Repeated Games

Debraj Ray, October 2006

1. Preliminaries

A *repeated game with common discount factor* is characterized by the following additional constraints on the infinite extensive form introduced earlier: $A_i(t) = A_i$ for all t, and $C_i(t, h(t)) = A_i$ for every t-history h(t). Moreover, for every i there is a one-period utility indicator $f_i : X_j A_j \mapsto \mathbb{R}$ and a (common) discount factor $\delta \in (0, 1)$ such that

$$F_i(\mathbf{a}) = (1 - \delta) \sum_{t=0}^{\infty} \delta^t f_i(a(t)).$$

for every path **a**.

The corresponding *one-shot* or *stage game* is denoted by $G = (\{A_i\}_{i=1}^n, \{f_i\}_{i=1}^n)$. The usual interpretation is that A_i is a set of pure actions.¹

The set of feasible payoff vectors of the stage game G is given by the set

$$F \equiv \{\mathbf{p} \in \mathbb{R}^n | f(a) = p \text{ for some } a \in A\}.$$

Let F^* be the convex hull of the set of feasible payoffs. It should be clear that any normalized payoff in the repeated game must lie in this set.

In what follows, assume

[G.1] Each A_i is compact metric, and $f_i : A \to \mathbb{R}$ is continuous for all *i*.

[G.1],together with $\delta \in (0, 1)$, implies that [C.1] and [F.1] are satisfied.

We also assume

[G.2] The stage game *G* has a one-shot Nash equilibrium in pure strategies.

Strategies and (perfect) equilibrium have already been defined for the infinite extensive form, and their restrictions to this special case form the corresponding definition for repeated games.

2. Example: The Repeated Prisoner's Dilemma

	C_2	D_2
C_1	2,2	0,3
D_1	3,0	1,1

Repeat this.

¹There is formally no loss in interpreting A_i to be a set of mixed strategies. But there are conceptual problems with this, as it requires that the strategies themselves (and not their realizations) be observed by all players.

OBSERVATION 1. The only SGPE in a finitely repeated PD is "defect forever".

How general is this observation? Well, it works for any game in which the stage Nash equilibrium is unique, such as the Cournot oligopoly.

What if there is more than one equilibrium in the stage game?

Example. Variant of the PD

	C_2	D_2	E_2
C_1	5,5	0,0	0,6
D_1	0,0	4,4	0,1
E ₃	6,0	1,0	1,1

Let T = 1; i.e., game is played at periods 0 and 1. Consider the strategy for player *i*:

Start with C_i .

Next, if both players have played *C* in period 0, play D_i in period 1.

Otherwise, play E_i in period 1.

Claim: if discount factor is no less than 1/3, these strategies form a SGPE.

The subgame strategies generate Nash equilibria. So all we have to do is check deviations in period 0. Given the opponent's strategy as described, if player *i* plays according to her prescribed strategy she gets $5 + \delta 4$. If she misbehaves, the best she can get is $6 + \delta$. The former is better than the latter if $\delta \ge 1/3$.

Calculus of cooperation and punishment, constrained by the requirement that the punishment has to be credible.

Finitely repeated games can therefore take us in interesting directions provided we are willing to rely on multiple stage equilibria.

Example. Prisoner's Dilemma again, this time infinitely repeated.

	C_2	D_2
C_1	2,2	0,3
D_1	3,0	1,1

Various strategy profiles possible.

Cooperate forever: start with C_i , and set $\sigma_i(h(t)) = C_i$ for all conceivable *t*-histories.

Defect forever: start with D_i , and set $\sigma_i(h(t)) = D_i$ for all conceivable *t*-histories.

Grim trigger: Start with C_i , thereafter $\sigma_i(h(t)) = C_i$ if and only if every *opponent* entry in h(t) is a " C_j ".

Modified grim trigger: Start with C_i , thereafter $\sigma_i(h(t)) = C_i$ if and only if *every* entry (mine or the opponent's) in h(t) is a "C".

2

Tit for tat: Start with C_i , thereafter $\sigma_i(h(t)) = x_i$, where x has the same letter as the opponent's last action under h(t).

Consider the modified grim trigger strategy: Start with C_i , thereafter $\sigma_i(h(t)) = C_i$ if and only if every entry in h(t) is a "C".

Claim: if discount factor is no less than 1/2, the modified grim trigger profile is a SGPE.

Two kinds of histories to consider:

1. Histories in which nothing but *C* has been played in the past (this includes the starting point).

2. Everything else.

In the second case, we induce the infinite repetition of the one-shot equilibrium which is subgame perfect.

So all that 's left is the first case. In that case, by complying you get $2/(1 - \delta)$. By disobeying, the best you can get is $3 + \delta/(1 - \delta)$. If $\delta \ge 1/2$, the former is no smaller than the latter.

Examine perfection of the grim trigger, the modified trigger with limited punishments, tit-fortat.

3. Equilibrium Payoffs

The set of equilibrium payoffs of a repeated game is often a simpler object to deal with than the strategies themselves. There are also interesting properties of this set that are related to the functional equation of dynamic programming.

Recall that F^* be the convex hull of the set of feasible payoffs. It should be clear that any normalized payoff in the repeated game must lie in this set. Define the class of sets \mathcal{F} by collecting all nonempty subsets of F^* . Formally,

$$\mathcal{F} \equiv \{ E | E \neq \emptyset \text{ and } E \subseteq F^* \}.$$

For any action profile *a* and any *i*, define $d_i(a) = \max f_i(a'_i, a_{-i})$ over all $a'_i \in A_i$. Pick any $E \subseteq \mathbb{R}^n$ and $\mathbf{p} \in \mathbb{R}^n$. Say that \mathbf{p} is *supported by E* if there exist n + 1 vectors (not necessarily distinct) $(\mathbf{p}', \mathbf{p}^1, \dots, \mathbf{p}^n) \in E$ and an action vector $a \in A$ such that for all *i*

(1)
$$p_i = (1 - \delta)f(a) + \delta p'_i,$$

and

(2)
$$p_i \ge (1-\delta)d_i(a) + \delta p_i^i \text{ for all } a_i' \in A_i.$$

We may think of *a* as the *supporting action* of *p*, of **p**' as the *supporting continuation payoff* of *p*, and so on, and define the entire collection $(a, \mathbf{p}', \mathbf{p}^1, \dots, \mathbf{p}^n)$ as the *supporter* of **p** ("in *E*").

Now define a map $\phi : \mathcal{F} \mapsto \mathcal{F}$ by

 $\phi(E) \equiv \{ \mathbf{p} \in \mathbb{R}^n | \mathbf{p} \text{ is supported by } E \}.$

It is easy to check that ϕ is well-defined and indeed maps into \mathcal{F} . [Use condition [G.2] as well as (1).]

We study several properties of this map. First, a set $W \in \mathcal{F}$ is *self-generating* (Abreu, Pearce and Stacchetti (1990)) if $W \subseteq \phi(W)$.

THEOREM 1. If W is self-generating, then $W \subseteq V$, where V is the set of all normalized perfect equilibrium payoffs.

Proof. Pick $\mathbf{p} \in W$. We exhibit an equilibrium σ that yields payoff \mathbf{p} . Proceed by induction on the length of *t*-histories. For h(0), pick $\mathbf{p}(h(0)) = \mathbf{p}$ and a(h(0)) to be any supporting action for \mathbf{p} drawn from a supporter in W.

Recursively, suppose that we have defined an action a(h(s)) as well as an equilibrium payoff vector $\mathbf{p}(h(s))$ for every *s*-history h(s) and all $0 \le s \le t$. Now consider a t + 1 history h_{t+1} , which we can write in the obvious way as $h_{t+1} = (h(t), a(t))$ for some *t*-history h(t) and some action vector $a(t) \in A$.

Let *a* be a supporting action for $\mathbf{p}(h(t))$. If a = a(t), or if *a* differs from a(t) in at least two components, define $\mathbf{p}(h(t+1))$ to be \mathbf{p}' , where \mathbf{p}' is a supporting continuation payoff for $\mathbf{p}(h_t)$ in *W*, and define a(h(t+1)) to be the corresponding action that supports $\mathbf{p}(h(t+1))$. If $a \neq a(t)$ in *precisely* one component *i*, then define $\mathbf{p}(h(t+1))$ to be the *i*th supporting punishment \mathbf{p}^i for $\mathbf{p}(h(t))$ in *W*, and a(h(t+1)) to be a supporting action for $\mathbf{p}(h(t+1))$.

Having completed this recursion, define a strategy profile by $\sigma(t)[h(t)] = a(h(t))$ for every *t* and every *t*-history. Use the one-shot deviation principle to show that σ is a subgame perfect equilibrium.

EXERCISE. Establish the following properties of the mapping ϕ .

[1] ϕ is isotone in the sense that if $E \subseteq E'$, then $\phi(E) \subseteq \phi(E')$.

[2] Under assumptions (G.1) and (G.2), ϕ maps compact sets to compact sets: that is, if *E* is a compact subset of *F*^{*}, then $\phi(E)$ is compact as well.

Our next theorem is an old result: the set of perfect equilibrium payoffs is compact. But the proof is new.

THEOREM 2. Under assumptions [G.1] and [G.2], the set of perfect equilibrium payoffs V is compact.

Proof. We begin by showing that $V \subseteq \phi(\operatorname{cl} V)$, where $\operatorname{cl} V$ denotes the closure of V. To this end, take any perfect equilibrium payoff **p**. Then there is a SGPE supporting **p**. Consider the action profile *a* prescribed in the first date of this equilibrium, as well as the prescribed paths and payoff vectors following every 1-history. These may be partitioned in the following way: (i) the payoff **p**' assigned if *a* is played as prescribed; (ii) for each *i*, a mapping that assigns a payoff vector $\mathbf{p}^i(a'_i)$ following each choice of a'_i at time period zero, *assuming that others are sticking to the prescription of a*, and (iii) payoff vectors that follow upon multiple simultaneous deviations of players from *a*.

Ignore (iii) in what follows.

Consider (ii). Note that $\mathbf{p}^i(a'_i) \in V$ for all a'_i , and that by the notion of SGPE,

$$p_i \ge (1-\delta)f_i(a'_i, a_{-i}) + \delta p_i^i(a'_i)$$

$$p_i \ge (1-\delta)f_i(a'_i, a_{-i}) + \delta p_i^i$$

for every action a'_i . Consequently,

$$p_i \ge (1-\delta)d_i(a) + \delta p_i^i.$$

Do this for every *i*, and combine with (i) to conclude that $(a, \mathbf{p}', \mathbf{p}^1, ..., \mathbf{p}^n)$ is a supporter of **p**. This proves that $\mathbf{p} \in \phi(\operatorname{cl} V)$, so that $V \subseteq \phi(\operatorname{cl} V)$.

Next, observe that because *V* is bounded, cl *V* is compact. Consequently, by (2) of the exercise above, $\phi(\text{cl } V)$ is compact as well. It follows from this and the claim of the previous paragraph that cl $V \subseteq \phi(\text{cl } V)$. But then by Theorem 1, cl $V \subseteq V$. This means that *V* is closed. Since *V* is bounded, *V* is compact.

These results permit the following characterization of *V* (note the analogy with the functional equation of dynamic programming).

THEOREM 3. Under assumptions [G.1] and [G.2], V is the largest fixed point of ϕ .

Proof. First we show that *V* is indeed a fixed point of ϕ . Since $V \subseteq \phi(\operatorname{cl} V)$ (see proof of Theorem 2) and since *V* is compact, it follows that $V \subseteq \phi(V)$. Let $W \equiv \phi(V)$, then $V \subseteq W$. By the exercise (1) above, it follows that $W = \phi(V) \subseteq \phi(W)$. Therefore *W* is self-generating, and so by Theorem 1, $W \subseteq V$. Combining, we see that W = V, which just means that *V* is a fixed point of ϕ .

To complete the proof, let *W* be any other fixed point of ϕ . Then *W* is self-generating. By Theorem 1, $W \subseteq V$, and we are done.

4. Equilibrium Paths

An alternative way to think about equilibria is to assign *paths* of play following every history, rather than continuation payoffs. Ultimately — except in very simple cases — it is the latter view that has come to be dominant in applications, but let us take a look at this alternative.

Informally, paths following every history may be pieced back together to form a strategy, provided that their description satisfies some minimal consistency requirements.

Thus think of a strategy as specifying (i) an "initial path" **a**, and (ii) paths **a**' following each *t*-history. We may think of the initial path as the "desired" outcome of the game, and of all other (noninitial) paths as "punishments".

A path *a* is (supportable as) a *perfect equilibrium path* if there exists perfect equilibrium σ such that $\mathbf{a} = \mathbf{a}(\sigma)$.

A full specification of all paths and punishments looks very complicated, but the compactness of the set of equilibrium values allows for a dramatic simplification, due to Abreu (1988).

Consider strategy profiles that have the good fortune to be completely described by an (n + 1)-vector of paths $(\mathbf{a}^0, \mathbf{a}^1, \dots, \mathbf{a}^n)$, and a simple rule that describes when each path is to be in effect.

Think of \mathbf{a}^0 as the initial or "desired" path and of \mathbf{a}^i as the punishment for player *i*. That is, any unilateral deviation of player *i* from any path will be followed by starting up the path \mathbf{a}^i . [Two or more simultaneous deviations are ignored; they are treated exactly as if no deviations have occurred.] Call the resulting strategy profile a *simple strategy profile* $\sigma(\mathbf{a}^0, \mathbf{a}^1, \dots, \mathbf{a}^n)$.

EXERCISE. Using an inductive argument on time periods, or otherwise, prove that $\sigma(\mathbf{a}^0, \mathbf{a}^1, \dots, \mathbf{a}^n)$ is uniquely defined.

Recall *V*; it is compact. Define $\mathbf{\bar{p}}^i$ to be any payoff vector \mathbf{p} that solves the problem: $\min_{\mathbf{p}\in V} p_i$. Let $\mathbf{\bar{a}}^i$ be a feasible path "generated" from $\mathbf{\bar{p}}^i$ by collecting in sequence the supporting actions for $\mathbf{\bar{p}}^i$ and for all its continuation payoffs of the form $\mathbf{p}', \mathbf{p}'', \ldots$

THEOREM **4.** (Abreu [1988]). A path **a** is supportable as a SGPE if and only if the simple strategy profile $\sigma = \sigma(\mathbf{a}; \mathbf{\bar{a}}^1, \dots, \mathbf{\bar{a}}^n)$ is a perfect equilibrium.

Moreover, to check that σ is perfect, it simply suffices to check the on-the-path inequalities

$$p_i(t) \ge (1 - \delta)d(a(t)) + \delta \bar{p}_i^i$$

for all *i* and *t*, where $p_i(t) = \sum_{s=t}^{\infty} \delta^{s-t} f_i(a(s))$.

Proof. All we need to do is observe that it is necessary and sufficient to use the worst perfect equilibrium punishments for deviations in all cases. □

The collection ($\bar{a}^1, ..., \bar{a}^n$) is called an *optimal penal code* (Abreu (1988)). The insight underlying an optimal penal code is that unless there is some extraneous reason to make the punishment fit the crime, a discounted repeated game sees no reason to use such tailored punishments.² It should be mentioned, however, that while these punishments appear to be "simple" in principle, they may be hard to compute in actual applications. Later on we shall specialize to symmetric games to obtain some additional insight into how these punishments work.

5. Comments on NASH Reversion

Is Nash reversion the worst imaginable credible punishment? To answer this question first look at the worst imaginable punishment you can inflict on a player: This is the idea of the *minmax* or *security level*: minimize the maximum payoff that a player can get from each of her actions. Formally, define

$$m_i \equiv \min_{a_i} \max_{a_i} f_i(a_i, a_{-i}) = \min_{a_i} d_i(a).$$

[Let's not worry about mixed strategies here.] Now if we didn't worry about credibility for the other players, the minmax is what they could restrict player *i* to. Simply choose the minimaxing action profile for the others and *i* will be nailed down to no more than m_i . In fact, it's more than that. In no way can we nail *i* down to anything *less* than her security level. For given the opponent strategy profile, *i* would know what is being played at every node of the game tree. She could respond accordingly to get at least m_i at every node, so her lifetime normalized payoff is also no less than m_i in any equilibrium. With this in mind, take another look at the prisoners' dilemma.

²Be warned: this result is *not* true of undiscounted repeated games.

	C_2	D_2
C_1	2,2	0,3
D_1	3,0	1,1

Notice that the minmax payoff is 1 for each player. Also, Nash reversion yields a payoff of 1 and therefore attains the security level. So the Nash reversion theorem is "enough" in the sense that anything that can be supported as a SGPE outcome can also be supported using Nash reversion.

Here is an economic example where Nash reversion is enough to support all the collusive outcomes that can conceivably be supported.

Example. Bertrand Competition. Looks just the same as the Cournot example studied earlier, except this time the firms will post *prices* rather than choose quantities. So: each firm's cost is linear in output x_i , given by cx_i . There is a demand curve P(x), except this time it will be more useful to write the demand curve in "inverse form" D(p), where D(p) is the total quantity demanded by the market when the price is p. There are n firms; each firm i chooses a price p_i . The firm with the lowest price supplies the whole market. If there are two or more firms with the lowest price, then they split the market.

Claim. If $n \ge 2$ there is a unique Nash equilibrium payoff outcome in which each firm makes zero profits. [If $n \ge 3$ there are several Nash equilibria but they all yield the same outcome in terms of payoffs.]

Proof. Left to you!

By the claim, we can see that Bertrand competition attains the security level for each firm, and therefore if we are looking to support any possible outcome it is enough to check whether a one-shot deviation from that outcome can be deterred by the threat of Nash punishment forever.

Let's work a bit more on the practicalities of doing this: look at a price p that is above the cost of production c and think about getting all firms to charge that price. First consider a price p between c (the unit cost of production) and p^* , the joint monopoly price. By conforming, each firm gets a lifetime normalized payoff of

$$(p-c)D(p)/n$$

If a firm deviates, the supremum payoff she can get is the whole market at price *p* (she needs to only deviate a tiny bit downwards). So the no-deviation condition reads:

$$(p-c)D(p)/n \ge (1-\delta)(p-c)D(p) + \delta \text{ Nash payoff } = (1-\delta)(p-c)D(p).$$

Or equivalently

$$\delta \ge 1 - (1/n).$$

Notice that this condition is independent of the price you want to support. This is a consequence of the linearity of the cost function.

Important. Just as the infinite repetition of Nash is a subgame-perfect equilibrium but at the same time inefficient, we do not have any "right" to assert that the best of the collusive equilibria will in fact come about: $p = p^*$. There are simply many equilibria, provided that

the condition (3) is satisfied, and choosing among them must involve considerations over and above the equilibrium notion itself.

Indeed, to play this theme some more, notice that even prices on the "wrong side of monopoly" can be supported as SGP equilibria. Pick some $p > p^*$. Once again, conformity yields a lifetime payoff of

$$(p-c)D(p)/n$$

Now if a firm deviates, what is the best payoff she can get? The firm will not want to undercut by a small bit: it will prefer to descend all the way to the monopoly price, earning a one-period payoff of $(p^* - c)D(p^*)$. So the no-deviation condition now reads

$$(p-c)D(p)/n \ge (1-\delta)(p^*-c)D(p^*) + \delta \operatorname{Nash} = (1-\delta)(p^*-c)D(p^*).$$

This gives us another condition on the discount factor:

$$(4) \qquad \qquad \delta \ge 1 - [\lambda(p)/n]$$

where $\lambda(p)$ is the *ratio* of collusive profits at the price *p* to monopoly profits; i.e.,

$$\lambda(p) = \frac{(p-c)D(p)}{(p^*-c)D(p^*)}$$

Note that $\lambda(p) < 1$ (because we are looking at prices higher than the monopoly price) and what is more, $\lambda(p)$ declines further as p increases, because we are in the inverted-U case. This means that higher and higher collusive prices (above the monopoly price) do create greater demands on patience, in contrast with the case in which prices are below the monopoly price.

6. PUNISHMENTS MORE SEVERE THAN NASH REVERSION

The example above notwithstanding, there are several games — the *Cournot* oligopoly being a good example — in which the one-shot Nash equilibrium fails to push agents down to their security level. This raises the question of whether more severe *credible* punishments are available.

Example. [Osborne, p. 456.] Consider the following game:

	A_2	B_2	C_2
A_1	4,4	3,0	1,0
B_1	0,3	2,2	1,0
C_1	0,1	0,1	0,0

The unique Nash equilibrium of the above game involves both parties playing *A*. But it is easy to check that each player's security level is 1. So are there equilibrium payoffs "in between"?

Consider the following strategy profile, described in two phases, which we describe as follows:

The Ongoing Path. (Phase O) Play (*B*₁, *B*₂) at every date.

The Punishment Phase. (Phase P) Play (C_1, C_2) for two periods; then return to Phase O.

Start with Phase O. If there is any deviation, start up Phase P. If there is any deviation from *that*, start Phase P again.

To check whether this strategy profile forms a SGPE, suffices to check one-shot deviations.

Phase O yields a lifetime payoff of 2. A deviation gets her the payoff

$$(1 - \delta)[3 + \delta 0 + \delta^2 0] + \delta^3 2.$$

Noting that $1 - \delta^3 = (1 - \delta)(1 + \delta + \delta^2)$, we see that a deviation in Phase O is not worthwhile if

$$2 + 2\delta + 2\delta^2 \ge 3$$

or $\delta \ge (\sqrt{3} - 1)/2$.

What about the first date of Phase P? Lifetime utility in this phase is $\delta^2 2$ (why?). If she deviates she can get 1 today, then the phase is started up again. So deviation is not worthwhile if

(5)
$$\delta^2 2 \ge (1-\delta)1 + \delta\delta^2 2$$

or if $\delta \ge \sqrt{2}/2$. This is a stronger restriction than the one for phase O so hold on to this one.

Finally, notice *without doing the calculations* that it is harder to deviate in date 2 of Phase P (why?). So these strategies form a SGPE if $\delta \ge \sqrt{2}/2$.

Several remarks are of interest here.

1. The equilibrium payoff from this strategy profile is 2. But in fact, the equilibrium bootstraps off another equilibrium: the one that actually starts at Phase P. The return to *that* equilibrium is even lower: it is $\delta^2 2$.

2. Indeed, at that lowest value of δ for which this second equilibrium is sustainable, the equilibrium exactly attains the minimax value for each player! And so *everything* that can be conceivable sustained in this example can be done with this punishment equilibrium, at least at this threshold discount factor.

3. Notice that the ability to sustain this security value as an equilibrium payoff is not exactly "monotonic" in the discount factor. In fact if the discount factor rises a bit above the minimum threshold you cannot find an equilibrium with security payoffs. But this is essentially an integer problem — you can punish for two periods but the discount factor may not be "good enough" for a three-period punishment. Ultimately, as the discount factor becomes close to 1 we can edge arbitrarily close to the security payoff *and stay in that close zone;* this insight will form the basis of the celebrated folk theorem.

Example. [Abreu (1988).] Here is a simple, stripped-down version of the Cournot example in which we can essentially try out the same sort of ideas. The nice feature about this example (in contrast to the previous one, the role of which was purely pedagogical) is that it has some collusive outcome better than the Nash which players are trying to sustain.

	L_2	M_2	H_2
L_1	10,10	3,15	0,7
M_1	15,3	7,7	-4,5
H_1	7,0	5,-4	-15, -15

Think of *L*, *M* and *H* as low, medium and high outputs respectively. Now try and interpret the payoffs to your satisfaction.

Notice that each player's maximin payoff is 0, but of course, no one-shot Nash equilibrium achieves this payoff.

1. You can support the collusive outcome using Nash reversion. To check when this works, notice that sticking to collusion gives 10, while the best deviation followed by Nash reversion yields

$$(1 - \delta)15 + \delta7.$$

It is easy to see that this strategy profile forms an equilibrium if and only if $\delta \ge 5/8$. For lower values of δ Nash reversion will not work.

2. But here is another one that works for somewhat lower values of δ . Start with (L_1, L_2) . If there is any deviation play (H_1, H_2) once and then revert to (L_1, L_2) . If there is any deviation from *that*, start the punishment up again. Check this out. The punishment value is

$$(6) \qquad -15(1-\delta) + 10\delta \equiv p,$$

and so the no-deviation constraint in the punishment phase is

$$p \ge (1 - \delta)0 + \delta p,$$

or $p \ge 0$. This yields the condition $\delta \ge 3/5$.

What about the collusive phase? In that phase, the no-deviation condition tells us that

$$10 \ge (1 - \delta)15 + \delta p$$

but (6) assures us that this restriction is always satisfied (why?). So the collusive phase is not an issue, so our restriction is indeed $\delta \ge 3/5$, the one that's needed to support the punishment phase.

3. For even lower values of δ , the symmetric punishment described above will not work. But here is something that else that will: punishments tailored to the deviator! Think of two punishment phases, one for player 1 and one for player 2. The punishment phase for player *i* (where *i* is either 1 or 2) looks like this:

$$(M_i, H_j); (L_i, M_j), (L_i, M_j), (L_i, M_j), \ldots$$

Now we have to be more careful in checking the conditions on the discount factor. First write down the payoff to players *i* and *j* from punishment phase P_i , the one that punishes *i*. It is

 $p \equiv -4(1 - \delta) + 3\delta$ in stage 1 and 3 in each stage thereafter

for the "punishee" player *i* and

 $5(1 - \delta) + 15\delta$ in stage 1 and 15 in each stage thereafter

for the "punisher" player *j*.

now, if *i* deviates in the first stage of his punishment he gets 0 and then is punished again. So the no-deviation condition is

$$v \ge (1-\delta)0 + \delta p,$$

or just plain $p \ge 0$, which yields the restriction $\delta \ge 4/7$.

What if *i* deviates in some future stage of his punishment? The condition there is

$$3 \ge (1-\delta)7 + \delta p = (1-\delta)7 - \delta(1-\delta)4 + \delta^2 3,$$

but it is easy to see that this is taken care of by the $\delta \ge 4/7$ restriction.

Now we must check *j*'s deviation from *i*'s punishment! In the second and later stages there is nothing to check (why?). In stage 1, the condition is

$$5(1-\delta) + 15\delta \ge (1-\delta)7 + \delta p.$$

[Notice how *j*'s punishment is started off if she deviates from *i*'s punishment!] Compared with the previous inequality, this one is easier to satisfy.

Finally, we must see that no deviation is profitable from the original cooperative path. This condition is just

$$10 \ge (1 - \delta)15 + \delta p,$$

and reviewing the definition of *p* we see that no further restrictions on δ are called for.

4. Can we do *still* better? We can! The following punishment exactly attains the minimax value for each agent for all $\delta \ge 8/15$.

To punish player *i*, simply play the path

$$(L_i, H_j); (L_i, H_j), \ldots$$

forever. Notice that this pushes player *i* down to minimax. Moreoover, player *i* cannot profitably deviate from this punishment.

But player *j* can! The point is, however, that in that case we will start punishing player *j* with the corresponding path

$$(L_{j}, H_{i}); (L_{j}, H_{i}), \ldots$$

which gives *her* zero. All we need to do now is to check that a one-shot deviation by j is unprofitable. Given the description above, this is simply the condition that

$$7 \ge (1 - \delta)15 + \delta$$
 punishment payoff $= (1 - \delta)15$.

This condition is satisfied for all $\delta \ge 8/15$.

So you see that in general, we can punish more strongly than Nash reversion, and what is more, this is a variety of such punishments, all involving *either* a nonstationary time structure ("carrot-and-stick", as in part 2) or a family of player-specific punishments (as in part 4) or both (as in part 3). This leads to the Pandora's Box of too many equilibria. The repeated game, in its quest to explain why players cooperate, also ends up "explaining" why they might fare even worse than one-shot Nash!

7. Symmetric Games: A Special Case

Finding individual-specific punishments may be a very complicated exercise in actual applications. See Abreu [1986] for just how difficult this exercise can get, even in the context of a simple game such as Cournot oligopoly. The purpose of this section is to identify the worst punishments in a subclass of cases when we restrict strategies to be symmetric in a strong sense. A game *G* is *symmetric* if $A_i = A_j$ for all players *i* and *j*, and the payoff functions are symmetric in the sense that for every permutation *p* of the set of players $\{1, ..., n\}$,

$$f_i(a) = f_{p(i)}(a_p)$$

for all *i* and action vectors *a*, where a_p denotes the action vector obtained by permuting the indices of *a* according to the permutation *p*.

A strategy profile σ is *strongly symmetric* if for every *t*-history h(t), $\sigma_i(t)[h(t)] = \sigma_j(t)[h(t)]$ for all *i* and *j*. Note that the symmetry is "strong" in the sense that players take the same actions after all histories, including asymmetric histories.

Now for some special assumptions. We will suppose that each A_i is the (same) interval of real numbers, unbounded above and closed on the left. Make the following assumptions:

[S.1] Every payoff function f_i is quasiconcave in a_i and continuous. Moreover, the payoff to symmetric action vectors (captured by the scalar *a*), denoted f(a), satisfies $f(a) \rightarrow -\infty$ as $a \rightarrow \infty$.

[S.2] The best payoff to any player when all other players take the symmetric action a, denoted by d(a), is nonincreasing in a, but bounded below.

[S.3] For every symmetric action *a* for the others, and 0 for *i*, $f_i(0, a)$ is bounded in *a* (for instance, can set $f_i(0, a) = 0$).

EXERCISE. As you can tell, Conditions [S.1]–[S.3] are set up to handle something like the case of Cournot oligopoly. Even though the action sets do not satisfy the compactness assumption, the equilibrium payoff set is nevertheless compact. How do we prove this?

[a] First prove that a one-shot equilibrium exists.

[b] This means that the set of strongly symmetric perfect equilibrium payoffs *V* is nonempty. Now, look at the infimum perfect equilibrium payoff. Show that it is bounded below, using S.2. Using S.1, show that the supremum perfect equilibrium payoff is bounded above.

[c] Now show that the paths supporting infimum punishments indeed are well-defined, and together they form a simple strategy profile which is a SGPE.

[d] Finally, prove the compactness of *V* by using part [c].

[For the answer to the exercise, see the appendix.]

With the above exercise worked out, we can claim that there exists best and worst symmetric payoffs v^* and v_* respectively, in the class of all strongly symmetric SGPE. The following theorem then applies to these payoffs.

THEOREM 5. Consider a symmetric game satisfying [S.1] and [S.2]. Let v^* and v_* denote the highest and lowest payoff respectively in the class of all strongly symmetric SGPE. Then

[a] The payoff v_* can be supported as a SGPE in the following way: Begin in phase I, where all players take an action a_* such that

$$(1-\delta)f(a_*) + \delta v^* = v_*.$$

If there are any defections, start up phase I again. Otherwise, switch to a perfect equilibrium with payoffs v^* .

[b] The payoff v^* can be supported as a SGPE using strategies that play a constant action a^* as long as there are no deviations, and by switching to phase 1 (with attendant payoffs v_*) if there are any deviations.

Proof. Part [a]. Fix some strongly symmetric equilibrium $\hat{\sigma}$ with payoff v_* . Because the continuation payoff can be no more than v^* , the first period action along this equilibrium must satisfy

$$f(a) \ge \frac{-\delta v^* + v_*}{1 - \delta}.$$

Using [S.1], it is easy to see that there exists a_* such that $f(a_*) = \frac{-\delta v^* + v_*}{1-\delta}$. By [S.2], it follows that $d(a_*) \le d(a)$. Now, because $\hat{\sigma}$ is an equilibrium, it must be the case that

$$v_* \ge (1-\delta)d(a) + \delta v_* \ge (1-\delta)d(a_*) + \delta v_*,$$

so that the proposed strategy is immune to deviation in Phase I. If there are no deviations, we apply some SGPE creating v^* , so it follows that this entire strategy as described constitutes a SGPE.

Part [b]. Let $\tilde{\sigma}$ be a strongly symmetric equilibrium which attains the equilibrium payoff v^* . Let $a \equiv a(\tilde{\sigma})$ be the path generated. Then *a* has symmetric actions a(t) at each date, and

$$v^* = (1 - \delta) \sum_{t=0}^{\infty} \delta^t f(a_t).$$

Clearly, for the above equality to hold, there must exist some date *T* such that $f(a_T) \ge v^*$. Using Condition 1, pick $a^* \ge a_T$ such that $f(a^*) = v^*$. By Condition 2, $d(a^*) \le d(a_T)$. Now consider the strategy profile that dictates the play of a^* forever, switching to Phase I if there are any deviations. Because $\tilde{\sigma}$ is an equilibrium, because v_* is the worst strongly symmetric continuation payoff, and because v^* is the largest continuation payoff along the equilibrium path at any date, we know that

$$v^* \ge (1 - \delta)d(a_T) + \delta v_*.$$

Because $d(a_T) \ge d(a^*)$,

 $v^* \ge (1 - \delta)d(a^*) + \delta v_*$

as well, and we are done.

The problem of finding the best strongly symmetric equilibrium therefore reduces, in this case, to that of finding two numbers, representing the actions to be taken in two phases.

Something more can be said about the punishment phase, under the assumptions made here.

THEOREM 6. Consider a symmetric game satisfying [S.1] and [S.2], and let (a_*, a^*) be the actions constructed to support v_* and v^* (see statement of Theorem 5). Then

$$d(a_*) = v_*$$

14

Proof. We know that in the punishment phase,

(7)
$$v_* \ge (1-\delta)d(a_*) + \delta v_*$$

while along the equilibrium path,

(8)
$$v_* = (1-\delta)f(a_*) + \delta v^*.$$

Suppose that strict inequality were to hold in (7), so that there exists a number $v < v_*$ such that

(9)
$$v \ge (1-\delta)d(a_*) + \delta v.$$

Using Condition 1, pick $a \ge a_*$ such that

(10)
$$v = (1 - \delta)f(a) + \delta v^*.$$

[To see that this is possible, use Condition 1, (8), and the fact that $v < v_*$.] Note that $d(a) \le d(a_*)$, by Condition 2. Using this information in (9), we may conclude that

(11)
$$v \ge (1-\delta)d(a) + \delta v.$$

Combining (10) and (11), we see from standard arguments (check) that v must be a strongly symmetric equilibrium payoff, which contradicts the definition of v_* .

8. The Folk Theorem

The folk theorem reaches a negative conclusion regarding repeated games. Repeated games came into being as a way of reconciling the observation of collusive (non-Nash) behavior with some notion of individual rationality. The folk theorem tells us that in "explaining" such behavior, we run into a dilemma: we end up explaining too much. Roughly speaking, *every* individually rational payoff is supportable as a SGPE, provided that the discount factor is sufficiently close to unity.

Recall that the *security level* of player *i* is given by the value

$$\hat{v}_i \equiv \min_{a \in A} d_i(a).$$

Let \hat{a}^i be an action vector such that \hat{v}_i is exactly attained; i.e., $f_i(\hat{a}^i) = \hat{v}_i$. Let \hat{v}^i_j be the payoff to j when this is happening. [For instance, $\hat{v}_i = \hat{v}^i_i$.] Normalize the security level to equal zero for each player.

For each δ , denote by $V(\delta)$ the set of all (normalized) perfect equilibrium payoffs.

THEOREM 7. Define F^{**} to be the set of all individually rational feasible payoffs, i.e.,

$$F^{**} \equiv F^* \cap \{\mathbf{p} \in \mathbb{R}^n | \mathbf{p} \ge 0 \text{ for all } i\},\$$

and assume that F^{**} is n-dimensional. Then for each $\tilde{\mathbf{p}}$ in F^{**} and each $\epsilon > 0$, there exists a payoff vector \mathbf{p} in F^{**} and also in the ϵ -neighborhood of $\tilde{\mathbf{p}}$ such that $\mathbf{p} \in V(\delta)$ for all δ sufficiently close to unity.

Proof. For simplicity we shall assume in this proof that any point in F^* , the convex hull of the set of feasible payoffs, can be attained by some pure strategy combination. Later, we indicate how the proof can be extended when this is not the case.

Pick any $\tilde{\mathbf{p}} \in F^{**}$, and $\epsilon > 0$. Because F^{**} has full dimension, it is possible to find \mathbf{p} in the ϵ -neighborhood of $\tilde{\mathbf{p}}$ such that $\mathbf{p} \in \operatorname{int} F^{**}$. Now pick *n* payoff vectors $\{\bar{\mathbf{p}}^i\}_{i \in N}$ (each in F^{**}) "around" \mathbf{p} as follows:

$$\bar{p}_i^i = p_i,$$
$$\bar{p}_j^i = p_j + \gamma \text{ for } j \neq i,$$

for some $\gamma > 0$. These vectors will be fixed throughout the proof. By our simplifying assumption, there are action vectors $a, \bar{a}^1, \ldots, \bar{a}^n$ such that $\mathbf{f}(a) = \mathbf{p}$ and $\mathbf{f}(\bar{a}^i) = \bar{\mathbf{p}}^i$ for each $i = 1, \ldots, n$.

The first of these action vectors is, of course, going to support the desired payoff, and the latter are going to serve as "rewards" to people who carry out punishments that may not be in their own short-term interests. The punishments, in turn, are going to be derived from the actions $\{\hat{a}^i\}$ that minimax particular players and drive them down to their security levels. Now for a precise statement. For each i = 0, 1, ..., n, consider the paths

$$\mathbf{a}^{0} \equiv (a, a, a, ...),$$
$$\mathbf{a}^{i} = (\hat{a}^{i}, \dots \hat{a}^{i}) \text{ for } T \text{ periods}$$
$$= (\bar{a}^{i}, \bar{a}^{i}, \dots) \text{ thereafter},$$

where *T* is soon going to be cunningly chosen (see below).

Consider the simple strategy profile $\sigma \equiv \sigma(\mathbf{a}^0, \mathbf{a}^1, \dots, \mathbf{a}^n)$. We claim that there exists an integer *T* and a $\delta^* \in (0, 1)$ such that for all $\delta \in (\delta^*, 1)$, σ is a perfect equilibrium.

For convenience, let us record the normalized payoff to player *i* along each of the paths. We have, first,

$$F_i(\mathbf{a}^0) = p_i$$

for all *i*, and for all $0 \le t \le T$,

$$F_i(\mathbf{a}^j, t) = \hat{v}_i^j (1 - \delta^{T+1-t}) + \delta^{T+1-t} p_i^j(\gamma),$$

where $p_i^j(\gamma) = p_i + \gamma$ if $i \neq j$, and $p_i^j(\gamma) = p_i$ if i = j. Of course, for all $t \ge T + 1$,

$$F_i(\mathbf{a}^j, t) = p_i^j(\gamma).$$

We must check the no-deviation constraints from each path. (That's enough, by the one shot deviation principle.)

Deviations from a⁰. Suppose that player *i* were to deviate from the path a^0 . Then he gets minimaxed for T + 1 periods, with 0 return, after which he gets p_i again forever. If *M* is the maximum absolute value of one shot payoffs in the game, the best deviation along the path is bounded above by *M*, so that the no-deviation condition is surely satisfied if

$$p_i \ge (1 - \delta)M + \delta^{T+1}p_i.$$

This inequality holds if

(12)
$$\frac{1-\delta^{l+1}}{1-\delta} \ge \frac{M}{p_i}$$

[Note that $p_i > 0$ (why?) so that this inequality makes perfect sense.] Now look at (12). As $\delta \rightarrow 1$, the LHS goes to T + 1 (why?). So if we take

(13)
$$T \ge \max_{i} \frac{M}{p_{i}},$$

then (12) is automatically satisfied for all δ sufficiently close to unity.

Deviations from a^{*j*}. First check player *j*'s deviation. If (12) is satisfied (for i = j), player *j* will never deviate from the second phase of his own punishment, because he just goes back to getting p_j . The deviations may have differnt values, of course, but we have bounded these above by *M* anyway to arrive at (12). In the first phase, note that by construction, player *j* is playing a one-shot best response. So there is no point in deviating, as there is no short-term gain and it will be followed by restarting his punishment.

It remains to check player *i*'s deviation from the path \mathbf{a}^j when $i \neq j$. By the same argument as in the previous paragraph, a deviation in the second phase is not worthwhile, provided that (12) is satisfied. We need to check, then, that player *i* will cooperate with *j*'s punishment in the first phase. He will do so if for each integer *t* that records the number of periods left in the first phase,

$$(1-\delta^t)\hat{v}_i^j + \delta^t(p_i+\gamma) \ge (1-\delta)M + \delta^{T+1}p_i.$$

Replace \hat{v}_i^j by -M on the LHS, and t by T + 1. On the RHS, replace $(1 - \delta)M$ by $(1 - \delta^{T+1})M$. Then noting that M and p_i are both positive, it is clear that the above inequality holds if

$$-(1 - \delta^{T+1})M + \delta^{T+1}(p_i + \gamma) \ge (1 - \delta^{T+1})M + \delta^{T+1}p_i,$$

or if

(14)
$$\frac{\delta^{T+1}}{1-\delta^{T+1}} \ge \frac{2M}{\gamma}$$

Now it should be clear that for any *T* satisfying (13), both conditions (12) and (14) are satisfied for all δ sufficiently close to unity. So we are done.

It remains to remark on the case where the payoffs **p** and its "constructed" neighbors are not exactly generated by pure action vectors. The proof then has to be slightly modified. First we perturb **p** a tiny bit if necessary and the choose δ close enough to unity so that in a sufficiently large number of finite periods, we can get **p** as the convex combination of payoffs from various pure actions (where the convexification is being carried out intertemporally). Then all we have to do is to use a nonstationary path (with a finite periodicity) to generate **p**. We do the same for each of the payoff vectors $\mathbf{\bar{p}}^i$ as well. The proof then goes through just the same way as before. The restrictions created by the choice of δ go "in the same direction" anyway.

The full-dimensionality of F^* is needed in general (though not for two-player games). Without it, the theorem is generally false. Consider the following

Example. Player 1 chooses rows, player 2 chooses columns, and player 3 chooses matrices:

Each player's minmax value is 0, but notice that there is no action combination that *simultaneously* minmaxes all three players. E.g., to minimax player 3, 12 play *UR*. To minmax 2, 13

	L	R		L	R
U	1,1,1	0,0,0	U	0,0,0	0,0,0
D	0,0,0	0,0,0	D	0,0,0	1,1,1

play *U*2. To minmax player 1, 23 play *L*2. Nothing works to simultaneously minmax all three players.

Let α be the lowest equilibrium payoff for any player. Note that

$$\alpha \ge (1-\delta)D + \delta\alpha,$$

where *D* is the largest deviation payoff to some player under any first-period action supporting α . It can be shown (even with the use of observable mixed strategies) that $D \ge 1/4$. So $\alpha \ge 1/4$. No folk theorem.

The problem is that we cannot separately minmax each deviator and provide incentives to the other players to carry out the minmaxing, because all payoffs are common. If there is enough "wiggle-room" to separately reward the players for going into the various punishment phase, then we can get around this problem, as noted in the proof of Theorem 7.

Appendix

Answer to the Exercise for Strongly Symmetric Games.

For each person, write the action set as $[0, \infty)$. Fix some symmetric action *a* for the other players and look at one player's best response. S.1 tells us that this player's payoff is quasiconcave in his own actions and S.2 tells us that the best payoff is well-defined. By quasiconcavity, the set of best responses A(a) to *a* is convex-valued. By continuity of payoffs, A(a) is upperhemicontinuous.

Now I claim that for large *a* we have a' < a for all $a' \in A(a)$. Suppose on the contrary that there is $a_m \to \infty$ and $a'_m \in A(a_m)$ for each a_m such that $a'_m \ge a_m$. So there is a sequence $\lambda_m \in (0, 1]$ such that $a_m = \lambda_m a'_m$ for all a_m . By quasiconcavity, $f(a_m) \ge \min\{f(0, a_m), f(a'_m, a_m)\}$. But the former term is bounded by S.3, and the latter term is bounded below by S.2. This contradicts the fact that $f(a_m) \to -\infty$ as $m \to \infty$.

This proves, by a slight variation on the intermediate value theorem, that there exists a^* such that $a^* \in A(a^*)$. Clearly, a^* is a strongly symmetric equilibrium, which proves [a].

[a] This means that the set of strongly symmetric perfect equilibrium payoffs *V* is nonempty. Simply repeat a^* regardless of history. Now define $d \equiv \inf_a d(a) > -\infty$ by assumption. *d* is like the strongly symmetric security level. Lifetime payoffs can't be pushed below this. Therefore the infimum of payoffs in *V* is at least as great as *d*.

Of course, the supremum is bounded because all one-shot payoffs are bounded by assumption, so in particular the symmetric equilibrium payoff is bounded above.

[c] Let p^m be a sequence of strongly symmetric equilibrium payoffs in *V* converging down to the infimum payoff *p*. For each such p^m let $a^m(t)$ be an action path supporting p^m using strongly symmetric action profiles at any date. Let *M* be the maximum strongly symmetric payoff in

the game. Then, because infimum payoffs in V are bounded below, say by d, it must be the case that

$$(1-\delta)f(a^m(t)) + \delta M \ge d$$

for every date *t*. But this means that there exists an upper bound \bar{a} such that $a^m(t) \le \bar{a}$ for every index *m* and every date *t*.

This bound allows us to extract a convergent subsequence of m —call it m' — such that for every t,

$$a^{m'}(t) \rightarrow a(t)$$

It is very easy to show that that the simple strategy profile defined as follows:

"Start up $\{\underline{a}(t)\}$. If there are any deviations, start it up again,"

is a simple penal code that supports the infimum punishment.

[d] Finally, prove the compactness of *V*. Take any payoff sequence p^m each of which lies in *V*, converging to some *p*. Each can be supported by some action path $a^m(t)$, with the threat of starting up the simple penal code of [a3] in case there is any deviation. Take a convergent subsequence of $a^m(t)$, call the pointwise limit path a(t), and show that it supports *p* with the threat of retreating to a(t) if there is any deviation.

By the way, why is condition [S.3] needed? Can you provide a counterexample to existence otherwise?