Finitely Repeated Games with Monitoring Options

Yasuyuki Miyahara\textsuperscript{a} * Tadashi Sekiguchi\textsuperscript{b}

Abstract

We study finitely repeated games where players can decide whether to monitor the other players' actions or not every period. Monitoring is assumed to be costless and private. We compare our model with the standard one where the players automatically monitor each other. Since monitoring other players never hurts, any equilibrium payoff vector of a standard finitely repeated game is an equilibrium payoff vector of the same game with monitoring options. We show that some finitely repeated games with monitoring options have sequential equilibrium outcomes which cannot be sustained under the standard model, even if the stage game has a unique Nash equilibrium. We also present sufficient conditions for a folk theorem, when the players have a long horizon.

\textit{JEL Classification:} C72; C73

Keywords: Finitely repeated games; Imperfect monitoring; Folk theorem

\textsuperscript{*}We thank Kenichi Amaya, Andreas Blume and Prajit Dutta for very helpful comments. We also thank Christian Hellwig (the editor) and two anonymous referees, whose comments significantly improved the paper. Miyahara is grateful to financial support from the Grants-in-Aid for Scientific Research (20730246, 24530199). Sekiguchi gratefully acknowledges financial supports from the Grants-in-Aid for Scientific Research (20223001, 20530153). This research has been developed when Miyahara stayed at Columbia University. Miyahara is thankful to the hospitality of Columbia University.

\textsuperscript{a}Graduate School of Business Administration, Kobe University, 2-1 Rokkodai-cho, Nada-ku, Kobe 657-8501, Japan

\textsuperscript{b}Institute of Economic Research, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan

\textsuperscript{*}Corresponding author. 
\textit{E-mail addresses:} miyahara@kobe-u.ac.jp (Y. Miyahara), sekiguchi@kier.kyoto-u.ac.jp (T. Sekiguchi)
1 Introduction

A standard assumption in the models of repeated games is that of perfect monitoring; every player directly observes the actions of the others every period. Early folk theorems for discounted infinitely repeated games (Fudenberg and Maskin [6]) and for finitely repeated games (Benoit and Krishna [2]) assume perfect monitoring. In some applications, however, this assumption is too strong. This criticism motivated a large body of literature on the repeated games with imperfect monitoring, where each player only receives partial information of the other players’ actions.

In this paper, we scrutinize the assumption of perfect observability from a different viewpoint. We regard the complete information about the past play as a consequence of the players’ conscious efforts. It needs attention to monitor somebody. Therefore, the acquisition of information is part of a player’s decision making, and an incentive problem arises. Some papers address this type of problem, under a (presumably reasonable) assumption that the information acquisition entails costs (Ben-Porath and Kahneman [4], Kandori and Obara [10], Miyagawa, Miyahara and Sekiguchi [14], Flesch and Perea [5]). In contrast, we assume that the information acquisition is costless. That is, complete data about the actions of the others are freely available every period, and every player simply decides whether to get them or not.

We also assume that each player’s information acquisition is completely unobservable to any other player. The other players neither directly observe the decision nor receive any signal of it. As a result, the information acquisition just enables a better decision making, without causing any punishment in itself. Therefore it is never suboptimal to monitor the other players. Based on this observation, we show that any (subgame perfect) equilibrium payoff vector in the standard model, where the players automatically monitor each other, is a (sequential) equilibrium payoff vector in our model with monitoring options.

We next ask whether the option not to observe the other players can expand the equilibrium payoff vector set. In particular, we ask the question in the context of finitely repeated games whose stage game has a unique Nash equilibrium. It is well known that in the standard model, which we will call automatic monitoring, the unique subgame perfect equilibrium for those games is repetition of the stage Nash equilibrium. However, we show that some finitely repeated games with monitoring options have sequential equilibrium outcomes which cannot be sustained under automatic monitoring. Furthermore, under some conditions, a folk theorem (Benoit and Krishna [2], Smith [16], Gossner [8]) holds if the players’ horizon is sufficiently long.

In our model, the players’ options not to monitor other players create two strategic possibilities. First, a potential deviator may be monitored only by a subset of the other players. Second, he may be randomly monitored by the other players. In the first scenario, only the monitors detect a deviation and punish the deviator, while the other players play as if no deviation has occurred. This punishment is enforceable because the non-monitors believe that no player has deviated.\(^1\)

\(^1\)This is true only when the non-monitors cannot detect the deviation from their payoffs. We thus assume that all players receive all stage payoffs at the end of the repeated game. An alternative assumption would be that the payoffs are stochastic and do not reveal a deviation for sure.
In the second scenario, each player either notices a deviation and responds to it, or plays as if no deviation has occurred, depending on his random monitoring decision. Consequently, the deviator faces mixed continuation strategies, where he does not know whether he is going to be punished or not. Again, this punishment is enforceable because a player who did not monitor the other players believes that no player has deviated. One important difference between the two scenarios is that the first one, based on monitoring by a subset of players, needs three or more players, while the second scenario, based on random monitoring, can work in two-player games.\footnote{Another difference is that the monitoring by a subset of players also works in the case where the players could observe their monitoring decisions, but the random monitoring does not.}

The ability not to monitor other players is, to some extent, related to a player’s inability to respond to past observations. In fact, some frictions and commitments introduced into the model also have that effect, and they serve as a vehicle to achieve cooperation in some finitely repeated games. For example, Lipman and Wang \cite{12} show that introduction of switching costs can make cooperation possible in stage games like prisoners’ dilemma.\footnote{Our conditions do not apply to the stage games that are solvable by the iterative eliminations of strongly dominated actions, and therefore do not cover the prisoners’ dilemma games.} Renou \cite{15} also shows that cooperation can be sustained for stage games including prisoners’ dilemma, if the game has an initial stage where players strategically decide to commit to a subset of their strategy sets. In contrast, our model does not rely on exogenous switching costs or the players’ ability to commit not to play certain strategies in the whole course of play. As for the latter ability, the players’ horizon must be long in order to achieve cooperation. The longer the horizon is, the harder may be to make credible commitments.

The literature usually assumes that the explicit information acquisition is costly. However, if it is just glimpsing, zero monitoring cost may be a reasonable approximation. Another scenario of costless and optional monitoring is delegated monitoring. Suppose a store owner has a regular customer who lives in front of a rival shop. Upon the regular’s visit, the owner can simply ask him if she wants to know the rival’s action. Otherwise, she can choose a different topic. If the choice of topic is costless (or more generally, equally costly), the situation can be modeled as optional monitoring.\footnote{This interpretation of delegated monitoring is interesting, when the player hiring the monitor sets up a default action. Suppose that the monitor’s default action is not to report the other players’ actions, so that the player must ask him if she wants to know the actions. If it is costly to ask him, the situation reduces to that of costly information acquisition. Alternatively, suppose that the monitor’s default is to report the actions. Then, the player must ask him not to report them, if she does not want to know them. Since the information never hurts, however, she would never refuse it. Therefore, the situation reduces to the standard, automatic monitoring.}

In this paper, we show how monitoring options allow different equilibria supported by a credible punishment. If we allow non-credible punishments, the monitoring options do not help because they do not change the minimax values of either the stage game or any finitely repeated game. Therefore, our setting does not improve existing Nash folk theorems by Benoît and Krishna \cite{3} and González-Díaz \cite{7}.

We are not the first to show existence of a nontrivial equilibrium in finitely repeated games whose stage game has a unique Nash equilibrium. If the monitoring is imperfect, some finitely repeated games with a unique stage Nash equilibrium have nontrivial equilibria. This is shown by Kandori \cite{9} for the case of imperfect private
monitoring, and by Mailath, Matthews and Sekiguchi [13] for the case of imperfect public monitoring. Note that our model is another class of private monitoring, because observations are private. We show that this class also admits nontrivial equilibria (and a folk theorem).

The rest of this paper is organized as follows. In Section 2, we introduce the model. We compare our model with the standard one in Section 3. In Section 4, we present two examples of games with nontrivial equilibria, based on the idea of monitoring by a subset of players and random monitoring, respectively. In Sections 5 and 6, we generalize the two ideas respectively, and provide sufficient conditions for existence of nontrivial equilibria and a folk theorem. We conclude in Section 7.

2 Model

In this section, we first describe the stage game of the model and then formulate the finitely repeated games. This section also introduces our solution concept.

2.1 Stage Game

A finite, \( n \)-player normal-form game \( G \) is given. Let \( A_i \) be a finite set of player \( i \)’s actions. Let \( \mathcal{A}_i \) be the set of probability distributions on \( A_i \), so that \( \mathcal{A}_i \) is the set of player \( i \)’s mixed actions. Define \( A = \prod_{i=1}^{n} A_i \) and \( \mathcal{A} = \prod_{i=1}^{n} \mathcal{A}_i \). Also for each \( i \), define \( A_{-i} = \prod_{j \neq i} A_j \) and \( \mathcal{A}_{-i} = \prod_{j \neq i} \mathcal{A}_j \). Player \( i \)’s stage payoff function is \( u_i : A \to \mathbb{R} \). We extend the domain of each \( u_i \) to \( \mathcal{A} \), in an obvious way.

The stage game of our model consists of playing this game \( G \) and then deciding whether to monitor the other players’ (realized) actions or not. We assume that the monitoring decision is binary: either to monitor all other players or not to monitor any player at all.\(^5\) Each player \( i \)'s monitoring decision is to choose from \( M_i = \{0, 1\} \), where 1 denotes observing all other players and 0 denotes monitoring no other player. In contrast with some papers which assume costly monitoring (Ben-Porath and Kahneman [4], Kandori and Obara [10], Miyagawa, Miyahara and Sekiguchi [14], Flesch and Perea [5]), we assume that the monitoring decision entails no cost.

Each player \( i \) first decides which action \( a_i \in A_i \) to choose, and then, depending on his actual choice \( a_i \), chooses \( m_i \in M_i \). A (pure) stage action of player \( i \) is defined as a pair \( s_i = (a_i, \mu_i) \), where \( a_i \in A_i \) and \( \mu_i \) is a function from \( A_i \) to \( \{0, 1\} \). For each \( a_i \in A_i \), \( \mu_i(a_i) \) specifies whether to observe the other players or not, given that player \( i \) has selected \( a_i \). We allow the players to randomize, so that each player can choose a mixed stage action, which is a probability distribution on the set of stage actions. Let \( S_i \) be the set of mixed stage actions of player \( i \).

For \( z \in \{0, 1\} \), let \( \mu_i^z : A_i \to \{0, 1\} \) be the function such that \( \mu_i^z(a_i) = z \) for any \( a_i \in A_i \). For example, \((a_i, \mu_i^1)\) is a stage action such that player \( i \) chooses \( a_i \) and then observes the other players irrespective of his action. For \( \alpha_i \in \mathcal{A}_i \) and \( z \in \{0, 1\} \), we

\(^5\)As we discuss in Section 7, our results extend to the case where each player can choose to monitor any subset of the other players.
abuse notation and denote the mixed stage action where the stage action \((a_i, \mu^z_i)\) is selected with the same probability as \(a_i\) chooses \(a_i\), by \((a_i, \mu^z_i)\).

### 2.2 Finitely Repeated Game

We consider a finitely repeated game where the stage game described in the previous subsection is played in periods \(t = 1, 2, \ldots, T\). We call the game a \(T\)-period repeated game, and denote it by \(G(T)\).

In each period \(t \geq 1\), each player \(i\) chooses a mixed stage action from \(S_i\). If he monitors the other players, he learns their realized actions. Otherwise, he learns nothing about their actions. We assume that the monitoring decision is unobservable at all. Namely, a player does not observe either the other players’ monitoring decisions or any signal of them.

Also, we want to assume that if a player does not monitor the others, he has no way to know their actions. We can think of two assumptions to that effect.

(A) The players receive all stage payoffs at the end of this game.

(B) The stage payoff of each player is stochastic, and its probability distribution depends on the action profile. Each \(u_i\) denotes the expected value of the realized payoff. Moreover, the set of realized payoffs is finite, and the support of the realized payoffs is independent of the action profile. In this case, he cannot know actions from the payoffs.\(^6\)

We think the assumption (B) is more plausible in many applications, but the notation becomes more complicated. Hence in what follows, we adopt the assumption (A). However, we emphasize that the subsequent analysis is valid under the assumption (B), too.

Player \(i\)’s history at period \(t \geq 2\) consists of all his past actions, all his past monitoring decisions, and the other players’ past actions in all periods he monitored them. For \(t \geq 2\), let \(H^t_i\) be the set of player \(i\)’s histories at period \(t\). For convenience, let \(H^1_i\) be an arbitrary singleton for each \(i\), whose sole element is interpreted as player \(i\)’s history at period 1. Then, the set of player \(i\)’s histories is

\[
H_i = \bigcup_{t=1}^{T} H^t_i.
\]

A (behavioral) strategy of player \(i\) in \(G(T)\) is a function \(\sigma_i\) from \(H_i\) to \(S_i\). Given a strategy profile \(\sigma = (\sigma_i)_{i=1}^n\), one can compute the probability distribution of the action profiles in all periods, denoted by \((a(t))_{t=1}^T\). We assume that player \(i\)’s expected payoff of the profile \(\sigma\) in \(G(T)\) is the average, undiscounted sum of the stage payoffs:

\[
\frac{1}{T} E \left[ \sum_{t=1}^{T} u_i(a(t)) \mid \sigma \right].
\]

\(\text{\textsuperscript{6}}\)We obtain the same conclusion even when the set of realized payoffs is an interval, if we additionally assume that the density of any payoff under any action profile is finite and bounded away from zero.
We assume no discounting only for easing exposition. Our analysis easily extends to the case of little discounting, and we can obtain similar results.

### 2.3 Solution Concept

Our solution concept is sequential equilibrium, although we also consider Nash equilibrium in Section 3, for the sake of comparison. Since the original definition of sequential equilibrium by Kreps and Wilson [11] is for finite extensive form games, we adapt their definition to our setting.

Let us start with some terminologies about beliefs. A *system of beliefs* is a function which maps each history $h^t_i$ to a probability distribution of the other players’ history profiles, $(h^t_j)_{j \neq i}$. A strategy profile is *completely mixed* if every stage action is selected with nonzero probability at any history. A completely mixed strategy profile uniquely specifies a system of beliefs obtained from applying Bayes’ rule to the profile. Given a strategy profile $\sigma$, a system of belief is *consistent* if there exists a sequence of completely mixed strategy profiles converging to $\sigma$ (we will call such a sequence a *tremble*) such that the corresponding sequence of the systems of beliefs, obtained from Bayes’ rule, converges to it.

A strategy profile $\sigma$ is a *sequential equilibrium*, if there exists a consistent system of beliefs $\Psi$ such that for each player $i$, his continuation strategy at any history is optimal, given $\sigma_{-i}$ and the belief about the other players’ histories specified by $\Psi$.

### 3 Comparisons with the Standard Model

This section compares our model with the standard model of finitely repeated games, where each player chooses his action and then automatically monitors the other players every period. Let us denote this $T$-period repeated game by $G^1(T)$.

As mentioned above, monitoring the other players never hurts in our model. Hence, intuitively, one should expect that our model can do anything the standard model can do. The following result confirms that intuition.

**Proposition 1.** Any subgame perfect equilibrium payoff vector of $G^1(T)$ is a sequential equilibrium payoff vector of $G(T)$.

**Proof.** See Appendix A.

In view of Proposition 1, our next question should be whether monitoring options increase the equilibrium payoff vector set. As we will see later, the answer is sometimes yes. However, before proceeding to the problem, we point out that a different result obtains if the solution concept is Nash equilibrium. Namely, what the players can attain as a Nash equilibrium of $G(T)$ is the same as what they can sustain as a Nash equilibrium of $G^1(T)$.

**Proposition 2.** A payoff vector is a Nash equilibrium payoff vector of $G(T)$, if and only if it is a Nash equilibrium payoff vector of $G^1(T)$.

**Proof.** See Appendix B.
The hard part of this result is to construct a Nash equilibrium of $G^1(T)$ with the same payoff vector as a given Nash equilibrium of $G(T)$. The idea is to consider each player’s strategy of $G^1(T)$, where the player decides in each period whether to remember or ignore the other players’ actions of that period. Remembering (ignoring, respectively) the other players’ actions in $G^1(T)$ corresponds to monitoring (not monitoring, respectively) the other players in $G(T)$. Hence if the player decides whether to remember or ignore the other players’ actions in the same way he decides whether to monitor or not the other players under the original equilibrium strategy of $G(T)$, we have an outcome-equivalent strategy of $G^1(T)$. The strategy profile constitutes a Nash equilibrium of $G^1(T)$, as is shown in the proof.

Proposition 2 shows that the monitoring options do not affect the Nash equilibrium payoff vector set. Therefore, we cannot improve existing Nash folk theorems by Benoît and Krishna [3] and González-Díaz [7].

4 Examples

This section provides two examples, which demonstrate that a finitely repeated game may have a nontrivial equilibrium, even if its stage game has a unique Nash equilibrium. The examples are based on the two ideas we mentioned in the introduction: monitoring by a subset of players, and random monitoring.

4.1 Monitoring by a Subset of Players

Let us assume that $G$ has $n = 3$, $A_i = \{C, D\}$ for any $i$, and the payoff matrices as follows.

\[
\begin{array}{cc|cc}
 & C & D \\
 C & 10, 10, 10 & 0, 11, 0 \\
 D & 11, 0, 0 & 1, 1, 0 \\
\end{array}
\quad
\begin{array}{cc|cc}
 & C & D \\
 C & 1, 1, 1 & 2, 0, 1 \\
 D & 0, 2, 1 & 3, 3, 1 \\
\end{array}
\]

Note first that $C$ is a best response of player 1 only if player 2 chooses $C$ with a probability no less than 1/2. The same argument holds for player 2. Therefore, any Nash equilibrium such that either player 1 or 2 plays $C$ with a nonzero probability prescribes both players 1 and 2 to choose $C$ with a probability no less than 1/2. But against any such strategy profile, $C$ is a unique best response of player 3. Given $a_3 = C$, $C$ is not a best response for either player 1 or 2. As a result, any Nash equilibrium of $G$ must have $a_1 = a_2 = D$. Hence $(D, D, D)$ is a unique Nash equilibrium of $G$.

For any $T$, the only subgame perfect equilibrium outcome of $G^1(T)$ is to play $(D, D, D)$ in all periods. We claim, however, that $G(2)$ has a sequential equilibrium where $(C, C, C)$ is played in the first period. Let us consider the following strategy profile, denoted by $\sigma^*$.

- In period 1, player $i \in \{1, 2\}$ chooses the stage action $(C, \mu_i^1)$, and player 3 chooses the stage action $(C, \mu_3^0)$. 
In period 2,

(i) if player \( i \) observed the other players in period 1 and if the action profile of players 1 and 2 in period 1 was \((C, C)\), then he chooses \((D, \mu^1_i)\),

(ii) if player \( i \) observed the other players in period 1 and if the action profile of players 1 and 2 in period 1 was not \((C, C)\), then he chooses \((C, \mu^1_i)\),

(iii) if player \( i \in \{1, 2\} \) did not observe the other players in period 1, then he chooses \((a_i, \mu^1_i)\), where \( a_i \) is the action he did not choose in period 1, and

(iv) if player 3 did not observe the other players in period 1, then he chooses \((D, \mu^3_3)\).

Note that the path under \( \sigma^* \) is \((C, C, C)\) in period 1, and \((D, D, D)\) in period 2. Let \( \Psi^* \) be a consistent system of beliefs given \( \sigma^* \), such that any player at any history at period 2 believes that the other players did not deviate in their monitoring decisions. \( \Psi^* \) is obtained from a tremble such that in period 1, any stage action entailing a deviation in the monitoring decision is much less likely than any stage action entailing a deviation only in the action.

Let us verify that \( \sigma^* \) is a sequential equilibrium given \( \Psi^* \). We first check sequential rationality for player \( i \in \{1, 2\} \). At any history at period 2, player \( i \) believes that player 3 did not observe the other players, and that player \( 3 - i \) observed the other players. Hence player \( i \) believes that player 3 chooses \( D \) in period 2. His belief about the action of player \( 3 - i \) depends on his action and observation (if any) in period 1.

(i) If player \( i \) observed the other players in period 1 and if the action profile of players 1 and 2 in period 1 was \((C, C)\), then he believes that player \( 3 - i \) plays \( D \) in period 2. Hence playing \( D \) with any monitoring decision is optimal.

(ii) If player \( i \) observed the other players in period 1 and if the action profile of players 1 and 2 in period 1 was not \((C, C)\), then he believes that player \( 3 - i \) plays \( C \) in period 2. Hence playing \( C \) with any monitoring decision is optimal.

(iii) If he did not observe the other players in period 1, then he believes that no other player deviated and that player \( 3 - i \) plays what player \( i \) did not play in period 1. Hence playing what he did not play in period 1 is optimal, together with any monitoring decision.

We next consider the history at period 1. If player \( i \) conforms to \( \sigma^*_i \), his stage payoff is 10 in period 1 and 3 in period 2, and the average is 13/2. If he chooses \( D \) and then conforms to his continuation strategy, \((C, C, D)\) is played in period 2. Thus the stage payoff is 11 in period 1 and 1 in period 2, and the average is 6. Hence \( \sigma^*_i \) is optimal.

Next, consider sequential rationality for player 3. At any history at period 2, player 3 believes that the other players observed the other players. We have two cases to consider.

(i) If player 3 (deviantly) observed the other players in period 1 and found that their action pair was not \((C, C)\), he believes that players 1 and 2 play \( C \). Hence playing \( C \) with any monitoring decision is optimal.
(ii) Otherwise, player 3 believes that players 1 and 2 play $D$. Hence playing $D$ with any monitoring decision is optimal.

Finally, consider the history at period 1. Player 3 plays his static best response in both periods. Since his stage action does not affect the future play at all, $\sigma_3^*$ is optimal.

Hence $G(2)$ has a sequential equilibrium where mutual cooperation is sustained in period 1. Note that the equilibrium payoff is different from the stage game equilibrium for any player. Thus $G(2)$ has multiple equilibrium payoffs, and the multiplicity can be used to sustain other payoff vectors in $G(T)$ with large $T$’s. In particular, a folk theorem holds for this stage game, as we will see in Section 5.

The sequential equilibrium constructed here does not work under automatic monitoring. This is simply because any deviation by player 1 or 2 would be detected by player 3, too. Then the idea of punishing player 1 or 2 by the profile $(C, C, D)$ would fail, since player 3, anticipating the punishment, would be willing to choose $C$. If the monitoring is an option, player 3 is fine with not monitoring the others, because playing $\sigma_3^*$ does not require information about their actions. If he indeed does not monitor, then the punishment by the profile $(C, C, D)$ works. A key fact here is that the game played only by players 1 and 2, on the premise that player 3’s action is fixed at his static equilibrium action, has an equilibrium giving each player a smaller payoff than the equilibrium of the original game. In Section 5, we generalize this logic.

### 4.2 Random Monitoring

Suppose $G$ is the following two-player game.

<table>
<thead>
<tr>
<th></th>
<th>$L$</th>
<th>$C$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U$</td>
<td>4,4</td>
<td>1,0</td>
<td>6,3</td>
</tr>
<tr>
<td>$M$</td>
<td>3,0</td>
<td>3,6</td>
<td>0,3</td>
</tr>
<tr>
<td>$D$</td>
<td>1,4</td>
<td>4,0</td>
<td>6,3</td>
</tr>
</tbody>
</table>

Let us show that $(U, L)$ is a unique Nash equilibrium of this game. We first note that $M$ is not a best response of player 1 when player 2 does not play $L$ with positive probability. This is because $D$ is strictly better than $M$ when player 2 plays such a strategy. By a similar argument, $M$ is not a best response of player 1 when player 2 does not play $C$ with positive probability. Therefore, in any Nash equilibrium where $M$ is played with positive probability, player 2 must play $L$ and $C$ with positive probability. Namely, player 2 must be indifferent between $L$ and $C$.

Note that player 2 is indifferent between $L$ and $C$ only if player 1 plays $M$ with probability 0.4. In this case, however, $R$ is a unique best response of player 2. Therefore, no Nash equilibrium exists such that player 1 plays $M$ with positive probability. Moreover, when player 1 does not play $M$ with positive probability, $L$ is a unique best response of player 2. Therefore, we have a unique Nash equilibrium $(U, L)$, which implies that $G^1(T)$ with any $T$ has a unique subgame perfect equilibrium.

Let us consider $G(2)$ and the following strategy profile $\sigma^*$. In period 1, player 1 chooses $(M, \mu_1^1)$, and player 2 chooses $(C, \mu_2^1)$ and $(C, \mu_2^0)$ with equal probability. In period 2, each player decides his stage action in the following way.
Irrespective of the monitoring decision in period 1, player 1 chooses \((U, \mu^1_1)\) if he played \(M\) in period 1, and chooses \((M, \mu^1_1)\) if he did not play \(M\).

Player 2 chooses \((C, \mu^2_1)\) if he observed the opponent and the observed action was not \(M\), and chooses \((L, \mu^2_1)\) otherwise.

On the path, \((M, C)\) is played in period 1, and \((U, L)\) is played in period 2. The average equilibrium payoff vector is \((3.5, 5)\), which is different from the stage Nash equilibrium payoff vector for each player. Let \(\Psi^*\) be a consistent system of beliefs given \(\sigma^*\), such that any player at any history at period 2 believes that the other player did not deviate in his monitoring decision in period 1. Again, \(\Psi^*\) is obtained from a tremble where any stage action entailing a deviation in the monitoring decision is much less likely than any stage action entailing a deviation only in the action.

It remains to check sequential rationality. We first check optimality of player 1’s play in period 2.

- If player 1 played \(M\) in period 1, then playing \(U\) with any monitoring decision is optimal for player 1. This is because player 1 believes that player 2 plays \(L\) in period 2 when player 1 played \(M\) in period 1.

- If player 1 played an action other than \(M\) in period 1, then player 1 believes that player 2 observed player 1 with probability 0.5, irrespective of whether player 1 observed player 2 or not. That is, player 1 believes that player 2 plays \(L\) and \(C\) with equal probability. Therefore, playing \(M\) with any monitoring decision is optimal for player 1.

Next, let us check player 1’s incentive to play \(M\) in period 1. If player 1 follows \(\sigma^*_1\), then he obtains an average payoff 3.5. On the other hand, if he chooses an action other than \(M\) in period 1, then he obtains at most 4 in period 1, and by the above argument, obtains 3 in period 2. Hence his expected average payoff is at most 3.5. Therefore, \(\sigma^*_1\) is optimal.

Let us consider player 2. First, we check optimality of his play in period 2.

- If player 2 observed an action other than \(M\) in period 1, then playing \(C\) with any monitoring decision is optimal for player 2, because he believes that player 1 plays \(M\) in period 2.

- If player 2 observed \(M\), or if player 2 did not observe player 1, then playing \(L\) with any monitoring decision is optimal for player 2. This is because he believes that player 1 did not deviate and plays \(U\).

Finally, we check player 2’s incentive to play \(C\) in period 1. Note that player 2’s action in period 1 does not affect player 1’s play in period 2 if player 1 follows the strategy. Since player 1’s action in period 1 is pure, he has nothing to learn from it. Therefore, choosing \(C\) together with any monitoring decision is optimal, because \(C\) is a short-run best response to \(M\).
Hence we have a nontrivial sequential equilibrium, and each player’s payoff is different from the stage Nash equilibrium payoff. Therefore, as in the previous example, a folk theorem holds for this stage game if the players’ horizon is long enough.

Again, this sequential equilibrium does not work under automatic monitoring. The reason is that player 1’s deviation would be surely detected by player 2. With monitoring options, player 2 may randomly decide whether to monitor player 1 or not. Thus player 1 who deviated in period 1 faces either a player 2 who knows that or a player 2 who does not. Player 1 therefore optimizes against a mixture of the actions selected by the two types of player 2. If player 1’s action is optimal against the mixture and if the action of player 2 who monitored is optimal against player 1’s action, the continuation play constitutes a credible punishment. This is exactly how the equilibrium works.

This time, a key game to consider is the one defined on the premise that player 2 is forced to follow his static equilibrium with probability 0.5 but can choose any action with the remaining probability. The game has multiple equilibrium payoffs, and $(M, C)$ is an equilibrium punishing player 1. Section 6 generalizes this logic.

5 General Analysis: Monitoring by a Subset of Players

This section generalizes the idea of monitoring by a subset of players, presented in Subsection 4.1. Our argument consists of two steps. First, we seek a sufficient condition for existence of a nontrivial equilibrium. Second, we strengthen the condition so that a folk theorem also holds if the horizon is sufficiently long.

5.1 Nontrivial Equilibria

In what follows, we assume that $G$ has a unique Nash equilibrium, which we denote by $\alpha^* \in \mathcal{A}$. Thus repeated play of $\alpha^*$ is the only subgame perfect equilibrium of $G^1(T)$ for any $T$. In Section 4, however, we have seen that under optional monitoring, this assumption is consistent with existence of a nontrivial equilibrium and a folk theorem.

Let us call a proper, nonempty subset of $N = \{1, 2, \ldots, n\}$ a group. For any group $K$, let $A_K = \prod_{i \in K} A_i$ and $A_{-K} = \prod_{i \in K^c} A_i$. Also, we write $-K$ for $N \setminus K$.

**Definition 1.** Let $K$ be a group. The $K$-reduced game, denoted by $G_K$, is the normal-form game such that

- the set of players is $K$,
- the set of actions of each player $i \in K$ is $A_i$, and
- the payoff for each player $i \in K$ of an action profile $a_K \in A_K$ is

$$\tilde{u}_i(a_K) = u_i(a_K, \alpha^*_{-K}),$$

where $\alpha^*_{-K} = (\alpha^*_j)_{j \notin K}$.

In other words, $G_K$ is the game where the players in $K$ play $G$, on the premise that the players outside the group follow the static equilibrium $\alpha^*$. Note that for any group $K$, $\alpha^*_K = (\alpha^*_i)_{i \in K}$ is a Nash equilibrium of $G_K$. 

10
Definition 2. A group \( K \) satisfies the **distinct Nash payoff condition** if the \( K \)-reduced game \( G_K \) satisfies the following two conditions.

(A) The set of feasible payoff vectors of \( G_K \) has a dimension of \(|K|\), and for each \( i \in K \), there exist two Nash equilibria of \( G_K \), \( \alpha^1_K \) and \( \alpha^2_K \), such that

\[
\tilde{u}_i(\alpha^1_K) \neq \tilde{u}_i(\alpha^2_K).
\]

(B) For each \( i \in K \),

\[
\tilde{u}_i(\alpha^*_K) > \min_{a_{K\setminus\{i\}} \in A_{K\setminus\{i\}}} \max_{a_i \in A_i} \tilde{u}_i(a_i, \alpha_{K\setminus\{i\}}).
\]

(A) in the distinct Nash payoff condition is a standard sufficient condition for the folk theorem in finitely repeated games (Benoît and Krishna [2], Gossner [8]). The condition (B) states that the “default” Nash equilibrium of \( G_K \), \( \alpha^*_K \), gives each player more than his minimax value. Those two conditions guarantee that for any finitely repeated game with a sufficiently long horizon and with a stage game \( G_K \), each player has an equilibrium payoff less than the default Nash equilibrium payoff. This prepares punishments and provides incentives for players in \( K \) to choose a cooperative action in the first period of \( G(T) \).

We point out that no **two-player** game satisfies the distinct Nash payoff condition. With \( n = 2 \), \( G_K \) is always a one-person game. Such games always have a unique equilibrium payoff, and therefore the condition (A) always fails. Another important class of games with no group satisfying the distinct Nash payoff condition is the ones solvable by the iterative eliminations of strongly dominated actions.\(^7\)

**Proposition 3.** Suppose some group satisfies the distinct Nash payoff condition. Then there exists \( T \) such that any \( G(T) \) with \( T \geq T \) has a sequential equilibrium whose action profile in period 1 is not \( \alpha^* \).

**Proof.** See Appendix C. \( \square \)

Note that the result says nothing about equilibrium payoffs. The nontrivial equilibrium may have the same payoff vector as \( \alpha^* \). In order to prove a folk theorem, we need an additional condition, which we detail in the next subsection.

### 5.2 A Folk Theorem

This subsection establishes a folk theorem for finitely repeated games with monitoring options. For each \( i \), let \( v_i \) be player \( i \)'s minimax value:

\[
v_i = \min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} u_i(a_i, \alpha_{-i}).
\]

\(^7\)This is because any reduced game of a game that is solvable by the iterative eliminations of strongly dominated actions is also solvable by the iterative eliminations of strongly dominated actions. Hence any reduced game has a unique equilibrium, which denies the condition (A).
Under the assumption that mixed actions are not observable, our minimax value concept differs from the one in the early literature on folk theorems for finitely repeated games (Benoit and Krishna [2], Smith [16]), in the sense that we allow the other players to (independently) randomize.

Let $F$ be the convex hull of the set $\{(u_i(a))_{i=1}^n : a \in A\}$, which is the set of feasible payoff vectors. Also, let $F^* = \{(v_i)_{i=1}^n \in F : v_i \geq \underline{v_i} \forall i\}$, which is the set of feasible and (weakly) individually rational payoff vectors.

The nontrivial equilibrium we constructed in the proof of Proposition 3 has a particular form such that an action profile $\alpha \neq \alpha^*$ is played in the first period and then the static equilibrium $\alpha^*$ is played from the second period on. Moreover, the proof indicates a set of action profiles that can be the first-period action profile of a nontrivial equilibrium. Suppose $K$ satisfies the distinct Nash payoff condition, and let $A(K)$ be the set of all elements in $A$ such that (i) $\alpha_j$ is a best response against $\alpha_{-j}$ for any $j \notin K$, and (ii) $\alpha_K$ is pure. The proof implies that for any $\alpha \in A(K)$, $G(T)$ with large enough $T$ has a sequential equilibrium where $\alpha$ is played in the first period.

**Definition 3.** $G$ satisfies the nondegeneracy condition if, for every player $i$, there exist a group $K$ satisfying the distinct Nash payoff condition and $\alpha \in A(K)$ such that $u_i(\alpha) \neq u_i(\alpha^*)$.

Clearly, the nondegeneracy condition is stronger than existence of a group satisfying the distinct Nash payoff condition. See Appendix E for an example which satisfies the latter condition but not the former.

**Proposition 4.** Assume that $G$ satisfies the nondegeneracy condition. Also assume that the dimension of $F$ is $n$. Then for any $\varepsilon > 0$, there exists $T$ such that for any $T \geq T_i$ and any $v \in F^*$, $G(T)$ has a sequential equilibrium with a payoff vector $(w_i)_{i=1}^n$ satisfying $|v_i - w_i| < \varepsilon$ for any $i \in N$.

**Proof.** Since $G$ satisfies the nondegeneracy condition, for any $i$, $T_i$ exists such that any $G(T)$ with $T \geq T_i$ has a sequential equilibrium where $\alpha$ with $u_i(\alpha) \neq u_i(\alpha^*)$ is played in the first period and then $\alpha^*$ is played in all subsequent periods. Clearly, player $i$’s payoff of this equilibrium is not $u_i(\alpha^*)$.

Let $T_0 = \max_i T_i$. Then $G(T_0)$ has distinct sequential equilibrium payoffs for each player. Therefore, if we consider finitely repeated games whose stage game is $G(T_0)$, the condition for Gossner’s [8] folk theorem is satisfied. Consequently, the statement of the proposition holds if $T$ is large enough and is a multiple of $T_0$. However, we need not restrict ourselves to the multiples of $T_0$. If $T$ is so large that its quotient by $T_0$ is large enough, the remainder part of $G(T)$ (which can be interpreted as $G(T')$ with $T' < T_0$) has a negligible influence on the equilibrium payoffs of $G(T)$. Hence the statement is true for any $G(T)$ with sufficiently large $T$.

The stage game studied in Subsection 4.1 satisfies the conditions for Propositions 3 and 4. Let $K = \{1, 2\}$. The $K$-reduced game has two Nash equilibrium payoff vectors $^8$\(\alpha_K\) is required to be pure, because we assume that the mixed actions are not observable.
(1, 1) and (3, 3). The minimax value of each player in $G_K$ is 1, which is less than 3, the payoff of the default Nash equilibrium $(D, D)$. Also $G_K$ satisfies the dimensionality condition, so that $K$ satisfies the distinct Nash payoff condition. To see that $G$ satisfies the nondegeneracy condition, choose $K = \{1, 2\}$ and $\alpha = (C, C, C)$ for any $i$.

5.3 Discussions

The results in this section do not depend on the assumption that monitoring decisions are unobservable. Even if we assumed that each player’s monitoring decision is observable to the others, Propositions 3 and 4 would hold. Our argument is based on the idea that a potential deviator is not monitored by some other players, and he believes so in equilibrium. Therefore the additional information about the monitoring decisions, in equilibrium, would just confirm that belief. Hence the same construction works, whether monitoring is private or not.\(^9\)

The condition (A) in the distinct Nash payoff condition allows us to use the folk theorem by Gossner [8]. Thus one can replace it with sufficient conditions for other folk theorems, such as Benoît and Krishna [2] and Smith [16]. However, the minimax value concept for those results is the one by pure actions. Therefore we must strengthen the condition (B) in the distinct Nash payoff condition. With this modification, we have two alternative conditions under which the statement of Proposition 3 holds.

First, suppose that a group $K$ exists such that

(a) the $K$-reduced game has nonequivalent utilities (Abreu, Dutta and Smith [1]), and there exists a sequence $(K_g)_{g=0}^h$ with

$$\emptyset = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_h = K$$

so that, for $g = 1, \ldots, h$ and for $k \in K_g \setminus K_{g-1}$, there exist $(\hat{K}, \hat{\alpha}_K)$ and $(\tilde{K}, \tilde{\alpha}_K)$, where $\hat{K} \subseteq K_{g-1}$ and $\tilde{K} \subseteq K_{g-1}$, such that

- $\hat{\alpha}_{\hat{K}}$ is pure, and for any $i \not\in \hat{K}$, $\hat{\alpha}_i$ is a best response against $\hat{\alpha}_{-i}$,
- $\tilde{\alpha}_{\tilde{K}}$ is pure, and for any $i \not\in \tilde{K}$, $\tilde{\alpha}_i$ is a best response against $\tilde{\alpha}_{-i}$, and
- $\hat{u}_k(\hat{\alpha}) \neq \tilde{u}_k(\tilde{\alpha})$,

and

(b) for each $i \in K$,

$$\hat{u}_i(\alpha^*_K) > \min_{a_{K \setminus \{i\}} \in A_{K \setminus \{i\}}} \max_{a_i \in A_i} \hat{u}_i(a_i, a_{K \setminus \{i\}}).$$

Then we can apply the folk theorem by Smith [16], and show that $G(T)$ with sufficiently large $T$ has a nontrivial equilibrium. Note that the condition (a) is weaker than the condition (A), because some players may have unique Nash payoffs.\(^{10}\)

---

\(^9\)However, the construction based on random monitoring depends on unobservability of monitoring decisions.

\(^{10}\)However, as is pointed out in [16], generic games satisfying the condition (a) satisfy the condition (A).
Second, suppose that a two-player group \( K \) exists such that each \( i \in K \) has distinct Nash equilibrium payoffs of \( G_K \), and the above condition (b) holds. This time, we can apply the folk theorem by Benoît and Krishna [2], and show that \( G(T) \) with sufficiently large \( T \) has a nontrivial equilibrium. Note that in this case, \( G_K \) need not satisfy the dimensionality condition.

6 General Analysis: Random Monitoring

This section generalizes the idea of construction we described in Subsection 4.2, which is based on random monitoring. As in the previous section, we first seek a sufficient condition for existence of a nontrivial equilibrium.

This time, it is helpful to define the following notion of constrained games.

**Definition 4.** Let \( \eta = (\eta_i)_{i=1}^n \) be a vector such that \( \eta_i \in [0, 1] \) for each \( i \). The \( \eta \)-constrained game, denoted by \( G^\eta \), is the normal-form game such that

- the set of players is \( N_\eta = \{ i | \eta_i < 1 \} \),
- the set of actions of each player \( i \in N_\eta \) is \( A_i \), and
- the payoff for each player \( i \in N_\eta \) of an action profile \( a_{N_\eta} = (a_i)_{i \in N_\eta} \) is
  \[
  u_i^\eta(a_{N_\eta}) = u_i(a_i, (\alpha_j^\eta)_{j \notin N_\eta, (\eta_j \alpha_j^\eta + (1 - \eta_j)a_j)_{j \in N_\eta \setminus \{i\}}}.
  \]

The \( \eta \)-constrained game is the normal-form game played on the premise that each player \( i \) is constrained to follow \( \alpha_i^* \) with probability \( \eta_i \). For any \( \eta \)-constrained game, \( \alpha_{N_\eta} = (\alpha_i^*)_{i \in N_\eta} \) is its Nash equilibrium. For the stage game analyzed in Subsection 4.2, if we set \( \eta = (0, 1/2) \), \((M, C)\) is a Nash equilibrium of \( G^\eta \). A special case of the \( \eta \)-constrained games is the one where \( \eta_i \in \{0, 1\} \) for any \( i \), \( \eta_j = 1 \) for some \( j \), and \( \eta_k = 0 \) for some \( k \). This \( \eta \)-constrained game is equivalent to the \( N_{\eta} \)-reduced game.

Given \( \eta = (\eta_i)_{i \in N} \) and \( k \in N \), we define \( \eta^k = (\eta_i^k)_{i \in N} \) as the vector such that

\[
\eta_i^k = \begin{cases} 0 & \text{if } k = i, \\ \eta_i & \text{if } k \neq i. \end{cases}
\]

Given \( \eta \), let \( \mathcal{A}^*(\eta) \) be the set of all \( \alpha \in \mathcal{A} \) such that

(i) for any \( i \), if \( \alpha_i \) is not pure, then \( \alpha_i \) is a best response against \( \alpha_{-i} \), and

(ii) for any \( i \) such that \( \alpha_i \) is not a best response against \( \alpha_{-i} \), there exists a Nash equilibrium of \( G^\eta \), denoted by \( \alpha = (\alpha_j^\eta)_{j \in N_\eta} \), such that

\[
\max_{a_i' \in A_i} u_i(a_i', \alpha_{-i}) - u_i(\alpha) \leq u_i(\alpha^*) - u_i^\eta(\alpha^*).
\]

The following result proves that for any \( \eta \) and any \( \alpha \in \mathcal{A}^*(\eta) \setminus \{\alpha^*\} \), \( G(T) \) has a nontrivial sequential equilibrium where \( \alpha \) is played in period 1. In this equilibrium,
each player \(i\) plays \(\alpha_i\) and then monitors the other players with probability \(1 - \eta_i\). For each player \(i\) who does not play a best response under \(\alpha\), the other players detect his deviation only with the probabilities \((1 - \eta_j)_{j \neq i}\). This allows them to punish him by an equilibrium of a corresponding constrained game. This is the idea of (1), but the punishment equilibrium \(\alpha^i\) is not an equilibrium of the \(\eta\)-constrained game but that of the \(\eta^j\)-constrained game. This is because player \(i\) who has deviated would optimally respond against the other players’ actions with probability 1 in period 2, and the other players anticipate that.

**Proposition 5.** Let \(\alpha \in \mathcal{A}^*(\eta) \setminus \{\alpha^*\}\) for some \(\eta\). Then \(G(2)\) has a sequential equilibrium where \(\alpha\) is played in period 1, and \(\alpha^*\) is played in period 2.

**Proof.** See Appendix D.

If for each \(i\) there exist \(\eta\) and \(\alpha \in \mathcal{A}^*(\eta)\) such that \(u_i(\alpha) \neq u_i(\alpha^*)\), \(G(2)\) has multiple equilibrium payoffs for each player. Therefore, the same type of folk theorem we obtained in Proposition 4 holds by exactly the same argument.

In the stage game in Subsection 4.2, \(\alpha = (M, C)\) belongs to \(\mathcal{A}^*(\eta)\) for \(\eta = (0, 1/2)\). To see that, let \(\alpha^1 = (M, C)\) (note that only player 1 does not play a best response under \(\alpha\)). Therefore, as we have constructed, \((M, C)\) can be sustained in period 1.

In Proposition 5, the assumption that the monitoring decision is not observable at all is important. In our construction, a potential deviator does not know whether he is monitored or not, and therefore he faces uncertainty about the other players’ actions after a deviation. If he could learn their monitoring decisions, he would respond accordingly, which would make the punishment ineffective. Also, it is important that the length of the punishment period may not be more than one. Since the actions in the second period would reveal whether they had monitored or not in the first period, the punishment from period 3 on (if any) would not be effective. Therefore, while the reduced games are special cases of the constrained games, the idea of construction in Proposition 5 is different from the one in Proposition 3.

7 Concluding Remarks

Our results do not depend on the assumption that the players have only two alternatives of monitoring all others and monitoring nobody. We have constructed equilibria where some player at some period does not need information about the other players’ actions and thus does not observe them. Then the player does not benefit from knowing part of their actions. Consequently, we obtain the same results if we assume that the players can choose any subset of the other players to monitor. The assumption of binary choice is made only for notational convenience.

As we emphasized before, monitoring all other players is never suboptimal. As a result, all nontrivial sequential equilibria presented in this paper are not trembling-hand perfect.\(^{11}\) Take the nontrivial equilibrium we constructed in Subsection 4.1 (monitoring by a subset of other players) as an example. If the equilibrium is subject to

\(^{11}\)We thank Andreas Blume for alerting us to this point.
perturbations, players 1 and 2 choose $D$ with a small probability in period 1, and therefore both $(a_1, a_2) = (C, C)$ and $(a_1, a_2) = (D, D)$ are played with positive probability in period 2. Given the zero monitoring cost, it is uniquely optimal for player 3 to observe them in period 1, since his optimal actions against the above two profiles are different. This upsets the equilibrium, which works because player 3 does not monitor.

We have two responses about this failure of trembling-hand perfection. First, our focus is on how the monitoring options produce new credible punishments which are not possible under the standard automatic monitoring. What we have shown is that one can design credible punishments under the criterion of sequential rationality, but not under the criterion of trembling-hand perfection. Whether our punishments are credible enough or not will depend on the context.

Second, the equilibrium exhibits robustness if not only actions but also monitoring costs are perturbed.\(^{12}\) Suppose that players 1 and 2 choose $D$ with a vanishing probability, and all players have vanishing monitoring costs. Also assume that players 1 and 2 have relatively small monitoring costs, while player 3’s monitoring cost is relatively large. Then players 1 and 2 would monitor, because they want to know each other’s actions and it is worth paying the cost. However for player 3, it is not worth paying. This perturbation, though specific, supports our equilibrium.

For a large part of analysis, we have assumed that the stage game has a unique Nash equilibrium. This is because we believe that the effects of monitoring options are most salient in this case. However, the sequential equilibrium payoff vectors of $G(T)$ and the subgame perfect equilibrium payoff vectors of $G^1(T)$ can be different, for any stage game $G$. In fact, this difference can arise even for the infinitely repeated games. We did not analyze general stage games or infinitely repeated games, partly because it is difficult to characterize the whole equilibrium payoff set. The characterization is beyond the scope of this paper.

Our sufficient conditions for existence of a nontrivial equilibrium and for a folk theorem are not necessary. In Appendix E, we provide a finitely repeated game with nontrivial equilibria, which is not covered by our analysis. Thus we have yet to fully understand sufficient conditions on the stage payoffs for any finitely repeated game with monitoring options to have a unique equilibrium payoff vector. This question is related to the necessary condition for the folk theorem, studied in Smith [16] for automatic monitoring, and it is left for future research.

### Appendix A  Proof of Proposition 1

For each $i$ and $t \geq 2$, let $\hat{H}_i t$ be the set of player $i$’s histories at period $t$ such that he has selected $m_i = 1$ in all previous periods. Let $\hat{H}_i 1 = H_i 1$, and define $\hat{H}_i$ as

$$\hat{H}_i = \bigcup_{t=1}^T \hat{H}_i t.$$

Let $\sigma^1 = (\sigma_i^1)_{i=1}^n$ be a subgame perfect equilibrium of $G^1(T)$. Then each $\sigma_i^1$ can

---

\(^{12}\)We are grateful to Christian Hellwig (the editor) for suggesting this line of argument.
be regarded as a function from $\hat{H}_i$ to $A_i$. Let us define a strategy of player $i$ in $G(T)$, denoted by $\sigma_i$, as follows.

- For any $h_i \in \hat{H}_i$, let $\sigma_i(h_i) = (\sigma_i^1(h_i), \mu_i^1)$.
- For any history in $H_i \setminus \hat{H}_i$, assign an arbitrary stage action.

Namely, $\sigma_i$ is a strategy where player $i$ plays $G$ in the same way as $\sigma_i^1$ and then always monitors the other players, at any history where he has observed the other players in all past periods. His behavior at the other histories is arbitrary.

Let $\Psi^*$ be a consistent system of beliefs given $\sigma \equiv (\sigma_i)_{i=1}^n$, such that any player at any history believes that the other players have monitored all players in all past periods. $\Psi^*$ is obtained from a tremble such that at any history any stage action entailing a deviation in the monitoring decision is far less likely than any stage action entailing a deviation only in the action.

Let $\sigma^* = (\sigma_i^*)_{i=1}^n$ be a strategy profile of $G(T)$ such that for each $i$,

- $\sigma_i^*$ coincides with $\sigma_i$ on $\hat{H}_i$, and
- the continuation strategy of $\sigma_i^*$ at any $h_i \in H_i \setminus \hat{H}_i$ is optimal, given $\sigma_{-i}$ and the belief $\Psi^*$ specifies at $h_i$.

At any history of player $i$, he believes under $\Psi^*$ that any other player $j \neq i$ is at a history in $\hat{H}_j$. Since $\sigma_j$ and $\sigma_j^*$ coincide on $\hat{H}_j$ for any $j$, consistency of $\Psi^*$ given $\sigma$ implies consistency of $\Psi^*$ given $\sigma^*$. For any $i$ and any $h_i \in H_i \setminus \hat{H}_i$, player $i$’s continuation strategy is optimal given $\Psi^*$, by the definition of $\sigma_i^*$. For any $h_i \in \hat{H}_i$, since $\sigma^1$ is a subgame perfect equilibrium of $G^1(T)$, playing in the same way as $\sigma_i^1$ is optimal. This is exactly what $\sigma_i^*$ does, and therefore the continuation strategy of $\sigma_i^*$ is optimal given $\Psi^*$. Hence $\sigma^*$ is a sequential equilibrium with the same play as $\sigma^1$, and thus with the same payoff vector as $\sigma^1$.

\section*{Appendix B \ Proof of Proposition 2}

Suppose $\sigma^1 = (\sigma_i^1)_{i=1}^n$ is a Nash equilibrium of $G^1(T)$. For each $i$, define a strategy of $G(T)$, denoted by $\sigma_i$, exactly in the same way as we defined $\sigma_i$ in the proof of Proposition 1 (Appendix A). Since (i) $\sigma_i^1$ is optimal against $\sigma_{-i}^1$, for each $i$ in $G^1(T)$, and (ii) monitoring other players never hurts in $G(T)$, $\sigma_i$ is optimal against $\sigma_{-i}$ for each $i$ in $G(T)$. This implies that $(\sigma_i)_{i=1}^n$ is a Nash equilibrium of $G(T)$ with the same payoff vector as $\sigma^1$, which proves the “if” part.

Next, suppose $\sigma = (\sigma_i)_{i=1}^n$ is a Nash equilibrium of $G(T)$. Fix $i$, and let us suppose that player $i$ plays $G^1(T)$ in the following way. In each period, after choosing an action of the period, player $i$ privately chooses a number $m_i' \in \{0, 1\}$. Player $i$’s action in each period $t \geq 2$ depends on the past action profile path $(a(\tau))_{\tau=1}^{t-1}$, as well as on the path of the numbers he has chosen, $(m_i'(\tau))_{\tau=1}^{t-1}$. More concretely, player $i$ decides actions and numbers, possibly randomly, in the following way.

- In period 1, player $i$ chooses an action and a number in the same way as the mixed stage action $\sigma_i(h_i^1)$ chooses an action and $m_i \in \{0, 1\}$.
• Let \( t \geq 2 \), \( (a(\tau))_{\tau=1}^{t-1} \), and \( (m_i(\tau))_{\tau=1}^{t-1} \) be given. Then in \( G(T) \), there exists player \( i \)'s history at period \( t \), denoted \( h_i^t \), such that he chose \( a_i(\tau) \) in any period \( \tau \), and observed \( a_{-i}(\tau) \) in any period \( \tau \) such that \( m_i(\tau) = 1 \). Player \( i \) chooses an action and a number in the same way as the mixed stage action \( \sigma_i(h_i^t) \) chooses an action and \( m_i \in \{0,1\} \).

Let \( \sigma_i^1 \) be player \( i \)'s strategy in \( G^1(T) \) which is equivalent to the above description of play. Define \( \sigma^1 = (\sigma_i^1)_{i=1}^n \).

In \( G^1(T) \), \( \sigma^1 \) achieves exactly the same action profile path as \( \sigma \) in \( G(T) \). Since monitoring the other players never hurts in \( G(T) \), the best response payoff of each player \( i \) against \( \sigma^1 \) in \( G^1(T) \) equals the best response payoff of player \( i \) against \( \sigma_{-i} \) in \( G(T) \). Since \( \sigma_i \) is optimal against \( \sigma_{-i} \) for each \( i \) in \( G(T) \), \( \sigma_i^1 \) is optimal against \( \sigma_{-i}^1 \) in \( G^1(T) \). This implies that \( \sigma^1 \) is a Nash equilibrium of \( G^1(T) \), which proves the “only if” part. \( \square \)

Appendix C  Proof of Proposition 3

Fix a group \( K \) satisfying the distinct Nash payoff condition. Let \( F^*_K \) be the set of feasible and (weakly) individually rational payoff vectors of \( G_K \). By the condition (B), there exists \( \varepsilon > 0 \) such that for any \( i \in K \), \( \tilde{v}^i \in F^*_K \) exists such that

\[
\tilde{u}_i(a^*_K) - \varepsilon > \tilde{v}^i \geq \min_{a_{K\setminus\{i\}} \in A_{K\setminus\{i\}}} \max_{a_i \in A_i} \tilde{u}_i(a_i, a_{K\setminus\{i\}}).
\]  

(2)

Fix \( a_K \in A_K \) such that \( a_K \neq a^*_K \). Then \( \alpha_{-K} \in A_{-K} \) exists such that, when we write \( \hat{a} = (a_K, \alpha_{-K}) \),

\[
u_j(\hat{a}) = \max_{a_j \in A_j} u_j(a_j, \hat{a}_{-j})
\]

(3)

for any \( j \notin K \). Since \( a_K \neq a^*_K \), \( \hat{a} \) is not a Nash equilibrium of \( G \). Therefore, (3) implies that

\[
\Delta \equiv \max_{i \in K} \left[ \max_{a_i \in A_i} u_i(a_i, \hat{a}_{-i}) - u_i(\hat{a}) \right] > 0.
\]

Let us choose a natural number \( T_0 \) so that

\[
\Delta < T_0 \varepsilon.
\]

(4)

For any \( T' \), let \( G^1_K(T') \) be the \( T' \)-period repeated game with automatic monitoring whose stage game is \( G_K \). By the condition (A), the folk theorem by Gossner [8] applies. Since \( \tilde{v}^i \in F^*_K \) and (2) holds for any \( i \in K \), \( T_1 \) exists such that for any \( T' \geq T_1 \) and any \( i \in K \), \( G^1_K(T') \) has a subgame perfect equilibrium giving player \( i \) a payoff less than \( u_i(\alpha^*) - \varepsilon \).

Let \( T = \max\{T_0, T_1\} + 1 \). Fix \( T \geq T \), and let \( G_K(T - 1) \) be the \( (T - 1) \)-period repeated game with monitoring options whose stage game is \( G_K \). Since \( T - 1 \geq T_1 \), Proposition 1 implies that for each \( i \in K \), \( G_K(T - 1) \) has a sequential equilibrium giving player \( i \) a payoff less than \( u_i(\alpha^*) - \varepsilon \). We denote the sequential equilibrium by \( \hat{\sigma}^i \), and for a later purpose, we explicitly denote the underlying system of beliefs by \( \hat{\Phi}^i \).
For any player \( i \in K \) and \( j \in K \), we define a strategy of \( G(T - 1) \), denoted by \( \hat{\sigma}^1_i \), as an extension of \( \tilde{\sigma}^i_j \) (note that \( \tilde{\sigma}^i_j \) is a strategy of \( G_K(T - 1) \)) in a natural way. Namely, (i) \( \hat{\sigma}^1_i \) does not depend on any observations of the players outside \( K \), and (ii) \( \hat{\sigma}^1_i \) depends on his own past actions and his observations about the players in \( K \) in the same way as \( \tilde{\sigma}^i_j \).

For each \( i \), let \( \sigma^0_i \) be the strategy of \( G(T - 1) \) such that at any history player \( i \) chooses the mixed stage action \( (\alpha^*_i, \mu^0_i) \). For \( k \in \{0\} \cup K \), let \( \hat{\sigma}^k \) be the strategy profile of \( G(T - 1) \) such that any player \( i \in K \) plays \( \tilde{\sigma}^i_k \) and any player \( i \notin K \) plays \( \sigma^0_i \). By construction, the payoff of \( \hat{\sigma}^k \) for each player \( i \in N \) is \( u_i(\alpha^*) \), and for \( k \in K \), \( \hat{\sigma}^k \) gives player \( k \) a payoff less than \( u_k(\alpha^*) - \varepsilon \).

Furthermore, for \( k \in K \), let \( \Psi^k \) be the consistent system of beliefs given \( \hat{\sigma}^k \), obtained from a tremble where the players in \( K \) ignore any observations of the players outside \( K \), and play in the same way that makes \( \Psi^k \) consistent given \( \hat{\sigma}^k \). Under \( \Psi^k \), the players in \( K \) have the same beliefs as \( \hat{\Psi}^k \) about their histories, when they ignore all observations about the players outside \( K \). More formally,

- any player \( i \in K \) at any history believes that the marginal distribution of partial histories of the players in \( K \setminus \{i\} \), where their observations about the players outside \( K \) are ignored, is independent of player \( i \)'s observations about the players outside \( K \), and

- the marginal distribution equals the belief given by \( \hat{\Psi}^k \).

Let us construct a strategy profile of \( G(T) \), denoted by \( \sigma^* \), as follows.

For any player \( i \notin K \), \( \sigma^*_i \) prescribes the mixed stage action \( (\hat{\alpha}_i, \mu^0_i) \) in period 1. His continuation strategies at histories at period 2 are determined as follows.

(A) Suppose player \( i \) (deviantly) chose \( m_i = 1 \) in period 1, and found that a player \( k \in K \) did not select \( \hat{\alpha}_k \) and any player \( j \in K \setminus \{k\} \) selected \( \hat{\alpha}_j \) (recall that \( \hat{\alpha}_K \) is pure). Then player \( i \)'s continuation strategy is a best response against \( \hat{\sigma}^k \) such that at any subsequent history, the continuation strategy is optimal, given \( \hat{\sigma}^k \) and the belief given by \( \Psi^k \).

(B) Suppose otherwise. Then his continuation strategy is \( \hat{\sigma}^0_i \).

For any player \( i \in K \), \( \sigma^*_i \) prescribes the mixed stage action \( (\hat{\alpha}_i, \mu^1_i) \) in period 1. His continuation strategies at histories at period 2 are determined as follows.

(a) If player \( i \) chose \( m_i = 1 \) in period 1, let \( a(1) \in A \) be the combination of his own action and his observations. If \( k \in K \) exists (it may be that \( k = i \)) such that \( a_k(1) \neq \hat{\alpha}_k \) and \( a_j(1) = \hat{\alpha}_j \) for any \( j \in K \setminus \{k\} \), player \( i \)'s continuation strategy is \( \hat{\sigma}^k_i \). If no such \( k \) exists, player \( i \)'s continuation strategy is \( \hat{\sigma}^0_i \).

(b) If player \( i \) chose \( m_i = 0 \) in period 1, let \( a_i(1) \) be his own action in period 1. Player \( i \)'s continuation strategy is \( \hat{\sigma}^0_i \) if \( a_i(1) = \hat{\alpha}_i \), and \( \hat{\sigma}^1_i \) if \( a_i(1) \neq \hat{\alpha}_i \).

Let \( \Psi^* \) be a consistent system of beliefs given \( \sigma^* \), such that
• any player at any history believes that the other players did not deviate in their monitoring decisions,

• at any history of player $i$ at period $t \geq 3$ where he chose $m_i = 0$ in period 1, he believes that the other players did not deviate in period 1, and

• at any history of player $i$ at period $t \geq 3$ where he chose $m_i = 1$ in period 1 and found that a player $k \in K$ did not select $\hat{\alpha}_k$ and any player $j \in K \setminus \{k\}$ selected $\hat{\alpha}_j$, the conditional probability distribution of the other players’ history profiles from period 2 to period $t - 1$ equals the belief specified by $\Psi^k$, given his history from period 2 to period $t - 1$.

$\Psi^*$ is obtained from a tremble such that

• at any history, any stage action entailing a deviation in the monitoring decision is much less likely than any stage action entailing a deviation only in the action,

• any deviation in period 1 is much less likely than any deviation at any history at period $t \geq 2$, and

• given that player $i$ chose $m_i = 1$ in period 1 and found that a player $k \in K$ did not select $\hat{\alpha}_k$ and any player $j \in K \setminus \{k\}$ selected $\hat{\alpha}_j$, the sequence of his continuation strategies from period 2 on coincides with the tremble from which we obtain $\Psi^k$ (see the definition of $\Psi^k$).

It remains to prove that $\sigma^*$ is a sequential equilibrium given $\Psi^*$. First, consider a history of player $i \notin K$ at period $t \geq 2$.

(A) Suppose player $i$ chose $m_i = 1$ in period 1, and found that a player $k \in K$ did not select $\hat{\alpha}_k$ and any player $j \in K \setminus \{k\}$ selected $\hat{\alpha}_j$. Under $\Psi^*$, player $i$ believes that no other player deviated in his monitoring decision in period 1, and therefore believes that the other players play $\hat{\sigma}_0^i$ from period 2. Hence his continuation strategy is optimal, by its definition (if $t \geq 3$, he believes under $\Psi^*$ that the conditional probability distribution of the other players’ history profiles from period 2 to period $t - 1$ equals the belief specified by $\Psi^k$, given his history from period 2 to period $t - 1$).

(B) Suppose otherwise. Under $\Psi^*$, player $i$ believes that no other player deviated in his monitoring decision in period 1, and if he chose $m_i = 1$ in period 1, he found no player within $K$ unilaterally deviated from $\hat{\alpha}_K$. Thus player $i$ believes that the other players play $\hat{\sigma}_0^i$ from period 2. Note that if $t \geq 3$, his information from period 2 to period $t - 1$ may be inconsistent with $\hat{\sigma}_0^i$. However, under $\Psi^*$, he believes that the other players have not deviated in period 1, and conform to $\hat{\sigma}_0^i$ from now on. Hence conforming to $\hat{\sigma}_0^i$ is optimal.

We next consider the history of player $i \notin K$ at period 1. In period 1, only players outside $K$ randomize, and their actions do not affect future play. Therefore player $i$ has nothing to learn in period 1. Also, player $i$ plays his static best response (see (3)), and his action does not affect future play. As a result, $\sigma^*_i$ is optimal.

Next, consider a history of player $i \in K$ at period $t \geq 2$. 20
(a) If player \( i \) chose \( m_i = 1 \) in period 1, let \( a(1) \in A \) be the combination of his own action and his observations. Player \( i \) believes that no other player deviated in his monitoring decision in period 1. Therefore if \( k \in K \) exists such that \( a_k(1) \neq \hat{a}_k \) and \( a_j(1) = \hat{a}_j \) for any \( j \in K \setminus \{k\} \), player \( i \) believes that the other players play \( \hat{\sigma}_k \) from period 2. Hence conforming to \( \hat{\sigma}_k \) is optimal (if \( t \geq 3 \), he believes under \( \Psi^* \) that the conditional probability distribution of the other players’ history profiles from period 2 to period \( t - 1 \) equals the belief specified by \( \Psi^k \), given his history from period 2 to period \( t - 1 \)). If no such \( k \in K \) exists, player \( i \) believes that the other players play \( \hat{\sigma}_k \) from period 2. Hence conforming to \( \hat{\sigma}_k \) is optimal.

(b) If player \( i \) chose \( m_i = 0 \) in period 1, let \( a_i(1) \) be his own action in period 1. Again, player \( i \) believes that no other player deviated in his monitoring decision in period 1. Therefore if \( a_i(1) \neq \hat{a}_i \), the same argument as in the case (a) proves that conforming to \( \hat{\sigma}_i \) is optimal. If \( a_i(1) = \hat{a}_i \), again the same argument shows that conforming to \( \hat{\sigma}_0 \) is optimal.

Finally, consider the history of player \( i \in K \) at period 1. If player \( i \) follows \( \sigma^*_i \), his payoff in \( G(T) \) is

\[
\frac{1}{T} u_i(\hat{\sigma}) + \frac{T-1}{T} u_i(\alpha^*).
\]  

(5)

If he plays \( a_i \neq \hat{a}_i \) in period 1, then his stage payoff increases at most by \( \Delta \). From the next period on, \( \hat{\sigma}^i \) is played, and his (average) continuation payoff is less than \( u_i(\alpha^*) - \varepsilon \). Therefore, his overall payoff is at most

\[
\frac{1}{T} \left( u_i(\hat{\sigma}) + \Delta \right) + \frac{T-1}{T} \left( u_i(\alpha^*) - \varepsilon \right).
\]  

(6)

By \( T \geq T_0 + 1 \) and (4), we conclude that (5) is greater than (6). Hence \( \sigma^*_i \) is optimal.

Hence \( \sigma^* \) is a sequential equilibrium whose action profile in period 1 is \( \hat{\sigma} \neq \alpha^* \), and the proof is complete. \( \square \)

Appendix D  Proof of Proposition 5

Fix \( \eta = (\eta_i)_{i \in N} \) and \( \alpha \in A^*(\eta) \setminus \{\alpha^*\} \). Let \( K \) be the set of players \( i \) such that \( \alpha_i \) is not a best response against \( \alpha_{-i} \). For each \( i \in K \), let \( \alpha^i = (\alpha^i_j)_{j \in N_{-i}} \) be a Nash equilibrium of the \( \eta^i \)-constrained game, satisfying (1).

This time, let \( \sigma^* \) be the following strategy profile of \( G(2) \). In period 1, each player \( i \) chooses \( (\alpha_i, \mu_i) \) with probability \( 1 - \eta_i \), and \( (\alpha_i, \mu_i^0) \) with probability \( \eta_i \).

In period 2, each player \( i \) chooses his stage action in the following way.

(i) If player \( i \) did not observe the other players in period 1, then

(a) he chooses \( (\alpha_i^*, \mu_i^*) \) if \( i \notin K \),

(b) he chooses \( (\alpha_i^*, \mu_i^*) \) if \( i \in K \) and if he played \( \alpha_i \) (in this case, \( \alpha_i \) is pure) in period 1, and

(c) he chooses \( (\alpha_i^*, \mu_i^*) \) if \( i \in K \) and if he did not play \( \alpha_i \) (note that \( i \in N_{\eta^i} \), and therefore \( \alpha_i^* \) is well-defined).
(ii) If player $i$ observed the other players in period 1, let $a(1) \in A$ be the combination of his own action and his observations.

(a) If $k \in K$ exists (it may be that $k = i$) such that $a_k(1) \neq \alpha_k$ (note that $\alpha_k$ is pure since $k \in K$) and $a_j(1)$ is in the support of $\alpha_j$ for any $j \in N \setminus \{k\}$, we have two cases to consider.

- If $\eta_i < 1$ or if $i = k$, player $i$ chooses $(\alpha^*_i, \mu^*_i)$ (in either case, we have $i \in N_{\eta^k}$ and therefore $\alpha^*_k$ is well-defined).
- If $\eta_i = 1$ and if $i \neq k$, player $i$ plays a best response against the following mixed action profile

$$\left(\alpha^*_i, \left(\eta^k_j \alpha^*_j + (1 - \eta^k_j)\alpha^k_j\right)_{j \in N_{\eta^k}}\right),$$

and then always monitors the other players.

(b) If no such $k \in K$ exists, player $i$ chooses $(\alpha^*_i, \mu^*_i)$.

Note that under $\sigma^*$, $\alpha$ is played in the first period and then $\alpha^*$ in the second.

Let $\Psi^*$ be a consistent system of beliefs given $\sigma^*$, such that each player believes at any history at period 2 that if some player $j$ selected an action not in the support of $\alpha_j$, player $j$ observed the other players with probability $1$.\(13\) $\Psi^*$ is obtained from a tremble such that given that player $i$ played an action not in the support of $\alpha_i$ in period 1, he is much more likely to monitor the other players in period 1 than not to do so.

In order to show that $\sigma^*$ is a sequential equilibrium given $\Psi^*$, we first consider the histories at period 2. Suppose player $i$ did not monitor the other players in period 1. Then player $i$ believes that no other player deviated in period 1.

- If (a) $i \notin K$ or (b) $i \in K$ and he played $\alpha_i$ in period 1, then he believes that any other player observed no deviation by a player in $K$. Hence he believes that the other players play $\alpha^*_{-i}$, and $(\alpha^*_i, \mu^*_i)$ is optimal.

- If $i \in K$ and if he did not play $\alpha_i$, he believes that the only deviation observed by any other player is his own. Hence he believes that the other players play

$$\left(\alpha^*_{-N_{\eta^k}\setminus \{i\}}, \left(\eta^k_j \alpha^*_j + (1 - \eta^k_j)\alpha^k_j\right)_{j \in N_{\eta^k}\setminus \{i\}}\right).$$

Since $\eta_j = \eta^k_j$ for any $j \neq i$, $(\alpha^*_i, \mu^*_i)$ is optimal by the definition of $\alpha^i$.

Next, suppose player $i$ observed the other players in period 1. Let $a(1) \in A$ be the combination of his own action and his observations.

- Suppose $k \in K$ exists (it may be that $k = i$) such that $a_k(1) \neq \alpha_k$ and $a_j(1)$ is in the support of $\alpha_j$ for any $j \in N \setminus \{k\}$. Then he believes that any other player $j \neq i$ observed the other players in period 1 with probability $1 - \eta^k_j$.

\(13\) We specify the system of beliefs differently from the one considered in Subsection 4.2. Here we employ this specification, because it makes easier to specify actions when multiple deviations were observed.
He also believes that the other players’ monitoring decisions in period 1 were independent. If \( \eta_i < 1 \) or if \( k = i \), then his belief about the action profile in period 2 is given by

\[
\left( \alpha^*_{N_{k}}, (\eta^k_j \alpha^*_j + (1 - \eta^k_j)\alpha^k_j)_{j \in N_{k} \setminus \{i\}} \right).
\]

Thus \((\alpha^k_i, \mu^1_i)\) is optimal by the definition of \(\alpha^k\). If \(\eta_i = 1\) and if \(i \neq k\), then his belief about the action profile in period 2 is given by (7). Thus the specified stage action is optimal by definition.

- Suppose no such \(k \in K\) exists. Then we have either (a) \(a_j(1) = a_j\) for any \(j \in K\), or (b) there exist two or more \(k\)'s such that \(a_k(1)\) is not in the support of \(\alpha_k\). In either case, he believes that any other player \(j \neq i\) plays \(\alpha^*_j\). Indeed, in case (b), player \(i\) believes that all players \(k\) such that \(a_k(1)\) is not in the support of \(\alpha_k\) observed the other players and therefore play \(\alpha^*\). Hence \((\alpha^*_i, \mu^1_i)\) is optimal.

Finally, we consider the history at period 1.

- For player \(i \notin K\), he plays his static best-response in both periods. Since his stage action does not affect the future play at all, his strategy is optimal.

- For player \(i \in K\), if he chooses \(\alpha_i\) in period 1, his monitoring decision of that period does not affect his future payoff. If he chooses an action other than \(\alpha_i\), it gives him at most \(\max_{a' \in A_i} u_i(a'_i, \alpha_{-i})\) in this period. The subsequent play in period 2 gives him \(u_i(\alpha^*)\). Hence by (1), his strategy is optimal.

Therefore, \(\sigma^*\) is a sequential equilibrium, and the proof is complete. \(\square\)

**Appendix E  A Further Example of Nontrivial Equilibria**

Suppose \(G\) has \(n = 3\), \(A_i = \{C, D\}\) for any \(i\), and the payoff matrices as follows.

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>3,3</td>
<td>0,4</td>
</tr>
<tr>
<td>D</td>
<td>4,0</td>
<td>1,1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>1,1</td>
<td>2,0</td>
</tr>
<tr>
<td>D</td>
<td>0,2</td>
<td>3,3</td>
</tr>
</tbody>
</table>

The same line of argument for the example in Subsection 4.1 proves that \((D, D, D)\) is a unique Nash equilibrium of \(G\). The distinct Nash payoff condition is satisfied for (and only for) \(K = \{1, 2\}\), so that Proposition 3 implies existence of a nontrivial equilibrium. However, \(G\) does not satisfy the nondegeneracy condition. This is because for any action profile \((a_1, a_2) \in A_K\), player 3’s best reply gives him the same payoff of 3. Thus Proposition 4 does not apply to this game.

Nevertheless, \(G(3)\) has a nontrivial sequential equilibrium in which player 3’s payoff is not 3. To see that, note first that by Proposition 3, \(G(2)\) has two nontrivial sequential equilibria which reward either player 1 or 2. The sequential equilibrium rewarding player 1 (player 2, respectively) has a path where \((D, C, C)\) ((\(C, D, C)\), respectively) is
played in the first period and then \((D, D, D)\) in the second. Let us consider a strategy profile of \(G(3)\), which satisfies the followings.

- In the first period, player 1 chooses \((D, \mu^1_1)\), player 2 chooses \((\frac{2}{5}C + \frac{3}{5}D, \mu^1_2)\), and player 3 chooses \((\frac{3}{5}C + \frac{1}{5}D, \mu^1_3)\).

- Suppose player \(i\) observed (or played, in case \(i = 3\)) \(a_3 = C\) in period 1. Then from period 2 on, player \(i\) chooses \((D, \mu^1_i)\) in all subsequent periods.

- Suppose player \(i\) observed (or played, in case \(i = 3\)) \(a_3 = D\) in period 1. Then if he observed (or played, in case \(i = 2\)) \(a_2 = C\) (\(a_2 = D\), respectively), player \(i\)’s continuation strategy from period 2 is the sequential equilibrium strategy of \(G(2)\) which rewards player 2 (player 1, respectively).

We claim that the profile forms a sequential equilibrium of \(G(3)\), if the behavior at the histories where a player deviated in his monitoring decision in period 1 is suitably specified. The prescribed actions of player 1 and player 3 in period 1 are their static best replies. Those players are willing to follow the strategies, because (i) player 1’s action does not affect future play, and (ii) player 3’s action does not affect his continuation payoff. As for player 2, if he selects \(D\) instead of \(C\) in the first period, he gains by 1 in that period. However, if player 3 chooses \(D\) (which occurs with probability 1/4), the continuation equilibrium rewards player 1 rather than player 2. The associated loss is the difference of his payoff from \((D, C, C)\) and that from \((C, D, C)\), which is 4. Therefore player 2 is indifferent between \(C\) and \(D\) in the first period. This proves that the profile is a sequential equilibrium.

The payoff of player 3 under this equilibrium is \(((\frac{9}{5}) + 3 + 3)/3 = 13/5\). Since \(G(3)\) has distinct equilibrium payoffs for each player, a folk theorem holds for this stage game if the horizon is long enough. While the nontrivial equilibrium we derived here is based on the idea of monitoring by a subset of players, the conditions in Section 5 are not satisfied.\textsuperscript{14}

\textbf{References}


\textsuperscript{14}The argument extends to the case where the players slightly discount future payoffs. If their discount factor is \(\delta < 1\) and \(\delta\) is large enough, simply modify the above strategy profile so that the probability with which player 3 chooses \((C, \mu^1_3)\) in the first period is \(1 - (1/4\delta)\). It is easy to prove that the modified profile is a sequential equilibrium of \(G(3)\) if \(\delta\) is large.


