Repeated Games II: Imperfect Public Monitoring

Jonathan Levin

May 2006

We now take up the problem of repeated games where players' actions may not be directly observable. This is a rich class of problems, with many economic applications. Abreu, Pearce and Stacchetti (1990) and Fudenberg, Levine and Maskin (1994) have developed a beautiful and powerful set of techniques for these games.

1 A Few Examples

1. Cournot competition with noisy demand (Green and Porter, 1984). Firms set outputs $q_{1t}, ...q_{It}$, chosen privately. Demand conditions then determine $p_t = P(q_{1t}, ..., q_{It}, \varepsilon)$, which is observed publicly.

2. "Reputation" for quality. A single firm sets price p_t and chooses an effort e_t at some cost $c(e_t)$. Product quality is "High" with probability $p(e_t) \in (0, 1)$, where p is increasing in e_t. Consumers are willing to pay more for high quality.

3. Noisy prisoners' dilemma. Players choose from $\{C, D\}$. But instead of these actions being observed, some noisy signal of these actions is observed instead (see below).

4. Team Production. Players choose efforts $e_1 \in \{e_L, e_H\}$, as part of a joint project that succeeds with probability $p(e_1 + e_2)$. Only the joint outcome is observed publicly.

5. Self-enforced agency contracts (Levin, 2003). Each period, the agent privately observes a cost parameter θ_t, and produces output y_t at cost $c(\theta_t, y_t)$. The output, but not the cost, is observed. Alternatively, the agent chooses an effort e_t, and output is stochastic $y_t \sim f(\cdot|e)$.

6. Consumption smoothing and insurance (Green, 1987). There are a continuum of consumers, who each period get income shocks z_{it}. They then report their incomes and make transfers among themselves. Transfers must be balanced.

2 The General Model

- Let $A_1, ..., A_I$ 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:

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

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

Example 1 In the Cournot game, a_i is quantity, y is price (here, one might want to have A, Y continuous rather than finite).

Example 2 In the PD game, a_i is intended action, y is actual actions.

Example 3 In the agency problem, a_i is a *vector* that specifies the agent's output y for each cost realization.

- The public information at the start of period $t : h^t = (y^0, ..., y^{t-1})$.
- Player i's private information is $h_i^t = (a_i^0, ..., a_i^{t-1})$.
- 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 σ_{it} : $h^t \to \Delta(A_i)$.*

We focus on public strategies because they are simple and lead to a nice structure for the game. More on this later, however.

- Player i's average discounted payoff for the game if he gets a sequence of payoffs $\{g_i^t\}$ is:

$$(1-\delta)\sum_{t=0}^{\infty}\delta^t g_i^t$$

Definition 2 *A profile* $(\sigma_1, .., \sigma_I)$ *is a* ***perfect public equilibrium*** *if*

(i) σ_i *is a public strategy for all* i.

(ii) *For each date* t *and history* h^t*, the strategies are a Nash equilibrium from that point on.*

A crucial point about PPE is that after any history h^t, a PPE induces a PPE in the remaining game. However, note that like the set of subgame perfect equilibrium payoffs in a repeated game model with perfect monitoring, the set of PPE payoffs is *stationary* — i.e. it's the same starting from any period t. The first point is perhaps a little subtle. With imperfect monitoring, there are often no proper subgames (i.e. 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 base their strategies on private information, all possible nodes have the same distribution over opponent play, so there's no need to distinguish. The perfection condition *(ii)* is well-defined because the public history is commonly known.

Note that if the game *happens* to have perfect monitoring, so $Y = A$, and $\pi(y|a)$ puts probability 1 on a, then PPE *coincides* with SPE. More generally, a PPE is a perfect bayesian equilibrium of the repeated game, but not all perfect bayesian equilibria are PPEs. What do I mean by this? 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 $j \neq i$ are using their private information, i might want to use his. This non-public strategy case is not nearly as well-understood.

Example 4 Green and Porter (1984) suggest the following type of "trigger strategies" for the noisy Cournot model:

1. Play $q_1, ..., q_I$. If $p_t < \underline{p}$, go to phase 2.
2. Play $q_1^c, ..., q_I^c$ (cournot) for T periods. Then return to phase I.

GP verify that if the players are sufficiently patient, there is an equilibrium of this form where $q_1, ..., q_I$ are less than the static Cournot levels (e.g. equal to q^M/I where q^M is the monopoly quantity) and $\underline{p}$ and T are chosen appropriately. This equilibrium is a PPE. Strategies are "public" and play is Nash from every time forward.

Note that a lower trigger price means less chance of punishment and more incentive to deviate. A longer punishment periods mean less incentive to deviate, but less efficiency. GP's main result is that firms can't achieve the first-best monopoly profits as there will be "price wars" in equilibrium.

3 Self-Generation

We now develop a set of powerful techniques for characterizing perfect public equilibria. In contrast to the Green-Porter analysis, we will think not in terms of *strategies*, but in terms of *payoffs*. The idea is that to "enforce" certain actions at time t, we will attach continuation payoffs from time $t+1$ on to each time t outcome. You can think of this as analogous to a principal–agent problem, where to motivate the agent, the principal promises certain rewards or punishments. The subtlety here is that the promised rewards and punishments must themselves correspond to payoffs in a PPE of the continuation game (rather than being monetary payoffs specified in a court-enforced contract).

Definition 3 *The pair (α, v) is* ***enforceable*** *with respect to δ and $W \subset \mathbb{R}^I$ if there exists a function $w : Y \to W$ such that for all i,*

(i) $v_i = (1-\delta)g_i(\alpha) + \delta \sum_y \pi(y|a)w_i(y)$

(ii) $\alpha_i \in \arg\max_{\alpha_i' \in \Delta(A_i)} (1-\delta)g_i(\alpha_i', \alpha_{-i}) + \delta \sum_y \pi(y|\alpha_i', \alpha_{-i})w_i(y)$

Condition *(i)* says that the target payoff v can be *decomposed* into today's payoff $g_i(\alpha)$ and the expected continuation payoff (macroeconomists call this the "promised utility"). Condition *(ii)* is essentially an incentive compatibility constraint. These conditions ought to remind you of Bellman's equation for dynamic programming.

Definition 4 *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 5 $E(\delta)$ *is the set of PPE payoffs.*

Proposition 1 $E(\delta) = B(\delta, E(\delta))$.

Proof. ($\supseteq$) Fix $v \in B(\delta, E(\delta))$. Pick α, $w : Y \to E(\delta)$ 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 $v \in B(\delta, E(\delta))$. *Q.E.D.*

Abreu, Pearce and Stacchetti (1986, 1990) call this factorization. The idea is that for any PPE, the payoffs can be decomposed or factored into today's payoffs and continuation payoffs. In a PPE, all the continuation payoffs have to themselves correspond to PPE profiles. So those can be decomposed, and so on. So they have a recursive structure.

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

The interpretation is that it is possible to sustain average payoffs in W by promising different continuation payoffs in W. Note that $E(\delta)$ is self-generating. The set of static Nash equilibrium payoffs is also self-generating.

Proposition 2 *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 such that there are no profitable deviations, and which by construction give payoff v from time 0. *Q.E.D.*

Corollary 1 $E(\delta)$ *is the largest self-generating set.*

We'll now go on to discuss some applications of these ideas.

4 Examples of Self-Generation

Let's try out self-generation in two variants of the prisoners' dilemma.

4.1 Prisoners' Dilemma

Consider the prisoner's dilemma with *perfect monitoring.*

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

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

Claim If $\delta \geq 1/2$, the set $W = \{(0,0),(1,1)\}$ is self-generating.

Proof. We want to 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

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

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

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

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

Exercise 1 *Try showing that if* $\delta = 1/2$, *then* $(3/2, 0)$ *is also in* $B(\delta, W)$.

4.2 Noisy Prisoners' Dilemma

There are two players $i = 1, 2$ and $A_i = \{C, D\}$. The observed outcomes are $Y = \{G, B\}$ (good and bad) where

$$\Pr(G \mid a) = \begin{cases} p & \text{if } a = (C,C) \\ q & \text{if } a = (C,D),(D,C) \\ r & \text{if } a = (D,D) \end{cases},$$

with $p > q > r$. We assume that $p - q > q - r$. Payoffs are given by

$$r_i(a_i, y) = \begin{cases} 1 + \frac{2-2p}{p-q} & \text{if } (C,G) \\ 1 - \frac{2p}{p-q} & \text{if } (C,B) \\ \frac{2-2r}{q-r} & \text{if } (D,G) \\ \frac{-2r}{q-r} & \text{if } (D,B) \end{cases},$$

which means that *expected payoffs* $g_i(a)$ are given by the standard prisoners' dilemma matrix above.

Claim If $\frac{1}{(p-r)+(q-r)} \geq \delta \geq \frac{1}{(p-q)+(p-r)}$, the set $W = \{\frac{\delta r}{1-\delta(p-r)}, \frac{1-\delta+\delta r}{1-\delta(p-r)}\}$ is self-generating.

Proof. Recall that to enforce a payoff v, we need a profile a and a map $w : Y \to W$ from outcomes to continuation payoffs such that:

$$v = (1-\delta) g_i(a) + \delta \mathbb{E}\left[w_i(y) \mid a_i', a_{-i}\right]$$

and for all $a_i' \neq a_i$

$$v \geq (1-\delta) g_i(a_i', a_{-i}) + \delta \mathbb{E}\left[w_i(y) \mid a_i', a_{-i}\right].$$

Now, let $v = \frac{\delta r}{1-\delta(p-r)}$, and $v' = \frac{1-\delta+\delta r}{1-\delta(p-r)}$ (the bad and good continuation payoffs). To enforce v, use the profile (D, D) and

$$w(y) = \begin{cases} v' & \text{if } y = G \\ v & \text{if } y = B \end{cases}$$

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 so long as:

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

that is, if $\delta \geq 1/(p+q-2r)$.

To enforce v', use the profile (C, C) and the *same* $w : Y \rightarrow W$. We need to check

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

The first condition is again just algebra, while the second holds so long as:

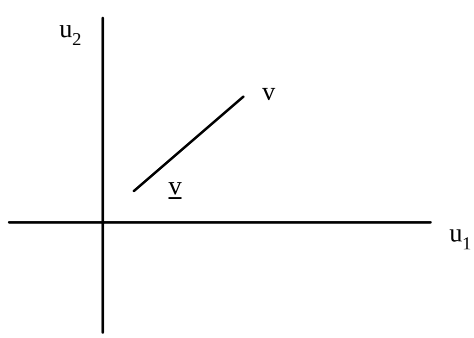
$$\delta(p-q)(v'-v) \geq 1-\delta$$

which requires that $\delta \geq 1/(2p-q-r)$. *Q.E.D.*

4.3 Strongly Symmetric PPE

The examples above are illustrative but they don't characterize the set of PPE. In general models, this tends to be quite hard. A useful simplification in symmetric games (introduced by Abreu, Pearce and Stacchetti, 1986) is to focus on PPE that are *strongly symmetric* in the sense that each player uses the same strategy after *every* history. In the two examples above, the strategies were strongly symmetric.

More generally, if we allow for public randomization, the set of strongly symmetric PPE payoffs will be an interval $[\underline{v}, \overline{v}]$, where $\underline{v}$ is the lowest and $\overline{v}$ the highest strongly symmetric PPE payoffs. Therefore, solving for the equilibrium set boils down to finding the best and worst equilibrium payoffs.



Set of Symmetric PPE

To actually characterize the highest and lowest payoffs, we need to solve a fixed point problem. We can think about fixing future symmetric payoffs to lie on some interval and finding maximal and minimal present payoffs subject to incentive compatibility, promise-keeping and the constraint that continuation payoffs are chosen from the allowable interval. If we start with a very large interval, we will find a somewhat smaller interval; continuing the process we will eventually converge to the equilibrium payoff set. Alternatively, we can jump right to the fixed point by solving the following problem:

$$\begin{aligned}
\overline{v} &= \max_{\overline{a},\underline{a},\overline{v},\underline{v},w:Y\to\mathbb{R}} (1-\delta)g(\overline{a}) + \delta\sum_y w(y)\pi(y|\overline{a}) \\
\text{s.t. } \overline{v} &= (1-\delta)g(\overline{a}) + \delta\sum_y w(y)\pi(y|\overline{a}) \\
\underline{v} &= (1-\delta)g(\underline{a}) + \delta\sum_y w(y)\pi(y|\underline{a}) \\
\overline{v} &\geq (1-\delta)g(a,\overline{a}) + \delta\sum_y w(y)\pi(y|a,\overline{a}) \text{ for all } a\in A \\
\underline{v} &\geq (1-\delta)g(a,\underline{a}) + \delta\sum_y w(y)\pi(y|a,\underline{a}) \text{ for all } a\in A \\
\overline{v} &\geq w(y) \geq \underline{v} \text{ for all } y\in Y
\end{aligned}$$

Note that here we find maximixal and minimal payoffs in one step using the fact that a lower minimum will automatically allow a higher minimum and vice versa.

While this is simpler than solving for PPE in general, it's still pretty complicated. Abreu, Pearce and Stacchetti (1986) characterize strongly symmetric equilibria in the Green-Porter oligopoly game where players chooses quantities and price is a noisy function of the aggregate quantity. Athey, Bagwell and Sanchirico (2004) study strongly symmetric equilibria in a repeated Bertrand pricing game where firms have private cost information.

5 The Folk Theorem

So far, we've argued that strongly symmetric strategies lead to inefficient outcomes (because of equilibrium "price wars"). Nevertheless, Fudenberg, Levine and Maskin (1994) show that this inefficiency arises because GP '84 and APS '86 limit the space of strategies, and go on to prove a version of the Folk Theorem.

Fudenberg, Levine and Maskin's result requires two "observability" or "identification" conditions.

(I1) For all i, and a_{-i}, the $|A_i|$ vectors $\pi(\cdot|a_i, a_{-i})$ are linearly independent.

(I2) For all i, j, there is some profile α such that the $|A_i| + |A_j|$ vectors $\pi(\cdot|a_i, \alpha_{-i})$ and $\pi(\cdot|a_j|\alpha_{-j})$ admit only one linear dependency.

The first condition requires that i's actions can be statistically identified — that is, they do not induce the same probability distribution on outcomes. Note that for (I1) to hold, it *must* be the case that $|Y| \geq |A_i|$ — which is arguably quite a strong requirement. The second condition says that if everyone is playing α, then not only can i's actions be distinguished, and j's actions being distinguished, but i's actions can be distinguished from j's actions. If you're 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. The second condition 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, $k \neq i, j$.

Let V^* be the set of feasible and strictly individually rational payoff vectors.

Proposition 3 *Suppose* $\dim V = I$*, and* $(I1), (I2)$ *hold. Then for any closed set* $W \subset int(V^*)$*, there exists some* $\underline{\delta} < 1$ *such that for all* $\delta \geq \underline{\delta}$*,* $W \subset E(\delta)$*.*

Proof. I'll sketch the idea in class; you'll have to read the paper for details! *Q.E.D.*

The folk theorem applies to payoff vectors in the interior of V^*. Generally you can't get *exact* efficiency with imperfect monitoring. The argument is simple and illustrative. Suppose that $\pi(\cdot|a)$ has a support that is independent of a (as in APS). And suppose that v is extremal but not a static Nash equilibrium payoff. Because v is extremal, the only sequence of payoffs that gives average value v must have payoffs v in *every* period. So if a PPE gives v, the first period strategies must specify a profile a with $g(a) = v$, and for any outcome y, the continuation payoffs must be $w(y) = v$. But then continuation payoffs are independent of today's outcome, so unless a happens to be a static Nash equilibrium (which it isn't by assumption), someone will want to deviate.

References

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

[2] Abreu, Dilip, David Pearce and Ennio Stacchetti (1990), "Toward of Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58, 1041–1063.

[3] Athey, Susan, Kyle Bagwell and Chris Sanchirico (2004), "Collusion and Price Rigidity," *Review of Economic Studies*.

[4] Fudenberg, Drew, David Levine and Eric Maskin (1994), "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62, 997–1039.

[5] Green, Edward, (1987), "Lending and Smoothing of Uninsurable Income," in *Contractual Arrangements for Intertemporal Trade,* ed. by Edward Prescott and Neil Wallace, Minneapolis: U of Minnesota Press.

[6] Green, Edward and Robert Porter (1984), "Noncooperative Collusion under Imperfect Price Information," *Econometrica*, 52, 87–100.

[7] Levin, Jonathan (2003), "Relational Incentive Contracts," *American Economic Review.*