Repeated Games II

Imperfect public signals

A limitation of the repeated games model studied so far is that we are assuming actions are observable.

In many interesting applications this is not the case.

An interesting and sufficiently simple case, is when actions are unobservable, but there are **imperfect public signals** correlated to the actions.

These signals can be used in a repeated game to sustain cooperation or more generally as inputs to the players' strategies.

Model

Let $A_1, \ldots, A_I$ be finite action sets.

Let Y be a finite set of public outcomes.

Let $\pi(y|a) = Pr(y|a)$.

Let $r_i(a_i, y)$ be i's payoff if she plays a_i and the public outcome is y.

Player i's expected payoff is:

$$g_i(a) = \sum_{y \in Y} \pi(y|a) \cdot r_i(a_i, y).$$

A mixed strategy is $\alpha_i \in \Delta(A_i)$. Payoffs are defined in the obvious way.

The *public information* at the start of period t:
$h^t = (y^0, \ldots, y^{t-1})$.

Player i's *private information* is $h_i^t = (a_i^0, \ldots, a_i^{t-1})$.

A strategy for i is a sequence of maps σ_i^t taking
$(h^t, h_i^t) \to \Delta(A_i)$

Definition 1 *A public strategy for player i is a sequence of maps $\sigma_i^t : h^t \to \Delta(A_i)$.*

We focus on public strategies because they are simple and lead to a nice structure for the game.

Player i's average discounted payoff for the game if he gets a sequence of payoffs $\{g_i^t\}$ is: $(1-\delta)E_\sigma \sum_{t=0}^{\infty} \delta^t g_i(\sigma(h^t))$.

Definition. *A profile* $(\sigma_1, \ldots, \sigma_I)$ *is a perfect public equilibrium if:*

- *σ_i is a public strategy for all i.*
- *For each date t and history h^t, the strategies are a Nash equilibrium from that point on.*

With imperfect monitoring, there are often no proper subgames.

A player may be uncertain as to which of many information nodes he is at, so SPE might have no bite.

However, since opponents don’t use private information in their strategies, all possible nodes have the same distribution over opponent play.

So there’s no need to distinguish.

Like the set of subgame perfect equilibrium payoffs in a repeated game model with perfect monitoring, the set of PPE payoffs is stationary.

A special case of Public Monitoring is when $Y = A$, and $\pi(y|a)$ puts probability 1 on a, then PPE coincides with SPE.

Note that a PPE is a perfect Bayesian equilibrium of the repeated game, but not all perfect Bayesian equilibria are PPEs.

In PPE, everyone uses a public strategy.

Given that opponent’s are using public strategies, it doesn’t help i to use a non-public strategy, since any private information he might have is not payoff-relevant (since preferences don’t depend on the private information).

However, if other players are using their private information, then a player might want to use his.

A motivating example: **Green and Porter (1984)**

Actions are interpreted as quantities $a_i = q_i \in [0, \overline{Q}]$.

Quantities are unobserved.

But they determine an observed marker price $p = P(q, \varepsilon)$.

p is a random variable: $\lambda(q) = \Pr(p \geq \widehat{p}|q)$

What equilibrium is reached in this market?

Green and Perter study an equilibrium in “trigger-price strategies”:

- Phase I. Start producing $\widehat{q}$. If $p \geq \widehat{p}$, stay in this phase. If not, go to Phase II.
- Phase II. If $p < \widehat{p}$, play a static equilibrium for T periods. Then return to Phase I

The value that can be achieved with these strategies in phase I is:

$$\widehat{v} = (1-\delta)g(\widehat{q}) + \delta[\lambda(\widehat{q}) + (1-\lambda(\widehat{q}))\delta^T]\widehat{v}$$

$$\Leftrightarrow \widehat{v} = \frac{(1-\delta)g(\widehat{q})}{1-\delta[\lambda(\widehat{q}) + (1-\lambda(\widehat{q}))\delta^T]}$$

Note that here we are normalizing the profit in the static Nash Equilibrium at 0.

An hypothetical designer of a collusive equilibrium would select $\widehat{q}$, $\widehat{p}$ and T to maximize $\widehat{v}$ under the constraint that the strategies are an equilibrium.

Obviously in Phase II, playing the static equilibrium for T periods is incentive compatible.

In Phase I, however, we need:

$$(1-\delta)g(q_i, \widehat{q}_{-i}) + \delta[\lambda(q_i, \widehat{q}_{-i}) + (1-\lambda(q_i, \widehat{q}_{-i})\delta^T)\widehat{v}]$$

$$\leq (1-\delta)g(\widehat{q}) + \delta[\lambda(\widehat{q}) + (1-\lambda(\widehat{q})\delta^T)\widehat{v}]$$

for all a_i.

Using the expression for $\widehat{v}$, we have:

$$(1-\delta)[g(q_i,\widehat{q}_{-i})-g(\widehat{q})] \leq \frac{\delta(1-\delta^T)[\lambda(\widehat{q})-\lambda(q_i,\widehat{q}_{-i})](1-\delta)g(\widehat{q})}{1-\delta[\lambda(\widehat{q})+(1-\lambda(\widehat{q}))\delta^T]}$$

for all a_i.

It can be shown that there are equilibria in which collusion is sustainable in the sense that the players can do better than in the static equilibrium.

Under some conditions, the payers can sustain collusion with a minimal deviation from the collusive quantity, by having a small punishment for long periods.

Dynamic programming and self generation

Abreu, Pierce and Stacchetti (1986, 1990) developed techniques to study repeated games with public monitoring that have more generally useful to study any type of repeated game.

The idea is to focus on promised utilities rather than on strategies.

For a given set of feasible utilities that can be promised in the future, we can then study what actions are taken in equilibrium (and so what utilities are credible).

Definition. *The pair* (α, v) *is **enforceable** with respect to* δ *and* $W \subset R^I$ *if there exists a function* $w : Y \to W$ *such that for all* i*:*

- $v_i = (1-\delta) g_i(\alpha) + \delta \sum_{y \in Y} \pi(y|a) \cdot w_i(y)$

- $\alpha_i \in \arg\max_{\alpha_i' \in \Delta(A_i)} \left[\begin{array}{c} (1-\delta) g_i(\alpha_i', \alpha_{-i}) \\ +\delta \sum_{y \in Y} \pi(y|\alpha_i', \alpha_{-i}) \cdot w_i(y) \end{array} \right].$

The first condition says that the target payoff v can be decomposed into today's payoff $g_i(\alpha)$ and the expected continuation payoff.

The second condition is an incentive compatibility constraint.

These conditions are similar to Bellman's equation for dynamic programming.

Definition. *Let* $B(\delta, W)$ *be the set of payoffs* v *such that for some* α*,* (α, v) *is enforced with respect to* δ *and* W*. Then* $B(\delta, W)$ *is the payoff set* ***generated*** *by* δ, W*.*

Definition $E(\delta)$ *is the set of PPE payoffs*.

Proposition $E(\delta) = B(\delta, E(\delta))$.

Proof. ($\supseteq$) Fix $v \in B(\delta, E(\delta))$. Pick $w : Y \to E(\delta)$ such that w enforces (α, v).

Now consider the following strategies: in period 0, play α. Then starting in period 1, play the perfect public equilibrium that gives payoffs $w(y_0)$.

This is a PPE, so $v \in E(\delta)$.

($\subseteq$) Fix $v \in E(\delta)$. There is some PPE that gives payoffs v.

Suppose in this PPE, play in period 0 is α, and continuation payoffs are $w(y_0) \in E(\delta)$, since continuation corresponds to PPE play.

The fact that no one wants to deviate means that (α, v) is enforced by $w : Y \to E(\delta)$, so $v \in B(\delta, E(\delta))$. ■

Abreu, Pearce and Stacchetti (1986, 1990) call the idea behind this result factorization.

The observation is that for any PPE, the payoffs can be factored (i.e. decomposed) into today's payoffs and continuation payoffs.

In a PPE, all the continuation payoffs have a recursive structure since they correspond to PPE profiles.

If it is possible to sustain average payoffs in W by promising different continuation payoffs in W, then W is self generating.

Definition. *W is self-generating if $W \subset B(\delta, W)$*.

Note that:

- $E(\delta)$ is self-generating.
- The set of static Nash equilibrium payoffs is also self-generating.

Proposition. *If W is self-generating, then $W \subset E(\delta)$.*

Proof. Fix $v \in W$. Then $v \in B(\delta, W)$, so there is some $w : Y \to W$ and some α such that (α, v) is enforced by w.

We construct an equilibrium that gives v.

In period 0, play α, and for an outcome y_0, set $v_1 = w(y_0)$.

Then $v_1 \in W \subset B(\delta, W)$, so again there is some α_1 and some $w_1 : Y \to W$ such that (α_1, v_1) is enforced by w_1.

Continue with this argument ad infinitum, to obtain recommended strategies.

After each public history, there are no profitable deviations, and by construction the payoff is v. $\blacksquare$.

Corollary. $E(\delta)$ *is the largest self-generating set.*

Let us now consider a couple of examples.

Examples

The Prisoners' dilemma with perfect monitoring

We noted that games with perfect monitoring are special examples of games with public information.

So we can use the techniques outlined above to study the classical prisoners' dilemma with perfect monitoring.

Consider this prisoners' dilemma game:

	Cooperate	Defect
Cooperate	1.1	$-1, 2$
Defect	$2, -1$	$0, 0$

With perfect monitoring $Y = \{(C,C),(C,D),(D,C),(D,D)\}$.

We now show that if $\delta \geq 1/2$, the set $W = \{(0,0),(1,1)\}$ is self-generating.

To this goal, we show that $(0,0) \in B(\delta, W)$, and $(1,1) \in B(\delta, W)$ for $\delta \geq 1/2$.

Consider $(0,0)$ first.

It is easy to see that the strategy profile (D,D), and payoff profile $(0,0)$ are enforced by any δ and the function $w(y) = (0,0)$ since

$$0 = (1-\delta)g_i(D,D) + \delta w_i(D,D)$$

and for all $a_i \in \{C,D\}$:

$$0 \geq (1-\delta)g_i(a_i,D) + \delta w_i(a_i,D)$$

Now consider $(1, 1)$.

We show that the strategy profile (C, C) and payoff profile $(1, 1)$ are enforced by $\delta \geq 1/2$ and W.

Let $w(C, C) = (1, 1)$ and $w(y) = (0, 0)$ for all $y \neq (C, C)$. Then

$$1 = (1 - \delta)g_i(C, C) + \delta w_i(C, C)$$

Moreover, for all $a_i \in \{C, D\}$, if $\delta \geq 1/2$:

$$1 \geq (1-\delta)g_i(a_i, C) + \delta wi(a_i, C).$$

So $W \subset B(\delta, W)$ for $\delta \geq 1/2$, meaning that W is self-generating. $\blacksquare$

Another example

Consider this game:

	Cooperate	Defect
Cooperate	$2, 2$	$-1, 3$
Defect	$3, -1$	$0, 0$

Again assume $Y = \{(C,C),(C,D),(D,C),(D,D)\}$.

We now intend to prove that if $\delta \geq 1/3$, then $W = \{v, \hat{v}\}$ is self-generating, where:

$$v = \left(\frac{3-\delta}{1+\delta}, \frac{3\delta-1}{1+\delta}\right)$$

$$\hat{v} = \left(\frac{3\delta-1}{1+\delta}, \frac{3-\delta}{1+\delta}\right)$$

Since it is a symmetric game and the continuation values in W are permutations, we just need to prove that v can be enforced with continuation in W.

Let the action profile α corresponding to v be (D, C) and the continuation payoffs be $w(D, C) = w(C, C) = \hat{v}$ and $w(D, D) = w(C, D) = v$.

If players follow α the payoffs are:

$$(1-\delta)(3,-1)+\delta\hat{v} = \begin{pmatrix} \frac{3(1-\delta^2+3\delta^2-\delta)}{1+\delta}, \\ -\frac{(1-\delta^2)+3\delta-\delta^2}{1+\delta} \end{pmatrix} = v$$

Clearly D maximizes a player's action, since the current action does not influence future payoffs.

If player 2 plays C as required by α, the payoff is $v_2 = \frac{3\delta-1}{1+\delta}$.

If 2 plays D, the payoff is 0 today and $v_2 = \frac{3\delta-1}{1+\delta}$.

So playing D is optimal if $v_2 \geq 0$, i.e. $\delta \geq 1/3$.

(Note $\delta \geq 1/3 \Rightarrow v_2 > \delta v_2$).

The folk theorem with imperfect public monitoring

Recall the a “Folk Theorem” aims at proving that all feasible, individually rational payoffs (i.e. V^*) are achievable in equilibrium.

A reasonable approximation is that any closed subset of V^* is achievable in equilibrium.

Can we prove this statement with imperfect public monitoring?

If the signal structure is that nothing is observed, then only static Nash equilibria are possible.

A reasonable assumption to make is that the signal structure is sufficiently rich that it is possible to *provide incentives* using expected payoffs.

Define $\boldsymbol{\pi}(a_i|\alpha_{-i})$ to be the row vector of probabilities on Y generated by a_i given α_{-i}, so it is a $|Y|$ dimensional vector.

Define $\Pi_i(\alpha_{-i})$, the $|A_i| \times |Y|$ dimensional matrix that stacks the $\boldsymbol{\pi}(a_i|\alpha_{-i})$ one on top of the other.

If we ignore feasibility constraints, we can implement a utility vector k if we can solve the system:

$$(1-\delta)G(a_i) + \delta\Pi_i(\alpha_{-i})w_i = k$$

where $G(a_i)$ is a column vector with generic element $g_i(a_i)$ for all $a_i \in A_i$.

The above system is solvable in w_i if $\Pi_i(\alpha_{-i})$ is invertible, i.e. it is full rank.

Definition. *The individual full rank condition is satisfied by a profile α if for each player i $\Pi_i(\alpha_{-i})$ is invertible, i.e. the vectors $\pi(a_i|\alpha_{-i})$ are linearly independent.*

The individual full rank condition, however is not sufficient for the Folk Theorem.

Consider this example (due to Radner, Myerson and Maskin (1986):

- Two players can work or shirk
- Work has a personal cost of 1; shirk costs 0
- Output has two levels, high or low
- If output is high then both players receive 4, if low, they receive 0.

- The probabilities for high output are:

$$\pi_H(H,H) = 9/16$$

$$\pi_H(H,L) = \pi_H(L,H) = 3/8$$

$$\pi_H(L,L) = 1/4$$

The individual full rank condition is satisfied at $\alpha = (H,H)$:

$$\Pi_i(\alpha_{-i}) = \begin{bmatrix} \pi_H(H,H) & 1-\pi_H(H,H) \\ \pi_H(L,H) & 1-\pi_H(L,H) \end{bmatrix}$$

$$= \begin{bmatrix} 9/16 & 7/16 \\ 3/8 & 5/8 \end{bmatrix}$$

which is full rank.

Despite the fact that the individual full rank condition is satisfied, the "folk theorem" does not hold.

To see this let v^* be the highest payoff in any symmetric equilibrium.

If the folk theorem is true, we should be able to have a payoff close to $4(\frac{9}{16}) - 1 = 5/4 > 1$

This because if players choose (H, H), then the expected payoff is $(5/4, 5/4)$.

We show that these payoff cannot be approximated.

We show they cannot be approximated in pure startegies for simplicity (the result is generalizable to mixed strategies)

So if the theorem hold $v^* > 1$ (note $5/4 > 1$)

Since the equilibrium is stationary, in equilibrium the players must choose (H,H) at least with positive probability, so:

$$v^* = (1-\delta)\left[\frac{9}{16}\cdot(4+\delta v_g)+\frac{7}{16}\cdot(0+\delta v_b)-1\right]$$

$$\geq (1-\delta)\left[\frac{3}{8}\cdot(4+\delta v_g)+\frac{5}{8}\cdot(0+\delta v_b)-1\right]$$

Implying: $v_g - v_b \geq \frac{4}{3}\frac{1-\delta}{\delta}$

But by definition $v_g \leq v^*$, so:

$$v^* \leq (1-\delta)\frac{5}{4} + \delta\left[\frac{9}{16} \cdot v^* + \frac{7}{16} \cdot \left(v^* - \frac{4}{3}\frac{1-\delta}{\delta}\right)\right]$$

$$\Leftrightarrow v^* \leq \frac{3}{2} < 1 \Rightarrow\Leftarrow$$

From this example, we learn that we need an additional assumption.

Define the matrix $\Pi_{i,j}(\alpha)$ to be the matrix formed by stacking the matrix $\Pi_i(\alpha_{-i})$ on top of the matrix $\Pi_j(\alpha_{-j})$.

It is a matrix with $|A_i| + |A_j|$ rows and $|Y|$ columns.

Definition. *The pairwise full-rank condition is satisfied at action* α *for players* i *and* j *if* $\Pi_{i,j}(\alpha)$ *has maximal rank (i.e. full column rank).*

Note that $\Pi_{i,j}(\alpha)$ cannot have full row rank, so the $|A_i| + |A_j|$ vectors admit at least one linear dependency.

To see this note that:

$$\boldsymbol{\pi}(\alpha) = \sum_{a_i \in A_i} \alpha_i(a_i)\boldsymbol{\pi}(a_i, \alpha_{-i})$$

$$= \sum_{a_j \in A_j} \alpha_j(a_j)\boldsymbol{\pi}(a_j, \alpha_{-j})$$

So:

$$\boldsymbol{\pi}(a_1, \alpha_{-i}) = \sum_{a_j \in A_j} \frac{\alpha_j(a_j)}{\alpha_1(a_1)}\boldsymbol{\pi}(a_j, \alpha_{-j}) - \sum_{a_i \in A_i \setminus a_1} \frac{\alpha_i(a_i)}{\alpha_1(a_1)}\boldsymbol{\pi}(a_i, \alpha_{-i})$$

We have:

Proposition. *Suppose $dim V = I$, and the Individual Full-Rank and Pairwise Full-Rank conditions hold. Then for any closed set $W \subset int(V^*)$, there exists some $\delta^* < 1$ such that for all $\delta \geq \delta^*$, $W \subset E(\delta)$.*

A couple of limitations of this result.

A necessary condition to satisfy Pairwise Full-rank is that $|A_i| + |A_j| - 1 \leq |Y|$, this may be demanding.

Indeed it is not satisfied in the moral hazard example, in which we have 2 signals but $|A_i| + |A_j| - 1 = 3$.

The signal structure, moreover needs to be rich enough.

For example, even if we have more than 2 signals it fails at symmetric profiles (say (H,H)).

We have:

$$\Pi_{i,j}(\alpha) = \begin{bmatrix} \pi_H(H,H) \\ \pi_H(L,H) \\ \pi_H(H,H) \\ \pi_H(H,L) \end{bmatrix}$$

where $\pi_H(H,H)$ has $|Y|$ dimensions. This matrix has rank $2 < |A_i| + |A_j| - 1 = 3$, since $\pi_H(L,H) = \pi_H(H,L)$.