# The Folk Theorem for Repeated Games with Observation Costs

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

October 28, 2007

## Abstract

The folk theorem literature has been relaxing the assumption on how much players know about each other's past action. Here we consider a general model where players can "buy" precise information. Every period, each player decides whether to pay a cost to accurately observe the actions chosen by other players in the previous period. When a player does not pay the cost, he obtains only imperfect private signals. Observational decisions are unobservable to others. Known strategies such as trigger strategies do not work since they fail to motivate players to pay for information. This paper shows that the folk theorem holds for any level of observation costs. Unlike existing folk theorems with private monitoring, ours imposes virtually no restriction on the nature of costless imperfect signals. The theorem does not use explicit or costless communication, thereby having implications on antitrust laws that rely on evidence of explicit communication. The main message is that accurate observation alone, however costly, enables efficient cooperation in general repeated games.

*JEL Classification*: C72, C73, D43, D82, K21, L13, L40.
*Keywords*: Cooperation, collusion, private information, information acquisition.

---

[*]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`.

[§]Corresponding author: fax +81 75 753 7118

# 1 Introduction

The folk theorem literature has been weakening the assumption on what players know about each other's past actions. A seminal version of the folk theorem by Fudenberg and Maskin [12] assumes *perfect monitoring*: players receive accurate information about each other's past actions. The theorem has been extended rather generally to the case of *imperfect public monitoring*, where information is imperfect but identical across players [1, 10, 11]. An example is where oligopolistic firms observe only market prices, which are public information but do not fully reveal the actions chosen by other firms [13]. More recent studies consider the case where information is not only imperfect but private. An example is where each firm observes only the sales of its own products, which are private information and are correlated only imperfectly with other firms' actions [29]. For this *imperfect private monitoring* case, the folk theorem is known only under restrictive assumptions. For example, many of the existing results pertain to the limit case where information is almost perfect [e.g., 7, 14, 28].

The basic idea of the current paper is that the quality of information usually depends on the cost spent for acquiring information. By spending more time and resources, one can usually obtain better information. To capture the feature, this paper considers an extreme case where one can obtain precise information by paying a sufficiently large cost. For community enforcement, for instance, the assumption means that you can obtain accurate information about your neighbors if you spend enough time for collecting information. For price-fixing collusion, it means that a firm can observe the prices charged by other firms if the firm spends enough resources for monitoring. Under the assumption, this paper clarifies conditions under which the folk theorem extends.

We distinguish observational decisions and other stage-game actions by assuming that observational decisions cannot be observed by other players at any cost. The assumption makes it difficult to motivate players to pay a cost for information. A player is willing to pay for information only if the cost does not exceed the expected benefit. Since a player's decision on information acquisition is unobservable and thereby has no effect on other players' future actions, the only benefit from paying the observation cost is that it enables the player to accurately predict the other players' future actions. An immediate but important implication is that if a player is expected to choose a certain action with probability one, other players have no incentive to pay a cost to observe the player.

Because of this feature, it is not trivial to extend known strategies to the present class of repeated games. As an illustration, consider a grim-trigger strategy, under which players start with cooperation and any deviation triggers a repetition of a static Nash equilibrium. Under the strategy, since players are expected to cooperate in the first period, the previous argument implies that players have no incentive to pay the observation cost in the first period. As a result, deviations in the first period are not observed by anyone. For the strategy to deter deviations, therefore,

2

punishments must be triggered on the basis of costless information (e.g., realized period-payoffs). However, if costless information is private, which is a case we allow, players cannot coordinate when to start punishments.[1] It is difficult to construct an equilibrium along these lines in general environments, as we know from the literature of imperfect private monitoring.

To construct equilibria in which players pay for information, we use strategies where players randomize. However, randomization alone does not solve the problem. As an illustration, consider a repeated prisoners' dilemma and suppose that an equilibrium strategy is such that each player defects with a positive probability at any history (as in Ely and Välimäki [7]). In such an equilibrium, even though players may be randomizing, no costly observation takes place. Indeed, nothing is gained by knowing the actions chosen by other players since defection is assigned a positive probability and therefore is an optimal action at any history. This argument says that, to motivate players to pay for information, the equilibrium strategy has to be such that one's optimal action varies with the other players' past actions. For the prisoners' dilemma, it means that there need to be a history at which a player's unique optimal action is to cooperate with probability one. But, if a player cooperates with probability one, he will not be observed by anyone and the problem discussed above persists.

The present paper shows, however, that the folk theorem holds generally. Here are some of the features. First, the theorem holds for any *n*-player finite stage game that satisfies the standard full dimensionality condition. Second, the theorem holds for *any* level of observation costs. That is, payoff vectors arbitrarily close to the Pareto frontier can be supported even if observation costs are extremely large. This is the case since there exists a rich class of equilibria in which players observe each other only periodically and the expected per-period observation cost is small. The level of observation cost affects only the threshold level of discount factor. Third, the theorem holds under a minimal assumption on the costless information—the information that players obtain when they do not pay the observation cost. All we need is that there is no one whose action has no influence over the other players' costless information. That is, it is sufficient that the action of each player has a non-zero influence on the probability distribution of at least another player's costless information under at least one action profile. Since the nature of influence is immaterial and costless information includes realized period-payoffs, the assumption would be satisfied in virtually all economic applications.

As mentioned above, the folk theorem for repeated games with private monitoring is known only under special cases. Early studies assume the prisoners' dilemma and almost perfect monitoring [5, 7, 25, 28]. Recent studies obtain more general results but are still limited to the prisoners' dilemma and its variants [24, 30], the case of almost perfect monitoring [14, 22], or a subclass of equilibria

---

[1]A player cannot deduce the action profile from his period-payoff because they are correlated only imperfectly.

that is not large enough to generate a general folk theorem [8]. The difficulty has led Matsushima [23] and other researchers to introduce explicit communication in the model so that players can exchange their private information [2, 3, 4, 6, 16]. Communication makes the analysis tractable because players can coordinate their continuation actions.

In contrast with those results, our folk theorem holds in general environments with private monitoring without explicit communication. A sufficient condition for the folk theorem is that obtaining precise information about other players' actions is possible, i.e., the cost of perfect monitoring is finite. By modeling players' monitoring decisions concretely, we find that the folk theorem holds if perfect monitoring is an *option* for each player, and it does not matter how costly the option is. The availability of the option enables players to coordinate their continuation actions when they need, and its cost can be controlled since the option does not have to be executed every period.

A key idea is that costly perfect monitoring enables players to communicate *implicitly*. Players can announce messages by their stage-game actions. For example, they may use last digits of prices for communication. Players obtain other players' messages via costly observation. Implicit communication of this kind differs from explicit communication considered in the literature. Unlike "cheap talk," implicit communication is costly for both senders and receivers. In particular, since acquiring information is costly, implicit communication works only if players are motivated to obtain the messages of other players. Moreover, if a player deviates and chooses not to acquire information, he becomes uncertain about other players' continuation actions but other players do not notice it. Therefore the continuation play after a deviation is not an equilibrium, which makes it difficult to apply the recursive method of dynamic programming. It should be also noted that in the case of implicit communication via stage-game actions, the message space is constrained by the action set, which may contain only two elements.

In the context of private monitoring, the series of work by Lehrer [18, 20, 21] are somewhat close to the present paper (see also Flesch and Monsuwé [9] for an extension). Lehrer considers two-player repeated games with no discounting where costless private signals are imperfect but deterministic. Lehrer shows that an action profile can be supported even if it admits short-run better replies that do not affect the opponent's signal and hence are undetectable, if the better replies make the player's own signal less informative. The result is relevant since not paying for information is precisely a short-run better reply that is undetectable and less informative. Lehrer's result therefore says that costly observation, even if unobservable, can be supported in equilibrium. However, his result relies on the assumption of no discounting and is limited to the two-player case.

The present paper also contributes to the literature of *public* monitoring since the assumption on costless signals subsumes imperfect public monitoring as a special case. For repeated games with public monitoring, the folk theorem by Fudenberg, Levine, and Maskin [11] depends on certain distinguishability assumptions on the

signal structure (Radner, Myerson, and Maskin [26]). Kandori [15] shows that, if communication is allowed, the folk theorem holds for a larger class of signal structures. However, there remain non-trivial restrictions. For instance, neither folk theorem works for games with two players, two actions, and two signals. The present paper shows that the folk theorem holds regardless of the numbers of players, actions, and signals, even without explicit communication, if the signal distribution is non-constant with respect to every player's action and the cost of perfect monitoring is finite.

There are a few papers that also study repeated games with costly observation. Ben-Porath and Kahneman [4] show that the folk theorem holds if explicit communication is allowed. The present paper shows that the theorem holds even if explicit communication is not allowed. On the other hand, we assume that players make an observational decision after they choose their stage-game action and observe the realization of public randomization. Thus, our model pertains to the case where a player's action can be observed ex post, at least within the period. Ben-Porath and Kahneman allow the case where players make both decisions at the same time. Kandori and Obara [17] consider a related setup with costly observation where what players observe may be wrong with a small probability. They assume that observational activities are observable with small errors. Their construction requires observation costs to be small and relies on the prisoners' dilemma and its variants.

The folk theorem in the present paper has significant advantages over the folk theorem that relies on explicit communication. First, the specific model of explicit communication used in the literature assumes that communication is costless. The assumption means that a message is costless not only to announce but also to receive and process. However, communication is never costless. It takes time and requires attention. In particular, it is costly to ensure that everyone gets your message. Firms spend a large portion of their resources for advertising, to ensure that their messages reach consumers. Our model can be considered as a model with costly communication, where the act of sending messages is included as part of stage-game action. The result obtained here thus shows how to design incentive-compatible communication schemes in the world of costly communication.

Our theorem also has practical implications for antitrust laws since it suggests that explicit communication is unnecessary for efficient cooperation, even in environments with private information. By dispensing with explicit communication, cartels can reduce the probability of antitrust charges, often almost to zero, because of the lack of evidence. In the US, an antitrust charge requires evidence of explicit communication among cartel members, and as a result, tacit collusion is legal. To see what the policy implies, we need to know to what extent collusion can be achieved via tacit coordination. The result obtained here shows that tacit coordination can achieve full collusion even in environments where information is costly and private. The result is shown in a fairly general model. The

theory therefore suggests that an antitrust policy that relies on evidence of explicit communication is not effective if cartels are patient.

One feature of the present paper is that the proof is constructive. We present a specific class of strategy profiles that can support any target payoff profile. The strategy profiles in this class have an intuitive general structure allowing for reasonable interpretations (Rubinstein [27]). Section 4 illustrates the construction for a prisoners' dilemma example.


## 2   Model

We consider a repeated game, where a set of players play the same game repeatedly over periods $t = 1, 2, \ldots$. Let $N = \{1, 2, \ldots, n\}$ denote a finite set of players, where $n \geqslant 2$, and let $A_i$ be a finite set of actions that player $i$ can choose in each period, where $|A_i| \geqslant 2$. Let $A \equiv A_1 \times \cdots \times A_n$ denote the set of action profiles.

Given a set $K$, let $\Delta(K)$ denote the set of probability distributions over $K$. Thus $\mathscr{A}_i \equiv \Delta(A_i)$ denotes the set of mixed actions of player $i$, $\mathscr{A} \equiv \mathscr{A}_1 \times \cdots \times \mathscr{A}_n$ denotes the set of mixed action profiles, and $\Delta(A)$ denotes the set of correlated action profiles.

At each period, after all players choose actions, each player $i$ observes a signal $\omega_i$ costlessly and privately. The set of signals that player $i$ might receive is given by a finite set $\Omega_i$. A signal profile $\omega = (\omega_1, \ldots, \omega_n) \in \Omega_1 \times \cdots \times \Omega_n$ is realized with probability $P(\omega \mid a)$ given an action profile $a$. Let $P_i(\omega_i \mid a)$ denote the marginal distribution of $\omega_i$ given $a$. We assume the following on the marginal distribution.

**Assumption 1.** For all $i \in N$, all $\omega_i \in \Omega_i$, and all $a \in A$,

$$P_i(\omega_i \mid a) > 0.$$

**Assumption 2.** There exists no player $i \in N$ such that for all pairs $\{a_i^1, a_i^2\} \subseteq A_i$, all $a_{-i} \in A_{-i}$, and all $r \in N \setminus \{i\}$,

$$P_r(\cdot \mid a_i^1, a_{-i}) = P_r(\cdot \mid a_i^2, a_{-i}).$$

Assumption 1 states that any $\omega_i \in \Omega_i$ is realized with a positive probability given any action profile. Since the full-support condition is required only for individual signal spaces, there may exist some $(\omega, a)$ such that $P(\omega \mid a) = 0$. Assumption 2 states that there exists no player who has no influence at all on any other player's signal. This means that, for each player, there exists a pair of actions that induce different probability distributions of at least one player's signal under at least one action profile. Assumption 2 is innocuous since, as we describe below, $\omega_j$ contains information about $j$'s own payoff, and usually, a player's action affects the probability distribution of other players' payoffs.[2] It should be also noted

---

[2]Assumption 2 is stated with pure actions, but rewriting it with mixed actions yields the same condition.

that Assumptions 1 and 2 allow the case of *public* signals, the case in which $\omega_1 = \cdots = \omega_n$ always holds.

The stage-game payoff for player $i$ is given by $\pi_i(a_i, \omega_i)$, which depends on his own action $a_i$ and the realized private signal $\omega_i$.[3] Since the payoff depends on what the player already knows, it gives no additional information on other players' actions or signals.[4] One special case is when the realized stage-game payoff is the sole information contained in the free signal, which is the case if $\pi_i(a_i, \cdot) \colon \Omega_i \to \mathbb{R}$ is one-to-one for each $a_i$.

Given an action profile $a \in A$, the expected stage-game payoff for player $i$ is

$$u_i(a) \equiv \sum_{\omega_i \in \Omega_i} \pi_i(a_i, \omega_i) P_i(\omega_i \mid a).$$

We write $u(a) = (u_i(a))_{i \in N}$. For a mixed action profile $\alpha \in \mathscr{A}$, we abuse notation and write $u(\alpha) = (u_i(\alpha))_{i \in N}$ to denote the expected payoff profile under $\alpha$. Similarly, for a correlated action profile $\rho \in \Delta(A)$, we write $u_i(\rho) = \sum_{a \in A} \rho(a) u_i(a)$ and $u(\rho) = (u_i(\rho))_{i \in N}$.

Observation takes place at the end of each period. After all players choose actions and receive signals, each player chooses the set of players to observe. Let $\lambda_i \colon 2^{N \setminus \{i\}} \to \mathbb{R}_+$ denote the observation cost function for player $i$. If player $i$ chooses $J \subseteq N \setminus \{i\}$, he incurs observation costs $\lambda_i(J)$ and obtains completely accurate information about the realized action profile $(a_j)_{j \in J}$ in the present period. We assume that $\lambda_i(\emptyset) = 0$, $\lambda_i(J) \geqslant 0$ for all $J$, and $\lambda_i(J) \leqslant \lambda_i(J')$ if $J \subseteq J'$.[5]

We assume that what players observe from their costly observational activities are their private information. We also assume that observational activities are stealthy. This means first that observational activities are not observable to other players. That is, whether player $i$ observes another player $j$ in a given period (let alone what $i$ observes) is unobservable to any player $k \neq i$ even if $k$ observes $i$ in the period (even if $k = j$). Second, players do not even receive any noisy information about other players' observational activities. These assumptions imply that one's observational decision itself does not affect other players' future actions. Therefore, deviations with respect to observational decisions cannot be punished directly. This feature makes it difficult to create monitoring incentives.

We assume that there exists a public randomization device (e.g., public lotteries, last digits of the Dow Jones, etc), which generates a sequence of independent

---

[3]Our result extends to the model where the stage-game payoff depends directly on the action profile $a$ and the signal profile $\omega$, i.e., $\pi_i(a, \omega)$, if we also assume that players do not observe their own stage-game payoffs (e.g., the repeated game ends stochastically and players collect repeated-game payoffs at the end).

[4]If the realized payoff does give additional information, we can redefine signals to include payoff information. That is, we can redefine a signal as a pair $(\omega_i, \pi_i)$ of the original signal and the realized payoff. If this pair satisfies the full-support condition, Assumption 1 is preserved and the payoff gives no additional information.

[5]The monotonicity condition is assumed only for simplicity and can be dispensed with, as we will note in Footnote 12.

random variables $(X_1, Y_1, X_2, Y_2, X_3, \dots)$ that are all uniformly distributed over $[0, 1]$. Random variable $X_t$ ($t = 1, 2, \dots$) is realized at the beginning of period $t$ before players choose actions, while $Y_t$ ($t = 1, 2, \dots$) is realized in the middle of period $t$ right before players make observational decisions. The realizations of the random variables ("sunspots") are observable publicly and costlessly and irrelevant to stage-game payoffs.

To summarize, the sequence of events within a given period $t$ is as follows. First, players observe the realization of public random variable $X_t$. Second, players simultaneously choose an action $a_i \in A_i$. Third, each player $i$ observes a signal $\omega_i$ privately, which determines $\pi_i(a_i, \omega_i)$. Fourth, players observe the realization of the middle-of-period public random variable $Y_t$. Fifth, each player $i$ chooses whom to observe, $J_i \subseteq N \setminus \{i\}$, and incurs a disutility of $\lambda_i(J_i)$. Finally, $i$ observes the realized action profile $(a_j)_{j \in J_i}$.

Note that players make observational decisions after observing the free signal $\omega_i$. Under the assumption, the amount of information that players can use for observational decisions is maximal: players can fine-tune their observational decisions based on $\omega_i$.[6] However, our result does not depend on the assumption. Indeed, as we show, there exists a sufficiently large class of equilibria in which observational decisions do not depend on the realized value of $\omega_i$.

Player $i$'s (private) history at the beginning of period $t \geqslant 2$ is a sequence of realizations of public random variables, his own actions, realizations of his private signals, and his observations about the other players' actions, all up to (including) period $t - 1$. Formally, it is a sequence

$$h_i^t = [x_k, a_{i,k}, \omega_{i,k}, y_k, (a_{j,k})_{j \neq i}]_{k=1}^{t-1}$$
$$\in \left[ [0, 1] \times A_i \times \Omega_i \times [0, 1] \times \prod_{j \in N \setminus \{i\}} (A_j \cup \{\phi\}) \right]^{t-1}.$$

In this sequence, $x_k \in [0, 1]$ is the realized value of random variable $X_k$, $a_{i,k} \in A_i$ is player $i$'s action in period $k$, $\omega_{i,k} \in \Omega_i$ is the realized private signal of $i$ in period $k$, $y_k \in [0, 1]$ is the realized value of random variable $Y_k$, and $a_{j,k} \in A_j \cup \{\phi\}$ is $i$'s observation about player $j$'s action in period $k$, where $a_{j,k} = \phi$ means that $i$ did not observe $j$ in period $k$.

For all $t = 1, 2, \dots$, let $H_i^t$ denote the set of all (private) histories for player $i$ at period $t$ ($H_i^1$ is an arbitrary singleton). A strategy of player $i$ is a pair of functions

---

[6]If players can choose the timing of observational decisions, they would weakly prefer to wait for the realization of $\omega_i$.

$\sigma_i = (\sigma_i^a, \sigma_i^m)$ such that

$$\sigma_i^a: \bigcup_{t=1}^{\infty} (H_i^t \times [0,1]) \to \Delta(A_i),$$

$$\sigma_i^m: \bigcup_{t=1}^{\infty} (H_i^t \times [0,1] \times A_i \times \Omega_i \times [0,1]) \to \Delta(2^{N\setminus\{i\}}).$$

A strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ generates a probability distribution over sequences $(a_t, (J_{i,t})_{i \in N})_{t=1}^{\infty}$, where $a_t \in A$ is the action profile in period $t$ and $J_{i,t} \subseteq N \setminus \{i\}$ is the set of players that $i$ observes in period $t$. Given the sequence, player $i$'s overall payoff is

$$(1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} \big[ u_i(a_t) - \lambda_i(J_{i,t}) \big],$$

where $\delta \in (0,1)$ is a discount factor common to all players. Players maximize the expected overall payoff. We are interested in sequential equilibria of the repeated game when the discount factor is close to one.

## 3  Result

Player $i$'s minmax payoff is defined by

$$\underline{u}_i \equiv \min_{\alpha_{-i} \in \mathscr{A}_{-i}} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i}),$$

where $\mathscr{A}_{-i} \equiv \prod_{j \neq i} \mathscr{A}_j$. Let

$$V \equiv \text{convex hull of } \{u(a) : a \in A\},$$
$$V^* \equiv \{v \in V : v_i \geqslant \underline{u}_i \text{ for all } i \in N\}.$$

Note that $\underline{u}_i$, $V$, and $V^*$ are all defined independently of the observation cost functions $\lambda_1, \ldots, \lambda_n$.

Our result is the following.

**Theorem.** *Suppose that $V^*$ has a dimension of n, i.e., $\text{int}\, V^* \neq \emptyset$. Then for any $v^* \in \text{int}\, V^*$, there exists $\underline{\delta} \in (0,1)$ such that, for any $\delta \in [\underline{\delta}, 1)$, there exists a sequential equilibrium whose payoff profile is $v^*$.*

*Proof.* See Appendix.

Our proof is constructive: for a given payoff profile $v^* \in \text{int}\, V^*$, we construct a specific strategy profile $\sigma$ that is a sequential equilibrium and yields $v^*$ if the discount factor is close to one.

9

## 4 An Illustration: Prisoners' Dilemma

In this section, we illustrate the proof of the folk theorem using a repeated prisoners' dilemma. At the end of the section, we comment on how to generalize the construction to general stage games.

Suppose that there are two players and each of them has two actions and two signals: $A_1 = A_2 = \{C,D\}$ and $\Omega_1 = \Omega_2 = \{C,D\}$. If $\omega_i \neq a_j$ ($j \neq i$), we call it an "error." The signal distribution $P(\omega \,|\, a)$ is such that a signal profile $\omega$ with a single error occurs with probability $p_1$ and the signal profile with two errors occurs with probability $p_2$. The right matrix in Figure 1 gives the signal distribution when $a = (C,C)$. Assumptions 1 and 2 are satisfied if and only if $0 < p_1 + p_2 < 1$ and $p_1 + p_2 \neq 0.5$, respectively. We here assume $p_1 + p_2 < 0.5$ for concreteness and will note how to deal with the reverse inequality (Footnote 8). The stage-game payoff function $\pi_i(a_i, \omega_i)$ is such that the expected payoff function $u_i(a)$ is given by the left matrix in Figure 1. The observation cost function is given by $\lambda_i(\emptyset) = 0$ and $\lambda_i(\{j\}) = \lambda > 0$ for each $i$ and $j \neq i$.

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

$u(a)$

|     | $C$             | $D$   |
| --- | --------------- | ----- |
| $C$ | $1 - 2p_1 - p_2$ | $p_1$ |
| $D$ | $p_1$           | $p_2$ |

$P(\omega \,|\, a)$ when $a = (C,C)$

**Figure 1:** The stage-game payoffs and the signal distribution

Fix an interior feasible payoff profile $v^* = (v_1^*, v_2^*) \gg (0,0)$ as a target (note that $\underline{u}_i = 0$ for each $i$). In what follows, we construct a specific strategy profile that achieves $v^*$ and constitutes a sequential equilibrium for sufficiently patient players.[7] We here describe only the strategy on the equilibrium path since it gives enough details to see the main idea. The technical details on the out-of-equilibrium path are given fully in Appendix.

The play on the equilibrium path is characterized by three types of periods: *cooperation* periods, *examination* periods, and *report* periods. We begin with describing those periods and the rule that governs the transition among them. A summary diagram is given in Figure 2.

---

[7]For some classes of signal distributions, the literature proves a folk theorem for repeated prisoners' dilemma without observation costs. Sekiguchi [28], Bhaskar and Obara [5], Piccione [25], and Ely and Välimäki [7] consider the case of almost perfect monitoring, where $p_1$ and $p_2$ are close to 0. Matsushima [24] extends the result to signal distributions where errors are not necessarily rare but signals are independent across players given an action profile, which is the case in the current example if $p_1 = \sqrt{p_2} - p_2$. The following exposition covers these cases and all remaining cases.
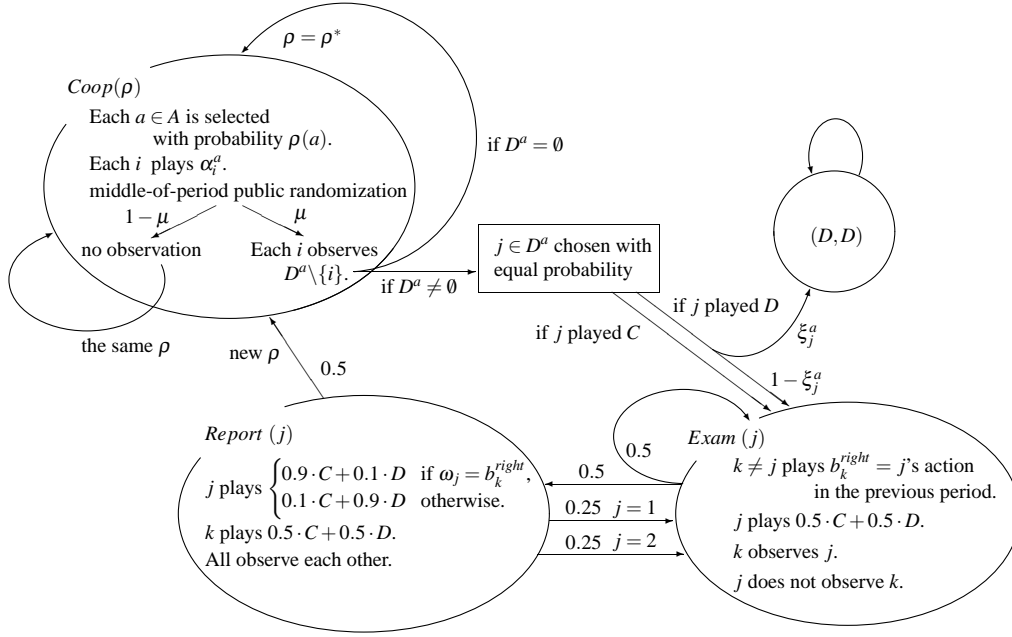
**Figure 2:** The strategy profile for the prisoners' dilemma.

## 4.1 Cooperation Periods

Cooperation periods are parameterized by $\rho \in \Delta(A)$ and denoted by $Coop(\rho)$. The initial period ($t = 1$) is $Coop(\rho^*)$ for a particular $\rho^*$ that will be specified below. In $Coop(\rho)$, the public random variable realized at the beginning of the period chooses each pure action profile $a \in A$ with probability $\rho(a)$. If $a$ is chosen, players play a mixed action profile $\alpha^a$ defined by

$$\alpha^{CC} = ((1-\eta)C + \eta D, (1-\eta)C + \eta D),$$
$$\alpha^{CD} = ((1-\eta)C + \eta D, D),$$
$$\alpha^{DC} = (D, (1-\eta)C + \eta D),$$
$$\alpha^{DD} = (D, D),$$

where $\eta \in (0,1)$ is a small probability. Note that, if $a_i$ is not a short-run best response to $\alpha_j^a$, then $\alpha_i^a$ assigns probability $\eta$ to the best response. If $a_i$ is a short-run best response, $\alpha_i^a$ assigns probability 1 to $a_i$. Let $D^a$ be the set of players who randomize under $\alpha^a$:

$$D^{CC} = \{1,2\}, \ D^{CD} = \{1\}, \ D^{DC} = \{2\}, \ D^{DD} = \emptyset.$$

Players use the middle-of-period public randomization to coordinate their observational decisions. Specifically, with probability $1 - \mu$ (i.e., if $y_t \geqslant \mu$), where $\mu \in (0,1)$ is small, no one observes the other player, and the next period continues to

11

be *Coop($\rho$)* with the same $\rho$. With probability $\mu$ (i.e., if $y_t < \mu$), on the other hand, each player observes the other player if the other player is in $D^a$. In this case, the play in the next period, say period $t+1$, is determined as follows. If $D^a = \emptyset$, which is the case if and only if $a = (D,D)$, period $t+1$ is set to *Coop($\rho^*$)*, i.e., the play is reset and returns to the initial period. If $D^a \neq \emptyset$, then a player, who will be denoted as $j$, is chosen equiprobably from $D^a$, with the help of public randomization at the beginning of period $t+1$. If $j$'s action in period $t$ was $C$, then period $t+1$ is an *examination period* denoted as *Exam($j$)*, which is described in the next subsection. If $j$'s action in period $t$ was $D$, then with probability $\xi_j^a$, the perpetual repetition of $(D,D)$ starts from period $t+1$, where $\xi_j^a$ is defined by

$$\xi_j^a = \frac{(1-\delta)|D^a|}{\mu \delta v_j^*}. \tag{1}$$

Note that, for a fixed $\mu$, we have $0 < \xi_j^a < 0.5$ if $\delta$ is sufficiently close to one. With the remaining probability (i.e., $1 - \xi_j^a$), period $t+1$ is an examination period *Exam($j$)*. As we will verify, the particular choice of $\xi_j^a$ equates the short-run gain from choosing $D$ and the long-run loss from the possibility of triggering the repetition of $(D,D)$, thereby motivating the players $j \in D^a$ to randomize between $C$ and $D$.

### 4.2 Examination Periods

In examination periods *Exam($j$)*, player $k \neq j$ is given a "test" where he has to state the action that player $j$ chose in the previous period. To state the answer without explicit communication, player $k$ ("examinee") uses his stage-game action. Specifically, player $k$ chooses the same action that player $j$ chose in the previous period, which is denoted $b_k^{right} \in A_k$. Player $j$ ("examiner"), on the other hand, plays $0.5 \cdot C + 0.5 \cdot D$. Regardless of the middle-of-period public randomization, player $k$ observes player $j$, while $j$ does not observe $k$. Player $j$ has no incentive to observe $k$ since $b_k^{right}$ is a pure action that player $j$ knows.

The state transition is determined by the public randomization at the beginning of the next period. With probability 0.5, the next period is again *Exam($j$)*, in which $k$ states what $j$ just played in this (examination) period. With the remaining probability, the next period is a *report* period denoted by *Report($j$)*.

Before we proceed to describe report periods, we briefly note the role of examination periods and why we need report periods. In examination periods, the action prescribed to the examinee depends on the realized action of the examiner in the previous period. This feature makes a player who ignored the instruction of observation uncertain of the prescribed action in the examination period. This feature, however, does not suffice for providing observational incentives, since, as described above, the examinee's action is *pure* and therefore is not observed by the examiner. We can surely create observational incentives if we make the examinee's

action stochastic. However, it makes the exam trivial since the stage-game has only two actions and therefore there will be an action that belongs to the support of the mixed action for any realized action of the examiner in the previous period.

Our solution is to use the costless signal that the examiner receives in the examination period. Since the distribution of the examiner's signal depends on the examinee's action, we can motivate the examinee to choose a pure action if we make the continuation strategy dependent on the signal received by the examiner. However, there remains a problem, which is that the examiner's signal is his private information and therefore does not allow the players to coordinate their continuation strategy. This is why we proceed to a report period, in which the examiner "reports" his private signal to the examinee by means of his stage-game action, which in turn determines the continuation strategy profile, as we now describe.

### 4.3 Report Periods

By construction, if the present period is $Report(j)$, the previous period is $Exam(j)$. Let $k \neq j$ denote the examinee. In the report period, player $j$ (examiner) uses his stage-game action to announce whether he "approves" or "disapproves" of player $k$. Specifically, $j$ plays the following mixed action:

$$
\begin{array}{ll}
0.9 \cdot C + 0.1 \cdot D & \text{if } \omega_j = b_k^{right}, \\
0.1 \cdot C + 0.9 \cdot D & \text{otherwise}
\end{array}
\tag{2}
$$

where $b_k^{right}$ is the action prescribed to $k$ in the previous period and $\omega_j$ is the signal received by $j$ in the same period.[8] Meanwhile, player $k$ plays $0.5 \cdot C + 0.5 \cdot D$. The players observe each other.

To understand $j$'s action, recall that $p_1 + p_2 < 0.5$ and hence the probability of $\omega_j = b_k^{right}$ is maximized if and only if player $k$ plays $b_k^{right}$. If $\omega_j = b_k^{right}$, player $j$ basically gives his approval for player $k$ by choosing $C$.[9] Player $j$ actually randomizes, giving his approval only with probability 0.9. Symmetrically, if $\omega_j \neq b_k^{right}$, player $j$ disapproves of player $k$ with probability 0.9. We make player $j$ randomize to ensure that his action ("report") is observed by $k$.[10]

This construction implies that the probability of getting $j$'s approval is maximized if and only if $k$ plays $b_k^{right}$. If $k$ is uncertain of $b_k^{right}$, he will play a wrong action in the examination period with a positive probability and hence the expected probability of getting $j$'s approval is strictly lower than if $k$ knows $b_k^{right}$.

---

[8]This part depends on the assumption $p_1 + p_2 < 0.5$. If the reverse inequality holds, we simply reverse the conditions in (2).

[9]We let $C$ signify $j$'s approval, but the roles of $C$ and $D$ can be reversed.

[10]Even if $j$ does not randomize, his action appears random to $k$ since $j$'s action depends on his private signal. However, if $p_1 = 0$ (which we allow), there is a perfect correlation between signals across the players, and therefore $k$ can infer $j$'s action without costly observation. This possibility introduces an unnecessary complication to the proof, which is why we introduced the randomization in $j$'s action. The exact way in which $j$ trembles is immaterial.

The state transition depends on the public randomization at the beginning of the next period. With probability 0.25, the next period is *Exam*(1), where player 1 is the examiner and player 2 has to state 1's action in the report period. With probability 0.25, the next period is *Exam*(2), where the players' roles are reversed. With the remaining probability, the next period is *Coop*($\rho$) where $\rho \in \Delta(A)$ is newly selected. The selection of $\rho$ is described in what follows.

### 4.4 Selection of $\rho$

The selection has two goals. The first is to offset the difference between the target payoffs $v^*$ and the realized payoffs during the previous two periods, which are an examination period and a report period. By doing this, we can make the players' continuation values from any examination period equal to $v^*$ regardless of the realized actions in the two periods. This in turn makes both players in the report period and the examiner in the examination period indifferent about their actions and willing to randomize as prescribed. The second goal of the selection rule for $\rho$ is to punish the examinee who did not get the examiner's approval.

To give a detailed description of the selection of $\rho$, we begin with its first goal. That is, we look for a selection rule for $\rho \in \Delta(A)$ that makes the continuation values from any examination period equal to $v^*$.

By construction, if period $t$ is a cooperation period and period $t-1$ is a report period, then period $t-2$ is an examination period. Let $k$ be the examinee, $b_k^{right}$ his prescribed action in the examination period, $b_j^{obs}$ the observed action of the examiner in the examination period, and $c^{obs}$ the observed action profile in the report period. Abusing notation, let $Coop_i(\rho)$ denote the continuation value of player $i$ from a cooperation period with $\rho$. Then our objective is to find a distribution $\rho' \in \Delta(A)$ that satisfies the following equation for all $i$:

$$
\begin{aligned}
v_i^* = {} & (1-\delta)\left[u_i(b_k^{right}, b_j^{obs}) - \lambda_i(N \setminus \{i,k\})\right] \\
& + \frac{1}{2}\delta v_i^* + \frac{1}{2}\delta(1-\delta)\left[u_i(c^{obs}) - \lambda\right] \\
& + \frac{1}{4}\delta^2 v_i^* + \frac{1}{4}\delta^2 Coop_i(\rho').
\end{aligned} \tag{3}
$$

Observe that $\rho'$ depends on $(k, b_k^{right}, b_j^{obs}, c^{obs})$, which are determined in the preceding two periods. If $\rho'$ is set to satisfy (3) for any given $(k, b_k^{right}, b_j^{obs}, c^{obs})$ and any $i$, then the continuation value from any examination period is indeed $v_i^*$ for

14

all players. Rearranging (3) yields

$$
\begin{aligned}
0 = {} & \left[ v_i^* - u_i(b_k^{right}, b_j^{obs}) + \lambda_i(N \setminus \{i,k\}) \right] \\
& + \left[ v_i^* - u_i(c^{obs}) + \lambda \right] \frac{\delta}{2} \\
& + \left[ v_i^* - Coop_i(\rho') \right] \frac{\delta^2}{4(1-\delta)}.
\end{aligned}
\tag{4}
$$

To identify $\rho'$ that satisfies (4), we need to compute the value $Coop_i(\rho')$. Since $Coop_i(\rho')$ is the continuation value from a cooperation period,

$$
\begin{aligned}
Coop_i(\rho') = {} & (1-\delta) \sum_{a \in A} \rho'(a) \left[ u_i(\alpha^a) - \mu \lambda_i(D^a \setminus \{i\}) \right] \\
& + (1-\mu)\delta Coop_i(\rho') \\
& + \mu \delta \left[ 1 - \sum_{a \in A: D^a \neq \emptyset} \rho'(a) \frac{1}{|D^a|} \eta \sum_{j \in D^a} \xi_j^a \right] v_i^*.
\end{aligned}
\tag{5}
$$

The equation is written based on the assumption that $Coop_i(\rho^*) = v_i^*$, which will be verified later.

Substituting (1) into (5) to eliminate $\xi_j^a$, we obtain

$$
Coop_i(\rho') = v_i^* + \frac{1-\delta}{1-\delta+\mu\delta} \left[ \hat{u}_i(\rho') - v_i^* \right]
\tag{6}
$$

where $\hat{u}_i \colon A \to \mathbb{R}$ is defined by

$$
\hat{u}_i(a) \equiv u_i(\alpha^a) - \mu \lambda_i(D^a \setminus \{i\}) - \eta \sum_{j \in D^a} \frac{v_i^*}{v_j^*}.
$$

The function $\hat{u}_i$, which we call the *virtual payoff function* for player $i$, represents his stage-game payoff in a cooperation period when $a$ is selected, after we take into account expected observation costs and expected losses from possible transition to the punishment stage. By choosing small numbers for $\mu$ and $\eta$, we can make the virtual payoff function $\hat{u}_i$ arbitrarily close to the true payoff function $u_i$.

Substituting (6) into (4), we obtain that for each $i$,

$$
\begin{aligned}
\hat{u}_i(\rho') = {} & v_i^* + \frac{1-\delta+\mu\delta}{(1/4)\delta^2} \left[ v_i^* - u_i(b_k^{right}, b_j^{obs}) + \lambda_i(N \setminus \{i,k\}) \right] \\
& + \frac{1-\delta+\mu\delta}{(1/2)\delta} \left[ v_i^* - u_i(c^{obs}) + \lambda \right].
\end{aligned}
\tag{7}
$$

There exists $\rho' \in \Delta(A)$ that satisfies this equation for all $i$ if $\mu$ and $\eta$ are close to 0 and $\delta$ is close to 1. Indeed, if $\delta$ is close to 1 and $\mu$ is close to 0, $1 - \delta + \mu\delta$

is close to 0 and hence $\hat{u}_i(\rho')$ is close to $v_i^*$. Since $v^*$ is an interior point of the feasible payoff vector set and $\hat{u}_i$ is close to $u_i$, a distribution $\rho' \in \Delta(A)$ that yields the equality for all $i$ exists regardless of the values of $(k, b_k^{right}, b_j^{obs}, c^{obs})$.

The preceding argument shows that if $\rho'$ is always chosen to satisfy (7), the continuation value from any examination period is $v_i^*$ for all players. Further, (3) shows that the continuation value is the same regardless of the randomizing players' realized actions in the examination and report periods (i.e., $(b_j^{obs}, c^{obs})$). This implies that all these randomizing players are completely indifferent about their actions.

The choice of $\rho'$, however, does not give a right incentive to the examinee (player $k$) in the examination period, since (7) depends only on what he is prescribed to do ($b_k^{right}$) and not what he does. This is where our second goal comes in. To deal with the examinee's incentive, we modify $\rho'$ slightly to punish or reward him depending on the report from the examiner. Specifically, let $AP_k$ denote the probability that player $k$ earns his examiner's approval given that $k$ plays as prescribed. For the current example,

$$AP_k = 0.9(1 - p_1 - p_2) + 0.1(p_1 + p_2).$$

Let $\varepsilon > 0$ be a small number. Then, finally, let $\rho \in \Delta(A)$ be such that

$$\hat{u}_k(\rho) = \begin{cases} \hat{u}_k(\rho') + \varepsilon(1 - AP_k) & \text{if } c_j^{obs} = C, \\ \hat{u}_k(\rho') - \varepsilon AP_k & \text{otherwise,} \end{cases} \tag{8}$$
$$\hat{u}_j(\rho) = \hat{u}_j(\rho').$$

This is the $\rho$ that is used in the new cooperation period. If $\varepsilon > 0$ is sufficiently small, there exists $\rho \in \Delta(A)$ that satisfies these equations. The equations mean that player $k$ receives a "bonus" of $\varepsilon(1 - AP_k)$ if he earns his examiner's approval, and pays a "penalty" of $\varepsilon AP_k$ otherwise. Since the examiner's approval is given with probability $AP_k$ in equilibrium, the expected net bonus is zero, and hence the continuation value from an examination period remains unchanged and equal to $v_k^*$.

### 4.5 Initial Play

The initial period is set as a cooperation period $Coop(\rho^*)$ where $\rho^* \in \Delta(A)$ is chosen to satisfy
$$\hat{u}(\rho^*) = v^*.$$

Such a $\rho^*$ exists if $\mu$ and $\eta$ are close to 0 and $\delta$ is close to 1. Then (6) implies $Coop_i(\rho^*) = v_i^*$ and hence $v^*$ is indeed the payoff profile for the entire repeated game under the strategy profile.

## 4.6 Incentives

We now discuss why players have incentives to follow the strategy described above. We start with incentives with respect to stage-game actions.

(i) *Cooperation periods.* Players $i \in D^a$ are willing to randomize between $C$ and $D$ because of the way the probabilities $\xi_i^a$ are defined. For our specific prisoners' dilemma, the short-run gain of playing $D$ is $1 - \delta$. The long-run loss of playing $D$ is given by $\xi_i^a (1/|D^a|) \mu \delta v_i^*$. The probabilities $\xi_i^a$ are set so that the long-run loss equals the short-run gain (see (1)). On the other hand, players $i \notin D^a$ are prescribed to play $D$ and, indeed, these players have no incentive to play $C$ since their action in this period is not observed by the other player and has no effect on the continuation play.

(ii) *Report periods.* In report periods, as discussed above, players are completely indifferent about their actions because of the way $\rho$ is chosen in the subsequent cooperation period (if reached).

(iii) *Examiner in an examination period.* For the same reason, the examiner in an examination period is also indifferent about his action.

(iv) *Examinee in an examination period.* Suppose that player $k$ is the examinee in an examination period and $b_k^{right} = C$ (the case when $b_k^{right} = D$ is omitted since it is similar and simpler). The short-run gain from playing $D$ is $1 - \delta$. Playing $D$, on the other hand, increases the probability that $j$'s signal is $\omega_j = D$, by $L = 1 - 2p_1 - 2p_2 > 0$. If $\omega_j = D$, then if the next period is a report period (which occurs with probability 0.5), the probability that the examiner gives his approval (choosing $C$) goes down from 0.9 to 0.1. This has three effects on player $k$'s payoffs. First, there is a direct effect on $k$'s stage-game payoff in the report period. Second, there is an effect on $\rho'$ since $\rho'$ depends on $c_j^{obs}$ (see (3)). However, (3) implies that these two effects are canceled out by each other. Finally, there is an effect on $\rho$ through the last term of (8). If the examiner gives his disapproval, $\hat{u}_k(\rho)$ goes down by $\varepsilon$, which means, by (6), that $Coop_k(\rho)$ goes down by $(1 - \delta)\varepsilon/(1 - \delta + \mu\delta)$. This effect matters if the report period is followed immediately by a cooperation period, which occurs with probability 0.5. Altogether, the long-run loss from playing $D$ is

$$L\frac{1}{2}(0.9 - 0.1)\frac{1}{2}\delta^2 \frac{(1-\delta)\varepsilon}{1 - \delta + \mu\delta}.$$

A sufficient condition for this to exceed the short-run gain is

$$1 < \frac{0.2L\delta^2\varepsilon}{1 - \delta + \mu\delta}.$$

This is satisfied if $\delta$ is close to 1 and $\mu$ is close to 0. Recall that $\mu$ is the probability that observation is prescribed in cooperation periods, which also determines how long a cooperative phase with the same $\rho$ is expected to continue. If $\mu$ is small, a

17

cooperative phase with the same $\rho$ is expected to continue for a large number of periods and therefore a slight effect on $\rho$ has a significant effect on the long-run payoff. This gives players strong incentives to answer correctly in examination periods since doing so maximizes the probability that $\rho$ is set favorably.

(v) *Observation.* Suppose that player $i$ did not observe player $j \neq i$ when he was prescribed to. Since $\eta$ and $\xi_k^a$ are set small, it follows that with a probability $p \geqslant 0.5$, the next period is an examination period. In the examination period, player $i$ is chosen as the examinee at least with probability 0.5. In this (worst) case, player $i$ is uncertain of $b_i^{right}$ and therefore there is a positive probability that his action turns out to be wrong: $b_i \neq b_i^{right}$. This probability is bounded below by some $F > 0$. In this event, the continuation payoff goes down strictly, as we just showed in the previous paragraph. A sufficient condition for the loss to exceed the saving of observation costs is

$$\lambda < \frac{1}{4} F \delta \Big[ \frac{0.2 L \delta^2 \varepsilon}{1 - \delta + \mu \delta} - 1 \Big],$$

which holds if $\delta$ is close to 1 and $\mu$ is close to 0.[11]

### 4.7 Extending to General Games

We now briefly describe how to extend the construction to general stage games and signal distributions. For cooperation periods, we can construct a mixed action $\alpha_i^a$ for each $a \in A$ and $i \in N$ to satisfy the following properties: (i) if $a_i$ is a short-run best response to $\alpha_{-i}^a$, then $\alpha_i^a$ assigns probability 1 to $a_i$, and (ii) if $a_i$ is not a short-run best response to $\alpha_{-i}^a$, then $\alpha_i^a$ is a randomization between $a_i$ and a short-run better response, assigning a small probability to the short-run better response. If observation is prescribed (i.e., if $y_t < \mu$), players observe only the players who randomize in $\alpha^a$.[12]

The play in examination periods is also modified. If there are three or more players, each examination period is parameterized by a pair of players $(j, k)$, where $j$ is one of the players who randomized in the previous period and $k \neq j$ is the examinee. By Assumption 2, there exists an action profile $a_{-k}$ under which player $k$'s action has a non-zero influence on the signal distribution of a player $r \neq k$.

---

[11]The value of $F$ is inversely related to the accuracy of costless signals. Thus, as the costless information becomes more accurate, the inequality gets tighter and therefore $\delta$ need to be higher for the strategy to work. When players can obtain fairly accurate information costlessly, a high level of patience is necessary to motivate them to undertake costly observation. The result is a by-product of our specific strategy, where the observational decision is made independently of the realized private signal.

[12]If the observation cost functions are not monotonic (i.e., $J' \subseteq J$ does not imply $\lambda_i(J') \leqslant \lambda_i(J)$, which is relevant if it may be easier to monitor multiple players at the same time), then it suffices to modify the strategy as follows: if $J$ is the set of randomizing players in the current period and monitoring is instructed, player $i$ monitors $J'$ that solves $\min_{J' \supseteq J} \lambda_i(J')$ and then ignores any observed deviation by players in $J' \setminus J$.

The influence of $k$'s action does not disappear if $a_{-k}$ is perturbed to a completely mixed action profile $\beta_{-k}$. Then there exists a subset $\Omega'_r \subseteq \Omega_r$ such that

$$\max_{a_k} P_r(\Omega'_r \,|\, a_k, \beta_{-k}) > \min_{a_k} P_r(\Omega'_r \,|\, a_k, \beta_{-k}),$$

which is equivalent to

$$\underset{a_k}{\mathrm{Argmax}}\, P_r(\Omega'_r \,|\, a_k, \beta_{-k}) \bigcap \underset{a_k}{\mathrm{Argmax}}\, P_r(\Omega_r \setminus \Omega'_r \,|\, a_k, \beta_{-k}) = \emptyset. \tag{9}$$

The prescribed action for player $k$, i.e., $b_k^{right}$ is chosen from a solution to one of the maximization problems in (9). Which maximization to choose is determined by the realized action of player $j$ in the previous period.[13] While $k$ plays $b_k^{right}$, the other players randomize according to $\beta_{-k}$. In the report period that follows, player $r$ (not $j$) decides whether to approve of player $k$ based on whether $\omega_r \in \Omega'_r$ or not, with trembles, as in (2). This construction ensures that $k$'s action in the examination period affects $r$'s signal distribution, and failing to observe $j$'s action in the previous period reduces the ex-ante probability of getting $r$'s approval.

The punishment phase also needs to be modified for general stage games since the repetition of a static Nash equilibrium does not necessarily give a sufficiently severe punishment. In the proof in Appendix, we adapt the punishment scheme of Fudenberg and Maskin [12] to our basic construction. Specifically, we introduce *minmax* periods, in which players play a perturbed minmax action profile where all players randomize except for the minmaxed player. The perturbation is introduced to create observational incentives. As in cooperation periods, with a positive probability, observation is prescribed, in which case the play returns to an examination period. However, in the continuation strategy, the target payoff profile $v^*$ (and the selection rule for $\rho$) is modified to make the minmaxing players indifferent about their actions in the minmax period and motivate them to randomize. This is possible since these players are observed when the play exits from the minmax period.

We conclude this section by noting that, for general stage games, the set $V^*$ may be a proper subset of feasible and individually rational payoff vectors in our class of repeated games. To see this, note that the set of feasible payoff vectors in our context is

$$\bar{V} \equiv \{(v_i - p_i \lambda_i(N \setminus \{i\}))_{i \in N} : v \in V \text{ and } p \in [0,1]^N\},$$

which is a superset of $V$ since $V$ deals only with the case where $p_i = 0$ for all $i$. While any $v \in V$ is feasible, players can also decrease their payoffs by paying observation costs, and the reduced payoff vector may not be in $V$. Our proof relies on a strategy profile that works only if the frequency of monitoring is close to zero,

---

[13]If the maximization has multiple solutions, choose one that maximizes player $k$'s stage-game payoff.

and it is not straightforward to modify the strategy to accommodate payoff profiles in $\bar{V} \setminus V$.[14]


## 5 Conclusion

There are a few assumptions that contribute to the extreme conclusion. First, we assume that it is possible to observe other players' actions without errors (e.g., obtaining hard evidence), although the cost of doing so can be arbitrarily large. If only noisy observations can be obtained at any cost, coordination is difficult and our construction does not extend easily. As mentioned in the introduction, Kandori and Obara [17] study a repeated prisoners' dilemma where observations are almost but not perfectly accurate, under the assumption that observation costs are not large and observational decisions are also observable with small errors. Second, the existence of middle-of-period public randomization plays a critical role in controlling the expected per-period observation costs. However, public random variables are abundant in real life and the use of them does not appear to be difficult.

Even if our assumptions are satisfied, a few concerns remain. First, the strategy profile in the proof may be too complex in the sense that it is difficult to believe that the carefully designed strategy profile emerges from some adjustment process. The strategy profile, however, has a rather intuitive recursive structure and a clear interpretation. Therefore, if the players discuss their collusion scheme explicitly in period 0, the strategy profile may not be too difficult for actual implementation. Finally, the critical discount factor may be too high, particularly when the observation cost is high. The theory is silent about the issue. A contribution of this paper is to confirm that the folk theorem is extremely robust.

The main message of our folk theorem is that accurate observation *per se*, however costly, enables efficient cooperation in general long-term relationships. The idea may give general insights in a wide range of applications. It may, for instance, provide guidance on how to avoid conflicts and facilitate cooperation among countries in the absence of a world government. The idea may also shed light on why only certain species of living creatures can cooperate. As mentioned in the introduction, the conclusion also poses a question on antitrust policies that rely on evidence of explicit communication.

---

[14]Another reason for the difference between $V^*$ and the set of feasible and individually rational payoff vectors is that the minmax value $\underline{u}_i$ is defined under the assumption that the other players randomize independently. Since actions and signals are private information, the other players can actually make their actions appear correlated to the player being punished, and the minmax value in correlated actions may be lower than that in mixed actions. For the idea of using private signals to induce correlations in repeated games, see Lehrer [19].

## Acknowledgments

## A  Appendix: Proof

### A.1  Preliminaries

We use the sup metric for Euclidean spaces: for all $v, w \in \mathbb{R}^{\ell}$, $\|v - w\| \equiv \max_{i \in \{1, \dots, \ell\}} |v_i - w_i|$. For all $\varepsilon > 0$ and all $v \in \mathbb{R}^N$, let $\bar{N}_{\varepsilon}(v)$ denote the closed $\varepsilon$-neighborhood of $v$.

Let $v^* \in \mathrm{int} V^*$ be an arbitrarily chosen target payoff profile. Then there exists $\varepsilon > 0$ such that

$$\bar{N}_{4\varepsilon}(v^*) \subseteq V^*, \tag{10}$$

$$3\varepsilon < \min_{i \in N} \left[ v_i^* - \underline{u}_i \right]. \tag{11}$$

Let $D \in \mathbb{R}$ be defined by

$$D \equiv \max_{i \in N} \left[ \lambda_i(N \setminus \{i\}) + \max_{a, a' \in A} |u_i(a) - u_i(a')| \right] > 0.$$

Then there exist $q > 0$ and $\bar{\eta} \in (0, 0.5)$ such that

$$qD < \varepsilon(1 - q), \tag{12}$$

$$2Dn\bar{\eta} \left[ 1 + \frac{D}{q\varepsilon} \right] < \varepsilon. \tag{13}$$

The following lemma defines a mixed action profile $\alpha^a$ for each $a \in A$. In the equilibrium we will construct, players play this mixed action profile in the cooperative stage when $a$ is chosen as the cooperative action profile.

**Lemma.** *For all $a \in A$, there exists a mixed action profile $\alpha^a$ such that for all $i \in N$:*

*(i) Either $\alpha_i^a = a_i$ or*

$$\alpha_i^a = (1 - \eta_i^a) \cdot a_i + \eta_i^a \cdot d_i^a \tag{14}$$

*where $d_i^a \neq a_i$ and $0 < \eta_i^a \leqslant \bar{\eta}$.*

*(ii) If $\alpha_i^a = a_i$, then $a_i$ is a best response to $\alpha_{-i}^a$.*

*(iii) If $\alpha_i^a \neq a_i$, i.e., $\alpha_i^a$ is given by (14), then $u_i(d_i^a, \alpha_{-i}^a) > u_i(a_i, \alpha_{-i}^a)$.*

*Proof.* Fix a pure action profile $a \in A$. For a given mixed action profile $\alpha \in \mathscr{A}$, define a set $Dev(\alpha)$ by

$$Dev(\alpha) \equiv \{i \in N : \max_{a_i' \in A_i} u_i(a_i', \alpha_{-i}) > u_i(a_i, \alpha_{-i})\}.$$

This is the set of players for whom $a_i$ is not a best response to $\alpha_{-i}$. Let $D^0 \equiv Dev(a)$. If $D^0 = \emptyset$, then we are done by setting $\alpha^a = a$. So suppose $D^0 \neq \emptyset$. Then for all $i \in D^0$, there exists an action $d_i^a \in A_i$ such that $u_i(d_i^a, a_{-i}) > u_i(a)$. Thus for any mixed action profile $\hat{\alpha}$ such that $\|\hat{\alpha} - a\|$ is sufficiently small, we have $u_i(d_i^a, \hat{\alpha}_{-i}) > u_i(a_i, \hat{\alpha}_{-i})$ for all $i \in D^0$. Hence, there exists $\eta_i^a \in (0, \bar{\eta}]$ for all $i \in D^0$, such that the mixed action profile $\alpha^1$ defined by

$$\alpha_i^1 \equiv \begin{cases} (1 - \eta_i^a) \cdot a_i + \eta_i^a \cdot d_i^a & \text{if } i \in D^0 \\ a_i & \text{if } i \notin D^0 \end{cases}$$

satisfies $u_i(d_i^a, \alpha_{-i}^1) > u_i(a_i, \alpha_{-i}^1)$ for all $i \in D^0$. Thus $D^0 \subseteq Dev(\alpha^1) \equiv D^1$. If $D^1 = D^0$, we are done by setting $\alpha^a = \alpha^1$. So suppose otherwise. Then for all $i \in D^1 \setminus D^0$, there exists an action $d_i^a \in A_i$ such that $u_i(d_i^a, \alpha_{-i}^1) > u_i(a_i, \alpha_{-i}^1)$. Therefore, for any mixed action profile $\hat{\alpha}$ such that $\|\hat{\alpha} - \alpha^1\|$ is sufficiently small, $u_i(d_i^a, \hat{\alpha}_{-i}) > u_i(a_i, \hat{\alpha}_{-i})$ for all $i \in D^1$. Hence, there exists $\eta_i^a \in (0, \bar{\eta}]$ for all $i \in D^1 \setminus D^0$, such that the mixed action profile $\alpha^2$ defined by

$$\alpha_i^2 \equiv \begin{cases} (1 - \eta_i^a) \cdot a_i + \eta_i^a \cdot d_i^a & \text{if } i \in D^1 \\ a_i & \text{if } i \notin D^1 \end{cases}$$

satisfies $u_i(d_i^a, \alpha_{-i}^2) > u_i(a_i, \alpha_{-i}^2)$ for all $i \in D^1$. Thus $D^1 \subseteq Dev(\alpha^2) \equiv D^2$. Since the number of players is finite, repeating this procedure yields $k$ such that $D^{k+1} = D^k$. We then set $\alpha^a = \alpha^{k+1}$. $\qquad\square$

For all $a \in A$, let $D^a$ be the set of randomizing players in $\alpha^a$. Define

$$\underline{\eta}^a \equiv \min_{i \in D^a} \eta_i^a \qquad (\text{set } \underline{\eta}^a = 1 \text{ if } D^a = \emptyset),$$

$$\underline{\eta}^1 \equiv \min_{a \in A} \underline{\eta}^a > 0. \tag{15}$$

We now modify the minmax profiles slightly in such a way that all players except for the minmaxed player randomize over all actions. Specifically, for all $i \in N$, there exists a mixed action profile $m^i \in \mathscr{A}$ such that

$$\text{supp}(m_j^i) = A_j \quad \text{for all } j \neq i, \tag{16}$$

$$u_i(m^i) = \max_{a_i \in A_i} u_i(a_i, m_{-i}^i) < \underline{u}_i + \varepsilon, \tag{17}$$

$$m_i^i \in A_i \quad (\text{i.e., } m_i^i \text{ is a pure action}), \tag{18}$$

22

where "supp" denotes the support of the probability distribution. By (16),

$$\underline{\eta}^2 \equiv \min_{i \in N} \min_{j \neq i} \min_{a_j \in A_j} m_j^i(a_j) > 0. \tag{19}$$

Assumption 2 implies that for all $k \in N$, there exist a completely mixed action profile $\beta_{-k}^k$, a player $r(k) \in N \setminus \{k\}$, and a partition of $\Omega_{r(k)}$, $\{\Omega_{r(k)}^{1,k}, \Omega_{r(k)}^{2,k}\}$, such that

$$\underset{b_k \in A_k}{\text{Argmax}} P_{r(k)}(\Omega_{r(k)}^{1,k} \mid b_k, \beta_{-k}^k) \bigcap \underset{b_k \in A_k}{\text{Argmax}} P_{r(k)}(\Omega_{r(k)}^{2,k} \mid b_k, \beta_{-k}^k) = \emptyset. \tag{20}$$

The player $r(k)$ will be called the *referee* for $k$. Let $B_k^1$ and $B_k^2$ be defined by

$$B_k^1 \equiv \underset{b_k \in A_k}{\text{Argmax}} P_{r(k)}(\Omega_{r(k)}^{1,k} \mid b_k, \beta_{-k}^k), \qquad B_k^2 \equiv \underset{b_k \in A_k}{\text{Argmax}} P_{r(k)}(\Omega_{r(k)}^{2,k} \mid b_k, \beta_{-k}^k).$$

By (20), $B_k^1 \cap B_k^2 = \emptyset$. Let $b_k^1$ and $b_k^2$ be such that

$$b_k^1 \in \underset{b_k \in B_k^1}{\text{Argmax}} u_k(b_k, \beta_{-k}^k), \qquad b_k^2 \in \underset{b_k \in B_k^2}{\text{Argmax}} u_k(b_k, \beta_{-k}^k).$$

The definitions of $B_k^1$ and $B_k^2$ imply

$$\begin{aligned}
L_k^1 &\equiv P_{r(k)}(\Omega_{r(k)}^{1,k} \mid b_k^1, \beta_{-k}^k) - \max_{b_k \notin B_k^1} P_{r(k)}(\Omega_{r(k)}^{1,k} \mid b_k, \beta_{-k}^k) > 0, \\
L_k^2 &\equiv P_{r(k)}(\Omega_{r(k)}^{2,k} \mid b_k^2, \beta_{-k}^k) - \max_{b_k \notin B_k^2} P_{r(k)}(\Omega_{r(k)}^{2,k} \mid b_k, \beta_{-k}^k) > 0.
\end{aligned} \tag{21}$$

Let $L \in \mathbb{R}$ be defined by

$$L \equiv \min_{k \in N} \min\{L_k^1, L_k^2\} > 0. \tag{22}$$

Since $\beta_{-k}^k$ is completely mixed, we have

$$\underline{\eta}^3 \equiv \min_{k \in N} \min_{j \neq k} \min_{a_j \in A_j} \beta_j^k(a_j) > 0. \tag{23}$$

Let $\underline{\eta} \in \mathbb{R}$ be defined by

$$\underline{\eta} \equiv \min\{\underline{\eta}^1, \underline{\eta}^2, \underline{\eta}^3, 0.1\} > 0. \tag{24}$$

Let $\underline{p} \in \mathbb{R}$ be defined by

$$\underline{p} \equiv \min_{i \in N} \min_{\omega_i \in \Omega_i} \min_{a \in A} P_i(\omega_i \mid a) > 0,$$

23

where the inequality holds by Assumption 1. Let $(\mu, \underline{\delta}) \in (0,1)^2$ be sufficiently close to $(0,1)$ so that for all $\delta \in [\underline{\delta}, 1)$,

$$\hat{\mu} \equiv \frac{1-\delta}{\delta} \frac{1-q}{q} < 1, \tag{25}$$

$$n(1-\delta)D < 0.5\mu\delta q\varepsilon, \tag{26}$$

$$D\mu < 0.5\varepsilon, \tag{27}$$

$$n(n-1)D(1-\delta+\mu\delta) < 0.05\delta^3 \underline{\eta}\, \underline{p} L\varepsilon. \tag{28}$$

In what follows, we fix $\delta \geqslant \underline{\delta}$.

For all $j \in N$, let $W^j \in \mathbb{R}^N$ be defined by

$$W_i^j \equiv \begin{cases} (1-q)(v_i^* - \varepsilon) + q\big[u_i(m^i) - \hat{\mu}\lambda_i(N \setminus \{i\})\big] & \text{if } i = j, \\[2mm] (1-q)(v_i^* + \varepsilon) + q\Big[\min\limits_{\hat{a} \in \text{supp}(m^j)} u_i(\hat{a}) - \hat{\mu}\lambda_i(N \setminus \{i,j\})\Big] & \text{if } i \neq j. \end{cases}$$

For all $v \in \bar{N}_\varepsilon(v^*)$, all $a \in A$, and all $i \in D^a$, let $\xi_i^v(a) \in \mathbb{R}$ be defined by

$$(1-\delta)\big[u_i(d_i^a, \alpha_{-i}^a) - u_i(a_i, \alpha_{-i}^a)\big] = \frac{1}{|D^a|}\xi_i^v(a)\mu\delta(v_i - W_i^i). \tag{29}$$

To see that $\xi_i^v(a)$ is a probability (i.e., falls in between 0 and 1), note that

$$v_i - W_i^i \geqslant v_i^* - \varepsilon - W_i^i > q(v_i^* - u_i(m^i) - \varepsilon) > q(v_i^* - \underline{u}_i - 2\varepsilon) > q\varepsilon \tag{30}$$

by (11) and (17). This implies $0 < \xi_i^v(a) \leqslant n(1-\delta)D/(\mu\delta q\varepsilon) < 0.5$ by (26).

For all $v \in \bar{N}_\varepsilon(v^*)$, we define a *virtual payoff function* $u^v : A \to \mathbb{R}^N$ as

$$u_i^v(a) \equiv u_i(\alpha^a) - \mu\lambda_i(D^a \setminus \{i\})$$
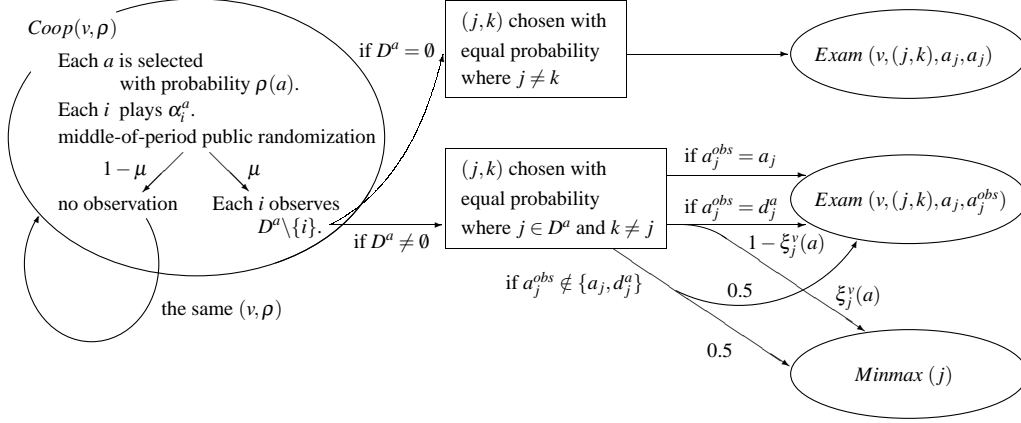$$- \sum_{j \in D^a} \eta_j^a \big[u_j(d_j^a, \alpha_{-j}^a) - u_j(a_j, \alpha_{-j}^a)\big]\frac{v_i - W_i^j}{v_j - W_j^j} \tag{31}$$

for all $i$. For all $v \in \bar{N}_\varepsilon(v^*)$ and all $a \in A$,

$$|u_i^v(a) - u_i(a)| \leqslant |u_i(\alpha^a) - u_i(a)| + \mu D + n\bar{\eta}D\frac{D}{q\varepsilon}$$
$$\leqslant D\Big[n\bar{\eta} + \mu + \frac{n\bar{\eta}D}{q\varepsilon}\Big]$$
$$< \varepsilon$$

by (13) and (27). This implies that there exists $\rho^* \in \Delta(A)$ such that

$$u^{v^*}(\rho^*) = v^*. \tag{32}$$

24

**Figure 3:** Cooperation states

For all $h \in N$ and all $a \in \text{supp}(m^h)$, let $V^h(a) \in \mathbb{R}^N$ be defined by

$$V_i^h(a) \equiv \begin{cases} v_i^* - \varepsilon & \text{if } i = h, \\ v_i^* + \varepsilon - \frac{q}{1-q}\left[u_i(a) - \min_{\hat{a}\in\text{supp}(m^h)} u_i(\hat{a})\right] & \text{if } i \neq h. \end{cases} \tag{33}$$

Then for all $i \neq h$, $v_i^* + \varepsilon \geqslant V_i^h(a) > v_i^*$ by (12). This implies

$$V^h(a) \in \bar{N}_\varepsilon(v^*).$$

Thus if we define $V^{**} \subseteq \mathbb{R}^N$ by

$$V^{**} \equiv \{v^*\} \cup \{V^h(a) : h \in N \text{ and } a \in \text{supp}(m^h)\},$$

then $V^{**}$ is a finite set and $V^{**} \subseteq \bar{N}_\varepsilon(v^*)$.

*A.2 Strategy*

We now construct a strategy profile that yields the target payoff profile $v^*$ and is a sequential equilibrium under $\delta \geqslant \underline{\delta}$. The strategy has four types of states: *cooperation*, *examination*, *report*, and *minmax*.

*Cooperation States* (Figure 3). A cooperation state, denoted *Coop*$(v,\rho)$, is indexed by a payoff profile $v \in V^{**}$ and a distribution $\rho \in \Delta(A)$. In particular, the initial period is in state *Coop*$(v^*,\rho^*)$ where $\rho^*$ is given by (32). In each period of state *Coop*$(v,\rho)$, the public randomization at the beginning of the period selects each $a \in A$ with probability $\rho(a)$ as the cooperative action profile of the period. Suppose $\rho$ selects $a$. Then players play the mixed action profile $\alpha^a$. Observational
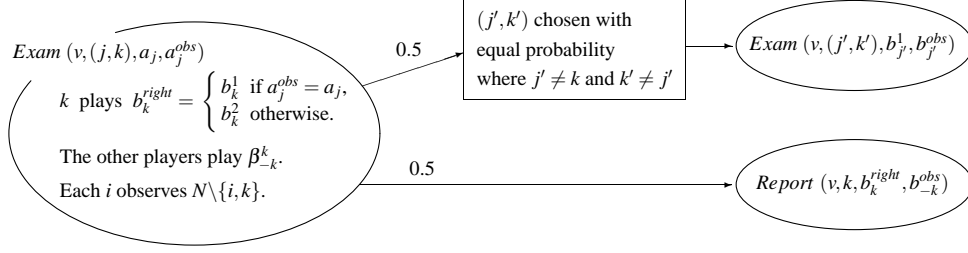
**Figure 4:** Examination states

decisions are coordinated by the public randomization in the middle of the period. If $y_t \geq \mu$, no one observes others, and the state remains $Coop(v,\rho)$ in the next period.

If $y_t < \mu$, on the other hand, each player $i$ observes all players in $D^a \setminus \{i\}$. If $D^a = \emptyset$, then the public randomization at the beginning of the next period selects a pair of players $(j,k)$ such that $j \neq k$ with equal probability, and the next period is in $Exam(v,(j,k),a_j,a_j)$. If $D^a \neq \emptyset$, let $(a_i^{obs})_{i \in D^a}$ denote the realized action profile of players in $D^a$. The public randomization at the beginning of the next period selects a pair of players $(j,k)$ such that $j \in D^a$ and $k \neq j$ with equal probability. If $a_j^{obs} = d_j^a$, then the state changes to $Minmax(j)$ with probability $\xi_j^v(a)$. If $a_j^{obs} \notin \{a_j, d_j^a\}$, the state changes to $Minmax(j)$ with probability 0.5. In all these cases (where $D^a \neq \emptyset$), if the state does not change to $Minmax(j)$, it changes to $Exam(v,(j,k),a_j,a_j^{obs})$.

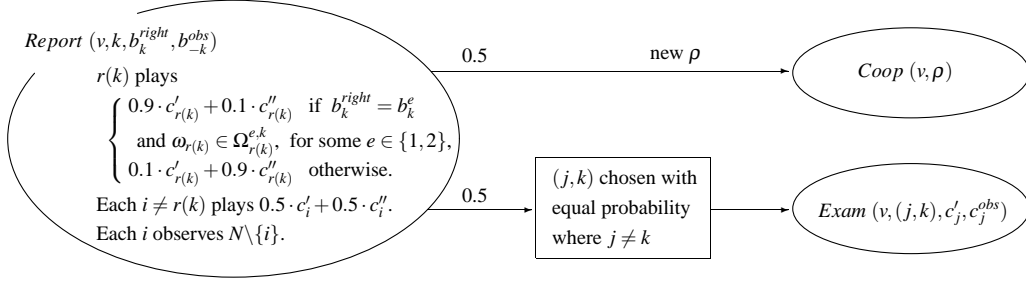*Examination States* (Figure 4). An examination state, denoted $Exam(v,(j,k),a_j,a_j^{obs})$, is indexed by a payoff profile $v \in V^{**}$, a pair of players $(j,k)$ such that $j \neq k$, $a_j \in A_j$, and $a_j^{obs} \in A_j$. In this state, player $k$ is prescribed to play a pure action $b_k^{right}$ determined by

$$b_k^{right} = \begin{cases} b_k^1 & \text{if } a_j^{obs} = a_j, \\ b_k^2 & \text{otherwise.} \end{cases}$$

The other players play a mixed action profile $\beta_{-k}^k$. Regardless of the public randomization in the middle of the period, each player $i$ observes $N \setminus \{i,k\}$. Let $b_{-k}^{obs} \in A_{-k}$ denote the realized action profile of players $N \setminus \{k\}$. The state transition depends on the public randomization at the beginning of the next period. With probability 0.5, a pair $(j',k')$ such that $j' \neq k$ and $k' \neq j'$ is selected with equal probability and the state changes to $Exam(v,(j',k'),b_{j'}^1,b_{j'}^{obs})$.[15] With the remaining probability, the state changes to $Report(v,k,b_k^{right},b_{-k}^{obs})$.

*Report States* (Figure 5). A report state, denoted $Report(v,k,b_k^{right},b_{-k}^{obs})$, is indexed by a payoff profile $v \in V^{**}$, a player $k \in N$, and an action profile

---

[15]Here the choice of a particular action $b_{j'}^1$ is arbitrary; any other action works since $\beta_{j'}^k$ assigns positive probability to all actions.

**Figure 5:** Report states

$(b_k^{right}, b_{-k}^{obs}) \in \{b_k^1, b_k^2\} \times A_{-k}$. Player $k$ is the one who was under examination in the last period (by construction, the previous period is an examination period). For each player $i$, choose a pair of distinct actions $\{c_i', c_i''\} \subseteq A_i$ arbitrarily in advance. In this period, player $r(k)$, i.e., the referee for $k$ defined in (20), plays the following mixed action:

$$
\begin{aligned}
0.9 \cdot c_{r(k)}' + 0.1 \cdot c_{r(k)}'' \qquad & \text{if } b_k^{right} = b_k^1 \text{ and } \omega_{r(k)} \in \Omega_{r(k)}^{1,k}, \\
0.9 \cdot c_{r(k)}' + 0.1 \cdot c_{r(k)}'' \qquad & \text{if } b_k^{right} = b_k^2 \text{ and } \omega_{r(k)} \in \Omega_{r(k)}^{2,k}, \\
0.1 \cdot c_{r(k)}' + 0.9 \cdot c_{r(k)}'' \qquad & \text{if } b_k^{right} = b_k^1 \text{ and } \omega_{r(k)} \notin \Omega_{r(k)}^{1,k}, \\
0.1 \cdot c_{r(k)}' + 0.9 \cdot c_{r(k)}'' \qquad & \text{if } b_k^{right} = b_k^2 \text{ and } \omega_{r(k)} \notin \Omega_{r(k)}^{2,k},
\end{aligned}
$$

where $\omega_{r(k)} \in \Omega_{r(k)}$ denotes the referee's private signal in the previous period. Any other player $i \neq r(k)$ plays a mixed action $0.5 \cdot c_i' + 0.5 \cdot c_i''$.

Regardless of the public randomization in the middle of the period, all players observe all the other players. Let $c^{obs} \in A$ denote the realized action profile (possibly $c_i^{obs} \notin \{c_i', c_i''\}$ for some players). The state transition depends on the public randomization at the beginning of the next period. With probability 0.5, a new pair of players $(j,k)$ such that $j \neq k$ is chosen with equal probability and the state changes to $Exam(v,(j,k),c_j',c_j^{obs})$. With the remaining probability, the state changes to $Coop(v,\rho)$ where $\rho \in \Delta(A)$ depends on $(v,k,b_k^{right},b_{-k}^{obs},c^{obs})$ and is determined as follows.

First, let $v' \in \mathbb{R}^N$ be defined by

$$
\begin{aligned}
v_i' = v_i &+ \frac{1-\delta+\mu\delta}{(1/4)\delta^2}\Big[v_i - u_i(b_k^{right}, b_{-k}^{obs}) + \lambda_i(N \setminus \{i,k\})\Big] \\
&+ \frac{1-\delta+\mu\delta}{(1/2)\delta}\Big[v_i - u_i(c^{obs}) + \lambda_i(N \setminus \{i\})\Big].
\end{aligned}
\tag{34}
$$

By (28),

$$
|v_i' - v_i| \leqslant \frac{(1-\delta+\mu\delta)D(3/2)}{(1/4)\delta^2} < \varepsilon.
$$

27
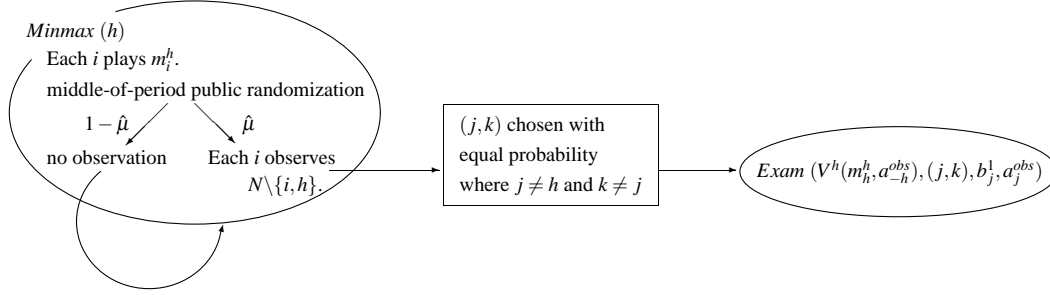
**Figure 6:** Minmax states

Thus $v' \in \bar{N}_\varepsilon(v) \subseteq \bar{N}_{2\varepsilon}(v^*)$. Now, we choose a distribution $\rho \in \Delta(A)$ such that

$$u_i^v(\rho) = v_i' \qquad\qquad\qquad \text{if } i \neq k, \tag{35}$$

$$u_k^v(\rho) = \begin{cases} v_k' + \varepsilon(1 - AP_k) & \text{if } c_{r(k)}^{obs} = c_{r(k)}', \\ v_k' - \varepsilon AP_k & \text{otherwise,} \end{cases} \tag{36}$$

where

$$AP_k \equiv \begin{cases} 0.9 P_{r(k)}(\Omega_{r(k)}^{1,k} \mid b_k^1, \beta_{-k}^k) + 0.1 P_{r(k)}(\Omega_{r(k)}^{2,k} \mid b_k^1, \beta_{-k}^k) & \text{if } b_k^{right} = b_k^1, \\ 0.9 P_{r(k)}(\Omega_{r(k)}^{2,k} \mid b_k^2, \beta_{-k}^k) + 0.1 P_{r(k)}(\Omega_{r(k)}^{1,k} \mid b_k^2, \beta_{-k}^k) & \text{if } b_k^{right} = b_k^2 \end{cases}$$

denotes the ex ante probability that $k$'s referee $r(k)$ plays $c_{r(k)}'$ in equilibrium. To see that $\rho$ exists, note that by construction, $u^v(\rho)$ is within $\varepsilon$ of $v'$ and so within $3\varepsilon$ of $v^*$. Since $u^v$ is within $\varepsilon$ of $u$, it follows that $u(\rho)$ is within $4\varepsilon$ of $v^*$. Hence, $\rho$ exists by (10).

*Minmax States* (Figure 6). A minmax state, denoted *Minmax*(*h*), is indexed by a player $h \in N$ who is to be punished. In this state, players play the modified minmax action profile $m^h$ (see (16)–(18)). Observational decisions are coordinated by the public randomization in the middle of the period. If $y_t \geq \hat{\mu}$, where $\hat{\mu}$ is defined in (25), no one observes others and the state remains the same in the next period.

If $y_t < \hat{\mu}$, on the other hand, each player $i$ observes $N \setminus \{i, h\}$. Let $a_{-h}^{obs} \in A_{-h}$ denote the realized action profile of players $i \neq h$. The public randomization at the beginning of the next period chooses a pair $(j, k)$ such that $j \neq h$ and $k \neq j$, and the state changes to

$$Exam(V^h(m_h^h, a_{-h}^{obs}), (j, k), b_j^1, a_j^{obs}),$$

where $V^h$ is defined by (33).[16]

---

[16]Again, $b_j^1$ was chosen arbitrarily as the determinant of $b_k^{right}$ for the examination stage. Any action will do since each player $j \neq h$ plays a completely mixed action in *Minmax*(*h*).

We have specified the strategy profile on the equilibrium path. To complete the specification of the strategy profile, we first add the following rules. (i) The prescribed observational decision for a player does not depend on the stage-game action he chose in the period. That is, a player's own deviation in terms of stage-game action does not change the prescribed observational decision for the player in the period. (ii) The prescribed behavior (action and observation) for a player does not depend on any information he obtained by observing players whom he was not prescribed to observe. That is, if a player $i$ observed a deviation of a player $j$ in a period when player $i$ was not prescribed to observe $j$, then $i$ is prescribed to ignore the deviation and behave as if he did not observe it.

Let $\hat{\sigma}$ be a strategy profile that follows the state-dependent play described above and satisfies rules (i) and (ii). Consider a sequence of behavioral strategy profiles $(\hat{\sigma}^k)_{k=1}^{\infty}$ with $\hat{\sigma}^k \to \hat{\sigma}$, such that each $\hat{\sigma}^k$ puts a positive probability to every move but puts far smaller weights on the trembles with respect to observational decisions than those with respect to actions. This sequence generates a sequence of belief systems $(\psi^1, \psi^2, \dots)$ whose limit $\psi$ is such that, at any history, each player believes that the other players have not deviated with respect to observational decisions.

For each player $i$, let $\hat{H}_i$ be the set of $i$'s (private) histories throughout which $i$ did observe every player he was prescribed to observe under the state-dependent play (with rules (i) and (ii)). Thus $\hat{H}_i$ includes histories in which $i$ deviated in terms of action, as well as histories in which $i$ observed players he was not prescribed to observe. It should be noted that, at all histories $h_i \in \hat{H}_i$, player $i$ knows the current state and can follow the state-dependent play.

For each player $i$, let $\sigma_i^*$ be a strategy that agrees with $\hat{\sigma}_i$ on $\hat{H}_i$ and such that, at all histories outside $\hat{H}_i$, the player plays a best response given the belief $\psi$ and given that the other players follow $\hat{\sigma}_{-i}$. Let $\sigma^* = (\sigma_1^*, \dots, \sigma_n^*)$. We show that $(\sigma^*, \psi)$ is a desired sequential equilibrium. To see that $\psi$ is consistent with $\sigma^*$, consider a sequence of behavioral strategy profiles $(\sigma^k)_{k=1}^{\infty}$ with $\sigma^k \to \sigma^*$, such that, for each $k$, $\sigma_i^k$ agrees with $\hat{\sigma}_i^k$ on $\hat{H}_i$ and, as before, puts a positive probability to every move but far smaller weights on the trembles with respect to observational decisions than those with respect to actions. Then, the associated sequence of belief systems also converges to $\psi$. In what follows, we show that $\sigma^*$ attains the target payoff profile $v^*$ and is sequentially rational given $\psi$.

### A.3 Values

In this section, we show that the strategy profile $\sigma^*$ attains the target payoff profile $v^*$. To do so, we also show that the continuation value from a state is given by

$$Coop(v, \rho) : \; \big[(1-\delta)u^v(\rho) + \mu\delta v\big]/(1-\delta+\mu\delta),$$
$$Exam(v, (j,k), a_j, a_j^{obs}) : \; v,$$
$$Minmax(h) : \; W^h.$$

To compute the continuation value for each state, we need to solve a system of equations. Since the set of states is finite in equilibrium (because the number of distributions $\rho$ used in the cooperation states is finite in equilibrium), the solution is unique. To identify the solution, we first assume that the continuation value from any state of the form $Exam(v, \cdot)$ is exactly equal to $v$. We then show that this indeed constitutes a solution.

Given the assumption, we first compute the continuation value from minmax states. Given $h \in N$, let $M(h) \in \mathbb{R}^N$ denote the continuation payoff profile at the beginning of the state $Minmax(h)$. The continuation payoff for player $h$ is given by

$$M_h(h) = (1 - \delta)\left[u_h(m^h) - \hat{\mu}\lambda_h(N \setminus \{h\})\right] + (1 - \hat{\mu})\delta M_h(h) + \hat{\mu}\delta(v_h^* - \varepsilon).$$

Since $\hat{\mu}\delta = (1 - \delta)(1 - q)/q$ by definition, reorganizing the equation gives

$$M_h(h) = W_h^h.$$

To compute the continuation payoff for players $i \neq h$, let $M_i^a(h)$ denote the continuation payoff of player $i$ evaluated at the beginning of the state given that $a \in \text{supp}(m^h)$ is the realized action profile in this period. Then

$$M_i^a(h) = (1 - \delta)\left[u_i(a) - \hat{\mu}\lambda_i(N \setminus \{i, h\})\right] + (1 - \hat{\mu})\delta M_i(h) + \hat{\mu}\delta V_i^h(a).$$

Using $\hat{\mu}\delta = (1 - \delta)(1 - q)/q$ and substituting the definition of $V_i^h(a)$ in (33) give

$$\begin{aligned}
M_i^a(h) = (1 - \delta)&\left[\min_{\hat{a} \in \text{supp}(m^h)} u_i(\hat{a}) - \hat{\mu}\lambda_i(N \setminus \{i, h\})\right] \\
&+ (1 - \hat{\mu})\delta M_i(h) + \hat{\mu}\delta(v_i^* + \varepsilon).
\end{aligned} \tag{37}$$

This implies that $M_i^a(h)$ does not depend on $a$ and hence $M_i^a(h) = M_i(h)$ for all $a$. Substituting this fact into (37) yields

$$M_i(h) = W_i^h \quad \text{for all } i \neq h.$$

Thus $M(h) = W^h$ for any $h \in N$.

Abusing notation, let $Coop(v, \rho) \in \mathbb{R}^N$ denote the continuation payoff profile at the beginning of the state $Coop(v, \rho)$. Then

$$\begin{aligned}
Coop_i(v, \rho) = (1 - \delta)\sum_{a \in A} \rho(a)\left[u_i(\alpha^a) - \mu\lambda_i(D^a \setminus \{i\})\right] + (1 - \mu)\delta Coop_i(v, \rho) \\
+ \mu\delta v_i - \mu\delta\sum_{a \in A} \rho(a)\sum_{j \in D^a} \frac{1}{|D^a|}\eta_j^a\xi_j^v(a)(v_i - W_i^j).
\end{aligned}$$

Substituting the definition of $\xi_j^v(a)$ (see (29)) yields

$$\begin{aligned}
(1 - \delta + \mu\delta)Coop_i(v, \rho) = (1 - \delta)\sum_{a \in A} \rho(a)\Bigg\{u_i(\alpha^a) - \mu\lambda_i(D^a \setminus \{i\}) \\
- \sum_{j \in D^a} \eta_j^a\left[u_j(d_j^a, \alpha_{-j}^a) - u_j(a_j, \alpha_{-j}^a)\right]\frac{v_i - W_i^j}{v_j - W_j^j}\Bigg\} + \mu\delta v_i.
\end{aligned}$$

30

Using virtual payoff functions $u^v$ defined in (31) and writing $u^v(\rho) = \sum_{a \in A} \rho(a) u^v(a)$, we obtain

$$Coop(v, \rho) = \frac{(1-\delta)u^v(\rho) + \mu\delta v}{1 - \delta + \mu\delta}. \tag{38}$$

Then by the definition of $\rho^*$ given by (32),

$$Coop(v^*, \rho^*) = v^*.$$

Since the initial state is $Coop(v^*, \rho^*)$, the target payoffs $v^*$ are indeed achieved as the repeated-game payoffs under $\sigma^*$.

We now verify that the continuation value from an examination state of the form $Exam(v, \cdot)$ is indeed $v$. Consider an examination state $Exam(v, (j,k), a_j, a_j^{obs})$. We need to show that for all $i \in N$,

$$
\begin{aligned}
v_i = E \Big\{ & (1-\delta)\big[u_i(b_k^{right}, b_{-k}^{obs}) - \lambda_i(N \setminus \{i,k\})\big] \\
& + \tfrac{1}{2}\delta v_i + \tfrac{1}{2}\delta(1-\delta)\big[u_i(c^{obs}) - \lambda_i(N \setminus \{i\})\big] \\
& + \tfrac{1}{4}\delta^2 v_i + \tfrac{1}{4}\delta^2 Coop_i(v, \rho) \Big\},
\end{aligned} \tag{39}
$$

where the expectation is taken over $(b_{-k}^{obs}, c^{obs})$, and $\rho$ is determined from $(b_{-k}^{obs}, c^{obs})$ by (34)–(36). For player $i \neq k$, substituting (34) into (35) and using (38) to replace $u^v$ with $Coop_i(v, \rho)$ yield

$$
\begin{aligned}
v_i = & (1-\delta)\big[u_i(b_k^{right}, b_{-k}^{obs}) - \lambda_i(N \setminus \{i,k\})\big] \\
& + \tfrac{1}{2}\delta v_i + \tfrac{1}{2}\delta(1-\delta)\big[u_i(c^{obs}) - \lambda_i(N \setminus \{i\})\big] \\
& + \tfrac{1}{4}\delta^2 v_i + \tfrac{1}{4}\delta^2 Coop_i(v, \rho)
\end{aligned} \tag{40}
$$

for all $b_{-k}^{obs} \in A_{-k}$ and all $c^{obs} \in A$. Taking the expectation of (40) implies (39). For player $k$, computation is the same except that $u_k^v(\rho) \neq v_k'$. Thus we obtain

$$
\begin{aligned}
v_k = & (1-\delta)\big[u_k(b_k^{right}, b_{-k}^{obs}) - \lambda_k(N \setminus \{k\})\big] \\
& + \tfrac{1}{2}\delta v_k + \tfrac{1}{2}\delta(1-\delta)\big[u_k(c^{obs}) - \lambda_k(N \setminus \{k\})\big] \\
& + \tfrac{1}{4}\delta^2 v_k + \tfrac{1}{4}\delta^2 Coop_k(v, \rho) - \tfrac{1}{4}\delta^2 \frac{1-\delta}{1-\delta+\mu\delta}\big[u_k^v(\rho) - v_k'\big],
\end{aligned}
$$

where the only non-trivial difference from (40) is the last term. However, this term is zero in expectation since the expected value of $u_k^v(\rho) - v_k'$ is $AP_k\varepsilon(1 - AP_k) - (1 - AP_k)\varepsilon AP_k = 0$. Thus (39) also holds for player $k$.

The fact that (40) holds for all $i \neq k$ and all $b_{-k}^{obs} \in A_{-k}$ implies that all players $i \neq k$ are completely indifferent over all actions in the examination period.

## A.4 Incentives

We now show that $\sigma^*$ is sequentially rational given $\psi$. We begin by showing that no player $i$ has an incentive to deviate at any history $h_i \in \hat{H}_i$. Recall that, at histories $h_i \in \hat{H}_i$, player $i$ knows the state, and believes that the other players also know the state and follow the state-dependent play. We start with incentives in terms of stage-game actions.

*Cooperation States.* When the public randomization selects an action profile $a$ as the cooperation action profile, players $i \in D^a$ are prescribed to randomize between $a_i$ and $d_i^a$, and are indeed indifferent between the actions by the definition of $\xi_i^v(a)$. These players also do not have incentives to play any other action $a_i' \notin \{a_i, d_i^a\}$; indeed, the long-run loss is at least

$$\mu \frac{1}{2n} \delta[v_i - W_i^i] > \mu \frac{1}{2n} \delta q \varepsilon$$

by (30), and this exceeds $(1-\delta)D$ by (26). On the other hand, players $i \notin D^a$ have no incentive to deviate from $a_i$ since by Lemma, $a_i$ is a short-run best response to $\alpha_{-i}^a$, and deviations are not observed and have no effects on the future play.

*Report States.* We show that in report states, all players are indifferent among all actions. Fix a report state $Report(v, k, b_k^{right}, b_{-k}^{obs})$. Let $R_i(c_i)$ denote player $i$'s continuation payoff from this period if $i$ chooses $c_i \in A_i$, where the expectation is taken with respect to $c_{-i}$ based on all the information that player $i$ has at the beginning of this period. For players $i \neq k$, (35) and (38) imply

$$R_i(c_i) = (1-\delta)\Big[E\big[u_i(c_i, c_{-i})\big] - \lambda_i(N \setminus \{i\})\Big] \tag{41}$$
$$+ \frac{1}{2}\delta v_i + \frac{1}{2}\delta \frac{(1-\delta)E\big[v_i' \,|\, c_i\big] + \mu \delta v_i}{1 - \delta + \mu \delta},$$

where $v_i'$ depends on $c_{-i}$ through (34). The right-hand side depends on $c_i$ because of the two terms with expectation. But by taking the expectation of (34), we can see that

$$E\big[u_i(c_i, c_{-i})\big] + \frac{1}{2}\delta \frac{E\big[v_i' \,|\, c_i\big]}{1 - \delta + \mu \delta}$$

as a whole does not depend on $c_i$. Thus $R_i(c_i)$ is actually constant in $c_i$. This implies that each player $i \neq k$ is completely indifferent about $c_i$ in this period and therefore willing to randomize as instructed by the strategy.

For player $k$, the argument is the same except that $v_i'$ in (41) for $i = k$ has to be replaced by $u_k^v(\rho)$ since $u_k^v(\rho) \neq v_k'$ for player $k$ by (36). Thus

$$R_k(c_k) = (1-\delta)\Big[E\big[u_k(c_k, c_{-k})\big] - \lambda_k(N \setminus \{k\})\Big]$$
$$+ \frac{1}{2}\delta v_k + \frac{1}{2}\delta \frac{(1-\delta)\Big[E\big[v_k' \,|\, c_k\big] + (AP_k' - AP_k)\varepsilon\Big] + \mu \delta v_k}{1 - \delta + \mu \delta}, \tag{42}$$

where $AP'_k$ denote $k$'s current belief about the probability that his referee $r(k)$ plays $c'_{r(k)}$ in this period.[17] The only difference between (42) and (41) is the term $(AP'_k - AP_k)\varepsilon$, which does not depend on $c_k$. Thus, the previous argument works for player $k$ as well. Hence $R_k(c_k)$ does not depend on $c_k$ and player $k$ is also indifferent about his action.

*Examination States.* At the end of Section A.3, we showed that, in examination periods, players who are not under examination are indifferent over all actions. Thus we now prove that the player under examination (player $k$) is willing to play the pure action prescribed by the strategy (i.e., $b_k^{right}$). Consider an examination state $Exam(v, (j,k), a_j, a_j^{obs})$. Let $\ell \in \{1, 2\}$ be such that $b_k^{right} = b_k^\ell$. Since $b_k^{right}$ is a short-run best response within $B_k^\ell$, it suffices to verify that player $k$ does not gain by playing any $b_k \notin B_k^\ell$. The short-run gains from playing any $b_k \notin B_k^\ell$ are at most $(1 - \delta)D$. On the other hand, by the definition of $B_k^\ell$, playing an action $b_k \notin B_k^\ell$ necessarily lowers the probability that player $r(k)$ receives a signal $\omega_{r(k)} \in \Omega_{r(k)}^{\ell,k}$, at least by $L > 0$ (see (21) and (22)). If $\omega_{r(k)} \notin \Omega_{r(k)}^{\ell,k}$ and if the next period is a report period, then the probability that the referee plays $c'_{r(k)}$ (i.e., approves of $k$'s answer) in the report period goes down from 0.9 to 0.1. If the referee indeed gives a disapproval and if the following period is a cooperation period, then the distribution $\rho$ used in the cooperation period changes in such a way that $u_i^v(\rho)$ goes down by $\varepsilon$ (see (36)), which in turn implies that the continuation value $Coop_i(v, \rho)$ goes down by $(1 - \delta)\varepsilon/(1 - \delta + \mu\delta)$.

Altogether, the long-run losses from playing $b_k \notin B_k^\ell$ are at least

$$L \frac{1}{2}(0.9 - 0.1) \frac{1}{2} \delta^2 \frac{(1 - \delta)\varepsilon}{1 - \delta + \mu\delta}.$$

This exceeds $(1 - \delta)D$ by (28).

*Minmax States.* Consider a state $Minmax(h)$. In this state, player $h$ has no incentive to deviate since the prescribed action $m_h^h$ is a short-run best response against $m_{-h}^h$. The other players $i \neq h$ are willing to play $m_i^h$ since $M_i^a(h)$ does not depend on $a$ and hence the players are completely indifferent.

*Observational Decisions.* We now verify that players have incentives to follow the strategy with respect to observational decisions. First, a player has no incentive to observe players who are not to be observed since they are expected to play a pure action. Suppose now that a player $k \in N$ chooses not to observe a player $j \neq k$ at the end of a period $t$ when the strategy prescribes him otherwise. Since $k$ is prescribed to observe $j$, player $j$ was prescribed to play some mixed action $\alpha_j \in \mathscr{A}_j$ in this period.

---

[17]We may have $AP'_k \neq AP_k$ since $AP_k$ is $k$'s belief at the beginning of the previous period and he has since then updated his belief based on his signal and observations. Moreover, player $k$ may have deviated in stage-game action in the previous period, which is possible since we are considering a history $h_k \in \hat{H}_k$.

By the definition of $\underline{\eta}$, any action in the support of $\alpha_j$ is assigned a probability at least as large as $\underline{\eta} > 0$ (see (15), (19), (23), and (24)). By construction, period $t+1$ is an examination period with a positive probability. The probability is at least as large as

$$\min\{0.5, 1 - \max_{a \in A} \max_{i \in N} \max_{v \in V^{**}} \xi_i^v(a)\} = 0.5 \tag{43}$$

since $\xi_i^v(a) < 0.5$.[18] If period $t+1$ is an examination period, then with at least probability $1/[n(n-1)]$, player $k$ is chosen to be examined and is prescribed to "state" $j$'s realized action in period $t$. In this case, the state in period $t+1$ is of the form $Exam(v,(j,k),a_j,a_j^{obs})$ where $\{a_j,a_j^{obs}\} \subseteq \text{supp}(\alpha_j)$. Since $k$ did not observe $j$, $k$ does not know $a_j^{obs}$ and hence is uncertain of $b_k^{right}$.

In this contingency, there is a positive probability bounded away from 0 with which player $k$ plays a "wrong" action, playing $b_k \notin B_k^\ell$ when $b_k^{right} = b_k^\ell$ ($\ell \in \{1,2\}$). To find a bound, let $\omega_k^t \in \Omega_k$ denote the private signal that $k$ received in the last period. Suppose, without loss of generality, that player $k$ plays an action $b_k \notin B_k^1$ (the case where $b_k \notin B_k^2$ works similarly). This action is "wrong" if $b_k^{right} = b_k^1$, i.e., $a_j^{obs} = a_j$. The conditional probability that $a_j^{obs} = a_j$ holds given that $k$ received $\omega_k^t$ is bounded below by $\underline{\eta}\,\underline{p}$, since $\underline{\eta}$ is the minimum probability assigned to each action in $\text{supp}(\alpha_j)$ and $\underline{p}$ is the minimum probability assigned to each signal $\omega_k \in \Omega_k$.

If player $k$ plays $b_k \notin B_k^\ell$ in the examination period when $b_k^{right} = b_k^\ell$ ($\ell \in \{1,2\}$), then the continuation payoff from the period goes down strictly as we showed above. The loss exceeds the short-run gain from not observing player $j$ in period $t$ if

$$(1-\delta)D < \frac{\underline{\eta}\,\underline{p}\,\delta}{2n(n-1)}\Big[L\frac{1}{2}(0.9-0.1)\frac{1}{2}\delta^2\frac{(1-\delta)\varepsilon}{1-\delta+\mu\delta} - (1-\delta)D\Big].$$

The inequality is satisfied by (28).

*Histories $h_i \notin \hat{H}_i$.*  It remains to consider each player $i$'s incentives at histories $h_i \notin \hat{H}_i$. By definition, the continuation play of $\sigma_i^*$ given $h_i$ prescribes an optimal decision for $i$ at the history given his belief $\psi(h_i)$ and given that the other players follow $\hat{\sigma}_{-i}$. By the construction of $\psi$, player $i$ believes that the other players $j \neq i$ are at some histories $h_j \in \hat{H}_j$, and hence, by the definition of $\sigma^*$, their continuation play coincides with that of $\hat{\sigma}_{-i}$ along any path that $i$ can induce. Therefore, following $\sigma_i^*$ is sequentially rational for player $i$ at $h_i$ given $\psi$.  □

---

[18]The lower bound (43) is valid even if the history is off the equilibrium path. For example, if period $t$ was a cooperation period and player $k$ played $a_k' \notin \{a_k, d_k^a\}$, then the deviation is ignored with probability 0.5 and therefore an exam is held next period with the same probability.

## References

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

[2] M. Aoyagi, Collusion in dynamic Bertrand oligopoly with correlated private signals and communication, J. Econ. Theory 102 (2002) 229–248.

[3] E. Ben-Porath, M. Kahneman, Communication in repeated games with private monitoring, J. Econ. Theory 70 (1996) 281–297.

[4] E. Ben-Porath, M. Kahneman, Communication in repeated games with costly monitoring, Games Econ. Behav. 44 (2003) 227–250.

[5] V. Bhaskar, I. Obara, Belief-based equilibria in the repeated prisoners' dilemma with private monitoring, J. Econ. Theory 102 (2002) 40–69.

[6] O. Compte, Communication in repeated games with imperfect private monitoring, Econometrica 66 (1998) 597–626.

[7] J. Ely, J. Välimäki, A robust folk theorem for the prisoner's dilemma, J. Econ. Theory 102 (2002) 84–105.

[8] J. C. Ely, J. Horner, W. Olszewski, Belief-free equilibria in repeated games, Econometrica 73 (2005) 377–415.

[9] J. Flesch, A. P. Y. Monsuwé, Repeated games with endogenous choice of information mechanisms, Math. Operations Res. 24 (1999) 785–794.

[10] D. Fudenberg, D. Levine, Efficiency and observability with long-run and short-run players, J. Econ. Theory 62 (1994) 103–135.

[11] D. Fudenberg, D. Levine, E. Maskin, The folk theorem with imperfect public information, Econometrica 62 (1994) 997–1039.

[12] D. Fudenberg, E. Maskin, The folk theorem in repeated games with discounting or with incomplete information, Econometrica 54 (1986) 533–556.

[13] E. Green, R. Porter, Noncooperative collusion under imperfect price information, Econometrica 52 (1984) 87–100.

[14] J. Horner, W. Olszewski, The folk theorem for all games with almost perfect monitoring, Econometrica 74 (2006) 1499–1544.

[15] M. Kandori, Randomization, communication, and efficiency in repeated games with imperfect public monitoring, Econometrica 71 (2003) 345–353.

[16] M. Kandori, H. Matsushima, Private observation, communication and collusion, Econometrica 66 (1998) 627–652.

[17] M. Kandori, I. Obara, Endogenous monitoring, 2004, mimeo.

[18] E. Lehrer, Lower equilibrium payoffs in two-player repeated games with non-observable actions, Int. J. Game Theory 18 (1989) 57–89.

[19] E. Lehrer, Internal correlation in repeated games, Int. J. Game Theory 19 (1991) 431–456.

[20] E. Lehrer, Correlated equilibria in two-player repeated games with nonobservable actions, Math. Operations Res. 17 (1992) 175–199.

[21] E. Lehrer, Two-player repeated games with nonobservable actions and observable payoffs, Math. Operations Res. 17 (1992) 200–224.

[22] G. Mailath, S. Morris, Repeated games with almost-public monitoring, J. Econ. Theory 102 (2002) 189–228.

[23] H. Matsushima, On the theory of repeated games with private information, part II: Revelation through communication, Econ. Letters 35 (1991) 257–261.

[24] H. Matsushima, Repeated games with private monitoring: Two players, Econometrica 72 (2004) 823–852.

[25] M. Piccione, The repeated prisoner's dilemma with imperfect private monitoring, J. Econ. Theory 102 (2002) 70–83.

[26] R. Radner, R. Myerson, E. Maskin, An example of a repeated partnership game with discounting and with uniformly inefficient equilibria, Rev. Econ. Stud. 53 (1986) 59–69.

[27] A. Rubinstein, Comments on the interpretation of repeated game theory, in: J.-J. Laffont (Ed.), Advances in Economic Theory: Sixth World Congress, vol. 1, Cambridge University Press, Cambridge, 1992.

[28] T. Sekiguchi, Efficiency in repeated prisoner's dilemma with private monitoring, J. Econ. Theory 76 (1997) 345–361.

[29] G. Stigler, A theory of oligopoly, J. Polit. Economy 72 (1964) 44–61.

[30] Y. Yamamoto, Efficiency results in $N$ player games with imperfect private monitoring, J. Econ. Theory 135 (2007) 382–413.