Repeated Games with Private Monitoring

Ichiro Obara

1 What is "Private Monitoring"?

I Common / Public Monitoring

action (not observable) $\qquad\qquad$ signal (observable)

$$(a_1, \cdots, a_n) \longrightarrow \omega \in Y$$

(a) Perfect Monitoring

$$\omega = (a_1 \cdots a_n)$$

(b) Imperfect Public Monitoring

ω is a random variable

II Private Monitoring

$$\begin{tikzcd} & \omega_1 \in Y_1 \\ (a_1, \cdots, a_n) \arrow[ur] \arrow[r] \arrow[ddr] & \omega_2 \\ & \vdots \\ & \omega_n \end{tikzcd}$$

Examples

noisy observations

- $\Omega_i = A_{-i}$, $P(\omega_i = a_{-i} \mid a) = 1 - \varepsilon$ (almost perfect if ε is small)

  but

  $P(\omega_i = a'_{-i} \mid a) > 0$ for all $a'_{-i} \neq a_{-i}$

Partial observations

- player i observes player $i-1$'s action only.

  (and player 1 observes player n's action)

Reference

Michihiro Kandori "Introduction to Repeated Games with Private Monitoring"

+

all the papers in the same special issue of JET, Vol 102, No 1, 2002.

2. Applications

(a) Secret Price Cutting

- (Repeated) Bertrand Competition (heterogeneous goods)
- Each demand is stochastic
- Price is not observable, each firm's own sale is the only signal (which is private information)

(b) Subjective assessment/evaluations

- Employer use some subjective measure to assess employee's performance.

or

- Employer monitors employee directly, but monitoring is not perfect, it is subject to private noise.

(c) trade with uncertain quality goods

- unobservable effort to produce goods
- Signal is quality of goods (which is stochastic).

Quality can be observed when the good is consumed, hence it is not producer's information.

- Imperfect Public Monitoring (Benchmark)

Characterization	Folk Theorem
Abreu, Pearce and Stacchetti (1990)	Fudenberg, Levine and Maskin (1994)

- Private Monitoring with Communication

Characterization	Folk Theorem
?	• Compte (1998) • Kandori and Matsushima (1998) • Ben-Porath and Kahneman (Partial Monitoring) (1996)

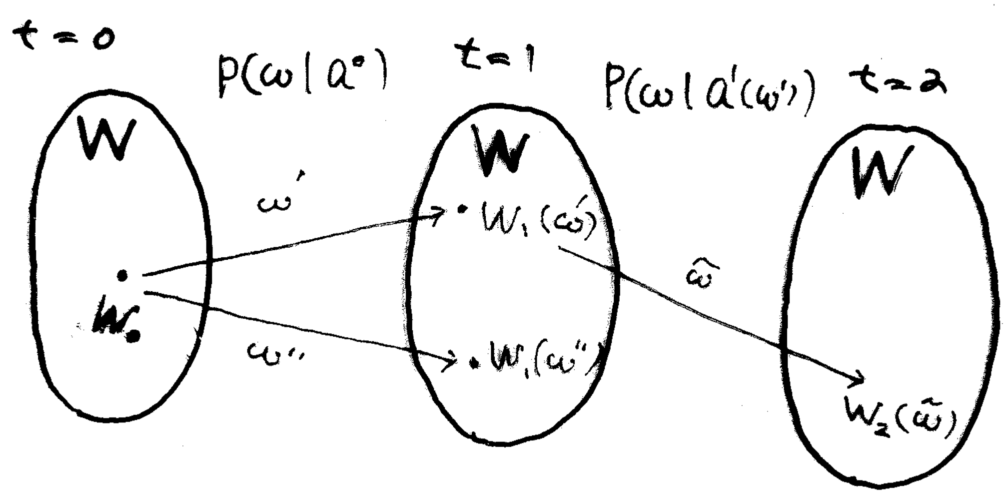
- Private Monitoring w/o Communication

	Characterization	Folk Theorem
PD	?	• Sekiguchi (1997) (Efficiency with AP monitoring) • Bhaskar and Obara (2002) • Ely and Valimaki (2002) (FT with AP monitoring) • Matsushima (mimeo) (FT with CI but not AP)
general games	?	?

I What is the problem?

(1) lack of recursive structure

◎ standard imperfect public monitoring



· W is the set of the equilibrium payoff ~~set~~ based on public strategies. (strategies which depend only on past ω).

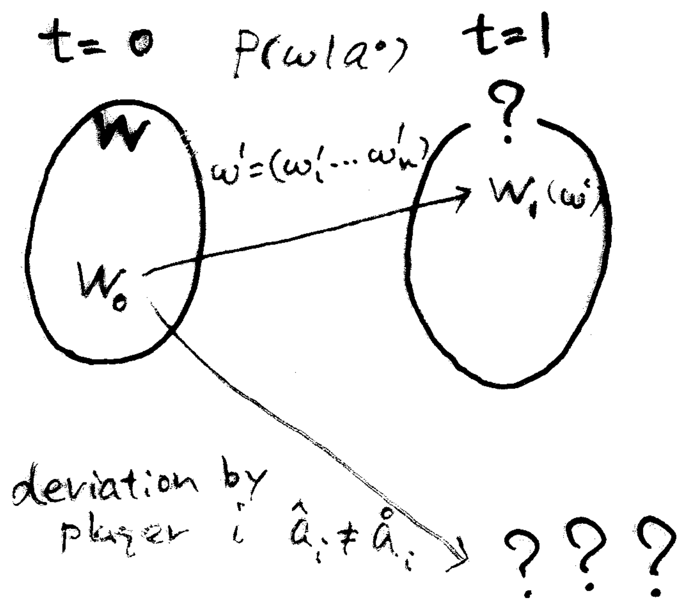
· Note that the set of the continuation equilibrium payoffs is also **W**. This is "recursive structure"

· Let $B(X)$ be the set of equilibrium payoffs of one shot game, which is defined as the stage game + "continuation payoffs" $W(\omega) \in X \ (\subset \mathbb{R}^n)$ for all $\omega \in \Omega$ (consider all such $W(\cdot)$)

Then, W is the maximal "fix point" of the operator B, i.e. $B(W) = W$

(Abreu, Pearce, and Stacchetti, 1990 Econometrica

⊙ Imperfect Private monitoring



Player i's belief is based on $P(\omega \mid \hat{a}_i, a^\circ_{-i})$

Other players' beliefs are based on $P(\omega \mid a^\circ)$

Two Issues

- Continuation game after 1st period is different from the original game. It is an incomplete information repeated game where the first period private signals serve as a correlation device.

  Hence, the ~~set~~ of the (exante) expected payoffs of continuation equilibria after period 2 is different from (larger than) W.

  $\Rightarrow$ recursive structure is lost. Characterization of W is open question.

- Moreover, off the equilibrium path in which player i deviate to play $\hat{a}_i$, the continuation game is ~~an~~ even an incomplete information repeated game for which there is no common prior distribution! We have only a limited knowledge about such games.

(2) Complex Statistical Inference

Each player's strategy depends on his/her private signals. So, to determine the best strategies of one player at each point of time, it is necessary to compute posterior distribution of the other players' private signals. Moreover, this task becomes more and more complex over time as more private information is accumulated.

⟹ It is difficult to construct even one (nontrivial) equilibrium for this class of games.

II Frontier of Research

Most of the recent contributions focuses on the following simple PD.

- $i = 1, 2$
- $A_i = \{C, D\}$
- $\Omega_i = \{c, d\}$
- Stage game expected payoffs

	C	D
C	$1, 1$	$-\ell, 1+g$
D	$1+g, -\ell$	$0, 0$

$g, \ell > 0$

$1 - g + \ell > 0$

- Almost perfect Monitoring

  $P((c, c) \mid CC) = 1 - \varepsilon$, $P((\omega_1, \omega_2) \neq (c, c) \mid C, C) > 0$

  ε small.

  (similar property for all the other action profiles)

- Conditionally independent monitoring.

  $P((\omega_1, \omega_2) \mid a_1, a_2) = p(\omega_2 \mid a_1) P(\omega_1 \mid a_2)$

III Approach 1: Partial Coordination

Sekiguchi + Bhaskar + Obara

(a) 2 period example (Bhaskar + Van Damme)

Period 1.

	C	D
C	1.1	-1.2
D	2.-1	0.0

Period 2

	X	Y
X	3.3	0.0
Y	0.0	1.1

$\omega_i \in \{c.d\}, i=1.2.$

- If private signals are perfectly correlated ($\omega_1 = \omega_2$ always), then (C.C) can be supported by "Play C, play X if the signal is good, ($\omega_1 = \omega_2 = c$) Y if the signal is bad" ($\omega_1 = \omega_2 = d$)

This is, of course, imperfect public monitoring.

- Consider the following information structure
  - $\omega_i = a_j$ with Probability $1-\varepsilon$, $\omega_i \neq a_j$ with ε
  - This is AP if ε is small, but conditionally independent.

- Then, there exists no pure strategy eq where $(c.c)$ is played in the first period. There is no correlation between private signals. If d is observed, it is interpreted as an "error".

- Mixed Strategy: "Play C or D with some probability and play X if and only if C is played & c is observed".

  This strategy works because players' private information $((a_1.\omega_1)(a_2.\omega_2))$ is correlated.

- Playing D leads to (Y.Y), hence we need to use public randomization device in the second period to make D more attractive.
  In the second period, players forget everything from the first period & play (X.X) with some probability

(b) Efficient Payoff in PD

	C	D
C	1 . 1	$-\ell$. $1+g$
D	$1+g$. $-\ell$	0 . 0

Assume conditionally independent & almost perfect private monitoring.

σ_c : Grim Trigger Strategy

$$\begin{cases} (Cc).(Cc)\cdots(Cc) \to \text{play } C \\ \text{otherwise play } D \end{cases}$$

σ_D : Play D all the time.

Perfect monitoring

Consider the following strategic form game

	σ_c	σ_D
σ_c	1 . 1	$(1-\delta)(\ell)$. $(1-\delta)(1+g)$
σ_D	$(1-\delta)(1+g)$. $(1-\delta)(-\ell)$	0 . 0

- If. $\delta \geq \frac{g}{1+g}$. there are 3 Nash equilibrium

$$\begin{cases} (\sigma_C . \sigma_C) \\ (1-\lambda)\sigma_C + \lambda\sigma_D . (1-\lambda)\sigma_C + \lambda\sigma_D \\ (\sigma_D . \sigma_D) \end{cases}$$

$$(\lambda \to 0 \text{ as } \delta \to \frac{g}{1+g})$$

- These 3 equilibria are indeed (subgame-perfect) equilibria of the original repeated game. if $\varepsilon = 0$ (Perfect monitoring)

- If $\varepsilon > 0$, $(\sigma_C . \sigma_C)$ is not an equilibrium of repeated game by the same reason as in the two period example.

  However, a mixed trigger strategy equilibrium and $(\sigma_D . \sigma_D)$ is robust to private noise $(\varepsilon > 0)$.

- If $\delta > \frac{g}{1+g}$ is close enough to $\frac{g}{1+g}$, the mixed trigger strategy equilibrium puts most of the weight on σ_C, hence achieve almost efficient payoff (given that ε is small enough)

※ Verifying that the mixed trigger strategy is an Nash equilibrium.

We need to check the incentive on the equilibrium path.

Case 1 $(C.c), \cdots, (C.c)$

Since the monitoring is almost perfect, you are sure that the other player is playing σ_c, hence you also have an incentive to be cooperative.

In precise, I need to show that σ_c is best response after such history. Instead, I prove that playing D cannot be optimal. Note that if this is done for every history on the equilibrium path (including case 2 and case 3), this is also a valid way to verify that some strategy is a Nash equilibrium.

Case 2 $(C.c), \cdots, (C.c), (C.d)$

Intuition :

(i) If $(C.d)$ is observed in the first period, "d" is interpreted as a signal that σ_D was chosen in the beginning, hence a player has incentive to defect.

(ii) If $(C.d)$ is observed in the middle of the game, than it could be because

(a) Error: the other player has been observing $(C.c), \cdots, (C.c)$ and continue to play σ_C.

(b) The other player has observed $(C.c), \cdots, (C.d), (D.c)$ and continue to play σ_D.

(c) Other, σ_D is being played.

The probability of (a) and (b) is the same because both history contains only one error.

$\Rightarrow$ Posterior probability of σ_C is less than $1/2$.

$\Rightarrow$ If δ is not large, (close to $\frac{g}{1+g}$) there is incentive to start defecting.

Case 3 ⋯⋯, (D.ω)

Pretty sure that the other player observed d.

⇒ Incentive to continue defecting.

Remark

- This mixed trigger strategy Nash equilibrium can be modified to be a sequential equilibrium with the same outcome distribution.

- Note that the other player's continuation strategy is always σ_C or σ_D after any history. So, a player needs to keep track of posterior probability on a finite set $\{\sigma_C, \sigma_D\}$. This minimizes complexity of statistical inference.

※ For high δ, it is still possible to approximate efficiency (when ε is very small) by effectively reducing δ in one of the following two ways.

(1) Public Randomization Device

Same as 2 period example. With some probability, players restart the game, which makes defection more attractive.

(2) Decomposing the original game

Period 1 2, ⋯ $N+1$ $N+2$, ⋯ $2N+1$ $2N+2$, ⋯

Treat $(1, N+1, 2N+1, \cdots)$, $(2, N+2, 2N+2, \cdots)$ $(3, N+3, 2N+3, \cdots)$ ⋯ as N separate repeated games.

Take N large so that the effective discount factor δ^N for each game is about $\frac{g}{1+g}$.

Then, efficiency can be approximated for each repeated game.

IV Approach 2: Uncoordinated Punishment

Piccione + Ely + Valimaki

(a) 2 period example

Period 1

	C	D
C	1.1	-1.2
D	2.-1	0.0

Period 2

	X	Y
X	4.4	0 4
Y	4 0	0 0

- In period 2, any action is optimal whatever the other player does, because a player's payoff only depends on the other's action, not on one's own action.

  ⇒ No need of complex statistical inference

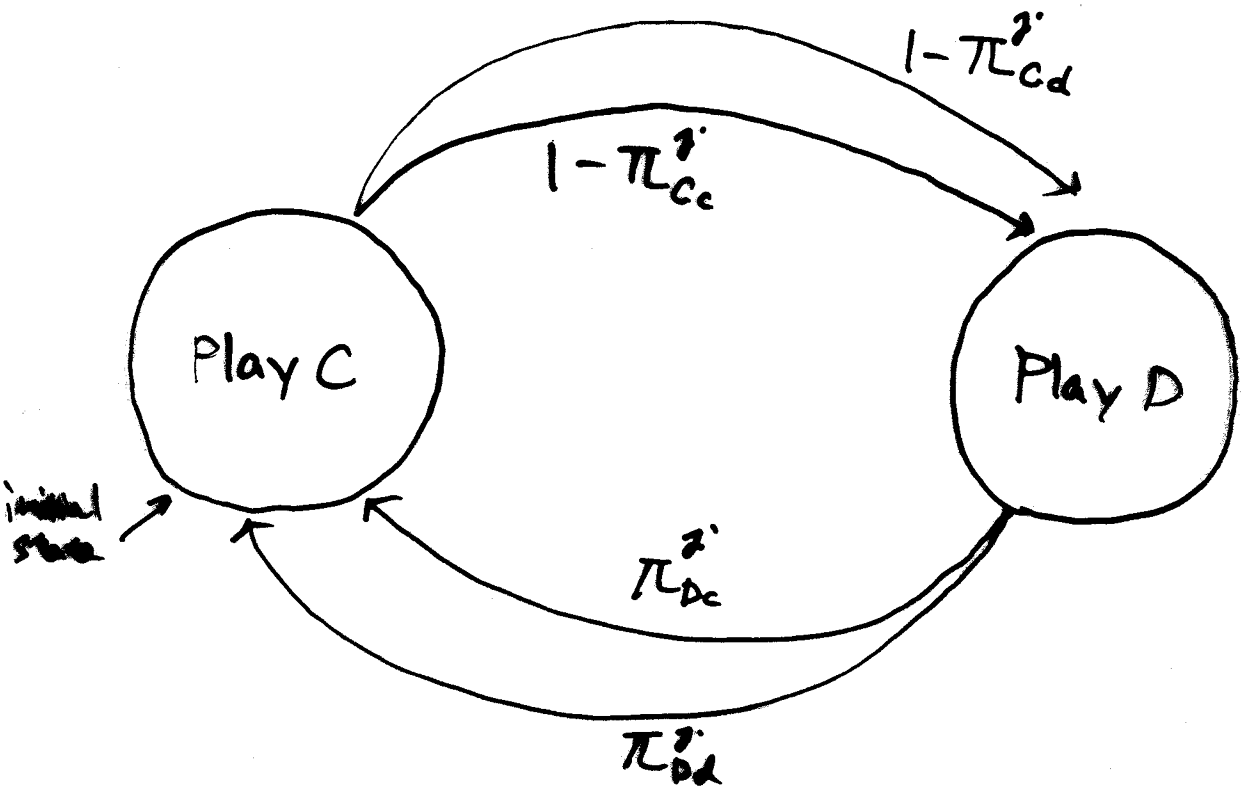
- So, " Play C. { Play X if c is observed / Punish by playing Y with some probability if d is observed " is an equilibrium.

⇒ A player punishes for herself $\boxed{\text{No coordinated punishment}}$

(b) Repeated PD with almost perfect monitoring.

<u>Assume perfect monitoring first.</u>

Consider the following strategy by player j.



- $\pi^j_{a\omega}$ is the probability to stay at or move to "Play C" when a is played and ω is observed.

$\{\pi_{aw}\}_{a,w}$ are chonsen to satisfy the following 4 equations

$$V_c^i = (1-\delta) + \delta\left[\pi^j_{Cc} V_c^i + (1-\pi_{Cc}) V_D^i\right]$$

$$V_c^i = (1-\delta)(1+g) + \delta\left[\pi^j_{Cd} V_c^i + (1-\pi^j_{Cd}) V_D^i\right]$$

$$V_D^i = -l(1-\delta) + \delta\left[\pi^j_{Dc} V_c^i + (1-\pi^j_{Dc}) V_D^i\right]$$

$$V_D^i = \delta\left[\pi^j_{Dd} V_c^i + (1-\pi^j_{Dd}) V_D^i\right]$$

Remark.

- V_c^i is player i's continuation payoff when player j is in state "Play C".

- Note that the other player's state completely determines one player's payoff. So, she is totally indifferent among all repeated game strategy. Any strategy is optimal. So is this machine itself

- You don't have to know which state the other player is in. Any strategy is optimal independent of the other player's state. It is precisely because of this property that this equilibrium is robust to introduction of private monitoring.

For almost perfect monitoring, there exists an equilibrium which is close to the equilibrium in the perfect monitoring case.

- π_{ce} can be chosen arbitrary close to 1 if monitoring is close to perfect. Then, this equilibrium is almost efficient.