Repeated Games with Imperfect Public Monitoring : A Survey*

太　田　勝　憲※
Ohta, Katsunori

ABSTRACT

In this paper, we investigate *infinitely repeated games* to illustrate long-term relationships among players. Specifically, we study repeated games with *imperfect public monitoring*. In this class of games, players cannot observe the others' actions directly, but can publicly observe imperfect signals about them. We introduce several prominent works in infinitely repeated games with imperfect public monitoring.

1　Introduction

Repeated games are the study of long-term relationships. From the economic point of view, examples of long-term relationships range from oligopolistic firms competing in a given market to people interacting with each other in a community. In this paper, we investigate *infinitely repeated games* to illustrate long-term relationships among players. Specifically, we study repeated games with *imperfect public monitoring*. In this class of games, players cannot observe the others' actions directly, but can publicly observe imperfect signals about them.(1) Hence, repeated games with perfect monitoring are the special case where the public

* This paper is a revised version of chapter 1 of my doctoral dissertation. I would like to thank my advisor, Takashi Yanagawa for encouragement. I would like to express my gratitude to Hajime Kobayashi and Tadashi Sekiguchi. I have been fortunate to learn repeated games from them.

※ ohta@eco.wakayama-u.ac.jp

signals consist of the realized actions themselves.

There are some reasons why we have to pay much attention to this class of information structure. First, it is a more reasonable assumption than perfect monitoring in many economic problems. "Secret price cutting" which is intended to explain emergence of "price wars" is an example of these games. In the analysis of an oligopolistic market, as suggested by the primary work of Stigler (1964), we cannot always expect firms to observe their rivals' actions. There are some industries where prices are negotiated between firms and customers on a case-by-case basis. For this reason, it is difficult to monitor secret price cutting by the rivals.

Second, the behavior on the equilibrium path may be drastically changed from the one in games with perfect monitoring. Suppose quantity-setting oligopoly.(2) Since firms cannot observe their rivals' output level, it is difficult for firms to maintain collusive agreement. Although market-clearing price can be used as the signals to infer the rivals' output level, this signals become imperfect information about their rivals' output level because of fluctuations of demand. When firms observe "low price", they cannot distinguish whether the low price is caused by their rivals' deviations from collusion or fluctuations of demand.

Consider a usual trigger strategy. When the model is the one with perfect monitoring, firms collude until the rivals deviate from collusion. Then, patient firms enjoy collusion; the price is high on average and non-collusive behavior does not appear. On the other hand, price wars (low price on average) occur, when the model is the one with imperfect monitoring. Consider that firms choose the collusive behavior until the low price is observed. Even if firms

↙(1) See Fudenberg and Tirole (1991), where chapter 6 gives a precise explanation of repeated games except recent studies. See also Tirole (1988), where chapter 5 gives applications of repeated games to the analysis of industrial organization.

(2) Green and Porter (1984) show the formal study of the games.

choose the collusive behavior, there are cases in which the low price realizes. Then, if firms do not switch their behavior to competitive one, deviations from collusion is profitable for the firms. That is, collusion cannot be sustained by such strategy. To support collusion, firms have to switch their behavior to competitive one when they observe the low price. Since firms cannot observe their rivals' output level, it is difficult for firms to maintain collusive agreement. Then inefficient punishment causes efficiency loss. To sum up, if we assume imperfect public monitoring, it creates a new issue: degree of value that is generated by cooperation.

We introduce several prominent works in infinitely repeated games with imperfect public monitoring. This paper is organized as follows. In section 2, we describe repeated games with imperfect public monitoring. In section 3, we refer to the work of Abreu, Pearce and Stacchetti (1990). This paper characterizes the set of equilibrium outcomes. In section 4, we explain the algorithm for finding the asymptotic equilibrium payoff set introduced by Fudenberg and Levine (1994). In section 5, we mention the folk theorem by Fudenberg, Levine and Maskin (1994). In section 6, we explore a formula introduced by the work of Abreu, Milgrom and Pearce (1991) (hereafter, we call the formula the AMP formula). This paper shows the formula measuring efficiency loss caused by imperfection of public signals. In the final section, we mention some recent studies for repeated games with imperfect public monitoring.

2 Modeling Imperfect Monitoring

2.1 General Models of Imperfect Monitoring

First, we describe imperfect monitoring generally. We define the stage game: $\Gamma = \{N, A, \Omega, \pi, \{r_i\}_{i=1}^{n}\}$. We denote the set of players by $N(|N| = n)$. Each player chooses an action simultaneously in each period. Let us denote the finite set of actions in Γ for player i by A_i. Let $A = A_1 \times A_2 \times \cdots \times A_n$.

Players cannot observe the others' actions. However they receive imperfect information about others' actions in the stage game. Let Ω_i be the finite set of signals for player i and Ω is the product set of Ω_i for all players. Let $\pi_i(\omega_i \mid a)$ be the marginal distribution of the signals. Note that player i can observe only $\omega_i \in \Omega_i$. The marginal distribution of the ω_i is represented by $\pi_i(\omega_i \mid a)$.

The realized payoff for player i is represented by $r_i : A_i \times \Omega_i \to \mathcal{R}$. Note that r_i is independent of a_{-i} and ω_{-i}. This means that there does not exist additional information about others' actions in the payoff function.

We represent the normal form $G = \{N, A, \{u_i\}_{i=1}^n\}$ where

$$u_i(a) = \sum_{w_i \in \Omega_i} r_i(a_i\ \omega_i)\pi_i(\omega_i \mid a)$$

For the notational convenience, we introduce some notations. We denote that $co\{g(a) \mid a \in A\}$ is the convex hull of $g(a)$ for any $a \in A$. Let $V = co\{g(a) \mid a \in A\}$ be the feasible payoff set. We denote the player i's *minimax value* by $\underline{v_i}$. That is, $\underline{v_i} = \min_{\alpha-i}[\max_{\alpha_i} g_i(\alpha_i, \alpha_{-i})]$. Let $V^* = \{v \in V \mid v_i \geq \underline{v_i}, \forall i\}$ be the set of feasible and individually rational payoffs.

Let us now turn to the infinitely repeated game under common discount factor δ, $\Gamma(\delta)$. In each period, the stage game is played and then the corresponding signals are revealed. The private history at the beginning of period t is $h_i(t) = \{a_i(t'), \omega_i(t')\}_{t'=1}^{t-1}$. We denote the set of all private histories for player i by H_i^t. Let $\mathcal{H}_i = \cup_{t=1}^{\infty} H_i^t$. A strategy of the player i, σ_i, is a map from $\mathcal{H}_i$ to the set of (randomized) actions. We call Σ_i the set of all payer i's strategies, and $\Sigma = \Sigma_i \times \cdots \times \Sigma_n$ the set of all strategy profiles. Given a strategy profile $\sigma \in \Sigma$ and a common discount factor δ, each player considers their discounted average payoff. The player i's average discounted sum of expected payoff is $g_i(\sigma) = (1-\delta)\Sigma_{t'=1}^{\infty} E[\delta^{t-1} u_i(a_i(t')) \mid \sigma, \pi]$.

Then, we define Nash equilibrium of $\Gamma(\delta)$ as follows.

Definition 1. $\sigma \in \Sigma$ *is Nash equilibrium if and only if*

$$g_i(\sigma) \geq g_i(\sigma_i', \sigma_{-i})$$

for any $i \in N$ and for any $\sigma_i' \in \Sigma_i$.

2.2 Imperfect Public Monitoring

Next, let us define imperfect *public* monitoring.

Definition 2. Γ *is the game with imperfect public monitoring when* $\Omega_1 = \Omega_2 = \cdots = \Omega_n$ *and* $\pi(\omega \mid a) > 0 \Rightarrow \omega_1 = \omega_2 = \cdots = \omega_n$ *hold for any* $a \in A$ *and for any* $\omega \in \Omega$.

In the rest of this paper, Let $\Omega = \Omega_i$ for any $i \in N$ be the set of *public signal* and $h(t) = \{\omega(t')\}_{t'=1}^{t-1} \in H^t$ be the public history at the beginning of periot t. We denote $\Gamma = \{N, A, \Omega, \pi, \{r_i\}_{i=1}^{n}\}$ where $\pi = \pi(\omega \mid a)$ and $r_i : A_i \times \Omega \to R$.

We focus on a special class of strategy, *public* strategy.(3) A strategy of player i is *public* if at each time, t, it depends on the public history and not on the private history. We define *public strategy* formally.

Definition 3. $\sigma_i \in \Sigma_i$ *is public if and only if for any* $h_i^t = \{a_i(\tau), \omega(\tau)\}_{\tau=0}^{t-1}$, $(t \geq 1)$ *and* $\hat{h}_i^t = \{\hat{a}_i, \hat{\omega}(\tau)\}_{\tau=0}^{t-1}$ *such that* $\omega(\tau) = \hat{\omega}(\tau)$ *for any* $\tau \leq t-1$,

$$\sigma_i(h_i^t) = \sigma_i(\hat{h}_i^t)$$

Then, we adopt *perfect public equilibrium* (hereafter, we call it PPE) as a solution concept. A PPE is a profile of public strategies such that at every date t and for any public history $h(t)$ the strategies constitute a Nash equilibrium

(3) Public strategy is a special class of strategy because it utilizes only public history. On the other hand, strategy that utilizes not only public history but also private history is called *private* strategy. See the final section in this paper.

from that date on.

For $\sigma_i \in \Sigma$, $t \geq 1$, $h^t = (\omega(1), \omega(2), \cdots, \omega(t-1)) \in H^t$, $\sigma|_{h_t}$ is continuation strategy which satisfies

$$\sigma|_{h_t}(\widehat{h}^t) = \sigma(h^t \oplus \widehat{h}^t),$$

where $\widehat{h}^t = (\widehat{\omega}(1), \widehat{\omega}(2), \cdots, \widehat{\omega}(t-1))$ and $h^t \oplus \widehat{h}^t = (\omega(1), \omega(2), \cdots, \omega(t-1), \widehat{\omega}(1), \widehat{\omega}(2), \cdots, \widehat{\omega}(t-1))$. Then, we can define PPE formally.

Definition 4. *A public strategy profile, $\sigma \in \Sigma$, is PPE if and only if (i) σ is Nash equilibrium and*

(ii) for any history h^t, $\sigma|_{h^t}$ is Nash Equilibrium for $\Gamma(\delta)$.

Note that in a perfect public equilibrium the players' beliefs about the opponents' past play are irrelevant. Therefore, perfect public equilibrium is a special class of sequential equilibrium. Under the full support assumption, $\pi(\omega|a) > 0$ for any $\omega \in \Omega$ and $a \in A$, a Nash equilibrium by public strategy is a sequential equilibrium.

Let $E(\delta)$ be the set of PPE when discount factor is δ. Let $V(\delta)$ be the set of the discounted average payoff vectors that are led by public perfect equilibrium when discount factor is δ. It is obvious that $V(\delta) \subseteq V^*$. Note that one shot deviation principal is sufficient to check whether a strategy constitutes PPE.

3. Dynamic Programming and Self-Generation

We can find a recursive structure in infinitely repeated games with imperfect public monitoring. Because players adopt public strategies, the continuation strategy after any realization of public signals needs to be PPE of the original game. By the concept of *self-generation*, Abreu, Pearce and Stacchetti (1990) characterize the payoff set of perfect public equilibrium. Self-generation is a sufficient condition that a payoff set is supportable by perfect public equilibrium.

We introduce some notations. Let $F(\Omega, W)$ be the set of functions from Ω to W where $W \subseteq \mathcal{R}^n$. The function $f \in F(\Omega, W)$ is interpreted as the continuation payoffs. Then, we define the normal form game $G^f = \{N, A, \{g_i^f\}_{i=1}^n\}$ where

$$g_i^f(a) = (1-\delta)u_i(a) + \delta \sum_{\omega \in \Omega} \pi(\omega \mid a) f_i(\omega),$$

given $f = (f_1, f_2, \cdots, f_n) \in F(\Omega, W)$.

Definition 5. *$v \in \mathcal{R}^n$ is generated by W, if for some $f \in F(\Omega, W)$ and s,*

$$v = (1-\delta)u_i(s) + \delta \sum_{\omega \in \Omega} \pi(\omega \mid s) f_i(\omega) \tag{1}$$

$$v \geq (1-\delta)u_i(s_i', s_{-i}) + \delta \sum_{\omega \in \Omega} \pi(\omega \mid s_i', s_{-i}) f_i(\omega) \tag{2}$$

Condition (1) says that the realized payoff vector is v when all players play s given f : $v = g^f(s)$. Condition (2) says that playing s_i is optimal given the continuation payoffs are f : $s \in NE(G^f)$ where $NE(G^f)$ is the set of Nash equilibrium in G^f.

Moreover, we must check that f is equilibrium payoffs. To see this, let $B(W;\delta)$ be the set of payoff vectors generated by W, given δ. Then, we introduce the concept of self-generation.

Definition 6. *W is self-generating if $W \subseteq B(W;\delta)$.*

W is self-generating if the payoff set that can be enforced with continuation payoffs in W.

Theorem 1. *If W is bounded and self-generating, then $B(W;\delta) \subseteq V(\delta)$.*

Proof. We will construct strategies for the repeated game in which payoff vector w is realized, and check that strategies are PPE.

Fix $w \in B(W;\delta)$. Because w is generated by W, there exists $f^0 \in F(Y, W)$ and $s^0 \in NE(G^{f^0})$ such that $w = g^{f^0}$. That is,

$$w = (1-\delta)u(s^0) + \delta\Sigma_{\omega \in \Omega}\pi(\omega \mid s^0) f^0(\omega)$$

where $f^0 \in F(\Omega, W)$ and $s^0 \in NE(G^{f^0})$. Because $f^0(\omega) \in W$ for any $\omega \in \Omega$, $f^0(\omega)$ is also generated by W. Hence, there exists $f^\omega \in F(\Omega, W)$ and $s^\omega \in NE(G^{f\omega})$ such that $f^0(\omega) = g^{f\omega}(s^\omega)$. That is,

$$f^0(\omega) = (1-\delta)u(s^{\{\omega\}}) + \delta \sum_{\omega' \in \Omega} \pi(\omega \mid s^{\{\omega\}}) f^{\{\omega\}}(\omega')$$

where $f^{\{\omega\}} \in F(\Omega, W)$, $s^{\{\omega\}} \in NE(G^{f^{\{\omega\}}})$. Furthermore, because $f^\omega(\omega') \in W$ for any $\omega \in \Omega$ and $\omega' \in \Omega$, $f^w(\omega')$ is also generated by W.

Hence, there exists $f^{\omega, \omega'} \in F(\Omega, W)$ and $s^{\omega, \omega'} \in NE(G^{f\omega, \omega'})$ such that $f^\omega(\omega') = g^{f\omega, \omega'}(s^{\omega, \omega'})$. For any sequence of signals $\{\omega(\tau)\}_{\tau=0}^{t}$, there exists $f^{\{\omega(\tau)\}_{\tau=0}^{t}} \in F(\Omega, W)$ and $s^{\{\omega(\tau)\}_{\tau=0}^{t}} \in NE(G^{f^{\{\omega(\tau)\}_{\tau=0}^{t}}})$ such that

$$f^{\{\omega(\tau)\}_{\tau=0}^{t-1}}(\omega(t)) = g^{f^{\{\omega(\tau)\}_{\tau=0}^{t}}}(s^{\{\omega(\tau)\}_{\tau=0}^{t}}). \tag{3}$$

Next, we define the strategy σ^* as following;

$$\sigma_i^*(h^t) = s_i^0, \text{ if } t = 0$$
$$= s_i^{\{\omega(\tau)\}_{\tau=0}^{t-1}}, \text{ if } t \geq 1.$$

Then, It is clear that σ^* is a public strategy profile.

Let $\{\omega(\tau)\}_{\tau=0}^{t-1}$ be the signal profile which corresponds to h^t. Then, from equation (3),

$$f^{\{\omega(\tau)\}_{\tau=0}^{t-2}}(\omega(t-1)) = (1-\delta)u(s^{\{\omega(\tau)\}_{\tau=0}^{t-1}})$$
$$+\delta \sum_{\omega \in \Omega} \pi(\omega \mid s^{\{\omega(\tau)\}_{\tau=0}^{t-1}}) f^{\{\omega(\tau)\}_{\tau=0}^{t-1}}(\omega) \tag{4}$$

Suppose the L.H.S. of the equation (4) is w when $t = 0$. By using the definition of σ^*, we obtain

$$f^{\{\omega(\tau)_{\tau=0}^{t-2}\}}(\omega(t-1)) = (1-\delta)E[\sum_{\tau=t}^{\infty} \delta^{\tau-t} u(a(\tau)) \mid \sigma^*] \tag{5}$$

Then, $f^{\{\omega(\tau)_{\tau=0}^{t-2}\}}(\omega(t-1)) = g(\sigma^* |_{h^t})$. We use the supposition $\delta \in (0, 1)$ and the property that W is bounded. Specifically, (5) implies $w = g(\sigma^*)$ when $t = 0$.

Moreover, (4) and $s^{\{\omega(\tau)\}_{\tau=0}^{t-1}} \in NE(G^{f^{\{\omega(\tau)\}_{\tau=0}^{t-1}}})$ implies

$$f_i^{\{\omega(\tau)\}_{\tau=0}^{t-2}} \geq (1-\delta)u_i(a_i', s_{-i}^{\{\omega(\tau)\}_{\tau=0}^{t-1}}) + \delta \sum_{\omega' \in \Omega} \pi(\omega' | a_i', s_{-i}^{\{\omega(\tau)\}_{\tau=0}^{t-1}}) f_i^{\{\omega(\tau)\}_{\tau=0}^{t-1}}(\omega'), \quad (6)$$

for any i and $a_i' \in A_i$.

Since $f_i^{\{\omega(\tau)\}_{\tau=0}^{t-1}}(\omega') = g_i(\sigma|_{h^t, \omega'})$, (6) implies that one shot deviations are not profitable. Hence, σ^* is PPE which yields the payoff vector w ∎.

Theorem 1 means that all payoffs in W are PPE payoffs.

Theorem 2. $B(V(\delta);\delta) = V(\delta)$ *for any* $\delta \in (0, 1)$.

proof Fix $\delta \in (0, 1)$ and $v \in V(\delta)$. Choose PPE σ^* such that $g(\sigma^*) = v$. We denote that s^* is the current period action profile by σ^*. Because $\sigma^* \in E(\delta)$

$$v_i = (1-\delta)u_i(s^*) + \delta \sum_{\omega \in \Omega} \pi(\omega | s^*) g_i(\sigma^*|_{(\omega)})$$

$$\geq (1-\delta)u_i(a_i', s_{-i}^*) + \delta \sum_{\omega \in \Omega} \pi(\omega | a_i', s_{-i}^*) g_i(\sigma^*|_{(\omega)})$$

for any $i \in N$ and $a_i' \in A_i$.

Because $\sigma^*|_{(\omega)} \in E(\delta)$ for any ω, $g(\sigma^*|_{(\omega)}) \in V(\delta)$ holds.

Let a function f be $f(\omega) = g(\sigma^*|_{(\omega)}) \in V(\delta)$. Because $f \in F(\Omega, V(\delta))$ and $s^* \in NE(G^f)$, $v \in B(V(\delta);\delta)$. Because $v \in V(\delta)$ is arbitrary, $V(\delta) \subseteq B(V(\delta);\delta)$.

Because $V(\delta) \subseteq V$, $V(\delta)$ is clearly bounded. Because $V(\delta)$ is self-generating by the above argument, $B(V(\delta);\delta) = V(\delta)$ holds by Theorem 1. ∎

4 Fudenberg and Levine's Algorithm

While APS characterizes the equilibrium payoff set given δ, Fudenberg and Levine (1994) characterize the asymptotic equilibrium payoff set, $\lim_{\delta \to 1} V(\delta)$. They find an algorithm to compute $\lim_{\delta \to 1} V(\delta)$ (hereafter we call the algorithm *FL algorithm*).

FL algorithm for finding $\lim_{\delta \to 1} V(\delta)$ is the study of the payoff that can be generated using continuation payoffs that lie in half-space $H \in \mathcal{R}^n$; $H(\lambda, k) = \{v | \lambda \cdot k \leq k\}$ for $\lambda \in \mathcal{R}^n$ and $k \in \mathcal{R}$.

For a given welfare weight vector $\lambda \in \mathcal{R}^n \setminus \{0\}$ and a mixed action profile $s \in S$ and $\delta \in (0, 1)$, we solve the following linear programming $LP(\lambda : s, \delta)$;

$$\max_{v, f(\cdot)} \lambda \cdot v$$

subject to

$$v_i = (1-\delta) u_i(a_i, s_{-i}) + \delta \sum_{\omega \in \Omega} \pi(w | a_i, s_{-i}) f_i(\omega) \text{ for } a_i \in \text{supp} s_i \tag{7}$$

$$v_i \geq (1-\delta) u_i(a_i, s_{-i}) + \delta \sum_{\omega \in \Omega} \pi(w | a_i, s_{-i}) f_i(\omega) \text{ for } a_i \notin \text{supp} s_i \tag{8}$$

$$\lambda_v \geq \lambda f(\omega) \tag{9}$$

for all $\omega \in \Omega$, $i = 1, 2 \cdots, n$

where supps_i is the support of s_i. The constraint (7) means $v = g^f(s)$. The constraint (8) means $s \in NE(G^f)$. The term f represents the variation in continuation payoffs, and the constraint (9) ensures that the continuation payoff lie in the equilibrium payoff set.

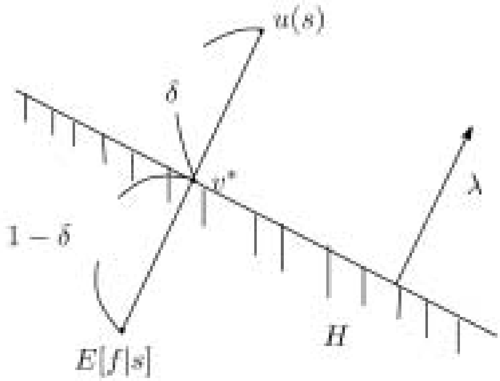

Figure 1: $LP(\lambda : s, \delta)$

Let $k^*(\lambda ; s, \delta)$ be the value of $LP(\lambda ; s, \delta)$.

Lemma 1. (*i*) $k^*(\lambda\,;s,\,\delta)$ *is independent of* δ ($k^*(\lambda\,;s,\,\delta) = k^*(\lambda\,;s)$). (*ii*)$k^*(\lambda\,;s) \leq \lambda u(s)$. (*iii*)$k^*(\lambda\,;s) = \lambda u(s) \Leftrightarrow s \in NE(G^f)$ *for some* $f \in F(\Omega,\,H)$ *where* $H = \{v \in \mathcal{R}^n\,;\lambda v = \lambda u(s)\}$.

Instead of a formal proof, we explain the statement intuitively. The idea of (*i*) is that we can find the continuation payoffs according to the degree of the discount factor. If (*ii*) does not hold, there exists $f_i(\omega)$ such that (7) does not hold. This is depicted by the following figure.

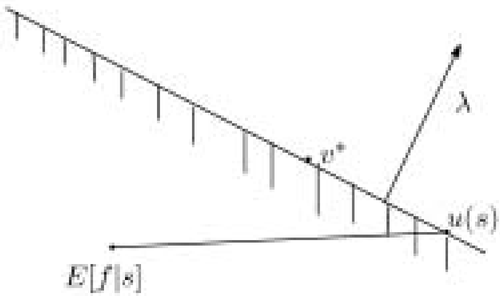

Figure 2: lemma 1-(*iii*)

Definition 7. *The maximal score in direction* λ *is* $k^*(\lambda) = sup_{s\in S}k^*(\lambda,\ s)$. *The maximal half-space in direction* λ *is* $H^*(\lambda) = \{v \in \mathcal{R}^n : \lambda v \leq k^*(\lambda)\}$

Definition 8. $Q = \cap_{\lambda\in\mathcal{R}^n\setminus\{0\}}H^*(\lambda)$.

Theorem 3. (*i*) *For any* $\delta \in (0,\ 1)$, $V(\delta) \subseteq Q$.

(*ii*) *If* $dim Q = n$, , *then* $lim_{\delta\to 1}V(\delta) = Q$.

The idea of Theorem 3−(*ii*) is as follows. Note that $v = (1-\delta)u(s)+\delta E[f(\omega)\,|\,s]$. When δ is sufficiently large, the interior closed balls in Q is self-generating. That is the continuation payoffs can be supportable by PPE. See Figure 3.

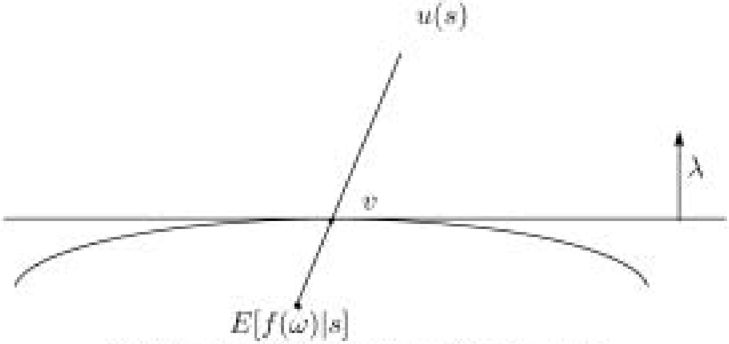

Figure 3: $LP(\lambda : s)$ when δ is close to 1

5　The Folk Theorem

5.1　A Partnership Game Example

Before analyzing the folk theorem, we demonstrate a simple example of the partnership game in Radner, Myerson and Maskin (1986). Consider two individuals are involved in a team production. The stage game, G, is as follows. $A_i = \{C, D\}$ for all $i = 1, 2$. C means cooperate, or work, and D means defect, or shirk. We assume the disutility of taking C is 3, and that of taking D is zero. In game G, two public outcomes, $\Omega = \{12, 0\}$, can be realized according to the actions taken by both players. These outcomes represent total output, and it is divided for two players equally. The distributions of the outcomes in G are following.

$$\pi(12 | C, C) = \frac{2}{3},$$

$$\pi(12 | D, C) = \pi_1(12 | C, D) = \frac{1}{3},$$

$$\pi(12 | D, D) = 0,$$

where the first argument in an action profile is the action of player 1, and the second is that of player 2.

Then we can compute the expected utility of G. The payoff structure is a Prisoners' Dilemma.

	C	D
C	1, 1	−1, 2
D	2, −1	0, 0

Table 1: Prisoners' Dilemma

We analyze $\Gamma(\delta)$ in which G is played infinitely with the common discount factor δ. Let (v_1, v_2) be the PPE payoff vectors which is the discounted average payoff in $\Gamma(\delta)$.

Proposition 1. $v_1+v_2 \leq 1$, *independent of* $\delta \in (0, 1)$.

Proof. Suppose that $v_1+v_2 > 1$. Then, both players take C with positive probability in the first period to support (v_1, v_2) such that $v_1+v_2 > 1$.

If not, then, from Table 1, the first period payoffs could sum to at most 1, implying that continuation payoffs, either $f_i(12)$ or $f_i(0)$, would sum to more than $\max\{v_1+v_2 | (v_1, v_2) \in V(\delta)\}$, a contradiction.

Let α_i be such probability for $i = 1, 2$. Then,

$$v_1 = (1-\delta)u_i(C, \alpha_j) + \delta\{\alpha_j(\frac{2}{3}f_i(12)+\frac{1}{3}f_i(0))+(1-\alpha_j)(\frac{1}{3}f_i(12)+\frac{2}{3}f_i(0))\} \quad (10)$$

where $f_i(\omega)$ is player i's continuation payoff after a realization ω.

Incentive condition for player i is

$$(1-\delta) \leq \delta\frac{f_i(12)-f_i(0)}{3}. \quad (11)$$

Because (10) can rewritten as

$$v_i = (1-\delta)u_i(C, \alpha_j) + \delta f_i(12) - (\alpha_j + 2(1-\alpha_j))\frac{f_i(12) - f_i(0)}{3},$$

(11) implies that

$$v_i = (1-\delta)u_i(C, \alpha_j) + \delta f_i(12) - (\alpha_j + 2(1-\alpha_j))(1-\delta).$$

Summing up these two equations to get

$$v_i + v_2 \leq (1-\delta)(u_1(C, \alpha_2) + u_2(C, \alpha_1))$$
$$+\delta(f_1(12) + f_2(12)) - (4-\alpha_1-\alpha_2)(1-\delta).$$

Since $u_1(C, \alpha_2) + u_2(C, \alpha_1) \leq 2$ and $f_1(12) + f_2(12) \leq v_1 + v_2$, we obtain

$$v_1 + v_2 \leq 2 - (4-\alpha_1-\alpha_2) \leq 0.$$

This is a contradiction. ∎

This result implies that efficiency loss is inevitable. The reason is that players have to use inefficient continuation payoff after a realization of the bad signal, 0, because of the symmetric information structure.

5.2 Fudenberg, Levine and Maskin (1994)

In the previous example, there is only one signal (0) that is informative about players' deviations. Fudenberg, Levine and Maskin (1994), FLM hereafter, shows that the folk theorem can be obtained under generic choice of signal distributions if the number of public signals is no less than the total number of elements of action sets. This condition enables players to detect deviations statistically so that they are able to punish the potential deviators and reward the others appropriately in the future. Since asymmetric punishment provides players with incentives to cooperate without loss of efficiency, we can obtain the folk theorem.

When we want to obtain the folk theorem under imperfect public monitoring, we need conditions that guarantee to enforce a targeted action profile and to punish the deviators efficiently.

FLM shows sufficient conditions that the interior points of V^* lies in $V(\delta)$ when the discount factor is close to one. We introduce the conditions that are a certain restriction on the probability distribution over public signals given the players' action. Let $\pi(\cdot|a)=\Sigma_{a\in A}\pi(\cdot|a)\alpha(a)$ be a row vectors which dimensions equal to the number of elements in Ω. Let $\Pi_i(\alpha_{-i})=\pi(\cdot|\cdot,\alpha_{-i})$ be a $|A_i|\times|\Omega|$ matrix whose rows are the elements in A_i and columns are the elements in Ω.

First, we introduce the condition related to enforceability.

Definition 9. *An action profile* α *has individual full-rank for player* i *if* $\Pi_i(\alpha_{-i})$ *has rank* $|A_i|$. *(There is no linear independence among* $|A_i|$ *vectors.)*
Individual full rank says that different deviations by the same player lead to different distributions over signals. This means that players can statistically detect "someone deviated". Then, they can enforce a target action by policing all of the suspicious players.

Second, we describe the remaining requirement. For the requirement, we introduce a notation. Let $\Pi_{ij}(\alpha)$ be

$$\Pi_{ij}(\alpha)=\begin{pmatrix}\Pi_j(\alpha_{-i})\\ \Pi_j(\alpha_{-j})\end{pmatrix}$$

Then, we state the condition related with distinguishability of deviations.

Definition 10. *An action profile* α *has pairwise full-rank for player* i, j $(i\neq j)$ *if* $\Pi_{ij}(\alpha)$ *has rank* $|A_i|+|A_j|-1$.
Pairwise full-rank condition is related to statistical distinguishability of each player's deviations. It guarantees players for punishing the deviators without efficiency loss.

Theorem 4. *If* (*i*) *the individual full-rank condition is satisfied for any player at any action profile,* (*ii*) *for each pair* $i,j\in N(i\neq j)$ *there exists an action profile which has pairwise full-rank for* i, j, *and* (*iii*) *the dimension of* V^* *is equal to* *the number of players, then for any closed and smooth set* $W \subseteq V^*$*, there exists* $\underline{\delta}$ *such that* $W \subseteq V(\delta)$ *for any* $\delta \in (\underline{\delta}, 1)$.

6 The AMP formula

We introduce Abreau, Milgrom and Pearce (1991) in this section. The model is a special type of repeated prisoners' dilemma.

Consider two players play the stage game over infinite periods $t = 1, 2, \cdots$. Each player has binary choices, $A_1 = A_2 = \{C, D\}$. Let $A = A_1 \times A_2$.

The players cannot observe the opponent's actions but receive imperfect public signals of the stage actions. Let Ω be the set of the signals. We assume Ω is countable, which may be finite or infinite. Given an action profile $a = (a_1, a_2) \in A$, a signal $\omega \in \Omega$ realizes with probability $\pi(\omega | a)$. We make the following assumption to focus on interesting situations.

Assumption 1. $p_i = \pi(\omega_i | C, C) > 0$ *and* $q_i = \pi(\omega_i | D, C) = \pi(\omega_i | C, D) > 0$ *for any* i.

Assumption 1 guarantees symmetry and the (partial) full support.

We assume that the stage expected payoff structure, $u(a)$, has a property of Prisoners' Dilemma.

For simplicity, we assume that the payoff structure is the following Prisoners' Dilemma game as in Table 2. The following assumption guarantees that (C, C) is the efficient action profile.

	C	D
C	$1, 1$	$-h, 1+g$
D	$1+g, -h$	$0, 0$

Table 2: Prisoners' Dilemma

Assumption 2. g, $h > 0$ and $g-h < 1$.

In each period, the stage game is played and then the corresponding public signal is revealed. We assume that a public randomization device is available at the end of each period. The randomization device selects a number $\lambda \in [0, 1]$ according to the uniform distribution on $[0, 1]$. The public history at the beginning of period t is represented by a sequence $h(t) = \{\omega(t'), \lambda(t')\}_{t'=1}^{t-1}$. The private history at the beginning of period t is $h_i(t) = \{a_i(t')\}_{t'=1}^{t-1}$. We denote the set of all public histories and private histories for firm i by H^t and H_i^t, respectively. Let $\mathcal{H}_i = \cup_{t=1}^{\infty}(H^t \times H_i^t)$. Following Abreu, Milgrom and Pearce (1991), we consider pure strategies only. A (pure) strategy of firm i, σ_i, is a map from $\mathcal{H}_i$ to A_i. We call Σ_i the set of all player i's strategies, and $\Sigma = \Sigma_1 \times \Sigma_2$ the set of all strategy profiles. Given a common discount factor δ and a sequence of action profiles, $\{a(t)\}_{t=1}^{\infty}$, generated by a strategy profile $\sigma \in \Sigma$, the player i's average discounted expected payoff is $(1-\delta)\Sigma_{t=1}^{\infty}\delta^{t-1}E[u_i(a_i(t))]$.

We focus on a special class of equilibrium, which is so-called *strongly symmetric perfect public equilibrium*. A strategy of player i is *public* if at each time, t, it depends on the public history $h(t)$ and not on the private history $h_i(t)$. A public strategy profile is *strongly symmetric* if at each time t and for any public history $h(t)$ the strategies prescribe the same action for both players. A strongly symmetric perfect public equilibrium is a profile of public strategies such that (i) it is strongly symmetric and (ii) at every date t and for any public history $h(t)$ the strategies constitute a Nash equilibrium from that date on.

The best equilibrium value in strongly symmetric perfect public equilibrium of a symmetric game can be expressed by applying the AMP formula. In order to derive the formula, without loss of generality, we limit our attention to a generalized trigger strategy. A generalized trigger strategy consists of the two phases:

C phase and D phase. In C phase each player plays a "cooperative" action, while plays the dominant action (D) in D phase. Transition of the phase is as follows. The phase in period 1 is arbitrary. If period $t-1$ is in C phase and if the signal in period $t-1$ is ω_i, then the phase switches to D with probability $\alpha_i \in [0, 1]$. With the remaining probability, the phase remains in phase C. Because of availability of public randomization device, the stochastic transition of the phase takes place via the realization of the public randomization device. If period $t-1$ is in D phase, then period t is in D phase with certainty. That is, D phase is the absorbing phase. Note that the profile (σ_D, σ_D), where the strategy prescribe the dominant action D at any history, is a special example of this type of strategy, which starts with D phase in period 1. Hence the infinitely repeated game has a strongly symmetric perfect public equilibrium for any δ.

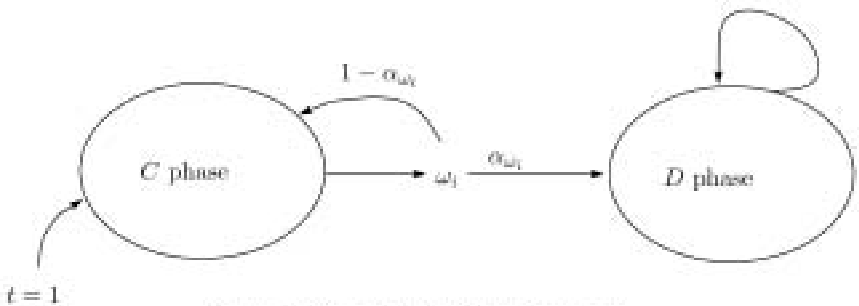

Figure 4: Generalized Trigger Strategy

We limit our attention to generalized trigger strategies which start with C phase. These strategies are characterized by a vector $\alpha = (\alpha_1, \alpha_2, \cdots)$. Given α, and the corresponding strategy σ, we can compute the continuation payoffs of the players which we denote v.

$$v = (1-\delta)+\delta \sum_i p_i(1-\alpha_i)v \tag{12}$$

This is the value recursion equation. For σ to be an equilibrium, it is necessary that playing D when it is in C phase is not a profitable deviation.

$$v \geq (1-\delta)(1+g)+\delta \sum_i q_i(1-\alpha_i)v \tag{13}$$

The one stage deviation principle implies that (13) is also sufficient for σ to be an equilibrium.

We explore the most efficient equilibrium and its value. Let σ^* be the most efficient strongly symmetric perfect public equilibrium with the associate vector $\alpha^* = (\alpha_1^*, \alpha_2^*, \cdots)$ and payoff v^*, given δ. If $\sigma^* \neq (\sigma_D, \sigma_D)$, then (13) must hold with equality. If not, then the punishment is so strong that the optimality of σ^* is not valid.(4) Then $\sigma^* \neq (\sigma_D, \sigma_D)$ exists, then σ^* must satisfy

$$v^* = (1-\delta)+\delta \sum_i p_i(1-\alpha_i^*)v^* \tag{14}$$

and

$$v^* = (1-\delta)(1+g)+\delta \sum_i q_i(1-\alpha_i^*)v^*. \tag{15}$$

By the incentive condition, (15),

$$v^*-\delta v^* = (1-\delta)(1+g)- \sum_i q_i \alpha_i^* v^*$$

$$\begin{aligned} v^* &= (1+g)-\frac{\delta}{1-\delta}\sum_i q_i \alpha_i^* v^* \\ &= (1+g)-\frac{\delta}{1-\delta}\sum_i p_i \alpha_i^* v^* \left(\frac{\Sigma_i q_i \alpha_i^*}{\Sigma_i p_i \alpha_i^*}\right) \\ &= (1+g)+(v^*-1)\left(\frac{\Sigma_i q_i \alpha_i^*}{\Sigma_i p_i \alpha_i^*}\right) \end{aligned}$$

The last equality comes from (14). Let $\left(\frac{\Sigma_i q_i \alpha_i^*}{\Sigma_i p_i \alpha_i^*}\right) = l^*$. This represents the likelihood ratio of the event in which the transition of the D phase takes place. Then the above equation becomes

$$-g = (v^*-1)(l^*-1).$$

Solving the above equation, we can obtain the AMP formula;

$$v^* = 1-\frac{g}{l^*-1}.$$

(4) Sekiguchi (2001) shows the proof of the statement.

If the value is positive, then it is the highest value of strongly symmetric equilibrium. On the other hand, if it is negative, then the repetition of the Nash equilibrium of the stage game is optimal and the value is zero. It states that the maximum payoff that yielded by symmetric public strategy equilibrium is cooperative payoff minus efficiency loss attributed to imperfect monitoring. The efficiency loss depends on the two terms, g and l^*. The larger deviation gain makes the loss large because a severe punishment is necessary. More interestingly, the larger l^* is, the lower the loss is. The larger l^* means that the deviation is more detectable, so a weaker punishment gives sufficient discipline.

Furthermore, we can predict the upper bound of the most efficient symmetric equilibrium payoff. Let $\hat{l} = \sup_i \frac{q_i}{p_i}$. If $\hat{l} > 1$, then $v^* \leq 1 - \frac{g}{\hat{l} - 1}$.

7 Some Recent Studies

7.1 The Folk Theorem without Pairwise Full-Rank

FLM shows that the folk theorem can be obtained under generic choice of signal distributions if the number of public signals is no less than the total number of elements of action sets. Kandori (2003) shows that the folk theorem also holds even when the number of signals is relatively small. He requires conditions (i) there are at least three players, (ii) players can publicly communicate with each other. Once allowing public communication, it performs the role of public signals. Then, even if the public signal space is relatively small, we can totally obtain enough information through public communication.

The enforceability of a given action profile on various hyperplanes is the crux of the argument to achieve the folk theorem. In the FLM folk theorem, pairwise full-rank condition ensures this by requiring that *linear* combinations of relevant signal distributions are distinct. Although this condition is clear, it is the restricted condition. Kandori and Matsushima (1998) point out that this can be weakened by requiring that the *convex* combinations of relevant signal distributions are distinct. Moreover, Kandori and Obara (2006b) suggest that we can obtain efficiency under weaker set of conditions for the enforceability on hyperplanes by imposing restrictions *jointly* on the information structure *and* the payoff structure.

7.2 Private Strategy

As we have seen, almost all works restrict their attention to public strategy, where the players' actions depend on the public observable signals. Kandori and Obara (2006a) explore *private strategy* which utilizes not only public history but also private history and show that efficiency in repeated games can often be drastically improved by private strategy. They consider mixed private strategy. Then, a pair of action and public signal may contain more information than just a public signal about the opponent's action.

References

[1] Abreu, D., P. Milgrom, and D. Pearce (1991) "Information and Timing in Repeated Partnerships," *Econometrica*, **59**, 1713-1733.

[2] Abreu, D., D. Pearce, and E. Stacchetti (1990) "Toward a theory of discounted repeated games with imperfect monitoring," *Econometrica*, **58**, 1041-1063.

[3] Fudenberg, D., D. Levine (1994) "Efficiency and Observability in Games with Long-Run and Short-Run Players," *Journal of Economic Theory*, **62**, 103-135.

[4] Fudenberg, D., D. Levine, and E. Maskin (1994) "The Folk Theorem with Imperfect Public Information," *Econometrica*, **62**, 997-1040.

[5] Fudenberg, D., J. Tirole (1991) *Game Theory*, MIT press, Cambridge, Massachusetts.

[6] Green, E. and R. Porter (1984) "Noncooperative Collusion under Imperfect Price Information," *Econometrica*, **52**, 87-100.

[7] Kandori, M. (2003) "Randomization, Communication and Efficiency in Repeated Games with Imperfect Public Monitoring," *Econometrica*, **71**, 345-353.

[8] Kandori, M and I. Obara (2006a) “Efficiency in Repeated Games Revisited: The Role of Private Strategies,” *Econometrica*, **74**, 499-519.

[9] Kandori, M and I. Obara (2006b) “Less is More: An Observability Paradox in Repeated Games,” *International Journal of Game Theory*, **34**, 475-493.

[10] Kandori, M and H. Matsushima. (1998) “Private Observation, Communication and Collusion,” *Econometrica*, **66**, 627-652.

[11] Radner, R. (1986) “Repeated Partnership Games with Imperfect Monitoring and no discounting,” *Review of Economic Studies*, **53**, 43-58.

[12] Radner, R., R. Myerson, and E. Maskin (1986) “An Example of a Repeated Partnership Game with Discounting and with Uniformly Inefficient Equilibria,” *Review of Economic Studies*, **53**, 59-69.

[13] Sekiguchi, T. (2001) “Resurgence of the Abreu-Milgrom-Pearce Formula,” *Kobe University Economic Review*, **47**, 43-60.

[14] Stigler, G. (1964) “A Theory of Oligopoly,” *Journal of Political Economy*, **72**, 44-61.

[15] Tirole, J. (1988) *The Theory of Industrial Organization*, MIT press, Cambridge, Massachusetts.