# Finitely Repeated Games with Automatic and Optional Monitoring⋆

Yasuyuki Miyahara[a] *
Kobe University

Tadashi Sekiguchi[b]
Kyoto University

April 15, 2016

**Abstract**

We extend a model of finitely repeated games with optional monitoring by our earlier paper, so that each player automatically receives complete information about the other players' actions with some exogenously given probability. Only when the automatic information did not arrive, the player privately decides whether to exercise a costless monitoring option or not. We show that a weak decrease in the vector of the players' probabilities of automatic monitoring is a necessary and sufficient condition for any repeated game with automatic and optional monitoring to have a weakly greater sequential equilibrium payoff vector set. This result considerably extends our earlier result, which only compares purely automatic monitoring and purely optional monitoring. We also verify that uniqueness of the stage game equilibrium is consistent with validity of a folk theorem under any automatic and optional monitoring structure.

*JEL Classification:* C72; C73
*Keywords:* Repeated games; Monitoring option

## 1. Introduction

This paper studies a class of repeated games where monitoring is part of the players' decision making. Particularly, we extend a model of repeated games with optional monitoring by our earlier paper (Miyahara and Sekiguchi [9]), where each player can costlessly decide whether to monitor the other players' actions or not. This paper introduces a possibility that each player may automatically learn the others' actions with some exogenously given probability, and the players' monitoring decisions are relevant only when the automatic information did not arrive.

⋆We thank Masatoshi Tsumagari for a stimulating conversation which originated this research. We are grateful to audience at the MEA meeting, Hitotsubashi University, Decentralization Conference in Japan, UECE Lisbon Meetings, and SING11-GTM2015. We gratefully acknowledge financial supports from the Grants-in-Aid for Scientific Research (Miyahara, 24530199; Sekiguchi, 23000001, 26245024, 26380238).

[a] *Graduate School of Business Administration, Kobe University, 2-1 Rokkodai-cho, Nada-ku, Kobe 657-8501, Japan*

*Corresponding author.

*E-mail addresses:* `miyahara@kobe-u.ac.jp` (Y. Miyahara), `sekiguchi@kier.kyoto-u.ac.jp` (T. Sekiguchi)

[b] *Institute of Economic Research, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan*

More concretely, we set up a model of finitely repeated games with the following structure. In each period, after the players have chosen their stage-game actions, each player automatically receives complete information about the other players' actions with some exogenous probability which does not depend on the actions. We assume that arrivals of the automatic information are independent across the players and over time. If the automatic information did not arrive, the player privately decides whether to exercise his monitoring option or not. If he exercises the option, he learns the others' actions without any noise. Otherwise, he does not get any information about them.[1] It is costless to monitor the other players, and the monitoring decision is completely unobservable. Namely, each player receives no signal as to whether any other player learnt his action either by automatic monitoring or his own monitoring. [9] corresponds to the case of purely optional monitoring, where the automatic information never arrives.

We have three results. The first two results examine how the sequential equilibrium payoff vector set of the *repeated games with automatic and optional monitoring* depends on the vector of the players' automatic monitoring probabilities, which we simply call the *monitoring vector*. We first show that for any stage game, any number of repetition, and any discount factor, the sequential equilibrium payoff vector set of the corresponding repeated game weakly decreases (in the sense of set inclusion) if the monitoring vector weakly increases. In other words, any sequential equilibrium payoff vector under a given monitoring vector continues to be a sequential equilibrium payoff vector if the monitoring vector decreases to any direction.

Second, and more interestingly, we show that a converse of the above result is also true. Fix two $n$-dimensional monitoring vectors $\lambda$ and $\lambda'$ such that $\lambda \geq \lambda'$ does *not* hold. Then, an $n$-player game exists such that the two-period repeated game under $\lambda'$ has a strictly smaller sequential equilibrium payoff vector set than the two-period repeated game under $\lambda$ if the players are sufficiently patient. Together with the first result, we see that a weak decrease in the monitoring vector is a necessary and sufficient condition for any repeated game with automatic and optional monitoring to have a weakly greater sequential equilibrium payoff vector set. Those results are a considerable extension of [9], who just compare the standard model of purely automatic monitoring and the case of purely optional monitoring.

Our third result is as to when a folk theorem (like Benoît and Krishna [1], Smith [10], and Gossner [4]) holds under *any* monitoring vector. It is easy to verify that our first result extends when we compare purely automatic monitoring and our automatic and optional monitoring. Therefore, if an existing folk theorem holds for some stage game under purely automatic monitoring, the same theorem holds for the same stage game under any monitoring vector. Our third result proves that its converse is not true. Namely, we show that for any $n$, an $n$-player stage game exists such that (i) it has a unique Nash equilibrium, (ii) the full dimensionality condition is satisfied, and (iii) for any monitoring vector, there exists $T$ such that the undiscounted $T$-period repeated game with that monitoring vector has a sequential equilibrium which gives each player a different average payoff from the stage-game equilibrium. The second and third conditions allow us to apply existing folk theorems for any monitoring vector. However, the first condition denies any folk theorem under purely automatic monitoring. Hence, this result highlights a fundamental difference

[1]This assumption denies a possibility that the player learns the others' actions from his stage payoff. To this end, we will assume that the players collect their stage payoffs in total, at the end of the repeated game.

of our model from the standard model.

One interpretation of this formulation of automatic and optional monitoring is players' overlooking. Namely, precise information about their actions is released (for example, sent by emails), and an absentminded player overlooks it. However, in case he overlooked the information, he has an opportunity to retrieve it (for example, by checking his email box). In this interpretation, the probability of automatic monitoring is the probability that the player is not absentminded.

Another interpretation is failure in avoiding information. As the above argument suggests, the players sometimes benefit from having a smaller monitoring probability, because it may create a more efficient equilibrium. Then they may want to install a device which prevents arrivals of the automatic information.[2] If the device is subject to random malfunction, however, then the probability of malfunction corresponds to the probability of automatic monitoring.

A key insight into our results is that since optional monitoring is costless and completely unobservable to the others, a player never hurts from having a smaller probability of automatic monitoring. The own monitoring simply compensates for lack of automatic information, because it is impossible to detect it and to inflict punishment on the monitor. Rather, the smaller probability of automatic monitoring expands their strategic flexibility. This is already pointed out by [9], but it only considers an extreme case where all players' probability of automatic monitoring changes from one to zero. Our contribution is to reveal that even a change that decreases only one player's automatic monitoring probability in a slightest way is, in some cases, sufficient to create new strategic possibilities.

Our second and third results are built on an example with a unique stage-game equilibrium. Namely, this is an example of a finitely repeated game with a nontrivial equilibrium despite that the stage game has a unique equilibrium, which never arises under the standard model.[3]

It is worthwhile to point out that the nontrivial equilibrium has a feature that a player randomizes over his monitoring decisions in case the automatic information did not arrive. Since the monitoring decision is private, a potential deviator does not know whether his deviation will be detected or not. The player who did not observe the others cannot respond to the deviation, and sequential rationality only limits behavior of the deviator and the other players knowing the deviation. This limited requirement of sequential rationality, caused by lack of common knowledge of the deviation, creates a credible punishment which works even if the stage game has a unique equilibrium. Note that a greater probability of automatic monitoring is burden for this construction, because it weakens the impact of the players who do not notice the deviation.

A main body of literature on repeated games with endogenous monitoring rather assumes that it is costly to monitor the others (for instance, Ben-Porath and Kahneman [2], Kandori and Obara [6], Miyagawa et al. [8], and Flesch and Perea [3]). In those papers, a central question is provision of incentives to monitor. Clearly, adding a possibility of automatic monitoring to those frameworks has quite different effects from ours, because the automatic information mitigates the incentive problem about costly monitoring. It would

[2] At the same time, the device must maintain their ability to monitor the others if they wish.

[3] This is not a novel feature, however, and [9] already highlights it. Further, other repeated games with imperfect monitoring, such as Kandori [5] and Mailath et al. [7], exhibit this type of results. See also a recent paper by Sugaya and Wolitzky [11].

be interesting to investigate those effects more thoroughly.

The rest of this paper is organized as follows. Section 2 introduces the model. Section 3 proves that a weak decrease in the monitoring vector is a necessary and sufficient condition for any repeated game with automatic and optional monitoring to have a weakly greater sequential equilibrium payoff vector set. Section 4 verifies that uniqueness of the stage Nash equilibrium is consistent with a folk theorem under any monitoring vector. We leave some proofs to the appendices.

## 2. Model

Let $G$ be a finite, $n$-player strategic form game with $n \geq 2$, and let $A_i$ be the set of player $i$'s pure actions. Player $i$'s payoff function is given by $u_i : A \to \mathbb{R}$, where $A \equiv A_1 \times \cdots \times A_n$ denotes the set of pure action profiles. Each player can choose a mixed action, and $\Delta A_i$ denotes the set of player $i$'s mixed actions.

After the players have chosen their actions, each player $i$ automatically receives complete information about the other players' actions with probability $\lambda_i \in (0, 1)$. We call $\lambda_i$ player $i$'s *automatic monitoring probability*. We assume that each $\lambda_i$ does not depend on the actions and that the arrivals of automatic information are independent across the players. Let $\lambda = (\lambda_1, \dots, \lambda_n)$ be the vector of automatic monitoring probabilities, called the *monitoring vector* hereafter. If the automatic information did not arrive, each player privately decides whether to monitor his opponents or not.[4] Monitoring the others is costless.

Fix $\delta \in (0, 1]$ and an integer $T \geq 2$. Let $G^{\delta}(\lambda, T)$ be the $\delta$-discounted $T$-period repeated game with automatic and optional monitoring where $G$ is played in periods $t = 1, \dots, T$. In each period $t \geq 1$, each player $i$ chooses an action $a_i \in A_i$ simultaneously. Then, player $i$ decides whether to monitor all the other players depending on his choice of action, in case the automatic information did not arrive. Each player can choose randomly whether to monitor them or not. We assume that the monitoring decision is not observable to the other players. Therefore, our model belongs to one with private monitoring. [9] corresponds to the case of $\lambda = (0, \dots, 0)$.

We also assume that if a player does not monitor the other players, then he receives no information about their actions. This amounts to assuming that the players receive all stage payoffs at the end of the repeated game, because the stage payoffs in general provide information about the actions. The players can monitor the actions in a period only at the end of that period; it is impossible to acquire information of any past period.

The information player $i$ obtains in each period $t$ is his action and information about the other players' actions in that period. We define $I_i = [A \times \{0, 1\}] \cup A_i$ as the set of information player $i$ obtains in one period. Here, (i) $(a, 0) \in I_i$ means that player $i$ chose $a_i$ and then automatically learnt that the other players' actions were $a_{-i}$, (ii) $(a, 1) \in I_i$ means that player $i$ chose $a_i$, did not receive the automatic information, and then found the other players playing $a_{-i}$ by his own monitoring, and (iii) $a_i \in I_i$ means that player $i$ chose $a_i$, did not receive the automatic information, and did not monitor the other players.

Player $i$'s *history* at period $t \geq 2$ consists of all his past information he obtains up to the end of period $t-1$. For $t \geq 2$, the set of all histories for player $i$ at period $t$ is $H_i^t = (I_i)^{t-1}$.

[4] We assume that the monitoring decision is binary in order to simplify the notations. It is easy to extend the analysis to the case where players can monitor any subset of the players.

Let $H_i^1$ be an arbitrary singleton set. The set of player $i$'s histories at all periods is

$$H_i = \bigcup_{t=1}^{T} H_i^t.$$

A strategy of player $i$ is denoted by $\sigma_i = (\sigma_i^a, \sigma_i^m)$. Here, $\sigma_i^a$ prescribes player $i$'s mixed action at each history at each period, namely, $\sigma_i^a : H_i \to \Delta A_i$. Then $\sigma_i^m$ prescribes a probability that player $i$ monitors all the other players when automatic information did not arrive, given any history at any period and any stage-game action chosen in that period, namely, $\sigma_i^m : H_i \times A_i \to [0,1]$. We call an element of $H_i \times A_i$ a *history-action pair*.

Given a strategy profile $\sigma = (\sigma_1, \dots, \sigma_n)$, player $i$'s payoff in $G^\delta(\lambda, T)$ is

$$\frac{1-\delta}{1-\delta^T} E\left[\sum_{t=1}^{T} \delta^{t-1} u_i\big(a(t)\big)\right],$$

where $a(t)$ is the action profile in period $t$, and the expectation is taken with respect to $\sigma$ and arrivals of automatic information.

The solution concept of this paper is sequential equilibrium adapted to our finitely repeated game. A *system of beliefs* is a function which maps each history $h_i^t$ to a probability distribution of the other players' history profiles $(h_j^t)_{j\neq i}$ and maps each history-action pair $(h_i^t, a_i)$ to a probability distribution of the other players' history-action pair profiles $(h_j^t, a_j)_{j\neq i}$. A strategy profile is *completely mixed* if at any history, any stage-game action is selected with positive probability and at any history-action pair, any monitoring decision is selected with positive probability. Given a strategy profile $\sigma$, a system of beliefs 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 system of beliefs, obtained from Bayes' rule, converges to it.[5] A strategy profile $\sigma$ is a *sequential equilibrium*, if there exists a consistent system of beliefs $\psi$ such that (a) at any history of any player $i$, his continuation strategy is optimal given $\sigma_{-i}$ and the belief about the other players' histories specified by $\psi$, and (b) at any history-action pair of any player $i$, his continuation strategy is optimal given $\sigma_{-i}$ and the belief about the other players' history-action pairs specified by $\psi$.

## 3. Monotonicity in the Monitoring Vector

This section proves the following result.

**Theorem 1.** *Fix $n \geq 2$ and two $n$-dimensional monitoring vectors $\overline{\lambda}$ and $\underline{\lambda}$. Then the following two conditions are equivalent.*

*(A) $\overline{\lambda} \geq \underline{\lambda}$.*

*(B) For any $n$-player stage game $G$, any $\delta \in (0,1]$, and any integer $T \geq 2$, any sequential equilibrium payoff vector of $G^\delta(\overline{\lambda}, T)$ is a sequential equilibrium payoff vector of $G^\delta(\underline{\lambda}, T)$.*

[5]For a consistent system of beliefs and for any $h_i^t$ and $a_i$, there is a close connection between the beliefs at $h_i^t$ and at $(h_i^t, a_i)$. Namely, the probability of any history-action pair profile $(h_j^t, a_j)_{j\neq i}$ at $(h_i^t, a_i)$ is the product of the probability of $(h_j^t)_{j\neq i}$ at $h_i^t$ and the probability of $(a_j)_{j\neq i}$ under the other players' mixed actions at $(h_j^t)_{j\neq i}$.

Theorem 1 implies that for any $G$, any $T$, and any $\delta$, the set of sequential equilibrium payoff vectors of $G^\delta(\lambda, T)$ is weakly decreasing in the monitoring vector, in the sense of set inclusion. The theorem also implies that for any change in the monitoring vector other than a weak decrease, a discounted finitely repeated game exists such that a payoff vector attained by a sequential equilibrium before the change cannot be sustained by a sequential equilibrium after the change. To summarize, a weak decrease in the monitoring vector is a necessary and sufficient condition for any repeated game with automatic and optional monitoring to have a weakly greater sequential equilibrium payoff vector set.

Appendix A proves that (A) implies (B). The intuition behind this part is simple. Since monitoring is costless and private, the players can always compensate a smaller probability of automatic monitoring by increasing the probability of his own monitoring, without affecting their payoffs and beliefs. Therefore, we can modify any sequential equilibrium of $G^\delta(\overline{\lambda}, T)$, so that the modified strategy profile of $G^\delta(\underline{\lambda}, T)$ has exactly the same play as the original equilibrium.

The rest of this section verifies that (B) implies (A). Fix $n \geq 2$ and two $n$-dimensional monitoring vectors, $\underline{\lambda}$ and $\overline{\lambda}$, such that $\underline{\lambda}_1 < \overline{\lambda}_1$. Note that, without loss of generality, this is a denial of $\underline{\lambda} \geq \overline{\lambda}$. It suffices to verify existence of an $n$-player stage game $G$, $\delta \in (0, 1]$, and $T \geq 2$ such that $G^\delta(\overline{\lambda}, T)$ has a strictly smaller sequential equilibrium payoff vector set than $G^\delta(\underline{\lambda}, T)$.[6] The next result, however, is somewhat stronger in the following senses. First, we can set $T = 2$, so that a very short horizon is sufficient to create a difference. Second, we can arbitrarily fix $\delta$, so that patience is not indispensable for creating a difference. For that matter, the same stage game which works for the fixed $\delta$ also works for any greater discount factor. Third, the choice of $G$ depends only on $\underline{\lambda}_1$ and $\overline{\lambda}_1$, so that the values of $\underline{\lambda}_i$ and $\overline{\lambda}_i$ for any $i \geq 2$ are irrelevant. For that matter, for any $\mu \in [\underline{\lambda}_1, \overline{\lambda}_1)$, we can choose $G$ so that it works not only for $\underline{\lambda}$ and $\overline{\lambda}$, but also for any two monitoring vectors $\underline{\lambda}'$ and $\overline{\lambda}'$ such that $\underline{\lambda}'_1 \leq \mu < \overline{\lambda}'_1$.

**Proposition 1.** *Fix $n \geq 2$, $\mu \in (0, 1)$, and $\underline{\delta} \in (0, 1]$. Then, an $n$-player strategic form game $G$ exists such that for any $\overline{\lambda}$ and $\underline{\lambda}$ such that $\underline{\lambda}_1 \leq \mu < \overline{\lambda}_1$, $G^\delta(\overline{\lambda}, 2)$ has a unique sequential equilibrium payoff vector for any $\delta$, and $G^\delta(\underline{\lambda}, 2)$ with any $\delta \geq \underline{\delta}$ has a sequential equilibrium whose payoff vector is different from the sequential equilibrium payoff vector of $G^\delta(\overline{\lambda}, 2)$.*

**Proof.** Fix $n \geq 2$, $\mu \in (0, 1)$ and $\underline{\delta} \in (0, 1]$. Define $G$ so that $A_1 = \{U, M, D\}$, $A_2 = \{L, C, R\}$, and $A_i = \{b_i, c_i\}$ for any $i \geq 3$ (if any). $u_1$ and $u_2$ depend only on $(a_1, a_2)$, and are represented by the following payoff matrix. For $i \geq 3$,

| | $L$ | $C$ | $R$ |
|---|---|---|---|
| $U$ | $4, 4$ | $0, 3$ | $4, 3 - \frac{\mu}{1-\mu}\underline{\delta}$ |
| $M$ | $0, 3 - \frac{1-\mu}{\mu}$ | $6, 3$ | $0, 3 + \underline{\delta}$ |
| $D$ | $3, 6$ | $3, 0$ | $3, 6$ |

Figure 1: Payoff matrix for players 1 and 2

[6] Here we reverse the roles of $\overline{\lambda}$ and $\underline{\lambda}$ in the statement of Theorem 1.

$$u_i(a) = \begin{cases} 3 & \text{if } (a_1, a_2) = (M, C) \text{ and } a_i = b_i, \\ 2 & \text{if } (a_1, a_2) \neq (M, C) \text{ and } a_i = b_i, \\ 0 & \text{if } a_i = c_i. \end{cases} \tag{1}$$

First, we show that $G$ has a unique Nash equilibrium $(U, L, b_3, \dots, b_n)$. It is sufficient to show that $C$ is not played with positive probability in any Nash equilibrium, because we have a unique outcome $(U, L, b_3, \dots, b_n)$ by iterated elimination of strictly dominated actions in the reduced game obtained after eliminating $C$ from the set of actions of player 2.

Suppose that player 1 plays $U$ with probability $x$, $M$ with probability $y$, and $D$ with probability $1 - x - y$ in a Nash equilibrium of $G$. A necessary and sufficient condition for $C$ to be a best response of player 2 is

$$\frac{(1-\mu)y}{\mu} + \frac{6(1-\mu)}{\underline{\delta}\mu}(1 - x - y) \leq x \leq \frac{(1-\mu)y}{\mu} - 6(1 - x - y).$$

This condition implies $1 - x - y = 0$ and $x = (1 - \mu)y/\mu$, which in turn imply $x = 1 - \mu$ and $y = \mu$. Since $x > 0$ and $y > 0$, player 1 must be indifferent between $U$ and $M$, and that holds only if player 2 plays $C$ with probability $2/5$. Then, $D$ is a unique best response of player 1. Therefore, $C$ is not played with positive probability in any Nash equilibrium. Hence, $(U, L, b_3, \dots, b_n)$ is a unique Nash equilibrium of $G$.

Fix $\overline{\lambda}$ and $\underline{\lambda}$ such that $\underline{\lambda}_1 \leq \mu < \overline{\lambda}_1$. In what follows, we show that $G^\delta(\underline{\lambda}, 2)$ with any $\delta \geq \underline{\delta}$ has a sequential equilibrium payoff vector which is different from the unique Nash equilibrium payoff vector of $G$, and $G^\delta(\overline{\lambda}, 2)$ with any $\delta \in (0, 1]$ has a unique sequential equilibrium payoff vector, which equals the unique Nash equilibrium payoff vector of $G$.

First, fix $\delta \geq \underline{\delta}$, and let $\hat{\sigma}$ be the following strategy profile of $G^\delta(\underline{\lambda}, 2)$.

- In period 1, player 1 plays $M$. Then if the automatic information did not arrive, he monitors the other players with probability $(\mu - \underline{\lambda}_1)/(1 - \underline{\lambda}_1)$, and does not monitor them with probability $(1 - \mu)/(1 - \underline{\lambda}_1)$ irrespective of his action. In period 2, player 1 plays $M$ and monitors the other players irrespective of his action if he found player 2 not playing $C$ in period 1. Otherwise, player 1 plays $U$ and monitors the other players irrespective of his action.

- In period 1, player 2 plays $C$, and monitors the other players irrespective of his action. In period 2, player 2 plays $L$ and monitors the other players regardless of his action if he played $C$ in period 1. Otherwise, player 2 plays $C$ and monitors the other players regardless of his action.

- Player $i \geq 3$ plays $b_i$ and monitors the other players regardless of his action at any history.

The play under $\hat{\sigma}$ is $(M, C, b_3, \dots, b_n)$ in period 1 and $(U, L, b_3, \dots, b_n)$ in period 2, and the average discounted payoff vector is

$$\left( \frac{6 + 4\delta}{1 + \delta}, \frac{3 + 4\delta}{1 + \delta}, \frac{3 + 2\delta}{1 + \delta}, \dots, \frac{3 + 2\delta}{1 + \delta} \right).$$

It is different from the Nash equilibrium payoff vector of the stage game. It remains to show that $\hat{\sigma}$ is a sequential equilibrium. Fix any consistent system of beliefs given $\hat{\sigma}$. First, we consider player 1.

- At any history of player 1 at period 2 such that player 1 found player 2 not playing $C$ in period 1, player 1 believes that player 2 plays $C$ with probability one in period 2. Hence, it is optimal for player 1 to play $M$ with any monitoring decision in period 2.

- At any history of player 1 at period 2 such that player 1 did not find a deviation of player 2 in period 1, player 1 believes that player 2 did not deviate and plays $L$ with probability one in period 2. Hence, it is optimal for player 1 to play $U$ with any monitoring decision in period 2.

- In period 1, player 1 plays a short-run best response, and his action does not affect future play. Further, he has nothing to learn from the others' actions, because their strategies are pure. Hence, it is optimal for player 1 to play $M$ and randomize between monitoring and not monitoring in the way specified by $\hat{\sigma}$ in period 1.

Next, we consider player 2.

- At any history of player 2 at period 2 such that he played $C$ in period 1, player 2 believes that player 1 plays $U$ with probability one. Hence, it is optimal for player 2 to play $L$ with any monitoring decision in period 2.

- At any history of player 2 at period 2 such that he did not play $C$ in period 1, player 2 believes that player 1 plays $U$ and $M$ with probabilities $1-\mu$ and $\mu$ respectively. This is because consistency requires that, whether player 2 found player 1's deviation or not, he believes that player 1 finds player 2's deviation (and therefore plays $M$ in period 2) with probability

$$\underline{\lambda}_1 + (1-\underline{\lambda}_1) \cdot \frac{\mu - \underline{\lambda}_1}{1-\underline{\lambda}_1} = \mu.$$

  Hence, it is optimal for player 2 to play $C$ with any monitoring decision in period 2, and his stage payoff is 3.

- When player 2 follows $\hat{\sigma}_2$, his average discounted payoff is $(3+4\delta)/(1+\delta)$. If he does not play $C$ in period 1, then his stage payoff is at most $3+\underline{\delta}$. We have seen that this deviation leads to a continuation play where his stage payoff is 3 in period 2. Therefore, his average discounted payoff when he does not play $C$ in period 1 is at most $(3+\underline{\delta}+3\delta)/(1+\delta)$. This implies that conforming to $\hat{\sigma}_2$ is optimal since $\delta \geq \underline{\delta}$.

Finally, for any player $i \geq 3$, his play does not affect future play at all. Hence, it is optimal to always play a static best response $b_i$, as is prescribed by $\hat{\sigma}_i$. Therefore, $\hat{\sigma}$ is a sequential equilibrium.

Next, fix $\delta \in (0,1]$. We show that $G^\delta(\overline{\lambda}, 2)$ has a unique sequential equilibrium payoff vector. Let $\sigma$ be a sequential equilibrium in $G^\delta(\overline{\lambda}, 2)$.

**Claim 1.** *For any $i \geq 3$, $\sigma_i$ prescribes to play $b_i$ at any history.*

At any history of player $i \geq 3$ at period 2, $b_i$ is uniquely optimal. At period 1, choosing $c_i$ instead of $b_i$ causes a loss of 2 or more in period 1, and the gain in period 2 is at most 1, because he can guarantee 2 by playing $b_i$. Therefore, playing $b_i$ is always optimal.

Let $\hat{A}_2 \subseteq A_2$ be the set of player 2's actions played with positive probability in period 1 under $\sigma$. Also, let $\hat{H}_2^2$ be the set of player 2's histories at period 2 such that he played $a_2 \in \hat{A}_2$, monitored the other players by either automatic or own monitoring, and found any player $i \geq 3$ (if any) playing $b_i$ in period 1. Note that player 2 may have observed a deviation of player 1 at $h_2^2 \in \hat{H}_2^2$.

**Claim 2.** *At any history $h_2^2 \in \hat{H}_2^2$, player 2 does not play $C$ with positive probability.*

Fix $h_2^2 \in \hat{H}_2^2$, and let $a_1$ and $a_2$ be the actions of players 1 and 2 in period 1 under $h_2^2$. Define $a = (a_1, a_2, b_3, \ldots, b_n)$. Then player 2 believes that player 1's history at period 2 is either $(a, 0)$, $(a, 1)$, or $a_1$.

Suppose, on the contrary to the claim, that $\sigma_2$ prescribes to play $C$ with positive probability at $h_2^2$. We have seen before that $C$ is optimal only if player 2 believes that player 1 plays $U$ with probability $1 - \mu$ and $M$ with probability $\mu$. For player 1, $U$ and $M$ are not simultaneously optimal at any history at period 2 (player 1 is indifferent between $U$ and $M$ only when player 2 chooses $C$ with probability 2/5, but then $D$ is uniquely optimal). Since $\overline{\lambda}_1 > 0$, player 2 believes that player 1 is at $h_1^2 = (a, 0)$ with positive probability, and also believes that player 1 at $h_1^2 = (a, 0)$ plays a pure action other than $D$, which we denote by $\hat{a}_1$.

Player 1 has the same belief about the other players' actions at $(a, 0)$ and $(a, 1)$. Thus, $\hat{a}_1$ is optimal at both $(a, 0)$ and $(a, 1)$. Consequently, in order to believe that both $U$ and $M$ are played with positive probability, player 2 must believe that player 1 is at $h_1^2 = a_1$ with positive probability and player 1 at $h_1^2 = a_1$ plays the pure action $\tilde{a}_1 \in \{U, M\} \setminus \{\hat{a}_1\}$.

It follows from consistency that player 1 who chose $a_1$ and did not receive automatic information in period 1 does not monitor the other players with positive probability. Hence, sequential rationality implies that player 1 who chose $a_1$ and did not receive automatic information finds it optimal not to monitor the others in period 1 and then play $\tilde{a}_1$ in period 2. Since $a_2 \in \hat{A}_2$, player 1 at the time of his monitoring decision believes that he reaches $h_1^2 = (a, 1)$ with positive probability if he monitors the others. As a result, $\tilde{a}_1$ must be optimal at $h_1^2 = (a, 1)$, which in turn implies that $\tilde{a}_1$ must be optimal at $h_1^2 = (a, 0)$. This is a contradiction, which establishes the claim.

**Claim 3.** *Fix $a_2 \in \hat{A}_2$, and suppose that $\sigma_2$ prescribes not to monitor the other players with positive probability if he played $a_2$ and did not receive automatic information in period 1. Then at the history $h_2^2 = a_2$, player 2 does not play $C$ with positive probability.*

On the contrary, suppose that $\sigma_2$ prescribes to play $C$ with positive probability at $h_2^2 = a_2$. From sequential rationality, it must be optimal not to monitor the others in period 1 and then to play $C$ in period 2, if he played $a_2$ and did not receive the automatic information in period 1. This implies that it must be also optimal to monitor the other players in period 1 and then to play always $C$ in period 2, if he played $a_2$ and did not receive the automatic information in period 1. However, player 2 believes that he will surely reach a history in $\hat{H}_2^2$, and we have seen in the proof of Claim 2 that $C$ is not optimal at any history in $\hat{H}_2^2$. Therefore, monitoring the others and then playing an optimal action for each realized history improve player 2's payoff. This is a contradiction.

**Claim 4.** *At any history of player 1 at period 2 such that he did not find any player's deviation in period 1, $\sigma_1$ prescribes $U$ with probability one.*

At any such history, consistency requires player 1 to believe that any player did not deviate in period 1. From Claims 2 and 3, player 2 never plays $C$ if he did not deviate in period 1 and did not find a deviation by player $i \geq 3$. When $C$ is not played with positive probability, $U$ is uniquely optimal.

**Claim 5.** *At any history of player 2 at period 2 such that he monitored the other players in period 1, player 2 does not play $C$ with positive probability.*

Let $a \in A$ be the combination of player 2's action and his observation in period 1, given the history. Then player 2 believes that player 1's history is either $(a, 1)$, $(a, 0)$, or $a_1$. From Claim 4, player 1 plays $U$ with probability one at the history $a_1$. Hence, for $C$ to be optimal, $\sigma_1$ must prescribe $M$ with positive probability at either $(a, 1)$ or $(a, 0)$. Since player 1's belief about the others' actions is the same at the two histories, sequential rationality requires that $M$ is optimal at both $(a, 1)$ and $(a, 0)$. Therefore, $\sigma_1$ does not prescribe $U$ with positive probability at the two histories, because $U$ and $M$ cannot be simultaneously optimal. Since player 1 receives the automatic information and reaches the history $(a, 0)$ with probability $\overline{\lambda}_1 > \mu$, the probability with which player 2 believes that player 1 plays $U$ is less than $1 - \mu$. Given the belief, $C$ is not optimal.

**Claim 6.** *Fix $a_2 \in A_2$, and suppose that $\sigma_2$ prescribes not to monitor the other players with positive probability if he played $a_2$ and did not receive automatic information in period 1. Then at the history $h_2^2 = a_2$, player 2 does not play $C$ with positive probability.*

On the contrary, suppose that $\sigma_2$ prescribes to play $C$ with positive probability at $h_2^2 = a_2$. From sequential rationality, it must be optimal not to monitor the others in period 1 and then to play $C$ in period 2, if he played $a_2$ and did not receive the automatic information in period 1. This implies that it must be also optimal to monitor the other players in period 1 and then to play always $C$ in period 2, if he played $a_2$ and did not receive the automatic information in period 1. However, we have seen in the proof of Claim 5 that $C$ is not optimal at any history at period 2 where he played $a_2$ and monitored the others in period 1. Therefore, monitoring the others and then playing an optimal action for each realized history at period 2 improve player 2's payoff. This is a contradiction.

**Claim 7.** *At any history of player 1 at period 2, $\sigma_1$ prescribes $U$ with probability one.*

Consistency requires player 1 to believe that player 2 conforms to his monitoring decision given his (possibly deviant) action. Hence, from Claims 5 and 6, player 1 believes that player 2 does not play $C$ with positive probability in period 2. Given the belief, $U$ is uniquely optimal.

**Claim 8.** *At any history of player 2 at period 2, $\sigma_2$ prescribes $L$ with probability one.*

By Claim 7, player 1 plays $U$ at any history at period 2. Thus, $L$ is uniquely optimal at any history at period 2.

From Claims 1 to 8, the players follow $(U, L, b_3, \ldots, b_n)$ at any history at period 2. Given that, play in the first period must form a Nash equilibrium of $G$. Therefore, the players play $(U, L, b_3, \ldots, b_n)$ in period 1. The payoff vector of this sequential equilibrium is $(4, 4, 2, \ldots, 2)$, which completes the proof. Q.E.D.

We have two remarks on our results. First, although we assume that each player's automatic monitoring probability is neither 0 nor 1, the results extend to the case where $\lambda_i = 0$ or $\lambda_i = 1$ for some $i$. This extension is straightforward, except for notational cumbersomeness. If $\lambda_i = 0$, player $i$ never receives automatic information. Thus, we should redefine $I_i$, the set of information player $i$ obtains in one period, as $I_i = A \cup A_i$. If $\lambda_i = 1$, the automatic information surely arrives. Thus, we should redefine $I_i$ as $I_i = A$. With these modifications, we can redefine each player's histories and strategies and study the repeated games accordingly. Then Theorem 1 continues to hold even if $\overline{\lambda}$ and $\underline{\lambda}$ contain 0 and/or 1. We omit a proof, because the intuition should be clear.

Second, we see from the proof of Proposition 1 that for any $n$-dimensional monitoring vector $\lambda$, an $n$-player stage game $G$ exists such that $G$ has a unique equilibrium but $G^\delta(\lambda, 2)$ has multiple sequential equilibrium payoff vectors if $\delta$ is sufficiently large. This further implies that for all large $\delta$, $G^\delta(\lambda, T)$ with any $T \geq 3$ has multiple sequential equilibrium payoff vectors.[7] In the standard model of purely automatic monitoring, if the stage game has a unique equilibrium payoff vector, any finitely repeated game has a unique subgame perfect equilibrium payoff vector. Therefore, our results imply that this well-known negative result does not extend to any automatic and optional monitoring environment.[8]

## 4. Purely Automatic Monitoring

In this section, we compare automatic and optional monitoring with purely automatic monitoring in view of the folk theorem (Benoît and Krishna [1], Smith [10], and Gossner [4]).

As the remark at the end of the previous section indicates, if an existing folk theorem applies to a given stage game under purely automatic monitoring, the same folk theorem applies to the same stage game under *any* monitoring vector. The following result, which is proved in Appendix B, shows that its converse does not hold.

**Theorem 2.** *For any $n \geq 2$, there exists an $n$-player strategic form game $G$ such that*

*(i) $G$ has a unique Nash equilibrium,*

*(ii) the full dimensionality condition is satisfied, and*

*(iii) for any monitoring vector $\lambda$, there exists $T$ such that $G^1(\lambda, T)$ has a sequential equilibrium which gives each player a different payoff from the Nash equilibrium of $G$.*

The conditions (ii) and (iii) allow us to apply any existing folk theorem under *any* monitoring vector.[9] In contrast, the condition (i) implies that no folk theorem holds under purely optional monitoring. Therefore, the sufficient condition for a folk theorem to hold under any automatic and optional monitoring structure is quite different from that in the standard model. Note that an arbitrarily small departure from purely automatic monitoring is sufficient for a folk theorem. Theorem 2 thus highlights a fundamental difference between the standard model and ours.

[7] For any $T \geq 3$, playing a static equilibrium in all periods $1, \ldots, T-2$ and then playing a nontrivial sequential equilibrium of $G^\delta(\lambda, 2)$ in the last two periods form a sequential equilibrium of $G^\delta(\lambda, T)$.

[8] [9] points out this failure of the negative result under purely optional monitoring.

[9] Note that the condition (iii) limits attention to no discounting. This restriction is not essential, and we impose it just because some folk theorems assume no discounting presumably for simplicity (e.g., Gossner [4]).

We have two more remarks. First, the stage game $G$ constructed in the proof of Proposition 1 satisfies the conditions in Theorem 2. Proposition 1 implies existence of $\lambda$ such that $G^\delta(\lambda, 2)$ has a unique sequential equilibrium payoff vector for any $\delta$. Further, Theorem 2 implies existence of $T$ such that $G^1(\lambda, T)$ has a nontrivial equilibrium. These results reveal another difference from the standard model. Under purely automatic monitoring, if $T$ exists such that the $T$-period repeated game has a unique subgame perfect equilibrium payoff vector under any discount factor, any finitely repeated game with any discount factor has a unique subgame perfect equilibrium payoff vector.

Second, Proposition 1 and Theorem 2 suggest that, for any two $n$-dimensional monitoring vectors $\overline{\lambda}$ and $\underline{\lambda}$, the difference between the sequential equilibrium payoff vector sets of $G^\delta(\overline{\lambda}, T)$ and $G^\delta(\underline{\lambda}, T)$ may disappear if $T$ and $\delta$ are large. Proposition 1 emphasizes the difference in the sequential equilibrium payoff vector sets of $G^\delta(\overline{\lambda}, 2)$ and $G^\delta(\underline{\lambda}, 2)$. However, Theorem 2 implies that existing folk theorems hold under both $\overline{\lambda}$ and $\underline{\lambda}$. It is an interesting topic for future research to examine whether a very long horizon and (nearly) complete patience nullify the effects caused by differences in monitoring vectors.[10]

## Appendix A. Proof of Theorem 1

Since Proposition 1 reveals that (B) implies (A), it suffices to verify that (A) implies (B).

Fix $n \geq 2$ and two $n$-dimensional monitoring vector $\overline{\lambda}$ and $\underline{\lambda}$ such that $\overline{\lambda} \geq \underline{\lambda}$. Also fix an $n$-player stage game $G$, $\delta \in (0, 1]$, and $T \geq 2$. Let $\overline{\sigma} \equiv (\overline{\sigma}_i)_{i=1}^n$ be a sequential equilibrium of $G^\delta(\overline{\lambda}, T)$, let $\overline{\psi}$ be a consistent system of beliefs supporting $\overline{\sigma}$, and let $(\overline{\sigma}^k)_{k=1}^\infty$ be a tremble of $\overline{\sigma}$ such that the sequence of the corresponding systems of beliefs converges to $\overline{\psi}$. For each $k$, we define a strategy profile of $G^\delta(\underline{\lambda}, T)$, $\underline{\sigma}^k \equiv (\underline{\sigma}_i^k)_{i=1}^n$, in the following way. For any $i$ such that $\underline{\lambda}_i = \overline{\lambda}_i$, we define $\underline{\sigma}_i^k = \overline{\sigma}_i^k$. For any $i$ such that $\underline{\lambda}_i < \overline{\lambda}_i$, suppose player $i$ plays $G^\delta(\underline{\lambda}, T)$ as follows.

- Player $i$ has a private randomization device, called *roulette*, which selects 0 with probability $x_i \equiv (\overline{\lambda}_i - \underline{\lambda}_i)/(1 - \underline{\lambda}_i)$ and selects 1 with probability $1 - x_i$. Note that $x_i \in (0, 1)$.

- In period 1, player $i$ plays the mixed action $\overline{\sigma}_i^k$ prescribes at $h_i^1$. If player $i$ did not obtain automatic information, he spins his roulette. If it selects 0, he monitors the other players with probability 1. If it selects 1, his monitoring decision is the one $\overline{\sigma}_i^k$ prescribes at $(h_i^1, a_i)$, where $a_i$ is his stage-game action in this period.

- Player $i$'s behavior in period $t \geq 2$ depends on his history and past realizations of his roulette (if any). Suppose $t \geq 2$ and he is at a history at period $t$, denoted by $h_i^t = (\omega_i^\tau)_{\tau=1}^{t-1}$, where $\omega_i^\tau \in I_i$ for any $\tau \leq t-1$. Define a new history $\overline{h}_i^t = (\overline{\omega}_i^\tau)_{\tau=1}^{t-1}$ so that for any $\tau \leq t - 1$,

  (i) if $\omega_i^\tau = (a, 1)$ and if his roulette selected 0 in period $\tau$, $\overline{\omega}_i^\tau = (a, 0)$, and

  (ii) otherwise, $\overline{\omega}_i^\tau = \omega_i^\tau$.

  Let us call $\overline{h}_i^t$ the *effective history*. Then, player $i$ plays the mixed action $\overline{\sigma}_i^k$ prescribes at $\overline{h}_i^t$. If player $i$ did not obtain automatic information, he spins his roulette. If

[10] This issue is absent in [9]. The difference between purely automatic monitoring and purely optional monitoring simply enlarges if $T$ gets larger.

it selects 0, he monitors the other players with probability 1. If it selects 1, his monitoring decision is the one $\overline{\sigma}_i^k$ prescribes at $(\overline{h}_i^t, a_i)$, where $a_i$ is his stage-game action in this period.

Let us define $\underline{\sigma}_i^k$ as the strategy of $G^\delta(\underline{\lambda}, T)$ which is equivalent to the above play. Note that $\underline{\sigma}_i^k$ is completely mixed because so is $\overline{\sigma}_i^k$.

Under $\underline{\sigma}^k$, whenever a player monitored the other players in a period where his roulette selected 0, he pretends that automatic information arrived in that period and accordingly follows $\overline{\sigma}^k$. In any period, player $i$ with $\underline{\lambda}_i < \overline{\lambda}_i$ either receives automatic information or pretends that he received it with probability $\underline{\lambda}_i + (1 - \underline{\lambda}_i)x_i = \overline{\lambda}_i$. Therefore, $\underline{\sigma}^k$ generates exactly the same action path as $\overline{\sigma}^k$. More precisely, if the players play $\underline{\sigma}^k$, then

(a) for any profile of histories at some period $t$, $\overline{h}^t \equiv (\overline{h}_i^t)_{i=1}^n$, the probability that their effective histories at period $t$ are $\overline{h}^t$ equals the probability that the play reaches to $\overline{h}^t$ when $\overline{\sigma}^k$ is played in $G^\delta(\overline{\lambda}, T)$, and

(b) for any profile of histories at some period $t$, $\overline{h}^t \equiv (\overline{h}_i^t)_{i=1}^n$, and any $a \in A$, the probability that their effective histories at period $t$ are $\overline{h}^t$ and their actions in period $t$ are $a$ equals the probability that the play reaches to $\overline{h}^t$ and $a$ is played in period $t$ when $\overline{\sigma}^k$ is played in $G^\delta(\overline{\lambda}, T)$.

Note that these equivalences imply that the payoffs of $\overline{\sigma}^k$ and $\underline{\sigma}^k$ coincide for any $k$.

For each $k$, let $\underline{\psi}^k$ be the system of beliefs derived from $\underline{\sigma}^k$ by Bayes' rule. Let $(\underline{\sigma}, \underline{\psi})$ be the limit of a convergent subsequence of $(\underline{\sigma}^k, \underline{\psi}^k)_{k=1}^\infty$. By definition, $\underline{\psi}$ is consistent given $\underline{\sigma}$. It suffices to prove that $\underline{\sigma}$ is a sequential equilibrium of $G^\delta(\underline{\lambda}, T)$, because $\overline{\sigma}^k \to \overline{\sigma}$, $\underline{\sigma}^k \to \underline{\sigma}$, and continuity imply that the payoffs of $\overline{\sigma}$ and $\underline{\sigma}$ coincide.

Fix a history of player $i$ or his history-action pair arbitrarily, which has the form of either $\iota_i^t = h_i^t$ or $\iota_i^t = (h_i^t, a_i)$. His continuation strategy under $\underline{\sigma}_i$ amounts to replacing $h_i^t$ with the effective history $\overline{h}_i^t$ and then playing the continuation strategy of $\overline{\sigma}_i$ at the new history, which has the form of either $\overline{\iota}_i^t = \overline{h}_i^t$ or $\overline{\iota}_i^t = (\overline{h}_i^t, a_i)$. Since $\overline{\sigma}$ is a sequential equilibrium of $G^\delta(\overline{\lambda}, T)$, the continuation strategy at $\overline{\iota}_i^t$ is optimal under the belief given by $\overline{\psi}$. Since $\underline{\sigma}$ in $G^\delta(\underline{\lambda}, T)$ effectively reproduces $\overline{\sigma}$ in $G^\delta(\overline{\lambda}, T)$, this implies that the continuation strategy of $\underline{\sigma}_i$ at $\iota_i^t$ when the effective history is $\overline{\iota}_i^t$ is optimal under the belief given by $\underline{\psi}$. Note that the beliefs under $\underline{\psi}$ do not depend on whether the history at period $t$ is $h_i^t$ or $\overline{h}_i^t$. This implies that the continuation strategy of $\underline{\sigma}_i$ at $\iota_i^t$ is optimal under the belief given by $\underline{\psi}$, independently of the effective history. This establishes sequential rationality, and the proof is complete. Q.E.D.

## Appendix B. Proof of Theorem 2

Fix $n \geq 2$, and we define $G$ as a special case of the stage game used in the proof of Proposition 1. Let $\mu \in (0, 1)$ be a parameter. The sets of each player's actions are $A_1 = \{U, M, D\}$, $A_2 = \{L, C, R\}$, and $A_i = \{b_i, c_i\}$ for any $i \geq 3$ (if any). As before, $u_1$ and $u_2$ depend only on $(a_1, a_2)$, and their payoff matrix is as follows.

| | $L$ | $C$ | $R$ |
|---|---|---|---|
| $U$ | $4, 4$ | $0, 3$ | $4, 3 - \frac{\mu}{1-\mu}$ |
| $M$ | $0, 3 - \frac{1-\mu}{\mu}$ | $6, 3$ | $0, 4$ |
| $D$ | $3, 6$ | $3, 0$ | $3, 6$ |

The payoff function of any player $i \geq 3$ is given by (1). As before, $G$ has a unique Nash equilibrium $(U, L)$. It is easy to show that the full dimensionality condition holds.

Now, we prove that, for any monitoring vector $\lambda$, $T \geq 2$ exists such that the repeated game $G^1(\lambda, T)$ has a sequential equilibrium payoff vector other than the Nash equilibrium payoff vector of $G$.

Fix $\lambda$, and choose $T \geq 2$ so that

$$\mu \geq \lambda_1^{T-1}. \tag{2}$$

Let $\sigma$ be the following strategy profile of $G^1(\lambda, T)$. For $i \geq 3$, player $i$ plays $b_i$ at any history at any period and observes the other players at any history-action pair at any period. Player 1 plays $\sigma_1$ as follows:

- At any history $h_1^t$ such that $t \leq T - 2$, player 1 plays $U$, and if the automatic information did not arrive, he monitors the other players with probability $\nu$, where

$$\nu = \frac{\mu^{\frac{1}{T-1}} - \lambda_1}{1 - \lambda_1}, \tag{3}$$

  irrespective of his action in period $t$ (note that we have $0 \leq \nu < 1$ by (2) and $\mu < 1$).

- In period $T - 1$, player 1 plays $U$ if he found that player 2 played an action other than $L$ in some period before period $T - 1$, and he plays $M$ otherwise. Whichever action player 1 plays in the period, he observes the other players with probability $\nu$, if the automatic information did not arrive.

- In period $T$, player 1 plays $M$ if he found that player 2 played $L$ in all periods $1, 2, \ldots, T-2$ and played an action other than $C$ in period $T-1$. Otherwise, player 1 plays $U$. Then, player 1 observes the other players, irrespective of his action in that period.

Player 2 plays $\sigma_2$ as follows:

- At any history $h_2^t$ such that $t \leq T-2$, player 2 plays $L$ and observes the other players regardless of his action in period $t$.

- In period $T - 1$, player 2 plays $C$ if he played $L$ in all past periods. Otherwise, let $\tau \geq 1$ be the number of past periods when he played an action other than $L$. Then

  (i) if $(1 - \lambda_1)^\tau (1 - \nu)^\tau \leq \mu$, he plays $L$ and observes the other players irrespective of his action, and

  (ii) if $(1 - \lambda_1)^\tau (1 - \nu)^\tau > \mu$, he plays $R$ and observes the other players irrespective of his action.

- In period $T$, player 2 plays $C$ if he played $L$ every period from period 1 to $T-2$ and played an action other than $C$ in period $T-1$. Otherwise, he plays $L$. Then, player 2 observes the other players, irrespective of his action in that period.

On the path of play, the players play $(U, L, b_3, \ldots, b_n)$ in all periods except period $T-1$, and play $(M, C, b_3, \ldots, b_n)$ in period $T-1$. Hence, the payoff vector of $\sigma$ differs from the Nash equilibrium payoff vector of $G$.

To prove that $\sigma$ is a sequential equilibrium, we employ a specific tremble and the corresponding system of beliefs. Let us consider a tremble such that player 2's deviation in period $T-1$ after playing $L$ in all past periods is far less likely than the deviations at all other histories at period $T-1$. Under the corresponding system of beliefs, at any history of player 1 at period $T$ such that

- in some period $t \leq T-2$, player 1 did not monitor the other players,
- player 2 played $L$ in any period $t \leq T-2$ in which player 1 monitored the other players, and
- player 1 found that player 2 played an action other than $C$ in period $T-1$,

player 1 believes that player 2 did not play $L$ in at least one period player 1 did not monitor the other players.

We show that $\sigma$ satisfies sequential rationality under this system of beliefs. For any player $i \neq 2$, his action and monitoring decision do not affect the other players' continuation strategies. Hence, at any history of player $i \geq 3$ at any period, it is always optimal to play $b_i$, a static best reply, and then monitor the others. This proves sequential rationality for any player $i \geq 3$. For player 1, note that any monitoring decision is optimal at any history-action pair of player 1 at any period, because the other players' strategies are pure. Thus, it suffices to show that at any history of player 1 at any period, the action $\sigma_1$ prescribes is his short-run best response given the belief. First, we consider the histories at period $T$.

1. Player 1 believes that player 2 plays $C$ at any history such that player 1 found that player 2 played $L$ in all periods $1, 2, \ldots, T-2$ and played an action other than $C$ in period $T-1$. Hence, $M$ is his short-run best response.

2. Player 1 believes that player 2 plays $L$ at any history such that player 1 found that player 2 played an action other than $L$ in some period $t \leq T-2$. Hence, $U$ is his short-run best response.

3. Player 1 believes that player 2 plays $L$ at any history such that player 1 found that player 2 played $C$ in period $T-1$. Hence, $U$ is his short-run best response.

4. Consider any history such that (i) player 1 did not monitor the other players in period $T-1$ and (ii) in any period $t \leq T-2$ he monitored the other players, player 2 played $L$. At that history, player 1 believes that player 2 did not deviate. Hence, player 1 believes that player 2 plays $L$ in period $T$, and $U$ is his short-run best response.

5. Consider any history such that (i) in some period $t \leq T-2$ he did not monitor the other players, (ii) in any period $t \leq T-2$ he monitored the other players, player 2

played $L$, and (iii) in period $T-1$ he found that player 2 played an action other than $C$. At that history, as we mentioned before, player 1 believes that player 2 did not play $L$ in at least one of the periods he did not monitor the other players. Therefore, player 1 believes that player 2 plays $L$ in period $T$, and $U$ is his short-run best response.

Next, we consider the histories at period $T-1$. When player 1 did not observe a deviation of player 2, player 1 believes that player 2 plays $C$ in period $T-1$. Hence, $M$ is his short-run best response. When player 1 observed that player 2 deviated in some period before period $T-1$, player 1 believes that player 2 never plays $C$ in period $T-1$. Hence, $U$ is his short-run best response.

Finally, at any history of player 1 at period $t \leq T-2$, player 1 believes that player 2 plays $L$ in period $t$. Hence, $U$ is his short-run best response.

Next, we examine sequential rationality of $\sigma_2$. Since player 2 is prescribed to monitor the other players at any history-action pair at any period, his monitoring decision is always optimal. Let us check the optimality of his actions. We start with the histories of player 2 at period $T$, and there are two cases to consider. First, at any history such that player 2 played $L$ every period from period 1 to $T-2$ and played an action other than $C$ in period $T-1$, player 2 believes that player 1's action in period $T$ is $M$ if player 1 monitored the other players in all past periods, and is $U$ otherwise. The probability with which player 1 monitored the other players in all past periods is

$$\{\lambda_1 + (1-\lambda_1)\nu\}^{T-1} = \mu,$$

where the equality follows from (3). Thus, player 2 believes that player 1 plays $U$ with probability $1-\mu$ and $M$ with probability $\mu$ in period $T$. Therefore, it is optimal for player 2 to play $C$ in period $T$. Second, at any other history, player 2 believes that player 1 plays $U$ in period $T$. Thus, it is optimal for player 2 to play $L$ in period $T$.

Let us examine the histories at period $T-1$. First, at any history such that player 2 played $L$ in all past periods, player 2 believes that player 1 plays $M$ for sure in period $T-1$. If player 2 plays $C$ in period $T-1$, his stage payoff is 3. Further, he reaches to a history at period $T$ where he believes that $(U, L, b_3, \ldots, b_n)$ is played and receives the stage payoff of 4. If player 2 does not play $C$ in period $T-1$, his stage payoff is at most 4. Further, he reaches to a history at period $T$ where he believes that player 1's action is $(1-\mu)U + \mu M$, and therefore his stage payoff is 3. Thus, the current gain from this deviation is at most 1, but the discounted loss in the next period is 1. Hence, it is optimal for player 2 to play $C$ in period $T-1$.

Next, consider a history such that player 2 did not always play $L$ in all past periods. Let $\tau \geq 1$ be the number of periods he did not play $L$. Now player 2 believes that player 1 plays $U$ in period $T$, irrespective of his action in period $T-1$. Hence, it suffices to show that the prescribed action at this history is a short-run best response of player 2, given his belief. Player 2 believes that player 1 did not observe any deviation with the probability $\eta_2 = \{(1-\lambda_1)(1-\nu)\}^{\tau}$. He thus expects that player 1's action in period $T-1$ is $U$ with probability $1-\eta_2$ and $M$ with probability $\eta_2$. His stage payoff in period $T-1$ is $\left(3 - \frac{1-\mu}{\mu}\right)\eta_2 + 4(1-\eta_2)$ if he plays $L$, 3 if he plays $C$, and $4\eta_2 + \left(3 - \frac{\mu}{1-\mu}\right)(1-\eta_2)$ if he plays $R$. If $\eta_2 \leq \mu$, it is optimal for player 2 to play $L$ in period $T-1$, as is prescribed. Otherwise, playing $R$ is optimal, again as is prescribed.

Finally, let us fix a history at period $t \leq T-2$. Note first that deviating to an action other than $L$ in period $t$ reduces his stage payoff by 1 or more. Note also that whether he deviated before or not, both conforming to the continuation strategy of $\sigma_2$ (and playing $C$ in period $T-1$ in case he has not deviated before) and deviating in the current period make player 1 play $U$ in period $T$. Thus the only future effect of a deviation in the current period is to change the outcome in period $T-1$. Whether player 2 deviates or not, he believes that the outcome in period $T-1$ is either (i) playing a static best response against player 1 who does not play $D$ and receiving a stage payoff not exceeding 4, or (ii) playing $(M, C, b_3, \ldots, b_n)$ and receiving the stage payoff of 3. Therefore, the gain in period $T-1$ when player 2 deviated in period $t$ is at most one. This establishes that no one-shot deviation at this history pays, which completes the proof. Q.E.D.

# References


[1] J.-P. Benoît, V. Krishna, Finitely repeated games, Econometrica 53 (1985) 905–922.

[2] E. Ben-Porath, M. Kahneman, Communication in repeated games with costly monitoring, Games and Economic Behavior 44 (2003) 227–250.

[3] J. Flesch, A. Perea, Repeated games with voluntary information purchase, Games and Economic Behavior 66 (2009) 126–145.

[4] O. Gossner, The folk theorem for finitely repeated games with mixed strategies, International Journal of Game Theory 24 (1995) 95–107.

[5] M. Kandori, Introduction to repeated games with private monitoring, Journal of Economic Theory 102 (2002) 1–15.

[6] M. Kandori, I. Obara, Endogenous monitoring (2004), mimeo.

[7] G. Mailath, S. Matthews, T. Sekiguchi, Private strategies in finitely repeated games with imperfect public monitoring, Contributions to Theoretical Economics 2 (2002) Issue 1, Article 2.

[8] E. Miyagawa, Y. Miyahara, T. Sekiguchi, The folk theorem for repeated games with observation costs, Journal of Economic Theory 139 (2008) 192–221.

[9] Y. Miyahara, T. Sekiguchi, Finitely repeated games with monitoring options, Journal of Economic Theory 148 (2013) 1929–1952.

[10] L. Smith, Necessary and sufficient conditions for the perfect finite horizon folk theorem, Econometrica 63 (1995) 425–430.

[11] T. Sugaya, A. Wolitzky, Perfect versus imperfect monitoring in repeated games (2014), mimeo.