A Folk Theorem for Infinitely Repeated Games with Equivalent Payoffs under Optional Monitoring*

Tadashi Sekiguchi†
Kyoto University

May 6, 2022

Abstract

The present paper studies infinitely repeated games where each player decides whether to monitor the other players' actions or not in each period. The monitoring is costless and private. We ask whether optional monitoring enlarges the equilibrium payoff vector set, in comparison with the standard model of automatic monitoring. For that purpose, we focus on a case where all players have an identical payoff function. Under automatic monitoring, some infinitely repeated games with equivalent payoffs are known to have the smallest equilibrium payoffs which are bounded away from the players' minmax values irrespective of the discount factor. We show that this failure of a folk theorem does not extend to optional monitoring. For a class of repeated games with identical payoffs, the discrepancy between the smallest equilibrium payoff and the minmax value remains under automatic monitoring and vanishes under optional monitoring, if the discount factor goes to one. In other words, optional monitoring expands the environment for the folk theorem to hold.

JEL Classification: C72, C73.
Keywords: repeated games, equivalent payoffs, folk theorem, optional monitoring

1 Introduction

The present paper studies repeated games with *optional monitoring*, first formulated and studied by Miyahara and Sekiguchi [5] for finitely repeated games, and extends their formulation to infinitely repeated games. Both papers assume that each player decides whether to monitor the other players' actions or not in each period. In comparison, the standard model of repeated games with perfect monitoring assumes that each player automatically learns the other players' actions in each period, so that we shall call this information structure *automatic monitoring*.

*The author gratefully acknowledges financial supports from the Grants-in-Aid for Scientific Research (17K03625, 18K01507, 17H00761).

†Institute of Economic Research, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan; sekiguchi@kier.kyoto-u.ac.jp.

An important assumption in Miyahara and Sekiguchi [5] and in the present paper is that monitoring is costless and private, where the latter assumption means that each player does not obtain any signal of the other players' monitoring decisions. A consequence of this assumption is that the players never hurt from exercising their monitoring options. Monitoring enables them to make a better decision in the future and never causes a future punishment. Hence, as [5] formally proves for the case of any finite horizon, any equilibrium payoff outcome of a repeated game with automatic monitoring is an equilibrium payoff outcome of the repeated game with the same stage game, the same number of periods, the same discount factor, and optional monitoring. This is because the players can reproduce the equilibrium strategies under automatic monitoring by strategies where they always exercise there monitoring options.

The present paper thus asks whether optional monitoring enlarges the equilibrium payoff vector set of an infinitely repeated game in the sense of set inclusion. Miyahara and Sekiguchi [5] addresses the same question for finitely repeated games and presents an affirmative answer. More concretely, [5] shows that some finitely repeated games with a unique stage-game equilibrium and optional monitoring possess multiple sequential equilibrium payoff vectors under a finite horizon and even admit a folk theorem under a finite but very long horizon. Note that a unique static equilibrium leads to a unique equilibrium under any finite horizon if the monitoring is automatic. Hence, the result implies that optional monitoring expands the environment for a folk theorem to hold in finitely repeated games.

The present paper focuses on a class of infinitely repeated games where the folk theorem under automatic monitoring by Fudenberg and Maskin [2] and Abreu, Dutta, and Smith [1] fails. [2] is the first to present an example of such failure of the folk theorem, where all three players have an identical stage-game payoff function. Irrespective of the discount factor, the smallest equilibrium payoff (common to all players because of the identical payoffs) is bounded away from their common minmax value, which implies that some feasible and individually rational payoffs are not (even approximately) sustained by an equilibrium under any level of patience. [1] shows that a necessary condition of the failure of the folk theorem is for the stage game to have equivalent payoffs, where a player's stage-game payoff function is a positive affine transformation of another player's. Wen [7] presents a stronger necessary condition, which is for the players with equivalent payoffs to have greater *effective minmax values* than their standard minmax values. [7] also proves a generalization of the folk theorem which applies to *any* stage game, by simply redefining individual rationality as dominance over the effective minmax values.

A main result of the present paper is to show that the failure of the folk theorem under automatic monitoring does not extend to optional monitoring. We prove a folk theorem under optional monitoring for a class of repeated games which includes the example of Fudenberg and Maskin [2], by showing that any feasible and individually rational payoff vector is sustained by an equilibrium if the players are sufficiently patient. Hence for this class, the discrepancy between the smallest equilibrium payoff and the minmax value remains under automatic monitoring and vanishes under optional monitoring, if the discount factor goes to one. In other words, as Miyahara and Sekiguchi [5] show for finitely repeated games, optional monitoring expands the environment for a folk theorem to hold in infinitely repeated games.

To understand our result and its difference from the automatic monitoring case, recall that a key to the folk theorem is to construct an equilibrium where a player to be punished is (nearly) held to his minmax value but the other players are rewarded in order to have an incentive to punish. This may require a punishing player to receive a continuation payoff bounded away from her minmax value. If the punished player and the punishing player have equivalent payoffs, however, the punished player also receives a continuation payoff bounded away from his minmax value. This "curse of equivalent payoffs" would deny the folk theorem.

Optional monitoring may overcome the curse, because we can construct a target equilibrium path where a potential deviator is monitored only by a proper subset of the other players.[1] If the potential deviator actually deviated, the continuation play is not commonly believed among all players because some player did not notice the deviation and consistently believes the play to be on the path. The deviator and the players who noticed the deviation mutually optimize, and payoff equivalence among them may impose some restriction on the equilibrium construction. However, the players who did not notice the deviation do not know the actual continuation play and therefore may not recognize a temptation to deviate. This weakens the incentive constraints of the equilibrium construction, which enables us to prove a folk theorem in our model.

Our result implies that the effective minmax values by Wen [7] are not a lower bound of the equilibrium payoffs of repeated games with optional monitoring. For the stage games studied in this paper, the standard minmax values are a tight lower bound of the repeated-game equilibrium payoffs. We do not know to what extent this property extends to other stage games.

The rest of the present paper is organized as follows. Section 2 sets up our model of repeated games. Section 3 starts with a basic result about the discrepancy between the standard and effective minmax values in our model and then proves a folk theorem under optional monitoring.

2 Model: Repeated Games with Identical Payoffs

The set of the players is $N \equiv \{1, 2, \ldots, n\}$. At the beginning of each period $t \in \{0, 1, 2, \ldots\}$, the players simultaneously choose their stage-game actions. The set of player i's stage-game actions is finite and is denoted by A_i. We define $A = \prod_{i=1}^{n} A_i$. All players have an identical stage-game payoff function, and we denote it by $U : A \to \mathbb{R}$.

At the end of each period t, the players simultaneously decide whether to exercise their monitoring options or not. This decision entails no additional benefit or cost and is completely unobservable to the other players. If player i exercises the option, she learns the other players' stage-game actions of that period, denoted by $a_{-i} = (a_j)_{j \neq i}$. If player i does not exercise it, she receives no information about a_{-i}. Note that this assumption implies that the players do not receive stage-game payoffs at the end of each period. This is because otherwise, a player who did not monitor the other players may still (at least partially) learn their stage-game actions from his stage-game payoff. More formally, we regard the players' discount factor (defined

[1] This idea also appears in Miyahara and Sekiguchi [5].

later) as a probability that the repeated game continues and assume that the players receive their stage-game payoffs in total when it terminates.

For each i, let I_i be the set of information player i obtains in one period. Formally, we have $I_i = A \cup A_i$, where (i) $a \in I_i$ means that player i selected a_i and learned that the other players selected a_{-i} by monitoring them, and (ii) $a_i \in I_i$ means that player i selected a_i and did not monitor the other players. For each i, let $H_i \equiv \bigcup_{t=0}^{\infty}(I_i)^t$, where $(I_i)^0$ is an arbitrary singleton which is disjoint with any $(I_i)^t$ with $t \geq 1$. Each element of H_i corresponds to player i's private history at the beginning of some period.[2] A strategy of player i maps each element of H_i to a mixed stage-game action, that is, a probability distribution over A_i, and a subsequent, random choice of his monitoring decision which depends on his stage-game action of that period.

Given a strategy profile of all players, each player's payoff of the repeated game is the average, discounted sum of all stage-game payoffs with a common discount factor $\delta \in (0, 1)$. That is, if the strategy profile yields a sequence of the expected stage-game payoffs $(U_t)_{t=0}^{\infty}$, each player's repeated-game payoff is $(1-\delta)\sum_{t=0}^{\infty}\delta^t U_t$. The equilibrium concept of the present paper is sequential equilibrium.

3 Results

Our analysis starts with minmax value concepts. Let us extend the domain of U to allow mixed stage-game actions. The standard minmax value for each player i is defined by

$$\underline{V}_i \equiv \min_s \max_{a_i} U(a_i, s_{-i}),$$

where $s = (s_i)_{i=1}^n$ and each s_i is a probability distribution on A_i. Any repeated-game equilibrium under a given discount factor guarantees each player i her minmax value $\underline{V}_i$. Since the payoffs are identical, this implies that any repeated-game equilibrium under a given discount factor gives at least $\max_i \underline{V}_i$ to each player. We define

$$\underline{V} \equiv \max_i \underline{V}_i$$

and simply call $\underline{V}$ the *minmax value*.

Wen [7] defines the following *effective minmax value*.

$$\underline{V}^E \equiv \min_s \max_i \max_{a_i} U(a_i, s_{-i}).$$

From the definitions, it follows that $\underline{V}^E \geq \underline{V}_i$ for any i, which implies that $\underline{V}^E \geq \underline{V}$.

Example 1 (Coordination games) *Suppose n is odd, and we sometimes write $n = 2k + 1$. Let $A_i \equiv \{-1, 1\}$ for each i. U is any function satisfying the following.*

(i) $U(a) \in [0, 1]$ for any $a \in A$.

(ii) $U(a) = 1$ if $a_1 = a_2 = \cdots = a_n$.

[2] More precisely, each element of $(I_i)^t$ corresponds to player i's private history at the beginning of period t.

(iii) $U(a) = 0$ *if and only if* $\left|\sum_{i=1}^n a_i\right| = 1$.

The stage games in Example 1 are a coordination game which severely punishes a split. Namely, each player's stage-game payoff attains a maximum when all players coordinate their stage-game actions. Further, each player's stage-game payoff attains a minimum when and only when the two stage-game actions are selected by k players and by $k+1$ players, respectively (recall that $n = 2k+1$). The payoffs under the other stage-game action profiles are arbitrary.[3]

Against $a_{-1} \in \prod_{i\geq 2} A_i$ such that $\sum_{i\geq 2} a_i = 0$, the assumption (iii) implies that player 1's stage-game payoff is 0 irrespective of his stage-game action. Combining this with the assumption (i), we see that $\underline{V}_1 = 0$. Due to symmetry, we have $\underline{V} = 0$.

We also have $\underline{V}^E > 0 = \underline{V}$ in this game. To see that, suppose on the contrary that $\underline{V}^E = 0$. Then $s = (s_j)_{j=1}^n$ exists such that for any i,

$$\max_{a_i} U(a_i, s_{-i}) = 0.$$

Due to the assumptions (i) and (iii), these hold only when each s_i is pure and, if we represent s by $a \in A$, we have $\sum_{j\neq i} a_j = 0$ for any i. Summing this up over i, we obtain $\sum_{i=1}^n a_i = 0$. However, this is impossible because n is odd.

We compare our model with the infinitely repeated games with automatic monitoring, whose sole difference from our model is that the players make no monitoring decision. That is, they automatically learn all other players' past stage-game actions. Then the following result is an easy consequence of Wen [7], so we omit the proof.

Proposition 1 *Let us assume automatic monitoring. Under any discount factor, any subgame-perfect equilibrium payoff of each player is at least* $\underline{V}^E$*. Consequently, any payoff vector of the form* $(V, V, \ldots, V)$*, where* $V \in (\underline{V}, \underline{V}^E)$*, is feasible and individually rational, but not a subgame-perfect equilibrium payoff vector under any discount factor.*

Proposition 1 implies that if $\underline{V}^E > \underline{V}$ as in the games in Example 1, the folk theorem by Fudenberg and Maskin [2] does not hold under automatic monitoring. The main contribution of the present paper is to show that the folk theorem sometimes holds under optional monitoring. The next proposition is a key result, which shows when any feasible and individually rational payoff vector which is sufficiently close to the minmax value is sustained by an equilibrium. We will later see how this result establishes a folk theorem. In what follows, the set of player i's mixed best responses against a mixed stage-game action profile s_{-i} is denoted by $\mathrm{BR}_i(s_{-i})$.

Proposition 2 *Let us assume optional monitoring. Further, suppose a proper subset of* N*, denoted by* N_1*, and four mixed stage-game action profiles, denoted by* $(\beta^H, \beta^L, \gamma^H, \gamma^L)$*, exist such that*

(i) $\beta_j^H = \gamma_j^H$ *and* $\beta_j^L = \gamma_j^L$ *for any* $j \in N_1$*,*

(ii) $U(\beta^H) > \underline{V} \geq \max\left\{U(\beta^L), U(\gamma^H), U(\gamma^L)\right\}$*,*

[3] The payoff function of the example in Fudenberg and Maskin [2] for which their folk theorem fails has $n = 3$ and satisfies the assumptions.

(iii) if $j \in N_1$ or if β_j^H is not pure, $\beta_j^H \in \mathrm{BR}_j(\beta_{-j}^H)$,

(iv) if $j \in N_1$ or if β_j^L is not pure, $\beta_j^L \in \mathrm{BR}_j(\beta_{-j}^L)$,

(v) for any $i \notin N_1$, $\gamma_i^H \in \mathrm{BR}_i(\gamma_{-i}^H)$ and $\gamma_i^L \in \mathrm{BR}_i(\gamma_{-i}^L)$.

Then for any $V \in \big(\underline{V}, U(\beta^H)\big)$, $\underline{\delta} \in (0,1)$ exists such that any repeated game with $\delta \geq \underline{\delta}$ has a sequential equilibrium whose payoff vector is $(V, V, \ldots, V)$.

Proof. Let $M = N \setminus N_1$, and fix $V \in \big(\underline{V}, U(\beta^H)\big)$. In Appendix A, we show existence of $\underline{\delta}_1 \in (0,1)$ such that for any $\delta \geq \underline{\delta}_1$, the repeated game with δ has a path of mixed stage-game action profiles $\pi = \big(s(t)\big)_{t=0}^{\infty}$ such that $s(t) \in \{\beta^H, \beta^L\}$ for any $t \geq 0$, and

$$(1-\delta)\sum_{t=0}^{\infty} \delta^t U\big(s(t)\big) = V, \tag{1}$$

$$\inf_{T \geq 1} (1-\delta)\sum_{t=T}^{\infty} \delta^{t-T} U\big(s(t)\big) > \underline{V}. \tag{2}$$

From (1) and (2), we see that the path π attains the target payoff vector $(V, V, \ldots, V)$ and has the continuation payoffs bounded away from the minmax value vector $(\underline{V}, \underline{V}, \ldots, \underline{V})$.[4]

Fix a repeated game with discount factor $\delta \geq \underline{\delta}_1$. For each i, let σ_i be the following strategy of player i.

(I) In any period, suppose player i found any player in M (including himself) selected a stage-game action in the support of the mixed stage-game action specified by π in all past periods he monitored the other players. Then he chooses his stage-game action according to π. Subsequently, he monitors the other players irrespective of his stage-game action if $i \in M$, and he does not monitor the other players regardless of his stage-game action if $i \in N_1$.

(II) In any period, suppose player $i \in M$ found a player in M (including himself) selected a stage-game action not in the support of the mixed stage-game action specified by π in a past period he monitored the other players. Then he chooses his stage-game action according to γ^H if π specifies β^H in the current period, and according to γ^L if π specifies β^L in the current period. Subsequently, he monitors the other players irrespective of his stage-game action.

(III) In any period, suppose player $i \in N_1$ found a player in M selected a stage-game action not in the support of the mixed stage-game action specified by π in a past period he monitored the other players. Then he chooses a best response against γ^H if π specifies β^H in the current period, and a best response against γ^L if π specifies β^L in the current period. Subsequently, he does not monitor the other players irrespective of his stage-game action.

[4] While the idea of this construction is attributed to Sorin [6] and Fudenberg and Maskin [3], Appendix A reproduces it for the readers' convenience.

Let $\sigma \equiv (\sigma_i)_{i=1}^n$. If the players follow σ, the stage-game action profile path is π.

Thus, the proof is complete if we show that σ forms a sequential equilibrium together with a consistent system of beliefs. Let us fix a consistent system of beliefs satisfying the following two conditions.

(A) At any player's any private history, he believes that any player in M monitored the other players in all past periods and that any player in N_1 never monitored the other players in all past periods.

(B) At any player's any private history such that he found any player in M (including himself) selected a stage-game action in the support of the mixed stage-game action specified by π in all past periods he monitored the other players, he believes that all other players in M followed π in all past periods.

These beliefs can be made consistent by trembles satisfying the following two conditions.

(a) Each player is far less likely to deviate in his monitoring decision than to deviate in his stage-game action.

(b) Any player in M is far less likely to deviate than any player in N_1.

The condition (a) supports the condition (A), and the condition (b) supports the condition (B).[5] In what follows, we verify sequential rationality given the system of beliefs.

Fix any period and any player i. First, suppose player i found any player in M (including himself) selected a stage-game action in the support of the mixed stage-game action specified by π in all past periods he monitored the other players. By the condition (B), he believes that all other players in M followed π in all past periods.[6] Hence by (I), player i believes that the continuation stage-game action profile path is the continuation path of π from the current period. Therefore, by (2), the continuation payoff exceeds $\underline{V}$.

We have two cases to consider.

- Suppose $i \in M$. A deviation from σ may increase his stage-game payoff at most by $\max_a U(a) - \min_a U(a)$. By (I), player i believes any other player $j \in M$ to notice the deviation and any other player $j \in N_1$ not to notice it. By the definition of π and (I)–(III), player i believes any other player to follow γ^H in any subsequent period π specifies β^H and to follow γ^L in any subsequent period π specifies β^L, regardless of his continuation strategy. By the assumption, any continuation strategy gives him at most $\underline{V}$. This establishes that any player i's deviation is not profitable for all sufficiently large δ. Hence, $\underline{\delta}_2 \geq \underline{\delta}_1$ exists such that player i's continuation play given the current history is optimal if $\delta \geq \underline{\delta}_2$.

[5]The condition (a) is an application of an idea in Miyagawa, Miyahara, and Sekiguchi [4], who study repeated games where it is costly to monitor the other players' actions.

[6]The condition (B) matters when player i found deviations by players in N_1, which may be interpreted as their responses to a deviation by a player in M in a period player i did not monitor the other players. The condition denies such an interpretation.

- Suppose $i \in N_1$. By the definition of β^H and β^L, player i is prescribed to choose a statically optimal stage-game action. Further, σ is defined so that he believes that his continuation play does not affect the other players' continuation play. Therefore, player i's continuation play given the current history is optimal.

Next, suppose player i found a player in M (including himself) selected a stage-game action not in the support of the mixed stage-game action specified by π in a past period he monitored the other players. By (A), player i believes any other player $j \in M$ to notice the deviation and any other player $j \in N_1$ not to notice any deviation by a player in M from π in any past period. By the definition of π and (I)–(III), player i believes any other player to follow γ^H in any subsequent period π specifies β^H and to follow γ^L in any subsequent period π specifies β^L, regardless of his continuation strategy. Given the belief, it is optimal to choose a static best response against the mixed stage-game action profile in each period. Since $\gamma_i^H \in \mathrm{BR}_i(\gamma_{-i}^H)$ and $\gamma_i^L \in \mathrm{BR}_i(\gamma_{-i}^L)$ for any $i \in M$, the prescribed continuation strategy of each player $i \in M$ is optimal. Further, the prescribed continuation strategy of each player $i \in N_1$ is optimal due to its construction.

Finally, we examine each player's monitoring decision given his private history and his stage-game action in that period. Any player $i \in M$ is prescribed to monitor the other players, which is optimal because he never suffers from exercising his monitoring option. Any player $i \in N_1$ believes that the other players' continuation strategies do not depend on the outcomes on the path. Hence he has nothing to learn from monitoring the other players. This makes his monitoring decision, not to monitor the other players, optimal.

This proves sequential rationality of σ combined with the system of beliefs, which completes the proof. Q.E.D.

Before proceeding to more discussions, we verify that all stage games in Example 1 satisfy the conditions in Proposition 2. Recall that the number of the players is odd, and hence we may write $n = 2k + 1$. Then we choose $N_1 = \{k+2, k+3, \ldots, n\}$, and

(i) β^H is a profile where all players choose 1, and β^L is a profile where the players in $\{1, 2, \ldots, k\}$ choose -1 and the players in $\{k+1, k+2, \ldots, n\}$ choose 1, and

(ii) γ^H and γ^L are such that the players in $\{1, 2, \ldots, k+1\}$ choose -1 and the players in $\{k+2, k+3, \ldots, n\}$ choose 1.

It is easy to see that all conditions in Proposition 2 hold. Appendix B provides another example.

We explain the idea of this equilibrium construction. We first construct the equilibrium path attaining the target payoff level, so that it consists of two mixed stage-game action profiles, β^H and β^L. Further, any player in a proper player subset N_1 does not monitor the other players in any period, and any player in the complement $M \equiv N \setminus N_1$ monitors the other players in any period.[7] The two mixed stage-game action profiles are chosen so that any player in N_1 or any player who is not prescribed a pure stage-game action is best responding and hence does not have a short-run incentive to deviate.[8]

[7] M denotes the set of the monitors.

[8] We require any randomizing player to be best responding because it is not easy to provide incentives to randomize over all stage-game actions on the support.

Given the path, any deviation in terms of stage-game actions is monitored only by the players in M. After such a deviation on the path by a player in M, the players in M switch to punishing mixed stage-game action profiles, γ^H and γ^L. γ^H is specified in any period the equilibrium path specifies β^H, and γ^L is specified in any period the equilibrium path specifies β^L. Since the players in N_1 do not monitor the other players, they are unable to respond to the deviation. Thus, β^H and γ^H specify a common mixed stage-game action to each player in N_1, and β^L and γ^L specify a common mixed stage-game action to each player in N_1. This is reflected in the condition (i) in Proposition 2.

The punishment must be sequentially rational. Since only the players in M know that the play is on the punishment path, we require them to choose a static best response. This establishes sequential rationality and is reflected in the condition (v) in Proposition 2.

This construction builds on optional monitoring and is not valid under standard automatic monitoring. This is because under automatic monitoring, all players automatically notice any deviation. Under any mixed stage-game action profile, at least one player can secure the effective minmax value. This makes it impossible to hold all players below the effective minmax value under automatic monitoring. However, in our construction under optional monitoring, we specify γ^H and γ^L so that any player who can secure the effective minmax value is in N_1. Since any player in N_1 is prescribed not to monitor the other players, she does not know the play is on the punishment path after a deviation by a player in M. We can thus ignore the incentive condition for that player, which cannot be ignored under automatic monitoring.

We next verify that Proposition 2 leads to a folk theorem, as is summarized by the following result.

Proposition 3 *Suppose the conditions in Proposition 2 all hold. Then for any* $V \in (\underline{V}, \overline{U}]$, *where* $\overline{U} \equiv \max_a U(a)$, $\underline{\delta} \in (0,1)$ *exists such that any repeated game with* $\delta \geq \underline{\delta}$ *has a sequential equilibrium whose payoff vector is* $(V, V, \ldots, V)$.

Proof. Note first that any solution of $\max_a U(a)$ is a Nash equilibrium of the stage game. Hence, $(\overline{U}, \overline{U}, \ldots, \overline{U})$ is a sequential equilibrium payoff vector under any discount factor.

Thus, let us assume $V \in (\underline{V}, \overline{U})$. The proof of Proposition 2 shows existence of $\underline{\delta}_1 \in (0,1)$, $\underline{v} \in (\underline{V}, V)$, and $\overline{v} \in (\underline{v}, V)$ such that any $(v, v, \ldots, v)$ with $v \in [\underline{v}, \overline{v}]$ is a sequential equilibrium payoff vector of any repeated game with $\delta \geq \underline{\delta}_1$. Now let us define

$$\underline{\delta} \equiv \max\left\{\frac{\overline{U} - \overline{v}}{\overline{U} - \underline{v}}, \frac{\overline{U} - V}{\overline{U} - \overline{v}}, \underline{\delta}_1\right\}. \tag{3}$$

Fix a repeated game with $\delta \geq \underline{\delta}$. By (3), an integer $T \geq 1$ exists such that

$$\delta^T \leq \frac{\overline{U} - V}{\overline{U} - \overline{v}}. \tag{4}$$

Suppose we also have

$$\delta^T < \frac{\overline{U} - V}{\overline{U} - \underline{v}}, \quad \delta^{T-1} > \frac{\overline{U} - V}{\overline{U} - \overline{v}}. \tag{5}$$

Then from (5), we obtain

$$\delta\frac{\overline{U}-V}{\overline{U}-\overline{v}} < \frac{\overline{U}-V}{\overline{U}-\underline{v}},$$

and this is a contradiction against (3). Hence, an integer $T \geq 1$ exists such that

$$\frac{\overline{U}-V}{\overline{U}-\underline{v}} \leq \delta^T \leq \frac{\overline{U}-V}{\overline{U}-\overline{v}}. \tag{6}$$

From (6), $z \in [\underline{v}, \overline{v}]$ exists such that

$$\frac{\overline{U}-V}{\overline{U}-z} = \delta^T, \quad \therefore V = (1-\delta^T)\overline{U} + \delta^T z.$$

Consider a strategy profile such that the players follow a solution of $\max_a U(a)$ in the first T periods regardless of their histories and then play an equilibrium whose payoff vector is $(z, z, \ldots, z)$ from period T on. Since any solution of $\max_a U(a)$ is a one-shot Nash equilibrium, this strategy profile is a sequential equilibrium whose payoff vector is $(V, V, \ldots, V)$. Since $\delta \geq \underline{\delta}$ is arbitrary, the proof is complete. Q.E.D.

To conclude, Proposition 2 presents a class of infinitely repeated games where the effective minmax value concept by Wen [7] does not provide an effective lower bound of the sequential equilibrium payoffs under optional monitoring. Further, the standard minmax value is a tight lower bound of the sequential equilibrium payoffs as the discount factor goes to one in the class. In other words, optional monitoring expands the environment for the folk theorem to hold in infinitely repeated games. A much harder problem is to characterize the stage games where the folk theorem fails under automatic monitoring and holds under optional monitoring, which is left for future research.

A Appendix: Construction of π

Suppose the conditions in Proposition 2 all hold. Fix $V \in \big(\underline{V}, U(\beta^H)\big)$. Recall that $U(\beta^L) \leq \underline{V}$. Then a positive integer M and a sequence of M mixed stage-game action profiles $B^L \equiv \big(s^L(0), s^L(1), \ldots, s^L(M-1)\big)$ exist such that $s^L(m) \in \{\beta^H, \beta^L\}$ for any $m \in \{0, 1, \ldots, M-1\}$, and

$$U\big(s^L(0)\big) \leq U\big(s^L(1)\big) \leq \cdots \leq U\big(s^L(M-1)\big), \tag{7}$$

$$\underline{V} < \frac{1}{M}\sum_{m=0}^{M-1} U\big(s^L(m)\big) < V. \tag{8}$$

Then $\varepsilon > 0$ and $\underline{\delta}_1 \in (0,1)$ exist such that

$$(\underline{\delta}_1)^M > \frac{1}{2}, \tag{9}$$

and for any $\delta \geq \underline{\delta}_1$, we have

$$\underline{V} + \varepsilon < \frac{1-\delta}{1-\delta^M} \sum_{m=0}^{M-1} \delta^m U\bigl(s^L(m)\bigr) < V. \tag{10}$$

Note that (10) follows from (8) and continuity.

Fix $\delta \geq \underline{\delta}_1$. Let $B^H \equiv \bigl(s^H(0), s^H(1), \ldots, s^H(M-1)\bigr)$ be a sequence of M mixed stage-game action profiles such that

$$s^H(0) = s^H(1) = \cdots = s^H(M-1) = \beta^H.$$

Then, the definition of B^H and (10) imply unique existence of $\zeta \in (0,1)$ such that

$$V = \zeta \frac{1-\delta}{1-\delta^M} \sum_{m=0}^{M-1} \delta^m U\bigl(s^L(m)\bigr) + (1-\zeta) \frac{1-\delta}{1-\delta^M} \sum_{m=0}^{M-1} \delta^m U\bigl(s^H(m)\bigr). \tag{11}$$

Consider the repeated game with δ. Define a path of mixed stage-game action profiles $\pi = \bigl(s(t)\bigr)_{t=0}^{\infty}$ in the following inductive way.

$$\bigl(s(0), s(1), \ldots, s(M-1)\bigr) = \begin{cases} B^H & \text{if } \zeta \geq 1/2, \text{ and} \\ B^L & \text{if } \zeta < 1/2. \end{cases}$$

Next, suppose $s(0), s(1), \ldots, s(TM-1)$ are already defined, where $T \geq 1$. Further, suppose for any $T' < T$, $\bigl(s(T'M), s(T'M+1), \ldots, s(T'M+M-1)\bigr) \in \{B^H, B^L\}$. Now define

$$\Theta_T^H = \Bigl\{T' < T : \bigl(s(T'M), s(T'M+1), \ldots, s(T'M+M-1)\bigr) = B^H\Bigr\},$$
$$\Theta_T^L = \Bigl\{T' < T : \bigl(s(T'M), s(T'M+1), \ldots, s(T'M+M-1)\bigr) = B^L\Bigr\}.$$

Then

$$\begin{aligned} &\bigl(s(TM), s(TM+1), \ldots, s(TM+M-1)\bigr) \\ &= \begin{cases} B^H & \text{if } \zeta - \sum_{T' \in \Theta_T^H} (1-\delta^M)\delta^{T'M} \geq 1 - \zeta - \sum_{T' \in \Theta_T^L} (1-\delta^M)\delta^{T'M}, \text{ and} \\ B^L & \text{if } \zeta - \sum_{T' \in \Theta_T^H} (1-\delta^M)\delta^{T'M} < 1 - \zeta - \sum_{T' \in \Theta_T^L} (1-\delta^M)\delta^{T'M}. \end{cases} \end{aligned}$$

For any $T \geq 1$, we have

$$\begin{aligned} &\zeta - \sum_{T' \in \Theta_T^H} (1-\delta^M)\delta^{T'M} - (1-\delta^M)\delta^{TM} + 1 - \zeta - \sum_{T' \in \Theta_T^L} (1-\delta^M)\delta^{T'M} - (1-\delta^M)\delta^{TM} \\ &= \delta^{TM}(2\delta^M - 1) > 0, \end{aligned}$$

where the last inequality follows from (9). This implies that if $\bigl(s(TM), s(TM+1), \ldots, s(TM+$

$M-1)\big) = B^H$, we have

$$\zeta - \sum_{T'\in\Theta_T^H} (1-\delta^M)\delta^{T'M} - (1-\delta^M)\delta^{TM} > 0,$$

and that if $\big(s(TM), s(TM+1), \ldots, s(TM+M-1)\big) = B^L$, we have

$$1-\zeta - \sum_{T'\in\Theta_T^L} (1-\delta^M)\delta^{T'M} - (1-\delta^M)\delta^{TM} > 0.$$

From these inequalities, we obtain

$$\min\left\{\zeta - \sum_{T'\in\Theta_{T+1}^H} (1-\delta^M)\delta^{T'M}, 1-\zeta - \sum_{T'\in\Theta_{T+1}^L} (1-\delta^M)\delta^{T'M}\right\} > 0 \tag{12}$$

for any $T \geq 1$. Define

$$\begin{aligned}\Theta^H &= \Big\{T : \big(s(TM), s(TM+1), \ldots, s(TM+M-1)\big) = B^H\Big\},\\ \Theta^L &= \Big\{T : \big(s(TM), s(TM+1), \ldots, s(TM+M-1)\big) = B^L\Big\}.\end{aligned}$$

Then, by letting $T \to \infty$ in (12), we obtain

$$\min\left\{\zeta - \sum_{T\in\Theta^H} (1-\delta^M)\delta^{TM}, 1-\zeta - \sum_{T\in\Theta^L} (1-\delta^M)\delta^{TM}\right\} \geq 0. \tag{13}$$

Note that $\Theta^H \cup \Theta^L = \{0, 1, 2, \ldots\}$. It thus follows from (13) that

$$\zeta = \sum_{T\in\Theta^H} (1-\delta^M)\delta^{TM}. \tag{14}$$

Now we can compute the repeated-game payoff of π as follows.

$$\begin{aligned}
&(1-\delta)\sum_{t=0}^{\infty}\delta^t U\big(s(t)\big)\\
=&(1-\delta)\left\{\sum_{T\in\Theta^H}\delta^{TM}\sum_{m=0}^{M-1}\delta^m U\big(s^H(m)\big) + \sum_{T\in\Theta^L}\delta^{TM}\sum_{m=0}^{M-1}\delta^m U\big(s^L(m)\big)\right\}\\
=&\sum_{T\in\Theta^H}(1-\delta^M)\delta^{TM}\frac{1-\delta}{1-\delta^M}\sum_{m=0}^{M-1}\delta^m U\big(s^H(m)\big) + \sum_{T\in\Theta^L}(1-\delta^M)\delta^{TM}\frac{1-\delta}{1-\delta^M}\sum_{m=0}^{M-1}\delta^m U\big(s^L(m)\big)\\
=&\zeta\frac{1-\delta}{1-\delta^M}\sum_{m=0}^{M-1}\delta^m U\big(s^H(m)\big) + (1-\zeta)\frac{1-\delta}{1-\delta^M}\sum_{m=0}^{M-1}\delta^m U\big(s^L(m)\big) = V,
\end{aligned} \tag{15}$$

where the equalities in (15) follow from (14) and (11). Thus (1) holds.

For any $t \geq 1$, let $\hat{T}$ be the largest T such that $t \geq TM$. If $\big(s(\hat{T}M), s(\hat{T}M+1), \ldots, s(\hat{T}M+M-1)\big) = B^H$, we have

$$(1-\delta)\sum_{\tau=t}^{\infty}\delta^{\tau-t}U\big(s(t)\big) \geq (1-\delta)\sum_{\tau=(\hat{T}+1)M}^{\infty}\delta^{\tau-(\hat{T}+1)M}U\big(s(t)\big) \geq \frac{1-\delta}{1-\delta^M}\sum_{m=0}^{M-1}\delta^m U\big(s^L(m)\big) > \underline{V}+\varepsilon,$$

where the last inequality follows from (10). If $\big(s(\hat{T}M), s(\hat{T}M+1), \ldots, s(\hat{T}M+M-1)\big) = B^L$, it follows from (7) that

$$(1-\delta)\sum_{\tau=t}^{\infty}\delta^{\tau-t}U\big(s(t)\big) \geq (1-\delta)\sum_{\tau=\hat{T}M}^{\infty}\delta^{\tau-\hat{T}M}U\big(s(t)\big) \geq \frac{1-\delta}{1-\delta^M}\sum_{m=0}^{M-1}\delta^m U\big(s^L(m)\big) > \underline{V}+\varepsilon,$$

where the last inequality follows from (10). Hence, (2) holds, which completes the proof.

B Another Example

Let us consider a three-player game where $A_1 = A_2 = A_3 = \{-1, 1\}$ and U is represented by the following matrices.

	1	-1
1	X	0
-1	0	X

1

	1	-1
1	0	X
-1	X	1

-1

We assume $X \in (0,1)$. In what follows, we denote any mixed stage-game action profile by $\mu = (\mu_1, \mu_2, \mu_3)$, where μ_i is the probability for player i to select 1.

We first show that the minmax value is $\underline{V} = X/2$. Note that any mixed stage-game action profile of the form $(1/2, \mu_2, \mu_3)$ guarantees the stage-game payoff $X/2$. Further, at the mixed stage-game action profile $(1/2, 1/2, 1)$, player 1 is best responding and obtains $X/2$. This proves $\underline{V} = X/2$.

We also show that the effective minmax value exceeds $\underline{V}$. Define

$$\mu^* = \frac{1+\sqrt{X}}{1+2\sqrt{X}}, \quad U^* = \frac{1+2\sqrt{X}+2X}{1+4\sqrt{X}+4X}X.$$

Note that $\mu^* > 1/2$ and $U^* > \underline{V}$. Fix a mixed stage-game action profile μ arbitrarily. Suppose for some i and $j \neq i$, $\min\{\mu_i, \mu_j\} \geq \mu^*$. Then, the remaining player's stage-game payoff when he chooses 1 is

$$\big\{\mu_i\mu_j + (1-\mu_i)(1-\mu_j)\big\}X \geq \big\{(\mu^*)^2 + (1-\mu^*)^2\big\}X = U^*,$$

where the inequality holds because the left-hand-side is increasing in μ_i and μ_j (note that $\min\{\mu_i, \mu_j\} \geq \mu^* > 1/2$).

Suppose otherwise. Then for some k and l, we have $\max\{\mu_k, \mu_l\} \le \mu^*$. The remaining player's stage-game payoff when he chooses -1 is

$$\big\{\mu_k(1-\mu_l) + (1-\mu_k)\mu_l\big\}X + (1-\mu_k)(1-\mu_l) \ge 2\mu^*(1-\mu^*)X + (1-\mu^*)^2 = U^*,$$

where the inequality holds because the left-hand-side is decreasing in μ_k and μ_l.

Consequently, for any μ, some player can secure a stage-game payoff not smaller than U^*. Further, any player is best responding at (μ^*, μ^*, μ^*) and obtains U^*. As a result, the effective minmax value is U^*.

Finally, it is easy to see that the conditions in Proposition 2 are all satisfied under the following specification. We set $N_1 = \{3\}$, and

$$\beta^H = (1,1,1), \quad \beta^L = (1, 1/2, 1/2), \quad \gamma^H = (1/2, 1/2, 1), \quad \gamma^L = (1, 1, 1/2).$$

References

[1] D. Abreu, P. K. Dutta, L. Smith, The folk theorem for repeated games: A neu condition, Econometrica 62 (1994) 939–948.

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

[3] D. Fudenberg, E. Maskin, On the dispensability of public randomization in discounted repeated games, Journal of Economic Theory 53 (1991) 428–438.

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

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

[6] S. Sorin, On repeated games with complete information, Mathematics of Operations Research 11 (1986) 147–160.

[7] Q. Wen, The "folk theorem" for repeated games with complete information, Econometrica 62 (1994) 949–954.