Finitely Repeated Games with Automatic and Optional Monitoring*

Yasuyuki Miyahara[a] *
Kobe University

Tadashi Sekiguchi[b]
Kyoto University

September 18, 2014

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 players privately decide whether to exercise a costless monitoring option or not. We show that if the probability of automatic monitoring decreases, the set of sequential equilibrium payoff vectors never shrinks, and it sometimes expands. This extends our earlier result, which only considers the case where the probability of automatic monitoring is zero.

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. We introduce 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.

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

*We thank Masatoshi Tsumagari for a stimulating conversation which originated this research. We are grateful to audience at the MEA meeting and Hitotsubashi University, as well as 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*

probability. We assume that arrivals of the automatic information are independent across the players and over time. In the event that 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, or not. [9] is regarded as a case where the automatic information never arrives.

Fixing the stage-game payoffs and the number of repetition, we examine how the sequential equilibrium payoff vector set of the *repeated games with automatic and optional monitoring* depends on the probabilities of automatic monitoring. Our results are twofold. First, we show that the sequential equilibrium payoff vector set is weakly decreasing (in the sense of set inclusion) in the probabilities of automatic monitoring. In other words, an increase in a player's probability of automatic monitoring never expands the sequential equilibrium payoff vector set.

Second, and more importantly, the increase in a player's probability of automatic monitoring sometimes shrinks the sequential equilibrium payoff vector set. Namely, we show that for any two automatic monitoring probability vectors, $\overline{\lambda}$ and $\underline{\lambda}$ with $\overline{\lambda} \geq \underline{\lambda}$ and $\overline{\lambda} \neq \underline{\lambda}$, there exists a stage game such that the two-period version of this game with the automatic monitoring probability vector $\underline{\lambda}$ has a sequential equilibrium payoff vector which cannot be sustained under the two-period version of this game with $\overline{\lambda}$. Those results are an extension of [9], who just consider the case of $\overline{\lambda} = (1, \ldots, 1)$ and $\underline{\lambda} = (0, \ldots, 0)$.

One interpretation of this formulation of automatic and optional monitoring is players' overlooking. That is, while the exact information about their actions is released (for example, sent by emails), an absentminded player may overlook 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 probability of automatic monitoring, because it creates new equilibrium possibilities. 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 monitoring options are costless, the players never hurt from having a smaller probability of automatic monitoring. Their own monitoring simply compensates for lack of automatic information. Rather, the smaller probability of automatic monitoring expands their strategic flexibility. This is already pointed out by [9], but their result is limited to an extreme case where the probability of automatic monitoring changes from one to zero. Our contribution is to reveal that even a slightest change in the probability of automatic monitoring is, in some cases, sufficient to create new strategic possibilities.

Our second result, strong monotonicity of the sequential equilibrium payoff vector set,

[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.

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

is shown in an example whose one-shot game has a unique 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 in the standard model where the automatic information arrives with certainty.[3] Moreover, the constructed nontrivial sequential equilibrium gives all players different payoffs from the stage-game equilibrium. This multiplicity allows us to prove a folk theorem (like Benoît and Krishina [1], Smith [10], and Gossner [4]) if the horizon is long enough, which does not hold for its standard monitoring version.

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. Thus there is a lack of common knowledge of the deviation, which allows them to design a punishment which works even if the stage game has a unique Nash equilibrium. Note that a greater probability of automatic monitoring is burden for this construction, because we have less flexibility about the probability that the potential deviator confronts the opponents 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 be interesting to investigate those effects more thoroughly.

The rest of this paper is organized as follows. Section 2 introduces the model. Section 3 shows that having a greater probability of automatic monitoring never expands the set of sequential equilibrium payoff vectors, and Section 4 shows that it sometimes shrinks the set.

2. Model

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

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 λ_i player i's *probability of automatic monitoring.* We assume that the arrivals of automatic information are independent across the players. Let $\lambda = (\lambda_1, \dots, \lambda_n)$ be the vector of probabilities of automatic monitoring. If the automatic information did not arrive, each player privately decides whether to monitor his opponents or not.[4] Monitoring the others

[3] However, this is not a novel feature and is already highlighted in [9]. Further, this type of results is known for other repeated games with imperfect monitoring, as in Kandori [5] and Mailath et al. [7]. See also a recent paper by Sugaya and Wolitzky [11].

[4] We assume that the monitoring decision is binary in order to simplify the notations. It can be extended

is costless.

Let $G(\lambda, T)$ be the T-period repeated game with automatic and optional monitoring where stage game G is played in periods $t = 1, \ldots, T$. 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 $\lambda = (0, \ldots, 0)$.

It is also assumed that if a player does not monitor the other players, then he receives no information about their actions. We thus assume that the players receive all stage payoffs at the end of the repeated game. The players can monitor actions in a period only at the end of that period; there is no opportunity 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 the beginning of period $t \geq 2$ consists of all his past information he obtains up to period $t - 1$. For $t \geq 2$, the set of all histories for player i at the beginning of period t is $H_i^t = (I_i)^{t-1}$. Let H_i^1 be an arbitrary singleton set. The set of player i's histories at the beginning of 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, σ_i^a prescribes a mixed action of player i at each history at the beginning of each period, that is, $\sigma_i^a : H_i \to \Delta A_i$. Then σ_i^m prescribes a probability that player i monitors all the other players when automatic information did not arrive, given any history at the beginning of any period and any stage-game action chosen in that period, that is,

$$\sigma_i^m : \bigcup_{t=1}^{T} (H_i^t \times A_i) \to [0, 1].$$

Given strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$, player i's average payoff is given by

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

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

As a solution concept of this paper, we use 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 (h_i^t, a_i) to a

to the case in which players can monitor any subset of the players.

probability distribution of the other players' history profiles $(h_j^t, a_j)_{j\neq i}$. A strategy profile is *completely mixed* if at any history, every stage-game action is selected with positive probability and any monitoring decision is selected with positive probability given any stage-game action. Given a strategy profile σ, a system of beliefs is *consistent* if there exists a sequence of completely mixed strategy profiles converging to σ (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 σ is a *sequential equilibrium*, if there exists a consistent system of beliefs ψ such that at any history of any player i, his continuation strategy is optimal given σ_{-i} and the belief about the other players' histories specified by ψ.

3. Weak Monotonicity

In this section, we show that for any T, the set of sequential equilibrium payoff vectors of $G(\lambda, T)$ is weakly decreasing with respect to probabilities of automatic monitoring in the sense of set inclusion.

Proposition 1. *For any $T \geq 1$, and any $\underline{\lambda}$ and $\overline{\lambda}$ such that $\underline{\lambda} \leq \overline{\lambda}$, any sequential equilibrium payoff vector of $G(\overline{\lambda}, T)$ is a sequential equilibrium payoff vector of $G(\underline{\lambda}, T)$.*

Proof. See Appendix A.

The intuition behind Proposition 1 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 $G(\overline{\lambda}, T)$, so that the modified strategy profile of $G(\underline{\lambda}, T)$ has exactly the same play as the original equilibrium.

4. Strong Monotonicity

In this section, we show that for any $\underline{\lambda}$ and $\overline{\lambda}$ such that $\underline{\lambda}_i \leq \overline{\lambda}_i$ for any i with strict inequality for some i, there exists a stage game such that the set of sequential equilibrium payoff vectors of $G(\overline{\lambda}, 2)$ is strictly smaller than that of $G(\underline{\lambda}, 2)$. That is, a very short horizon, $T = 2$, is sufficient to create a difference.

Proposition 2. *Fix $\underline{\lambda}$ and $\overline{\lambda}$ such that $\underline{\lambda}_i \leq \overline{\lambda}_i$ for all i with strict inequality for some i. Then, there exists a strategic form game G such that $G(\overline{\lambda}, 2)$ has a unique sequential equilibrium payoff vector and $G(\underline{\lambda}, 2)$ has a sequential equilibrium whose payoff vector is different from the equilibrium payoff vector of $G(\overline{\lambda}, 2)$.*

Proof. Without loss of generality, we can fix $\underline{\lambda}$ and $\overline{\lambda}$ so that $\underline{\lambda}_1 < \overline{\lambda}_1$ and $\underline{\lambda}_i \leq \overline{\lambda}_i$ for any $i \geq 2$. Define G so that $A_1 = \{U, M, D\}$, $A_2 = \{L, C, R\}$, and $A_i = \{b_i, c_i\}$ for any $i \geq 3$. u_1 and u_2 depend only on (a_1, a_2), and are represented by the following payoff matrix.

[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 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 $(h_j^t)_{j\neq i}$.

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

Figure 1: Payoff matrix for players 1 and 2

Here, $\beta = 1 - (\overline{\lambda}_1 + \underline{\lambda}_1)/2$. Note that $\beta \in (0, 1)$. For $i \geq 3$,

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

First, we show that strategic form game G has the unique Nash equilibrium $(U, L, b_3, \dots, b_n)$. It is sufficient to show that C is not played in equilibrium with positive probability, because we have 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 order that C is a best response of player 2, the following condition must be satisfied.

$$\frac{\beta y}{1-\beta} + \frac{6\beta}{1-\beta}(1 - x - y) \leq x \leq \frac{\beta y}{1-\beta} - 6(1 - x - y).$$

The above condition is satisfied only if $1 - x - y = 0$ and $x = \beta y/(1 - \beta)$, that is, $x = \beta$ and $y = 1 - \beta$. 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 cannot be played with positive probability in equilibrium. Hence, we have the unique Nash equilibrium $(U, L, b_3, \dots, b_n)$.

In what follows, we show that in $G(\underline{\lambda}, 2)$, there exists a sequential equilibrium payoff vector which is different from the unique Nash equilibrium payoff vector of G, and in $G(\overline{\lambda}, 2)$, the sequential equilibrium payoff vector is unique, which equals the unique Nash equilibrium payoff vector of G.

First, let us consider $G(\underline{\lambda}, 2)$, and define the following strategy profile $\hat{\sigma}$.

- In period 1, player 1 plays M. Then if the automatic information did not arrive, he monitors the other players with probability $(1 - \underline{\lambda}_1 - \beta)/(1 - \underline{\lambda}_1)$, and does not monitor them with probability $\beta/(1 - \underline{\lambda}_1)$ 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 payoff vector is $(5, 3.5, 2.5, \dots, 2.5)$. 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 the beginning of 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 the beginning of period 2 such that player 1 did not find a deviation of player 2 in period 1, player 1 believes that player 2 plays L with probability one in period 2. 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 in period 1, 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 in period 1 and randomize between monitoring and not monitoring in the way specified by $\hat{\sigma}$.

Next, we consider player 2.

- At any history of player 2 at the beginning of period 2 such that he played C in period 1, player 2 believes that player 1 plays U with probability one. Then, it is optimal for player 2 to play L with any monitoring decision in period 2.
- At any history of player 2 at the beginning of period 2 such that he did not play C in period 1, player 2 believes that player 1 plays U and M with probabilities β and $1 - \beta$ 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

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

  It is optimal for player 2 to play C with any monitoring decision in period 2, which gives him the stage payoff 3.
- When player 2 follows $\hat{\sigma}_2$, his average payoff is 3.5. If he does not play C in period 1, then his stage payoff is at most 4. We have seen that this deviation leads to a history at period 2 from which he obtains the stage payoff 3. Therefore, his average payoff when he does not play C in period 1 is at most 3.5, which proves that conforming to $\hat{\sigma}_2$ is optimal.

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.

Finally, we show $G(\overline{\lambda}, 2)$ has a unique sequential equilibrium payoff vector. Let σ be a sequential equilibrium in $G(\overline{\lambda}, 2)$.

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

At any history of player $i \geq 3$ at the beginning of 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 σ. Also, let $\hat{H}_2^2$ be the set of player 2's histories at the beginning of period 2 such that he played $a_2 \in \hat{A}_2$ and found any player $i \geq 3$ 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 the beginning of period 2 is either $(a, 0)$, $(a, 1)$ or a_1.

Suppose, on the contrary to the claim, that σ_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 β and M with probability $1 - \beta$. For player 1, U and M are not simultaneously optimal at any history at the beginning of 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). Hence player 2 must believe that player 1 at $h_1^2 = (a, 0)$ plays a pure action other than D, which we denote by $\hat{a}_1$ (since $\overline{\lambda}_1 > 0$, player 2 believes that player 1 is at $h_1^2 = (a, 0)$ with positive probability).

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 $\tilde{a}_1 \in \{U, M\}$ such that $\tilde{a}_1 \neq \hat{a}_1$.

Sequential rationality implies that after choosing a_1 and not receiving automatic information, player 1 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 σ_2 prescribes not to monitor the other players with positive probability if he played a_2 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 σ_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 the beginning of period 2 such that he did not find any player's deviation in period 1, σ_1 prescribes U with probability one.*

At any history of player 1 at the beginning of period 2 such that he did not find any player's deviation in period 1, consistency requires that he believes 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 the beginning of 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, σ_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, σ_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 with probability $\overline{\lambda}_1 > 1 - \beta$, the probability with which player 2 believes that player 1 plays U is less than β. Given the belief, C is not optimal.

Claim 6. *Fix $a_2 \notin \hat{A}_2$, and suppose that σ_2 prescribes not to monitor the other players with positive probability if he played a_2 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 σ_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 the beginning of period 2 where he played a_2 and monitored the others in period 1. Again, monitoring the others and then playing an optimal action for each realized history at the beginning of period 2 improve player 2's payoff. This is a contradiction.

Claim 7. *At any history of player 1 at the beginning of period 2, σ_1 prescribes U with probability one.*

By Claim 4, it suffices to consider a history where player 1 found some player's deviation in period 1. Consistency requires that player 1 believes that player 2 conforms to his monitoring decision given his own deviation. 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 the beginning of period 2, σ_2 prescribes L with probability one.*

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

From Claims 1 to 8, the players follow $(U, L, b_3, \ldots, b_n)$ at any history at the beginning of 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 sequential equilibrium is $(4, 4, 2, \ldots, 2)$, which completes the proof. Q.E.D.

We have four remarks on this proposition. First, the constructed nontrivial sequential equilibrium of $G(\underline{\lambda}, 2)$ gives each player a different payoff from the static equilibrium. This multiplicity can be used to prove a folk theorem when the horizon is long enough.[6] In contrast, $G(\overline{\lambda}, 2)$ has a unique sequential equilibrium payoff vector, and this is true for any $G(\hat{\lambda}, 2)$ with $\hat{\lambda} \geq \overline{\lambda}$. Also, any finitely repeated game where G is played under the standard perfect monitoring (the automatic information arrives with certainty) has a unique subgame perfect equilibrium, which is repeated play of the one-shot equilibrium. Thus, this is an example where a well-known negative result, which states that uniqueness of stage-game equilibrium implies uniqueness in any finitely repeated game, does not hold under optional monitoring.

Second, what is important for the result is that $\overline{\lambda}_i > \underline{\lambda}_i$ for some i. The relationship between $\overline{\lambda}_j$ and $\underline{\lambda}_j$ for any $j \neq i$ is irrelevant. Hence, we have the following strengthening of Proposition 2. If the probabilities of automatic monitoring change from λ to λ' and if some player's probability of automatic monitoring gets smaller, then there exists a stage game with a unique equilibrium such that a nontrivial equilibrium exists in a finitely repeated game with λ' but not in the finitely repeated game with λ and the same horizon.

Third, although we assumed that $\lambda_i \in (0, 1)$ for any i, the results extends to the case where $\lambda_i = 0$ or $\lambda_i = 1$ for some i. This extension is straightforward, except for notational inconvenience. If $\lambda_i = 0$, player i never receives automatic information. Thus, we redefine I_i, the set of information player i obtains in one period, as $I_i = A \cup A_i$. Instead, if $\lambda_i = 1$, the automatic information surely arrives. Thus, we would have $I_i = A$. With these modifications, we can define player i's histories and strategies, and study the repeated games accordingly. We omit a proof of the extension, but the intuition should be clear. Whether $\overline{\lambda}$ and $\underline{\lambda}$ contain 0 or 1 or not, each player of $G(\underline{\lambda}, T)$ can reproduce play in $G(\overline{\lambda}, T)$ by exercising his monitoring option when necessary.

Fourth, we have only compared $G(\overline{\lambda}, 2)$ and $G(\underline{\lambda}, 2)$, and we have shown that the former game has a unique sequential equilibrium action path. Does the uniqueness of the sequential equilibrium action path extend to $G(\overline{\lambda}, T)$ with any T? The answer is no. We can show that the type of stage games in the proof of this proposition has a nontrivial equilibrium under *any* probability vector of automatic monitoring if T is large enough. See Appendix B for details. This observation suggests difficulty in obtaining a sharp result about the relationship between the automatic monitoring probabilities and the equilibrium payoff set. At the same time, it reveals the extent to which equilibrium possibilities in repeated games with automatic and optional monitoring can be diverse.

[6] For example, one can apply Gossner's [4] folk theorem for finitely repeated games, a strongest result in the literature. We omit the details here, in order to avoid a mere repetition of the argument in [9].

Appendix A. Proof of Proposition 1

Suppose that $\overline{\sigma} \equiv (\overline{\sigma}_i)_{i=1}^n$ is a sequential equilibrium of $G(\overline{\lambda}, T)$, and that $\overline{\psi}$ is a consistent system of beliefs supporting $\overline{\sigma}$. 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 $G(\underline{\lambda}, T)$, $\underline{\sigma}^k \equiv (\underline{\sigma}_i^k)_{i=1}^n$, 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 $\underline{\lambda}_i < \overline{\lambda}_i$, consider the following play of $G(\underline{\lambda}, T)$.

- 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 $x_i \in (0, 1)$. We suppose that the roulette automatically spins whenever automatic information did not arrive.

- 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, then his roulette spins. If his roulette selects 0, he monitors the other players with probability 1. If his roulette 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 the beginning of 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 τ, $\overline{\omega}_i^\tau = (a, 0)$, 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, then his roulette spins. If his roulette selects 0, he monitors the other players with probability 1. If his roulette 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 $G(\underline{\lambda}, T)$ 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 follow the above play defining $\underline{\sigma}^k$, then

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

(b) for any profile of histories at the beginning of 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 the beginning of 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 $G(\overline{\lambda}, T)$.

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. By taking a subsequence if necessary, we can assume that the sequence $(\underline{\sigma}^k, \underline{\psi}^k)_{k=1}^{\infty}$ converges, and let $(\underline{\sigma}, \underline{\psi})$ be its limit. By definition, $\underline{\psi}$ is consistent given $\underline{\sigma}$. It suffices to prove that $\underline{\sigma}$ is a sequential equilibrium of $G(\underline{\lambda}, T)$, because $\overline{\sigma}^k \to \overline{\sigma}$ and $\underline{\sigma}^k \to \underline{\sigma}$ imply that the payoffs of $\overline{\sigma}$ and $\underline{\sigma}$ coincide.

Fix a history of player i 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 an 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 $G(\overline{\lambda}, T)$, the continuation strategy at $\overline{\iota}_i^t$ is optimal under the belief given by $\overline{\psi}$. Since $\underline{\sigma}$ in $G(\underline{\lambda}, T)$ effectively reproduces $\overline{\sigma}$ in $G(\overline{\lambda}, T)$, this implies that the continuation strategy of $\underline{\sigma}_i$ at ι_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 the beginning of period t is h_i^t or $\overline{h}_i^t$. This implies that the continuation strategy of $\underline{\sigma}_i$ at ι_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. Nontrivial Equilibria under a Long Horizon

Let us reexamine the type of stage games we considered in the proof of Proposition 2. Namely, let G be an n-player strategic form game with $A_1 = \{U, M, D\}$, $A_2 = \{L, C, R\}$, and $A_i = \{b_i, c_i\}$ for any $i \geq 3$. As before, u_1 and u_2 depend only on (a_1, a_2), and are represented by the same payoff matrix, reproduced here.

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

A difference from the previous argument is that β does not depend on a given pair of automatic monitoring probability vectors; we simply assume that $0 < \beta < 1$. The payoff function of any player $i \geq 3$ is given by (1). Let us now fix an automatic monitoring probability vector $\lambda \in (0, 1)^n$ arbitrarily. In what follows, we show that there exists T such that $G(\lambda, T)$ has a nontrivial sequential equilibrium.

Let us choose an integer T such that

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

Proposition 2 covers the case where we can set $T = 2$ for $\lambda = \underline{\lambda}$. We thus confine attention to the case of $T \geq 3$.

Let us define the following strategy profile σ. For $i \geq 3$, player i plays b_i and observes the other players at any history of player i. Player 1 plays σ_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 μ_1, where

$$\mu_1 = \frac{(1-\beta)^{\frac{1}{T-1}} - \lambda_1}{1 - \lambda_1}, \tag{3}$$

irrespective of his action in period t (note that we have $\mu_1 \geq 0$ by (2) and $\mu_1 < 1$ by the fact that $(1-\beta)^{1/(T-1)} < 1$).

- 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 μ_1, if the automatic information did not arrive.
- In period T, player 1 plays M if he found that player 2 played L in any period $t \leq T-2$ 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 σ_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 $(1-\lambda_1)^\tau(1-\mu_1)^\tau \leq 1-\beta$, he plays L and observes the other players irrespective of his action, and
  - (ii) if $(1-\lambda_1)^\tau(1-\mu_1)^\tau > 1-\beta$, 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 $(U, L, b_3, \ldots, b_n)$ from period 1 to $T-2$, $(M, C, b_3, \ldots, b_n)$ in period $T-1$, and $(U, L, b_3, \ldots, b_n)$ in period T. Hence, σ is a nontrivial sequential equilibrium if it is a sequential equilibrium.

To prove sequential rationality, we employ a specific tremble and the corresponding system of beliefs. Let us consider trembles 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 the beginning of period $T-1$. Choose any system of beliefs which is made consistent by this type of trembles. Note that under this system of beliefs, at any history of player 1 at the beginning of 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 σ satisfies sequential rationality under this system of beliefs. For any player $i \neq 2$, his chosen actions do not affect the other players' continuation strategies.

Hence, at any history of player $i \geq 3$, it is always optimal to play b_i, a static best reply, and monitor the others. This proves sequential rationality for any player $i \geq 3$. For sequential rationality of player 1, note first that any monitoring decision is optimal at any history, because the other players' strategies are pure. Thus, it is enough to show that, at any history, the action that σ_1 prescribes is a short-run best response of player 1 given the belief. First, we check the optimality of the choices of actions 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 every period from period 1 to $T-2$ and played an action other than C in period $T-1$. Hence, it is optimal for player 1 to play M.

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, it is optimal for player 1 to play U.

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, it is optimal for player 1 to play U.

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. That is, player 1 believes that player 2 plays L in period T, and it is optimal for player 1 to play U.

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 it is optimal for player 1 to play U.

Next, we check the optimality of the choices of actions 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, it is a short-run best response of player 1 to play M in period $T-1$. 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 $T-1$. Then, it is a short-run best response of player 1 to play U in period $T-1$.

Finally, consider any history of player 1 before period $T-1$. At any history at period $t \leq T-2$, player 1 believes that player 2 plays L in period t. Thus, it is a short-run best response of player 1 to play U in period t.

Next, we examine sequential rationality of σ_2. Since it never hurts player 2 to monitor the other players, it is optimal for him to do so at any history. Let us check the optimality of the choices of 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 given by

$$\{\lambda_1 + (1-\lambda_1)\mu_1\}^{T-1} = 1-\beta,$$

where the equality follows from (3). Thus, player 2 believes that player 1 plays U with probability β and M with probability $1-\beta$ 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 optimality in 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 $(U, L, b_3, \dots, b_n)$ 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 $\beta U + (1-\beta)M$ and therefore his stage payoff is 3. Hence, it is optimal for player 2 to play M 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 $\eta_2 = \{(1-\lambda_1)(1-\mu_1)\}^\tau$. He thus expects that player 1's action in period $T-1$ is U with probability $1-\eta_2$ and M with probability η_2. His stage payoff in period $T-1$ is $\{3-\beta/(1-\beta)\}\,\eta_2 + 4(1-\eta_2)$ if he plays L, 3 if he plays C, and $4\eta_2 + \{3-(1-\beta)/\beta\}\,(1-\eta_2)$ if he plays R. If $\eta_2 \leq 1-\beta$, 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 examine the optimality at any history of player 2 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 σ_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 $(M, C, b_3, \dots, b_n)$ 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.