Belief-Free Review-Strategy Equilibrium without Conditional Independence

Takuo Sugaya*†
Princeton University

April 4, 2010

Abstract

We study repeated games with imperfect private monitoring. We obtain the characterization, which only depends on the parameters of the stage game, of the set of belief-free review-strategy equilibrium payoffs in the limit as the discount factor converges to one. Our characterization is valid for a generic monitoring technology if the number of private signals is sufficiently large and the number of players is no less than four.

In addition, we show a sufficient condition so that the characterized set contains a Pareto efficient payoff profile.

*tsugaya@princeton.edu

†The author is thankful to Stephen Morris for his encouragement and guidance. The author is also grateful to Dilip Abreu, Marco Battaglini, Jeffrey Ely, Kyna Fong, Johannes Hörner, Andriy Norets, Ichiro Obara, Yuliy Sannikov, Satoru Takahashi and Yuichi Yamamoto for their valuable comments. The usual disclaimer of course applies.

1 Introduction

The main finding in the theory of repeated games is that long term relationships enhance cooperation. In fact, the central result is probably the folk theorem: any feasible and individually rational payoff can be sustained in equilibrium when players are sufficiently patient. Fudenberg and Maskin (1986) establish the folk theorem under perfect monitoring, that is, when players can observe the opponents' actions directly. Fudenberg, Levine and Maskin (1994) extend the folk theorem to imperfect public monitoring, where players cannot observe the opponents' actions directly, but observe public noisy signals about the opponents' actions.[1] Recently, it has been shown that these results are robust to the introduction of private monitoring.[2] Hörner and Olszewski (2006) show the robustness of the folk theorem to almost perfect monitoring. Hörner and Olszewski (2009) establish the robustness of the folk theorem to almost public monitoring.[3]

On the other hand, when the monitoring is neither almost perfect nor almost public, almost all the results are attained only with conditionally independent monitoring: players can obtain no information on what their opponents have observed by observing their own private signals conditional on an action profile. Based on the seminal earlier works by Matsushima (2004), Ely, Hörner and Olszewski (2004 and 2005) and Yamamoto (2007),[4] Yamamoto (2009b) obtains the characterization of the equilibrium payoff set implementable by so-called belief-free review-strategy equilibria (henceforth BFRSE) for a general N-player game.[5]

[1] See also Fudenberg and Levine (1994).

[2] See Kandori (2002) and Mailath and Samuelson (2006) for the survey of private monitoring.

[3] See also Mailath and Morris (2002 and 2006).

[4] Matsushima (2004) obtains the folk theorem for two-player prisoners' dilemma with conditionally independent monitoring. Ely, Hörner, and Olszewski (2005) characterize the set of payoffs attainable by "belief-free review-strategy" for two-by-two games with conditionally independent monitoring. Yamamoto (2007) shows that efficiency (the payoff profile close to mutual cooperation) is obtained by a "belief-free review-strategy" for an N-player prisoners' dilemma with conditionally independent monitoring.

[5] The idea of the review strategy also appears in Radner (1985 and 1986) and Abreu, Milgrom and Pearce (1991).

The belief-free equilibrium is firstly introduced by Piccione (2002) and significantly modified by Ely and Välimäki. Yamamoto (2009a) characterizes the belief-free equilibrium payoff set when the discount factor converges to one.

The main idea of BFRSE is as follows. The infinite periods are regarded as a sequence of T-period review blocks. In each review block, suppose it is optimal to take a constant action. Then, however noisy the monitoring is, when T is sufficiently large, players can statistically infer the opponents' actions with arbitrarily high power by aggregation of information across T periods.

The remaining task is to verify the optimality of a constant action. It has been done with conditionally independent monitoring. This paper is the first to attain the characterization of the set of BFRSE payoffs for a generic monitoring structure.

The importance of conditional independence for Matsushima (2004), Ely, Hörner and Olszewski (2005) and Yamamoto (2007 and 2009b) is explained as follows. To statistically infer the opponents' actions by pooling the information, it is important to create an equilibrium where each player takes a constant action in each review block. Consider a prisoners'-dilemma example, in which player i is likely to be punished in the following review block if a lot of signals indicating defection are pooled in the current review block by player $j \neq i$. Suppose player i takes cooperation initially in the current review block. If the signals are conditionally independent, after any realization of the signals, it is optimal to stick to the cooperation since the observed signals have no information about whether player i will be punished or not. With conditionally dependent signals, however, in the periods near to the end of the current review block, under some realization of the signals, player i considers that she will be punished regardless of actions in the remaining periods of the current review block, which destroys the incentive to continue cooperation. Let us call this problem the "statistical inference problem".

Specifically, in Matsushima (2004), Ely, Hörner and Olszewski (2005) and Yamamoto (2007 and 2009b), the signals are pooled as follows. Consider the situation where player j reviews player i. Conditional on that player j plays a_j during the review block, she reviews player i's action by a random event whose realizations depend on (a_j, ω_j), where ω_j is player

There are several papers that extend the belief-free equilibrium. Kandori (2008) generalizes the belief-free equilibrium. Miyagawa, Miyahara and Sekiguchi (2008) apply the belief-free equilibrium to the situation where players can increase the precision of monitoring by paying the cost. Takahashi (2010) applies the belief-free equilibrium to the community enforcement.

j's private signal. Since the distribution of ω_j depends on player i's action, player j can infer player i's action by pooling the realized random events. From player i's perspective, the statistical inference on player j's review depends on the distribution of player j's random events conditional on player i's signals and actions. If the conditional probability of the random events does not depend on player i's signals but only on player i's actions, the statistical inference problem does not occur. However, to satisfy this requirement for a generic monitoring structure, it is necessary to have $|A_i| \times |\Omega_i| \leq |\Omega_j|$, which cannot be satisfied for all (i, j). Here, $|\Omega_i|$ is the number of player i's private signals and $|A_i|$ is the number of player i's actions.

The idea of this paper is that, without assuming conditionally independent monitoring directly, player j has sufficient information to create random events with the conditional independence property against player i if player j can also use the signals of the other players. We formalize this idea *without* introducing any communication device, that is, after main rounds corresponding to Matsushima's review blocks, we introduce the communication stages where players communicate about signals observed in the main rounds by taking actions.

There are two main difficulties to construct an equilibrium based on this idea. Firstly, if players communicate the realized signals of each period, the communication stages are too long to attain the efficiency. Therefore, after each main round, players use only randomly picked periods for monitoring and they communicate about signal observations in these periods.[6] Since this randomization is done after the main rounds, this does not affect player's incentive to take a constant action. Secondly, we have to give an incentive for players to tell the truth in the communication stages. As a usual claim in mechanism design, if we have no less than three players to monitor a player, it is possible to give the incentive to tell the truth. That is why we assume the number of players is no less than four.

Let us finish the introduction by reviewing other papers to attain the efficiency in re-

[6] We do *not* assume the existence of public randomization devices. See the equilibrium construction in Section 6.4 for the details of the randomization.

peated games with private monitoring.[7] Fong, Gossner, Hörner and Sannikov (2007) attain the efficiency result in a two-player prisoners' dilemma with a not-nongeneric monitoring structure. Neither their result nor ours contains the other as a special case. Moreover, the main ideas are different. Fong, Gossner, Hörner and Sannikov (2007) derive the condition that the statistical inference problem occurs only with small probability on the equilibrium path. On the other hand, this paper constructs a strategy that is completely free from the statistical inference problem.

Several papers such as Sekiguchi (1997), Bhaskar and Obara (2002) and Chen (2009) focus on belief-based techniques. In such equilibria, players' strategies involve statistical inference about the past history of the play. Results in those papers are limited to prisoners' dilemma with almost perfect monitoring.

Another approach is to introduce explicit communication. Versions of the folk theorem have been proven by Compte (1998), Kandori and Matsushima (1998), Aoyagi (2002), Fudenberg and Levine (2002) and Obara (2009). Since some practical economic settings make communication impossible or costly, their result might not be applicable. For example, in Stigler (1964)'s oligopoly example, anti-trust laws make communication illegal. Note that this paper does not introduce explicit communication.

The similar results are obtained without discounting in Lehrer (1989) and Fudenberg and Levine (1991). Although the average payoff in BFRSE under discounting with T being the length of each review block converges to T-period time-average payoff when the discount factor converging to 1, our equilibrium construction is significantly different from papers without discounting since we need to take care of the incentives for the gain and loss in finite periods.

The rest of the paper is organized as follows. Section 2 introduces the model. Section 3 defines the belief-free review-strategy equilibrium. Section 4 states the main result. Section 5 briefly overviews the basic theoretical ideas behind the equilibrium construction. Sections 6 explains the equilibrium strategy. Section 7 presents sufficient conditions for the efficiency

[7] Our review is close to that in Fong, Gossner, Hörner and Sannikov (2007).

and the folk theorem.

2 Model

The stage game is given by $\left\{I,(A_i,\Omega_i,g_i)_{i\in I},q\right\}$. $I=\{1,2,\ldots,N\}$ is the set of players, A_i is the finite set of player i's pure actions, Ω_i is the finite set of player i's private signals, and $g_i:A_i\times\Omega_i\to\mathbb{R}$ is player i's ex post utility function. We assume $N\geq 4$ and $|A_i|\geq 2$. Let $A\equiv\prod_{i\in I}A_i$ and $\Omega\equiv\prod_{i\in I}\Omega_i$ be the set of action profiles and signal profiles, respectively. For $\mathcal{I}\subset I$, $A_{-\mathcal{I}}$, $\Omega_{-\mathcal{I}}$, $a_{-\mathcal{I}}\in A_{-\mathcal{I}}$, $\omega_{-\mathcal{I}}\in\Omega_{-\mathcal{I}}$ are defined as usual.

In every stage game, players choose an action profile $a\equiv(a_1,\ldots,a_N)\in A$ and then a signal profile $\omega=(\omega_1,\ldots,\omega_N)\in\Omega$ is distributed according to the conditional probability function $q(\cdot\mid a)$. Given an action $a_i\in A_i$ and a private signal $\omega_i\in\Omega_i$, player i receives a profit $g_i(a_i,\omega_i)$. Thus, her expected payoff conditional on an action profile $a\in A$ is denoted by $\pi_i(a)=\sum_{\omega\in\Omega}q(\omega\mid a)\,g_i(a_i,\omega_i)$. For each $a\in A$, let $\pi(a)$ represent the payoff vector $(\pi_i(a))_{i\in I}$.

This paper imposes three assumptions on the monitoring technology q. Firstly, we assume the full support condition, which ensures that the set of Nash equilibrium payoffs is equivalent to that of sequential equilibrium payoffs as Sekiguchi (1997) shows:

Assumption 1 *Monitoring satisfies the full support condition if* $q(\omega\mid a)>0$ *for all* $a\in A$ *and* $\omega\in\Omega$.

Secondly, we require the following: the distribution of player j's signals given (a_j,a_{-j}) are linearly independent for fixed a_j. This implies that player j can statistically infer the opponents' actions a_{-j}. Mathematically, let us define $Q_j(a_j)$ as follows. $Q_j(a_j,a_{-j})$ is a row vector $(q(\omega_j\mid a_j,a_{-j}))_{\omega_j\in\Omega_j}$. $Q_j(a_j)$ is a matrix stacking $Q_j(a_j,a_{-j})$ for all a_{-j} vertically. Then, we can formally state the second assumption:

Assumption 2 *Monitoring satisfies the identifiability condition if*

$$\operatorname{rank}[Q_j(a_j)]=|A_{-j}|$$

for all $j \in I$ *and* $a_j \in A_j$*.*

Finally, we require the following. Consider the distribution of all the players but (i, l)'s signals, given an action profile (a_j, a_{-j}), player i's signal observation ω_i and player l's signal observation ω_l. Consider all these distributions when letting a_{-j} and ω_i vary and when exchanging the roles of player i and l. We assume all these distributions are linearly independent. This implies that player j, using the information from players $-(j, i, l)$, can monitor player i keeping the conditional independence property for players i and l. Mathematically, let $Q_i(a_j, a_{-j}, \omega_i)$ be a matrix with rows $\left(q\left(\omega_{-(i,l)} \mid a_j, a_{-j}, \omega_i\right)\right)_{\omega_{-(i,l)} \in \Omega_{-(i,l)}}$. $Q_i(a_j, a_{-j})$ is a matrix stacking $Q_j(a_j, a_{-j}, \omega_i)$ for all ω_i vertically. $Q_i(a_j)$ is a matrix stacking $Q_i(a_j, a_{-j})$ for all a_{-j} vertically. $Q_l(a_j)$ is similarly defined with i replaced with l. For $Q_{il}(a_j) = \begin{bmatrix} Q_i(a_j) \\ Q_l(a_j) \end{bmatrix}$, we require the following:

Assumption 3 *Monitoring satisfies the* $(N-2)$*-identifiability condition if*

$$\operatorname{rank}\left[Q_{il}(a_j)\right] = |A_{-j}| \times (|\Omega_i| + |\Omega_l|)$$

for all $j, i, l \in I$ *with* $j \neq i \neq l \neq j$ *and* $a_j \in A_j$*.*

For the rest of the paper, if we say j, i, l, we implicitly assume $j \neq i \neq l \neq j$. We show that if the number of private signals is sufficiently large, the above three assumptions are generically satisfied.

Lemma 4 *If* $|A_{-j}| \leq |\Omega_j|$ *for all* j *and* $|A_{-j}| \times (|\Omega_i| + |\Omega_l|) \leq \left|\Omega_{-(i,l)}\right|$ *for all* $j, i, l \in I$*, then the full support condition, the identifiability condition and the* $(N-2)$*-identifiability condition are generically satisfied.*

Proof. See Appendix. ■

Consider the infinitely repeated game in which the discount factor is $\delta \in (0, 1)$. Let $a_{i,\tau}$ and $\omega_{i,\tau}$ denote the performed action and the observed private signal respectively in period

τ by player i. Player i's private history up to period $t \geq 1$ is given by $h_i^t \equiv (a_{i,\tau}, \omega_{i,\tau})_{\tau=1}^t$. Let $h_i^0 = \emptyset$ and for each $t \geq 0$, let H_i^t be the set of all h_i^t. Then, a strategy for player i is defined to be a mapping $\sigma_i : \bigcup_{t=0}^{\infty} H_i^t \to \triangle A_i$. Let Σ_i be the set of all strategies for player i and let $\Sigma \equiv \prod_{i \in I} \Sigma_i$. Also, for any strategy $\sigma_i \in \Sigma_i$, history $h_i^t \in H_i^t$ and action $a_i \in A_i$, let $\sigma_i \mid h_i^t$ be player i's continuation strategy after h_i^t and $\sigma_i(h_i^t)[a_i]$ be the probability for σ_i to take a_i after h_i^t. In addition, let $\sigma_i \mid (h_i^t, a_i)$ represent player i's strategy $\tilde{\sigma}_i \in \Sigma_i$ such that $\tilde{\sigma}_i(h_i^0) = a_i$ and for any $h_i^1 \in H_i^1$, $\tilde{\sigma}_i \mid h_i^1 = \sigma_i \mid h_i^{t+1}$ where $h_i^{t+1} = (h_i^t, h_i^1)$. In words, $\sigma_i \mid (h_i^t, a_i)$ denotes the continuation strategy after history h_i^t but the action $a_{i,t+1}$ is replaced with the pure action a_i. Finally, let $w_i(\sigma)$ represent player i's expected average payoff by a strategy profile $\sigma \in \Sigma$, that is, $w_i(\sigma) = (1-\delta) E\left[\sum_{t=1}^{\infty} \delta^{t-1} \pi_i(a_t) \mid \sigma\right]$.

Our objective is to characterize the belief-free review-strategy equilibrium (BFRSE) payoff set under the above assumptions. The definition of BFRSE is given by Yamamoto (2009b), followed by the characterization with conditionally independent monitoring. Instead of conditionally independent monitoring, we assume the $(N-2)$-identifiability, which is satisfied generically. We show that we can "restore" the conditional independence property using the $(N-2)$-identifiability. We use the same notation and explanation as Yamamoto (2009b) wherever possible to make comparison easier.

3 Belief-Free Review-Strategy Equilibrium

This section states a notion of the belief-free review-strategy equilibrium (BFRSE), on which we concentrate. Firstly, we define a *review strategy profile.*

Definition 5 *Let $(t_l)_{l=0}^{\infty}$ be a sequence of integers with $t_0 = 0$, $t_l > t_{l-1}$ for all $l \geq 1$. A strategy profile $\sigma \in \Sigma$ is a review strategy profile with $(t_l)_{l=0}^{\infty}$ if $\sigma_i\left(h_i^{t-1}\right)[a_{i,t-1}] = 1$ for all l, $t \in \{t_{l-1}+2, \ldots, t_l\}$, i and $h_i^t = (a_{i,\tau}, \omega_{i,\tau}) \in H_i^t$.*

In this definition, an infinitely repeated game is divided into infinitely repeated review rounds. The lth review round is from $t_{l-1}+1$ to t_l. Once player i decides an action a_{i,t_l+1} in the initial period of each round, player i is stick to that action within that round.

Secondly, we define a *belief-free review-strategy equilibrium (BFRSE).*

Definition 6 *Let $\sigma \in \Sigma$ be a review strategy profile with a sequence $(t_l)_{l=0}^{\infty}$. σ is belief-free if*

$$\sigma_i \mid h_i^{t_{l-1}} \in BR(\sigma_{-i} \mid h_{-i}^{t_{l-1}}, a_{-i}) \text{ for all } a_{-i} \text{ with } \sigma_{-i}(h_{-i}^{t_{l-1}})\,[a_{-i}] > 0$$

for all $i \in I$, $l \geq 1$ and $h^{t_{l-1}} \in H^{t_{l-1}}$.

Intuitively, a review strategy profile to be belief-free requires that, in each lth review round, conditional on the past history and what constant action the other players choose in the current round, player i's continuation strategy is a best response. Note that a belief-free review-strategy profile is a Nash equilibrium by definition. Note also that our definition of a belief-free review-strategy is the same as Yamamoto (2009b).

4 Characterization

Our objective is to characterize the set of payoffs attainable in BFRSE. To state the characterization, we introduce several notations.

A non-empty set $\mathcal{A} \subset A$ is a *regime generated from* A if $\mathcal{A}$ has a product structure, that is, $\mathcal{A} = \prod_{i \in I} \mathcal{A}_i$ with $\mathcal{A}_i \subset A_i$ for all $i \in I$. As we will see, $\mathcal{A}$ is the set of actions taken on the equilibrium path.

Let $\mathcal{J}$ be the set of all regimes generated from A. Suppose players can access to a public randomization device, which selects a "recommended action set" $\mathcal{A}$ from $\mathcal{J}$ with probability $p(\mathcal{A})$.[8] Let $V(p)$ be the feasible payoff set when players obey this recommendation, that is, for any probability distribution $p \in \triangle \mathcal{J}$, define $V(p)$ as

$$V(p) \equiv \text{co}\left\{ \sum_{\mathcal{A} \in \mathcal{J}} p(\mathcal{A})\, \pi(a(\mathcal{A})) \mid a(\mathcal{A}) \in \mathcal{A}, \forall \mathcal{A} \in \mathcal{J} \right\},$$

where co$V$ denote the convex hull of a set V.

[8] We do *not* assume the existence of public randomization devices for the equilibrium construction.

For each i and $\mathcal{A} \in \mathcal{J}$, consider the situation that players $-i$ punish player i. Suppose players $-i$ are required to take actions included in $\mathcal{A}$ but player i is free to deviate to any action. Then, the lowest payoff that players $-i$ can guarantee regardless of player i's action $a_i \in A_i$ is

$$\underline{v}_i(\mathcal{A}) \equiv \min_{a_{-i}\in\mathcal{A}_{-i}} \max_{a_i\in A_i} \pi_i(a).$$

On the other hand, consider the situation that players $-i$ reward player i. Suppose every player is required to take actions included in $\mathcal{A}$. Then, the highest payoff that players $-i$ can guarantee regardless of player i's action $a_i \in \mathcal{A}_i$ is

$$\overline{v}_i(\mathcal{A}) \equiv \max_{a_{-i}\in\mathcal{A}_{-i}} \min_{a_i\in\mathcal{A}_i} \pi_i(a).$$

Also, for each i and $\mathcal{A} \in \mathcal{J}$, let $\underline{a}^i(\mathcal{A}) \in \mathcal{A}$ and $\overline{a}^i(\mathcal{A}) \in \mathcal{A}$ be such that $\underline{a}^i_{-i}(\mathcal{A}) \in \mathcal{A}_{-i}$ and $\overline{a}^i_{-i}(\mathcal{A}) \in \mathcal{A}_{-i}$ solve the above problem, that is,

$$\underline{v}_i(\mathcal{A}) = \max_{a_i\in A_i} \pi_i\left(a_i, \underline{a}^i_{-i}(\mathcal{A})\right), \overline{v}_i(\mathcal{A}) = \min_{a_i\in\mathcal{A}_i} \pi_i\left(a_i, \overline{a}^i_{-i}(\mathcal{A})\right). \tag{1}$$

Note that $\underline{a}^i_i(\mathcal{A}) \in \mathcal{A}_i$ and $\overline{a}^i_i(\mathcal{A}) \in \mathcal{A}_i$ are arbitrary.

For all $i \in I$, for notational convenience, let

$$p\underline{v}_i \equiv \sum_{\mathcal{A}\in\mathcal{J}} p(\mathcal{A})\,\underline{v}_i(\mathcal{A}), p\overline{v}_i \equiv \sum_{\mathcal{A}\in\mathcal{J}} p(\mathcal{A})\,\overline{v}_i(\mathcal{A}).$$

As we will see, $p\underline{v}_i$ is the payoff used for punishing player i and $p\overline{v}_i$ is that for rewarding player i. For this to work, we assume the following full dimensionality condition[9]:

Assumption 7 *A stage game payoff structure satisfies the full dimensionality condition if*

$$\dim \bigcup_{p\in\triangle\mathcal{J}} \left(V(p) \cap \prod_{i\in I} [p\underline{v}_i, p\overline{v}_i]\right) = N.$$

[9]See empty, negative and abnormal cases in Yamamoto (2009b) for the cases without the full dimensionality.

Now we can state our main result.

Theorem 8 *If the full support, identifiability, $(N-2)$-identifiability and full dimensionality conditions are satisfied, then* $\bigcup_{p\in\triangle\mathcal{J}}\left(V(p)\cap\prod_{i\in I}[p\underline{v}_i,p\overline{v}_i]\right)$ *is the limit set of BFRSE payoffs as the discount factor converges to one, that is,*

$$\bigcup_{p\in\triangle\mathcal{J}}\left(V(p)\cap\prod_{i\in I}[p\underline{v}_i,p\overline{v}_i]\right)=\lim_{\delta\to 1}E(\delta), \tag{2}$$

where $E(\delta)$ is the set of BFRSE payoffs with discount factor δ.

Note that the characterization (2) is the same as Yamamoto (2009b). Yamamoto (2009b) shows the result *with conditional independence monitoring*. Our objective is to show that the same characterization is valid *for a generic monitoring structure* when $N\geq 4$.

The proof of $\bigcup_{p\in\triangle\mathcal{J}}\left(V(p)\cap\prod_{i\in I}[p\underline{v}_i,p\overline{v}_i]\right)\supset\lim_{\delta\to 1}E(\delta)$ is exactly the same as Yamamoto (2009b) and so is omitted.[10] We explain that any $v\in\operatorname{int}\bigcup_{p\in\triangle\mathcal{J}}\left(V(p)\cap\prod_{i\in I}[p\underline{v}_i,p\overline{v}_i]\right)$ can be attained by BFRSE.

5 Basic Ideas: Recovery of the Conditional Independence Property

To explain the idea, we consider a prisoners'-dilemma example in this section. The idea of the review strategy is as follows. Even if the monitoring is far from perfect, if we aggregate information for a long review round and make players take a constant action within a round, we can statistically infer the opponents' actions accurately. However, to make a constant action optimal on the equilibrium path after any history, it is important to prevent the statistical inference problem as explained in Introduction.

[10] Note that Yamamoto (2009b) does not use the conditional independence property to show $\bigcup_{p\in\triangle\mathcal{J}}\left(V(p)\cap\prod_{i\in I}[p\underline{v}_i,p\overline{v}_i]\right)\supset\lim_{\delta\to 1}E(\delta)$.

Since **Definition 6** requires that the strategy is optimal *conditional on* the opponents' action in a current round, we conditional on that player j takes a_j. Suppose player j tries to infer player i's action. To prevent the statistical inference problem, it is enough to show the existence of the solution for

$$\begin{bmatrix} q(\omega_j^1 \mid a_j, C_i, a_{-(j,i)}, \omega_i^1) & \cdots & q(\omega_j^{|\Omega_j|} \mid a_j, C_i, a_{-(j,i)}, \omega_i^1) \\ \vdots & & \vdots \\ q(\omega_j^1 \mid a_j, C_i, a_{-(j,i)}, \omega_i^{|\Omega_i|}) & \cdots & q(\omega_j^{|\Omega_j|} \mid a_j, C_i, a_{-(j,i)}, \omega_i^{|\Omega_i|}) \\ q(\omega_j^1 \mid a_j, D_i, a_{-(j,i)}, \omega_i^1) & \cdots & q(\omega_j^{|\Omega_j|} \mid a_j, D_i, a_{-(j,i)}, \omega_i^1) \\ \vdots & & \vdots \\ q(\omega_j^1 \mid a_j, D_i, a_{-(j,i)}, \omega_i^{|\Omega_i|}) & \cdots & q(\omega_j^{|\Omega_j|} \mid a_j, D_i, a_{-(j,i)}, \omega_i^{|\Omega_i|}) \end{bmatrix} \begin{bmatrix} \psi(a_j, \omega_j^1) \\ \vdots \\ \psi(a_j, \omega_j^{|\Omega_j|}) \end{bmatrix} = \begin{bmatrix} q_C \\ \vdots \\ q_C \\ q_D \\ \vdots \\ q_D \end{bmatrix}$$

for any $a_{-(j,i)} \in A_{-(j,i)}$ with $\psi(a_j, \omega_j) \in [0,1]$ and $q_C > q_D$ for all $\omega_j \in \Omega_j$ and $a_j \in A_j$. Note that the matrix expression is equivalent to

$$E\left[\psi(a_j, \omega_j)|a, \omega_i\right] = \begin{cases} q_C & \text{if } a_i = C_i, \\ q_D & \text{if } a_i = D_i. \end{cases}$$

To see why this is sufficient, consider that player j reviews player i's action as follows. After each period τ in the current round, after playing a_j and observing $\omega_{j,\tau}$, player j draws a random variable from uniform $[0,1]$. If the realization of this random variable is less than $\psi(a_j, \omega_{j,\tau})$, we say $\Psi_\tau = 1$ and otherwise $\Psi_\tau = 0$. Note that $\Pr\left(\{\Psi = 1\}|a_j, \omega_j\right) = \psi\left(a_j, \omega_j\right)$. Hence, (i) since $\Pr\left(\{\Psi = 1\}|C_i, a_{-i}\right) = q_C > q_D = \Pr\left(\{\Psi = 1\}|D_i, a_{-i}\right)$ for any a_{-i}, by counting the number of periods where $\Psi_\tau = 1$, player j can statistically distinguish player i's action. (ii) In addition, player i cannot infer player j's counting from ω_i since $\Pr\left(\{\Psi = 1\}|a, \omega_i\right) = \Pr\left(\{\Psi = 1\}|a\right)$ for all a and ω_i. However, so that the above system has a solution generically, it is necessary to have $|A_i| \times |\Omega_i| \leq |\Omega_j|$, which cannot be satisfied for all (j,i).

The idea of this paper is as follows. Suppose player j can also use the signal $\omega_{-(i,l)}$ with

$l \neq j, i$ to review player i and

$$\begin{bmatrix} q(\omega^1_{-(i,l)} \mid a_j, C_i, a_{-(j,i)}, \omega_i^1) & \cdots & q(\omega^{|\Omega_{-(i,l)}|}_{-(i,l)} \mid a_j, C_i, a_{-(j,i)}, \omega_i^1) \\ \vdots & & \vdots \\ q(\omega^1_{-(i,l)} \mid a_j, C_i, a_{-(j,i)}, \omega_i^{|\Omega_i|}) & \cdots & q(\omega^{|\Omega_{-(i,l)}|}_{-(i,l)} \mid a_j, C_i, a_{-(j,i)}, \omega_i^{|\Omega_i|}) \\ q(\omega^1_{-(i,l)} \mid a_j, D_i, a_{-(j,i)}, \omega_i^1) & \cdots & q(\omega^{|\Omega_{-(i,l)}|}_{-(i,l)} \mid a_j, D_i, a_{-(j,i)}, \omega_i^1) \\ \vdots & & \vdots \\ q(\omega^1_{-(i,l)} \mid a_j, D_i, a_{-(j,i)}, \omega_i^{|\Omega_i|}) & \cdots & q(\omega^{|\Omega_{-(i,l)}|}_{-(i,l)} \mid a_j, D_i, a_{-(j,i)}, \omega_i^{|\Omega_i|}) \end{bmatrix} \begin{bmatrix} \psi(a_j, \omega^1_{-(i,l)}) \\ \vdots \\ \psi(a_j, \omega^{|\Omega_{-(i,l)}|}_{-(i,l)}) \end{bmatrix} = \begin{bmatrix} q_C \\ \vdots \\ q_C \\ q_D \\ \vdots \\ q_D \end{bmatrix} \tag{3}$$

for all $a_{-(j,i)}$ has a solution for all a_j. Equivalently,

$$E\left[\psi(a_j, \omega_{-(i,l)})|a, \omega_i\right] = \begin{cases} q_C & \text{if } a_i = C_i, \\ q_D & \text{if } a_i = D_i. \end{cases}$$

Suppose player j reviews player i's action as follows: After each period τ, after playing a_j and observing $\omega_{-(i,l),\tau}$,[11] player j draws a random variable from uniform $[0, 1]$. If the realization of this random variable is less than $\psi(a_j, \omega_{-(i,l),\tau})$, we say $\Psi_\tau = 1$ and otherwise $\Psi_\tau = 0$. Then, the similar properties as (i) and (ii) are satisfied.

As we will see, it is important to have one player l whose signal is not used when player j reviews player i and player l cannot infer player j's counting about player i, which is

[11] Here, for simplicity, we do not take into account how player j collects the information of $\omega_{-(i,l)}$. See Section 6.4 for the details.

expressed by the following condition:

$$\begin{bmatrix} q(\omega^1_{-(i,l)} \mid a_j, C_i, a_{-(j,i)}, \omega^1_l) & \cdots & q(\omega^{|\Omega_{-(i,l)}|}_{-(i,l)} \mid a_j, C_i, a_{-(j,i)}, \omega^1_l) \\ \vdots & & \vdots \\ q(\omega^1_{-(i,l)} \mid a_j, C_i, a_{-(j,i)}, \omega^{|\Omega_l|}_l) & \cdots & q(\omega^{|\Omega_{-(i,l)}|}_{-(i,l)} \mid a_j, C_i, a_{-(j,i)}, \omega^{|\Omega_l|}_l) \\ q(\omega^1_{-(i,l)} \mid a_j, D_i, a_{-(j,i)}, \omega^1_l) & \cdots & q(\omega^{|\Omega_{-(i,l)}|}_{-(i,l)} \mid a_j, D_i, a_{-(j,i)}, \omega^1_l) \\ \vdots & & \vdots \\ q(\omega^1_{-(i,l)} \mid a_j, D_i, a_{-(j,i)}, \omega^{|\Omega_l|}_l) & \cdots & q(\omega^{|\Omega_{-(i,l)}|}_{-(i,l)} \mid a_j, D_i, a_{-(j,i)}, \omega^{|\Omega_l|}_l) \end{bmatrix} \begin{bmatrix} \psi(a_j, \omega^1_{-(i,l)}) \\ \vdots \\ \psi(a_j, \omega^{|\Omega_{-(i,l)}|}_{-(i,l)}) \end{bmatrix} = \begin{bmatrix} q_C \\ \vdots \\ q_C \\ q_D \\ \vdots \\ q_D \end{bmatrix} \tag{4}$$

for all $a_{-(j,i)}$, or,

$$E\left[\psi(a_j, \omega_{-(i,l)})|a, \omega_l\right] = \begin{cases} q_C & \text{if } a_i = C_i, \\ q_D & \text{if } a_i = D_i. \end{cases}$$

This condition says that the conditional independence property holds for player l when player j tries to monitor player i, which means the conditional independence property holds not only while herself being monitored but also while someone else monitoring another player.

Note that if the $(N-2)$-identifiability is satisfied, there exists $\psi : A_j \times \Omega_{-(j,l)} \to [0,1]$ satisfying (3) and (4).

6 Equilibrium Construction

6.1 Overview

Take any $v \in \text{int} \bigcup_{p \in \triangle \mathcal{J}} (V(p) \cap \prod_{i \in I} [p\underline{v}_i, p\overline{v}_i])$. Let $p \in \triangle \mathcal{J}$ be such that $v \in \text{int}(V(p) \cap \prod_{i\in I} [p\underline{v}_i, p\overline{v}_i])$. Let $(\underline{w}_i)_{i\in I}$ and $(\overline{w}_i)_{i\in I}$ be vectors of real numbers such that $\underline{w}_i < v_i < \overline{w}_i$ and $\prod_{i\in I} [\underline{w}_i, \overline{w}_i] \subset \text{int}(V(p) \cap \prod_{i\in I} [p\underline{v}_i, p\overline{v}_i])$. It suffices to show that $\prod_{i\in I} [\underline{w}_i, \overline{w}_i]$ is sustained in BFRSE for sufficiently large δ.

In the rest of the paper, let "player $i-1$" refer to player $i-1$ for each $i \in \{2, \ldots, N\}$ and to player N for $i = 1$. Likewise, let "player $i+1$" refer to player $i+1$ for each $i \in \{1, \ldots, N-1\}$ and to player 1 for $i = N$. In addition, let $X_i \equiv \{G, B\}$ and $X \equiv \prod_{i\in I} X_i$. As we will see,

X_i is the set of player i's possible states in the equilibrium.

In the equilibrium construction below, we see the infinite repeated game as a sequence of T_b-period *block games.* T_b will be specified later. In each block game, each player i is either in the "good state" G or in the "bad state" B. When player i is in state G, she plays a T_b-period repeated game strategy σ_i^G during the current block game. On the other hand, when player i is in state B, she plays a T_b-period repeated game strategy σ_i^B. These "block-game strategies" σ_i^G and σ_i^B are chosen in such a way that player i earns high block-game payoffs $\overline{w}_i$ if player $i-1$ is in state G, and obtains low block-game payoffs $\underline{w}_i$ if player $i-1$ is in state B, that is, player $i-1$'s state controls player i's payoff. Therefore, the strategy σ_i^G is a "good" strategy and σ_i^B is a "bad" strategy.

At the end of each block game, a player transits over two states G and B so that the following conditions are satisfied: (i) given any players $-i$'s current state, player i is indifferent between being in good state and being in bad state and is not willing to deviate to a block-game strategy $\sigma_i^{T_b} \neq \sigma_i^G, \sigma_i^B$; (ii) for each $j \neq i-1$, player j's state does not affect player i's continuation payoff from that block game; and (iii) player i's continuation payoff from a block game is high if player $i-1$'s current state is good and it is low if player $i-1$'s current state is bad.

From (i), a player is indifferent between being in state G and in state B independently of the opponents' states. This assures that a continuation play from the beginning of each block game is an equilibrium given any state profile and in this sense players do not coordinate a choice of states. Moreover, (ii) and (iii) imply that player $i-1$ can solely control player i's continuation payoff through a choice of states and hence player $i-1$ does not need to know the state of the other players in order to punish or reward player i. (In particular player $i-1$ chooses state G if he wants to reward player i, while she chooses state B if he wants to punish.)

Note that the above structure of BFRSE is the same as Yamamoto (2009b). However, Yamamoto (2009b) constructs an equilibrium to satisfy the above properties when monitoring is conditionally independent. The contribution of our paper is to give a way to construct an equilibrium when monitoring itself is conditionally *dependent*. Our basic idea is to recover the "conditional independence property." That is, the important property that Yamamoto (2009b) derives from conditionally independent monitoring is that each player i's signal observations are, conditional on the action profile, not informative about the other players continuation play. We construct an equilibrium such that even though the monitoring is conditionally dependent, the statistics used to determine the continuation play satisfies the following *conditional independence property*: each player i's signal observations are not informative about the continuation play. After achieving the conditional independence property, we can use Yamamoto (2009b)'s construction. Hence, we firstly explain Yamamoto (2009b)'s construction, which works only for (almost) conditionally independent monitoring: for each i, ω_i and a, $q(\omega_{-i}|a, \omega_i) = q(\omega_{-i}|a)$. Secondly, we explain how to restore the conditional independence property with conditionally dependent monitoring.

6.1.1 With Conditional Independence

In Yamamoto (2009b)'s construction, a T_b-period block game is divided into rounds and each round is further divided into review phases. Specifically, a block game consists of a coordination round, a confirmation round, K pairs of main rounds and supplemental rounds and a report round. The coordination, confirmation, supplemental and report rounds are regarded as "communication stages," where players communicate via a choice of actions. These rounds consist of a series of T-period review phases. As we will see, for each i, the state space X_i is also player i's message space in each phase. Arbitrarily pick two different elements from A_i and call them a_i^G and a_i^B respectively. If player i wants to send message $x_i \in X_i$ in a phase, she constantly takes $a_i^{x_i}$ in that phase. Since T is sufficiently large, the communication in these rounds is almost perfect. We will explain player i's strategy given a state $x_i \in \{G, B\}$.

States As we have seen, each player has two possible states $x_i \in \{G, B\}$. When player i's state is G, player i takes the good strategy σ_i^G. On the other hand, when player i's state is B, player i takes the bad strategy σ_i^B. As we will see, σ_i^G and σ_i^B only differ in the coordination round and confirmation round.

Coordination Round This round is regarded as a communication stage, where each player i reveals whether she is in state G or in state B. The set of player i's possible messages is $X_i = \{G, B\}$ and player i with state $x_i \in X_i$ sends message x_i in order to reveal her state. To send message x_i, player i chooses action $a_i^{x_i}$ constantly in the coordination round.

Confirmation Round This round is also regarded as a communication stage, in which each player i reports her inferences about the messages from players $-i$ in the coordination round. The purpose of this communication is to make a confirmation and consensus about what has happened in the coordination round.

Main Rounds and Supplemental Rounds Players' play in the main rounds depends on the result of the communication in the confirmation round. If a player has confirmed that the current state profile is $\tilde{x} = (\tilde{x}_i)_{i=1}^N \in X = \{G, B\}^N$, then the player plays an action $a_i^{\tilde{x},k}$ until some player's deviation is confirmed in the supplemental round. $\tilde{x}$ can be different from the true state profile x. However, as we will see, on the equilibrium path, it is likely that $\tilde{x} = x$. Therefore, the action profile $a^{x,k}$ is played with high probability, which implies that player i's expected block-game payoff is high if player $i-1$'s current state is G and it is low if player $i-1$'s state is B.

In each supplemental round, every player reports whether or not they think the opponents have deviated in the previous main round. Based on the messages in the supplemental round, players try to make a confirmation and consensus about whether player i has deviated in the main round or not.

Report Round This round is also regarded as a communication stage and each player reports what she has observed in the confirmation round and the supplemental rounds. The information revealed in the report round is utilized to determine the transition probability between states G and B at the beginning of the next block game.

States in the Next Review Block Given the current state x_i and the history in the current block $h_i^{T_b}$, player i's state in the next review block, (that is, whether player i will take the good strategy σ_i^G or the bad strategy σ_i^B,) is determined so that the requirements (i), (ii) and (iii) are satisfied.

6.1.2 Without Conditional Independence

So far, the idea of the equilibrium construction is the same as Yamamoto (2009b). However, so that this construction works, it is important to have the conditionally independence property for each round, that is, during the round, player i's signal observations should not reveal the other players' inference of player i's action. To "restore" the conditional independence property if the monitoring itself does not satisfy conditional independence, we consider the following modification:

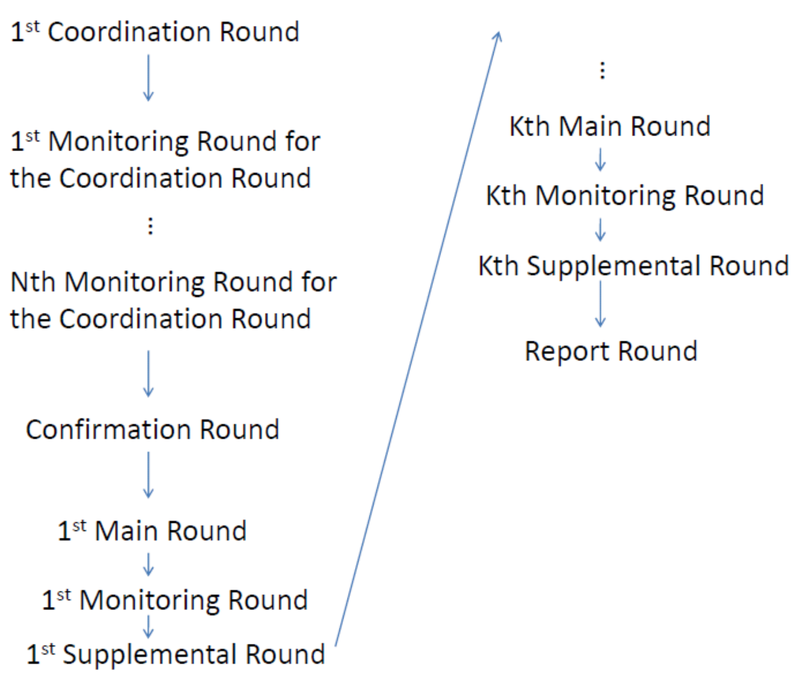
Monitoring Rounds for the Coordination Round After the coordination round is over, where each player i sends the message x_i, we insert the "ith monitoring round for the coordination round" for each i. In this round, each player $j \neq i$ tries to infer player i's message in the coordination round by collecting information from players $n \neq i, j$. The specific way to collect the information will be explained below. By allowing player j to use the information from players $n \neq i, j$, we can restore the conditional independence property in the coordination round as explained in Section 5.

kth Monitoring Round After the kth main round, we insert the "kth monitoring round." In this round, each player $j \neq i$ tries to infer whether player i has deviated in the kth main round by collecting information from players $n \neq i, j$. Again, by allowing player j to use the information from players $n \neq i, j$, we can restore the conditional independence property in the main rounds.

Both the monitoring rounds for coordination round and kth monitoring rounds can be seen as communication stages since players exchange their signal observations. In addition, the substantial changes are also necessary for the confirmation and supplemental rounds, which will be explained later.

One may notice that unlike cheap-talk games, actions in these communication stages are payoff-relevant. However, the duration of the communication stages is much shorter than that of the main rounds, so that payoffs in the communication stages are almost negligible. This property is the same as Yamamoto (2009b), but for our equilibrium, this requirement is harder to be satisfied since players need to exchange the messages about the signal observations in the main round. After each kth main round, we let players randomly select only a small fraction of periods in the kth main round and use only these periods for the monitoring, which makes the kth monitoring round sufficiently short. Further, since the randomization occurs after the main round is over, this does not affect the incentive in the main round.

Therefore, in summary, we consider the BFRSE illustrated in the picture below.



6.2 Regimes and Payoffs

Let us formally construct block-game strategies. Take any $v = (v_1, \ldots, v_N) \in \operatorname{int} \bigcup_{p \in \triangle \mathcal{J}} (V(p) \cap \prod_{i \in I} [p\underline{v}_i, p\overline{v}_i])$. Let $p \in \triangle \mathcal{J}$ be such that $v \in \operatorname{int}(V(p) \cap \prod_{i \in I} [p\underline{v}_i, p\overline{v}_i])$. Let $(\underline{w}_i)_{i \in I}$ and $(\overline{w}_i)_{i \in I}$ be vectors of real numbers such that $\underline{w}_i < v_i < \overline{w}_i$ and such that the hyper-rectangle $\prod_{i \in I} [\underline{w}_i, \overline{w}_i] \subset \operatorname{int}(V(p) \cap \prod_{i \in I} [p\underline{v}_i, p\overline{v}_i])$. It suffices to show that $\prod_{i \in I} [\underline{w}_i, \overline{w}_i]$ is sustained in BFRSE for sufficiently large δ.

First, we define the action profile $a^{x,k}$ that is taken in the kth main round with high probability.[12] There are a natural number $\tilde{K}$ and a sequence of regimes $(\mathcal{A}^1, \ldots, \mathcal{A}^{\tilde{K}})$ such that

$$\frac{1}{\tilde{K}} \sum_{k=1}^{\tilde{K}} \underline{v}_i \left(\mathcal{A}^k\right) < \underline{w}_i < v_i < \overline{w}_i < \frac{1}{\tilde{K}} \sum_{k=1}^{\tilde{K}} \overline{v}_i \left(\mathcal{A}^k\right) \tag{5}$$

for all $i \in I$ and such that for each $x \in X$, there exists a sequence of action profiles $(a^{x,1}, \ldots, a^{x,\tilde{K}})$ such that

$$a^{x,k} \in \mathcal{A}^k$$

for all $k \in \{1, \ldots, \tilde{K}\}$ and

$$w_i^x \equiv \frac{1}{\tilde{K}} \sum_{k=1}^{\tilde{K}} \pi_i \left(a^{x,k}\right) \begin{cases} < \underline{w}_i & \text{if } x_{i-1} = B \\ > \overline{w}_i & \text{if } x_{i-1} = G \end{cases} \tag{6}$$

for all $i \in I$. In words, (5) implies the following: if players $-i$ play $\underline{a}^i_{-i}(\mathcal{A}^1), \ldots \underline{a}^i_{-i}(\mathcal{A}^{\tilde{K}})$, player i's time-average payoff cannot exceed $\underline{w}_i$. Hence, this action sequence is used to punish player i. Similarly, if players $-i$ play $\overline{a}^i_{-i}(\mathcal{A}^1), \ldots \overline{a}^i_{-i}(\mathcal{A}^{\tilde{K}})$, player i's time-average payoff is at least $\overline{w}_i$ as long as player i takes an action from the "recommended set" $\mathcal{A}^k_i$. Hence, this action sequence is used to reward player i. In addition, (6) implies that if player $i-1$'s state is good, player i's time-average payoff is high and if player $i-1$'s state is bad, player i's time-average payoff is low.

Given a natural number K, let $\left(\mathcal{A}^1, \ldots, \mathcal{A}^K\right)$ be a cyclic sequence of $(\mathcal{A}^1, \ldots, \mathcal{A}^{\tilde{K}})$ with

[12] Since this action is the same as Yamamoto (2009b), our explanations are also the same.

length K, that is, $\mathcal{A}^{k+n\tilde{K}} = \mathcal{A}^k$ for all $k \in \{1, \ldots, \tilde{K}\}$ and $n \geq 0$. Also, let $(a^{x,1}, \ldots, a^{x,K})$ be a cyclic sequence of $(a^{x,1}, \ldots, a^{x,\tilde{K}})$ for all x.

We will construct BFRSE such that, in each block game, each player i takes an automaton strategy with two possible states $x_i \in \{G, B\}$. When the state is G, player i takes the good strategy σ_i^G and when it is B, player i takes the bad strategy σ_i^B. If the state profile is x, players play $a^{x,k}$ in the kth main round "almost always," which gives us an average payoff close to w_i^x satisfying (6).

6.3 Message Protocol

Second, let us explain how players send the messages about the signal observations since in our equilibrium, players exchange the messages about their signal observations via taking actions.[13]

Message Protocol for ω_i To send the message $\omega_{i,t} = \omega_i$, consider the following procedure. Let us attach a sequence of actions $a_i(\omega_i) = (a_{i,1}(\omega_i), \ldots, a_{i,M}(\omega_i))$ to each $\omega_i \in \Omega_i$ with $a_{i,m}(\omega_i) \in \{a_i^G, a_i^B\}$ for all m, $\omega_i \in \Omega_i$ and $i \in I$ as follows:

- order ω_i such that $\Omega_i = \{\omega_{i,1}, \ldots, \omega_{i,|\Omega_i|}\}$.
- $a_{i,1}(\omega_i) = a_i^G$ if $\omega_i \in \{\omega_{i,1}, \ldots, \omega_{i,|\Omega_i|/2}\}$ and $a_{i,1}(\omega_i) = a_i^B$ otherwise. That is, $a_{i,1}(\omega_i) = a_i^G$ if ω_i is in the first half of Ω_i and $a_{i,1}(\omega_i) = a_i^B$ otherwise.[14]
  - For ω_i with $a_{i,1}(\omega_i) = a_i^G$, $a_{i,2}(\omega_i) = a_i^G$ if $\omega_i \in \{\omega_{i,1}, \ldots, \omega_{i,(|\Omega_i|/2)/2}\}$ and $a_{i,2}(\omega_i) = a_i^B$ otherwise. That is, among ω_i in the first half of Ω_i, $a_{i,2}(\omega_i) = a_i^G$ if ω_i is in the first quarter and $a_{i,2}(\omega_i) = a_i^B$ otherwise.
  - Similarly, for ω_i with $a_{i,1}(\omega_i) = a_i^B$, $a_{i,2}(\omega_i) = a_i^G$ if $\omega_i \in \{\omega_{i,|\Omega_i|/2+1}, \ldots, \omega_{i,(|\Omega_i|-|\Omega_i|/2)/2}\}$ and $a_{i,2}(\omega_i) = a_i^B$ otherwise.
  - Keep this procedure until we can identify ω_i uniquely from $a_i(\omega_i)$.

[13] This is unique to our equilibrium. Sections 6.3 and 6.4 are our new contribution.
[14] In the rest of the paper, we neglect the integer problems since they are fixed straightforwardly.

Without loss of generality, we can assume M is constant among players. To send message $\omega_{i,t} = \omega_i$, player i takes $a_{i,1}(\omega_i)$ for S periods, then takes $a_{i,2}(\omega_i)$ for S periods and so forth. Hence, to send $\omega_{i,t} = \omega_i$, it takes MS periods.

Let us consider player $j(\neq i)$'s inference of player i's message. For this purpose, the following lemma is useful:

Lemma 9 *Suppose that monitoring satisfies the identifiability condition. Then, there exist $q_2 < q_3$ such that for all $i, j \in I$, $a \in A$ and $\mathcal{A}_i \subset A_i$, there exists a function $\phi_{\mathcal{A}_i} : A_j \times \Omega_j \to [0,1]$ such that*

$$E\left[\phi_{\mathcal{A}_i}(a_j, \omega_j) | a\right] = \begin{cases} q_3 & \text{if } a_i \in \mathcal{A}_i, \\ q_2 & \text{otherwise.} \end{cases}$$

Proof. Analogous to **Lemma 1** in Yamamoto (2009b). ■

Suppose player j tries to infer whether player i sends $a_{i,m}(\omega_i) = a_i^G$ or a_i^B. After player j takes a_j and observes ω_j, she draws a random variable according to the uniform distribution on $[0,1]$. If the random variable is less than $\phi_{\{a_i^G\}}(a_j, \omega_j)$, we say $\Phi_j\left(\{a_i^G\}\right) = 1$. Otherwise, $\Phi_j\left(\{a_i^G\}\right) = 0$. **Lemma 9** guarantees no matter what action players $-(i,j)$ take, $\Pr(\{\Phi_j\left(\{a_i^G\}\right) = 1\}|a) = q_3$ if $a_i = a_i^G$ and $\Pr(\{\Phi_j\left(\{a_i^G\}\right) = 1\}|a) = q_2$ if $a_i = a_i^B$.

Player j infers that player i sends the message $a_{i,m}(\omega_i) = a_i^G$ if $\Phi_j\left(\{a_i^G\}\right) = 1$ occurs $\frac{q_2+q_3}{2}S$ times or more. Otherwise, she infers the message is a_i^B. After inferring $\{a_{i,m}(\omega_i)\}_{m=1}^M$, player j can infer the message about ω_i since $\{a_{i,m}(\omega_i)\}_{m=1}^M$ uniquely identifies ω_i. The law of large numbers guarantees that the message transmits correctly with high probability for sufficiently large S.

In addition, as we will see, players need to send the message about $t \in \{1, \ldots, KT\}$. Let us explain the protocol to send $t \in \{1, \ldots, KT\}$.

Message Protocol for t Let us attach a sequence of actions $a_i(t) = (a_{i,1}(t), \ldots, a_{i,\log_2 KT}(t))$ to each $t \in \{1, \ldots, KT\}$ with $a_{i,m}(t) \in \left\{a_i^G, a_i^B\right\}$ for all $m \in \{1, \ldots, \log_2 KT\}$ as follows:

- $a_{i,1}(t) = a_i^G$ if $t \in \{1, \ldots, KT/2\}$ and $a_{i,1}(t) = a_i^B$ otherwise. Similarly to the above, we attach $a_{i,1}(t) = a_i^G$ if t is in the first half of KT and $a_{i,1}(t) = a_i^B$ otherwise.

- Keep this procedure similarly to the above until we can identify t uniquely from $a_i(t)$. Note that we can make sure that $|a_i(t)| \leq \log_2 KT$.

When player i wants to send the message about t, she takes $a_{i,1}(t)$ for $T^{\frac{1}{2}}$ periods, then $a_{i,2}(t)$ for $T^{\frac{1}{2}}$ periods and so forth. Therefore, to send message about t, it takes $(\log_2 KT)\, T^{\frac{1}{2}}$ periods. Similarly to the above, player $j \neq i$ infers $a_{i,m}(t) = a_i^G$ if $\Phi_j\left(\{a_i^G\}\right)$ occurs $\frac{q_2+q_3}{2}T^{\frac{1}{2}}$ times or more while player i sends the message $a_{i,m}(t)$. Otherwise, she infers the message is a_i^B. Combining $\{a_{i,m}(t)\}_{m=1}^{\log_2 KT}$, player j can infer t.

Lemma 9 guarantees the following: (i) players $-(i,j)$ cannot change player j's inference about player i's message. (ii) Since $T^{\frac{1}{2}}$ is sufficiently larger than the total number of messages $\log_2 KT$, the law of large numbers guarantees that message t transmits correctly with high probability.

6.4 Equilibrium Strategies of the Block Game

With these protocols in hand, we can fully specify the equilibrium strategy of the block game.

States As before, each player has two possible states $x_i \in \{G, B\}$. When player i's state is G, player i takes the good strategy σ_i^G. On the other hand, when player i's state is B, player i takes the bad strategy σ_i^B. As we will see, σ_i^G and σ_i^B only differ in the coordination round and confirmation round.

Coordination Round As explained, in this round, player i reveals her own states x_i. If $x_i = G$, player i constantly takes a_i^G for T periods. If $x_i = B$, player i constantly takes a_i^B for T periods. Note that this round takes T periods and that each player sends the message simultaneously.

ith Monitoring Round for the Coordination Round As explained, in this round, player $j \neq i$ tries to infer player i's message in the coordination round by collecting the information from $n \neq i, j$ such that the conditional independence property is restored.

- For each j and $i \neq j$,
  - for each $l \neq i, j$, player j collects information from players $-(j, i, l)$ as follows:
    - for each $n \neq j, i, l$, player n alternatively sends the message about $\omega_{n,t}$ with t included in the coordination round. Let $\hat{\omega}_{n,t}$ be player j's inference about $\omega_{n,t}$.
    - In total, $(N-3)$ players $-(j, i, l)$ send the message. Let $\hat{\omega}_{-(j,i,l),t}$ be player j's inference about $\omega_{-(j,i,l),t}$, that is, $\hat{\omega}_{-(j,i,l),t} = \{\hat{\omega}_{n,t}\}_{n \neq j,i,l}$.
    - Player j infers player i's action in the coordination round based on the information from players $-(j, i, l)$ as explained below.
  - Hence, player j has $(N-2)$ inferences of player i's message in the coordination round depending on which player l is excluded.

Since it takes MST periods for each player n to send the message about $\omega_{n,t}$ and all the pairs exchange the messages alternatively, this round lasts for $(N-1)(N-2)(N-3)MST$ periods.

Now, we explain player j's inference about player i's message in the coordination round based on the information from players $-(j, i, l)$. The following lemma is useful:

Lemma 10 *Suppose that monitoring satisfies the identifiability and* $(N-2)$*-identifiability conditions. Then, there exists* S *such that there exist* $q_2 < q_3$ *such that, for all* $j, i, l \in I$*,*

$a \in A$ and $\mathcal{A}_i \subset A_i$, there exists a function $\gamma_{\mathcal{A}_i} : \Omega_j \times \Omega_{-(j,i,l)} \times A_j \to [0,1]$ such that

$$
\begin{aligned}
& \sum_{\omega} q\left(\omega_j, \omega_{-(j,i,l)} | a\right) \sum_{\hat{\omega}_{-(j,i,l)} \in \Omega_{-(j,i,l)}} \Pr(\hat{\omega}_{-(j,i,l)} | \omega_{-(j,i,l)}) \gamma_{\mathcal{A}_i}\left(\omega_j, \hat{\omega}_{-(j,i,l)}, a_j\right) \\
= & \sum_{\omega} q\left(\omega_j, \omega_{-(j,i,l)} | \omega_i, a\right) \sum_{\hat{\omega}_{-(j,i,l)} \in \Omega_{-(j,i,l)}} \Pr(\hat{\omega}_{-(j,i,l)} | \omega_{-(j,i,l)}) \gamma_{\mathcal{A}_i}\left(\omega_j, \hat{\omega}_{-(j,i,l)}, a_j\right) \\
= & \sum_{\omega} q\left(\omega_j, \omega_{-(j,i,l)} | \omega_l, a\right) \sum_{\hat{\omega}_{-(j,i,l)} \in \Omega_{-(j,i,l)}} \Pr(\hat{\omega}_{-(j,i,l)} | \omega_{-(j,i,l)}) \gamma_{\mathcal{A}_i}\left(\omega_j, \hat{\omega}_{-(j,i,l)}, a_j\right) \\
= & \begin{cases} q_3 & \text{if } a_i \in \mathcal{A}_i, \\ q_2 & \text{otherwise,} \end{cases}
\end{aligned}
$$

where $\Pr(\hat{\omega}_{-(j,i,l)} | \omega_{-(j,i,l)})$ is the probability that in the above message protocol, player j infers the message is $\hat{\omega}_{-(j,i,l)}$ when players $-(j,i,l)$ send the message $\omega_{-(j,i,l)}$.

Proof. The $(N-2)$-identifiability condition guarantees the existence for the case with $\Pr(\omega_{-(j,i,l)} | \omega_{-(j,i,l)}) = 1$, that is, for the case where the message transmits perfectly. Since the law of large numbers guarantees that the miscommunication occurs with sufficiently law probability for sufficiently large S, the result holds. ■

In **Lemma 10**, with abuse of notation, we assume q_2, q_3 are the same as **Lemma 9** for a simple notation. It is straightforward to consider the case where these are different.

Each player j constructs a random variable Γ_j as follows. After playing a_j, observing ω_j, and receiving the message $\hat{\omega}_{-(j,i,l)}$ from the players $-(j,i,l)$, player j draws a random variable according to the uniform distribution on $[0,1]$. If the realization is less than $\gamma_{\mathcal{A}_i}\left(\omega_j, \omega_{-(j,i,l)}, a_j\right)$, we say $\Gamma_{j,i,-(i,j,l)}(\mathcal{A}_i) = 1$. Otherwise, $\Gamma_{j,i,-(i,j,l)}(\mathcal{A}_i) = 0$.

Note that the first two equalities in **Lemma 10** imply the conditional independence property for player i sending the message in the coordination round and player l excluded from the players sending message to player j in the ith monitoring round for the coordination round. The last equality implies that no matter what actions players $-i$ take in the coordination round, the distribution of $\Gamma_{j,i,-(i,j,l)}(\mathcal{A}_i)$ is not changed assuming that every player tells the truth in the monitoring rounds.

Each player j, depending on the choices of $-(j,i,l)$, has $(N-2)$ pairs of two inferences $x_j^0(j,i,-(j,i,l))[1]$ and $x_j^0(j,i,-(j,i,l))[2]$. Player j infers player i takes a_i^G in the coordination round based on the information from players $-(j,i,l)$ if $\Gamma_{j,i,-(i,j,l)}(\{a_i^G\}) = 1$ happens more than $\frac{q_2+2q_3}{3}T$ times. Let $x^0(j,i,-(j,i,l))[1] = G$ denote this event. Similarly, player j infers player i takes a_i^B in the coordination round based on the information from players $-(j,i,l)$ if $\Gamma_{j,i,-(i,j,l)}(\{a_i^G\}) = 1$ happens less than $\frac{2q_2+q_3}{3}T$ times. Let $x^0(j,i,-(j,i,l))[1] = B$ denote this event. Otherwise, that is, if the frequency of $\Gamma_{j,i,-(i,j,l)}(\{a_i^G\}) = 1$ is in $\left[\frac{2q_2+q_3}{3}T, \frac{q_2+2q_3}{3}T\right]$, then player j infers she is not sure about player i's message: $x^0(j,i,-(j,i,l))[1] = E$.

In addition to $x^0(j,i,-(j,i,l))[1]$, player j has another inference $x^0(j,i,-(j,i,l))[2]$. Player j infers player i takes a_i^G if $\Gamma_{j,i,-(i,j,l)}(\{a_i^G\}) = 1$ happens more than $\frac{q_2+q_3}{2}T$ times: $x^0(j,i,-(j,i,l))[2] = G$. Otherwise, player j infers player i takes a_i^B, that is, $x^0(j,i,-(j,i,l))[2] = B$. Note that $x^0(j,i,-(j,i,l))[1] \in \{G,B\}$ implies $x^0(j,i,-(j,i,l))[1] = x^0(j,i,-(j,i,l))[2]$.

Confirmation Round As explained, this round is also regarded as a communication stage, in which each player j sends player j's own message in the coordination round again and reports her inferences about the messages from players $-j$ in the coordination round. That is, every player j reveals

$$\left(x_j, \left(x^0\left(j,i,-(j,i,l)\right)[1], x^0\left(j,i,-(j,i,l)\right)[2]\right)_{l\neq j,i,i\neq j}\right)$$

to the other players as follows:

- for each j, player j sends x_j. To sends the message x_j, player j takes $a_j^{x_j}$ for T periods.

  Then,

  - for each i,

    * for each l, player j sends $x^0\left(j,i,-(j,i,l)\right)[1] \in \{G,B,E\}$ using $2T$ periods. To send the message $x^0\left(j,i,-(j,i,l)\right) \in \{G,B\}$, player j takes $a_j^{x^0(j,i,-(j,i,l))[1]}$

for $2T$ periods. To send $x^0(j,i,-(j,i,l)) = E$, player j takes a_j^G for the first T periods and a_j^B for the second T periods.

* Then, for each l, player j sends $x^0(j,i,-(j,i,l))[2] \in \{G,B\}$ using T periods. To send the message $x^0(j,i,-(j,i,l))[2] \in \{G,B\}$, player j takes $a_j^{x^0(j,i,-(j,i,l))[2]}$ for T periods.

Since each player sends the message alternatively, it takes $(N + 3N(N-1)(N-2))T$ periods for all the players to send all the messages.

While player j sends the message x_j, each player $n \neq j$ takes a_n^G and infers that player j's message x_j is G if $\Phi_n(\{a_j^G\}) = 1$ happens more than $\frac{q_2+q_3}{2}T$ times. Otherwise, player n infers x_j is B. The inference for $x^0(j,i,-(j,i,l))[2]$ is similarly defined.

Let us consider the inference of $x^0(j,i,-(j,i,l))[1]$. While player j sends the message $x^0(j,i,-(j,i,l))[1]$, each player $n \neq j$ takes a_n^G and infers that player j's message x_j is G if $\Phi_n(\{a_j^G\}) = 1$ happens more than $\frac{q_2+q_3}{2}T$ times for both the first T periods and the second T periods. If $\Phi_n(\{a_j^G\}) = 1$ happens less than $\frac{q_2+q_3}{2}T$ times for both the first T periods and the second T periods, then player n infers $x^0(j,i,-(j,i,l))[1]$ is B. Otherwise, player n infers $x^0(j,i,-(j,i,l))[1]$ is E.

Let $\hat{x}_j$ be the inference of x_j. The inference for $x^0(j,i,-(j,i,l))[1]$ and $x^0(j,i,-(j,i,l))[2]$ are similarly denoted as $\hat{x}^0(j,i,-(j,i,l))[1]$ and $\hat{x}^0(j,i,-(j,i,l))[2]$. For notational simplicity, let $\hat{x}_n = x_n$, $\hat{x}^0(n,i,-(n,i,l))[1] = x^0(n,i,-(n,i,l))[1]$ and $\hat{x}^0(n,i,-(n,i,l))[2] = x^0(n,i,-(n,i,l))[2]$ be player n's inference of player n's own messages.

Note that player n's inferences in total are expressed as

$$\left(\hat{x}_j, \left(\left(\hat{x}^0(j,i,-(j,i,l))[1], \hat{x}^0(j,i,-(j,i,l))[2]\right)_{l\neq j,i}\right)_{i\neq j}\right)_j .$$

Let m_n^0 be this inference profile of player n. Let M_n^0 be the set of inference profiles m_n^0.

From the inference profile m_n^0, player n confirms the state profile $\tilde{x}$ as follows:

- for each i,

- if $\hat{x}^0(j,i,-(j,i,l))[1] = E$ for all $(j,i,-(j,i,l))_{j,l\neq i}$, then $\tilde{x}_i = G$.
- If there exists (j,l) with $\hat{x}^0(j,i,-(j,i,l))[1] = G$ and there does not exist (j,l) with $\hat{x}^0(j,i,-(j,i,l))[1] = B$, then $\tilde{x}_i = G$.
- Similarly, if there exists (j,l) with $\hat{x}^0(j,i,-(j,i,l))[1] = B$ and there does not exist (j,l) with $\hat{x}^0(j,i,-(j,i,l))[1] = G$, then $\tilde{x}_i = B$.
- If there exist (j,l) with $\hat{x}^0(j,i,-(j,i,l))[1] = B$ and (j',l') with $\hat{x}^0(j',i,-(j,i,l'))[1] = G$, then
  * if

$$
\begin{aligned}
& ((N-2)(N-3)+1) \times 1\{\hat{x}_i = G\} \\
& + \sum_{j\neq i}\sum_{l\neq j,i} 1\left\{\hat{x}^0(j,i,-(j,i,l))[2] = G\right\} \\
> \; & ((N-2)(N-3)+1) \times 1\{\hat{x}_i = B\} \\
& + \sum_{j\neq i}\sum_{l\neq j,i} 1\left\{\hat{x}^0(j,i,-(j,i,l))[2] = B\right\},
\end{aligned}
$$

    then $\tilde{x}_i = G$.
  * Otherwise, $\tilde{x}_i = B$.

Let $M_n^0(\tilde{x})$ denote the set of all $m_n^0 \in M_n^0$ such that $\tilde{x}$ is confirmed.

Consider the effect of player i's action on the confirmation of $\tilde{x}_i$ by player n. Since the conditional independence property holds for $(x^0(j,i,-(j,i,l)))_{j,l}$, after any player i's history at the end of the coordination round, the probability that there exist (j,l) with $\hat{x}^0(j,i,-(j,i,l))[1] = B$ and (j',l') with $\hat{x}^0(j',i,-(j,i,l'))[1] = G$ is negligible since $\frac{2q_2+q_3}{3}T$ and $\frac{q_2+2q_3}{3}T$ are sufficiently far away. Therefore, player i's message in the confirmation round cannot affect the confirmation of $\tilde{x}_i$ by player n with non-negligible probability. In addition, since x_i controls only player $i+1$'s continuation payoff, player i is indifferent whether $\tilde{x}_i = G$ is confirmed or $\tilde{x}_i = G$ is confirmed.

Consider the effect of player $l(\neq i,n)$'s action on the confirmation of $\tilde{x}_i$ by player n. Since the conditional independence property holds for $(x^0(j,i,-(j,i,l)))_j$, if player i's state is x_i, the probability that $x^0(j,i,-(j,i,l)) = x_i$ for all (j,l) is almost equal to 1 after any player l's history at the end of the coordination round. Then, x_i is confirmed with probability almost equal to 1 regardless of player l's messages in the monitoring round and confirmation round. To see this, note that $x^0(j,i,-(j,i,l)) = x_i$ for all (j,l) implies that we have either

- there exists (j,l) with $\hat{x}^0(j,i,-(j,i,l))[1] = x_i$ and there does not exist (j',l') with $\{G,B\} \ni \hat{x}^0(j',i,-(j',i,l'))[1] \neq x_i$, or
-

$$\begin{aligned} & ((N-2)(N-3)+1) \times 1\{\hat{x}_i = x_i\} \\ & + \sum_{j \neq i} \sum_{l \neq j,i} 1\left\{\hat{x}^0(j,i,-(j,i,l)) = x_i\right\} \\ > & ((N-2)(N-3)+1) \times 1\{\hat{x}_i \neq x_i\} \\ & + \sum_{j \neq i} \sum_{l \neq j,i} 1\left\{\hat{x}^0(j,i,-(j,i,l))[2] \neq x_i\right\}. \end{aligned}$$

In both cases, x_i is confirmed.

To give the incentive to tell the truth, we make player j's continuation payoff constant *whenever* there exists (n,i) such that player j's messages have an impact on the confirmation of $\tilde{x}_i$ by player n either in the ith monitoring round for the coordination round or the confirmation round. As we have seen, after any history at the end of the coordination round, the probability that player j's messages has an impact is negligible and this does not affect the equilibrium payoff or incentive to follow the equilibrium strategy in the coordination round.

kth Main Round This round lasts for KT periods. As we will see, if K is sufficiently large, the duration of the main round is much longer than that of the other rounds, so that the behavior in the main rounds almost solely determines the average payoff profile in the block game. Players' behavior in the main rounds is determined by the history in the confirmation round and the previous supplemental rounds. Thus, we postpone the explanation.

kth Monitoring Round This round is also a communication stage. In this round, player $j \neq i$ tries to infer player i's action in the kth main round collecting the information from $n \neq i, j$ such that the conditional independence property is restored.

- For each j and $i \neq j$,
  - for each $l \neq i, j$, player j collects information from players $-(j, i, l)$ as follows:
    - player j randomly picks $t_{-(j,i,l)} \in \{1, \ldots, KT\}$ and sends the message about $t_{-(j,i,l)}$ as explained above.
    - Each player $n \neq i, j, l$ infers $t_{-(j,i,l)}$ as explained above. Let $\hat{t}_{-(j,i,l)}(n)$ be player n's inference of $t_{-(j,i,l)}$.
    - Each player n alternatively sends message about the signal observations in the kth main round.
    - In player n's turn, player n firstly sends the message about $\omega_{n,t}$ such that t is the $\hat{t}_{-(j,i,l)}(n)$th period of the kth main round. She secondly sends the message about $\omega_{i,t}$ such that t is the $\hat{t}_{-(j,i,l)}(n)+1$th period of the kth main round and so forth. In total, player n sends the message $\omega_{n,t}$ such that t is the τth period in the kth main round with $\tau \in \{\hat{t}_{-(j,i,l)}(n), \hat{t}_{-(j,i,l)}(n)+1, \ldots, \hat{t}_{-(j,i,l)}(n)+\varepsilon KT\} \pmod{KT}$ with ε being sufficiently small.[15,16]
    - Player j believes that $t_{-(j,i,l)}$ transmits correctly after any history. While player j thinks player n sends the message $\omega_{n,t}$, player j infers $\omega_{n,t}$ as explained above. Note that there exists a small chance of miscoordination about t.
    - In total, $(N-3)$ players $-(j, i, l)$ send the message. Let $\hat{\omega}_{-(j,i,l),t}$ be player j's inference about the message.

For each j, i and l, player j sends the message about $t_{-(j,i,l)}$ for $(\log_2 KT)\, T^{\frac{1}{2}}$ periods. Responding to player j, each n sends the message about the signal observations in εKT

[15] Hence, if $\hat{t}_{-(i,j,l)} > (1-\varepsilon)\, KT$, we identify, for example, $\hat{t}_{-(i,j,l)}(n) + \varepsilon KT$ as $\hat{t}_{-(i,j,l)}(n) + \varepsilon KT - KT$.

[16] Note that we count periods from the beginning of the kth main round for τ and $t_{-(j,i,l)}$ while we count periods from the beginning of the block game for t.

periods in the kth main round, which takes $\varepsilon KTMS$ periods. Therefore, in total, the kth monitoring round lasts for $N(N-1)(N-2)\left((\log_2 KT)T^{\frac{1}{2}}+(N-3)\varepsilon KTMS\right)$ periods. Note that the message $t_{-(j,i,l)}$ determines which periods are used for players j and $-(j,i,l)$ to monitor player i. We only use ε fraction of the kth main round for monitoring. It is important since if players communicate about all the periods in the kth main round, then the kth monitoring round becomes too long and affects the equilibrium payoff.

Let us explain how player j infers player i's deviation The following lemma is useful:

Lemma 11 *Suppose that monitoring satisfies the identifiability and $(N-2)$-identifiability conditions. Then, there exists S such that there exist $0<q_1<q_2<q_3<1$ such that there exists $\overline{T}$ such that for all $T\geq\overline{T}$, for all $j,i,i',l\in I$ with $j\neq i,i'$ and $i\neq i'$, $a,\tilde{a}\in A$, $a_i\in A_i$, $a_{i'}\in A_{i'}$ and $\mathcal{A}_i\subset A_i$, there exist functions $\psi_{\mathcal{A}_i}:\Omega_j\times\Omega_{-(j,i,l)}\times A_j\to[0,1]$ and $\psi_{(a_i,a_{i'})}:\Omega_j\times\Omega_{-(j,i,i')}\times A_j\to[0,1]$ such that*

$$\begin{aligned}
&\sum_{\omega}q\left(\omega_j,\omega_{-(j,i,l)}|a\right)\sum_{\hat{\omega}_{-(j,i,l)}\in\Omega_{-(j,i,l)}}\Pr(\hat{\omega}_{-(j,i,l)}|\omega_{-(j,i,l)})\psi_{\mathcal{A}_i}\left(\omega_j,\hat{\omega}_{-(j,i,l)},a_j\right)\\
=&\sum_{\omega}q\left(\omega_j,\omega_{-(j,i,l)}|\omega_i,a\right)\sum_{\hat{\omega}_{-(j,i,l)}\in\Omega_{-(j,i,l)}}\Pr(\hat{\omega}_{-(j,i,l)}|\omega_{-(j,i,l)})\psi_{\mathcal{A}_i}\left(\omega_j,\hat{\omega}_{-(j,i,l)},a_j\right)\\
=&\sum_{\omega}q\left(\omega_j,\omega_{-(j,i,l)}|\omega_l,a\right)\sum_{\hat{\omega}_{-(j,i,l)}\in\Omega_{-(j,i,l)}}\Pr(\hat{\omega}_{-(j,i,l)}|\omega_{-(j,i,l)})\psi_{\mathcal{A}_i}\left(\omega_j,\hat{\omega}_{-(j,i,l)},a_j\right)\\
=&\begin{cases}q_3 & \text{if } a_i\in\mathcal{A}_i,\\ q_2 & \text{otherwise,}\end{cases}
\end{aligned}$$

$$
\begin{aligned}
&\sum_{\omega} q\left(\omega_j, \omega_{-(j,i,i')}|\tilde{a}\right) \sum_{\hat{\omega}_{-(j,i,i')}\in\Omega_{-(j,i,i')}} \Pr(\hat{\omega}_{-(j,i,i')}|\omega_{-(j,i,i')})\psi_{(a_i,a_{i'})}\left(\omega_j, \hat{\omega}_{-(j,i,i')}, \tilde{a}_j\right) \\
=& \sum_{\omega} q\left(\omega_j, \omega_{-(j,i,i')}|\omega_i, \tilde{a}\right) \sum_{\hat{\omega}_{-(j,i,i')}\in\Omega_{-(j,i,i')}} \Pr(\hat{\omega}_{-(j,i,i')}|\omega_{-(j,i,i')})\psi_{(a_i,a_{i'})}\left(\omega_j, \hat{\omega}_{-(j,i,i')}, \tilde{a}_j\right) \\
=& \sum_{\omega} q\left(\omega_j, \omega_{-(j,i,i')}|\omega_{i'}, \tilde{a}\right) \sum_{\hat{\omega}_{-(j,i,i')}\in\Omega_{-(j,i,i')}} \Pr(\hat{\omega}_{-(j,i,i')}|\omega_{-(j,i,i')})\psi_{(a_i,a_{i'})}\left(\omega_j, \hat{\omega}_{-(j,i,i')}, \tilde{a}_j\right) \\
=& \begin{cases} q_1 & \text{if } \tilde{a}_i = a_i \text{ and } \tilde{a}_{i'} \neq a_{i'}, \\ q_3 & \text{if } \tilde{a}_i \neq a_i \text{ and } \tilde{a}_{i'} = a_{i'}, \\ q_2 & \text{otherwise,} \end{cases}
\end{aligned}
$$

where $\Pr(\hat{\omega}_{-(j,i,l)}|\omega_{-(j,i,l)})$ *is defined in the same way as in* ***Lemma 10****.*

Proof. The same as **Lemma 10**. ■

Here, again, with abuse of notation, we assume q_2, q_3 are the same as **Lemma 9**. It is straightforward to consider the case where these are different. In addition, for the rest of the paper, if we say j, i, i', we implicitly assume $j \neq i, i'$ and we identify (i, i') and (i', i).

Let us assume the message about $t_{-(j,i,l)}$ transmits perfectly for a while. Consider the following construction of random variables. Suppose player j tries to monitor player i's action. After playing $a_{j,t}$ in the τth period of the main round, (i) if $\tau \in \{t_{-(j,i,l)}, t_{-(j,i,l)} + 1, \ldots, t_{-(j,i,l)} + \varepsilon KT\}$, that is, if j decides to use the τth period to monitor player i based on players $-(j,i,l)$'s information, then, after observing $\omega_{j,t}$ and receiving the message $\hat{\omega}_{-(j,i,l),t}$ from the players $-(j,i,l)$, player j draws a random variable according to the uniform distribution on $[0,1]$. If the realization is less than $\psi_{\{a_i\}}(\omega_{j,t}, \hat{\omega}_{-(j,i,l),t}, a_{j,t})$, we say $\tilde{\Psi}_{j,i,-(j,i,l),t}(\{a_i\}) = 1$. Otherwise, $\tilde{\Psi}_{j,i,-(j,i,l),t}(\{a_i\}) = 0$. (ii) if $\tau \notin \{t_{-(j,i,l)}(n), t_{-(j,i,l)}(n) + 1, \ldots, t_{-(j,i,l)}(n) + \varepsilon KT\}$, $\tilde{\Psi}_{j,i,-(j,i,l),t}(\{a_i\}) = 0$.

Similarly, consider player j tries to monitor a pair (i, i') based on the information from players $-(j,i,i')$. (i) if $\tau \in \{t_{-(j,i,i')}, t_{-(j,i,i')} + 1, \ldots, t_{-(j,i,i')} + \varepsilon KT\}$, that is, if j decides to use the τth period to monitor players (i,i') based on players $(-j,i,i')$'s information, then, after observing $\omega_{j,t}$ and receiving the message $\hat{\omega}_{-(j,i,i'),t}$ from the players $-(j,i,i')$, player j draws a random variable according to the uniform distribution on $[0,1]$. If the realization is less than $\psi_{(a_i,a_{i'})}(\omega_{j,t}, \hat{\omega}_{-(j,i,i'),t}, a_{j,t})$, we say $\tilde{\Psi}_{j,(i,i'),-(j,i,i'),t}(a_i, a_{i'}) = 1$. Otherwise, $\tilde{\Psi}_{j,(i,i'),-(j,i,i'),t}(a_i, a_{i'}) = 0$. (ii) if $\tau \notin \{t_{-(j,i,i')}(n), t_{-(j,i,i')}(n)+1, \ldots, t_{-(j,i,i')}(n) + \varepsilon KT\}$, $\tilde{\Psi}_{j,(i,i'),-(j,i,i'),t}(a_i, a_{i'}) = 0$.[17]

To consider how player j infers the other players' actions in the kth main round based on $\{\Psi_{j,i,-(j,i,l),t}\}$ and $\{\Psi_{j,(i,i'),-(j,i,i'),t}\}$, we define the following variables as Yamamoto (2009b). Let $F_1(\tau, T, r)$ be the probability that $\tilde{\Psi}_{j,i,-(j,i,l),t}(\{a_i\}) = 1$ is counted r times during a T-period interval when player i chooses some $\tilde{a}_i \neq a_i$ in the first τ periods and then a_i in the remaining $T - \tau$ periods.[18] Let $F_2(\tau, T, r)$ be the probability that $\tilde{\Psi}_{j,(i,i'),-(j,i,i'),t}(a_i, a_{i'}) = 1$ is counted r times during a T-period interval when player i chooses $\tilde{a}_i \neq a_i$ in the first τ periods and then chooses a_i in the remaining $T - \tau$ periods, while player i' chooses $a_{i'}$ constantly.[19] Let $(Z_T)_{T=1}^{\infty}$, $(Z'_T)_{T=1}^{\infty}$ and $(Z''_T)_{T=1}^{\infty}$ be sequences of integers such that

$$\begin{aligned}
Z''_T &< \varepsilon q_2 < Z'_T, \\
Z_T &\leq \varepsilon q_3 T, \\
\lim_{T\to\infty} \sum_{r=Z''_T+1}^{Z'_T} F_1(T, T, r) &= \lim_{T\to\infty} \sum_{r=Z''_T+1}^{Z'_T} F_2(0, T, r) = \lim_{T\to\infty} \sum_{r>Z_T} F_1(0, T, r) = 1, \\
\lim_{T\to\infty} \left| \frac{Z''_T}{T} - \varepsilon q_2 \right| &= \lim_{T\to\infty} \left| \frac{Z'_T}{T} - \varepsilon q_2 \right| = \lim_{T\to\infty} \left| \frac{Z_T}{T} - \varepsilon q_3 \right| = 0,
\end{aligned}$$

and

$$\lim_{T\to\infty} T F_1(0, T-1, Z_T) = \infty.$$

The existence of $(Z_T)_{T=1}^{\infty}$, $(Z'_T)_{T=1}^{\infty}$ and $(Z''_T)_{T=1}^{\infty}$ is assured by the law of large numbers as Yamamoto (2009b). Notice that, even though the probability of $\tilde{\Psi}_{j,i,-(j,i,l),t}(\{a_i\}) = 1$ is q_2 *given that* j decides to use that period to monitor player i based on $-(j,i,l)$, since the probability that the specific period t is selected is ε, the overall probability is εq_2. Therefore, compared to Yamamoto (2009b), we multiply q_2 and q_3 by ε. The same is true

[17] Note that we count periods from the beginning of the kth main round for τ and $t_{-(j,i,l)}$ while we count periods from the beginning of the block game for t.

[18] Note that this probability is independnet of j, i, l and a_{-i}.

[19] Again, this probability is independnet of j, i, i' and $a_{-(i,i')}$.

for $\tilde{\Psi}_{j,(i,i'),-(j,i,i'),t}(a_i, a_{i'})$.

Then, player j monitors player i's action as follows. Suppose player j confirms x, that is, $m_j^0 \in M_j^0(x)$, in the confirmation round. Then, player j tries to infer whether a pair of players (i, i') has taken $a_i^{x,k}, a_{i'}^{x,k}$ or not in the kth main round. If $\tilde{\Psi}_{j,(i,i'),-(j,i,i'),t}(a_i, a_{i'}) = 1$ occurs more than Z'_{KT} times, player j infers that player i does not take $a_i^{x,k}$ in the kth main round while player i' takes $a_{i'}^{x,k}$ based on the information from players $-(j, i, i')$. Let $x^k(j, (i, i'), -(j, i, i')) = i$ denote this event. Similarly, if $\tilde{\Psi}_{j,(i,i'),-(j,i,i'),t}(a_i, a_{i'}) = 1$ occurs less than Z''_{KT} times, player j infers that player i takes $a_i^{x,k}$ while player i' does not take $a_{i'}^{x,k}$ based on the information from players $-(j, i, i')$. Let $x^k(j, (i, i'), -(j, i, i')) = i'$ denote this event. Otherwise, we say $x^k(j, (i, i'), -(j, i, i')) = 0$ (nobody deviates).

So far, we assume $t_{-(j,i,l)}$ *transmits perfectly.* **Lemma 10** implies that the conditional independence property is restored for $\tilde{\Psi}_{j,i,-(j,i,l),t}(\{a_i\})$ and $\tilde{\Psi}_{j,(i,i'),-(j,i,i'),t}(a_i, a_{i'})$. Now, let us consider the situation where players send the messages and infer $t_{-(j,i,l)}$ as explained above. Even though there exists a positive probability of miscoordination, since $T^{\frac{1}{2}}$ is sufficiently large compared to $\log_2 KT$, the miscoordination occurs with very small probability. Mathematically, we can prove the following. Let $\Psi_{j,i,-(j,i,l),t}(\{a_i\})$ and $\Psi_{j,(i,i'),-(j,i,i'),t}(a_i, a_{i'})$ be the random variables constructed in the same way as $\tilde{\Psi}_{j,i,-(j,i,l),t}(\{a_i\})$ and $\tilde{\Psi}_{j,(i,i'),-(j,i,i'),t}(a_i, a_{i'})$ but assuming $t_{-(j,i,l)}$ does not transmit perfectly. Then, the following lemma holds:

Lemma 12 *Under the inference protocol explained above, for any $j, i, i', l \in I$, $a_i \in A_i$, $a_{i'} \in A_{i'}$, $t \in \{1, \ldots, KT\}$ and any action sequence a^{KT} and signal observations ω^{KT} in the kth main round,*

$$\left| \begin{array}{c} \Pr\left(\left\{\Psi_{j,i,-(j,i,l),t}(\{a_i\}) = 1\right\} \mid a^{KT}, \omega^{KT}\right) \\ -\Pr\left(\left\{\tilde{\Psi}_{j,i,-(j,i,l),t}(\{a_i\}) = 1\right\} \mid a^{KT}, \omega^{KT}\right) \end{array} \right| = o\left(T^2\right),$$

$$\left| \begin{array}{c} \Pr(\{\Psi_{j,(i,i'),-(j,i,i'),t}(a_i, a_{i'}) = 1\} \mid a^{KT}, \omega^{KT}) \\ -\Pr(\{\tilde{\Psi}_{j,(i,i'),-(j,i,i'),t}(a_i, a_{i'}) = 1\} \mid a^{KT}, \omega^{KT}) \end{array} \right| = o\left(T^2\right)$$

Proof. From the Hoeffding's inequality, for any $\varepsilon > 0$, the probability that $t_{-(i,j,l)}$ does not transmit perfectly is less than

$$\log_2 KT \exp\left(T^{-1+\varepsilon}\right)$$

for sufficiently large T. Therefore, since $\log_2 KT \exp\left(T^{-1+\varepsilon}\right) = o(T^2)$, the lemma holds. ■

Since the conditional independence property holds for $\tilde{\Psi}_{j,i,-(j,i,l),t}\left(\{a_i\}\right)$ and $\tilde{\Psi}_{j,(i,i'),-(j,i,i'),t}\left(a_i, a_{i'}\right)$, the "almost conditional independence" property holds for $\Psi_{j,i,-(j,i,l),t}\left(\{a_i\}\right)$ and $\Psi_{j,(i,i'),-(j,i,i'),t}\left(a_i, a_{i'}\right)$. As Yamamoto (2009b) shows, this is sufficient for the equilibrium construction to work.

kth Supplemental Round As explained, this round is also regarded as a communication stage, in which each player i reports her inferences in the kth monitoring round. That is, every player j reveals $\left(x^k\left(j,(i,i'),-(j,i,i')\right)\right)_{i,i'\neq j}$ to the other players as follows:

- For each j,
  - for each (i,i'), player j sends $x^k\left(j,(i,i'),-(j,i,i')\right) \in \{i,i',0\}$ using $2T$ periods. To sends the message $x^k\left(j,(i,i'),-(j,i,i')\right) = i$ with $i < i'$,[20] player i takes a_i^G for the first T periods and then a_i^B for the second T periods. To send $x^k\left(j,(i,i'),-(j,i,i')\right) = i'$, player i takes a_i^B for the first T periods and then a_i^G for the second T periods. To send $x^k\left(j,(i,i')-(j,i,i')\right) = 0$, player j takes a_i^B for $2T$ periods constantly.

Each player sends the message alternatively here. Each player j has $(N-1)(N-2)/2$ inference depending on which pair (i,i') is monitored. Hence, it takes $N(N-1)(N-2)T$ periods for all the players to send the message.

While player j sends the message $x^k\left(j,(i,i'),-(j,i,i')\right)$ with $i < i'$, each player $n \neq j$ takes a_n^G and infers that player j's message is i if $\Phi_n(\{a_j^G\}) = 1$ happens more than $\frac{q_2+q_3}{2}T$ times for the first T periods and less than $\frac{q_2+q_3}{2}$ for the second T periods. On the other hand, player n infers that player j's message is i' if $\Phi_n(\{a_j^G\}) = 1$ happens less than $\frac{q_2+q_3}{2}T$ times

[20] Since the order for (i,i') does not matter, for notational simplicity, we assume $i < i'$ here.

for the first T periods and more than $\frac{q_2+q_3}{2}$ for the second T periods. Otherwise, player n infers player j's message is 0.

Since player n receives the message from all the players $-n$, player n's inferences in total are expressed as $\left(\left(\hat{x}^k\left(j,\left(i,i'\right),-\left(j,i,i'\right)\right)\right)_{i,i'\neq j}\right)_j$ (As before, $\left(\hat{x}^k\left(n,\left(i,i'\right),-\left(n,i,i'\right)\right)\right)_{i,i'\neq n}$ is player n's own message). Let $m_n^k \equiv \left(\left(\hat{x}^k\left(j,\left(i,i'\right),-\left(j,i,i'\right)\right)\right)_{i,i'\neq j}\right)_j$ be the inference profile of player n. Let M_n^k be the set of inference profiles m_n^k.

We say that player i's deviation is confirmed by player n if and only if the following is true: according to the inference profile m_n^k,

$$\#\left\{j,i' : i\neq j\neq i'\neq i \text{ and } \hat{x}^k\left(j,\left(i,i'\right),-\left(j,i,i'\right)\right)=i\right\} > (N-2)^2$$

and for all $\iota\neq i$,

$$\#\left\{j,i' : \iota\neq j\neq i'\neq \iota \text{ and } \hat{x}^k\left(j,\left(\iota,i'\right),-\left(j,\iota,i'\right)\right)=\iota\right\} \leq (N-2)^2 .$$

Then, for each $i\in I$, let $M_n^k(i)$ denote the set of all $m_n^k\in M_n^k$ such that player i's deviation is confirmed. Notice that there is at most one player whose deviation is confirmed by n. If no deviation is confirmed, we say $m_n^k\in M_n^k(0)$, that is, $M_n^k(0)$ is the set of inferences according to which no deviation is confirmed.

Consider intuitively when player i's deviation is confirmed. There are $(N-1)$ possible choices about who monitors i (the choice of j in $x^k\left(j,\left(i,i'\right),-\left(j,i,i'\right)\right)$). Each monitor has $(N-2)$ choices about with whom player i is paired (the choice of i' in $x^k\left(j,\left(i,i'\right),-\left(j,i,i'\right)\right)$). Therefore, there are $(N-1)(N-2)$ "votes" for player i's deviation in the supplemental round. We confirm player i's deviation if and only if player i is an only person who collects more than $(N-2)^2$ votes.

Let us verify that, as long as the other players follow the equilibrium strategy, player i cannot create the situation that for some $\iota\neq i$,

$$\#\left\{j,i' : \iota\neq j\neq i'\neq \iota \text{ and } \hat{x}^k\left(j,\left(\iota,i'\right),-\left(j,\iota,i'\right)\right)=\iota\right\} > (N-2)^2$$

with non-negligible probability. For each $\iota \neq i$, player i can control $\left(\hat{x}^k\left(i,\left(\iota,i'\right),-\left(i,\iota,i'\right)\right)\right)_{i'}$ votes. Therefore, player i can control $(N-2)$ votes. In addition, by changing the message in the monitoring round, player i can control $\left(\hat{x}^k\left(j,\left(\iota,i'\right),-\left(j,\iota,i'\right)\right)\right)_{j,i'\neq i}$. Therefore, player i can also control $(N-2)(N-3)$ votes. In total, player i can control $(N-2)^2$ votes. However, this is not enough to induce $\#\left\{j,i' : \iota \neq j \neq i' \neq \iota \text{ and } \hat{x}^k\left(j,\left(\iota,i'\right),-\left(j,\iota,i'\right)\right)=\iota\right\} > (N-2)^2$, which implies that player i cannot prevent others from confirming player i's deviation if player i deviates in the main rounds.

To give the incentive to tell the truth in the kth monitoring and supplemental rounds, whenever player i's messages in these rounds affect the confirmation of the deviation of any player, we make player i's continuation payoff constant. Whether player i's messages have an impact or not depends only on $\hat{x}\left(j,\left(\iota,\iota'\right),-\left(j,\iota,\iota'\right)\right)_{j\neq i,i\in(\iota,\iota')}$, the conditional independence property holds for player i.

kth main round Now we can specify the action in the kth main round. For player i with $m_i^0 \in M_i^0(\tilde{x})$, that is, if $\tilde{x}$ is confirmed, player i takes $a_i^{\tilde{x},1}$ in the first main round. For $k \geq 2$, if there exists $\tilde{k} \leq k-1$ such that $(m_i^0,\ldots,m_i^{\tilde{k}-1}) \in M_i^0(\tilde{x}) \times M_i^1(0) \times \cdots \times M_i^{\tilde{k}-1}(0)$ and $m_i^{\tilde{k}} \in M_i^{\tilde{k}}(j)$, then player i takes $\underline{a}_i^j(\mathcal{A})$ and $\overline{a}_i^j(\mathcal{A})$ if $\tilde{x}_{j-1}=B$ and G, respectively. Intuitively, if $\tilde{x}$ is confirmed in the confirmation round and if there is player j whose unilateral deviation from $a_j^{\tilde{x}}$ was firstly confirmed in a previous round, then player i takes the punishment action $\underline{a}_i^j(\mathcal{A})$ if player $j-1$, whose state controls player j's equilibrium payoffs, is in the bad state and takes the reward action $\overline{a}_i^j(\mathcal{A})$ if player $j-1$ is in the good state.

Report round This round is also regarded as a communication stage and each player reports what inference profiles they had in the confirmation round and the supplemental rounds according to her private history. Thus, the message space for player i is $J_i \equiv (M_i^0 \times M_i^1 \times \cdots \times M_i^K)$. Here, players send their messages alternately and send M_i^1, ..., and M_i^K similarly to before. Note that M_i^0 is included into M_i^1 and so is omitted. How to send messages about M_i^1, ..., and M_i^K is analogous to the confirmation round and the supplemental rounds. Therefore, each player spends $N+3N(N-1)(N-2)T+N(N-1)(N-2)KT$

periods. Thus, this round lasts for $N^2 + 3N^2(N-1)(N-2)T + N^2(N-1)(N-2)KT$ periods in total.

Determinant of T_b Therefore, in total,

$$\begin{aligned} T_b &= \{1 + N(N-1)(N-2)(N-3)MS + 3(N-1)(N-2)\}T \\ &\quad + K\left\{K + N(N-1)(N-2)\left((\log_2 KT)T^{-\frac{1}{2}} + (N-3)MS\varepsilon K\right) + N(N-1)(N-2)\right\}T \\ &\quad + \left\{N^2 + 3N^2(N-1)(N-2) + N^2(N-1)(N-2)K\right\}T. \end{aligned}$$

As we will see in the Appendix, each player i is

- indifferent for any message in the initial period of the coordination round conditional on the other players' actions. Once player i decides which message to send, it is optimal to take a constant action. Therefore, the initial period of the coordination round is included in $(t_l + 1)_{l=0}^{\infty}$ in **Definition 6**.

- indifferent for any message in each monitoring round for the coordination round, confirmation round, monitoring rounds, supplemental rounds and report round conditional on the other players' actions. Therefore, each period in these rounds is included in $(t_l + 1)_{l=0}^{\infty}$ in **Definition 6**.

- indifferent for any action in $\mathcal{A}_i$ in the initial period in each main round conditional on the other players' actions, but a constant action is optimal in a given main round once player i decides which action she takes. Therefore, the initial period in each main round is included in $(t_l + 1)_{l=0}^{\infty}$ in **Definition 6**.

6.5 Equilibrium Strategy in the Repeated Game

So far, we focused on the strategies in the block game. Here, let us explain how to construct the equilibrium strategy in the infinitely repeated games from the strategies in the block game. Since we have recovered the conditional independence property for the block game, the rest of the paper is the same as Yamamoto (2009b).

Before the analysis of infinitely repeated games, it is useful to consider the following T_b-period repeated game with transfers. Let $\Sigma_i^{T_b}$ be the set of player i's strategies in the T_b-period block game. Suppose that player i receives a transfer $U_i : H_{i-1}^{T_b} \to \mathbb{R}$ after the T_b-period block game. Let $w_i^A\left(\sigma^{T_b}, U_i\right)$ denote player i's average payoff in this *auxiliary scenario* when the players perform a block strategy profile $\sigma^{T_b} \in \Sigma^{T_b}$, that is,

$$w_i^A\left(\sigma^{T_b}, U_i\right) \equiv \frac{1-\delta}{1-\delta^{T_b}}\left[\sum_{t=1}^{T_b} \delta^{t-1} E\left[\pi_i\left(a_t\right) \mid \sigma^{T_b}\right] + \delta^{T_b} E[U_i(h_{i-1}^{T_b}) \mid \sigma^{T_b}]\right].$$

Let $\sigma_i^{T_b} \mid h_i^t$ denote player i's continuation strategy after history $h_i^t \in H_i^t$ induced by $\sigma_i^{T_b} \in \Sigma_i^{T_b}$. Also, let $BR^A(\sigma_{-i}^{T_b} \mid h_{-i}^t, U_i)$ be the set of player i's best replies in the auxiliary scenario continuation game from period $t+1$ on, given that opponents play $\sigma_{-i}^{T_b} \in \Sigma_{-i}^{T_b}$ in the block game and that their past history is $h_{-i}^t \in H_{-i}^t$.

Let U_i^B and U_i^G be as in **Lemmas 13** and **14**.

Lemma 13 *Suppose that monitoring satisfies the full support, identifiability and* $(N-2)$*-identifiability conditions and that the stage game satisfies the full dimensionality condition. Then, there exist* K *and* $\bar{T}$ *such that for all* $T > \bar{T}$*, there exists* $\overline{\delta} < 1$ *such that for all* $\delta \in \left(\overline{\delta}, 1\right)$ *and for all* $i \in I$*, there exists* $U_i^B : H_{i-1}^{T_b} \to \mathbb{R}$ *such that for all* $l \geq 0$*,* $h^{t_l} \in H^{t_l}$*,* $h_{i-1}^{T_b} \in H_{i-1}^{T_b}$ *and* $x \in X$ *with* $x_{i-1} = B$*,*

$$\sigma_i^{x_i} \mid h_i^{t_l} \in BR^A(\sigma_{-i} \mid h_{-i}^{t_l}, a_{-i}) \text{ for all } a_{-i} \in \operatorname{supp}\left(\sigma_{-i}^{x_{-i}} \mid h_i^t\right), \tag{7}$$

$$w_i^A\left(\sigma^x, U_i^B\right) = \underline{w}_i, \tag{8}$$

$$0 \leq U_i^B(h_{i-1}^{T_b}) < \tfrac{\overline{w}_i - \underline{w}_i}{1-\delta}. \tag{9}$$

Lemma 14 *Suppose that monitoring satisfies the full support, identifiability and* $(N-2)$*-identifiability conditions and that the stage game satisfies the full dimensionality condition. Then, there exist* K *and* $\bar{T}$ *such that for all* $T > \bar{T}$*, there exists* $\overline{\delta} < 1$ *such that for all*

$\delta \in (\overline{\delta}, 1)$ *and for all* $i \in I$*, there exists* $U_i^G : H_{i-1}^{T_b} \to \mathbb{R}$ *such that for all* $l \geq 0$*,* $h^{t_l} \in H^{t_l}$*,* $h_{i-1}^{T_b} \in H_{i-1}^{T_b}$ *and* $x \in X$ *with* $x_{i-1} = G$*,*

$$\sigma_i^{x_i} \mid h_i^{t_l} \in BR^A(\sigma_{-i} \mid h_{-i}^{t_l}, a_{-i}) \text{ for all } a_{-i} \in \operatorname{supp}\left(\sigma_{-i}^{x_{-i}} \mid h_i^t\right), \tag{10}$$

$$w_i^A\left(\sigma^x, U_i^G\right) = \overline{w}_i, \tag{11}$$

$$-\tfrac{\overline{w}_i - \underline{w}_i}{1-\delta} < U_i^G(h_{i-1}^{T_b}) \leq 0. \tag{12}$$

Proof. See Appendix. ■

Notice that **Lemma 13** ensures that, if player $i-1$ is taking the bad strategy, (that is, if player $i-1$'s state is B), the other players $-i$ take either the good strategy or bad strategy, and player $i-1$ uses the reward function U_i^B, then both the good strategy and the bad strategy are optimal for player i and give the payoff $\underline{w}_i$. On the other hand, **Lemma 14** ensures that, if player $i-1$ is taking the good strategy, (that is, if player $i-1$'s state is G), the other players $-i$ take either the good strategy or bad strategy, and player $i-1$ uses the reward function U_i^G, then both the good strategy and the bad strategy are optimal for player i and give the payoff $\overline{w}_i$.

Hence, to prove **Theorem 8**, it suffices to construct each player $i-1$'s strategy σ_{i-1} in the infinitely repeated game so that the state transition ensures that player i's continuation payoff in the infinitely repeated game without auxiliary scenario is actually equal to the continuation payoff in the block game with auxiliary scenario.

Formally, player $i-1$'s strategy $\sigma_{i-1}(v)$ in the infinitely repeated game is specified by the following automaton with the initial "intermediate state" $v \in \prod_{i \in I} [\underline{w}_i, \overline{w}_i]$.

Intermediate State w_i Go to state B with probability p_{i-1} and go to state G with probability $1 - p_{i-1}$ such that $w_i = p_{i-1}\underline{w}_i + (1 - p_{i-1})\overline{w}_i$.

State B Play the block strategy σ_{i-1}^B for T_b periods. After that, go to the intermediate state w_i given by $w_i = \underline{w}_i + (1-\delta) U_i^B(h_{i-1}^{T_b})$.

State G Play the block strategy σ_{i-1}^{G} for T_b periods. After that, go to the intermediate state w_i given by $w_i = \overline{w}_i + (1-\delta)\, U_i^G(h_{i-1}^{T_b})$.

From **Lemmas 13** and **14** and the one shot deviation principle, $\sigma(v) = (\sigma_i(v))_{i\in I}$ is a Nash equilibrium. From Sekiguchi (1997), there exists a realization equivalent sequential equilibrium.

6.6 Comparison with Yamamoto (2009b)

Let us comment on how our equilibrium construction is different from Yamamoto (2009b) to clarify our new contribution. Our equilibrium construction solves two problems: *restoring the conditional independence property* and *constructing BFRSE*.

Restoring the Conditional Independence Property To restoring the conditional independence property, the monitoring rounds for the coordination round and the main rounds are introduced in addition to the coordination, confirmation, main, supplemental and report rounds that are shared with Yamamoto (2009b).

Player i's continuation payoff is determined by the following four events: (i) what state profile is confirmed by others, (ii) whether there exists (n, j) such that player i's messages in the monitoring rounds for the coordination round and the confirmation round have an impact on the confirmation of x_j by player n, (iii) whether player i's deviation is confirmed, and (iv) whether player i's messages have an impact on the confirmation of the other players' deviation. Give that player i's messages do not have an impact on the confirmation of the states as explained in (ii), (i) is out of player i's control.[21] As explained, if either (ii) or (iv) is true, to give the incentive to tell the truth, player i's continuation payoff is constant. Since (ii), (iii) and (iv) are determined by the message exchange where player i is excluded. Therefore, to restore the conditional independence property for all (i), (ii), (iii) and (iv) based on the exchanges of the messages, we construct random variables to monitor player i

[21] More precisely, player i can control whether $x_i = G$ is confirmed by all the players with probability close to 1 or B is confirmed by all the players with probability close to 1 via changing player i's action in the coordination round. However, player i is indifferent since x_i controls only player $i+1$'s continuation payoff.

with the following properties:

- player i's signal observations have no information about the random variables used to monitor herself, that is, $x(j,i,-(j,i,l))$ with $j,l \neq i$ and $x(j,(\iota,\iota'),-(j,\iota,\iota'))$ with $i \in (\iota,\iota')$. This restores the conditional independence property about *whether* $x_i = G_i$ *is confirmed or* $x_i = B_i$ *is confirmed* and *whether player i's deviation is confirmed.*

- player i's signal observations have no information about the random variables that player $n \neq i$ uses to monitor player $j \neq i, n$ when player i is excluded from the players to send the message to n, that is, $x(n,j,-(n,j,i))$ and $x(n,(i,j),-(n,j,i))$. This restores the conditional independence property about *whether player i's messages have an impact on the confirmation of* x_j and *whether player i's messages have an impact on the confirmation of player j's deviation.*

Constructing BFRSE Once players obtain the conditionally independent signals, we can use Yamamoto (2009b)'s construction with the following modifications: in both Yamamoto (2009b)'s equilibrium and ours, we give the incentive for each player i to send the message truthfully as follows: if player i's messages have an impact on the results of the communication stages,[22] player i's continuation payoff is constant. To keep the efficiency, this cannot happen with non-negligible probability. We construct the equilibrium such that if the messages transmit almost perfectly on the equilibrium path, no player can affect the result of the communication stages. Then, by the law of large numbers, no player can affect the result of the communication stages with non-negligible probability.

Since we introduce two additional communication stages, the number of votes under each player's control increases, which makes it harder to sustain the above property. This is the reason why we require $N \geq 4$ and our "voting mechanism" in the confirmation and supplemental rounds are more complicated than in Yamamoto (2009b).

[22] As before, more precisely, player i can control whether $x_i = G$ is confirmed by all the players with probability close to 1 or B is confirmed by all the players with probability close to 1.

7 Efficiency Result and Folk Theorem

Since we construct BFRSE to attain the same equilibrium payoff set as Yamamoto (2009b), a sufficient condition for the efficiency result in Yamamoto (2009b) is also valid for our equilibrium:

Proposition 15 *Suppose that the feasible payoff set is full dimensional and that there are profiles $a^* \in A$ and $a^{**} \in A$ such that $\max_{a_i \in A_i} \pi_i\left(a_i, a_{-i}^{**}\right) < \pi_i\left(a^*\right) < \pi_i\left(a_i^*, a_{-i}^{**}\right)$ for all $i \in I$. Then the full dimensionality condition is satisfied and the payoff vector $\pi\left(a^*\right)$ is an element of $\bigcup_{p \in \triangle \mathcal{J}} \left(V\left(p\right) \cap \prod_{i \in I} \left[p\underline{v}_i, p\overline{v}_i\right]\right)$. Therefore, if the monitoring structure satisfies the full support condition, the identifiability condition and the $(N-2)$-identifiability condition, then $\pi\left(a^*\right) \in \lim_{\delta \to 1} E\left(\delta\right)$.*

Proof. Similar to Yamamoto (2009b). ■

In addition, the next proposition assures that the folk theorem holds for the games with prisoners'-dilemma structure.

Definition 16 *The stage game is an N-player prisoners' dilemma if $A_i = \{C_i, D_i\}$ for all $i \in I$, $\pi_i\left(D_i, a_{-i}\right) \geq \pi_i\left(C_i, a_{-i}\right)$ for all $i \in I$ and $a_{-i} \in A_{-i}$, $\pi_i\left(C_j, a_{-j}\right) \geq \pi_i\left(D_j, a_{-j}\right)$ for all $j \neq i$ and $a_{-j} \in A_{-j}$ and $\pi_i\left(C_1, \ldots, C_N\right) > \pi_i\left(D_1, \ldots, D_N\right)$ for all $i \in I$.*

Proposition 17 *Suppose that the stage game is an N-player prisoners' dilemma with $N \geq 4$ and that the full support, the identifiability and the $(N-2)$-identifiability conditions are satisfied. Then, $\lim_{\delta \to 1} E\left(\delta\right)$ is equal to the feasible and individually rational payoff set.*

Proof. Similar to Yamamoto (2009b). ■

8 Appendix

8.1 Proof of Lemma 4

Note that the full support condition is generic. The identifiability condition is generically satisfied if $|A_{-j}| \leq |\Omega_j|$. The $(N-2)$-identifiability condition is generically satisfied for

$(j,i,l) \subset I$ with $i \neq j \neq l \neq i$ if $|A_{-j}| \times (|\Omega_i| + |\Omega_l|) \leq \left|\Omega_{-(i,l)}\right|$.

8.2 Proof of Theorem 8

From Section 6.5, it suffices to show that **Lemmas 13** and **14** hold. The basic structure and explanation is the same as Yamamoto (2009b) since we restore the conditional independence property by introducing the monitoring round for the communication rounds and main rounds.

8.2.1 Average Payoff with Perfect Monitoring

Let $\mathcal{S}_i^{T_b}$ be the set of all $\sigma_i^{T_b} \in \Sigma_i^{T_b}$ such that player i chooses a constant action from the recommended set $\mathcal{A}_i^k$ in the kth main round, for each $k \in \{1, \dots, K\}$ and for each history up to the beginning of the kth main round. In addition, let $w_i^P\left(\sigma^{T_b} \mid \delta\right)$ denote player i's average payoff in the block game with discount factor $\delta \in [0,1]$ when a strategy profile $\sigma^{T_b} \in \Sigma^{T_b}$ is performed under perfect monitoring and when payoffs in the periods other than the main rounds are replaced with 0.[23]

Lemma 18 *There exist $\bar{\varepsilon} > 0$ and $\overline{K}$ such that for all $\varepsilon \leq \bar{\varepsilon}$ and $K \geq \overline{K}$, there exists $\overline{T}$ such that for all $T \geq \overline{T}$, there exists $\overline{\delta} < 1$ such that for all $\delta \in \left(\overline{\delta}, 1\right]$, $i \in I$, $x_{-i} \in X_{-i}$ with $x_{i-1} = B$, $\tilde{x}_{-i} \in X_{-i}$ with $\tilde{x}_{i-1} = G$, $\sigma_i^{T_b} \in \Sigma_i^{T_b}$ and $\tilde{\sigma}_i^{T_b} \in \mathcal{S}_i^{T_b}$,*

$$\max_{\sigma_i^{T_b} \in \Sigma_i^{T_b}} w_i^P(\sigma_i^{T_b}, \sigma_{-i}^{x_{-i}} \mid \delta) < \underline{w}_i < \overline{w}_i < \min_{\sigma_i^{T_b} \in \mathcal{S}_i^{T_b}} w_i^P(\sigma_i^{T_b}, \sigma_{-i}^{\tilde{x}_{-i}} \mid \delta).$$

Proof. Analogous to **Lemma 4** of Yamamoto (2009b). The difference is the existence of the monitoring rounds for the coordination round and kth monitoring round for $k \in \{1, \dots, K\}$. However, if ε is sufficiently small, the length of the monitoring rounds is negligible compared to that of the kth main round. ■

For notational convenience, define $\min_{\sigma_{-i}^{x_{-i}} | x_{i-1} = B} \max_{\sigma_i^{T_b} \in \Sigma_i^{T_b}} w_i^P(\sigma_i^{T_b}, \sigma_{-i}^{x_{-i}} \mid \delta) \equiv \underline{w}_i^P(\delta)$ and $\max_{\sigma_{-i}^{x_{-i}} | x_{i-1} = G} \min_{\sigma_i^{T_b} \in \mathcal{S}_i^{T_b}} w_i^P(\sigma_i^{T_b}, \sigma_{-i}^{x_{-i}} \mid \delta) \equiv \overline{w}_i^P(\delta)$. Note that so far, the definitions

[23] If $\delta = 1$, the average payoff is defined as the time-average payoff.

are the same as Yamamoto (2009b).

8.2.2 Determination of K and ε

Fix $\overline{u}$ such that for all i,

$$-\overline{u} < \underline{w}_i < \overline{w}_i < \overline{u},$$
$$\max_i \max_a |\pi_i(a)| < \overline{u},$$

and there exists $u_i : A_{i-1} \times \Omega_{i-1} \to \mathbb{R}$ satisfying

$$\pi_i(a) + E\left[u_i(a_{i-1}, \omega_{i-1}) \mid a\right] = 0 \text{ for all } a, \tag{13}$$

and

$$-\overline{u} < u_i(a_{i-1}, \omega_{i-1}) < \overline{u} \text{ for all } a_{i-1}, \omega_{i-1}. \tag{14}$$

Note that the identifiability condition guarantees the existence.

We will define K and ε so that the payoffs during the main round dominates the payoffs during the other rounds.

Lemma 19 *There exist K, $\varepsilon > 0$ and $\eta > 0$ such that there exists $\overline{T}$ such that for all $T \geq \overline{T}$, there exists $\overline{\delta} < 1$ such that for all $\delta \in \left(\overline{\delta}, 1\right]$ and i*

$$\left(1 - \frac{K^2 T}{T_b}\right) 3\overline{u} \quad < \quad 3\eta < \underline{w}_i - \underline{w}_i^P(\delta), \tag{15}$$
$$\left(1 - \frac{K^2 T}{T_b}\right) (|A_i| + 2)\,\overline{u} \quad < \quad (|A_i| + 2)\,\eta < \overline{w}_i^P(\delta) - \overline{w}_i. \tag{16}$$

Proof. Since $\lim_{K\to\infty,\varepsilon\to 0} \lim_{T\to\infty} \frac{K^2 T}{T_b} = 1$, the result holds from **Lemma 18**. ■

From now on, we fix K, $\varepsilon > 0$ and $\eta > 0$ such that the above lemma holds. Note that, so far, except for $\varepsilon > 0$, the definitions and explanations are the same as Yamamoto (2009b).

8.2.3 Proof of Lemma 13

Throughout this subsubsection, for each $k \in \{1, \ldots, K\}$, let $h_i^{[k]}$ denote player i's private history up to the end of the kth supplemental round. Also, let $h_i^{[k,m]}$ be player i's history up to the end of the kth main round, $h_i^{[0]}$ be player i's history up to the end of the confirmation round, and $h_i^{[-1]}$ be player i's history up to the end of the monitoring round for the coordination round. For each $k \in \{-1, \ldots, K\}$, let $H_i^{[k]}$ represent the set of all $h_i^{[k]}$ and $H_i^{[k,m]}$ represent the set of all $h_i^{[k,m]}$.

Given $h_{i-1}^{T_b} \in H_{i-1}^{T_b}$, let $I_{-i} \in (M_j^0 \times M_j^1 \times \cdots \times M_j^K)_{j \neq i}$ represent player $i-1$'s inference on the messages from players $-i$ in the report round. As Yamamoto (2009b) mentions, player i's actions in the report round cannot affect the realization of I_{-i} since player $i-1$ makes her inferences on player j's messages using the random event $\Phi_{i-1}(\{a_j^G\})$ and **Lemma 9** asserts that player i cannot manipulate the distribution of this random event. Let I_{-i}^k be the projection of I_{-i} onto $(M_j^0 \times M_j^1 \times \cdots \times M_j^k)_{j \neq i}$ for $k \in \{0, \ldots, K\}$.

Without loss of generality, consider a particular $i \in I$. Suppose that U_i^B is decomposable into real-valued functions $\left(\theta^{-1}, \theta^0, \ldots, \theta^{K+1}\right)$ such that

$$U_i^B(h_{i-1}^{T_b}) = \frac{1}{\delta^{T_b}} \left[\begin{array}{c} \delta^{T_{-1}} \theta^{-1}(h_{i-1}^{[-1]}) + \delta^{T_0} \theta^0(h_{i-1}^{[0]}) \\ + \sum_{k=1}^{K} \delta^{T_k} \theta^k(h_{i-1}^{[k]}, I_{-i}^k) + \delta^{T_b} \theta^{K+1}(h_{i-1}^{T_b}) \end{array} \right],$$

where T_{-1} is the last period of the monitoring round for the coordination round, T_0 is the last period of the confirmation round, and T_k is the last period of the kth supplemental round.

Intuitively, player i receives a transfer θ^{-1} after the coordination round, θ^0 after the confirmation round, θ^k after the kth supplemental round for each $k \in \{1, \ldots, K\}$, and θ^{K+1} after the report round.

In this transfer scheme, the transfers for the past rounds are irrelevant to player i's incentive compatibility. For example, consider the report round. Note that the transfer θ^{-1} is a function of $h_{i-1}^{[-1]}$, which does not depend on the history in the report round. Similarly, one can check that the transfers $\left(\theta^{-1}, \theta^0, \ldots, \theta^K\right)$ are irrelevant to player i's incentive compatibility in the report round. Likewise, the transfers $\left(\theta^{-1}, \theta^0, \ldots, \theta^k\right)$ are irrelevant to player

i's incentive compatibility in the continuation game from the kth main round.

In what follows, it is shown that there are transfers $\left(\theta^{-1}, \theta^{0}, \ldots, \theta^{K}\right)$ satisfying (7) through (9). To simplify the notation, let X^{B} denote the set of all $x \in X$ satisfying $x_{i-1} = B$. Likewise, let X_{-i}^{B} be the set of all $x_{-i} \in X_{-i}$ satisfying $x_{i-1} = B$.

Notice that so far, the definition of the functions and variables and explanations are the same as Yamamoto (2009b). To restore the conditional independence property, the functional form of $\left(\theta^{-1}, \theta^{0}, \ldots, \theta^{K+1}\right)$ is significantly different from Yamamoto (2009b).

Construction of θ^{K+1} Let

$$\theta^{K+1}(h_{i-1}^{T_b}) = \sum_{t=1}^{T_r} \frac{u_i\left(a_{i-1,t}, \omega_{i-1,t}\right)}{\delta^{T_r+1-t}},$$

where T_r is the length of the report round and $(a_{i-1,t}, \omega_{i-1,t})$ is player $(i-1)$'s action and signal in the tth period of the report round.

Intuitively, θ^{K+1} is a transfer for the report round. Our definition of θ^{K+1} is the same as Yamamoto (2009b). In addition, player i's messages in the report round do not affect $\theta^{-1}, \theta^{0}, \ldots, \theta^{K}$ as Yamamoto (2009b). Since player i's messages in the report round do not affect player i's utility, *regardless of the conditional independence property*, player i has the incentive to follow the equilibrium strategy. Therefore, the similar proof in A.1 in Yamamoto (2009b) can establish the following:

Lemma 20 *1. Any action is optimal conditional on the opponents' actions in each period of the report round, that is, (7) holds for each period in the report round.*

2. For all K and T, there exists $\overline{\delta} < 1$ such that for all $\delta \in \left(0, \overline{\delta}\right)$,

$$-T_r \overline{u} < \theta^{K+1}(h_{i-1}^{T_b}).$$

Construction of θ^{k} with $k \in \{1, \ldots, K\}$ In this construction, the following notation is useful. For each $x \in X$, let $H_{-i}^{[0]}(x)$ be the set of $h_{-i}^{[0]} \in H_{-i}^{[0]}$ such that for each $j \neq i$, regardless of player i's messages in the monitoring rounds for the coordination round and regardless of player i's messages in the confirmation round,

$$m_j^0 \in M_j^0 (x),$$

that is, players $-i$ confirm x and play $a_{-i}^{x,1}$ in the first main round regardless of player i's actions in the monitoring rounds for the coordination round and the confirmation round. Define

$$\bar{H}_{-i}^{[0]} \equiv \bigcup_{x \in X^B} H_{-i}^{[0]} (x).$$

Notice that if all the players $-i$ play the block strategy profile $\sigma_{-i}^{x_{-i}}$ up to the end of the confirmation round, then it is likely that the resulting history profile $h_{-i}^{[0]}$ is an element of $H_{-i}^{[0]} (\tilde{x})$ with $\tilde{x}_{-i} = x_{-i}$. Note that, if player i deviates from $\sigma_i^{x_i}$, $\tilde{x}_i$ can be different from x_i with high probability. However, as we have seen in Section 6.4, it is likely that either G_i or B_i is confirmed and that player i's messages in the ith monitoring round for the coordination round and the confirmation round do not have an impact on the confirmation of $\tilde{x}_i$.

Likewise, for each $k \in \{1, \dots, K-1\}$ and $x \in X$, let $H_{-i}^{[k]} (x)$ be the set of $h_{-i}^{[k]} \in H_{-i}^{[k]}$ such that for each $j \neq i$, regardless of player i's messages in the kth monitoring and supplemental rounds, no one's deviation is inferred and players $-i$ play $a_{-i}^{x,k+1}$ in the $(k+1)$ main round. Define

$$\bar{H}_{-i}^{[k]} \equiv \bigcup_{x \in X^B} H_{-i}^{[k]} (x).$$

Note that, as before, if all the players perform the block strategy profile σ^x up to the kth main round and if the other players than player i perform σ^x up to the kth supplemental round, it is likely that the resulting history profile $h_{-i}^{[k]}$ is an element of $H_{-i}^{[k]} (x)$ as explained in Section 6.4.

Also, for each $k \in \{1, \dots, K\}$ and $x \in X$, let $H_{-i}^{[k]} (x, i)$ be the set of $h_{-i}^{[k]} \in H_{-i}^{[k]}$ such that there exists $\tilde{k} \in \{1, \dots, k\}$ such that for each $j \neq i$, regardless of player i's messages in the $\tilde{k}$th monitoring and supplemental rounds, player i's unilateral deviation is firstly confirmed in the $\tilde{k}$th supplemental round. Note that $h_{-i}^{[k]}$ is likely to be an element of $H_{-i}^{[k]} (x, i)$ if the other players than player i play the block strategy profile $\sigma_{-i}^{x_{-i}}$ but player i deviates from $a^{x,\tilde{k}}$ in the $\tilde{k}$th main round as we have seen in Section 6.4.

Define

$$\bar{H}_{-i}^{[k]} \equiv \bigcup_{x \in X^B} \left(H_{-i}^{[k]}(x) \cup H_{-i}^{[k]}(x, i) \right).$$

We specify the transfers $\left(\theta^{-1}, \theta^0, \ldots, \theta^{K+1}\right)$ are specified by backward induction. To define θ^k, assume that $\left(\theta^{k+1}, \ldots, \theta^{K+1}\right)$ have already been determined so that player i's continuation payoff after history $h_{-i}^{[k]} \in H_{-i}^{[k]}$, augmented by $\left(\theta^{k+1}, \ldots, \theta^{K+1}\right)$, is equal to $V_i(h_{-i}^{[k]})$ and that (7) holds for the initial period of $k+1, \ldots, K$th main round. Here, for each $k \in \{1, \ldots, K\}$ and $h_{-i}^{[k]} \in H_{-i}^{[k]}$, the value $V_i(h_{-i}^{[k]})$ denotes the maximum of player i's actual continuation payoff (i.e., the discounted sum of stage game payoffs) after history $h_{-i}^{[k]}$ over all her continuation strategies, subject to the constraints that monitoring is perfect and that payoffs in the communication stages are replaced with zero. For each $h_{-i}^{[0]} \in \bar{H}_{-i}^{[0]}$, the value $V_i(h_{-i}^{[0]})$ denotes the maximum of player i's actual continuation payoff after history $\tilde{h}_{-i}^{[0]}$ over all $\tilde{h}_{-i}^{[0]} \in \bar{H}_{-i}^{[0]}$ and over all her continuation strategies, subject to the constraints that monitoring is perfect and that payoffs in the communication stages are replaced with zero. For each $k \in \{0, \ldots, K\}$ and $h_{-i}^{[k]} \notin \bar{H}_{-i}^{[k]}$, the value $V_i(h_{-i}^{[k]})$ denotes player i's actual continuation payoff when she earns $\max_{a \in A} \pi_i(a)$ in periods of the main rounds and zero in other periods. Notice that the transfers $\left(\theta^1, \ldots, \theta^K\right)$ are specified in such a way that player i's continuation payoff $V_i(h_{-i}^{[k]})$ from the kth main round is high and is the same for all histories $h_{-i}^{[k]} \notin \bar{H}_{-i}^{[k]}$. This "constant continuation payoff" property is used to show that player i has a truth-telling incentive in the $k-1$th monitoring and supplemental rounds, even when $h_{-i}^{[k]} \notin \bar{H}_{-i}^{[k]}$ so that player i's messages in those rounds can affect the opponents' continuation play.

Notice that so far, the explanation is the same as Yamamoto (2009b). However, as explained in Section 6.4, since we insert the monitoring rounds to recover the conditional independence property, the definition of $H_{-i}^{[k]}(x)$ and $H_{-i}^{[k]}(x, i)$ are more complicated. However, note that the key property for $H_{-i}^{[k]}(x)$ and $H_{-i}^{[k]}(x, i)$ still holds. That is, whether $h_{-i}^{[k]} \in H_{-i}^{[k]}(x)$, $h_{-i}^{[k]} \in H_{-i}^{[k]}(x, i)$ or $h_{-i}^{[k]} \in \left(\bar{H}_{-i}^{[k]}\right)^c$ is fully determined by $\left(x^0(j, i, -(j, i, l))\right)_{j,l \neq i}$,

$\left(x^k\left(j,i,-\left(j,i,l\right)\right)\right)_{j,l\neq i}$ and $\left(x^k\left(j,\left(\iota,\iota'\right),-\left(j,\iota,\iota'\right)\right)\right)_{j\neq i,i\in(\iota,\iota')}$. Therefore, the conditional independence property holds for player i with respect to whether $h_{-i}^{[k]}\in H_{-i}^{[k]}(x)$, $h_{-i}^{[k]}\in H_{-i}^{[k]}(x,i)$ or $h_{-i}^{[k]}\in\left(\bar{H}_{-i}^{[k]}\right)^c$.

Now, let us specify $\left(\theta^1,\ldots,\theta^K\right)$ by backward induction. Given that $\left(\theta^{k+1},\ldots,\theta^{K+1}\right)$ have already been determined so that player i's continuation payoff after history $h_{-i}^{[k]}\in H_{-i}^{[k]}$, augmented by $\left(\theta^{k+1},\ldots,\theta^{K+1}\right)$, is equal to $V_i(h_{-i}^{[k]})$ and that (7) holds for the initial period of $k+1,\ldots,K$th main round.

For $K+1$, we are done by **Lemma 20**. Suppose we have constructed $\theta^{\tilde{k}}$ with $\tilde{k}\geq k+1$. Consider k. Suppose that θ^k is decomposable as

$$\theta^k(h_{i-1}^{[k]},I_{-i}^k)=\tilde{\theta}^k(h_{i-1}^{[k,m]},\hat{\Omega}_{-(i-1,i,l)},I_{-i}^{k-1})+\sum_{t=1}^{\tilde{T}_k-KT}\frac{u_i\left(a_{i-1,t},\omega_{i-1,t}\right)}{\delta^{\tilde{T}_k-KT+1-t}},\tag{17}$$

where $\tilde{T}_k$ is the total length of the kth main, monitoring and supplemental rounds and $\hat{\Omega}_{-(i-1,i,l)}$ denotes the realized messages from players $-(i-1,i,l)$ to player $i-1$ during the kth monitoring round with some fixed $l\neq i-1,i$. Here, $(a_{i-1,t},\omega_{i-1,t})$ is player $(i-1)$'s action and signal in the tth period of the kth monitoring and supplemental rounds.

Intuitively, θ^k is a transfer for the kth main, monitoring and supplemental rounds. The explanation of θ^k is the same as Yamamoto (2009b) except that we have $\hat{\Omega}_{-(i-1,i,l)}$ to restore the conditional independence property and that I_{-i}^{k-1} is more complicated due to the existence of the monitoring rounds. Since I_{-i}^{k-1} and $\hat{\Omega}_{-(i-1,i,l)}$ are independent of player i's action in the kth monitoring and supplemental rounds, similarly to **Lemma 9** of Yamamoto (2009b), we have the following:

Lemma 21 *Any action is optimal conditional on the opponents' action in each period of the kth monitoring and supplemental rounds, that is, (7) holds for each period of the kth monitoring and supplemental rounds.*

Next, we specify a real valued function $\tilde{\theta}^k$. The following notation is useful. For each $h_{-i}^{[k-1]}\in H_{-i}^{[k-1]}$ and $a_i\in A_i$, let $\tilde{W}_i(h_{-i}^{[k-1]},a_i)$ denote player i's continuation payoff from the

kth main round, augmented by $\left(\theta^{k+1}, \ldots, \theta^{K+1}\right)$ and by the second term of (17), when player i plays a_i constantly in the kth main round and plays a best reply thereafter. That is,

$$\tilde{W}_i(h_{-i}^{[k-1]}, a_i) \equiv \sum_{t=1}^{KT} \delta^{t-1} \pi_i(a_i, \sigma_{-i}^{x_{-i}}(h_{-i}^{[k-1]})) + \delta^{\tilde{T}_k} \sum_{h_{-i}^{[k]} \in H_{-i}^{[k]}} \Pr(h_{-i}^{[k]} | h_{-i}^{[k-1]}, a_i) V_i(h_{-i}^{[k]}),$$

where $\Pr(h_{-i}^{[k]} | h_{-i}^{[k-1]}, a_i)$ is the probability that $h_{-i}^{[k]}$ realizes when player i plays a_i constantly in the kth main round and sends $x(i, (\iota, \iota'), -(i, \iota, \iota')) = 0$ for all (ι, ι') in the kth supplemental round while the opponents' play $\sigma_{-i}^{x_{-i}} \mid h_{-i}^{[k-1]}$. Note that the first term is the stage game payoff in the kth main round and that the second is the continuation payoff after the kth monitoring round augmented by $\left(\theta^{k+1}, \ldots, \theta^{K+1}\right)$. The stage game payoffs in the kth monitoring and supplemental rounds do not appear here, as the second term of (17) offsets them.

For each $j \neq i$ and $h_j^{[k-1]} \in H_j^{[k-1]}$, let J_j^{k-1} denote the inference profiles calculated from the history in the confirmation round and the past monitoring and supplemental rounds, that is, $J_{-i}^k = (m_j^0, m_j^1, \ldots, m_j^k)_{j \neq i}$. Note that $\tilde{W}_i(h_{-i}^{[k-1]}, a_i)$ only depend on (J_{-i}^{k-1}, a_i). Therefore, we can write $\tilde{W}_i(J_{-i}^{[k-1]}, a_i)$ instead of $\tilde{W}_i(h_{-i}^{[k-1]}, a_i)$. Recall that everyone tells the truth in the report round from **Lemma 20** and that I_{-i}^k denotes the realized messages corresponding to J_{-i}^k in the report round. Hence, for any history up to the end of the Kth supplemental round, $I_{-i}^K = J_{-i}^K$ with high probability.

For each J_{-i}^{k-1}, let $\{a_i^1\left(J_{-i}^{k-1}\right), \ldots, a_i^{|A_i|}\left(J_{-i}^{k-1}\right)\}$ be a sequence of all the elements of A_i such that

$$\lim_{T \to \infty} \lim_{\delta \to 1} \frac{\tilde{W}_i\left(J_{-i}^{k-1}, a_i^1\left(J_{-i}^{k-1}\right)\right)}{T} \geq \cdots \geq \lim_{T \to \infty} \lim_{\delta \to 1} \frac{\tilde{W}_i(J_{-i}^{k-1}, a_i^{|A_i|}\left(J_{-i}^{k-1}\right))}{T}.$$

For each J_{-i}^{k-1} and $n \in \{1, \ldots, |A_i|\}$, let $1_{\left[J_{-i}^{k-1}, n\right]} : H_{i-1}^{[k,m]} \times \hat{\Omega}_{-(i-1,i,l)} \to \{0, 1\}$ is the indicator function such that $1_{\left[J_{-i}^{k-1}, n\right]}(h_{i-1}^{[k,m]}, \hat{\Omega}_{-(i-1,i,l)}) = 1$ if the random event $\Psi_{i-1,i,-(i-1,i,l)}(\{a_i^n(J_{-i}^{k-1}), \ldots, a_i^{|A_i|}(J_{-i}^{k-1})\})$ is counted more than Z_{KT} times in the kth main round when player $i-1$ monitors player i based on the messages from players $-(i-1, i, l)$ and $1_{\left[J_{-i}^{k-1}, n\right]}(h_{i-1}^{[k,m]}, \hat{\Omega}_{-(i-1,i,l)}) = 0$ otherwise. Likewise, for each $a_i \in A_i$, let $1_{a_i} : H_{i-1}^{[k,m]} \times \hat{\Omega}_{-(i-1,i,l)} \to \{0, 1\}$ is the indicator function such that $1_{a_i}(h_{i-1}^{[k,m]}, \hat{\Omega}_{-(i-1,i,l)}) = 1$ if the random event $\Psi_{i-1,i,-(i-1,i,l)}(\{a_i\})$ occurs more than Z_{KT} times in the kth main round when player $i-1$ monitors player i based on the messages from players $-(i-1,i,l)$ and $1_{a_i}(h_{i-1}^{[k,m]}, \hat{\Omega}_{-(i-1,i,l)}) = 0$ otherwise.

Then define $\tilde{\theta}^k$ to be

$$\begin{aligned} & \tilde{\theta}^k(h_{i-1}^{[k,m]}, I_{-i}^{k-1}, \hat{\Omega}_{-(i-1,i,l)}) \\ = & \sum_{a_i \in A_i} 1_{a_i}(h_{i-1}^{[k,m]}, \hat{\Omega}_{-(i-1,i,l)}) KT\eta + \sum_{n=1}^{|A_i|} 1_{\left[I_{-i}^{k-1}, n\right]}(h_{i-1}^{[k,m]}, \hat{\Omega}_{-(i-1,i,l)}) \lambda^k \left(I_{-i}^{k-1}, n\right), \end{aligned}$$

where $\lambda^k \left(I_{-i}^{k-1}, n\right)$ solves

$$\begin{aligned} V_i\left(J_{-i}^{k-1}\right) &= \tilde{W}_i\left(J_{-i}^{k-1}, a_i\right) + \delta^{\tilde{T}_k} \sum_{\tilde{a}_i \in A_i} \Pr\left(1_{\tilde{a}_i} | J_{-i}^{k-1}, a_i\right) KT\eta \\ &+ \delta^{\tilde{T}_k} \sum_{I_{-i}^{k-1}} \Pr\left(I_{-i}^{k-1} | J_{-i}^{k-1}\right) \sum_{n=1}^{|A_i|} \Pr(1_{\left[I_{-i}^{k-1}, n\right]} | J_{-i}^{k-1}, a_i) \lambda^k \left(I_{-i}^{k-1}, n\right) \end{aligned} \tag{18}$$

for all J_{-i}^{k-1} and $a_i \in A_i$. Since $V_i(h_i^{[k-1]})$ only depends on J_{-i}^{k-1}, with abuse of notation, we write $V_i\left(J_{-i}^{k-1}\right)$ instead of $V_i(h_{-i}^{[k-1]})$. Here, $\Pr\left(1_{[\cdot]} \mid J_{-i}^{k-1}, a_i\right)$ denotes the probability that the indicator function takes 1 conditional on that player i chooses the constant action a_i while players $-i$ play the action $\sigma_{-i}^{x_{-i}}\left(J_{-i}^{k-1}\right)$ constantly.[24] Note that $\sigma_{-i}^{x_{-i}}$ only depends on J_{-i}^{k-1}. In addition, $\Pr(I_{-i}^{k-1} | J_{-i}^{k-1})$ denotes the probability that I_{-i}^{k-1} realizes given that the history at the beginning of the kth main round is J_{-i}^{k-1}. In words, the values $\left(\lambda^k\left(I_{-i}^{k-1}, n\right)\right)_{I_{-i}^{k-1}, n}$ are determined so that player i's unnormalized continuation payoff after J_{-i}^{k-1}, augmented by $(\theta^k, \ldots, \theta^{K+1})$ equals $V_i(J_{-i}^{k-1})$, no matter what constant action player i chooses in the kth main round.

So far, the definition and the explanation of $\tilde{W}_i(J_{-i}^{k-1}, a_i)$ are the same as Yamamoto (2009b) except that we insert $\hat{\Omega}_{-(i-1,i,l)}$ to recover the conditional independence property, that we use $\Psi_{i-1,i,-(i-1,i,l)}$ to restore the conditional independence property instead of ψ_{i-1},[25]

[24] More precisely, $\sigma_{-i}^{x_{-i}}(h_{-i}^{[k-1]})$ with $h_{-i}^{[k-1]}$ corresponding to J_{-i}^{k-1}.
[25] For the definition of ψ_{i-1} in Yamamoto (2009b), see the original paper.

and that the definition of J_{-i}^{k-1} implicitly includes the result of the message exchanges in the monitoring rounds. Then, the conditional independence property is restored. In particular, from **Lemma 12**,

$$\lim_{T\to\infty}\max_{\{\omega_{i,\tau}\}_{\tau=1}^{KT}}\left|\begin{array}{c}\sum_{J_{-i}^{[k]}\in J_{-i}^{[k]}}\Pr(J_{-i}^{k}|J_{-i}^{k-1},\{a_{i,\tau}\}_{\tau=1}^{KT})V_i(J_{-i}^{k})\\ -\sum_{J_{-i}^{[k]}\in J_{-i}^{[k]}}\Pr(J_{-i}^{k}|J_{-i}^{k-1},\{a_{i,\tau}\}_{\tau=1}^{KT},\{\omega_{i,\tau}\}_{\tau=1}^{KT})V_i(J_{-i}^{k})\end{array}\right|=0,$$

where $\omega_{i,\tau}$ is the signal observation in the τth period of the kth main round. Similarly,

$$\lim_{T\to\infty}\max_{\{\omega_{i,\tau}\}_{\tau=1}^{KT}}\left|\begin{array}{c}\sum_{\tilde{a}_i\in A_i}\Pr\left(1_{\tilde{a}_i}|J_{-i}^{k-1},\{a_{i,\tau}\}_{\tau=1}^{KT}\right)KT\eta\\ -\sum_{\tilde{a}_i\in A_i}\Pr\left(1_{\tilde{a}_i}|J_{-i}^{k-1},\{a_{i,\tau}\}_{\tau=1}^{KT},\{\omega_{i,\tau}\}_{\tau=1}^{KT}\right)KT\eta\end{array}\right|=0$$

and

$$\lim_{T\to\infty}\max_{\{\omega_{i,\tau}\}_{\tau=1}^{KT}}\left|\begin{array}{c}\sum_{I_{-i}^{k-1}}\Pr\left(I_{-i}^{k-1}|J_{-i}^{k-1}\right)\sum_{n=1}^{|A_i|}\Pr(1_{\left[I_{-i}^{k-1},n\right]}|J_{-i}^{k-1},\{a_{i,\tau}\}_{\tau=1}^{KT})\lambda^k\left(I_{-i}^{k-1},n\right)\\ -\sum_{I_{-i}^{k-1}}\Pr\left(I_{-i}^{k-1}|J_{-i}^{k-1}\right)\sum_{n=1}^{|A_i|}\Pr(1_{\left[I_{-i}^{k-1},n\right]}|J_{-i}^{k-1},\{a_{i,\tau}\}_{\tau=1}^{KT},\{\omega_{i,\tau}\}_{\tau=1}^{KT})\lambda^k\left(I_{-i}^{k-1},n\right)\end{array}\right|=0.$$

Hence, the "almost conditional independence" property in Yamamoto (2009b) holds. Therefore, similarity to **Lemma 10** in Yamamoto (2009b), we have the following.

Lemma 22 *For all K, there exists $\overline{T}$ such that for all $T>\overline{T}$, there exists $\bar{\delta}\in(0,1)$ such that for all $\delta\in[\bar{\delta},1)$, (18) has a unique solution and for all $(h_{i-1}^{[k]},I_{-i}^{k-1})$,*

$$-\left(\tilde{T}_k-KT\right)\overline{u}-2KT\eta<\theta^k(h_{i-1}^{[k]},I_{-i}^{k-1}).$$

Also, using this transfer scheme, (7) holds for the initial period of the kth main round with $x\in X^B$ and player i's continuation payoff after history $h_{i-1}^{[k-1]}$ is equal to $V_i(h_{-i}^{[k-1]})$.

Construction of θ^0 Let

$$\theta^0(h_{i-1}^{[0]})=\sum_{t=1}^{\tilde{T}_0}\frac{u_i\left(a_{i-1,t},\omega_{i-1,t}\right)}{\delta^{\tilde{T}_0+1-t}},$$

where $\tilde{T}_0$ is the length of the confirmation round and $(a_{i-1,t}, \omega_{i-1,t})$ is player $(i-1)$'s action and signal in the tth period of the confirmation round.

Intuitively, θ^0 is a transfer for the confirmation round. Our definition of θ^0 is the same as Yamamoto (2009b). In addition, player i's messages in the confirmation round do not affect her continuation payoff as Yamamoto (2009b). Therefore, regardless of the conditional independence, the similar proof for A.3 in Yamamoto (2009b) can establish the following:

Lemma 23 *1. Any action is optimal conditional on the opponents' action in each period of the confirmation round, that is, (7) holds for the confirmation round.*

2. For all T, there exists $\overline{\delta} < 1$ such that for all $\delta \in \left(0, \overline{\delta}\right)$,

$$-\tilde{T}_0 \overline{u} < \theta^0(h_{i-1}^{[0]}).$$

Construction of θ^{-1} Suppose that θ^{-1} is decomposable as

$$\theta^{-1}(h_{i-1}^{[-1]}) = 3T_b\eta + \tilde{\theta}^{-1}(h_{i-1}^{[-1]}) + \sum_{t=1}^{\tilde{T}_{-1}} \frac{u_i\left(a_{i-1,t}, \omega_{i-1,t}\right)}{\delta^{\tilde{T}_{-1}+1-t}}, \tag{19}$$

where $\tilde{T}_{-1}$ is the total length of the coordination round and monitoring rounds for the coordination round and $(a_{i-1,t}, \omega_{i-1,t})$ is player $(i-1)$'s action and signal in the tth period of the block game.

For each $x \in X$, let $H_{i-1}^{[-1]}(x)$ denote the set of all $h_{i-1}^{[-1]} \in H_{i-1}^{[-1]}$ such that for $h_{i-1}^{[-1]}$,

- $x\left(i-1, j, -\left(i-1, j, i\right)\right) = x_j$ for all $j \neq i, i-1$,
- player $i-1$ takes $a_{i-1}^{x_{i-1}}$ constantly for the coordination round, and
- $\Gamma_{i-1,i,-(i-1,i,l)}\left(\{a_i^{x_i}\}\right)$ occurs more than $\frac{1}{\varepsilon} Z_T$ times when player $i-1$ monitors player i's message in the coordination round based on the messages from players $-\left(i-1, i, l\right)$ with arbitrarily fixed l.[26]

[26] In the main round, only ε-fraction of all the periods are used for the monitoring to keep the efficiency. On the other hand, in the coordination round, each period is used. Since Z_T is originaly defined for the main round, we need to adjust the probability by multiplying $\frac{1}{\varepsilon}$ here for the coordination round.

Then, for each $x \in X$, let $1_x : H_{i-1}^{[-1]} \to \{0,1\}$ denote the indicator function of $H_{i-1}^{[-1]}(x)$, that is, $1_x(h_{i-1}^{[-1]}) = 1$ if and only if $h_{i-1}^{[-1]} \in H_{i-1}^{[-1]}(x)$.

Let

$$\tilde{\theta}^{-1}(h_{i-1}^{[-1]}) = \sum_{x \in X^B} 1_x(h_{i-1}^{[-1]}) \lambda^{-1}(x),$$

where $\left(\lambda^{-1}(x)\right)_{x \in X_B}$ solve

$$\begin{aligned} \sum_{t=1}^{T_b} \delta^{t-1} \underline{w}_i &= \delta^{\tilde{T}_{-1}} \left[3T_b\eta + \delta^{\tilde{T}_{-1}} \sum_{\tilde{x} \in X^B} \sum_{h_{i-1}^{[-1]} \in H_{i-1}^{[-1]}(x)} \Pr(h_{i-1}^{[-1]}(x) | \sigma^x) \lambda^{-1}(\tilde{x}) \right] \\ &\quad + \delta^{\tilde{T}_{-1} + \tilde{T}_0} \sum_{h_{-i}^{[0]} \in H_{-i}^{[0]}} \Pr(h_{-i}^{[0]} | \sigma^x) V_i(h_{-i}^{[0]}) \end{aligned} \tag{20}$$

for all $x_{-i} \in X_{-i}^B$. Here, $\Pr(h_{i-1}^{[-1]}(x) | \sigma^x)$ denotes the probability that $h_{i-1}^{[-1]}(x)$ realizes when players perform the block strategy profile σ^x and $\Pr(h_{-i}^{[0]} | \sigma^x)$ denotes the probability of $h_{-i}^{[0]}(x)$. Intuitively, the values $\left(\lambda^{-1}(x)\right)_{x \in X_B}$ are determined so that (8) holds. Indeed, the right-hand side of (20) denotes player i's auxiliary scenario payoff from the block strategy profile σ^x. Note that so far, the definitions and the explanations of θ^{-1} is the same as Yamamoto (2009b) except that we insert the monitoring rounds for the coordination round to restore the conditional independence property, that the confirmation of $(x_j)_j$ is more complicated due to the existence of the monitoring rounds, and that we use $\Gamma_{i-1,i,-(i,j,l)}$ instead of ψ_{i-1}.[27] As we will see, since we restore the conditional independence property by inserting the monitoring round, it is possible to apply Yamamoto (2009b)'s proof.

We can interpret θ^{-1} as a transfer for the coordination round and monitoring rounds for the coordination round. Since player i's messages except for the coordination round do not affect player i's continuation payoff, any action is optimal for player i conditional on the opponents' action for each period of each monitoring round for the coordination round.

Let us consider the coordination round. Without loss of generality, consider a particular $x_{-i} \in X_{-i}^B$. Let $W_i(\sigma_{-i}^{x_{-i}}, \{a_{i,\tau}\}_{\tau=1}^T)$ denote player i's unnormalized payoff in the auxiliary

[27]For the definition of ψ_{i-1} in Yamamoto (2009b), see the original paper.

scenario against $\sigma_{-i}^{x_{-i}}$ when player i plays $a_{i,\tau}$ in the τth period of the coordination round. Specifically,

$$
\begin{aligned}
& W_i(\sigma_{-i}^{x_{-i}}, \{a_{i,\tau}\}_{\tau=1}^{T}) \\
= \; & \delta^{\tilde{T}_{-1}+\tilde{T}_0} \sum_{h_{-i}^{[0]} \in H_{-i}^{[0]}} \Pr(h_{-i}^{[0]} | \sigma_{-i}^{x_{-i}}, \{a_{i,\tau}\}_{\tau=1}^{T}) V_i(h_{-i}^{[0]}) + \delta^{\tilde{T}_{-1}+\tilde{T}_0} \sum_{\tilde{x} \in X^B} \Pr(1_{\tilde{x}} | \sigma^{x}, \{a_{i,\tau}\}_{\tau=1}^{T}) \lambda^{-1}(\tilde{x}) \\
= \; & \delta^{\tilde{T}_{-1}+\tilde{T}_0} \sum_{J_{-i}^{0}} \Pr(J_{-i}^{0} | \sigma_{-i}^{x_{-i}}, \{a_{i,\tau}\}_{\tau=1}^{T}) V_i(J_{-i}^{0}) + \delta^{\tilde{T}_{-1}+\tilde{T}_0} \sum_{\tilde{x} \in X^B} \Pr(1_{\tilde{x}} | \sigma^{x}, \{a_{i,\tau}\}_{\tau=1}^{T}) \lambda^{-1}(\tilde{x})
\end{aligned}
$$

where $\Pr(h_{-i}^{[0]} | \sigma_{-i}^{x_{-i}}, \{a_{i,\tau}\}_{\tau=1}^{T})$, $\Pr(1_{\tilde{x}} | \sigma_{-i}^{x_{-i}}, \{a_{i,\tau}\}_{\tau=1}^{T})$ and $\Pr(J_{-i}^{0} | \sigma_{-i}^{x_{-i}}, \{a_{i,\tau}\}_{\tau=1}^{T})$ are probabilities that $h_{-i}^{[0]}$, $1_{\tilde{x}}(h_{i-1}^{[-1]}) = 1$ and J_{-i}^{0} respectively occur conditional on $\sigma_{-i}^{x_{-i}}, \{a_{i,\tau}\}_{\tau=1}^{T}$. The second line follows since $V_i(h_{-i}^{[0]})$ only depends on J_{-i}^{0}.

Note that $W_i(\sigma_{-i}^{x_{-i}}, \{a_{i,\tau}\}_{\tau=1}^{T})$ is similarly determined as Yamamoto (2009b). In particular, from **Lemma 10,** for any signal observation $\{\omega_{i,\tau}\}_{\tau=1}^{T}$ in the τth round of the coordination round,

$$
\Pr(J_{-i}^{0} | \sigma_{-i}^{x_{-i}}, \{a_{i,\tau}\}_{\tau=1}^{T}) = \Pr(J_{-i}^{0} | \sigma_{-i}^{x_{-i}}, \{a_{i,\tau}\}_{\tau=1}^{T}, \{\omega_{i,\tau}\}_{\tau=1}^{T})
$$

and

$$
\Pr(1_{\tilde{x}} | \sigma^{x}, \{a_{i,\tau}\}_{\tau=1}^{T}) = \Pr(1_{\tilde{x}} | \sigma^{x}, \{a_{i,\tau}\}_{\tau=1}^{T}, \{\omega_{i,\tau}\}_{\tau=1}^{T}).
$$

Hence, the conditional independence property holds. Therefore, similarity to **Lemma 11** in Yamamoto (2009b), we have the following:

Lemma 24 *For all K, there exists $\overline{T}$ such that for all $T > \overline{T}$, there exists $\bar{\delta} \in (0,1)$ such that for all $\delta \in [\bar{\delta}, 1)$, (20) has a unique solution and for all $h_{i-1}^{T_b}$,*

$$
3T_b\eta - \tilde{T}_{-1}\overline{u} < \theta^{-1}(h_{i-1}^{T_b}).
$$

Also, using this transfer scheme, (7) holds for the initial period of coordination round with $x \in X^B$ and for any period of the each monitoring round for the coordination round.

Determination of $\overline{\delta}$ From the above argument, there exists $\overline{T}$ such that for all $T > \overline{T}$, there exists $\overline{\delta} < 1$ such that for all $\delta \in \left(\overline{\delta}, 1\right)$ and for all $i \in I$, $l \geq 0$, $h^{t_l} \in H^{t_l}$, $h_{i-1}^{T_b} \in H_{i-1}^{T_b}$ and $x \in X$ with $x_{i-1} = B$, (7) and (8) hold. In addition, from (15),

$$U_i(h_{i-1}^{T_b}) \geq 0 \text{ for all } h_{i-1}^{T_b}.$$

Fix some $T > \overline{T}$. Then, since $U_i(h_{i-1}^{T_b}) \leq 2T_b\overline{u}$ for all $h_{i-1}^{T_b}$, for sufficiently large δ,

$$U_i(h_{i-1}^{T_b}) < \frac{\overline{w}_i - \underline{w}_i}{1-\delta} \text{ for all } h_{i-1}^{T_b} \in H_{i-1}^{T_b},$$

which implies (9).

8.2.4 Reward Functions for the Good Strategy

Let us prove **Lemma 14**. As we will see, the basic structure of the proof is similar to the above. Hence, except that we insert the monitoring rounds to restore the conditional independence property and that we use $\hat{\Omega}_{-(i-1,i,l)}$ and $\Psi_{i-1,i,-(i-1,i,l)}$ for θ^k and $x\left(i-1, j, -\left(i-1, j, l\right)\right)$ and $\Gamma_{i-1,i,-(i-1,i,l)}$ for θ^{-1},[28] the definitions and explanations are the same as Yamamoto (2009b).

To simplify the notation, let X^G be the set of all $x \in X$ satisfying $x_{i-1} = G$ and X_{-i}^G be the set of all $x_{-i} \in X_{-i}$ satisfying $x_{i-1} = G$. Let $\bar{H}_{-i}^{[0]} \equiv \bigcup_{x \in X^G} H_{-i}^{[0]}(x)$ and $\bar{H}_{-i}^{[k]} \equiv \bigcup_{x \in X^G} \left(H_{-i}^{[k]}(x) \cup H_{-i}^{[k]}(x, i)\right)$ for $k \in \{1, \ldots, K\}$. See Section 8.2.3 for the definition of $H_{-i}^{[0]}(x)$ and $H_{-i}^{[k]}(x) \cup H_{-i}^{[k]}(x, i)$ for $k \in \{1, \ldots, K\}$.

Without loss of generality, consider a particular $i \in I$. Suppose that U_i^G is decomposable into real-valued functions $\left(\theta^{-1}, \ldots, \theta^{K+1}\right)$ as in Section 8.2.3. We specify $\left(\theta^{-1}, \ldots, \theta^{K+1}\right)$ so that **Lemma 14** holds.

Let θ^0 and θ^{K+1} be as in the proof of **Lemma 13**, that is, these transfers are the discounted sums of u_i. Then, (10) holds for the confirmation round and the report round.

[28] As we will see, U_i^G is decomposable into real-valued functions $\left(\theta^{-1}, \ldots, \theta^{K+1}\right)$ as in Section 8.2.3.

The sequence of transfers $\left(\theta^1, \ldots, \theta^K\right)$ is specified by backward induction. Note that θ^{K+1} is determined so that **Lemma 14** holds. To define θ^k, assume that the sequence $\left(\theta^{k+1}, \ldots, \theta^{K+1}\right)$ is determined so that player i's continuation payoff after history $h_i^{[k]}$, augmented by $\left(\theta^{k+1}, \ldots, \theta^{K+1}\right)$, is equal to $V_i(h_{-i}^{[k]})$. Here, for each $k \in \{1, \ldots, K\}$ and $h_{-i}^{[k]} \in \bar{H}_{-i}^{[k]}$, the value $V_i(h_{-i}^{[k]})$ is defined to be the minimum of player i's continuation payoff after history $h_{-i}^{[k]}$ over all her continuation strategies in the set $\{\sigma_i^{T_b} \mid h_i^{[k]} : \forall h_i^{[k]} \in H_i^{[k]}, \forall \sigma_i^{T_b} \in \mathcal{S}_i^{T_b}\}$, subject to the constraints that monitoring is perfect and that payoffs in the communication stages are replaced with 0. For each $h_{-i}^{[0]} \in \bar{H}_{-i}^{[0]}$, the value $V_i(h_{-i}^{[0]})$ is defined to be the minimum of player i's continuation payoff after history $\tilde{h}_{-i}^{[0]}$ over all $\tilde{h}_{-i}^{[0]} \in \bar{H}_{-i}^{[0]}$ and over all her continuation strategies in the set $\{\sigma_i^{T_b} \mid h_i^{[0]} : \forall h_i^{[0]} \in H_i^{[0]}, \forall \sigma_i^{T_b} \in \mathcal{S}_i^{T_b}\}$, subject to the constraints that monitoring is perfect and that payoffs in the communication stages are replaced with 0. For each $k \in \{0, \ldots, K\}$ and $h_{-i}^{[k]} \notin \bar{H}_{-i}^{[k]}$, the value $V_i(h_{-i}^{[k]})$ is defined to be player i's continuation payoff when she earns $-\overline{u}$ in periods of the main rounds and zero in the other periods.

Suppose that θ^k is decomposed as (17). Then, from the analogous argument to **Lemma 21**, (10) holds for each period of the kth monitoring and supplemental rounds.

To specify $\tilde{\theta}^k$, the following notion is useful. For each $h_{-i}^{[k-1]} \in H_{-i}^{[k-1]}$ and $a_i \in A_i$, let $\tilde{W}_i(h_{-i}^{[k-1]}, a_i)$ denote player i's continuation payoff from the kth main round, augmented by $\left(\theta^{k+1}, \ldots, \theta^{K+1}\right)$ and by the second term of (17) when player i plays a_i constantly in the kth main round and plays a best reply thereafter. As before, we can write $\tilde{W}_i\left(J_{-i}^{k-1}, a_i\right)$ instead of $\tilde{W}_i(h_{-i}^{[k-1]}, a_i)$.

For each J_{-i}^{k-1}, let $\{a_i^1\left(J_{-i}^{k-1}\right), \ldots, a_i^{\left|\mathcal{A}_i^k\right|}\left(J_{-i}^{k-1}\right)\}$ be a sequence of all the elements of $\mathcal{A}_i^k$ such that

$$\lim_{T\to\infty}\lim_{\delta\to 1}\frac{\tilde{W}_i\left(J_{-i}^{k-1}, a_i^1\left(J_{-i}^{k-1}\right)\right)}{T} \geq \cdots \geq \lim_{T\to\infty}\lim_{\delta\to 1}\frac{\tilde{W}_i(J_{-i}^{k-1}, a_i^{\left|\mathcal{A}_i^k\right|}\left(J_{-i}^{k-1}\right))}{T}.$$

For each J_{-i}^{k-1} and $n \in \{1, \ldots, \left|\mathcal{A}_i^k\right|\}$, let $1_{\left[J_{-i}^{k-1}, n\right]} : H_{i-1}^{[k,m]} \times \hat{\Omega}_{-(i-1,i,l)} \to \{0,1\}$ is the indicator function such that $1_{\left[J_{-i}^{k-1}, n\right]}(h_{i-1}^{[k,m]}, \hat{\Omega}_{-(i-1,i,l)}) = 1$ if the random event $\Psi_{i-1,i,-(i-1,i,l)}(\{a_i^n(J_{-i}^{k-1}),$

$\ldots, a_i^{|\mathcal{A}_i|}(J_{-i}^{k-1})\})$ is counted more than Z_{KT} times in the kth round when player $i-1$ monitors player i based on the messages from players $-(i-1,i,l)$ and $1_{[J_{-i}^{k-1},n]}(h_{i-1}^{[k,m]}, \hat{\Omega}_{-(i-1,i,l)}) = 0$ otherwise. Likewise, for each $a_i \in A_i$, let $1_{a_i} : H_{i-1}^{[k,m]} \times \hat{\Omega}_{-(i-1,i,l)} \to \{0,1\}$ is the indicator function such that $1_{a_i}(h_{i-1}^{[k,m]}, \hat{\Omega}_{-(i-1,i,l)}) = 1$ if the random event $\Psi_{i-1,i,-(i-1,i,l)}(\{a_i\})$ occurs more than Z_{KT} times in the kth round when player $i-1$ monitors player i based on the messages from players $-(i-1,i,l)$ and $1_{a_i}(h_{i-1}^{[k,m]}, \hat{\Omega}_{-(i-1,i,l)}) = 0$ otherwise.

Let C be a real number satisfying $C > 2\overline{u}$ and $\tilde{\theta}^k$ be

$$
\begin{aligned}
& \tilde{\theta}^k(h_{i-1}^{[k,m]}, I_{-i}^{k-1}, \hat{\Omega}_{-(i-1,i,l)}) \\
= \;& -T_b C + \sum_{a_i \in \mathcal{A}_i^k} 1_{a_i}(h_{i-1}^{[k,m]}, \hat{\Omega}_{-(i-1,i,l)}) KT\eta + \sum_{n=1}^{\left|A_i^k\right|} 1_{\left[I_{-i}^{k-1},n\right]}(h_{i-1}^{[k,m]}, \hat{\Omega}_{-(i-1,i,l)}) \lambda^k\left(I_{-i}^{k-1}, n\right),
\end{aligned}
$$

where $\lambda^k\left(I_{-i,k-1}, n\right)$ solves

$$
\begin{aligned}
& V_i\left(J_{-i}^{k-1}\right) \\
= \;& \tilde{W}_i\left(J_{-i}^{k-1}, a_i\right) - \delta^{\tilde{T}_k} C T_b + \delta^{\tilde{T}_k} \sum_{\tilde{a}_i \in A_i^k} \Pr\left(1_{\tilde{a}_i} \mid J_{-i}^{k-1}, a_i\right) KT\eta \\
& + \delta^{\tilde{T}_k} \sum_{I_{-i}^{k-1}} \Pr\left(I_{-i}^{k-1} \mid J_{-i}^{k-1}\right) \sum_{n=1}^{\left|A_i^k\right|} \Pr(1_{\left[I_{-i}^{k-1},n\right]} \mid J_{-i}^{k-1}, a_i) \lambda^k\left(I_{-i}^{k-1}, n\right) \qquad (21)
\end{aligned}
$$

for all J_{-i}^{k-1} and $a_i \in \mathcal{A}_i^k$. Similarly to **Lemma 22**, we have the following:

Lemma 25 *For all K, there exists $\overline{T}$ such that for all $T > \overline{T}$, there exists $\bar{\delta} \in (0,1)$ such that for all $\delta \in [\bar{\delta}, 1)$, (21) has a unique solution and for all $(h_{i-1}^{[k]}, I_{-i}^{k-1})$,*

$$
\theta^k(h_{i-1}^{[k]}, I_{-i}^{k-1}) < \left(\tilde{T}_k - KT\right)\overline{u} + \left|\mathcal{A}_i^k\right| KT\eta.
$$

Also, using this transfer scheme, (10) holds for the initial period of the kth main round with $x \in X^G$ and player i's continuation payoff after history $h_{i-1}^{[k-1]}$ is equal to $V_i(h_{-i}^{[k-1]})$.

Proof. Analogous to **Lemma 22**. ■

Then, the remaining task is to specify θ^{-1}. Suppose that θ^{-1} is decomposable as

$$\theta^{-1}(h_{i-1}^{[-1]}) = \sum_{x\in X^G} 1_x(h_{i-1}^{[-1]})\lambda^{-1}(x) + \sum_{t=1}^{\tilde{T}_{-1}} \frac{u_i(a_{i-1,t},\omega_{i-1,t})}{\delta^{\tilde{T}_{-1}+1-t}},$$

where $(a_{i-1,t},\omega_{i-1,t})$ is player $(i-1)$'s action and signal in the tth period of the block game and $\left(\lambda^{-1}(x)\right)_{x\in X_G}$ solve

$$\begin{aligned}\sum_{t=1}^{T_b}\delta^{t-1}\overline{w}_i &= -T_b\left(\eta+\lim_{\delta\to 1}\left(\overline{w}_i^P(\delta)-\overline{w}_i\right)\right)\\ &\quad +\delta^{\tilde{T}_{-1}}\sum_{\tilde{x}\in X^G}\Pr\left(1_{\tilde{x}}|\sigma^x\right)\lambda^{-1}(\tilde{x})+\delta^{\tilde{T}_{-1}+\tilde{T}_0}\sum_{h_{-i}^{[0]}\in H_{-i}^{[0]}}\Pr(h_{-i}^{[0]}|\sigma^x)V_i(h_{-i}^{[0]}) \quad (22)\end{aligned}$$

for all $x_{-i}\in X_{-i}^G$. Then, the analogous argument to **Lemma 24** establishes the following lemma.

Lemma 26 *1. Under U_i^G, (10) holds for the initial period of the coordination round and any period of each monitoring round for the coordination round.*

2. There exists $\overline{T}$ such that for all $T>\overline{T}$, there exists $\bar{\delta}\in(0,1)$ such that for all $\delta\in[\bar{\delta},1)$, (22) has a unique solution and

$$\theta^{-1}(h_{i-1}^{[-1]}) < \tilde{T}_{-1}\overline{u} - (|A_i|+1)T_b\eta.$$

From the above argument, there exists $\overline{T}$ such that for all $T>\overline{T}$, there exists $\overline{\delta}<1$ such that for all $\delta\in\left(\overline{\delta},1\right)$ and for all $i\in I$, $l\geq 0$, $h^{t_l}\in H^{t_l}$, $h_{i-1}^{T_b}\in H_{i-1}^{T_b}$ and $x\in X$ with $x_{i-1}=G$, (10) and (11) hold. In addition, from (16),

$$U_i^G(h_{i-1}^{T_b})\leq 0 \text{ for all } h_{i-1}^{T_b}.$$

Fix some $T > \overline{T}$. Then, since $U_i(h_{i-1}^{T_b}) \geq -2T_b\overline{u}$ for all $h_{i-1}^{T_b}$, for sufficiently large δ,

$$U_i^G(h_{i-1}^{T_b}) > -\frac{\overline{w}_i - \underline{w}_i}{1-\delta} \text{ for all } h_{i-1}^{T_b} \in H_{i-1}^{T_b},$$

which means (12) holds.

References

[1] Abreu, D., P. Milgrom and D. Pearce (1991): "Information and Timing in Repeated Partnerships," *Econometrica*, 59, 1713-33

[2] Aoyagi, M. (2002): "Collusion in Dynamic Bertrand Oligopoly with Correlated Private Signals and Communication," *Journal of Economic Theory*, 102, 229-248.

[3] Bhaskar, V. and I. Obara (2002): "Belief-Based Equilibria in the Repeated Prisoners' Dilemma with Private Monitoring," *Journal of Economic Theory*, 102, 49-69.

[4] Chen, B. (2009): "A Belief-Based Approach to the Repeated Prisoners' Dilemma with Asymmetric Private Monitoring," *Journal of Economic Theory*, 145, 402-420.

[5] Compte, O. (1998): "Communication in Repeated Games with Imperfect Private Monitoring," *Econometrica*, 66, 597-626.

[6] Ely, J., J. Hörner and W. Olszewski (2004): "Dispensability of Public Randomization Device," *mimeo.*

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

[8] Ely, J. and J. Välimäki (2002): "A Robust Folk Theorem for the Prisoners' Dilemma," *Journal of Economic Theory*, 102, 84-105.

[9] Fong, K., O. Gossner, J. Hörner and Y. Sannikov (2007): "Efficiency in a Repeated prisoners' Dilemma with Imperfect Private Monitoring," *mimeo.*

[10] Fudenberg, D. and D. Levine (1991): "Approximate Equilibria in Repeated Games with Imperfect Private Information," *Journal of Economic Theory*, 54, 26-47.

[11] Fudenberg, D. and D. Levine (1994): "Efficiency and Observability with Long-Run and Short-Run Players," *Journal of Economic Theory*, 62, 103-135.

[12] Fudenberg, D. and D. Levine (2007): "The Nash-Threats Folk Theorem with Communication and Approximate Common Knowledge in Two Player Games," *Journal of Economic Theory*, 132, 461-473.

[13] Fudenberg, D., D. Levine and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62, 997-1040.

[14] Fudenberg, D. and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting and with Incomplete Information," *Econometrica*, 54, 533-554.

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

[16] Hörner, J. and W. Olszewski (2009): "How Robust is the Folk Theorem with Imperfect Public Monitoring?," *Quarterly Journal of Economics*, 124, 1773-1814.

[17] Kandori, M. (2002): "Introduction to Repeated Games with Private Monitoring," *Journal of Economic Theory*, 102, 1-15.

[18] Kandori, M. (2008): "Weakly Belief-Free Equilibria in Repeated Games with Private Monitoring," *mimeo.*

[19] Kandori, M. and H. Matsushima (1998): "Private Observation, Communication and Collusion," *Econometrica*, 66, 627-652.

[20] Kandori, M. and I. Obara (2006): "Efficiency in Repeated Games Revisited: the Role of Private Strategies," *Econometrica*, 72, 499-519.

[21] Lehrer, E. (1989): "Lower Equilibrium Payoffs in Two-Player Repeated Games with Non-Observable Actions," *International Journal of Game Theory*, 18, 57-89.

[22] Mailath, G. and S. Morris (2002): "Repeated Games with Almost-Public Monitoring," *Journal of Economic Theory* 102, 189-228.

[23] Mailath, G. and S. Morris (2006): "Coordination Failure in Repeated Games with Almost-Public Monitoring," *Theoretical Economics*, 1, 311-340.

[24] Mailath, G. and L. Samuelson (2006): *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, New York, NY.

[25] Matsushima, H. (2004): "Repeated Games with Private Monitoring: Two Players," *Econometrica*, 72, 823-852.

[26] Miyagawa, E., Y. Miyahara and T. Sekiguchi (2008): "The Folk Theorem for Repeated Games with Observation Costs," *Journal of Economic Theory*, 139, 192-221.

[27] Obara, I. (2009): "Folk Theorem with Communication," *Journal of Economic Theory*, 144, 120-134.

[28] Piccione, M. (2002): "The Repeated prisoners' Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory*, 102, 70-83.

[29] Radner, R. (1985): "Repeated Principal-Agent Games with Discounting," *Econometrica*, 53, 1173-1198.

[30] Radner, R. (1986): "Repeated Partnership Games with Imperfect Monitoring and No Discounting," *Review of Economic Studies*, 53, 43-57.

[31] Sekiguchi, T. (1997): "Efficiency in Repeated prisoners' Dilemma with Private Monitoring," *Journal of Economic Theory*, 76, 345-361.

[32] Stigler, G. (1964): "A Theory of Oligopoly," Journal of Political Economy, 72, 44-61.

[33] Takahashi, S. (2010): "Community Enforcement when Players Observe Partners' Past Play," *Journal of Economic Theory*, 145, 42-62.

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

[35] Yamamoto, Y. (2009a): "A Limit Characterization of Belief-Free Equilibrium Payoffs in Repeated Games," forthcoming in *Journal of Economic Theory*.

[36] Yamamoto, Y. (2009b): "Repeated Games with Private and Independent Monitoring," *mimeo*.