, \overline{\lambda})$, *there exists a sequential equilibrium whose payoff vector* (v_1^*, v_2^*) *satisfies* $v_1^* = 0$ *and* $v_2^* \geq \frac{1+g+\ell}{1+\ell} - \varepsilon$.

Proof of Corollary 3.1 . Use Lemma 1. □

Hence, we have shown that two kinds of payoff vector can be approximated by sequential equilibria (Proposition 1 and Proposition 3) when the discount factor is sufficiently large and the observation cost is sufficiently small.

By utilizing interim public randomization again, we obtain the folk theorem below.

Theorem 2. *Suppose that an interim public randomization is available, and Assumptions 1 and 2 are satisfied. Fix any interior point* $v = (v_1, v_2)$ *of* $\mathcal{F}^*$*. There exist a discount factor* $\underline{\delta} \in \left(\frac{g}{1+g}, 1\right)$ *and observation cost* $\overline{\lambda} > 0$ *such that for any* $\delta \in [\underline{\delta}, 1)$ *and* $\lambda \in (0, \overline{\lambda})$*, there exists a sequential equilibrium whose payoff vector is* v*.*

Proof of Theorem 2. Without loss of generality, let us assume that $v_1 \leq v_2$. By Corollary 3.1, there exists a sequential equilibrium whose payoff vector $v^* = (v_1^*, v_2^*)$ is sufficiently close to $\left(0, \frac{1+g+\ell}{1+\ell}\right)$ and satisfies $\delta v_1^* > v_1$ when discount factor δ is sufficiently large. We can also find a sequential equilibrium whose payoff vector $v^{**} = (v_1^{**}, v_2^{**})$ is sufficiently close to $(1, 1)$ and satisfies $\delta v_2^{**} > v_2$ by Proposition 2.

The desired payoff vector v can be expressed uniquely as a convex combination of δv^*, δv^{**} and $(0, 0)$ as below.

$$v = \alpha_1 \delta v^* + \alpha_2 \delta v^{**} + (1 - \alpha_1 - \alpha_2) \cdot 0.$$

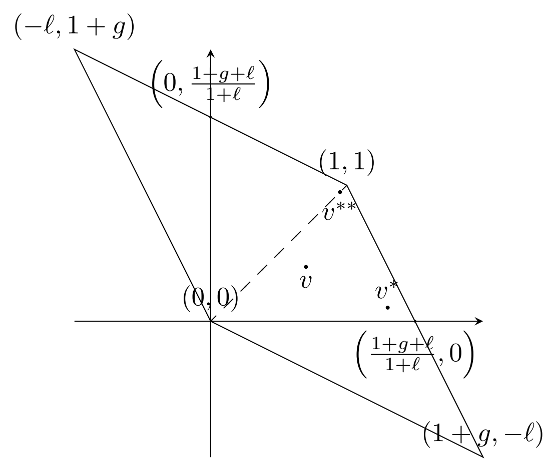

Figure 2: v, v^*, v^{**}

Let us consider the following strategy. In period 1, each player chooses $(D_i, 0)$. If the realized interim public randomization device is smaller than α_1, players play a sequential equilibrium strategy whose payoff vector is v^* from period 2 onwards. If the realized interim public randomization device is not smaller than α_1 but smaller than $\alpha_1 + \alpha_2$, players play a sequential equilibrium strategy whose payoff vector is v^{**} from period 2 on. Otherwise, players play a repetition of the stage game Nash equilibrium every period. This strategy is a sequential equilibrium and its payoff vector is exactly v. $\square$

Remark 2. Our result holds under the monitoring structure of Flesch and Perea (2009) if an interim public randomization device is available. Our result is a variant of the grim trigger strategy. Therefore, each player does not have an incentive to acquire information about the action chosen in the past.

7 Discussion

We have proved efficiency results and folk theorem in repeated symmetric prisoner's dilemma. In this section, we discuss what happens if the prisoner's dilemma is asymmetric as in Table 3.

		Player 2	
		C_2	D_2
Player 1	C_1	1 , 1	$-\ell_1$, $1+g_2$
	D_1	$1+g_1$, $-\ell_2$	0 , 0

Table 3: Asymmetric prisoner's dilemma

In the proofs of any propositions and theorems, we require that the discount factor δ is sufficiently close to $\frac{g}{1+g}$. This condition is required to approximate an Pareto-efficient payoff vector. If $g_1 \neq g_2$, it is impossible to satisfy that the discount factor δ is sufficiently close to both $\frac{g_1}{1+g_1}$ and $\frac{g_2}{1+g_2}$. Therefore, we have to confine our attention to the case of $g_1 = g_2 = g$.

Let us consider Propositions 1 and 2. In the construction of the strategy, the randomization probability of player i is defined based on the incentive constraint of the opponent only. In other words, the randomization probability is determined independently of the payoffs of player i. It means that the randomization probability of player i is determined based on δ, g, ℓ_j and independent of ℓ_i. Therefore, we can discuss the randomization probabilities of player 1 and 2 independently. Hence, if $g_1 = g_2$ and Assumptions 1 and 2 for each ℓ_i $(i = 1, 2)$ hold, our efficiency result and our folk theorem under small observation cost.

8 Concluding Remarks

The ways of cooperation in a two-player, two-action prisoner's dilemma is most limited even though it is a meaningful model. First, the number of actions is limited. This means that players cannot communicate by using a variety of actions. Second, the number of players is limited. If there are three players A, B, C, it is easy to check the observation deviation of the opponents. Player A can monitor the observation decisions of players B and C by comparing the actions of B and

C. If players B and C choose inconsistent actions toward each other, player A finds that players B or C do not observe some player. Third, there is no free-cost informative signal. Players have to observe to obtain the information about the action chosen by their opponents.

Originally, the prisoner's dilemma has these constraints. Despite the above limitation, we have shown an efficiency result without any randomization device. Our paper is the first result that shows an efficiency holds without public randomization under infinitely repeated prisoner's dilemma with costly monitoring, although it is the simplest model among those with costly monitoring considered in the previous literature (e.g.,Miyagawa et al. (2003) and Flesch and Perea (2009)).

We considered interim public randomization device and obtained a folk theorem. It is worth mentioning that our folk theorem holds in asymmetric prisoner's dilemma. Our results might be applied to more general games.

A Proofs of Proposition 1 and its corollaries

Proof. We prove Proposition 1 and its corollaries.

Strategy

We define a grim trigger strategy σ^*, and then we define a consistent system of beliefs ψ^*. Strategy σ^* is represented by an automaton that has four kind of states: initial state ω_i^1, cooperation state $(\omega_i^t)_{t=2}^{\infty}$, transition state ω_i^E and defection state ω_i^D. For any period $t \geq 2$, there is a unique cooperation state. Let ω_i^t be the cooperation state in period $t \geq 2$.

At initial state ω_i^1, each player i chooses $(C_i, 1)$ with probability $(1-\beta_1)(1-\beta_2)$, chooses $(C_i, 0)$ with probability $(1-\beta_1)\beta_2$, and chooses $(D_i, 0)$ with probability β_1. We call (a_i, o_i) an action–observation pair. The state moves from the initial state to cooperation state ω_i^2 if the action–observation pair in period 1 is (C_i, C_j). The state moves to transition state ω_i^E in period 2 when (a_i^1, o_i^1) is (C_i, ϕ_i) realized in period 1. Otherwise, the state moves to a defection state in period 2.

At cooperation state ω_i^t, each player i chooses $(C_i, 1)$ with probability $1-\beta_{t+1}$ and $(C_i, 0)$ with probability β_{t+1}. That is, the randomization probability β_{t+1} depends on calendar time t. The state moves to the next cooperation state ω_i^{t+1} if the action–observation pair in period t is (C_i, C_j). The state moves to transition state ω_i^E in period $t+1$ when (a_i^t, o_i^t) is (C_i, ϕ_i) realized in period t. Otherwise, the state moves to a defection state in period $t+1$.

At transition state ω_i^E in period t, each player i chooses $(D_i, 0)$ with certainty. The state moves to defection state ω_i^D in period $t+1$ when $a_i^t = D_i$ or $o_i^t = D_j$ is realized. If player i chooses $(C_i, 0)$, the state remains the same. When player i chooses C_i and observes C_j in period t, the state in period $t+1$ moves to cooperation state ω_i^{t+1}.

Players choose $(D_i, 0)$ and the state remains the same ω_i^D at defection state ω_i^D irrespective of the action–observation pair.

The state-transition rule is summarized in Figure 1. Let strategy σ^* be the strategy represented by the above automaton.

We define a system of beliefs consistent with strategy σ^* by the same tremble as the one in Miyagawa et al. (2008). That is, we consider a sequence of behavioral strategy profile $(\hat{\sigma}^n)_{n=1}^{\infty}$ such that each strategy profile puts a positive probability to every move but puts far smaller weights on the trembles with respect to the observational decisions than those with respect to actions[4]. Each behavioral strategy profile $\hat{\sigma}^n$ induces a the system of belief ψ^n and we define the consistent system of beliefs ψ^* as the limit of $\lim_{n\to\infty} \psi^n$.

Selection of discount factor and observation cost

Fix any $\varepsilon > 0$. We define $\overline{\varepsilon}$, $\underline{\delta}$, $\overline{\delta}$ and $\overline{\lambda}$ as follows

$$\begin{aligned}
\overline{\varepsilon} \equiv & \frac{\ell^2}{54(1+g+\ell)^3}\frac{\varepsilon}{1+\varepsilon}, \\
\underline{\delta} \equiv & \frac{g}{1+g} + \overline{\varepsilon}, \\
\overline{\delta} \equiv & \frac{g}{1+g} + 2\overline{\varepsilon}, \\
\overline{\lambda} \equiv & \frac{1}{16}\frac{g}{1+g}\frac{1}{1+g+\ell}\overline{\varepsilon}^2.
\end{aligned}$$

We fix an arbitrary discount factor $\delta \in [\underline{\delta}, \overline{\delta}]$ and an arbitrary observation cost $\lambda \in (0, \overline{\lambda})$. We will show that there exists a sequential equilibrium whose payoff vector (v_1^*, v_2^*) satisfies $v_i^* \geq 1 - \varepsilon$ for each $i = 1, 2$.

Specification of strategy

Let us define $\varepsilon' \equiv \delta - \frac{g}{1+g}$. We set $\beta_1 = \frac{1+g+\ell}{g+\ell}\varepsilon'$. Given β_1, we define β_2 as the solution of the following indifference condition between $(C_i, 0)$ and $(D_i, 0)$ in period 1.

$$(1-\beta_1)\cdot 1 - \beta_1 \cdot \ell + \delta(1-\beta_1)(1-\beta_2)(1+g) = (1-\beta_1)(1+g). \qquad (1)$$

Next, we define $(\beta_t)_{t=3}^{\infty}$. We choose β_{t+2} so that player j at state ω_i^t is indifferent between choosing $(C_i, 1)$ and choosing $(C_i, 0)$.

To define $\beta_t (t \geq 3)$, let W_t $(t \geq 1)$ be the sum of the stage game payoffs from state ω_i^t. That is, payoff W_t is given by

$$W_t = \left[\sum_{s=1}^{\infty} \delta^{s-1} u_i(a^{t+s-1}) \middle| \sigma^*, \psi^*, h_i^t \right],$$

[4]See Miyagawa et al. (2008) for further information.

where h_i^t is a history associated with cooperation state ω_i^t. In cooperation state $\omega_i^t (t \geq 2)$, player i weakly prefers to play $(C_i, 0)$. Therefore, the payoff W_t is given by

$$W_t = (1-\beta_t) \cdot 1 - \beta_t \ell + \delta(1-\beta_t)(1-\beta_{t+1})(1+g), \quad \forall\, t \geq 2. \tag{2}$$

Then, β_3 is given by

$$W_1 = (1-\beta_1) \cdot 1 - \beta_1 \ell - \lambda + \delta(1-\beta_1) W_2. \tag{3}$$

Note that W_2 is a function of β_3 by (2).

Next, let us consider the indifference condition between $(C_i, 1)$ and $(C_i, 0)$ at cooperation state $\omega_i^t (t \geq 2)$. Let us consider the belief for each player i at cooperation state ω_i^t in period t. Assume that $\beta_t \in (0,1)$ for any $t \in \mathbb{N}$, which will be proved later. Then, we show by mathematical induction that, for any period $t \geq 2$, player i at cooperation state ω_i^t in period t believes that the state of his opponent is a cooperation state with positive probability $1-\beta_t$. The state moves to cooperation state ω_i^2 in period 2 only when player i has observed the action–observation pair $(a_i^1, o_i^1) = (C_i, C_j)$ in period 1. Therefore, player i believes that the state of his opponent is a cooperation state with positive probability $1-\beta_2$ by Bayes' rule. Thus, the statement is true in period 2. Next, suppose that the statement is true until period t and consider a player i at cooperation state ω_i^{t+1}. This means that player i has observed action–observation pair $(a_i^t, o_i^t) = (C_i, C_j)$ in period t. Player i believes that the state of his opponent in period t was a cooperation state with certainty. Therefore, he believes that the state of his opponent in period $t+1$ is a cooperation state with positive probability $1-\beta_t$ by Bayes' rule. Hence, the statement is true.

Taking the belief at cooperation state $\omega_i^t (t \geq 2)$ into account, β_{t+2} is defined as the solution of the equation below.

$$W_t = (1-\beta_t) \cdot 1 - \beta_t \ell - \lambda + \delta(1-\beta_t) W_{t+1}. \tag{4}$$

Note that W_{t+1} is a function of β_{t+2} by (2).

Specifically, $(\beta_t)_{t=2}^{\infty}$ is determined by the following equations.

$$\begin{aligned}
\beta_2 =& \frac{(1-\beta_1)\{\delta(1+g) - g\} - \beta_1 \ell}{\delta(1-\beta_1)(1+g)} \\
=& \frac{g + g^2 - \ell^2 - (1+g+\ell)\varepsilon'}{(g+\ell)\{g + (1+g)\varepsilon'\}\left(1 - \frac{1+g+\ell}{g+\ell}\varepsilon'\right)} \varepsilon' \\
\beta_{t+2} =& \frac{(1-\beta_{t+1})\{\delta(1+g) - g\} - \beta_{t+1}\ell + \frac{\lambda}{\delta(1-\beta_t)}}{\delta(1-\beta_{t+1})(1+g)}, \quad \forall\, t \in \mathbb{N}.
\end{aligned}$$

Before we proceed to the proof, we will show that $(\beta_t)_{t=1}^{\infty}$ is well defined. To prove it, we will show that $\frac{\ell}{2g} < -\frac{\beta_{t+2} - \beta_{t+1}}{\beta_{t+1} - \beta_t} < 1$ for any $t \in \mathbb{N}$ because β_{t+2} can be expressed by using β_t, β_{t+1}, and $-\frac{\beta_{t+2}-\beta_{t+1}}{\beta_{t+1}-\beta_t}$ as follows.

$$\begin{aligned}\beta_{t+2} =&\beta_t + (\beta_{t+1}-\beta_t) + (\beta_{t+2}-\beta_{t+1})\\ =&\beta_t + (\beta_{t+1}-\beta_t)\left\{1-\left(-\frac{\beta_{t+2}-\beta_{t+1}}{\beta_{t+1}-\beta_t}\right)\right\}\\ =&\left(-\frac{\beta_{t+2}-\beta_{t+1}}{\beta_{t+1}-\beta_t}\right)\beta_t + \left\{1-\left(-\frac{\beta_{t+2}-\beta_{t+1}}{\beta_{t+1}-\beta_t}\right)\right\}\beta_{t+1}.\end{aligned}$$

Therefore, if $\beta_t, \beta_{t+1} \in [0,1]$, and $\frac{\ell}{2g} < -\frac{\beta_{t+2}-\beta_{t+1}}{\beta_{t+1}-\beta_t} < 1$ hold, we obtain $\beta_{t+2} \in (\min\{\beta_t, \beta_{t+1}\}, \max\{\beta_t, \beta_{t+1}\})$ because β_{t+2} is a convex combination of β_t and β_{t+1}.

Lemma 2. *Suppose that Assumptions 1 and 2 are satisfied. Fix any discount factor* $\delta \in [\underline{\delta}, \overline{\delta}]$ *and observation cost* $\lambda \in (0, \overline{\lambda})$*. Then, for any* $t \in \mathbb{N}$*, it holds that*

$$0 < \frac{\ell}{2g} < -\frac{\beta_{t+2}-\beta_{t+1}}{\beta_{t+1}-\beta_t} < \frac{g+\ell}{2g} < 1.$$

Proof of Lemma 2. First, let us derive $-\frac{\beta_3-\beta_2}{\beta_2-\beta_1}$. By (1), we have

$$0 = -(1-\beta_1)g - \beta_1\ell + \delta(1+g)(1-\beta_1)(1-\beta_2). \tag{5}$$

Furthermore, by (2) and (3), we have

$$\frac{\lambda}{\delta(1-\beta_1)} = -(1-\beta_2)g - \beta_2\ell + \delta(1+g)(1-\beta_2)(1-\beta_3) \tag{6}$$

By (5) and (6), we obtain

$$(\beta_2-\beta_1)(g-\ell) - \delta(1+g)(1-\beta_2)\left\{(\beta_3-\beta_2)+(\beta_2-\beta_1)\right\} = \frac{\lambda}{\delta(1-\beta_1)}.$$

The definition of $\overline{\varepsilon}$ ensures

$$\frac{1}{2}\frac{1+g-\ell}{g+\ell}\varepsilon' < \beta_2 < \frac{1+g}{g+\ell}\varepsilon'.$$

As $\beta_2 < \frac{1+g}{g+\ell}\varepsilon' < \beta_1$, we can divide both sides by $\beta_2-\beta_1$, and obtain $-\frac{\beta_3-\beta_2}{\beta_2-\beta_1}$.

$$-\frac{\beta_3-\beta_2}{\beta_2-\beta_1} = \frac{\ell + \delta(1+g)(1-\beta_2) - g + \frac{\lambda}{\delta(1-\beta_1)(\beta_2-\beta_1)}}{\delta(1+g)(1-\beta_2)}.$$

As Assumption 2, $\beta_1, \beta_2 < 1$, and $\beta_2 - \beta_1 < 0$ holds, we find an upper bound of $-\frac{\beta_3-\beta_2}{\beta_2-\beta_1}$.

$$-\frac{\beta_3-\beta_2}{\beta_2-\beta_1} \leq \frac{\delta(1+g)(1-\beta_2) + \frac{\lambda}{\delta(1-\beta_1)(\beta_2-\beta_1)}}{\delta(1+g)(1-\beta_2)} < 1.$$

Taking $\beta_1 = \frac{1+g+\ell}{g+\ell}\varepsilon'$, $\beta_2 < \frac{1+g}{g+\ell}\varepsilon'$, and $-(\beta_2 - \beta_1) > \frac{\ell}{g+\ell}\varepsilon'$ into account, we have a lower bound of $-\frac{\beta_3-\beta_2}{\beta_2-\beta_1}$ as follows.

$$\begin{aligned}-\frac{\beta_3 - \beta_2}{\beta_2 - \beta_1} >& \frac{\left(\frac{g}{1+g} + \varepsilon'\right)(1+g)\left(1 - \frac{1+\ell}{2\ell}\varepsilon'\right) - g + \ell - \frac{\ell}{\left(\frac{g}{1+g}+\varepsilon'\right)\left(1-\frac{1+g+\ell}{g+\ell}\varepsilon'\right)}\frac{\lambda}{\varepsilon'}}{\left(\frac{g}{1+g} + \varepsilon'\right)(1+g)} \\ >& \frac{\ell}{2g}.\end{aligned}$$

The last inequality is ensured by $\varepsilon' < 2\overline{\varepsilon}$ and $\lambda < \overline{\lambda}$. Therefore, we have obtained $\frac{\ell}{2g} < -\frac{\beta_3-\beta_2}{\beta_2-\beta_1} < 1$ and $\beta_3 \in (\beta_2, \beta_2)$. That is, $\beta_3 - \beta_2 > 0$.

Next, let us derive $-\frac{\beta_{t+3}-\beta_{t+2}}{\beta_{t+2}-\beta_{t+1}}$ inductively. Suppose that $\frac{\ell}{2g} < -\frac{\beta_{s+2}-\beta_{s+1}}{\beta_{s+1}-\beta_s} < 1$ and $\beta_{s+2} \in (\min\{\beta_s, \beta_{s+1}\}, \max\{\beta_s, \beta_{s+1}\})$ holds for period $s = 1, 2, 3 \ldots, t$. We have shown that this supposition holds for $t = 1$. We will show that $\frac{\ell}{2g} < -\frac{\beta_{t+3}-\beta_{t+2}}{\beta_{t+2}-\beta_{t+1}} < 1$ and $\beta_{t+3} \in (\min\{\beta_{t+1}, \beta_{t+2}\}, \max\{\beta_{t+1}, \beta_{t+2}\})$ holds.

By (2), (3), and (4), for any $t \in \mathbb{N}$, we have

$$\begin{cases}\frac{\lambda}{\delta(1-\beta_t)} = -(1-\beta_{t+1})g - \beta_{t+1}\ell + \delta(1-\beta_{t+1})(1-\beta_{t+2})(1+g) \\ \frac{\lambda}{\delta(1-\beta_{t+1})} = -(1-\beta_{t+2})g - \beta_{t+2}\ell + \delta(1-\beta_{t+2})(1-\beta_{t+3})(1+g),\end{cases}$$

or, equivalently,

$$\begin{aligned}&-\frac{\beta_{t+1} - \beta_t}{\delta(1-\beta_t)(1-\beta_{t+1})}\lambda \\ =&(\beta_{t+2} - \beta_{t+1})(g-\ell) - \delta(1-\beta_{t+2})\left\{(\beta_{t+3} - \beta_{t+2}) + (\beta_{t+2} - \beta_{t+1})\right\}(1+g).\end{aligned}$$

The suppositions ensure $\beta_{t+2}-\beta_{t+1} \neq 0$. Divide both sides of the above equation by $\beta_{t+2} - \beta_{t+1}$. Then, we obtain

$$-\frac{\beta_{t+3} - \beta_{t+2}}{\beta_{t+2} - \beta_{t+1}} = \frac{\ell + \delta(1-\beta_{t+2})(1+g) - g - \frac{1}{\delta(1-\beta_t)(1-\beta_{t+1})\frac{\beta_{t+2}-\beta_{t+1}}{\beta_{t+1}-\beta_t}}\lambda}{\delta(1+g)(1-\beta_{t+2})}.$$

By Assumption 2, $\beta_t, \beta_{t+1} < 1$, and $\frac{\beta_{t+2}-\beta_{t+1}}{\beta_{t+1}-\beta_t} < 0$ hold, $-\frac{\beta_{t+3}-\beta_{t+2}}{\beta_{t+2}-\beta_{t+1}}$ is bounded above by

$$-\frac{\beta_{t+3} - \beta_{t+2}}{\beta_{t+2} - \beta_{t+1}} \leq \frac{\delta(1+g)(1-\beta_{t+2}) + \frac{1}{\delta(1-\beta_t)(1-\beta_{t+1})\frac{\beta_{t+2}-\beta_{t+1}}{\beta_{t+1}-\beta_t}}\lambda}{\delta(1+g)(1-\beta_{t+2})} < 1.$$

Taking $0 < \beta_{t+1}, \beta_{t+2} < \frac{1+g+\ell}{g+\ell}\varepsilon' = \beta_1$, and $\frac{\ell}{2g} < -\frac{\beta_{t+2}-\beta_{t+1}}{\beta_{t+1}-\beta_t} < 1$ into account, we find the following lower bound of $-\frac{\beta_{t+3}-\beta_{t+2}}{\beta_{t+2}-\beta_{t+1}}$.

$$\begin{aligned}-\frac{\beta_{t+3} - \beta_{t+2}}{\beta_{t+2} - \beta_{t+1}} >& \frac{\ell + \left(\frac{g}{1+g} + \varepsilon'\right)(1+g)\left(1 - \frac{2+g}{2\ell}\varepsilon'\right) - g - \frac{1}{\left(\frac{g}{1+g}+\varepsilon'\right)\left(1-\frac{2+g}{2\ell}\varepsilon'\right)^2\frac{2g}{\ell}}\lambda}{\left(\frac{g}{1+g} + \varepsilon'\right)(1+g)} \\ >& \frac{\ell}{2g}.\end{aligned}$$

Therefore, we obtained $\frac{\ell}{2g} < -\frac{\beta_{t+3}-\beta_{t+2}}{\beta_{t+2}-\beta_{t+1}} < 1$ and $\beta_{t+3} \in (\min\{\beta_{t+1}, \beta_{t+2}\}, \max\{\beta_{t+1}, \beta_{t+2}\})$. $\square$

Corollary 1.2 (Corollary of Lemma 2)**.** *Suppose that Assumptions 1 and 2 are satisfied. Fix any discount factor $\delta \in [\underline{\delta}, \overline{\delta}]$ and observation cost $\lambda \in (0, \overline{\lambda})$. Then, it holds that*

$$\frac{1}{2}\frac{1+g-\ell}{g+\ell}\varepsilon' < \beta_2 < \beta_4 < \beta_6 \cdots < \beta_5 < \beta_3 < \beta_1 = \frac{1+g+\ell}{g+\ell}\varepsilon'.$$

Proof of Corollary 1.2. Let us compare β_1, β_2, and β_3. As we have already shown, β_1 is greater than $\beta_2 > \left(\frac{1}{2}\frac{1+g-\ell}{g+\ell}\varepsilon'\right)$. Furthermore, we have $\beta_2 < \beta_3 < \beta_1$ because β_3 is a convex combination of β_1 and β_2. Next, let us compare β_2, β_3, and β_4. As we know, β_2 is smaller than β_3. Therefore, we have $\beta_2 < \beta_4 < \beta_3$ because β_4 is a convex combination of β_2 and β_3. Similarly, for any $s \in \mathbb{N}$, it holds that $(\beta_{2s} <)\beta_{2s+1} < \beta_{2s-1}$, and $\beta_{2s} < \beta_{2s+2}(< \beta_{2s+1})$. $\square$

Lastly, let us consider what happens if Assumption 2 is not satisfied.

Corollary 1.3 (Corollary of Lemma 2)**.** *Suppose that Assumption 1 is satisfied, but 2 is not satisfied. Then, $(\beta_t)_{t=1}^{\infty}$ is not well defined for small observation cost λ.*

Proof of Corollary 1.3 . We have

$$-\frac{\beta_{t+3}-\beta_{t+2}}{\beta_{t+2}-\beta_{t+1}} = 1 - \frac{g-\ell}{\delta(1-\beta_{t+2})(1+g)} - \frac{1}{\delta^2(1-\beta_t)(1-\beta_{t+1})(1-\beta_{t+1})\frac{\beta_{t+2}-\beta_{t+1}}{\beta_{t+1}-\beta_t}}\lambda.$$

Therefore, if $g - \ell \leq 0$ and λ is small, then $-\frac{\beta_{t+3}-\beta_{t+2}}{\beta_{t+2}-\beta_{t+1}} > 1$, and $|\beta_t|$ goes to infinity as t goes to infinity. That is, we have obtained a necessary condition for the efficiency result. $\square$

Now, let us show that the grim trigger strategy σ^* is a sequential equilibrium.

Sequential rationality at the initial state

At the initial state, the indifference condition between $(C_i, 0)$ and $(D_i, 0)$ is ensured by the construction of β_2. The indifference condition between $(C_i, 1)$ and $(C_i, 0)$ is ensured by the construction of β_3. Furthermore, if player i chooses action D_i, then his opponent chooses action D_j with certainty from the next period on, irrespective of his observation result. Thus, player i has no incentive to choose $(D_i, 1)$. Therefore, it is optimal for player i to follow strategy σ^* at the initial state.

Sequential rationality in the cooperation state

Next, consider a history associated with a cooperation state in period t (≥ 2). Then, strategy σ^* prescribes to randomize $(C_i, 1)$ and $(C_i, 0)$. The definition of β_{t+2} ensures that $(C_i, 1)$ and $(C_i, 0)$ are indifferent for player i in period t. When player i chooses $(D_i, 0)$ or $(D_i, 1)$, then the continuation payoff is bounded above by $(1-\beta_t)(1+g)$. The equation (4) implies that, for any $t \in \mathbb{N}$, it holds that

$$W_{t+1} = (1-\beta_{t+1})(1+g) + \frac{\lambda}{\delta(1-\beta_t)}. \tag{7}$$

The above equality ensures that, for any period $t \geq 1$, $(1-\beta_{t+1})(1+g)$ is strictly smaller than W_{t+1}, which is the payoff when player i chooses $(C_i, 1)$ in period $t+1$. Thus, both $(D_i, 0)$ and $(D_i, 1)$ are suboptimal in any period $t \geq 2$. Therefore, it is optimal for player i to follow strategy σ^* in a cooperation state.

Sequential rationality at the defection state

Consider any history associated with a defection state. Then, σ^* prescribes $(D_i, 0)$. Player i is certain that the state of his opponent is a defection state, and player i's action in that period does not affect the continuation play of his opponent. Furthermore, player i believes that player j chooses action D_j with certainty and has no incentive to observe his opponent. Therefore, it is optimal for player i to follow strategy σ^* in a defection state.

Sequential rationality in the transition state

We consider any history in period t (≥ 2) associated with a transition state. Strategy σ^* prescribes $(D_i, 0)$ in a transition state.

Let us consider a continuation payoff when player i chooses action C_i in period t. Let p be the belief of player i in the transition state in period t that his opponent is in a cooperation state. If player i observes his opponent, then $(a_i^t, o_i^t) = (C_i, C_j)$ is realized with probability p and the state moves to cooperation state (ω_i^{t+1}). The continuation payoff in the cooperation state in period $t+1$ is bounded above by W_{t+1}. This is because W_{t+1} is a continuation payoff when player i chooses action C_i from ω_i^{t+1}, and W_{t+1} is strictly greater than payoff $(1-\beta_{t+1})(1+g)$, which is the upper bound of the payoff when player i chooses action D_i at ω_i^{t+1}. Therefore, the upper bound of the payoff when player i chooses action C_i in period t is given by

$$p - (1-p)\ell + \delta p W_{t+1}.$$

The payoff when player i chooses D_i is bounded above by $p(1+g)$. Therefore, action D_i is profitable if the following value is negative.

$$p - (1-p)\ell + \delta p W_{t+1} - p(1+g).$$

We can rewrite the above value as follows.

$$\begin{aligned}
&p-(1-p)\ell+\delta pW_{t+1}-p(1+g)\\
=&(1-\beta_t)-\beta_t\ell-\lambda+\delta(1-\beta_t)W_{t+1}-(1-\beta_t)(1+g)\\
&\quad+\lambda+\{p-(1-\beta_t)\}\{1+\ell+\delta W_{t+1}-(1+g)\}\\
=&W_t-(1-\beta_t)(1+g)+\lambda+\{p-(1-\beta_t)\}\{\delta W_{t+1}-(g-\ell)\}\\
=&\frac{\lambda}{\delta(1-\beta_{t-1})}+\lambda+\{p-(1-\beta_t)\}\{\delta W_{t+1}-(g-\ell)\}. \qquad (8)
\end{aligned}$$

The second equality follows from equation (4) in period t. The last equality is ensured by (7) in period $t-1$.

Using equation (7), we obtain the lower bound of $\delta W_{t+1}-(g-\ell)$ as follows.

$$\begin{aligned}
\delta W_{t+1}-(g-\ell)\ \geq&\delta(1-\beta_{t+1})(1+g)-(g-\ell)\\
\geq&\{g+(1+g)\varepsilon'\}\left(1-\frac{1+g+\ell}{g+\ell}\varepsilon'\right)-(g-\ell)\\
\geq&\frac{\ell}{2}. \qquad (9)
\end{aligned}$$

The second inequality follows from $\beta_t\leq\frac{1+g+\ell}{g+\ell}\varepsilon'$. The last inequality is ensured by $\varepsilon'\leq 2\overline{\varepsilon}$. The maximum value of p is $(1-\beta_{t-1})(1-\beta_t)$. Taking (9) into account, we show that (8) is negative as follows.

$$\begin{aligned}
&\frac{\lambda}{\delta(1-\beta_{t-1})}+\lambda-\{(1-\beta_t)-p\}\{\delta W_{t+1}-(g-\ell)\}\\
\leq&\frac{\lambda}{\delta(1-\beta_{t-1})}+\lambda-(1-\beta_t)\beta_{t-1}\frac{\ell}{2}\\
\leq&\frac{1+g}{g}\frac{1}{1-\frac{1+g+\ell}{g+\ell}\varepsilon'}\lambda+\lambda-\left(1-\frac{1+g+\ell}{g+\ell}\varepsilon'\right)\frac{1}{2}\frac{1+g-\ell}{g+\ell}\varepsilon'\frac{\ell}{2}\\
<&0.
\end{aligned}$$

The second inequality is ensured by $\delta\in[\underline{\delta},\overline{\delta}]$ and $\beta_t,\beta_{t-1}\in\left[\frac{1}{2}\frac{1+g-\ell}{g+\ell}\varepsilon',\frac{1+g+\ell}{g+\ell}\varepsilon'\right]$. Therefore, player i prefers D_i to C_i. Hence, it has been proved that it is optimal for player i to follow strategy σ^*. The strategy σ^* is a sequential equilibrium.

The payoff

Finally, we show that the sequential equilibrium payoff v_i^* is strictly greater than $1-\varepsilon$. Player i chooses $(D_i,0)$ in period 1 at the initial state. Therefore, the equilibrium payoff v_i^* is given by

$$v_i^*=(1-\delta)(1-\beta_1)(1+g)=\{1-(1+g)\varepsilon'\}\left(1-\frac{1+g+\ell}{g+\ell}\varepsilon'\right)>1-\varepsilon.$$

Therefore, Proposition 1 has been proved. □

B Proof of Proposition 3

Proof. Fix any $\varepsilon > 0$. We define $\overline{\varepsilon}$, $\underline{\delta}$, $\overline{\delta}$ and $\overline{\lambda}$ as follows:

$$\begin{aligned}
\overline{\varepsilon} \equiv & \frac{\ell^2}{54(1+g+\ell)^2}\frac{\varepsilon}{1+\varepsilon}, \\
\underline{\delta} \equiv & \frac{g}{1+g} + \overline{\varepsilon}, \\
\overline{\delta} \equiv & \frac{g}{1+g} + 2\overline{\varepsilon}, \\
\overline{\lambda} \equiv & \frac{1}{16}\frac{g}{1+g}\frac{1}{1+g+\ell}\overline{\varepsilon}^2.
\end{aligned}$$

Fix any $\delta \in \left[\underline{\delta}, \overline{\delta}\right]$ and $\lambda \in (0, \overline{\lambda})$. We will show a sequential equilibrium whose payoff vector (v_1^*, v_2^*) satisfies $v_1^* = 0$ and $v_2^* \geq \frac{1+g+\ell}{1+\ell} - \varepsilon$.

We define a grim trigger strategy $\tilde{\sigma}$. Strategy $\tilde{\sigma}$ is represented by an automaton that has four kinds of state: initial state $\tilde{\omega}_i^1$, cooperation state $(\tilde{\omega}_i^t)_{t=2}^{\infty}$, transition state ω_i^E, and defection state ω_i^D. Players use the sunspot only at the initial state.

At initial state $\tilde{\omega}_1^1$, player 1 chooses C_1 with probability $1 - \beta_{1,1}$, and chooses D_1 with probability $\beta_{1,1}$. Player 1 does not observe player 2 irrespective of his action. The transition state depends on a realized sunspot. If the realized sunspot is greater than $\hat{x}$, the state remains the same. If the realized sunspot is not greater than $\hat{x}$ and player 1 chooses C_1, then the state in the next period moves to cooperation state $\tilde{\omega}_1^2$. If the realized sunspot is not greater than $\hat{x}$ where player 1 chooses D_1, then the state in the next period moves to defection state ω_1^D.

At initial state $\tilde{\omega}_2^1$, player 2 chooses D_2. Player 2's observational decision depends on the sunspot. If the realized sunspot is greater than $\hat{x}$, player 2 does not observe his opponent. If the realized sunspot is not greater than $\hat{x}$, player 2 randomizes his observational decision. Irrespective of his action, player 2 observes player 1 with probability $1 - \beta_{2,2}$ and does not observe him with probability $\beta_{2,2}$. The transition state also depends on the realized sunspot. If the realized sunspot is greater than $\hat{x}$, the state remains the same. Suppose that the realized sunspot is not greater than $\hat{x}$. If player 2 observes C_1, then the state in the next period moves to cooperation state $\tilde{\omega}_2^2$. If player 2 observes D_1, then the state in the next period is defection state ω_2^D. If player 2 does not observe his opponent in period 1, then the state in the next period is transition state ω_2^E.

At cooperation state $\tilde{\omega}_1^2$, player 1 chooses action C_1 with probability $1 - \beta_{1,2}$. When player 1 chooses action C_1, he observes the opponent with probability $1 - \beta_{1,3}$. When player 1 chooses action D_1, he does not observe. If player 1 chooses action D_1, he does not observe his opponent. If player 1 chooses C_1 and observes C_2, then the state in the next period is cooperation state $\tilde{\omega}_1^3$. If player 1 chooses D_1 or observes D_2, then the state in the next period is defection state ω_1^D. If player 1 chooses C_1 but does not observe, then the state in the next period is transition state ω_1^E.

At cooperation state $\tilde{\omega}_1^t (t \geq 3)$, player 1 chooses action C_1. Player 1 observes his opponent with probability $1-\beta_{1,t+1}$. If player 1 chooses C_1 and observes C_2, then the state in the next period is cooperation state $\tilde{\omega}_1^{t+1}$. If player 1 chooses D_1 or observes D_2, then the state in the next period is defection state ω_1^D. If player 1 chooses C_1 but does not observe, then the state in the next period is transition state ω_1^E.

At cooperation state $(\tilde{\omega}_2^t)_{t=2}^{\infty}$, player 2 chooses action C_2. He observes player 1 with probability $1 - \beta_{2,t+1}$. If player 2 chooses C_2 and observes C_1, then the state in the next period is cooperation state $\tilde{\omega}_2^{t+1}$. If player 2 chooses D_2 or observes D_1, then the state in the next period is defection state ω_2^D. If player 2 chooses C_2 but does not observe, then the state in the next period is transition state ω_2^E.

The output function and transition function at the transition state and the defection state is defined in the same manner as in the previous section. At transition state ω_i^E in period t, each player i chooses D_i and does not observe irrespective of his action. The state moves to defection state ω_i^D in period $t+1$ when $a_i^t = D_i$ or $o_i^t = D_j$ is realized. If player i chooses $(C_i, 0)$, the state remains the same. When player i chooses C_i and observes C_j in period t, the state in period $t+1$ moves to cooperation state $\tilde{\omega}_i^{t+1}$. At defection state ω_i^D, the state remains the same; defection state ω_i^D, irrespective of the action–observation pair.

The belief ψ_i^* for player i is determined in the same manner in Section 5. We consider a tremble that puts far less weight on the deviations with respect to observation at any history h_t^i than those with respect to action for any i and any $t \in \mathbb{N}$. The above tremble induces the unique belief ψ_j^* for player j for each j. We denote by ψ^* the system of beliefs (ψ_1^*, ψ_2^*). The belief ψ^* has a similar property to the one in Section 5. That is, given ϕ^*, player i is certain that the state of his opponent is a defection state when player i chose D_i or observed D_j in the past.

We define $(\beta_{1,t})_{t=1}^{\infty}$ and $(\beta_{2,t})_{t=2}^{\infty}$. First, let us define $\beta_{1,1}$ and $\beta_{1,2}$. We define $\varepsilon' \equiv (1+g)\delta - g$. It is obvious that $\varepsilon' \in [\bar{\varepsilon}, 2\bar{\varepsilon}]$ We set $\beta_{1,1} = \frac{1+g+\ell}{g+\ell}\varepsilon'$. We define $\beta_{1,2}$ as follows.

$$\beta_{1,2} = \frac{(1-\beta_{1,1})\left\{\delta(1+g) - g\right\} - \beta_{1,1}\ell}{\delta(1-\beta_{1,1})(1+g)}.$$

Let $W_{i,t} (t \geq 2)$ be the continuation payoff from cooperation state ω_i^t for player i. At any cooperation state $\omega_2^{t+1} (t \in \mathbb{N})$, player 2 believes that the state of his opponent is cooperation state ω_1^{t+1} with probability $1 - \beta_{1,t+1}$, and with the remaining probability $\beta_{1,t+1}$, the state is either ω_1^E or ω_1^D. Therefore, $W_{2,t+1}$ is given by

$$W_{2,t+1} = (1-\beta_{1,t+1}) - \beta_{1,t+1}\ell + \delta(1-\beta_{1,t+1})(1-\beta_{1,t+2})(1+g).$$

At the initial state and any cooperation state, player 2 is indifferent between

$m_2 = 1$ and $m_2 = 0$. Therefore, for any $t \in \mathbb{N}$, $\beta_{1,t+2}$ is given by

$$\frac{\lambda}{\delta(1-\beta_{1,1})} = W_{2,2} - (1-\beta_{1,2})(1+g).$$

Note that $W_{2,2}$ is a function of $\beta_{1,3}$.

At any cooperation state, player 2 is indifferent between $m_2 = 1$ and $m_2 = 0$. Therefore, for any $t \in \mathbb{N}$, $\beta_{1,t+2}$ is given by

$$\frac{\lambda}{\delta(1-\beta_{1,t})} = W_{2,t+1} - (1-\beta_{1,t+1})(1+g). \tag{10}$$

Note that $W_{2,t+1}$ is a function of $\beta_{1,t+2}$.

Next, we define $(\beta_{2,t})_{t=2}^{\infty}$. We define $\beta_{2,2}$ so that player 1 is indifferent between choosing $(C_1, 0)$ and choosing $(D_1, 0)$ at the initial state. That is, $\beta_{2,2}$ is given by the equation below.

$$-\ell + \hat{x}\delta(1-\beta_{2,2})(1+g) = 0.$$

Player 1 randomizes $(C_1, 0)$ and $(D_1, 0)$ at cooperation state $\tilde{\omega}_1^2$. Hence, $\beta_{2,3}$ is given by the following equation.

$$(1-\beta_{2,2}) - \beta_{2,2}\ell + \delta(1-\beta_{2,2})(1-\beta_{2,3})(1+g) = (1-\beta_{2,2})(1+g).$$

In cooperation state $\tilde{\omega}_1^t$ $(t \geq 2)$, player 1 believes that the state of his opponent is a cooperation state with probability $1-\beta_{2,t}$. Therefore, $W_{1,t}(t \geq 2)$ is given by

$$W_{1,t} = (1-\beta_{2,t}) - \beta_{2,t}\ell + \delta(1-\beta_{2,t})(1-\beta_{2,t+1})(1+g).$$

Furthermore, player 1 randomizes $(C_1, 1)$ and $(C_1, 0)$ at cooperation state $\tilde{\omega}_1^2$. At cooperation state $\tilde{\omega}_1^2$, player 1 believes that the state of player 2 is $\tilde{\omega}_2^2$ with probability $1-\beta_{2,2}$. Therefore, $\beta_{2,4}$ is determined as the solution of the following equation.

$$\frac{\lambda}{\delta(1-\beta_{2,2})} = W_{1,3} - (1-\beta_{2,3})(1+g).$$

Note that $W_{1,3}$ is a function of $\beta_{2,4}$.

In addition, player 1 randomizes $(C_1, 1)$ and $(C_1, 0)$ at cooperation state $\tilde{\omega}_1^t$ $(t \geq 3)$. At cooperation state $\tilde{\omega}_1^t$ $(t \geq 3)$, player 1 believes that the state of player 2 is $\tilde{\omega}_2^t$ with probability $1-\beta_{2,t}$. We choose $\beta_{2,t+1}$ as the solution of the equation below so that player 1 is indifferent between choosing $(C_1, 1)$ and $(C_1, 0)$.

$$\frac{\lambda}{\delta(1-\beta_{2,t})} = W_{1,t+1} - (1-\beta_{2,t+1})(1+g). \tag{11}$$

Note that $W_{1,t+1}$ is a function of $\beta_{2,t+2}$.

Taking into account the definition of δ, $(\beta_{1,t})_{t=2}^{\infty}$ and $(\beta_{2,t})_{t=2}^{\infty}$ are chosen as follows.

$$\beta_{1,1} = \frac{1+g+\ell}{g+\ell}\varepsilon'$$
$$\beta_{1,2} = \frac{(1-\beta_{1,1})\{\delta(1+g)-g\} - \beta_{1,1}\ell}{\delta(1-\beta_{1,1})(1+g)}$$
$$\beta_{1,t+2} = \frac{(1-\beta_{1,t+1})\{\delta(1+g)-g\} - \beta_{1,t+1}\ell - \frac{\lambda}{\delta(1-\beta_{1,t})}}{\delta(1+g)(1-\beta_{1,t+1})}, \quad \forall\, t \geq 1.$$
$$\beta_{2,2} = \frac{\hat{x}\delta(1+g)-\ell}{\hat{x}\delta(1+g)}$$
$$\beta_{2,3} = \frac{(1-\beta_{2,2})\{\delta(1+g)-g\} - \beta_{2,2}\ell}{\delta(1+g)(1-\beta_{2,2})}$$
$$\beta_{2,t+2} = \frac{(1-\beta_{2,t+1})\{\delta(1+g)-g\} - \beta_{2,t+1}\ell - \frac{\lambda}{\delta(1-\beta_{2,t})}}{\delta(1+g)(1-\beta_{2,t+1})}, \quad \forall\, t \geq 2.$$

Therefore, the sunspot $\hat{x}$ has an effect on $\beta_{2,2}$ only.

Finally, we choose $\hat{x}$. We define $\hat{x}$ as the solution below.

$$\frac{\hat{x}\delta(1+g)-\ell}{\hat{x}\delta(1+g)} = \frac{1+g+\ell}{g+\ell}\varepsilon'.$$

When $\hat{x} = \frac{\ell}{g}$, the left-hand side is greater than the right-hand side.

$$\frac{\frac{\ell}{g}(1+g)\varepsilon'}{\frac{\ell}{g}\delta(1+g)} = \frac{1}{\delta}\varepsilon' > \frac{1+g+\ell}{g+\ell}\varepsilon'.$$

Furthermore, if $\hat{x} = \frac{\ell}{\delta(1+g)}$, then the left-hand side is smaller than the right-hand side. Therefore, $\hat{x} \in \left(\frac{\ell}{\delta(1+g)}, \frac{\ell}{g}\right)$ is well defined.

The above definition ensures that $\beta_{1,t} = \beta_{2,t+1}$ for any $t \in \mathbb{N}$. In addition, $\beta_{1,t}$ is equal to β_t in Section 5. Therefore, by Corollary 1.2, we have

$$\frac{1}{2}\frac{1+g-\ell}{g+\ell}\varepsilon' < \beta_{1,t} < \frac{1+g+\ell}{g+\ell}\varepsilon', \qquad \forall t \in \mathbb{N}, \text{ and}$$
$$\frac{1}{2}\frac{1+g-\ell}{g+\ell}\varepsilon' < \beta_{2,t+1} < \frac{1+g+\ell}{g+\ell}\varepsilon', \qquad \forall t \in \mathbb{N}.$$

As the same with the proof of Proposition 1, we show the sequential rationality and show that the equilibrium payoff vector (v_1^*, v_2^*) satisfies $v_1^* = 0$ and $v_2^* \geq \frac{1+g+\ell}{1+\ell} - \varepsilon$.

Sequential rationality at the defection state

Let us confine our attention to show sequential rationality. At the defection state, player i is certain that the state of his opponent is a defection state, and the opponent chooses $(D_j, 0)$ with certainty from the current period onwards. Player i has no incentive to choose C_i or $m_i = 1$. Therefore, it is optimal for player i to choose $(D_i, 0)$.

Sequential rationality at the initial state and the cooperation state

Let us consider cooperation state $\tilde{\omega}_i^t (t \geq 2)$. Once player i chooses D_i, the strategy σ^* prescribes D_i every period irrespective of his observation. Therefore, at any cooperation state, each player i has no incentive to choose $(D_i, 1)$.

First, let us consider player 1's sequential rationality at initial state $\tilde{\omega}_1^1$. The definition of $\beta_{2,2}$ ensures that player 1 is indifferent between $(C_1, 0)$ and $(D_1, 0)$. It is obvious that player 1 has no incentive to observe player 2 because player 2 chooses action D_2 with certainty.

Next, let us consider the decision of player 1 at cooperation states. At cooperation state $\tilde{\omega}_1^2$, the stage behaviors $(C_1, 1)$, $(C_1, 0)$ and $(D_1, 0)$ are indifferent by the definitions of $\beta_{2,3}$ and $\beta_{2,4}$. At cooperation state $\tilde{\omega}_1^{t+2}$ $(t \geq 1)$, the definition of $\beta_{2,t+4}$ ensures that $(C_1, 1)$ and $(C_1, 0)$ are indifferent. In addition, the equation (11) in period $t+1$ implies that the payoff $W_{1,t+2}$ for choosing action C_1 is greater than the payoff $(1-\beta_{2,t+2})(1+g)$ when he chooses action D_2. It is optimal for player 1 to follow the strategy σ^* at cooperation state $(\tilde{\omega}_{1,t})_{t=2}^{\infty}$.

Lastly, let us consider player 2's choice at initial state $\tilde{\omega}_2^1$. By the definition of $\beta_{1,3}$, player 2 is indifferent between choosing $(C_2, 1)$ and choosing $(C_2, 0)$. Player 2 does not prefer action D_2 because player 1 never observes him. Next, let us confine our attention to player 2's choice at cooperation state $\tilde{\omega}_2^t$ $(t \geq 2)$. By the definition of $\beta_{1,t+2}$, player 2 is indifferent between choosing $(C_2, 1)$ and choosing $(C_2, 0)$. If player 2 chooses $(D_2, 0)$, his payoff is $(1-\beta_{1,t})(1+g)$. The inequality (10) in period $t-1$ ensures that the payoff $W_{2,t}$ for choosing C_1 is greater than $(1-\beta_{1,t})(1+g)$. That is, action D_2 is suboptimal.

Thus, it is optimal for both players to follow strategy σ^* in a cooperation state.

Sequential rationality in the transition state

We consider sequential rationality at any period t (≥ 2) associated with a transition state.

First, let us consider the transition state for player 1 in period t $(t \geq 3)$. Let p be the probability with which player 1 believes that the state of his opponent is a cooperation state. Therefore, the upper bound of the payoff when player 1 chooses action C_1 in period t is given by

$$p - (1-p)\ell + \delta p W_{1,t+1}.$$

Furthermore, the payoff for $(D_1, 0)$ is bounded above by $p(1+g)$. Therefore, $(D_1, 0)$ is profitable if the following value is negative.

$$p - (1-p)\ell + \delta p W_{1,t+1} - p(1+g).$$

Using (11), we can rewrite the above value as follows.

$$\begin{aligned}
&p-(1-p)\ell+\delta p W_{1,t+1}-p(1+g)\\
=&(1-\beta_{2,t})-\beta_{2,t}\ell-\lambda+\delta(1-\beta_{2,t})W_{1,t+1}-(1-\beta_{2,t})(1+g)\\
&+\lambda+\{p-(1-\beta_{2,t})\}\{1+\ell+\delta W_{1,t+1}-(1+g)\}\\
=&\frac{\lambda}{\delta(1-\beta_{2,t-1})}+\lambda-\{(1-\beta_{2,t})-p\}\{\delta W_{1,t+1}-(g-\ell)\}. \qquad (12)
\end{aligned}$$

The second equality follows from equation (11) for $t-1$.

Furthermore, the payoff $\delta W_{1,t+1}$ is greater than the payoff for choosing $(D_1, 0)$. Therefore, the payoff $\delta W_{1,t+1}$ is bounded below by

$$\begin{aligned}
\delta W_{1,t+1}-(g-\ell)\geq&\delta(1-\beta_{2,t+1})(1+g)-(g-\ell)\\
\geq&\{g+(1+g)\varepsilon'\}\left(1-\frac{1+g+\ell}{g+\ell}\varepsilon'\right)-(g-\ell)\\
\geq&\frac{\ell}{2}. \qquad (13)
\end{aligned}$$

The second inequality follows from $\delta=\frac{g}{1+g}+\varepsilon'$ and $\beta_{2,t+1}\leq\frac{1+g+\ell}{g+\ell}\varepsilon'$.

The maximum value of p in period t ($t\geq 3$) is $(1-\beta_{2,t-1})(1-\beta_{2,t})$. Taking (13) into account, the value of (12) has the following upper bound.

$$\begin{aligned}
&\frac{\lambda}{\delta(1-\beta_{2,t-1})}+\lambda-\{(1-\beta_{2,t})-p\}\delta W_{1,t+1}\\
<&\frac{1+g}{g}\frac{\lambda}{1-\beta_{2,t-1}}+\lambda-(1-\beta_{2,t-1})\beta_{2,t}\frac{\ell}{2}\\
<&\frac{1+g}{g}\frac{\lambda}{1-\frac{1+g+\ell}{g+\ell}\varepsilon'}+\lambda-\left(1-\frac{1+g+\ell}{g+\ell}\varepsilon'\right)\frac{1}{2}\frac{1+g-\ell}{1+g+\ell}\varepsilon'\frac{\ell}{2}\\
<&0.
\end{aligned}$$

The second inequality follows from $\frac{1}{2}\frac{1+g-\ell}{1+g+\ell}\varepsilon'<\beta_{2,t-1},\beta_{2,t}<\frac{1+g+\ell}{g+\ell}\varepsilon'$. Therefore, choosing $(D_1, 0)$ is optimal at transition state ω_1^E.

Next, let us consider the transition state for player 2 in period 2. Then, player 2 believes that the state of his opponent is cooperation state $\tilde{\omega}_1^2$ with probability $1-\beta_{1,1}$. If player 2 chooses C_2, the continuation payoff is bounded above by

$$(1-\beta_{1,1})W_{2,2}-\beta_{1,1}\ell.$$

However, the payoff of choosing $(D_2, 0)$ is given by $(1-\beta_{1,1})(1-\beta_{1,2})(1+g)$. Therefore, it is optimal for player 2 to choose $(D_2, 0)$ if the following value is negative.

$$(1-\beta_{1,1})W_{2,2}-\beta_{1,1}\ell-(1-\beta_{1,1})(1-\beta_{1,2})(1+g).$$

Or, equivalently

$$\begin{aligned}
&(1-\beta_{1,1})\left\{W_{2,2}-(1-\beta_{1,2})(1+g)\right\}-\beta_{1,1}\ell \\
=&(1-\beta_{1,1})\frac{\lambda}{\delta(1-\beta_{1,1})}-\beta_{1,1}\ell \\
=&\frac{\lambda}{\delta}-\beta_{1,1}\ell<0.
\end{aligned}$$

Therefore, it is optimal for player 2 to choose $(D_2, 0)$.

Finally, let us consider the transition state for player 2 in period t ($t \geq 3$). Let us denote by p the probability with which player 2 believes that the state of his opponent is a cooperation state. Then, the upper bound of the payoff when player 2 chooses action C_2 in period t is given by

$$p-(1-p)\ell+\delta p W_{2,t+1}.$$

The payoff for $(D_2, 0)$ is given by $p(1+g)$. Therefore, $(D_2, 0)$ is profitable if the following value is negative.

$$p-(1-p)\ell+\delta p W_{2,t+1}-p(1+g).$$

We can rewrite the above value as follows.

$$\begin{aligned}
&p-(1-p)\ell+\delta p W_{2,t+1}-p(1+g) \\
=&(1-\beta_{1,t})-\beta_{1,t}\ell-\lambda+\delta(1-\beta_{1,t})W_{2,t+1}-(1-\beta_{1,t})(1+g) \\
&\quad+\lambda+\left\{p-(1-\beta_{1,t})\right\}\left\{1+\ell+\delta W_{2,t+1}-(1+g)\right\} \\
=&W_{2,t}-(1-\beta_{1,t})(1+g)+\lambda+\left\{p-(1-\beta_{1,t})\right\}\left\{\delta W_{2,t+1}-(g-\ell)\right\} \\
=&\frac{\lambda}{\delta(1-\beta_{1,t-1})}+\lambda-\left\{(1-\beta_{1,t})-p\right\}\left\{\delta W_{2,t+1}-(g-\ell)\right\}. \qquad (14)
\end{aligned}$$

The third equality follows from equation (10) for $t-1$.

Furthermore, $\delta W_{2,t+1}$ is bounded below by

$$\begin{aligned}
\delta W_{2,t+1}-(g-\ell) \geq&\delta(1-\beta_{1,t+1})(1+g)-(g-\ell) \\
\geq&\left\{g+(1+g)\varepsilon'\right\}\left(1-\frac{1+g+\ell}{g+\ell}\varepsilon'\right)-(g-\ell) \\
\geq&\frac{\ell}{2}.
\end{aligned}$$

The second inequality follows from $\delta = \frac{g}{1+g} + \varepsilon'$ and $\beta_{1,t+1} \leq \frac{1+g+\ell}{g+\ell}\varepsilon'$.

The maximum value of p in period t is $(1-\beta_{1,t-1})(1-\beta_{1,t})$. Taking (13) into account, we can show that (14) is negative as follows.

$$\begin{aligned}
&\frac{\lambda}{\delta(1-\beta_{1,t-1})}+\lambda-\{(1-\beta_{1,t})-p\}\,\delta W_{2,t+1}\\
\leq&\frac{\lambda}{\delta(1-\beta_{1,t-1})}+\lambda-(1-\beta_{1,t})\beta_{1,t-1}\frac{\ell}{2}\\
\leq&\frac{1+g}{g}\frac{1}{1-\frac{1+2g}{2g}\varepsilon'}\lambda+\lambda-\left(1-\frac{1+g+\ell}{g+\ell}\varepsilon'\right)\frac{1}{2}\frac{1+g-\ell}{1+g+\ell}\varepsilon'\frac{\ell}{2}\\
<&0.
\end{aligned}$$

The second inequality is ensured by $\beta_{1,t},\beta_{1,t-1}\in\left(\frac{1}{2}\frac{1+g-\ell}{1+g+\ell}\varepsilon',\,\frac{1+g+\ell}{g+\ell}\varepsilon'\right)$. Therefore, player 2 prefers D_2 to C_2 at the transition state.

Hence, it has been proved that it is optimal for both players to follow strategy σ^*. The strategy σ^* is a sequential equilibrium.

The payoff

Finally, let us consider the equilibrium payoff. The equilibrium payoff for player 1 is 0 because player 1 weakly prefers $(D_1,0)$ in period 1.

Similarly, player 2 weakly prefers $(D_2,0)$ in period 2. Thus, his equilibrium payoff v_2^* is given by

$$\begin{aligned}
v_2^*=&(1-\delta)(1-\beta_{1,1})\{(1+g)+\hat{x}\delta(1-\beta_{1,2})(1+g)\}+(1-\hat{x})\delta v_2^*\\
=&\frac{(1-\delta)(1-\beta_{1,1})\{(1+g)+\hat{x}\delta(1-\beta_{1,2})(1+g)\}}{1-(1-\hat{x})\delta}\\
=&\frac{(1-\beta_{1,1})\{1+\hat{x}\delta(1-\beta_{1,2})\}}{1+\hat{x}\frac{\delta}{1-\delta}}(1+g).
\end{aligned}$$

Taking $\hat{x}\in\left(\frac{\ell}{\delta(1+g)},\frac{\ell}{g}\right)$ into consideration, we obtain a lower bound of v_2^* below.

$$\begin{aligned}
v_2^*>&\frac{(1-\beta_{1,1})\left\{1+\frac{\ell}{1+g}(1-\beta_{1,2})\right\}}{1+\frac{\ell}{g}\frac{\delta}{1-\delta}}(1+g)\\
>&\frac{\left(1-\frac{1+g+\ell}{g+\ell}\varepsilon'\right)\left(1+g+\ell-\frac{1+g+\ell}{g+\ell}\varepsilon'\ell\right)}{1+\frac{\ell}{g}\frac{g+\varepsilon'}{1-(1+g)\varepsilon'}}>\frac{1+g+\ell}{1+\ell}-\varepsilon.
\end{aligned}$$

The second inequality follows from the upper bound of $\beta_{1,1}$ and $\beta_{1,2}$. Therefore, Proposition 3 has been proved. □

References

[1] Ben-Porath, Elchanan and Michael Kahneman (2003) "Communication in repeated games with costly monitoring," *Games and Economic Be-*

havior, Vol. 44, No. 2, pp. 227–250, DOI: https://doi.org/10.1016/S0899-8256(03)00022-8.

[2] Bhaskar, V. and Eric van Damme (2002) "Moral hazard and private monitoring," *Discussion Paper*, Vol. 102, No. 1, pp. 16–39, DOI: https://doi.org/10.1006/jeth.2001.2861.

[3] Bhaskar, V. and Ichiro Obara (2002) "Belief-Based Equilibria in the Repeated Prisoners' Dilemma with Private Monitoring," *Journal of Economic Theory*, Vol. 102, No. 1, pp. 40–69, DOI: https://doi.org/10.1006/jeth.2001.2878.

[4] Chen, Bo (2010) "A belief-based approach to the repeated prisoners' dilemma with asymmetric private monitoring," *Journal of Economic Theory*, Vol. 145, No. 1, pp. 402–420, DOI: https://doi.org/10.1016/J.JET.2009.05.006.

[5] Ely, Jeffrey C. and Juuso Välimäki (2002) "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory*, Vol. 102, No. 1, pp. 84–105, DOI: https://doi.org/10.1006/jeth.2000.2774.

[6] Ely, Jeffrey C., Johannes Horner, and Wojciech Olszewski (2005) "Belief-Free Equilibria in Repeated Games," *Econometrica*, Vol. 73, No. 2, pp. 377–415, DOI: https://doi.org/10.1111/j.1468-0262.2005.00583.x.

[7] Flesch, János and Andrés Perea (2009) "Repeated Games with Voluntary Information Purchase," *Games and Economic Behavior*, Vol. 66, No. 1, pp. 126–145, DOI: https://doi.org/10.1016/j.geb.2008.04.015.

[8] Hörner, Johannes and Wojciech Olszewski (2006) "The Folk Theorem for Games with Private Almost-Perfect Monitoring," *Econometrica*, Vol. 74, No. 6, pp. 1499–1544, DOI: https://doi.org/10.1111/j.1468-0262.2006.00717.x.

[9] ——— (2009) "How Robust Is the Folk Theorem?," *Quarterly Journal of Economics*, Vol. 124, No. 4, pp. 1773–1814, DOI: https://doi.org/10.1162/qjec.2009.124.4.1773.

[10] Kandori, Michihiro and Ichiro Obara (2004) "Endogenous Monitoring," *UCLA Economics Online Papers*, Vol. 398, URL: http://www.econ.ucla.edu/people/papers/Obara/Obara398.pdf.

[11] Lehrer, Ehud and Eilon Solan (2018) "High Frequency Repeated Games with Costly," *Theoretical Economics*, Vol. 13, No. 1, pp. 87–113, DOI: https://doi.org/10.3982/TE2627.

[12] Mailath, George J. and Stephen Morris (2002) "Repeated Games with Almost-Public Monitoring," *Journal of Economic Theory*, Vol. 102, No. 1, pp. 189–228, DOI: https://doi.org/10.1006/jeth.2001.2869.

[13] Mailath, George J and Stephen Morris (2006) "Coordination failure in repeated games with almost-public monitoring," *Theoretical Economics*, Vol. 1, pp. 311–340, URL: https://econtheory.org/ojs/index.php/te/article/viewFile/20060311/791/30.

[14] Mailath, George J. and Wojciech Olszewski (2011) "Folk theorems with bounded recall under (almost) perfect monitoring," *Games and Economic Behavior*, Vol. 71, No. 1, pp. 174–192, DOI: https://doi.org/10.1016/j.geb.2010.11.002.

[15] Miyagawa, Eiichi, Yasuyuki Miyahara, and Tadashi Sekiguchi (2003) "Repeated Games with Observation Costs," *Columbia University Academic Commons*, No. 203-14, DOI: https://doi.org/https://doi.org/10.7916/D8VX0TRW.

[16] ——— (2008) "The folk theorem for repeated games with observation costs," *Journal of Economic Theory*, Vol. 139, No. 1, pp. 192–221, DOI: https://doi.org/10.1016/j.jet.2007.04.001.

[17] Piccione, Michele (2002) "The Repeated Prisoner's Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory*, Vol. 102, No. 1, pp. 70–83, DOI: https://doi.org/10.1006/JETH.2001.2819.

[18] Sekiguchi, Tadashi (1997) "Efficiency in Repeated Prisoner's Dilemma with Private Monitoring," *Journal of Economic Theory*, Vol. 76, No. 2, pp. 345–361, DOI: https://doi.org/10.1006/jeth.1997.2313.

[19] Sugaya, Takuo (2011) "Folk Theorem in Repeated Games with Private Monitoring," *Economic Theory Center Working Paper*, No. 011, DOI: https://doi.org/10.2139/ssrn.1789775.

[20] Yamamoto, Yuichi (2007) "Efficiency results in N player games with imperfect private monitoring," *Journal of Economic Theory*, Vol. 135, No. 1, pp. 382–413, DOI: https://doi.org/10.1016/J.JET.2006.05.003.

[21] ——— (2009) "A limit characterization of belief-free equilibrium payoffs in repeated games," *Journal of Economic Theory*, Vol. 144, No. 2, pp. 802–824, DOI: https://doi.org/10.1016/J.JET.2008.07.005.

[22] ——— (2012) "Characterizing belief-free review-strategy equilibrium payoffs under conditional independence," *Journal of Economic Theory*, Vol. 147, No. 5, pp. 1998–2027, DOI: https://doi.org/10.1016/J.JET.2012.05.016.