Repeated Games with Imperfect Public Monitoring

Niccolò Lomys*

March 29, 2019

Preamble

- These notes heavily draw upon Hörner (2016), Mailath and Samuelson (2006), Levin (2006), Obara (2003), and Fudenberg and Tirole (1991). All errors are my own. Please bring any error, including typos, to my attention.
- These notes are only a first introduction to the theory of repeated games with imperfect public monitoring. If you want (or need) to learn more, an excellent starting point is Part II in Mailath and Samuelson (2006).

1 Introduction

We now study another information structure, *imperfect public monitoring*. In this class of models, players cannot observe the other players' actions directly, but can observe imperfect and public signals about them. The link between current actions and future play is now indirect, and in general, deviations cannot be unambiguously detected. However, equilibrium play will affect the distribution of the commonly observed signals, allowing intertemporal incentives to be created by attaching punishments to signals that are especially likely to arise in the event of a deviation. This in turn will allow us to support behavior in which players do not myopically optimize.

There are at least two reasons why this information structure is worth special attention. First, it is simply a more reasonable assumption than perfect monitoring in some settings. Second, the equilibrium behavior can be significantly different from the equilibrium behavior in models with perfect monitoring, and this sometimes has significant economic implications. The following examples illustrate these points.

An Example

This is an (overly) simplified version of the dynamic quantity competition model in Green and Porter (1984).[1] Consider two firms, 1 and 2, producing the same product. Each firm can either "collude" (play C, i.e. produce a small amount of goods) or "not collude" (play NC, i.e. produce

*Toulouse School of Economics, University of Toulouse Capitole; niccolo.lomys@tse-fr.eu.

[1] You can also interpret the game as a prisoners' dilemma.

a large amount of goods). Each firm's profit is determined by its choice and the realization of a common price. If both firms choose C, the price is "high" with probability $1 - p_0$ and "low" with probability p_0; if only one firm chooses C, the price is "high" with probability $1 - p_1$ and "low" with probability p_1; finally, if both firms choose NC, the price is "high" with probability $1 - p_2$ and "low" with probability p_2. We assume that $p_2 > p_1 > p_0$, so that a low price is more likely to realize when more firms choose NC. A firm's action is not observable, but the price is public information, hence this is a model of imperfect public monitoring. Suppose the expected rewards of the stage game are as in the following matrix

		Firm 2	
		C	NC
Firm 1	C	$4, 4$	$-1, 6$
	NC	$6, -1$	$1, 1$

We restrict firms' strategies to be functions from past realizations of prices to $\{C, NC\}$ (i.e. firms do not condition their strategies on their own past actions). As we will show, under this restriction recursive methods are applicable to infinitely repeated games with imperfect monitoring.[2] Consider, for instance, the following grim trigger strategy: both firms play (C, C) until a low price is observed, and play (NC, NC) forever once a low price is observed. Each firm's incentive constraint is

$$
\begin{aligned}
(1-\delta)6 + \delta\big[(1-p_1)\overline{V} + p_1\underline{V}\big] &\leq (1-\delta)4 + \delta\big[(1-p_0)\overline{V} + p_0\underline{V}\big] \\
\text{or} \quad (1-\delta)2 &\leq \delta(p_1 - p_0)\big(\overline{V} - \underline{V}\big), \\
\text{where} \quad \overline{V} &= (1-\delta)4 + \delta\big[(1-p_0)\overline{V} + p_0\underline{V}\big] \\
\text{and} \quad \underline{V} &= 1.
\end{aligned}
$$

You can check that if $p_1/p_0 > 5/3$ (the random price is informative enough) and firms are sufficiently patient, then the above constraints are satisfied.

What is interesting in this grim trigger strategy equilibrium is that (NC, NC) is played on the equilibrium path (punishments happen): there is probability one that a low price eventually occurs and both firms play NC thereafter. This contrasts with models with perfect monitoring. If firms play the above stage game over time with perfect monitoring, (NC, NC) may not be observed at all (if they play the best symmetric collusive equilibrium). Indeed, the most important feature of this model is that NC has to be played on the equilibrium path in any nontrivial equilibrium. This is because firms have to punish themselves after a realization of low price to support any degree of collusion, which requires firms to play more NC following a realization of low price. If nontrivial intertemporal incentives are to be created, then over the course of equilibrium play the two firms will find themselves in a punishment phase infinitely often. This happens despite the fact that firms know, when the punishment is triggered by a low price, that both have in fact followed the equilibrium prescription of colluding. Then why

[2] Why do we need to impose such restriction? Is it without loss of generality? We will answer these questions throughout these notes.

do they carry through the punishment? Given that the other firm is entering the punishment phase, it is a best response to do likewise. But why would equilibria arise that routinely punish firms for offenses not committed? Because the expected payoffs in such equilibria can be higher than those produced by simply playing a Nash equilibrium of the stage game.

This might affect the way we interpret price dynamics observed in some particular market. Suppose there is enough reason to believe that firms are playing the best strongly symmetric collusive equilibrium (i.e. the equilibrium in which both firms produce the same output after every history). If you believe the true model is one with perfect monitoring, then an episode of "price war" would be taken as a proof that there is no collusion in the market. In contrast, if you believe that the true model is one with imperfect public monitoring, you would not be able to reach that conclusion. On the contrary, you might take an episode of "price war" as a proof of collusion if such regime-switching is a part of the best strongly symmetric collusive equilibrium.

The punishment supporting collusion, consisting of permanent reversion to no collusion after the first low price, is often more severe than necessary for efficiency (from the viewpoint of the two firms' payoffs). That is, strongly symmetric behavior may be inefficient. This was no problem in the perfect-monitoring case, where the punishment was safely off the equilibrium path and hence need never be carried out. Here, imperfect monitoring requires that punishments will occur. The firms would thus prefer the punishments to be as lenient as possible, consistent with creating the appropriate incentives to collude. In other words, forgiving strategies may sustain higher payoffs.

Given the inevitability of punishments, one might suspect that the set of feasible outcomes in games of imperfect monitoring is rather limited. In particular, it appears as if the inevitability of some periods in which firms do not collude makes efficient outcomes impossible. In fact, we will show that the Folk Theorem need not hold once monitoring is imperfect. However, we will also characterize (quite general) conditions under which we again have a folk-theorem result. The key is to work with asymmetric punishments, sliding along the frontier of efficient payoffs so as to reward some players as others are penalized. There may thus be a premium on asymmetric strategies, despite the lack of any asymmetry in the game. This requires that the monitoring structure gives players (noisy) indications not only that a deviation from equilibrium play has occurred but also who might have been the deviator.[3]

2 Model

Stage Game and Repeated Game

Attention is restricted throughout to infinitely repeated games. A repeated game with imperfect public monitoring specifies, in addition to the set of players $N := \{1, \dots, n\}$ and action profiles A, a set of signals Y (finite) and, for each action profile $a \in A$, a distribution $\pi(\cdot \mid a)$ on Y; we refer to (Y, π), where $\pi := \pi(\cdot \mid a)_{a \in A}$, as the *monitoring structure* of the game. The interpretation is straightforward: as a function of the action profile a played in a given period, the signal $y \in Y$ is drawn according to the distribution $\pi(\cdot \mid a)$. Actions are not observed by other players; on

[3] As we will see, the two-signal example we just studied fails this condition.

the other hand, the signal y is *publicly* observed. We write $\pi(y \mid \alpha) := \sum_{a\in A} \pi(y \mid a)\alpha(a)$ for the distribution of signals induced by a mixed action $\alpha \in \Delta(A)$.

For consistency, rewards are first defined as functions

$$g_i \colon A_i \times Y \to \mathbb{R},$$

so that a player's realized reward $g_i(a_i, y)$ carries no more information than what he already knows, or observes, namely a_i and y. Given some action profile a, player i's expected reward $u_i \colon A \to \mathbb{R}$ is defined pointwise as

$$u_i(a) := \sum_{y\in Y} g_i(a_i, y)\pi(y \mid a)$$

Each player seeks to maximize average discounted sum of his expected rewards. It has been customary to use the function u as the primitive of the repeated game, rather than g. In this fashion, fixing u, we can examine how the quality of the monitoring affects the equilibrium payoff set without having to worry about how the change in monitoring affects the set of feasible payoffs, as it "mechanically" would if we were to take g as a primitive. Therefore, we shall take u as a primitive, and ignore g from now on, but it is important to keep in mind that players cannot infer anything from $u_i(a)$ beyond what they already know, a_i and y.

An *infinitely repeated game with imperfect public monitoring*, then, is a collection

$$G := (N, A, (Y, \pi), u),$$

referred to as the the stage game, along with a discount factor $\delta \in [0, 1)$. Again, we denote the repeated game as G^δ. Note that, up to period t, player i has observed an element of $H_i^t := (A_i \times Y)^t$, corresponding to the actions he has played and the public signals he has observed.[4] This is the set of player i's , where we denote h_i^t a generic such history. Players share some information, namely the sequence of public signals, or *public history* $h^t \in H^t := Y^t$. As usual, define $H_i := \cup_{t=0}^{\infty} H^t$ and $H := \cup_{t=0}^{\infty} H^t$

Perfect monitoring is the special case in which $Y = A$ and $\pi(y \mid a) = 1$ iff $y = a$, so that action profiles are perfectly observed. Of course, our interest primarily lies in the case in which the monitoring is not perfect, though everything we shall prove applies to perfect monitoring as well.

Equilibrium Notion for the Repeated Game

In perfect monitoring games, there is a natural isomorphism between histories and information sets. Consequently, in perfect monitoring games, every history, h, induces a continuation game that is strategically identical to the original repeated game, and for every strategy σ in the original game, h induces a well-defined continuation strategy $\sigma|_h$. Moreover, any Nash equilibrium induces Nash equilibria on the induced equilibrium outcome path.

[4]If a public correlating device is assumed, it is understood that each player also has observed its realizations in all previous periods.

Unfortunately, none of these observations hold for public monitoring games. Since players have private information (their own past action choices), a player's information sets are naturally isomorphic to the set of his own private histories, H_i, not to the set of public histories, H. Thus there is no continuation game induced by any history—a public history is clearly insufficient, and i's private history will not be known by the other players. This lack of a recursive structure is a significant complication, not just in calculating Nash equilibria but in formulating and applying a tractable notion of sequential rationality.

A recursive structure does hold, however, on a restricted strategy space.

Definition 1 (Public and Private Strategies)**.** *A strategy σ_i is* public *if, in every period t, it depends only on the public history $h^t \in H$ and not on i's private history. That is, for all $h_i^t, \hat{h}_i^t \in H_i$ satisfying $y^\tau = \hat{y}^\tau$ for all $\tau \leq t-1$, we have*

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

A strategy σ_i is private *if it is not public.*

Let $h_i^t \in H_i^t$ be a private history for player i. Player i's actions before t may well be relevant in determining i's beliefs over the actions chosen by the other players before t. However, because the other players' continuation behavior in period t is only a function of the public history $h^t \in H^t$ and not their own past behavior, player i's expected payoffs are independent of i's beliefs over the past actions of the other players, and so i has a best reply in public strategies. Thus, we have the following.

Lemma 1. *If all players other than i are playing a public strategy, then player i has a public strategy as a best reply.*

Note, however, that a player need not have a public best reply to a nonpublic strategy profile of his opponents.

Restricting attention to public strategy profiles, every public history h^t induces a continuation game that is strategically identical (in terms of public strategies) to the original repeated public monitoring game, and for any public strategy σ_i in the original game, h^t induces a well-defined continuation public strategy $\sigma|_{h^t}$. Moreover, any Nash equilibrium in public strategies induces Nash equilibria (in public strategies) on the induced equilibrium outcome path.

Exercise 1. Show that every pure strategy in a repeated public monitoring game is realization equivalent to a public pure strategy. [Recall: Two strategies σ_i and $\hat{\sigma}_i$ are realization equivalent if, for all strategies for the other players, σ_{-i}, the distributions over outcomes induced by (σ_i, σ_{-i}) and $(\hat{\sigma}_i, \sigma_{-i})$ are the same.]

Solution. See Mailath and Samuelson (2006), Lemma 7.1.2 (and its proof), pages 229–230. ■

Definition 2 (Perfect Public Equilibrium)**.** *The strategy profile σ is a* public perfect equilibrium (hereafter, PPE) *of the repeated game, if, for all $i \in N$, σ_i is public, and for all public histories $h^t \in H$, $\sigma|_{h^t}$ is a Nash equilibrium of the subgame beginning at h^t.*

This solution concept is a natural extension of subgame-perfection to imperfect monitoring: if players' strategies only condition on public events, we require that they are Nash equilibria conditional on any such event. Clearly, PPE are sequential equilibria. However, there are sequential equilibria that are not PPE, and there are well-known examples of repeated games in which, as $\delta \to 1$, efficient payoffs can be approximated by sequential equilibria, but not by PPE: the power of statistical tests to detect deviations can be improved by using all information a player has available, which includes his own privately observed actions.[5]

We let E_δ denote the set of PPE payoffs of G^δ. The main benefit of this solution concept is that the set of PPE payoffs is *stationary*—i.e. it is the same starting from any period public history: of course, which PPE is selected as a continuation strategy profile depends on the public history, in general, but the set of PPE to select from does not. This is not true for sequential equilibria, as private histories provide private correlation devices whose structure depends, among others, on the period considered.

One-Shot Deviation Principle

The proof of the next theorem is a straightforward modifications of its perfect-monitoring analog, and so is omitted.

Theorem 1 (One-Shot Deviation Principle)**.** *A public strategy profile σ is a public perfect equilibrium of G^δ if and only if no player has a profitable one-shot deviation.*

3 Dynamic Programming and Self-Generation

In this section, we describe a general method (introduced by Abreu, Pearce and Stacchetti (1986, 1990), hereafter APS) for characterizing E_δ, the set of PPE payoffs. We have already seen a preview of this method when studying repeated games with perfect monitoring.[6] The essence of this approach is to view a PPE as describing after each public history the specified action to be taken after the history and continuation promises. The continuation promises are themselves required to be equilibrium values.[7] The recursive properties of PPE described in the previous section provide the necessary structure for this approach.

The next two definitions are the public monitoring versions of the notions of enforceability and decomposability we already encountered. Hereafter, let us fix $\delta \in [0, 1)$.

Definition 3 (Enforceability)**.** *Given $W \subseteq \mathbb{R}^n$, the mixed action profile α is* enforceable on W *if there exists $\gamma \in W^Y$ such that α is a Nash equilibrium of the game static $\Gamma_\delta(\gamma)$ with action sets A_i and payoff function*

$$(1-\delta)u(\cdot) + \delta \sum_{y \in Y} \pi(y \mid \cdot)\gamma(y).$$

[5] Kreps and Wilson (1982) define sequential equilibrium for finite games only. Throughout these notes, the definition is extended to infinitely repeated games by equipping both the set of strategies and the set of systems of beliefs with the uniform topology of uniform convergence over information sets.

[6] This recursive method was originally developed for repeated games with imperfect public monitoring. We applied this method to repeated games with perfect monitoring for pedagogical reasons.

[7] Again, see also Mertens and Parthasarathy (2003) and Shapley (1953).

In this case, we say that γ enforces α on W *and define*

$$V(\alpha, \gamma) := (1 - \delta)u(\alpha) + \delta \sum_{y \in Y} \pi(y \mid \alpha)\gamma(y).$$

We interpret the function γ as describing expected payoffs from future play ("continuation promises") as a function of the public signal y. Enforceability is then essentially an incentive compatibility requirement. The profile α is enforceable if it is optimal for each player to choose α, given some γ describing the implications of current signals for future payoffs.

Definition 4 (Decomposability). *A payoff vector* $v \in \mathbb{R}^n$ *is* decomposable *on* $W \subseteq \mathbb{R}^n$ *if there exists a mixed action profile* α*, enforced by* γ *on* W*, such that*

$$v = V(\alpha, \gamma). \tag{1}$$

In this case, we say that the payoff v *is* decomposed *by the pair* (α, γ) *on* W*.*

Let $\mathcal{P}(\mathbb{R}^n)$ denote the set of all subsets of $\mathbb{R}^n$. We define the function

$$B_\delta \colon \mathcal{P}(\mathbb{R}^n) \to \mathcal{P}(\mathbb{R}^n), \qquad W \mapsto B_\delta(W),$$

where

$$B_\delta(W) := \{v \in \mathbb{R}^n : v \text{ is decomposable on } W \text{ with discount } \delta\}.$$

We can think of $B_\delta(W)$ as the set of all the payoffs that can be obtained from (decomposed) using an enforceable choice α in the current period and with the link γ between current signals and future equilibrium payoffs when the discount factor is δ. It is easy to show that the function B_δ has the following properties.

1. B_δ is a monotone function; that is, $W \subseteq W' \Longrightarrow B_\delta(W) \subseteq B_\delta(W')$.
2. B_δ maps compact sets into compact sets; that is, W compact $\Longrightarrow B_\delta(W)$ compact.
3. $B_\delta(\mathcal{F}) \subseteq \mathcal{F}$.

Definition 5 (Self-Generation). *The set* $W \subseteq \mathbb{R}^n$ *is* self-generating *if every payoff in* W *is decomposable on* W*, that is, if* $W \subseteq B_\delta(W)$*.*

The interest in self-generating sets follows from the next result.

Theorem 2 (Self-Generation and PPE Payoffs). *If the set* $W \subseteq \mathbb{R}^n$ *is bounded and self-generating, then* $B_\delta(W) \subseteq E_\delta$ *(and hence* $W \subseteq E_\delta$*).*

Proof. Pick any $v \in B_\delta(W)$. We will exhibit a public strategy for the repeated game that yields payoff v, and check that the strategy is a PPE.

Since W is self-generating, each $v \in B_\delta(W)$ is decomposed by some $(\alpha_v, \gamma_v) \in \times_i \Delta(A_i) \times W^Y$. For each $v \in B_\delta(W)$, define the strategy σ, parametrized by $v' \in B_\delta(W)$ that starts at the beginning of the game in state v, and after any history h^t, given the current state $v' \in B_\delta(W)$, specifies $\sigma(h^t) = \alpha_{v'}$ and moves to state $v'' = \gamma_{v'}(y)$ in the next period, as a function of the realized signal y. We first show that states truly correspond to payoffs, i.e. that the payoff from playing σ is indeed v. By construction of σ,

$$\begin{aligned} v &= (1-\delta)u(\sigma(\varnothing)) + \delta \sum_{y^0 \in Y} \pi\big(y^0 \mid \sigma(\varnothing)\big)\gamma_v(y^0) \\ &= (1-\delta)u(\sigma(\varnothing)) + \delta \sum_{y^0 \in Y} \pi\big(y^0 \mid \sigma(\varnothing)\big)\left[(1-\delta)u\big(\sigma\big(y^0\big)\big) + \delta \sum_{y^1 \in Y} \pi\big(y^1 \mid \sigma\big(y^1\big)\big)\gamma_{\gamma_v(y^0)}\big(y^1\big)\right] \\ &= \dots \\ &= (1-\delta)\sum_{s=0}^{t-1}\delta^s \sum_{h^s \in Y^s} u(\sigma(h^s))\,\mathbb{P}_\sigma[h^s] + \delta^t \sum_{h^t \in Y^t} \gamma_v[h^t]\,\mathbb{P}_\sigma[h^t], \end{aligned}$$

where $\mathbb{P}_\sigma[h^t]$ is the probability that the sequence of public signals h^t arises under σ, and $\gamma_v[h^t]$ is the state that is obtained after public history h^t, starting from state v, given the strategy σ. Because $\gamma_v[h^t] \in W$ and W is bounded, $\sum_{h^t \in Y^t} \gamma_v[h^t]\,\mathbb{P}_\sigma[h^t]$ is bounded. Therefore, taking $t \to \infty$ yields

$$\begin{aligned} v &= \lim_{t\to\infty}\left[(1-\delta)\sum_{s=0}^{t-1}\delta^s \sum_{h^s \in Y^s} u(\sigma(h^s))\,\mathbb{P}_\sigma[h^s] + \delta^t \sum_{h^t \in Y^t} \gamma_v[h^t]\,\mathbb{P}_\sigma[h^t]\right] \\ &= \lim_{t\to\infty}(1-\delta)\sum_{s=0}^{t-1}\delta^s \sum_{h^s \in Y^s} u(\sigma(h^s))\,\mathbb{P}_\sigma[h^s] + \lim_{t\to\infty}\delta^t \sum_{h^t \in Y^t} \gamma_v[h^t]\,\mathbb{P}_\sigma[h^t] \\ &= (1-\delta)\sum_{s=0}^{\infty}\delta^s \sum_{h^s \in Y^s} u(\sigma(h^s))\,\mathbb{P}_\sigma[h^s] + 0 \\ &= (1-\delta)\mathbb{E}_\sigma\left[\sum_{t=0}^{\infty}\delta^t u(a^t)\right]. \end{aligned}$$

That is, σ yields payoff v, as was to be shown.

For any h^t, $\gamma_v[h^t]$ is the continuation payoff under σ given the history of public signals h^t. By construction of σ, for any state $v' \in B_\delta(W)$ and any $i \in N$, $\alpha_{i,v'}$ is a best response to $\alpha_{-i,v'}$ given continuation payoffs $\gamma_{v'}$. Thus, there is no history where a player can gain by deviating once and conforming thereafter. That σ is a PPE follows by the one-shot deviation principle. ■

The previous theorem gives us a criterion for identifying subsets of the set of PPE payoffs, because any self-generating and bounded set is such a subset. The next proposition tells us that the set of PPE payoffs is a fixed point of the function B_δ.[8] By Theorem 2, every bounded fixed point of B_δ must be a subset of E_δ; thus, E_δ is actually the largest bounded fixed point of E_δ.

Proposition 1 (Factorization). *It holds that* $E_\delta = B_\delta(E_\delta)$.

Proof. If E_δ is self-generating (i.e. $E_\delta \subseteq B_\delta(E_\delta)$), then (because E_δ is clearly bounded, being a subset of $\mathcal{F}^*$) by Theorem 2, $B_\delta(E_\delta) \subseteq E_\delta$, and so $E_\delta = B_\delta(E_\delta)$. Thus, it suffices to show that $E_\delta \subseteq B_\delta(E_\delta)$.

[8] We say that a set of payoffs can be *factorized* if it is a fixed point of B_δ. Hence the name of Proposition 1.

Pick any $v \in E_\delta$ and let σ be a PPE with value $v(\sigma) = v$. Set $\alpha \coloneqq \sigma(\varnothing)$ and $\gamma(y) \coloneqq v(\sigma|_y)$ for all $y \in Y$. It is enough to show that α is enforced by γ in E_δ and $V(\alpha, \gamma) = v$. But,

$$\begin{aligned} V(\alpha,\gamma) &\coloneqq (1-\delta)u(\alpha) + \delta \sum_{y\in Y} \pi(y \mid \alpha)\gamma(y) \\ &= (1-\delta)u(\sigma(\varnothing)) + \delta \sum_{y\in Y} \pi(y \mid \sigma(\varnothing))v(\sigma|_y) \\ &= v(\sigma) \\ &= v. \end{aligned}$$

Moreover, since σ is a PPE, $\sigma|_y$ is also a PPE, and so $v(\sigma|_y) \in E_\delta$ for all $y \in Y$; that $\gamma \in {E_\delta}^Y$ follows. Finally, since σ is a PPE, there are no profitable one-shot deviations, and so α is enforced by γ on E_δ. Thus, $v \in B_\delta(E_\delta)$, as was to be shown. ■

Fix $W \subseteq \mathbb{R}^n$. Suppose $B_\delta^k(W)$ for $k \in \mathbb{N}$, $k > 1$, is defined recursively by

$$B_\delta^k(W) \coloneqq \{v \in \mathcal{F} : v \text{ is decomposable on } B_\delta^{k-1}(W) \text{ with discount } \delta\}.$$

The set of feasible payoffs $\mathcal{F}$ is clearly compact. Moreover, we have seen that every payoff that can be decomposed on the set of feasible payoffs must itself be feasible, that is, $B_\delta(\mathcal{F}) \subseteq \mathcal{F}$. Since E_δ is a fixed point of B_δ and B_δ is monotonic, for every integer $k \geq 1$ we have

$$E_\delta \subseteq B_\delta^k(\mathcal{F}) \subseteq \mathcal{F}.$$

In fact, $\{B_\delta^k(\mathcal{F})\}_{k\in\mathbb{N}}$ is a decreasing sequence. Define

$$\mathcal{F}_\infty \coloneqq \bigcap_{k=1}^{\infty} B_\delta^k(\mathcal{F}).$$

Each $B_\delta^k(\mathcal{F})$ is compact and so $\mathcal{F}_\infty$ is compact and nonempty (because $E_\delta \subseteq \mathcal{F}_\infty$). Therefore, we have

$$E_\delta \subseteq \mathcal{F}_\infty \subseteq \cdots \subseteq B_\delta^2(\mathcal{F}) \subseteq B_\delta(\mathcal{F}) \subseteq \mathcal{F}.$$

The following proposition implies that the algorithm of iteratively calculating $B_\delta^k(\mathcal{F})$ computes the set of PPE payoffs. See Judd, Yeltekin and Conklin (2003) for an implementation.

Proposition 2. *The set $\mathcal{F}^\infty$ is self-generating. Thus, $\mathcal{F}^\infty = E_\delta$ and E_δ is compact.*

Proof. We need to show that $\mathcal{F}^\infty \subseteq B_\delta^p(\mathcal{F}^\infty)$. Pick any $v \in \mathcal{F}^\infty$. Thus, $v \in B_\delta^k(\mathcal{F})$ for all $k \in \mathbb{N}$, and so there exists (α^k, γ^k) such that $v = V(\alpha^k, \gamma^k)$ and $\gamma^k \in B_\delta^{k-1}(\mathcal{F})$ for all $y \in Y$. By extracting convergent subsequences if necessary, we can assume the sequence $\{(\alpha^k, \gamma^k)\}_{k\in\mathbb{N}}$ converges to a limit (α^*, γ^*). It remains to show that α^* is enforced by γ^* on $\mathcal{F}^\infty$ and $v = V(\alpha^*, \gamma^*)$. We only verify that $\gamma^*(y) \in \mathcal{F}^\infty$ (since the other parts are trivial). Suppose then that there is some $y \in Y$ such that $\gamma^*(y) \notin \mathcal{F}^\infty$. As $\mathcal{F}^\infty$ is closed, there is an $\varepsilon > 0$ such that

$$\overline{B}_\varepsilon(\gamma^*(y)) \cap \mathcal{F}^\infty = \varnothing,$$

where $\overline{B}_\varepsilon(v)$ is the closed ball of radius ε centered at v. Because $\gamma^k \to \gamma^*$, there exists $K \in \mathbb{N}$ such that for all $k > K$, $\gamma^k(y) \in \overline{B}_\varepsilon(\gamma^*(y))$, which implies

$$\overline{B}_\varepsilon(\gamma^*(y)) \cap \left(\cap_{k' \le k} B_\delta^{k'}(\mathcal{F}) \right) \neq \varnothing \qquad \text{for all } \ k > K$$

because $\gamma^{k+1}(y) \in B_\delta^k(\mathcal{F}) = \cap_{k' \le k} B_\delta^{k'}(\mathcal{F})$). Thus, the collection

$$\left\{ \overline{B}_\varepsilon(\gamma^*(y)) \right\} \cup \{ B_\delta^k(\mathcal{F}) \}_{k > K}$$

has the finite intersection property, and so, by the compactness of $\overline{B}_\varepsilon(\gamma^*(y)) \cup \mathcal{F}$,

$$\overline{B}_\varepsilon(\gamma^*(y)) \cap \mathcal{F}^\infty \neq \varnothing,$$

a contradiction.[9] ■

We now turn to the monotonicity of PPE payoffs with respect to the discount factor. Intuitively, as players become more patient, it should be easier to enforce an action profile because myopic incentives to deviate are now less important. Consequently, we should be able to adjust continuation promises so that incentive constraints are still satisfied, and yet players' total payoffs have not been affected by the change in weighting between flow and continuation values. However, there is a discreteness issue: if the set of available continuations is disconnected, it may not be possible to adjust the continuation value by a sufficiently small amount that the incentive constraint is not violated. On the other hand, if continuations can be chosen from a convex set, then the above intuition is valid.

Proposition 3. *Suppose $W \subseteq W' \subseteq \mathbb{R}^n$ and $W \subseteq B_\delta(W')$. Then, $W \subseteq B_{\delta'}(\mathrm{co}(W'))$ for all $\delta' > \delta$.*

Proof. Fix $v \in W$ and suppose v is decomposed by (α, γ) on W', given δ. Define γ' pointwise as

$$\gamma'(y) := \frac{\delta' - \delta}{\delta'(1-\delta)} v + \frac{\delta(1-\delta')}{\delta'(1-\delta)} \gamma(y).$$

Since $v, \gamma(y) \in W'$ for all $y \in Y$, we have $\gamma' \in \mathrm{co}(W')^Y$. Moreover,

$$\begin{aligned} &(1-\delta')u(\alpha) + \delta' \sum_{y \in Y} \pi(y \mid \alpha)\gamma'(y) \\ &= (1-\delta')u(\alpha) + \frac{\delta' - \delta}{1-\delta} v + \frac{\delta(1-\delta')}{1-\delta} \sum_{y \in Y} \pi(y \mid \alpha)\gamma(y) \\ &= \frac{\delta' - \delta}{1-\delta} v + \frac{1-\delta'}{1-\delta} \left[(1-\delta)u(\alpha) + \delta \sum_{y \in Y} \pi(y \mid \alpha)\gamma(y) \right]. \end{aligned}$$

[9] A collection $\{F_\lambda\}_{\lambda \in \Lambda}$ of subsets of a set X is said to have the finite intersection property if the intersection over any finite subcollection of $\{F_\lambda\}_{\lambda \in \Lambda}$ is nonempty. Moreover, recall that If $\{F_\lambda\}_{\lambda \in \Lambda}$ is a collection of closed sets in a compact metric space and has the finite intersection property, then $\cap_{\lambda \in \Lambda} F_\lambda \neq \varnothing$.

Thus, since α is enforced by γ on W' given δ, it is also enforced by γ' on $\text{co}(W')$ given δ'.[10] Moreover, the term in $[\cdot]$ is the last expression equals v, and so

$$(1-\delta')u(\alpha) + \delta' \sum_{y \in Y} \pi(y \mid \alpha)\gamma'(y) = v.$$

Hence, $v \in B_{\delta'}(\text{co}(W'))$. ■

The next result immediately follows.

Corollary 1. *If $W \subseteq \mathbb{R}^n$ is bounded, convex, and $W \subseteq B_\delta(W)$, then $W \subseteq E_{\delta'}$ for all $\delta' > \delta$. In particular, if E_δ is convex, then $E_\delta \subseteq E_{\delta'}$ for all $\delta' > \delta$.*

When players have access to a public correlating device, E_δ is convex and so weakly increasing in δ. Without public correlating device, there are well-known examples in which E_δ is not convex, no matter how large δ is (see, e.g., Yamamoto (2010)).

Exercise 2. Let G^δ be an infinitely repeated game with imperfect public monitoring with n players. Show that $\mathbb{R}^n \subseteq B_\delta(\mathbb{R}^n)$.

Solution. To generate an arbitrary payoff $v \in \mathbb{R}^n$, one need only couple the current play of a Nash equilibrium of the stage game (thus ensuring enforceability), with payoff profile v^N, with the function $\gamma(y) := v'$ for all y, where v' is such that $(1-\delta)v^N + \delta v' = v$. [Note that the unboundedness of the reals plays a key role in this construction.] ■

Exercise 3. Suppose the following stage game is repeated infinitely often with discount δ.

- Players: $i \in N = \{1, 2\}$.
- Actions: $a_i \in A_i = \{E, S\}$ (exert effort or shirk, unobservable).
- Signals: $y \in Y = \{y^b, y^g\}$ (bad or good, publicly observable).
- Monitoring distribution: $\pi(y^g \mid (E,E)) = p$, $\pi(y^g \mid (E,S)) = \pi(y^g \mid (S,E)) = q$, and $\pi(y^g \mid (S,S)) = r$. Assume that $1 > p > q > r > 0$ and $p - q > q - r$.
- Rewards:

$$g_i(a_i, y) = \begin{cases} \frac{2-p-q}{p-q} & \text{if } (a_i, y) = (E, y^g) \\ -\frac{p+q}{p-q} & \text{if } (a_i, y) = (E, y^b) \\ \frac{2-2r}{q-r} & \text{if } (a_i, y) = (S, y^g) \\ -\frac{2r}{q-r} & \text{if } (a_i, y) = (S, y^b) \end{cases}.$$

1. Find the matrix of expected rewards for the stage game.

[10] *Quiz:* You see why, right?

2. Consider the set $W := \{(v, v), (v', v')\}$, where

$$v := \frac{\delta r}{1 - \delta(p - r)} \qquad \text{and} \qquad v' := \frac{1 - \delta + \delta r}{1 - \delta(p - r)}.$$

For which values of δ is the set W pure-action self-generating?

Solution.

1. The expected rewards of the stage game are as in the following matrix

		Player 2	
		E	S
Player 1	E	$1, 1$	$-1, 2$
	S	$2, -1$	$0, 0$

2. To enforce (v, v), use the action profile (S, S) and $\gamma \in W^Y$ defined pointwise as

$$\gamma(y) := \begin{cases} (v', v') & \text{if } y = y^g \\ (v, v) & \text{if } y = y^b \end{cases}.$$

We need to check that

$$\begin{aligned} v &= (1 - \delta)(0) + \delta v + \delta r(v' - v) \\ v &\geq (1 - \delta)(-1) + \delta v + \delta q(v' - v). \end{aligned}$$

The first constraint is just algebra. The second constraint holds as long as

$$\delta(q - r)(v' - v) \leq 1 - \delta,$$

or, equivalently,

$$\delta \leq \overline{\delta} := \frac{1}{p + q - 2r}.$$

To enforce (v', v'), use the action profile (E, E) and the same $\gamma \in W^Y$ as before. We need to check that

$$\begin{aligned} v' &= (1 - \delta)(1) + \delta v + \delta p(v' - v) \\ v' &\geq (1 - \delta)(2) + \delta v + \delta q(v' - v). \end{aligned}$$

The first constraint is again just algebra. The second constraint holds as long as

$$\delta(p - q)(v' - v) \geq 1 - \delta,$$

or, equivalently,

$$\delta \geq \underline{\delta} := \frac{1}{2p - q - r}.$$

Given our assumptions, we have $0 < \underline{\delta} < \overline{\delta}$. To sum up, the set W is pure-action self-generating for any $\delta \in [\underline{\delta}, \overline{\delta}]$ if $\overline{\delta} < 1$, and for any $\delta \in [\underline{\delta}, 1)$ if $\overline{\delta} \geq 1$. ■

Exercise 4. Suppose the following stage game is repeated infinitely often with discount δ.

- Players: $i \in N = \{1, 2\}$.
- Actions: $a_i \in A_i = \{E, S\}$ (exert effort or shirk, unobservable).
- Signals: $y \in Y = \{y^b, y^g\}$ (bad or good, publicly observable).
- Monitoring distribution: $\pi(y^g \mid (E,E)) = p$, $\pi(y^g \mid (E,S)) = \pi(y^g \mid (S,E)) = q$, and $\pi(y^g \mid (S,S)) = r$. Assume that $1 > p > q > r > 0$ and $3p - 2q > 1$.
- Rewards:

$$g_i(a_i, y) = \begin{cases} \frac{3-p-2q}{p-q} & \text{if } (a_i, y) = (E, y^g) \\ -\frac{p+2q}{p-q} & \text{if } (a_i, y) = (E, y^b) \\ \frac{3(1-r)}{q-r} & \text{if } (a_i, y) = (S, y^g) \\ -\frac{3r}{q-r} & \text{if } (a_i, y) = (S, y^b) \end{cases}.$$

The expected rewards of the stage game are as in the following matrix

		Player 2	
		E	S
Player 1	E	$2, 2$	$-1, 3$
	S	$3, -1$	$0, 0$

The public profile (σ_1, σ_2) is strongly symmetric if, for all public histories h^t, $\sigma_1(h^t) = \sigma_2(h^t)$. This question asks you to prove that, for sufficiently large δ, any payoff in the interval $[0, \overline{v}]$, is the payoff of some strongly symmetric PPE, where

$$\overline{v} := 2 - \frac{1-p}{p-q}$$

and that no payoff larger than $\overline{v}$ is the payoff of some strongly symmetric pure-strategy PPE equilibrium. Strong symmetry implies it is enough to focus on player 1, and the player subscript will often be omitted.

(a) The action profile (S, S) is trivially enforced by any constant continuation $\gamma \in [0, \overline{\gamma}]$ independent of y. Let W^{SS} be the set of values that can be obtained by (S, S) and a constant continuation $\gamma \in [0, \overline{\gamma}]$, i.e.

$$W^{SS} := \{v \in \mathbb{R} : v = (1-\delta)u_1(S,S) + \delta\gamma \text{ for some } \gamma \in [0, \overline{\gamma}]\}.$$

Prove that $W^{SS} = [0, \delta\overline{\gamma}]$.

(b) Let W^{EE} be the set of values that can be decomposed by (E,E) on $[0,\overline{\gamma}]$. It is clear that $W^{EE} := [\gamma', \gamma'']$ for some $\gamma', \gamma'' \in \mathbb{R}$. Calculate γ' by using the smallest possible choices of $\gamma(y^b)$ and $\gamma(y^g)$ in the interval $[0,\overline{\gamma}]$ to enforce (E,E).

(c) Similarly, give an expression for γ'' (that will involve $\gamma(y^g)$) by using the largest possible choices of $\gamma(y^b)$ and $\gamma(y^g)$ in the interval $[0,\overline{\gamma}]$ to enforce (E,E). Argue that $\delta\overline{\gamma} < \gamma''$.

(d) We would like all continuations in $[0,\overline{\gamma}]$ to be themselves decomposable using continuations in $[0,\overline{\gamma}]$, i.e. we would like

$$[0,\overline{\gamma}] \subseteq W^{SS} \cup W^{EE}.$$

Since $\delta\overline{\gamma} < \gamma''$, we then would like $\overline{\gamma} \leq \gamma''$. Moreover, since we would like $[0,\overline{\gamma}]$ to be the largest such interval, we have $\overline{\gamma} = \gamma''$. What is the relationship between γ'' and $\overline{v}$?

(e) For what values of δ do we have $[0,\overline{\gamma}] = W^{SS} \cup W^{EE}$?

Solution.

(a) Since $u_1(S,S) = 0$, we have

$$\begin{aligned} W^{SS} &:= \{v \in \mathbb{R} : v = (1-\delta)u_1(S,S) + \delta\gamma \text{ for some } \gamma \in [0,\overline{\gamma}]\} \\ &= \{v \in \mathbb{R} : v = \delta\gamma \text{ for some } \gamma \in [0,\overline{\gamma}]\} \\ &= [0, \delta\overline{\gamma}]. \end{aligned}$$

as desired.

(b) The payoff vector v is decomposable by (E,E) on $[0,\overline{\gamma}]$ id there exists $\gamma \in Y^A$ such that

$$v = (1-\delta)u_1(E,E) + \delta[p\gamma(y^g) + (1-p)\gamma(y^b)] \quad (2)$$

$$v \geq (1-\delta)u_1(S,E) + \delta[a\gamma(y^g) + (1-q)\gamma(y^b)] \quad (3)$$

From (2) and (2) we have

$$(1-\delta)2 + \delta[p\gamma(y^g) + (1-p)\gamma(y^b)] \geq (1-\delta)3 + \delta[a\gamma(y^g) + (1-q)\gamma(y^b)],$$

which implies

$$\gamma(y^g) - \gamma(y^b) \geq \frac{1-\delta}{\delta(p-q)}. \quad (4)$$

Since $v = (1-\delta)2 + \delta[p\gamma(y^g) + (1-p)\gamma(y^b)] = (1-\delta)2 + \delta p(\gamma(y^g) - \gamma(y^b)) + \delta\gamma(y^b)$, we have

$$\gamma' = \min_{\gamma(y^b),\gamma(y^g)\in[0,\overline{\gamma}]} (1-\delta)2 + \delta[p\gamma(y^g) + (1-p)\gamma(y^b)]$$

subject to (4). This gives

$$\gamma' = (1-\delta)\left(2 + \frac{p}{p-q}\right),$$

with

$$\gamma(y^g) = \frac{1-\delta}{\delta(p-q)} \qquad \text{and} \qquad \gamma(y^b) = 0.$$

(c) Similarly,

$$\gamma'' = \max_{\gamma(y^b),\gamma(y^g)\in[0,\overline{\gamma}]} (1-\delta)2 + \delta[p\gamma(y^g) + (1-p)\gamma(y^b)].$$

subject to (4). This gives

$$\gamma'' = (1-\delta)\left(2 - \frac{1-p}{p-q}\right) + \delta\gamma(y^g) = (1-\delta)\overline{v} + \delta\overline{\gamma},$$

with

$$\gamma(y^g) = \overline{\gamma} \qquad \text{and} \qquad \gamma(y^b) = \overline{\gamma} - \frac{1-\delta}{\delta(p-q)}.$$

Under the assumption that $3p - 2q > 1$, we have

$$\overline{v} = 2 - \frac{1-p}{p-q} = \frac{3p-2q-1}{p-q} > 0,$$

and so $\gamma'' > \delta\overline{\gamma}$.

(d) We have

$$\gamma'' = (1-\delta)\overline{v} + \delta\overline{\gamma} = (1-\delta)\overline{v} + \delta\gamma'',$$

which implies

$$\gamma'' = \overline{\gamma} = \overline{v}.$$

(e) (i) Since $\overline{\gamma} > \delta\overline{\gamma}$, we have $W^{SS} \subseteq [0,\overline{\gamma}]$. Moreover, since $\gamma' = (1-\delta)(2 + p/(p-q)) > 0$ and $\gamma'' = \overline{\gamma}$, we have $W^{EE} \subseteq [0,\overline{\gamma}]$. Therefore

$$W^{SS} \cup W^{EE} \subseteq [0,\overline{\gamma}].$$

(ii) Note that

$$\delta\overline{\gamma} \geq \gamma' \iff \delta\left(2 - \frac{1-p}{p-q}\right) \geq (1-\delta)\left(2 + \frac{p}{p-q}\right)$$
$$\iff \delta \geq \underline{\delta} := \frac{3p-2q}{6p-4q-1} \in (0,1).$$

Thus, for any $\delta \in (\underline{\delta}, 1)$ we have

$$[0,\overline{\gamma}] \subseteq W^{SS} \cup W^{EE}.$$

From (i) and (ii) we conclude that for any $\delta \in (\underline{\delta}, 1)$ we have $[0,\overline{\gamma}] = W^{SS} \cup W^{EE}$.

3.1 Bang-Bang Property of PPE

Let $\text{ext}(W)$ be the set of extreme points of $\text{co}(W)$.[11] The following holds true.

Proposition 4. *Let $W \subseteq \mathbb{R}^n$ be compact. Then, if a public correlating device is available, $B_\delta(W) = B_\delta(\text{ext}(W))$.*

The intuition behind the result is simple. First, since W is compact, $\text{ext}(W) \subseteq W$. Thus, by monotonicity of B_δ,

$$B_\delta(\text{ext}(W)) \subseteq B_\delta(W). \tag{5}$$

Second, any point in $\text{co}(W)$ is some convex combination of extreme points of $\text{co}(W)$ (that is, by the Krein-Milman theorem, we have $\text{co}(W) = \text{co}(\text{ext}(W))$). Therefore the public correlating device can induce an appropriate randomization over the points on $\text{ext}(W)$ to generate any point in $\text{co}(W)$ after any realization of public signals, so that $B_\delta(\text{co}(W)) = B_\delta(\text{ext}(W))$. Since $W \subseteq \text{co}(W)$ and B_δ is monotone, it follows that

$$B_\delta(W) \subseteq B_\delta(\text{ext}(W)). \tag{6}$$

The statement in Proposition 4 thus follows from (5) and (6).

There are two things to notice. First, the result also holds for games with perfect monitoring. Second, the result is similar to the one on "simple strategies" for repeated games with perfect monitoring, in the sense that one can restrict attention to only the extreme points of the equilibrium payoff set to support any PPE payoff. That is, we have the following.

Corollary 2. *It holds that $E_\delta = B_\delta(\text{ext}(E_\delta))$.*

Proof. Since E_δ is compact, by Proposition 4 we have $B_\delta(E_\delta) = B_\delta(\text{ext}(E_\delta))$. Moreover, by Proposition 1 we have $E_\delta = B_\delta(E_\delta)$. The desired result follows. ∎

In the original paper by Abreu et al. (1990), the range Y of public signal is a subset of a finite dimensional Euclidean space and $\pi(\cdot \mid a)$ is a density function on Y derived from an absolutely continuous measure on Y (and σ-algebra of Y). They show that a public randomization is not required in such setting. That is, they establish the following.

Theorem 3. *Suppose the signals are distributed absolutely continuously with respect to the Lebesgue measure on a subset of $\mathbb{R}^k$, for some k. Suppose $W \subseteq \mathbb{R}^n$ is compact. Then, $B_\delta(W) = B_\delta(\text{ext}(W))$.*

Again, it follows immediately that $E_\delta = B_\delta(\text{ext}(E_\delta))$: any equilibrium payoff can be generated by using only extreme points of the equilibrium payoff set as continuation values. Thus, any equilibrium *can* have a so-called *bang-bang structure*. What is even more surprising is another result in Abreu et al. (1990), which says that, under some additional conditions, continuation PPE payoff *must* take values in $\text{ext}(E_\delta)$.

[11] Let C be a convex subset of a real vector space. A point $x \in C$ is called an extreme point of C if it does not lie in any open line segment joining two points of C. That is, x an extreme point of C if the following holds: $\langle x = \lambda y + (1-\lambda)z,\ \lambda \in (0,1),\ y \neq z \rangle \implies \langle \text{either } y \notin C \text{ or } z \notin C \text{ (or both)} \rangle$.

Exercise 5. Suppose the signals are distributed absolutely continuously with respect to the Lebesgue measure on a subset of $\mathbb{R}^k$, for some k. Suppose $W \subseteq \mathbb{R}^n$ is compact. Show that $B_\delta(W) = B_\delta(\text{co}(W))$. [Feel free to use the fact that the results in these notes—which assume Y is finite—also apply to signals that are distributed absolutely continuously with respect to the Lebesgue measure on a subset of $\mathbb{R}^k$, for some k.]

Solution. Since players have access to a public correlating device, $B_\delta(\text{co}(W)) = B_\delta(\text{ext}(W))$. Moreover, by Theorem 3, $B_\delta(W) = B_\delta(ext(W))$. The desired result follows. ■

3.2 Impact of Increased Precision

In this section, we present a result, due to Kandori (1992), showing that improving the precision of the public signals cannot reduce the set of equilibrium payoffs.

A natural ranking of the informativeness of the public signals is provided by Blackwell (1950)'s *partial ordering of experiments*. We can view the realized signal y as the result of an experiment about the underlying space of uncertainty, the space A of pure-action profiles. Two different public monitoring distributions (with different signal spaces) can then be viewed as two different experiments. Let R denote the $|A| \times |Y|$-matrix whose a-th row corresponds to $\pi(\cdot \mid a)$, the probability distribution over Y conditional on the action profile a. We can construct a noisier experiment from π by assuming that when y is realized under π, y is observed with probability $1-\varepsilon$ and a uniform (thus uninformative) draw from Y is observed with probability ε. Denoting this distribution by π' and the corresponding probability matrix R', we have $R' = RQ$, where

$$Q = \begin{bmatrix} (1-\varepsilon)+\varepsilon/|Y| & \varepsilon/|Y| & \dots & \varepsilon/|Y| \\ \varepsilon/|Y| & (1-\varepsilon)+\varepsilon/|Y| & & \vdots \\ \vdots & & \ddots & \varepsilon/|Y| \\ \varepsilon/|Y| & \dots & \varepsilon/|Y| & (1-\varepsilon)+\varepsilon/|Y| \end{bmatrix}$$

Note that Q is a stochastic matrix, i.e. a nonnegative matrix whose rows sum to 1. More generally, we define the following.

Definition 6 (Garbling)**.** *The public monitoring distribution (Y', π') is a* garbling *of (Y, π) if there exists a stochastic matrix Q such that $R' = RQ$.*

Note that there is no requirement that the signal spaces Y and Y' bear any particular relationship (in particular, Y' may have more or less elements than Y).

Denote as $B_\delta(W'; \pi)$ the set of payoffs that can be decomposed on $W' \subseteq \mathbb{R}^n$ when the discount factor is δ and the public monitoring distribution is π. Then, we have the following.

Proposition 5. *Suppose the public monitoring distribution (Y', π') is a garbling of (Y, π), and $W \subseteq B_\delta(W'; \pi')$. Then, $W \subseteq B_\delta(\text{co}(W'); \pi)$.*

For instance, if players use a public correlating device, the set of continuation payoffs is convex. Hence, a more informative signal structure must give at least as large a set of self-generating payoffs. An immediate corollary is that the set of PPE payoffs is at least weakly increasing as the monitoring becomes more precise.

Proof. For any $\gamma \in \mathbb{R}^{nY}$, let γ_i be the vector in $\mathbb{R}^{|Y|}$ describing player i's continuation values under γ after different signals. Pick any $v \in W$ and suppose v is decomposed by (α, γ') on W' under the public monitoring distribution π'. For any action profile α', denote the implied vector of probabilities on A also by α'. Player i's expected continuation value under π' and γ' from any action profile π' is then

$$\sum_{y \in Y} \pi'(y \mid \alpha')\gamma'(y) = \alpha' R' \gamma_i'.$$

(Y', π') is a garbling of (Y, π), there exists a stochastic matrix Q such that $R' = RQ$. Defining $\gamma_i \coloneqq Q\gamma_i'$, we get

$$\sum_{y \in Y} \pi'(y \mid \alpha')\gamma'(y) = \alpha' RQ\gamma_i' = \alpha' R\gamma_i = \sum_{y \in Y} \pi(y \mid \alpha')\gamma(y).$$

In other words, for all action profiles α', player i's expected continuation value under π and γ is the same as that under π' and γ'. Hence, α is enforced by γ under π.

To show that v is decomposed by (α, γ') on $\text{co}(W')$, it remains to show that $\gamma \in \text{co}(W')^Y$. Given the vectors $\gamma_i \in \mathbb{R}^{|Y|}$ for all i, let $\gamma(y) \in \mathbb{R}^n$ describe the vector of continuation values for all the players for $y \in Y$. Since Q is independent of i,

$$\gamma(y) = \sum_{y' \in Y'} q_{yy'}\gamma'(y'),$$

where $Q = [q_{yy'}]$. Finally, because Q is a stochastic matrix, implying $\gamma(y)$ is a convex combination of the $\gamma'(y')$, we have $\gamma(y) \in \text{co}(W')$, as desired. ■

4 Folk Theorem with Imperfect Public Monitoring

4.1 An Example

This is a simplified version of the partnership game in Radner, Myerson and Maskin (1986). Suppose the following stage game is repeated infinitely often with discount δ.

- Players: $i \in N = \{1, 2\}$.
- Actions: $a_i \in A_i = \{E, S\}$ (exert effort or shirk, unobservable).
- Signals: $y \in Y = \{y^b, y^g\}$ (bad or good, publicly observable).
- Monitoring distribution: $\pi(y^b \mid (E, E)) = 1/3$, $\pi(y^b \mid (E, S)) = \pi(y^b \mid (S, E)) = 2/3$, and $\pi(y^b \mid (S, S)) = 3/4$.
- Rewards: $g_i(a_i, y) = \pi(y)/2 - c(a_i)$, where: $\pi(y^b) = 0$ and $\pi(y^g) = 12$ are revenues; $c(E) = 3$ and $c(S) = 0$ are costs.

Thus, the expected rewards of the stage game are as in the following matrix (this is a prisoners' dilemma)

		Player 2	
		E	S
Player 1	E	$1, 1$	$-1, 2$
	S	$2, -1$	$0, 0$

Claim 1. Let $(\overline{v}_1, \overline{v}_2)$ the PPE equilibrium payoff vector that maximizes $v_1 + v_2$.[12] Then, $\overline{v}_1 + \overline{v}_2 \leq 1$ for any $\delta \in [0, 1)$. That is, PPE payoffs are bounded away from efficiency.

Proof. Suppose $\overline{v}_1 + \overline{v}_2 > 1$. Both players have to play E with positive probability in the first period to support $(\overline{v}_1, \overline{v}_2)$.[13] Let α_i be such probability for $i = 1, 2$. Then,

$$\overline{v}_i = (1-\delta)u_i(E, \alpha_j) + \delta \left\{ \alpha_j \left(\frac{1}{3}\gamma_i(y^b) + \frac{2}{3}\gamma_i(y^g) \right) + (1-\alpha_j) \left(\frac{2}{3}\gamma_i(y^b) + \frac{1}{3}\gamma_i(y^g), \right) \right\} \tag{7}$$

where $\gamma_i(y)$ is player i's continuation payoff after the realization of signal y. The incentive constraint for player i is

$$(1-\delta) \leq \delta \frac{\gamma_i(y^g) - \gamma_i(y^b)}{3}. \tag{8}$$

Since (7) can be rewritten as

$$\overline{v}_i = (1-\delta)u_i(E, \alpha_j) + \delta \left\{ \gamma_i(y^g) - (\alpha_j + 2(1-\alpha_j)) \frac{\gamma_i(y^g) - \gamma_i(y^b)}{3} \right\} v \tag{9}$$

(8) and (9) imply that

$$\overline{v}_i \leq (1-\delta)u_i(E, \alpha_j) + \delta\gamma_i(y^g) - (\alpha_j + 2(1-\alpha_j))(1-\delta). \tag{10}$$

Summing up (10) for $i = 1, 2$ yields

$$\overline{v}_1 + \overline{v}_2 \leq (1-\delta)[u_1(E, \alpha_2) + u_2(E, \alpha_1)] + \delta[\gamma_1(y^g) + \gamma_2(y^g)] - (4 - \alpha_1 - \alpha_2)(1-\delta). \tag{11}$$

Since $u_1(E, \alpha_2) + u_2(E, \alpha_1) \leq 2$ and $\gamma_1(y^g) + \gamma_2(y^g) \leq \overline{v}_1 + \overline{v}_2$, (11) implies that

$$\overline{v}_1 + \overline{v}_2 \leq 2 - (4 - \alpha_1 - \alpha_2) \leq 0,$$

which is a contradiction. ■

PPE payoffs are bounded away from efficiency in this example despite the fact that the profile (E, E), as well as any other profile, is enforceable. Thus, we cannot obtain a Folk Theorem in this particular model. The problem is that to induce both players to exert effort, we must threaten them both with a low continuation payoff if $y = y^b$ (i.e. we must have $\gamma_i(y^b) < \gamma_i(y^g)$ for $i = 1, 2$). This need to "punish" both players for low output (which occurs with positive probability even when both players exert effort) is what creates the inefficiency.

[12]Recall that E_δ is compact.

[13]If not, the first-period rewards could sum to at most 1, implying either $\gamma_1(y^b) + \gamma_2(y^b) > \overline{v}_1 + \overline{v}_2$ or $\gamma_1(y^g) + \gamma_2(y^g) > \overline{v}_1 + \overline{v}_2$, a contradiction.

Remark 1. Radner (1986) shows that the efficient payoff vector $(1, 1)$ can be achieved as PPE when players do not discount the future ($\delta = 1$). So, this example also shows a discontinuity of E_δ at $\delta = 1$.

In the above example, it is critical that there is only one signal, y^b, which is informative about players' deviations. Now suppose that there are two public signals, y_i for $i = 1, 2$, such that y_i is more informative about player i's deviation. In such case, players may be able to punish player i by transferring continuation payoffs from player i to player j when y_i is observed. This is a more efficient punishment because it is based on a transfer of utility, not on a waste of it. To illustrate this point, let us modify the above model as follows (this example is due to Fudenberg, Levin and Maskin (1994)).

- Signals: $y \in Y = \{y_1, y_2, y^g\}$.
- Monitoring distribution:

$$
\begin{array}{lll}
\pi(y^g \mid (E,E)) = 2/3 & \pi(y_1 \mid (E,E)) = 1/6 & \pi(y_2 \mid (E,E)) = 1/6 \\
\pi(y^g \mid (S,E)) = 1/3 & \pi(y_1 \mid (S,E)) = 1/2 & \pi(y_2 \mid (S,E)) = 1/6 \\
\pi(y^g \mid (E,S)) = 1/3 & \pi(y_1 \mid (E,S)) = 1/6 & \pi(y_2 \mid (E,S)) = 1/2 \\
\pi(y^g \mid (S,S)) = 1/4 & \pi(y_1 \mid (S,S)) = 3/8 & \pi(y_2 \mid (S,S)) = 3/8.
\end{array}
$$

- Revenues: $\pi(y^g) = 12$, $\pi(y_1) = \pi(y_2) = 0$.

The matrix of expected rewards of the stage game is the same as before.

Let $\mathcal{F}^*$ be the set of feasible and individually rational (not strictly so) expected rewards of the stage game. Pick any closed ball $W \subseteq \text{Int}(\mathcal{F}^*)$. We will now argue that, for δ sufficiently close to 1, we have $W \subseteq E_\delta$. Hence, the Folk Theorem extends to this game.

From Theorem 2, it is enough to show that $W \subseteq B_\delta$ for δ sufficiently close to 1. Divide W into four parts, A, B, C, and D [figure in class!]. Consider $v = (v_1, v_2)$ on the boundary of A. To decompose v, choose $a = (E, E)$. To ensure that $\gamma(y^g) := (\gamma_1(y^g), \gamma_2(y^g))$, $\gamma(y_1) := (\gamma_1(y_1), \gamma_2(y_1))$, and $\gamma(y_2) := (\gamma_1(y_2), \gamma_2(y_2))$ are in W, we select them to lie along a hyperplane P'_v parallel to the hyperplane P_v which is tangent to W at v. Thus, (E, E) will be *enforceable with respect to the hyperplane* P'_v. Note that $P'_v := \{(v'_1, v'_2) \in \mathbb{R}^2 : \beta_1 v'_1 + \beta_2 v'_2 = c\}$ for some $\beta_1, \beta_2, c \in \mathbb{R}$. Hence, selecting $\gamma(y^g)$, $\gamma(y_1)$, and $\gamma(y_2)$ amounts to solving the system

$$v_1 \geq (1-\delta)(2) + \delta\left(\frac{1}{3}\gamma_1(y^g) + \frac{1}{2}\gamma_1(y_1) + \frac{1}{6}\gamma_1(y_2)\right) \tag{12}$$

$$v_2 \geq (1-\delta)(2) + \delta\left(\frac{1}{3}\gamma_2(y^g) + \frac{1}{6}\gamma_2(y_1) + \frac{1}{2}\gamma_2(y_2)\right) \tag{13}$$

$$v_1 = (1-\delta) + \delta\left(\frac{2}{3}\gamma_1(y^g) + \frac{1}{6}\gamma_1(y_1) + \frac{1}{6}\gamma_1(y_2)\right) \tag{14}$$

where: (12) and (13) ensure that player 1 and player 2 do not gain from shirking; (14) ensures that player 1's expected payoff is v_1. We can omit the requirement that player 2's expected payoff be v_2 since, given (14) and the stipulation that $\gamma(y^g)$, $\gamma(y_1)$, and $\gamma(y_2)$ lie in P'_v, this is ensured automatically.

If δ is sufficiently close to 1, $\gamma(y^g)$, $\gamma(y_1)$, and $\gamma(y_2)$ will lie near the intersection of P'_v and the hyperplane through $u(E, E)$ and v, and so will line in W. All that is required is that W is smooth (i.e. has no "kinks") at v (which is trivially the case for a closed ball). Because continuation payoffs $\gamma(y^g)$, $\gamma(y_1)$, and $\gamma(y_2)$ lie in P'_v, we can replace each $\gamma_2(y)$ in (13) by $(c - \beta_1\gamma_1(y))/\beta_2$ and rewrite (12)-(14) as equalities to obtain

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{2} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{6} & \frac{1}{6} \end{bmatrix} \begin{pmatrix} \gamma_1(y^g) \\ \gamma_1(y_1) \\ \gamma_1(y_2) \end{pmatrix} = \begin{pmatrix} \frac{v_1 - 2(1-\delta)}{\delta} \\ \frac{\beta_2(2(1-\delta) - v_2) + c\delta}{\beta_1\delta} \\ \frac{v_1 - (1-\delta)}{\delta} \end{pmatrix}. \tag{15}$$

Now the matrix in (15) has full rank, and so (15) can be solved with equality. In the terminology of the next section, the profile (E, E) has *pairwise full rank*. Roughly speaking, a profile α has this property for players i and j if every deviation that i or j could make from α induces a different distribution over public signals. Pairwise full rank ensures that the profile (E, E) is enforceable with respect to the hyperplane P'_v (or, for that matter, with respect to almost any other hyperplane). From [figure in class!], we see that this means that, when y_1 or y_2 occur, it is not necessary to punish both players, in contrast with the two-signal partnership game. The earlier game had too few public signals to have profiles with pairwise full rank.

We have shown that any point $v \in A$ can be decomposed so that the continuation payoffs lie in W for δ sufficiently close to 1. Points in B, C, and D can be decomposed similarly. Indeed, W is decomposable with respect to hyperplanes that are parallel translates of the tangent hyperplanes (since each boundary point v is decomposable in such a way that, as in [figure in class!], the continuation payoffs lie on a hyperplane P'_v parallel to the tangent hyperplane P_v to W at v). The key trick to establish the Folk Theorem in more general (than the partnership example) repeated games with imperfect public monitoring is to show that any set W that is decomposable on tangent hyperplanes belongs to E_δ as $\delta \to 1$.

4.2 Public Monitoring Folk Theorem

We now present the formal result due to Fudenberg et al. (1994).

If an action profile α is to be enforceable, and it is not an equilibrium of the stage game, then deviations must lead to different expected continuations. That is, the distribution over signals should be different if a player deviates. A sufficient condition is that the distribution over signals induced by α be different from the distribution induced by any profile (α'_i, α_{-i}) with $\alpha'_i \neq \alpha$, i.e. any two deviations by the same player induce different distributions. This is clearly implied by the following.

Definition 7 (Individual Full Rank)**.** *The profile α has* individual full rank for player i *if the $|A_i| \times |Y|$ matrix $R_i(\alpha_{-i})$ with elements $R_i(\alpha_{-i})_{a_i y} := \pi(y \mid a_i, \alpha_{-i})$ has full row rank (i.e. the collection of probability distributions $\{\pi(\cdot \mid a_i, \alpha_{-i})\}_{a_i \in A_i}$ is linearly independent). If this holds for all players i, then α has* individual full rank*.*

Individual full rank requires $|Y| \geq \max_{i \in N} |A_i|$, and ensures that the signals generated by any (possibly mixed) action α_i are statistically distinguishable from those generated by any other action $\alpha'_i \neq \alpha_i$. Because both α_i and α'_i are convex combinations of pure actions, it would suffice that each $\pi(y \mid \alpha)$ is a unique convex combination of the distributions in $\{\pi(\cdot \mid a_i, \alpha_{-i})\}_{a_i \in A_i}$ (such a condition is discussed in Kandori and Matsushima (1998), who establish the public monitoring Folk Theorem under slightly weaker conditions than those presented below).

Definition 8 (Pairwise Full Rank)**.** *The profile α has* pairwise full rank for players i and j *if the $(|A_i| \times |A_j|) \times |Y|$ matrix*

$$R_{ij}(\alpha) \coloneqq \begin{bmatrix} R_i(\alpha_{-i}) \\ R_j(\alpha_{-j}) \end{bmatrix}$$

has rank $|A_i| + |A_j| - 1$ (i.e. the collection of probability distributions

$$\left\{\{\pi(\cdot \mid a_i, \alpha_{-i})\}_{a_i \in A_i}, \{\pi(\cdot \mid a_j, \alpha_{-j})\}_{a_j \in A_j}\right\}$$

admit only one linear dependency).

Note that $|A_i| + |A_j| - 1$ is the maximum feasible rank for $R_{ij}(\alpha)$ because $\alpha_i R_i(\alpha_{-i}) = \alpha_j R_i(\alpha_{-j})$. Moreover, if α has pairwise full rank for players i and j, then α has individual full rank for players i and j.

Pairwise full rank says that a deviation by player i leads to a distribution over signals that is different from that induced by any deviation by player j. Thus, if everyone is playing α, then not only can i's actions be distinguished, and j's actions be distinguished, but i's actions can be distinguished from j's actions. If you are a statistics/econometrics type, you can think of this just like statistical identification. Here, i's action is the parameter. To identify it, you need the probability distribution over observables to change when it changes. Pairwise full rank is what you need to identify both a_i and a_j at the same time—we only require this for some profile played by the others, $k \neq i, j$.

Theorem 4 (Public Monitoring Folk Theorem)**.** *Let $\mathcal{F}^*$ be the set of feasible and individually rational rewards. Suppose: (i) $\dim(\mathcal{F}^*) = n$; (ii) every pure action profile has individual full rank; (iii) for any $i, j \in N$ with $i \neq j$, there exists an action profile which has pairwise full-rank for i and j. Then, for any closed set $W \subseteq \mathrm{Int}(\mathcal{F}^*)$, there exists $\underline{\delta} < 1$ such that $W \subseteq E_\delta$ for all $\delta \in (\underline{\delta}, 1)$.*

The Folk Theorem applies to payoff vectors in the interior of $\mathcal{F}^*$. Generally, we cannot get exact efficiency with imperfect monitoring. The argument is simple and illustrative. Suppose that $\pi(\cdot \mid a)$ has a support that is independent of a (as in our example). Moreover, suppose that v is an extreme point of $\mathcal{F}^*$ but not a static Nash equilibrium reward (as $(1,1)$ in our example). Because v is an extreme point, the only sequence of payoffs that gives expected payoff v must have payoffs v in every period. So if a PPE gives v, the first period strategies must specify a profile α with $u(\alpha) = v$, and for any signal $y \in Y$, the continuation payoffs must be $\gamma(y) = v$. But then continuation payoffs are independent of today's signal, so unless α happens to be a static Nash equilibrium (which it is not by assumption), someone will want to deviate.

5 Private Strategies: An Example

What would happen if we allow for private strategies? Here, we just study a simple two-period game based on Kandori and Obara (2006) to see some possible role of private strategies. In particular, this example shows that players can sometimes make better use of information by using private strategies and that efficiency in repeated games can be improved by using such strategies.

There are two players and each player's payoff is the sum of the rewards he gets in each period. Suppose that the first-stage game is given by

		Player 2: E	Player 2: S
Player 1	E	$3, 3$	$-1, 4$
	S	$4, -1$	$0, 0$

At the end of the first period, either signal y^b or signal y^g is observed, and players play the following second-stage game

		Player 2: L	Player 2: R
Player 1	T	K, K	$0, K$
	B	$K, 0$	$0, 0$

In this second-stage game, player i's action completely determines player j's reward. Also, note that any action profile is a Nash equilibrium of the second-stage game and any reward vector in $[0, K] \times [0, K]$ can be an equilibrium reward vector in the second stage.

Let's assume the following monitoring distribution: $\pi(y^b \mid (E, E)) = p^0 = 1/20$, $\pi(y^b \mid (E, S)) = \pi(y^b \mid (S, E)) = p^1 = 1/10$, and $\pi(y^b \mid (S, S)) = p^2 = 8/10$. That is, y^b is more likely to be observed as more players play S.

Best strongly symmetric pure PPE. By the usual analysis, we have

$$\begin{aligned} \overline{V}_{pure} &= 3 + (1 - p^0)K + p^0(K - d) \\ 1 &\leq (p^1 - p^0)d, \qquad \text{(incentive constraint)} \end{aligned}$$

where d is a degree of punishment. Since d should be minimized, the incentive constraint should hold with equality. Hence,

$$\overline{V}_{pure} = 3 + K - \frac{1}{\frac{p^1}{p^0} - 1} = 2 + K.$$

Note that the likelihood ratio

$$\frac{p^1}{p^0} = \frac{\pi(y^b \mid (S, E))}{\pi(y^b \mid (E, E))} = 2$$

is the critical factor to determine this bound.

Best strongly symmetric mixed PPE. Since players are indifferent between E and S, they can mix between them if they like. Then, we again have two equations:

$$\begin{aligned} \overline{V}_{mixed} &= 3(1-q) - q + [1-(1-q)p^0 - qp^1]K + [(1-q)p^0 + qp^1](K-d) \\ 1 &\leq [(1-q)(p^1-p^0) + q(p^2-p^1)]d, \qquad \text{(incentive constraint)} \end{aligned}$$

Again, since d should be minimized, the incentive constraint should hold with equality. Hence,

$$\overline{V}_{mixed} = 3 - 4q + K - \frac{1}{\frac{(1-q)p^1+qp^2}{(1-q)p^0+qp^1} - 1}.$$

Is there any reason to mix between E and S in this way? Note that the likelihood ratio with S being played is

$$\frac{p^2}{p^1} = \frac{\pi(y^b \mid (S,S))}{\pi(y^b \mid (E,S))} = 8,$$

which is much higher than $p^1/p^0 = 2$. When $q = 0$, $\overline{V}_{mixed}$ clearly coincides with $\overline{V}_{pure}$; as q increases from 0, the likelihood ratio goes up to reduce the last term of $\overline{V}_{mixed}$. If this monitoring effect dominates the stage-game reward-reducing effect (i.e. $3-4q$ goes down as q increases), it can be that $\overline{V}_{mixed} > \overline{V}_{pure}$. In fact, you can verify that this is so for $q \in (0, 0.15)$. So, this is an example in which mixed strategies can make a difference within PPE.

Private strategies and private equilibrium. Note that public signal is more informative about player j's action when when player i plays S than when player i plays E. This suggests that the punishment after $(a_i, y) = (E, y^b)$ is a waste of efficiency. Now consider the following strategy: (i) play E with probability $(1-q)$ and S with probability q as in the mixed PPE, but (ii) punish player j in the second stage only when $(a_i, y) = (S, y^b)$, otherwise play T (if $i = 1$) or L (if $i = 2$). This is clearly a private strategy. We can get the equilibrium payoff by solving

$$\begin{aligned} \overline{V}_{private} &= 3(1-q) - q + (1-qp^1)K + qp^1(K-d) \\ 1 &\leq q(p^2-p^1)d, \qquad \text{(incentive constraint)} \end{aligned}$$

where d is a degree of punishment. Again, since d should be minimized, the incentive constraint should hold with equality. Hence,

$$\overline{V}_{private} = 3 - 4q + K - \frac{1}{\frac{p^2}{p^1} - 1},$$

which is clearly larger than $\overline{V}_{mixed}$ for every q. So, for example, for $q = 0.05$ we have

$$\overline{V}_{pure} < \overline{V}_{mixed} < \overline{V}_{private}$$

(as long as K is large enough to generate a level of punishment we need).

Remark 2. After y^b is observed in the first stage, each player does not know the other player's continuation strategy because it depends on his private information (the realization of his action in the first stage). This does not create any problem here because the second stage game has a very peculiar payoff structure. However, it could create many technical problems if we try to apply the idea of this particular equilibrium to infinitely repeated games (or long, but finite horizon repeated games). If you are interested in repeated games with private monitoring (which we do not cover in this course), an excellent starting point is Part III in Mailath and Samuelson (2006).

References

Abreu, Dilip, David Pearce, and Ennio Stacchetti (1986), "Optimal Cartel Equilibria with Imperfect Monitoring." *Journal of Economic Theory*, 39, 251–269.

——— (1990), "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring." *Econometrica*, 58, 1041–1063.

Blackwell, David (1950), "The Comparison of Experiments." *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability.* University of California Press.

Fudenberg, Drew, David K. Levin, and Eric S. Maskin (1994), "The Folk Theorem with Imperfect Public Information." *Econometrica*, 62, 997–1093.

Fudenberg, Drew and Jean Tirole (1991), "Game Theory." *MIT Press*.

Green, Edward J. and Robert H. Porter (1984), "Noncooperative Collusion under Imperfect Private Formation." *Econometrica*, 52, 87–100.

Hörner, Johannes (2016), "Repeated Games, Part II: Imperfect Public Monitoring." *Lecture Notes*.

Judd, Kenneth L., Sevin Yeltekin, and James Conklin (2003), "Computing Supergame Equilibria." *Econometrica*, 71, 1239–1254.

Kandori, Michihiro (1992), "The Use of Information in Repeated Games with Imperfect Monitoring." *The Review of Economic Studies*, 59, 581–594.

Kandori, Michihiro and Hitoshi Matsushima (1998), "Private Observation, Communication and Collusion." *Econometrica*, 66, 627–652.

Kandori, Michihiro and Ichiro Obara (2006), "Efficiency in Repeated Games Revisited: The Role of Private Strategies." *Econometrica*, 74, 499–519.

Kreps, David M. and Robert Wilson (1982), "Sequential Equilibria." *Econometrica*, 50, 863–894.

Levin, Jonathan (2006), "Repeated Games II: Imperfect Public Monitoring." *Lecture Notes*.

Mailath, George J. and Larry Samuelson (2006), "Repeated Games and Reputations: Long-Run Relationships." *Oxford University Press.*

Mertens, Jean-François and T. Parthasarathy (2003), "Equilibria for Discounted Stochastic Games." Sylvain Sorin and Abraham Neyman eds., *Stochastic Games and Applications. NATO Science Series (Series C: Mathematical and Physical Sciences).* Springer, Dordrecht.

Obara, Ichiro (2003), "Repeated Games with Imperfect Public Monitoring." *Lecture Notes.*

Radner, Roy (1986), "Repeated Partnership Games with Imperfect Monitoring and no Discounting." *The Review of Economic Studies*, 53, 43–57.

Radner, Roy, Roger Myerson, and Eric Maskin (1986), "An Example of a Repeated Partnership Game with Discounting and with Uniformly Inefficient Equilibria." *The Review of Economic Studies*, 53, 59–69.

Shapley, Lloyd Stowell (1953), "Stochastic Games." *Proceedings of National Academy of Science*, 39, 1095–1100.

Yamamoto, Yuichi (2010), "The Use of Public Randomization in Discounted Repeated Games." *International Journal of Game Theory*, 39, 431–443.