Repeated Games with Observation Costs*

Eiichi Miyagawa
Department of Economics
Columbia University

Yasuyuki Miyahara
Graduate School of Business Administration
Kobe University

Tadashi Sekiguchi
Institute of Economic Research
Kyoto University

January 27, 2003

Abstract

This paper analyzes repeated games in which it is possible for players to observe the other players' past actions without noise but it is costly. One's observation decision itself is not observable to the other players, and this private nature of monitoring activity makes it difficult to give the players proper incentives to monitor each other. We provide a sufficient condition for a feasible payoff vector to be approximated by a sequential equilibrium when the observation costs are sufficiently small. We then show that this result generates an approximate Folk Theorem for a wide class of repeated games with observation costs. The Folk Theorem holds for a variant of prisoners' dilemma, partnership games, and any games in which the players have an ability to "burn" small amounts of their own payoffs.

Journal of Economic Literature Classification Numbers: C72, C73, D43, D82.

Key Words: repeated games, private monitoring, costly monitoring, Folk Theorem.

*This research was started while Miyagawa was visiting the Graduate School of Economics at Kobe University; he thanks the school for its hospitality. Sekiguchi thanks financial supports from Grant-in-Aid for Scientific Research and the Japan Economic Research Foundation. All of us thank the participants of the 8th Decentralization Conference in Japan and the workshops at the University of Tokyo and Hitotsubashi University for helpful comments.

1 Introduction

In the theory of repeated games, the benchmark assumption is that of perfect monitoring, i.e., the players obtain perfect information about the other players' past actions. Under the assumption, the theory shows that the players can sustain a large set of payoff vectors as equilibria by making their actions contingent on the other players' past actions.[1] The recent literature relaxes the assumption of perfect monitoring and considers the case in which players receive only imperfect (public or private) information about the other players' past actions.[2]

The present paper relaxes the assumption of perfect monitoring in a different direction. We consider the case in which it is possible for the players to obtain perfect information about the other players' past actions but it is *costly*. We assume that at the end of each period, each player decides whether to obtain information about the actions chosen by the other players in the period. Obtaining the information costs a certain amount of utility, which is referred to as the observation cost. If a player chooses not to pay the observation cost at the end of a period, then she obtains no information about the other players' actions chosen in the period.[3] We also assume that a player's observation decision itself is not observable to the other players. Perfect monitoring can be considered as the limit case in which the observation costs are zero for all players.

It is important to note that the model of costly monitoring differs considerably from that of perfect (costless) monitoring even when the observation costs are arbitrarily small as long as they are positive. To see this, consider a repeated prisoners' dilemma with costly monitoring. Suppose that the players use the trigger strategy profile, in which each player starts with cooperation but switches to perpetual defection if (and only if) a defection is observed in the past. If the observation costs are zero and the players are sufficiently patient, then the trigger strategy profile is an equilibrium. However, this strategy profile is *not* an equilibrium when the observation

[1]See, e.g., Abreu (1986, 1988) and Fudenberg and Maskin (1986).

[2]For the case of imperfect public monitoring, see, e.g., Abreu, Pearce, and Stacchetti (1990), Fudenberg, Levine, and Maskin (1994), and Fudenberg and Levine (1994). For the case of imperfect private monitoring, see, e.g., Sekiguchi (1997), Bhaskar and van Damme (2002), Bhaskar and Obara (2002), Ely and Valimaki (2002), Mailath and Morris (2002), Matsushima (2002), and Piccione (2002).

[3]This formulation raises a subtle issue on when the players receive payoffs. While this is discussed in detail in subsequent sections, for the time being imagine that the payoffs are received as a whole when the game "ends," interpreting the discount factor as the probability with which the game continues.

costs are strictly positive even when they are arbitrarily small. The reason is simply that since the strategy profile is deterministic, each player knows the other player's past and future actions on the equilibrium path and has no reason to pay the observation costs. Therefore, in equilibrium, no one monitors the other player. However, then deviations from the strategy profile are not detected and hence cooperation is not sustained as an equilibrium. This argument generalizes and we can show that at any pure-strategy equilibrium of a repeated game with observation costs, the players play a stage-game equilibrium in every period (with no observation activity).[4]

Therefore, a construction of non-trivial cooperative/efficient equilibria must use strategy profiles in which some of the players randomize. The main contribution of the present paper is to show that such a construction is possible and therefore some positive results are obtained in a wide class of situations. First, we provide a sufficient condition for a payoff vector to be approximated by a sequential equilibrium when the observation costs are sufficiently small and the players are sufficiently patient. Using the condition, we then prove an approximate Folk Theorem for several classes of repeated games with observation costs. The approximate Folk Theorem is shown to hold for a variant of prisoners' dilemma, partnership games, and any game in which the players have an ability to "burn" small amounts of their own payoffs.

An important assumption for the positive results is small observation costs. The results say only that a large set of payoff vectors can be sustained when the observation costs are small and the players are patient. We are unable to prove a general Folk Theorem/efficiency result for a *given* level of observation cost. However, we believe that our result is of some economic relevance because in many interesting economic applications, the observation costs can be considerably small. For example, consider two firms competing in terms of prices. If these firms compete in a small local market, it can be a matter of walking several blocks to see the rival's prices. The cost of such activity can be indeed small in comparison with the magnitude of their business.

More theoretically, it can be argued that the approximate Folk Theorem

[4]The unobservability of monitoring decisions plays an important role in the result. If monitoring decisions are observable, then the situation does not differ very much from perfect monitoring. Indeed, in the repeated prisoners' dilemma with observation costs, cooperation can be sustained by a modified trigger strategy profile in which punishment is triggered by not only a single defection but also a single failure to observe the other player's action. On the other hand, the perfect observability of monitoring decisions is difficult to imagine when the monitoring activity takes the form of spying or glimpsing.

demonstrates the robustness of the assumption of perfect monitoring. On the one hand, many regard perfect monitoring as an extreme assumption. In reality, information about the past comes at a (possibly small) cost. On the other hand, as we have seen before, the model with zero observation costs and the one with positive costs differ significantly in terms of the incentive for monitoring. Thus, it is theoretically an interesting question whether the two models yield qualitatively similar results, justifying our approach to regard perfect monitoring as a limit of costly monitoring.

A few papers have studied repeated games with costly monitoring. Ahn and Suominen (2001) consider a random matching game (like the one in Kandori (1992)) with a twist that each player is given an opportunity to invest in a monitoring technology *in the initial period*. If the player invests in the technology in the initial period, she can observe her neighbors' actions in all subsequent periods. Thus the costly monitoring activity in their model has a once-and-for-all nature. In our model, on the other hand, the player has to engage in costly monitoring in every period if she wants to keep track of the other players' behavior completely.

A paper more closely related to ours is Miyahara (2002), who considers a repeated prisoners' dilemma in which monitoring at the end of a period gives a player information not only about the period but also about some of the previous periods. Miyahara (2002) shows that efficiency can be approximated if the monitoring costs are sufficiently small. It is important to point out that the result in the present paper does *not* subsume that of Miyahara (2002) since the latter uses a construction that takes advantage of the assumption that more than one period in the past can be observed.

Our model is a special example of repeated games with private monitoring. Since each player's observation (if any) is not observable to the other players, it is private information, which makes our model private monitoring. The literature of repeated games with private monitoring has focused on the case when players receive *noisy* signals of the other players' actions *costlessly*, while we examine the case when players obtain *complete* information if they pay observation costs. Thus, the results and the construction of cooperative/efficient equilibria in the literature do not apply in our model. However, this does not mean that our model has no bearing on repeated games with noisy costless private monitoring. In Section 6, we briefly discuss what happens if costly monitoring is introduced into repeated games with noisy private monitoring.

The remaining part of this paper is organized as follows. Section 2 introduces the model. Section 3 provides important definitions. Section 4 states our main result, describes the strategy profile used in the proof, and sketches the proof. Section 5 applies the result to prove an approximate Folk Theorem in a variant of prisoners' dilemma, partnership games, and games with an opportunity of utility burning. Section 6 discusses possible extensions of our model. The Appendix proves the main result.

2 Model

The stage game is a finite n-player game $G = \{n, A, (u_i)_{i=1}^n\}$, where $A = \times_{i=1}^n A_i$ and $u_i \colon A \to \mathbb{R}$ is player i's stage payoff function. We often write $u(a) = (u_i(a))_{i=1}^n$. For each i, let S_i be the set of all mixed actions for player i and let $S = \times_{i=1}^n S_i$. For a mixed action profile $s \in S$, we abuse notation and let $u_i(s)$ denote the expected payoff of player i under s. Let "co" denote convex hull, and define $V = \text{co}\{u(a) : a \in A\}$, which is the set of feasible payoff vectors.

Game G itself does not include monitoring activity. Thus, precisely speaking, G is not the game played in every period. It is meant to describe the basic strategic interaction within each period.

The infinitely repeated version of G (plus monitoring activity) with discounting and observation costs is denoted by $\Gamma(\delta, \lambda)$, where $\delta = (\delta_1, \ldots, \delta_n) \in (0,1)^n$ is a vector of discount factors and $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n_{++}$ is a vector of observation costs. We permit differential discount factors.[5] In each period, each player simultaneously chooses an action $a_i \in A_i$ and then decides whether to privately observe the actions that the other players chose in the period. For each player i, λ_i denotes the cost of observing the others' actions. We also assume that if player i does not monitor the other players at the end of a period, then *no* information about the action profile of the other players in the period is revealed to player i. Each player's monitoring decision itself is assumed to be private and not observable to the other players. Hence player i's private information on the play of a given past period can be represented by a pair of her chosen action and observations, $(a_i, \omega_i) \in A_i \times (A_{-i} \cup \{\phi\})$. Here, $(a_i, \omega_i) = (a_i, a_{-i})$ means that player i chose a_i, monitored the other players, and observed a_{-i}. On the other hand, $(a_i, \omega_i) = (a_i, \phi)$ means that player i chose a_i and did not monitor the other players.

[5] As Lehrer and Pauzner (1999) show, when discount factors are heterogeneous, payoff vectors outside V might be feasible. However, the present paper concentrates on sustaining payoff vectors in V for expositional simplicity. We consider differential discount factors only to demonstrate that our analysis does not require identical discount factors, although our construction can be used to sustain payoff vectors outside V.

We assume that monitoring is the only way to obtain information about the other players' past actions. This implies that the players do *not* receive the stage payoffs in each period, but receive them in total at the "end" of the repeated game. Of course, the infinitely repeated game never ends under the basic interpretation of the game. However, if we regard each δ_i as a (subjective) probability with which the game continues, then the interpretation about the timing of receiving payoff is less problematic. Anyway, this assumption is extreme, and it is assumed to make the issue of costly monitoring as stark as possible (and partly for analytical simplicity). In Section 6, we briefly comment on what happens if payoffs are received in each period, in which case realized payoffs give players information about the others' actions.

We also assume that there exists a public randomization device which generates a sunspot according to the uniform distribution over the unit interval $[0, 1]$. At the beginning of each period, a sunspot is realized and observed by the players before they choose their actions. The (private) history for player i at the beginning of period t before she chooses an action is denoted by h_i^t and defined as the sequence of her private information and realized sunspots up to the beginning of the period. Thus, the set of all possible histories for player i at period t is defined by $H_i^t = [0, 1]^t \times (A_i \times (A_{-i} \cup \{\phi\}))^{t-1}$, where H_i^1 is equivalent to the set of sunspots. Then the set of possible histories for player i is $H_i = \cup_{t=1}^{\infty} H_i^t$.

Player i's strategy σ_i is a function from H_i to $S_i \times [0, 1]^{|A_i|}$ where $|A_i|$ is the cardinality of A_i. Thus, for any history $h_i^t \in H_i$, we have $\sigma_i(h_i^t) = (s_i(t), \{l_i(a_i, h_i^t)\}_{a_i \in A_i})$, where $s_i(t)$ is player i's (possibly mixed) action in period t given h_i^t, and $l_i(a_i, h_i^t)$ is the probability that player i monitors the other players given that the history is h_i^t and she played a_i in period t. Player i's payoffs in $\Gamma(\delta, \lambda)$ are the average (expected) discounted sum of the stage game payoffs minus observation costs. Formally, player i's payoffs under a strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ are denoted by $g_i(\sigma)$ and given by

$$g_i(\sigma) = (1 - \delta_i) \sum_{t=1}^{\infty} \delta_i^{t-1} E\big[u_i(a(t)) - \lambda_i \cdot l_i(a_i(t), h_i^t) \mid \sigma\big],$$

where $E[\ \cdot \mid \sigma]$ denotes the expectation with respect to the probability measure over histories induced by strategy profile σ.

3 Definitions

This section introduces some definitions which facilitate subsequent analysis.

Let the stage game $G = \{n, A, (u_i)_{i=1}^n\}$ be given. For a given (possibly mixed) action profile $s \in S$ for G, we define

$$BR_i(s) = \{a'_i \in A_i : u_i(a'_i, s_{-i}) \geq u_i(a_i, s_{-i}) \quad \forall a_i \in A_i\},$$

which is the set of (pure-action) best responses of player i against s_{-i}. For a given (pure) action profile $a \in A$, we define

$$B_i(a) = \{a'_i \in A_i : u_i(a'_i, a_{-i}) > u_i(a)\},$$

which is the set of (strictly) better replies to a_{-i};

$$D(a) = \{i : a_i \notin BR_i(a)\},$$

which is the set of players for whom a_i is not a best response to a_{-i};

$$D^w(a) = \{i : BR_i(a) \neq \{a_i\}\},$$

which is the set of players for whom a_i is not a unique best response to a_{-i}; and

$$SD(a) = \{i : BR_i(a') = \{a_i\} \quad \forall a' \in A\},$$

which is the set of players for whom a_i is the strictly dominant action.

Let $NE(G)$ be the set of (mixed) Nash equilibria of G. A *penal code* is a profile of Nash equilibria, $(\hat{s}(i))_{i=1}^n$, where $\hat{s}(i) \in NE(G)$ for each $i \in \{1, \ldots, n\}$.[6] We allow $\hat{s}(i) = \hat{s}(j)$ for some i and $j \neq i$.

Given a penal code $(\hat{s}(i))_{i=1}^n$, let $E_1 \subseteq A$ be the set of all action profiles $a \in A$ such that for some player i,

(1-i) $D(a) = \{i\}$,

(1-ii) for all $j \neq i$, $BR_j(a) \cap BR_j(\hat{s}(i)) = \emptyset$, and

(1-iii) there exist $a'_i \in B_i(a)$ and $\zeta \in (0, 1)$ such that for all $j \neq i$,

$$a_j \in BR_j((1 - \zeta)a_i + \zeta a'_i, a_{-i}). \tag{1}$$

Next, let $E_2 \subseteq A$ be the set of all action profiles $a \in A$ such that for some players i and $j \neq i$,

(2-i) $\{i, j\} \subseteq D(a)$,

[6] This terminology follows Abreu (1988), although our use of the term is slightly different: while Abreu (1988) used the term for a profile of *repeated game* equilibria, we use it for a profile of *stage game* equilibria.

(2-ii) there exist $a'_i \in B_i(a)$ and $a'_j \in B_j(a)$ such that for all $k \in \{i, j\}$,

$$\{a_k, a'_k\} \cap \Big[BR_k(\hat{s}(i)) \cup BR_k(\hat{s}(j))\Big] = \emptyset,$$

(2-iii) for all $k \notin \{i, j\} \cup SD(a)$, $a_k \notin BR_k(\hat{s}(i)) \cap BR_k(\hat{s}(j))$.

For a given $a \in E_2$, players i and j for whom (2-i)–(2-iii) hold are called *associated players*. For a given $a \in E_2$, there may exist more than one pair of associated players, but we select a pair $\{i, j\}$ arbitrarily for each $a \in E_2$ and denote the selected pair by $AP(a)$. Similarly, for each $a \in E_1$ and i such that $\{i\} = D(a)$, we let $AP(a) = \{i\}$. Let $E = E_1 \cup E_2$.

Given a penal code $(\hat{s}(i))_{i=1}^n$, we say that a payoff vector $v = (v_i)_{i=1}^n \in \mathbb{R}^n$ is *supportable* with respect to $(\hat{s}(i))_{i=1}^n$, if there exists a probability distribution on E, denoted $(\rho(a))_{a \in E}$, such that

(s-i) for any i, $v_i = \sum_{a \in E} \rho(a) u_i(a)$,

(s-ii) for any i, if there exists $a \in E_1$ such that $\rho(a) > 0$ and $D(a) = \{i\}$, then $v_i > u_i(\hat{s}(i))$,

(s-iii) for any i, if there exists $a \in E_2$ such that $\rho(a) > 0$ and $i \in D^w(a)$, then $v_i > u_i(\hat{s}(i))$, and

(s-iv) for any $a \in E_2$ such that $\rho(a) > 0$ and any $k \in SD(a)$, if there exists $\hat{a} \in E$ such that $\rho(\hat{a}) > 0$ and $\hat{a}_k \neq a_k$, then there exists such an $\hat{a}$ that satisfies either $\hat{a} \in E_2$ or $[\hat{a} \in E_1(k)$ and $\hat{a}'_k \neq a_k]$ where $\hat{a}'_k$ is a better reply that satisfies (1-iii) with respect to $\hat{a}$.

Note that in (s-iii), i is not required to be an associated player. Condition (s-iv) says that if there exists a player k who plays her dominant action in some $a \in E_2$ in the support of ρ but does not play it in some $\hat{a} \in E$ in the support, then either there exists such an $\hat{a}$ in E_2, or there exists such an $\hat{a} \in E_1(k)$ such that an associated better reply that satisfies (1-iii) is not the dominant action of the player. This technical condition is irrelevant for many cases. For example, if no one plays her dominant action (if any) in the support of ρ, then (s-iv) is trivially satisfied. Note also that if two or more players have a dominant action in the stage game, then E_1 is empty by (1-ii) and therefore (s-iv) holds for any ρ.

Let $V^* \subseteq V$ denote the set of supportable payoff vectors with respect to a given penal code $(\hat{s}(i))_{i=1}^n$.[7]

[7] When the penal code is understood, we simply call V^* the set of supportable payoff vectors.

4 A Characterization of Equilibrium Payoff Vectors

We are now ready to state our main result, which gives a sufficient condition for a given payoff vector to be approximated by a sequential equilibrium when the observation costs $(\lambda_i)_{i=1}^n$ are sufficiently small and the discount factors $(\delta_i)_{i=1}^n$ are sufficiently close to one.

Proposition 1 *Let $(\hat{s}(i))_{i=1}^n$ be a penal code and V^* be the set of supportable payoff vectors with respect to the penal code. Then for any $v \in V^*$ and any $\epsilon > 0$, there exist $\bar{\lambda} = (\bar{\lambda}_i)_{i=1}^n \in \mathbb{R}_{++}^n$ and $\underline{\delta} = (\underline{\delta}_i)_{i=1}^n \in (0,1)^n$ such that, for any game $\Gamma(\delta, \lambda)$ with $\delta \geq \underline{\delta}$ and $\lambda \leq \bar{\lambda}$, there exists a sequential equilibrium σ^* that satisfies $|g_i(\sigma^*) - v_i| < \epsilon$ for any i.*

Proof. See the Appendix.

While the proof in the Appendix provides a general construction of an equilibrium that approximates a given supportable payoff vector, we here give its main idea, restricting ourselves to special examples of supportable payoff vectors. Let us begin with the simplest case, which is to approximate a payoff vector that is equal to $u(a)$ for some $a \in E_2$ such that $D^w(a) = \{1, 2, \ldots, n\}$. Since $a \in E_2$, there exist a pair of associated players $\{i, j\} = AP(a)$ and better replies $\{a_i', a_j'\}$ of them such that (2-i)–(2-iii) hold. Since $D^w(a) = \{1, 2, \ldots, n\}$, supportability implies $v_k > u_k(\hat{s}(k))$ for all k. Thus the penal code Nash equilibrium for player k, $\hat{s}(k)$, indeed makes k suffer.

To simplify our exposition, we call action a_k "cooperation," action a_k' given by (2-ii) "minor-deviation," and any other action "major-deviation." Note that only the associated players $\{i, j\} = AP(a)$ can minor-deviate.

We construct a strategy profile that uses $n + 1$ states. The set of states is $\{0, 1, \ldots, n\}$. State 0 is regarded as the cooperative state, in which (i) each player $k \in \{i, j\}$ randomizes between cooperation and minor-deviation, where the probability of cooperation is sufficiently close to 1, (ii) all other players $k \notin \{i, j\}$ cooperate, and (iii) all players monitor the other players. On the other hand, in state $i \geq 1$, player i is punished; the players play $\hat{s}(i)$ and no one monitors the others.

We now specify the rule that governs the transition of states. The initial state is 0 (cooperative state). If the state is 0 in period $t - 1$, then period t is in

(i) state k if player $k \in \{1, \ldots, n\}$ is the only player who major-deviated,

(ii) state k if player $k \in \{i, j\}$ minor-deviated and all other players cooperated, and

(iii) state 0 otherwise.

For any $k \geq 1$, if the state moves to k because of a unilateral major-deviation of player k (case (i)), then the state remains k for all subsequent periods. On the other hand, if the state moves to k because of a unilateral minor-deviation of player $k \in \{i, j\}$ (case (ii)), then state remains k for a certain number of periods and then moves back to 0. The length of the k-state periods is set so that the gain from the minor-deviation is exactly equal to the loss from playing $\hat{s}(k)$. Since $u_k(a) > u_k(\hat{s}(k))$, the appropriate length of the k-state periods can be found if the players are sufficiently patient. Since the appropriate length is not necessarily an integer, the public randomization device is used to make the transition from state k to 0 contingent on sunspots. Moreover, since when the state moves back to 0 depends only on sunspots, the state is common knowledge although the players do not monitor each other during the punishment periods.

This specification is sufficient to determine what happens on the path. Since the players cooperate with a probability sufficiently close to 1, the path approximates the payoff vector $u(a)$ as long as the observation costs are sufficiently small. Note that the above specification also determines the continuation play at off-the-path histories if the player did not deviate in terms of monitoring in the previous periods (since then she knows the state). To define the equilibrium strategy formally, it remains to specify how a player behaves after she deviates in terms of monitoring. However, since this specification does not affect the following argument, we do not complete the specification of strategy here.

Let us now examine the incentive to follow the state-dependent play described above. First, for δ sufficiently close to the unit vector, no player has an incentive to major-deviate in state 0. This is because $u_k(a) > u_k(\hat{s}(k))$ for all k and once player k major-deviates, the resulting outcome is the perpetual play of $\hat{s}(k)$. Second, players $k \in \{i, j\}$ are indifferent between cooperation and minor-deviation because of the way in which the number of k-state periods is set. Third, in state $k \geq 1$, no player has an incentive to deviate in terms of action or monitoring since (i) a stage game Nash equilibrium is played, (ii) the play does not affect the transition of the state, and (iii) no monitoring is required.

The remaining step is to examine the monitoring incentive in state 0. We start with players $k \notin \{i, j\}$. Suppose that the state was 0 in period t and player $k \notin \{i, j\}$ did not monitor at the end of the period. Then, in period $t+1$, she is uncertain about the state, which is either 0, i, or j depending on whether i or j (or both) minor-deviated in period t. By (2-iii), playing a_k in period $t+1$ is not optimal if the state is either i or j. On the other hand, if player k plays an action other than a_k in period $t+1$, the action is considered as a major-deviation and triggers a perpetual punishment if the state is actually 0. Therefore, if λ_k is sufficiently small, the gain from eliminating the uncertainty exceeds the cost of monitoring.

Let us now consider a player $k \in \{i, j\}$. Without loss of generality, let $k = i$. Suppose that the state was 0 in period t and player i did not monitor at the end of the period, so i is uncertain about the state in period $t+1$. First, consider the case in which player i cooperated in period t. Then the state in period $t+1$ is either j or 0 depending on the action of player j in period t. Thus, if i cooperates or minor-deviates in period $t+1$, then by (2-ii), the action is suboptimal if the state is j. On the other hand, if i major-deviates in period $t+1$, a perpetual punishment follows if the actual state is 0. Hence, player i suffers strictly from the uncertainty and it is optimal for her to eliminate the uncertainty if her monitoring cost is sufficiently small.

Let us now consider the case when player i minor-deviated in period t. Then the state in period $t+1$ is either i or 0; the latter occurs if j also minor-deviated. Since the latter case occurs with a small probability, the state in period $t+1$ is almost surely i, so by (2-ii), it is suboptimal for i to either cooperate or minor-deviate in period $t+1$. However, if she chooses an action other than cooperation and minor-deviation in period $t+1$, then it is regarded as a major-deviation if the state in this period is actually 0, which occurs with a small but positive probability. Thus, player i suffers strictly from the uncertainty, which she is willing to avoid if her monitoring cost is sufficiently small.

In this way, we can prove that it is not profitable for players to deviate in terms of monitoring (on the path). This together with the previous arguments shows that the state-dependent play is an equilibrium when the players are patient and monitoring costs are small.

It is less straightforward to approximate other supportable payoff vectors. For example, let us consider a payoff vector that is equal to $u(a)$ for some $a \in E_1$. By (1-i), only one player has a short-run incentive to deviate from a. The state-dependent play described above cannot be used since it requires two players to minor-deviate (to give monitoring incentives to each of them). Therefore, we consider a different type of behavior in this case. Specifically, in the cooperative state, the player i such that $D(a) = \{i\}$ randomizes between cooperation and minor-deviation and does not monitor the other players, while all other players cooperate and monitor the others. The state transition is specified similarly. Then, all agents $j \neq i$ have a monitoring incentive in the cooperative state since the future play is either to cooperate or to punish i and the state transition depends on i's action. On the other hand, since the state transition depends only on i's action, i can identify the current state even if she does not monitor the other players. Thus, at equilibrium, the state is common knowledge among the players although not all players observe the past actions.

The construction of an equilibrium is more complicated if the payoff vector to be approximated can be generated only by a randomization among elements of E_1 and E_2, or when some players play a dominant action in the cooperative stage. We deal with these cases in the Appendix.

A final remark on Proposition 1 is that if the players use sunspots wisely, many other payoff vectors can be approximated. Let $NE^*(G)$ be the set of Nash equilibrium *payoff vectors* of G. Then it is easily seen that any payoff vector in the convex hull of $V^* \cup NE^*(G)$ can be approximated. Moreover, as we vary the penal code $(\hat{s}(i))_{i=1}^n$, we obtain different V^* and therefore different $V^* \cup NE^*(G)$, and all elements of those sets can be approximated. Thus, if G has a number of Nash equilibria, the set of payoff vectors that our construction can approximate can be large.[8] In the next section, we demonstrate that the set is indeed large and yields an approximate Folk Theorem.

5 Application: Approximate Folk Theorem

This section examines three examples and shows that Proposition 1 generates an approximate Nash Folk Theorem in each of the examples. In these examples, we consider a penal code in which the same Nash equilibrium is used for all players. Denoting the stage Nash equilibrium by $\hat{s}$, we will show that all efficient payoff vectors that Pareto-dominate $u(\hat{s})$ are supportable with respect to $\hat{s}$. Then, Proposition 1 proves that all those payoff vectors are approximated by equilibria if the monitoring costs are sufficiently small. Since sunspots are available, all interior payoff vectors that Pareto-dominate $u(\hat{s})$ are also attainable as equilibria. In this way, we obtain an approximate Nash Folk Theorem. A minimax version of approximate Folk Theorem may also hold if, in addition, $u_i(\hat{s})$ is the minimax value of player i for all i. We indeed obtain an approximate minimax Folk Theorem in the example of

[8] Furthermore, there may be payoff vectors that can be supported by other strategy profiles than the ones we consider in this paper.

linear partnership examined below.

5.1 A Variant of Prisoners' Dilemma

We begin our discussion with the following standard prisoners' dilemma.

	C	D
C	$1, 1$	$-1, 2$
D	$2, -1$	$0, 0$

If this is the stage game, then our construction of strategy profile cannot support cooperation. Indeed, since the Nash equilibrium is unique, the only possible penal code is $\hat{s}(1) = \hat{s}(2) = (D, D)$. Then, (C, C) violates (1-i) and (2-ii). Condition (2-ii) is simply impossible to satisfy if there are only two actions. Similarly, (C, D) and (D, C) violate (1-ii) and (2-i). Thus E_1 and E_2 are empty for the prisoners' dilemma.

On the other hand, the result changes considerably if the stage game has a slightly larger action set. For prisoners' dilemma, our construction can easily support cooperation if each player has another action. This is illustrated by the following stage game.

	C	D	E
C	$1, 1$	$-1, 2$	$-1, -1$
D	$2, -1$	$0, 0$	$-1, -1$
E	$-1, -1$	$-1, -1$	$0, 0$

This is a simplified version of the bilateral trade game with moral hazard in Bhaskar and van Damme (2002). This game has two pure Nash equilibria, (D, D) and (E, E), as well as a mixed Nash equilibrium, $\underline{s}$, where each player randomizes between D and E with equal probabilities. Clearly, C is strictly dominated, and (C, C) Pareto-dominates all Nash equilibria.

We present an approximate Nash Folk Theorem for the expanded prisoners' dilemma. We set $\hat{s}(1) = \hat{s}(2) = (E, E)$ as a penal code. Then, since neither C nor D is a best response to E, we have $(C, C) \in E_2$. Furthermore, since D is a unique best response to C, we also have $(C, D) \in E_1$ and $(D, C) \in E_1$. Since no player has a dominant action, Condition (s-iv) of supportability holds trivially. Since (C, C), (D, C), and (C, D) are the only efficient action profiles, any efficient payoff vector that Pareto-dominates $(0, 0)$ is supportable and therefore approximated by an equilibrium. Thus an approximate Nash Folk Theorem holds.

However, this is not an approximate *minimax* Folk Theorem since the minimax value in this game is $-1/2$ for each player. The fact that the minimax value is attained by the mixed Nash equilibrium $\underline{s}$ does not enable us to prove a minimax Folk Theorem. Indeed, if we set $\hat{s}(1) = \hat{s}(2) = \underline{s}$ as a penal code, then E_1 and E_2 are both empty, and so is the set of supportable payoff vectors with respect to the penal code. This argument also demonstrates that the set of supportable payoff vectors depends on the penal code.

5.2 Linear Partnership Games

This subsection further explores the idea that our construction of strategy profile can support cooperation if the action space is sufficiently rich. We consider a class of linear partnership games where each game is parameterized by the richness of action set. The assumption of linearity plays an important expositional role; it ensures that the set of feasible payoff vectors, V, does not depend on the richness of action set. At the cost of complication, the idea can be extended to more general games of partnership.

The linear n-player partnership game is defined as follows. There are n players, and each player has $m+1$ actions where $m \geq 2$. The set of actions for each player is $A_i = \{0, 1/m, \ldots, (m-1)/m, 1\}$. The production function is linear and given by $f(a) = \sum_{i=1}^{n} a_i$. Let $c_i(a_i) = \alpha a_i$ be the cost that player i has to pay if she chooses a_i. The output is divided equally among the players. Player i's payoffs are therefore $u_i(a) = (1/n)\sum_{k=1}^{n} a_k - \alpha a_i$. We impose the non-triviality assumption that $1/n < \alpha < 1$. This implies that $a_i = 0$ is a dominant action for each player, while $(1, 1, \ldots, 1)$ is the efficient action profile. Hence, this game is also a variant of prisoners' dilemma. Note also that any $a_i \geq 1/m$ is strictly dominated by $a_i - (1/m)$. The minimax value for each player is 0, which is attained in the unique Nash equilibrium $s^0 = (0, \ldots, 0)$, independently of m.

Since the partnership game has a unique Nash equilibrium, the only possible penal code is $\hat{s}(i) = s^0$ for all i. With respect to the penal code, $E_1 = \emptyset$ by (1-ii). This implies that (s-iv) holds trivially. On the other hand, E_2 is characterized as follows.

Proposition 2 $E_2 = \{a \in A : \exists i, j \neq i \text{ s.t. } \min\{a_i, a_j\} \geq 2/m\}$.

Proof. Let $a \in A$ be such that $\min\{a_i, a_j\} \geq 2/m$ for some i and $j \neq i$. Then, $\{i, j\} \subseteq D(a)$, and for all $k \in \{i, j\}$, $1/m \in B_k(a) \setminus \{0\}$. For all $k \notin \{i, j\} \cup SD(a)$, $a_k \geq 1/m$ and hence $a_k \notin BR_k(s^0)$. Therefore (2-i)–(2-iii) hold and $a \in E_2$.

To prove the converse, let $a \in E_2$. Then there exist associated players i and $j \neq i$ for whom there exist $a'_i \in B_i(a) \setminus \{0\}$ and $a'_j \in B_j(a) \setminus \{0\}$. Hence

$\min\{a_i, a_j\} \geq 2/m$. Q.E.D.

Let $\bar{V}$ be the boundary of V, and $V_{IR} \subseteq V$ be the set of feasible payoff vectors that are strictly individually rational, i.e., $V_{IR} = \{v \in V : v_i > 0 \text{ for all } i\}$. Since payoff functions are linear, V, $\bar{V}$, and V_{IR} are all independent of m. The following result proves that all feasible, boundary, and strictly individually rational payoff vectors are supportable if m is sufficiently large. In view of Proposition 1 and the availability of sunspots, the result implies that if m is large, any $v \in V_{IR}$ can be approximated by an equilibrium. Therefore, we have an approximate minimax Folk Theorem.

Proposition 3 *If $m \geq 2/(n\alpha - 1)$, any $v \in V_{IR} \cap \bar{V}$ is supportable.*

Proof. Assume $m \geq 2/(n\alpha - 1)$ and let $v \in V_{IR} \cap \bar{V}$. Then there exists a probability distribution on A, $(\rho(a))_{a \in A}$, such that $v = \sum_{a \in A} \rho(a) u(a)$. Let $\rho_i^* = \sum_{a \in A} \rho(a) a_i$, which is the expected action level of player i. Since $v \in V_{IR} \cap \bar{V}$, there exists a player i such that $\rho_i^* = 1$ (otherwise, a Pareto improvement can be achieved by multiplying everyone's expected action level by some $\beta > 1$). Without loss of generality, we assume $\rho_1^* = 1$. If $\sum_{k \geq 2} \rho_k^* < 2/m$, then the expected utility of player 1 is

$$v_1 = (1/n) \sum_{i=1}^{n} \rho_i^* - \alpha < (1/n)(1 + 2/m) - \alpha \leq 0,$$

where the last inequality follows from $m \geq 2/(n\alpha - 1)$. The inequalities imply that v is not strictly individually rational, a contradiction. Thus $\sum_{k \geq 2} \rho_k^* \geq 2/m$.

We have to show that there exists a probability distribution over E_2 that generates payoff vector v. Since payoff functions are linear, it suffices to prove that the convex hull of E_2 includes $\rho^* = (\rho_i^*)_{i=1}^n$. To prove this, let β^H and β^L be defined by

$$\beta^H = 1/(\max_{k \geq 2} \rho_k^*) \geq 1, \tag{2}$$

$$\beta^L = (2/m)/\sum_{k \geq 2} \rho_k^* \leq 1, \tag{3}$$

where the inequality in (3) is proved in the previous paragraph. Let $\rho^H, \rho^L \in [0,1]^n$ be defined by $\rho_1^H = \rho_1^L = 1$, and for all $k \geq 2$, $\rho_k^H = \beta^H \rho_k^*$ and $\rho_k^L = \beta^L \rho_k^*$. Clearly, ρ^* is a convex combination of ρ^H and ρ^L. By (2), ρ^H has at least two components of 1. Thus, it follows easily from Proposition 2 that ρ^H is in the convex hull of E_2. We now prove that ρ^L is also in the convex hull of E_2. For each $k \geq 2$, let $a^k \in A$ be the action profile defined by $a_1^k = 1$, $a_k^k = 2/m$, and $a_j^k = 0$ for all $j \notin \{1, k\}$. By Proposition 2, $a^k \in E_2$ for each $k \geq 2$. By (3), ρ^L is a convex combination of $(a^k)_{k \geq 2}$ where the weight assigned to a^k is $\rho_k^* / (\sum_{j \geq 2} \rho_j^*)$. Q.E.D.

5.3 Games with Utility Burning

The objective of this subsection is to demonstrate that an approximate Nash Folk Theorem holds if the players are able to "burn" small amounts of their own payoffs.

Let a stage game $G = \{n, A, (u_i)_{i=1}^n\}$ be given. For a given number $z > 0$, we define the *game with z-utility burning* as $G^z = \{n, A', (u_i')_{i=1}^n\}$ where $A_i' = A_i \times \{0, 1, 2\}$ for each i, and for any action profile $a' = (a_i, k_i)_{i=1}^n \in A'$,

$$u_i'(a') = u_i(a) - k_i z.$$

In this game, each player chooses an action and *at the same time* chooses the amount of her payoffs to burn. It is assumed that a player can decrease her payoffs without affecting the others'. We also assume that if a player monitors the other players, she learns the amounts of payoffs that the other players burnt.

It is easily seen that none of the Nash equilibria in G^z involves utility burning. Note also that G and G^z have the same Pareto frontier. Moreover, if we define $V^z = \mathrm{co}\{u'(a') : a' = (a_i, 2)_{i=1}^n\}$, then V^z converges to V as $z \to 0$. Let us fix a penal code (in G^z), $(\hat{s}'(i))_{i=1}^n$, arbitrarily and consider an action profile $a' \in A'$ of the form $a' = (a_i, 2)_{i=1}^n$. Then for all i, we have $(a_i, 1) \in B_i(a')$ and for all $j \in \{1, \ldots, n\}$ and all $k \in \{1, 2\}$, $(a_i, k) \notin BR_i(\hat{s}'(j))$, which implies $a' \in E_2$. Therefore, any $v \in V^z$ that Pareto-dominates $(u_i(\hat{s}'(i)))_{i=1}^n$ is supportable.[9] Hence, if the unit of utility burning, z, is small, an approximate Nash Folk Theorem holds.[10] Note that this result holds for any game G.

6 Concluding Remarks

This section discusses possible extensions of our model.

[9]Since v is represented by a convex combination of elements of E_2, Conditions (s-ii) and (s-iv) hold trivially.

[10]However, to sustain cooperation, the observation costs have to be small in comparison with the already small level of utility burning.

6.1 Fixed Observation Costs

An important assumption in our characterization of equilibrium payoff vectors and approximate Folk Theorems is that the observation costs are sufficiently small. The results say nothing if the levels of observation costs are fixed. A simple, alternative framework in which we can deal with fixed levels of observation costs is one in which monitoring at the end of a period gives information about not only the present period but all the previous periods. This framework is a variation of that in Miyahara (2002), who examines the case when at least the last *two* periods can be observed. However, Miyahara's efficiency result for repeated prisoners' dilemma also requires small observation costs.

When all previous periods are observable (with costs), we can use Miyahara's construction of strategy profile to support a large set of payoff vectors for fixed observation costs. To see this, assume that there exists an action profile $\hat{a}$ that attains a given target payoff vector. Let us also assume the existence of an action profile a in which there exist at least two potential deviators, i.e., $|D(a)| \geq 2$. As in our construction, select two players $\{i, j\} \subseteq D(a)$ and call them the associated players. Then consider the following strategy profile for a given $T \in \{2, 3, \ldots\}$: (i) the players play $\hat{a}$ in the first $T-1$ periods without monitoring each other; (ii) in period T, the players play a, except that the associated players mix between a and their minor-deviations, and all players monitor; (iii) the play in the next T periods is either another sequence of (i) and (ii), or a repetition of a penal code Nash equilibrium, depending on the presence of a deviator in the first T periods, and so on.

Under the strategy profile, the players do not monitor the other players in $\hat{a}$ state. But they have no incentive to deviate in terms of actions since deviations are detected at least T periods later and regarded as major-deviations. The incentive for monitoring in period T is guaranteed if for each player k, $\hat{a}_k$ is not a best response to the penal code Nash equilibria designed for the associated players. Under this condition, the above strategy profile constitutes an equilibrium and approximates the target payoff vector for a *given* vector of observation costs if T is sufficiently large and discount factors are close to 1.[11]

[11] In the strategy profile, the action profile is the same for the first $T-1$ periods. Alternatively, we could consider a strategy profile in which the action profile during these periods is time-dependent. The advantage of using the larger class of strategy profiles is that the corresponding condition on the relation between "$\hat{a}$" and the penal code can be weakened considerably.

This construction for multi-period observation technology can sustain cooperation even for stage games whose action sets are small. Indeed, the stage game examined in Miyahara (2002) is the standard two-action prisoners' dilemma and he obtains an efficiency result for the game.

In our future research, we will elaborate the strategy profile (along the line mentioned in footnote 11) to obtain a characterization of payoff vectors that can be approximated by equilibria and derive conditions on stage games for which a Folk Theorem with fixed observation costs holds.

6.2 Timing of Receiving Payoffs

We have assumed that the players do not receive payoffs in each period but they receive the total payoffs when the game ends. We need these assumptions in order to keep consistency with the assumption that, without paying monitoring costs, a player receives *no* information about the others' actions. If payoffs are received in every period, then they generally provide players with some information about the other players' actions.

However, we can imagine a framework in which payoffs are received in every period but monitoring remains important because realized payoffs are only a noisy signal of the other players' actions. For example, let a stage game $G = \{n, A, (u_i)_{i=1}^n\}$ be given. Suppose that at the end of each period, player i receives payoffs of $u_i(a) + \epsilon_i$ if $a \in A$ is played in that period, where ϵ_i is a noise term which follows a normal distribution with mean 0. Assume also that the noise terms are independent across the players.

In this formulation, the realized payoff is not a sure indicator of the other players' actions while it is informative. If we ignore the issue of costly observation, the standard model of repeated games with imperfect private monitoring (like Sekiguchi (1997)) falls into this category if the realized payoff is a sufficient statistic of the privately observed signal about the other players' actions.

Even in this framework, we can use the state-dependent strategy profile considered in Section 4. Under this strategy profile, players monitor each other and do not use the information contained in the realized payoffs. The monitoring incentive is weaker under this strategy profile since realized payoffs also give players information about the state. However, players who do not pay observation costs are not able to determine the state with certainty. Therefore, if the likelihood ratio of any pair of action profiles that generate the same level of payoffs is bounded away from zero, then the players do have an incentive to pay observation costs, given any payoff realization, if the observation costs are sufficiently small.[12] Hence the basic idea of our construction also applies to the case in which payoffs are received in each period.

This observation is important because it suggests a possibility that costly observation is one comprehensive solution to the private monitoring problem. The literature of repeated games with imperfect private monitoring has shown difficulty in constructing a cooperative/efficient equilibrium, and the existing positive results (reported in the papers cited in footnote 2) are limited to simple specific games, e.g., repeated prisoners' dilemma and its variations.[13] It is still unknown whether a Folk Theorem or an efficiency result holds in general settings with private monitoring. In contrast, our result and the above discussions show that an approximate Folk Theorem does hold in general environments if the players have an ability to observe the other players' actions directly and the observation costs are sufficiently small.

The literature also identifies communication among the players as a driving force to cooperation in general environments with private monitoring (Compte (1998), Kandori and Matsushima (1998), and Aoyagi (2002)).[14] Thus our analysis may as well be seen as demonstrating that costly observation is a convenient substitute for communication. This interpretation has a strong implication on antitrust laws since they control communication among firms in the belief that communication is a major tool that facilitates cartels.

6.3 Partial Monitoring

The monitoring activity that we have considered has a binary aspect in that each player has to decide whether to obtain *complete* information about the action profile of the other players in the period or to obtain *no* information.

A more realistic formulation is that each player can choose to what extent she observes the other players' actions, and the more she spends for

[12] Precisely speaking, the likelihood ratio is not bounded away from zero when the noise terms are normally distributed. However, the likelihood ratio is close to zero only at the tails. Hence, we conjecture that there exists a cooperative equilibrium in which the players monitor each other unless the realized payoffs take extreme values. Moreover, the likelihood ratio condition can be satisfied for other specifications of the noise term.

[13] Mailath and Morris (2002) consider more general stage games, but they assume that private signals are correlated across the players. Amarante (2002) also conducts a general analysis. For some negative results, see Matsushima (1991) and Compte (2002).

[14] See also Ben-Porath and Kahneman (1996) for the role of communication in related environments with private monitoring.

monitoring, the more information she obtains. For example, suppose that λ_i is the *unit* monitoring cost and player i incurs the cost of $m\lambda_i$ if she monitors m of the other players. This alternative framework is relevant in the price-setting oligopoly if the goods are sold at each firm's own outlet. In this framework, each firm decides the set of firms to monitor and the total observation cost depends on the number of firms to monitor. Such *partial monitoring* is relevant even in the case of duopoly if the firms operate in multiple markets. In this case, each firm decides the set of markets to monitor, and the total observation cost depends on the number of markets to monitor. Thus the price-setting oligopoly is a prominent example of partial monitoring since the firms often compete in a large number of dimensions. As Stigler (1964) concluded, collusion is hard to implement since it requires an ability to detect any possible secret price-cuts in any market.

In general, partial monitoring is relevant whenever the action profile of $n-1$ players is multi-dimensional (this is trivially the case when $n \geq 3$) and a player can choose to observe only a subset of the coordinates in the profile of the other players' actions. A basic difficulty in analyzing the case when partial monitoring is feasible is that the players have an incentive to economize on observation expenses by not monitoring some of the players (or markets). In the strategy profile used in our proof, some of the players do not randomize in the cooperative state, but this is not a problem in the proof since these players are also monitored by the other players. Under the binary nature of our monitoring technology, any player who has an incentive to monitor at least one of the other players has no choice but to monitor all other players. However, if partial monitoring is feasible, players would monitor only those who randomize, but then deviations of non-randomizing players are not detected. Therefore, if partial monitoring is feasible, the cooperative equilibria that we constructed are upset.

Nevertheless, our construction can be modified to deal with partial monitoring if the payoff vector to be approximated can be generated by action profiles in which *all* players have proper minor-deviations. Formally, let a stage game G and a penal code $(\hat{s}(k))_{k=1}^{n}$ be given. Let $E_n \subseteq A$ be the set of all $a \in A$ such that

(n-i) $D(a) = \{1, 2, \ldots, n\}$,

(n-ii) for each player i, there exists $a'_i \in B_i(a)$ such that

$$\{a_i, a'_i\} \cap \left[\cup_{k=1}^{n} BR_i(\hat{s}(k))\right] = \emptyset.$$

It is then not very difficult to see that for all $a \in E_n$, if $u_i(a) > u_i(\hat{s}(i))$ for all i, then $u(a)$ can be approximated by an equilibrium, in which all players randomize between a_i and a_i' in the cooperative state. Thus, any convex combination of such $u(a)$'s can be also approximated. We have seen in Subsection 5.2 that the finer the action set is, the more payoff vectors are approximated using action profiles where all players have short-run incentives to deviate. Therefore, our result extends to the case of partial monitoring when the underlying strategic situation involves sufficiently fine action sets.

This idea also applies to the case of duopoly with multiple markets. If a price profile is such that each firm has a short-run incentive to deviate *in every market*, then the price profile can be supported by an equilibrium where the firms randomize between cooperation and minor-deviation in every market. Again, if the price space is sufficiently fine, many levels of collusion can be sustained, so an approximate Folk Theorem will be obtained.

6.4 Finite Repetition

Assuming that the horizon is finite has both an advantage and a disadvantage. An advantage is that the finite horizon makes it easier to interpret the assumption that the payoffs are received in total at the end of the repeated game. A disadvantage is that the finite horizon makes cooperation unsustainable if the stage game has a unique equilibrium.

On the other hand, it might be possible to obtain an approximate Folk Theorem under a finite horizon if the stage game has multiple equilibrium payoffs for each player as in Benoit and Krishna (1985). We conjecture that if the number of periods is sufficiently large, an action profile that Pareto-dominates a stage-game equilibrium can be sustained in early periods. This is another topic of our future research.

Appendix: Proof of Proposition 1

Let $v \in V^*$ and $\epsilon > 0$. Since v is supportable, there exists a probability distribution on E, denoted $(\rho(a))_{a\in E}$, such that (s-i)–(s-iv) hold. Let us define $E_1^*(i) = \{a \in E_1 : \rho(a) > 0 \text{ and } D(a) = \{i\}\}$, $E_1^* = \cup_i E_1^*(i)$, $E_2^* = \{a \in E_2 : \rho(a) > 0\}$, and $E^* = E_1^* \cup E_2^*$. Thus, E^* is the support of ρ.

We can choose a sufficiently small $\epsilon > 0$ so that for all players i, if there exists $a \in E^*$ such that either $[a \in E_1$ and $D(a) = \{i\}]$ or $[a \in E_2$ and $i \in D^w(a)]$, then $v_i - \epsilon > u_i(\hat{s}(i))$. This is possible since v is supportable.

Let $SD^* = \{i : i \in SD(a) \quad \forall a \in E^*\}$ be the set of players i such that for all $a, \hat{a} \in E^*$, $a_i = \hat{a}_i$ and it is the strictly dominant action for i.

For each i and $a \in E_1^*(i)$, fix $a_i' \in B_i(a)$ such that (1-iii) and (s-iv) hold. Similarly, for each $a \in E_2^*$, we fix $a_i' \in B_i(a)$ and $a_j' \in B_j(a)$ that satisfy (2-ii), where $\{i, j\} = AP(a)$.

Since G is a finite game, there exists $\bar{\zeta} \in (0, 1)$ such that for all $a \in E_1^*$, (1-iii) holds with $\zeta = \bar{\zeta}$, and for all $a \in E_2$ and all $k \notin D^w(a)$,

$$\{a_k\} = BR_k((1-\bar{\zeta})a_i + \bar{\zeta}a_i', (1-\bar{\zeta})a_j + \bar{\zeta}a_j', a_{-i-j}), \tag{4}$$

where $\{i, j\} = AP(a)$. Note that these conditions also hold when $\bar{\zeta}$ is replaced with any $\zeta \in [0, \bar{\zeta})$.

Let $\eta \in [0, \bar{\zeta}]$, $\lambda_i \geq 0$, and $\delta_i \in (0, 1)$ for all i. For any $a \in E^*$ and any $k \in AP(a)$, we define a mixed action a_k^η as

$$a_k^\eta = (1-\eta)a_k + \eta a_k',$$

with the obvious interpretation that a_k^η assigns probability $1 - \eta$ to a_k and the remaining probability to a_k'. We now define the following vectors and numbers: $V^0 \in \mathbb{R}^n$, $V^0(a) \in \mathbb{R}^n$, $V^k(a) \in \mathbb{R}^n$, and $\nu^k(a) \in (0, 1)$ for any $a \in E^*$ and any $k \in AP(a)$. We define them as a solution of the following system:

$$V^0 = \sum_{a\in E^*} \rho(a) V^0(a); \tag{5}$$

$$V_i^0(a) = (1-\delta_i)u_i(a) + \delta_i V_i^0 \tag{6}$$

$$= (1-\delta_i)u_i(a_i', a_{-i}) + \delta_i V_i^i(a) \tag{7}$$

for any i and any $a \in E_1^*(i)$;

$$V_k^0(a) = (1-\delta_k)\big[u_k(a_i^\eta, a_{-i}) - \lambda_k\big] + \delta_k\big[(1-\eta)V_k^0 + \eta V_k^i(a)\big] \tag{8}$$

for any i, any $a \in E_1^*(i)$, and any $k \neq i$;

$$V^i(a) = [1 - \nu^i(a)]u(\hat{s}(i)) + \nu^i(a)V^0 \tag{9}$$

for any i and any $a \in E_1^*(i)$;

$$\begin{aligned} V_i^0(a) &= (1-\delta_i)\big[u_i(a_j^\eta, a_{-j}) - \lambda_i\big] + \delta_i\big[(1-\eta)V_i^0 + \eta V_i^j(a)\big] & (10)\\ &= (1-\delta_i)\big[u_i(a_i', a_j^\eta, a_{-i-j}) - \lambda_i\big] + \delta_i\big[(1-\eta)V_i^i(a) + \eta V_i^0\big] & (11)\end{aligned}$$

and

$$\begin{aligned} V_j^0(a) &= (1-\delta_j)\big[u_j(a_i^\eta, a_{-i}) - \lambda_j\big] + \delta_j\big[(1-\eta)V_j^0 + \eta V_j^i(a)\big] & (12)\\ &= (1-\delta_j)\big[u_j(a_j', a_i^\eta, a_{-i-j}) - \lambda_j\big] + \delta_j\big[(1-\eta)V_j^j(a) + \eta V_j^0\big] & (13)\end{aligned}$$

for any $a \in E_2^*$ and the associated players $\{i, j\} = AP(a)$;

$$\begin{aligned} V_k^0(a) = {} & (1-\delta_k)\big[u_k(a_i^\eta, a_j^\eta, a_{-i-j}) - \lambda_k\big] \\ & + \delta_k\Big\{\big[(1-\eta)^2 + \eta^2\big]V_k^0 + \eta(1-\eta)\big[V_k^i(a) + V_k^j(a)\big]\Big\} \end{aligned} \tag{14}$$

for any $a \in E_2^*$, any $k \notin AP(a) \cup SD^*$, and $\{i, j\} = AP(a)$;

$$\begin{aligned} V_k^0(a) = {} & (1-\delta_k)u_k(a_i^\eta, a_j^\eta, a_{-i-j}) \\ & + \delta_k\Big\{\big[(1-\eta)^2 + \eta^2\big]V_k^0 + \eta(1-\eta)\big[V_k^i(a) + V_k^j(a)\big]\Big\} \end{aligned} \tag{15}$$

for any $a \in E_2^*$, any $k \in SD^*$ (note that $AP(a) \cap SD^* = \emptyset$ by (2-i)), and $\{i, j\} = AP(a)$; and

$$V^k(a) = [1 - \nu^k(a)]u(\hat{s}(k)) + \nu^k(a)V^0 \tag{16}$$

for any $a \in E_2^*$ and any $k \in AP(a)$.

Regarding system (5)–(16), we prove the following lemma.

Lemma 1 *There exist $\hat{\eta} \in (0, \bar{\zeta}]$, $\hat{\lambda} > 0$, $\hat{\kappa} > 0$, and $\hat{\delta} \in (0, 1)$ such that*

(i) System (5)–(16) has a solution if $\eta \leq \hat{\eta}$, $\lambda_i \leq \hat{\lambda}$, and $\delta_i \geq \hat{\delta}$ for any i.

(ii) The solution satisfies

$$|V_i^0 - v_i| < \epsilon \tag{17}$$

for any i, and

$$\frac{1 - \nu^k(a)}{1 - \delta_k} > \hat{\kappa} \tag{18}$$

for any $a \in E^*$ and any $k \in AP(a)$.

Proof. We first examine system (5)–(16) when $\eta = \lambda_i = 0$ for any i. By (5), (6), (8), (10), (12), (14), and (15), any solution satisfies

$$\begin{aligned} V_i^0 &= v_i \\ V_i^0(a) &= (1-\delta_i)u_i(a) + \delta_i v_i \end{aligned} \tag{19}$$

for any $a \in E^*$ and any i. Therefore, (6), (7), and (9) reduce to

$$(1-\delta_i)\big[u_i(a_i', a_{-i}) - u_i(a)\big] = \delta_i(1-\nu^i(a))\big[v_i - u_i(\hat{s}(i))\big] \tag{20}$$

for any i and $a \in E_1^*(i)$. Since v is supportable, $v_i > u_i(\hat{s}(i))$. Since $a_i' \in B_i(a)$, (20) has a unique solution $\nu^i(a) \in (0,1)$ if δ_i is sufficiently close to 1. By (9) and (19), this solution uniquely determines $V^i(a)$. Furthermore, (20) yields

$$\frac{1-\nu^i(a)}{1-\delta_i} > \frac{u_i(a_i', a_{-i}) - u_i(a)}{v_i - u_i(\hat{s}(i))}. \tag{21}$$

Similarly, for any $a \in E_2^*$, (10)–(13) and (16) reduce to

$$(1-\delta_k)\big[u_k(a_k', a_{-k}) - u_k(a)\big] = \delta_k(1-\nu^k(a))\big[v_k - u_k(\hat{s}(k))\big] \tag{22}$$

for all $k \in \{i,j\} = AP(a)$. Since $\{i,j\} \subseteq D(a)$ and v is supportable, $v_k > u_k(\hat{s}(k))$ for all $k \in \{i,j\}$. Since $a_k' \in B_k(a)$ for all $k \in \{i,j\}$, (22) has a unique solution $(\nu^i(a), \nu^j(a)) \in (0,1)^2$ if δ_i and δ_j are both sufficiently close to 1. By (16), the solution uniquely determines $V^i(a)$ and $V^j(a)$. Moreover, (22) yields

$$\frac{1-\nu^k(a)}{1-\delta_k} > \frac{u_k(a_k', a_{-k}) - u_k(a)}{v_k - u_k(\hat{s}(k))} \tag{23}$$

for all $k \in AP(a)$.

Hence, if $\eta = \lambda_i = 0$ for all i, then there exists $\hat{\delta} \in (0,1)$ such that if $\delta_i \geq \hat{\delta}$ for all i, then system (5)–(16) has a unique solution satisfying (17). Moreover, if we choose

$$\hat{\kappa} = (1/2) \min_{\substack{a \in E^* \\ k \in AP(a)}} \Big[\frac{u_k(a_k', a_{-k}) - u_k(a)}{v_k - u_k(\hat{s}(k))}\Big] > 0,$$

then (21) and (23) imply that (18) holds for all $a \in E^*$ and all $k \in AP(a)$.

By the standard continuity argument, there exist $\hat{\eta} \in (0, \bar{\zeta}]$, $\hat{\lambda} > 0$, and $\hat{\delta} \in (0,1)$ such that if $\eta \leq \hat{\eta}$, $\lambda_i \leq \hat{\lambda}$, and $\delta_i \geq \hat{\delta}$ for all i, then system (5)–(16) has a nearby solution satisfying (17) and (18). Q.E.D.

For any player k, if $BR_k(\hat{s}(i)) \neq A_k$ for some i, then let

$$p_k = \min\bigl\{u_k(\hat{s}(i)) - u_k(a_k, \hat{s}_{-k}(i)) : \\ i \in \{1, \ldots, n\} \text{ and } a_k \notin BR_k(\hat{s}(i))\bigr\} > 0.$$

That is, any player k who plays a suboptimal action against a penal code Nash equilibrium loses at least p_k in terms of per-stage payoffs.

For any player k, if $k \notin D(a)$ for some $a \in E_1^*$, then let

$$r_k = \min\bigl\{u_k(a) - u_k(\hat{a}_k, a_{-k}) : \\ a \in E_1^*,\ k \notin D(a), \text{ and } \hat{a}_k \notin BR_k(a)\bigr\} > 0.$$

This means that if player k such that $k \notin D(a)$ for some $a \in E_1^*$ plays a suboptimal action against a_{-k}, then she loses at least r_k in terms of per-stage payoffs.

For any player k, if $k \notin (D^w(a) \cup SD(a))$ for some $a \in E_2^*$, then let

$$q_k = \min\bigl\{u_k(a) - u_k(\hat{a}_k, a_{-k}) : \\ a \in E_2^*,\ k \notin (D^w(a) \cup SD(a)), \text{ and } \hat{a}_k \neq a_k\bigr\} > 0.$$

Thus, if player k does not play a_k against a_{-k}, then she loses at least q_k in terms of per-stage payoffs.

Let $\underline{\rho} = \min_{a \in E^*} \rho(a)$ and $\Delta_i = \max_{a, a' \in A}(u_i(a) - u_i(a'))$. We then choose $\bar{\eta} \in (0, \hat{\eta}]$, $\bar{\lambda} = (\bar{\lambda}_i)_{i=1}^n \in (0, \hat{\lambda}]^n$, and $\underline{\delta} = (\underline{\delta}_i)_{i=1}^n \in [\hat{\delta}, 1)^n$ that satisfy the following inequalities:

$$(1 - \underline{\delta}_i + 2\bar{\eta})(\Delta_i + \bar{\lambda}_i) \leq \underline{\delta}_i (1 - \bar{\eta})^2 \bigl[v_i - \epsilon - u_i(\hat{s}(i))\bigr] \tag{24}$$

for any i such that $v_i - \epsilon > u_i(\hat{s}(i))$;

$$\bar{\lambda}_i \leq \underline{\delta}_i \underline{\rho} \hat{\kappa}(\bar{\eta}/2)(1 - \bar{\eta}) p_i \tag{25}$$

for any $i \notin SD^*$ such that $BR_i(\hat{s}(k)) \neq A_i$ for some k;

$$(1 - \underline{\delta}_i)\bar{\lambda}_i \leq \underline{\delta}_i \underline{\rho} \hat{\kappa}(\bar{\eta}/2) \Bigl[-(1 - \underline{\delta}_i + \bar{\eta})(\Delta_i + \bar{\lambda}_i) + \underline{\delta}_i \bigl[v_i - \epsilon - u_i(\hat{s}(i))\bigr]\Bigr] \tag{26}$$

for any i such that $i \in AP(a)$ for some $a \in E_2^*$;

$$\bar{\lambda}_i \leq \underline{\delta}_i \underline{\rho} \hat{\kappa}(1 - \bar{\eta}) \Bigl[(1 - \bar{\eta}) r_i - \bar{\eta} \Delta_i\Bigr] \tag{27}$$

for any i such that $i \notin D(a)$ for some $a \in E_1^*$; and

$$\bar{\lambda}_i \leq \underline{\delta}_i \underline{\rho} \hat{\kappa}(1 - \bar{\eta})^2 \Bigl[(1 - \bar{\eta})^2 q_i - [1 - (1 - \bar{\eta})^2] \Delta_i\Bigr] \tag{28}$$

for any i such that $i \notin (D^w(a) \cup SD(a))$ for some $a \in E_2^*$.

Let us consider $\Gamma(\delta, \lambda)$ with $\delta \geq \underline{\delta}$ and $\lambda \leq \bar{\lambda}$. Fix $\eta \in [\bar{\eta}/2, \bar{\eta}]$ and consider the system (5)–(16) associated with η, λ, and δ. By Lemma 1, the system has a solution satisfying (17) and (18).

We now construct an equilibrium of $\Gamma(\delta, \lambda)$ that approximates v. We start with describing the play on the path. The equilibrium play in each period depends on the "state" of that period. The set of states is $\{0, 1, \ldots, n\}$. State 0 is called the "cooperative state," while states $i \geq 1$ are called "punishment states." State $i \geq 1$ is the state in which player i is punished.

The play in state 0 is determined by the "substate" of the period, which is determined by the sunspot in that period. The set of substates is identified with E^*, and each substate $a \in E^*$ is realized with probability $\rho(a)$. In substate $a \in E_1^*(i)$, (a_i^η, a_{-i}) is played, while in substate $a \in E_2^*$, $a^\eta = (a_i^\eta, a_j^\eta, a_{-i-j})$ is played where $\{i, j\} = AP(a)$. In substate $a \in E_1^*(i)$, all players except for i monitor the other players. In substate $a \in E_2^*$, all players *not* in SD^* monitor the other players. This describes the play in the cooperative state. In state $i \geq 1$, on the other hand, the players play $\hat{s}(i)$ and do not monitor the other players.

For future reference, given a-substate, we call action a_i "cooperation," action a_i' "minor-deviation," and any other action "major-deviation." Note that, according to this terminology, only player i has a minor-deviation in substate $a \in E_1^*(i)$. Likewise, in substate $a \in E_2^*$, only players i and j associated with a can minor-deviate.

Next, we define the rule that determines the transition of states. The initial state is 0 (cooperative). If period $t-1$ is in state 0 and substate $a \in E_1^*(i)$, then period t is in:

(i) state i if player i major- or minor-deviated in period $t-1$,

(ii) state 0 otherwise.

Note that deviations of players $j \neq i$ are ignored. If period $t-1$ is in state 0 and substate $a \in E_2^*$, then period t is in:

(i) state k if player $k \in D^w(a)$ is the only player in $D^w(a)$ who major-deviated in period $t-1$,

(ii) state k if player $k \in \{i, j\} = AP(a)$ is the only player who minor-deviated in period $t-1$ and all other players in $D^w(a)$ cooperated in period $t-1$,

(iii) state 0 otherwise.

Note that, if $a \in E_2^*$ and the associated players i and j both minor-deviated, then their deviations are ignored and the state remains 0. Any deviation of player $k \notin D^w(a)$ is also ignored.

If the state changes to k because of player k's major deviation, then the state remains k in all subsequent periods. If the state changes to k because of k's minor deviation in a-substate, then the state remains k for $\tau^k(a)$ periods irrespective of the actions during the periods. The state of the subsequent period depends on the sunspot at the beginning of the period: with probability $\xi^k(a)$, the state remains k for one more period and then moves back to 0; with probability $1-\xi^k(a)$, the state moves back to 0 immediately. The numbers $\tau^k(a) \in \{0, 1, 2, \ldots\}$ and $\xi^k(a) \in [0, 1)$ are determined uniquely by

$$(1 - \xi^k(a))\delta_k^{\tau^k(a)} + \xi^k(a)\delta_k^{\tau^k(a)+1} = \nu^k(a), \tag{29}$$

where $\nu^k(a)$ is the relevant part of the solution of (5)–(16).

Note that $\tau^k(a) = 0$ is possible in the unique solution of (29). If this is the case, then the state transition after player k's unilateral minor-deviation is that with probability $1 - \xi^k(a)$, the state remains 0, and with probability $\xi^k(a)$, the state stays in k for one period and then moves back to 0.

A subtle issue that arises when $\tau^k(a) = 0$ is that the players seemingly need *two* sunspots: one that determines whether the state stays at 0 or moves to k, and the other that decides, if the state remains 0, which element of E^* is played. However, a single sunspot suffices since one can always let a sunspot play two roles at the same time.[15]

The above description specifies only the play on the path. However, each player can follow the description as long as she does not deviate from the description in terms of monitoring. To see this, note first that if the current period is in a-substate of the cooperative state with $a \in E_1^*(i)$, then all players except for i monitor the other players, and the future state (on the path) depends only on i's action (recall that deviations by player $j \neq i$ are ignored). Thus the state in the next period is common knowledge even though player i does not monitor the other players. Second, if the current period is in substate $a \in E_2^*$ and if $SD^* = \emptyset$, then all players monitor each other and the state in the next period is common knowledge. If $SD^* \neq \emptyset$, then the future state is common knowledge only among the players outside SD^*. However, this does not cause a problem since each player in SD^* plays

[15] For example, suppose $E^* = \{a', a''\}$ and $\rho(a') = \rho(a'') = 0.5$. If $\tau^k(a) = 0$ and $\xi^k(a) = 0.5$, then after a minor deviation by player k in substate a, the players can arrange the next-period play so that $\hat{s}(k)$ is played if the sunspot is in $[0, 0.5]$, a' if it is in $(0.5, 0.75]$, and a'' if it is in $(0.75, 1]$.

the (strictly) dominant action in all states, including the punishment states. Thus, the players in SD^* can play consistently with the above description without knowing the state.[16] Finally, if the current state is $i \geq 1$, then the future state depends only on sunspots and hence is common knowledge.

Let us define a strategy profile $\hat{\sigma}$ which is consistent with the state-dependent play described above. For each i, let $\hat{H}_i$ be the set of all histories for player i such that i monitored the other players whenever it was prescribed by the state-dependent play. Note that, as long as the players follow the state-dependent play, any history on the path is in $\hat{H}_i$. On the other hand, $\hat{H}_i$ includes other histories as well. It includes histories in which some player(s) (including i) deviated in terms of actions, and those in which i monitored the other players when she was not required to. Anyway, at any history $h_i^t \in \hat{H}_i$, player i knows the state and therefore can follow the state-dependent play described above. For $i \notin SD^*$, let $\hat{\sigma}_i$ be a strategy that is consistent with the above state-dependent play at any $h_i^t \in \hat{H}_i$. The behavior at the remaining histories can be arbitrary. For $i \in SD^*$, let $\hat{\sigma}_i$ be the strategy in which i always plays the dominant action and never monitors the other players. Then, $\hat{\sigma}$ generates a path consistent with the state-dependent play.

For $\hat{\sigma}$ defined as above, (5)–(16) and (29) imply that

(i) V^0 is the expected payoff vector under $\hat{\sigma}$,

(ii) $V^0(a)$ is the expected continuation payoff vector at substate a, and

(iii) $V^i(a)$ is the expected continuation payoff vector if the previous period is in substate a and player i unilaterally minor-deviates in that period.

We are now ready to prove the following:

Lemma 2 *Strategy profile* $\hat{\sigma}$ *is a Nash equilibrium of* $\Gamma(\delta, \lambda)$ *and satisfies* $|g_i(\hat{\sigma}) - v_i| < \epsilon$ *for any* i.

Proof. The second part follows directly from (17). Hence, it suffices to prove that $\hat{\sigma}$ is a Nash equilibrium of $\Gamma(\delta, \lambda)$.

We first show that it is not profitable for any player i to major-deviate in the cooperative state. To see this, note first that player i's major-deviation at substate a is not profitable in the short-run term if either $a \in E_1^*(j)$ for $j \neq i$ (by (1-iii) and $\eta \leq \bar{\zeta}$) or $a \in E_2^*$ and $i \notin D^w(a)$ (by (4) and $\eta \leq \bar{\zeta}$). In

[16] It is useful to note that if there exists a player in SD^*, then $E_1^* = \emptyset$, since (1-i) and (1-ii) cannot hold simultaneously.

either case, a major-deviation is not profitable since it yields no gain in the short run and does not change the future path. In any other case, we have $v_i - \epsilon > u_i(\hat{s}(i))$. In this case, if i major-deviates unilaterally, then the future play is the infinite repetition of $\hat{s}(i)$. While the major-deviation increases the current period payoff at most by $\Delta_i + \lambda_i$ (note that i can also deviate in monitoring), the continuation payoff decreases at least by $v_i - \epsilon - u_i(\hat{s}(i))$. By (24),

$$(1 - \delta_i)(\Delta_i + \lambda_i) \leq \delta_i\big[v_i - \epsilon - u_i(\hat{s}(i))\big].$$

Therefore, a major-deviation does not pay.

By (6), (7), and (10)–(13), any player who can minor-deviate given a cooperative substate is indifferent between cooperation and minor-deviation. Thus, in the cooperative state, no player has an incentive to deviate in terms of action. It is easy to see that, in state $i \geq 1$, no player has an incentive to deviate in terms of action or monitoring since a stage game Nash equilibrium is played and the future state depends only on sunspots (until the state moves back to 0). Thus, the proof is complete if we show that no player has an incentive to deviate in terms of monitoring at any substate a of state 0. We start with substates in E_1^*.

Substates in E_1^*. Let us examine the monitoring incentive in state 0 when the substate is $a \in E_1^*(i)$ for some i. In the substate, all players except for i monitor the other players. Since the future play depends only on player i's action, player i has no incentive to monitor the others.

Consider player $j \neq i$. Suppose that the substate was $a \in E_1^*(i)$ in period t and player j cooperated as prescribed by $\hat{\sigma}$ but did not monitor the other players. Then, in period $t+1$, she is uncertain whether the state is 0 or i, since she does not know whether the action of player i in period t was cooperation or minor-deviation. In what follows, we show that player j has an incentive to eliminate the uncertainty if the observation cost is sufficiently small.

If player i (and the others) cooperated in period t, then the state remains cooperative in period $t+1$. The substate remains a with probability $\rho(a) \geq \underline{\rho}$, in which case the other players play (a_i^η, a_{-i-j}) again. On the other hand, if player i minor-deviated in period t, then the other players play $\hat{s}_{-j}(i)$ in period $t+1$ with a positive probability. More precisely, if $\tau^i(a) \geq 1$, $\hat{s}_{-j}(i)$ is played with probability 1. If $\tau^i(a) = 0$, then $\hat{s}_{-j}(i)$ is played with probability $\xi^i(a)$ and (a_i^η, a_{-i-j}) is played with the remaining probability. (29) tells us that when $\tau^i(a) = 0$, $\xi^i(a)$ is given by

$$\xi^i(a) = \frac{1 - \nu^i(a)}{1 - \delta_i} > \hat{\kappa}, \tag{30}$$

where the inequality follows from (18).

Hence, if player j did not monitor the others at the end of period t, then in period $t+1$, with a probability no less than $\underline{\rho}\hat{\kappa}$, she is in a situation where she believes that the other players play $\hat{s}_{-j}(i)$ with probability η and (a_i^η, a_{-i-j}) with probability $1-\eta$. In this situation, since the state is either i or 0 with substate $a \in E_1^*(i)$, player j's action in period $t+1$ does not change the future play. By (1-ii), player j has no action that is a (static) best response against both $\hat{s}_{-j}(i)$ and a_{-j}. If she plays an action $\hat{a}_j \notin BR_j(\hat{s}(i))$, then she loses at least p_j in the period if the other players coordinate on $\hat{s}_{-j}(i)$, which occurs with probability $\eta \geq \bar{\eta}/2 > 0$. On the other hand, if she plays $\hat{a}_j \notin BR_j(a)$, then she loses at least r_j if the other players coordinate on a_{-j}, which occurs with probability $(1-\eta)^2 \geq (1-\bar{\eta})^2 > 0$.

Thus, by not monitoring the other players in period t, player j saves $\lambda_j \leq \bar{\lambda}_j$ in the period but suffers an expected loss of at least

$$\underline{\rho}\hat{\kappa} \min\Big\{(\bar{\eta}/2)p_j, (1-\bar{\eta})\big[(1-\bar{\eta})r_j - \bar{\eta}\Delta_j\big]\Big\}$$

in period $t+1$ and possibly more in the future. By (25) and (27), the deviation is not profitable.

Substates in E_2^*. We now consider the cooperative state with substate $a \in E_2^*$. In this substate, $\hat{\sigma}$ tells all players outside SD^* to monitor the other players. Let $\{i,j\} = AP(a)$. We examine the monitoring incentive of a given player k in the substate.

Case 1: $k \in SD^*$. Under $\hat{\sigma}$, player $k \in SD^*$ plays the dominant action a_k in every period regardless of the state (note that the dominant action must be played in any penal code Nash equilibrium). Since her play never affects the future play of the other players, she has no reason to monitor the other players.

Case 2: $k \notin D^w(a) \cup SD(a)$. This implies that a_k is a unique best response to a_{-k} while it is not the dominant action. Note that $k \notin AP(a)$. Suppose that in period t, the substate was a and player k cooperated but did not monitor. If the other players all cooperated or the associated players $\{i,j\}$ both minor-deviated, then with probability $\rho(a) \geq \underline{\rho}$, the other players play a_{-k}^η again in period $t+1$. On the other hand, if either i or j (but not both) minor-deviated, then $\hat{s}_{-k}(h)$ is played in period $t+1$ where $h \in \{i,j\}$ is the minor-deviator.

By the same argument as above, if player k did not monitor the other players in period t, then in period $t+1$, with a probability no smaller than $\underline{\rho}\hat{\kappa}$, player k is in a situation where she believes that the other players play

a^{η}_{-k} with probability $(1-\eta)^2+\eta^2$, $\hat{s}_{-k}(i)$ with probability $\eta(1-\eta)$, and $\hat{s}_{-k}(j)$ with probability $\eta(1-\eta)$.

In this situation, since the state is either i, j, or 0 with substate a, player k's action does not affect the future play. If player k plays a_k, then by (2-iii), she loses at least p_k in the period if either the other players coordinate on $\hat{s}_{-k}(i)$ or they coordinate on $\hat{s}_{-k}(j)$, each of which occurs with probability $\eta(1-\eta) \geq (\bar{\eta}/2)(1-\bar{\eta})$. On the other hand, if she plays any other action, then she loses at least q_k if the other players coordinate on a_{-k}, which occurs with probability $[(1-\eta)^2+\eta^2](1-\eta)^2 \geq (1-\bar{\eta})^4$.

Therefore, by not observing the other players in period t, player k saves $\lambda_k \leq \bar{\lambda}_k$ but suffers an expected loss of at least

$$\underline{\rho}\hat{\kappa}\min\Big\{(\bar{\eta}/2)(1-\bar{\eta})p_k, (1-\bar{\eta})^2\Big[(1-\bar{\eta})^2 q_k - [1-(1-\bar{\eta})^2]\Delta_k\Big]\Big\}$$

in period $t+1$ (in terms of stage-game payoffs) and possibly more in the future. Therefore, by (25) and (28), the deviation is not profitable.

Case 3: $k \in D^w(a)\backslash\{i,j\}$. Suppose that in period t, the state was 0 with substate a, and player $k \in D^w(a)\backslash\{i,j\}$ cooperated but did not monitor the other players. Then by a similar argument, it follows that in period $t+1$, with a probability no smaller than $\underline{\rho}\hat{\kappa}$, player k is in a situation where she believes that the other players play a^{η}_{-k} with probability $(1-\eta)^2+\eta^2$, $\hat{s}_{-k}(i)$ with probability $\eta(1-\eta)$, and $\hat{s}_{-k}(j)$ with probability $\eta(1-\eta)$.

We first show that, under the uncertainty, it is suboptimal for player k to play any action $a''_k \neq a_k$. Indeed, she gains by choosing a_k and monitoring the other players. To see this, note first that playing a_k and monitoring the other players decreases the current payoff at most by $\Delta_k + \lambda_k$. On the other hand, it does not decrease the continuation payoff if the state is either i or j. If the state is 0 with substate a, which occurs with probability $1-2\eta(1-\eta)$, then playing a_k and monitoring the other players increases the continuation payoff (from period $t+2$ on) at least by

$$\begin{aligned}&[1-2\eta(1-\eta)]V^0_k + \eta(1-\eta)[V^i_k(a)+V^j_k(a)] - u_k(\hat{s}(k))\\ &\quad > [v_k - \epsilon - u_k(\hat{s}(k))] - 2\eta(1-\eta)(\Delta_k+\lambda_k).\end{aligned}$$

Therefore, by (24), the overall payoff evaluated at $t+1$ increases by playing a_k and monitoring the other players.

However, (2-iii) implies that playing a_k in period $t+1$ causes a loss of at least p_k to player k in the period if the state is either i or j, each of which occurs with probability $\eta(1-\eta)$. Thus, by not monitoring the other players in period t, player k saves $\lambda_k \leq \bar{\lambda}_k$ but suffers an expected loss of at least $\underline{\rho}\hat{\kappa}(\bar{\eta}/2)(1-\bar{\eta})p_k$ in period $t+1$ (in terms of stage game payoffs) and possibly more in the future. Therefore, by (25), not monitoring in period t is not optimal.

Case 4: $k \in \{i, j\}$. Without loss of generality, let $k = i$. We first consider the case in which player i cooperated but did not monitor the other players in period t. Then, again, with a probability no less than $\underline{\rho}\hat{\kappa}$, player i in period $t+1$ is in a situation where she believes that the other players play a^{η}_{-i} with probability $1-\eta$ and $\hat{s}_{-i}(j)$ with probability η. Using a similar argument and (24), it can be shown that, under the uncertainty, it is suboptimal for i to play any $a''_i \notin \{a_i, a'_i\}$. However, by (2-ii), playing a_i or a'_i causes a loss of at least p_i in the period if the other players coordinate on $\hat{s}_{-i}(j)$. Therefore, by (25), not monitoring the other players is suboptimal.

We also have to consider the case in which player i *minor-deviated* and then did not monitor the other players in period t. Then, in period $t+1$, player i is uncertain whether the state is i or 0; the latter occurs if j also minor-deviated in period t. Then, by a similar argument, it follows that with a probability no smaller than $\underline{\rho}\hat{\kappa}$, player i in period $t+1$ is in a situation where she believes that the other players play $\hat{s}_{-i}(i)$ with probability $1-\eta$ and a^{η}_{-i} with probability η. Under the uncertainty, if player i chooses an action $a''_i \in \{a_i, a'_i\}$, then by (2-ii), it causes a loss of at least p_i in the period if the other players coordinate on $\hat{s}_{-i}(i)$. On the other hand, if she plays any $a''_i \notin \{a_i, a'_i\}$, it is regarded as a major-deviation if the other players coordinate on (a^{η}_j, a_{-j-i}), which occurs with probability η.

Thus, if player i does not monitor the other players at period t, she saves $\lambda_i \leq \bar{\lambda}_i$ in the period but the continuation payoff evaluated at period t decreases at least by the minimum of the following values:

$$\delta_i \underline{\rho}\hat{\kappa}(1-\eta)(1-\delta_i)p_i,$$
$$\delta_i \underline{\rho}\hat{\kappa}\eta\Big[-(1-\delta_i)(\Delta_i+\lambda_i)+\delta_i\big[\eta V_i^j(a)+(1-\eta)V_i^0-u_i(\hat{s}(i))\big]\Big].$$

The second value is greater than the right-hand side of (26). Therefore, by (25) and (26), not monitoring at period t is suboptimal.

Case 5: $k \in SD(a) \setminus SD^*$. Then since $k \notin SD^*$, there exist $\hat{a} \in E^*$ such that $\hat{a}_k \neq a_k$. By Condition (s-iv) of supportability, there exists such an $\hat{a}$ that satisfies either $\hat{a} \in E_2^*$ or $[\hat{a} \in E_1^*(k)$ and $\hat{a}'_k \neq a_k]$, where $\hat{a}'_k$ is the minor-deviation from $\hat{a}$ for player k. As before, suppose that in period t, the state was 0 with substate a and player k cooperated but did not monitor the other players. Then, in the next period, with positive probability no smaller than $\underline{\rho}\hat{\kappa}$, the player is uncertain whether the state is 0 with substate $\hat{a}$ or is either i or j. As before, we show that the loss that player k suffers from the uncertainty exceeds the monitoring cost. We distinguish the following two cases.

First, consider the case when $\hat{a} \in E_2^*$. An argument similar to the previous ones can show that, under the uncertainty, choosing a major-deviation from $\hat{a}$ is not optimal for player k. That is, it is suboptimal for her to play $a_k'' \notin \{\hat{a}_k, \hat{a}_k'\}$ if $k \in AP(\hat{a})$ and $a_k'' \neq \hat{a}_k$ if $k \notin AP(\hat{a})$. However, by $\hat{a}_k \neq a_k$ and (2-ii), neither $\hat{a}_k$ nor $\hat{a}_k'$ (if well defined) is the dominant action, and therefore, playing either action causes a loss of at least p_k in the period if the other players coordinate on $\hat{s}_{-k}(i)$ or $\hat{s}_{-k}(j)$, which occurs with probability $2\eta(1-\eta)$. Thus, the desired conclusion follows from (25).

Second, consider the case when $\hat{a} \in E_1^*(k)$ and $\hat{a}_k' \neq a_k$. A similar argument shows that, under the uncertainty, a major-deviation from $\hat{a}$ is not optimal for player k. But, since neither $\hat{a}_k$ nor $\hat{a}_k'$ is her dominant action, playing either action causes a loss of at least p_k in the period if the other players coordinate on $\hat{s}_{-k}(i)$ or $\hat{s}_{-k}(j)$, which occurs with probability $2\eta(1-\eta)$. Thus, the desired conclusion follows from (25).

These arguments prove that no player has an incentive to deviate in terms of monitoring on the equilibrium path. Q.E.D.

The next lemma completes the proof of Proposition 1.

Lemma 3 *There exists a sequential equilibrium in $\Gamma(\delta, \lambda)$ that generates the same path as $\hat{\sigma}$.*

Proof. We start with the construction of a system of beliefs. First, consider a sequence of totally mixed strategy profiles that converges to $\hat{\sigma}$ and puts far smaller weights on the trembles regarding monitoring, in comparison with the trembles regarding actions. This generates a sequence of systems of beliefs, whose limit is a system of belief that is consistent with $\hat{\sigma}$ and such that, at any history (on or off the path), each player believes that the other players did not deviate in terms of monitoring. Let us denote the system of beliefs by μ and let $\mu(h_i^t)$ denote the belief of player i at history h_i^t about the other players' (private) histories.

We now construct each player's strategy, σ_i^*. For players $i \in SD^*$, we simply set $\sigma_i^* = \hat{\sigma}_i$. For players $i \notin SD^*$, we first set

$$\sigma_i^*(h_i^t) = \hat{\sigma}_i(h_i^t) \quad \text{for any } h_i^t \in \hat{H}_i. \tag{31}$$

For any $h_i^t \notin \hat{H}_i$, let $\sigma_i^*(h_i^t)$ be an optimal continuation strategy at h_i^t given her belief $\mu(h_i^t)$ and the other players' strategies $\hat{\sigma}_{-i}$.

(31) means that, for any player i, σ_i^* and $\hat{\sigma}_i$ coincide at all histories h_i^t in which i never deviates in terms of monitoring. This together with the construction of μ implies that μ is also consistent with σ^*.

We now verify sequential rationality. We start with a player $i \notin SD^*$ and a history $h_i^t \in \hat{H}_i$. At the history, player i knows the state. Moreover, by the construction of μ, player i believes that the other players not in SD^* also know the state since i believes that these players never deviated in terms of monitoring. In other words, i believes that any other player $j \notin SD^*$ is at a history in $\hat{H}_j$ and any other player $j \in SD^*$ plays her dominant action in all subsequent periods. Then, by Lemma 2 and (31), it is optimal for player i to follow the state-dependent play, i.e., $\hat{\sigma}_i$. By (31), this is exactly what σ_i^* prescribes. Thus the continuation play of σ_i^* at history h_i^t is sequentially rational.

We now consider a player $i \notin SD^*$ at a history $h_i^t \notin \hat{H}_i$. We defined $\sigma_i^*(h_i^t)$ as an optimal continuation strategy at the history given that i's belief is $\mu(h_i^t)$ and the other players follow $\hat{\sigma}_{-i}$ (not σ_{-i}^*). Sequential rationality then follows since player i believes that all other players $j \notin SD^*$ are at some histories $h_j^t \in \hat{H}_j$ and therefore, by (31), their continuation strategies coincide with $\sigma_j^*(h_j^t)$.

Finally, consider player $i \in SD^*$. At any history, she believes that the other players never deviated in terms of monitoring and thus she believes that her play does not affect the future play. Thus, given σ_{-i}^* and μ, it is sequentially rational for i to play a short-run best response at any history. This is nothing but following σ_i^*.

Therefore, $(\sigma^*,\ \mu)$ is a sequential equilibrium of $\Gamma(\delta, \lambda)$. By (31), it is outcome-equivalent to $\hat{\sigma}$. This completes the proofs of the lemma and Proposition 1. Q.E.D.

References

Abreu, D. (1986): "Extremal Equilibria of Oligopolistic Supergames," *Journal of Economic Theory*, 39, 191–225.

Abreu, D. (1988): "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56, 383–396.

Abreu, D., D. Pearce, and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58, 1041–1063.

Ahn, I., and M. Suominen (2001): "Word-of-Mouth Communication and Community Enforcement," *International Economic Review*, 42, 399–415.

Amarante, M. (2002): "Recursive Structure and Equilibria in Games with Private Monitoring," forthcoming in *Economic Theory*.

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

Benoit, J.-P., and V. Krishna (1985): "Finitely Repeated Games," *Econometrica*, 53, 905–922.

Ben-Porath, E., and M. Kahneman (1996): "Communication in Repeated Games with Private Monitoring," *Journal of Economic Theory* 70, 281–297.

Bhaskar, V., and E. van Damme (2002): "Moral Hazard and Private Monitoring," *Journal of Economic Theory*, 102, 16–39.

Bhaskar, V., and I. Obara (2002): "Belief-based Equilibria in the Repeated Prisoners' Dilemma with Private Monitoring," *Journal of Economic Theory*, 102, 40–69.

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

Compte, O. (2002): "On Failing to Cooperate When Monitoring Is Private," *Journal of Economic Theory*, 102, 151–188.

Ely, J., and J. Valimaki (2002): "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory*, 102, 84–105.

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

Fudenberg, D., and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54 533–556.

Fudenberg, D., D. Levine, and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62, 997–1039.

Kandori, M. (1992): "Social Norms and Community Enforcement," *Review of Economic Studies*, 59, 63–80.

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

Lehrer, E., and A. Pauzner (1999): "Repeated Games with Differential Time Preferences," *Econometrica*, 67, 393–412.

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

Matsushima, H. (1991): "On the Theory of Repeated Games with Private Information, Part 1: Anti-folk Theorem without Communication," *Economics Letters*, 35, 253–256.

Matsushima, H. (2002): "The Folk Theorem for the Prisoner-Dilemma with Imperfect Private Monitoring," forthcoming in *Econometrica*.

Miyahara, Y. (2002): "Near Efficiency in Repeated Prisoner's Dilemma with Costly Monitoring," mimeo.

Piccione, M. (2002): "The Repeated Prisoner's Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory*, 102, 70–83.

Sekiguchi, T. (1997): "Efficiency in Repeated Prisoner's Dilemma with Private Monitoring," *Journal of Economic Theory*, 76, 345–361.

Stigler, G. (1964): "A Theory of Oligopoly," *Journal of Political Economy*, 72, 44–61.