22 Long-Term Competition—A Game-Theoretic Analysis

with L. S. Shapley

1 Introduction

There have been continuing expressions of interest from a variety of quarters in the development of techniques for modelling national behavior in a long-term context of continuing international rivalry—for short, "long term competition." The most characteristic feature of these models is that they extend over time in a fairly regular or repetitive manner. The underlying structure of possible actions and consequences remains the same, though parameters may vary and balances shift, and the decisions and policies of the national decision makers are by no means constrained to be constant or smoothly-varying, or even "rational" in any precisely identifiable sense. The use of game theory or an extension thereof is obviously indicated, and considerable theoretical progress has been made in this area. But the ability of the theory to handle real applications is still far from satisfactory. The trouble lies less with the *descriptive modelling*, i.e., formulating the "rules of the game" in a dynamic setting, than with the choice of a *solution concept* that will do dynamic justice to the interplay of motivations of the actors. (Game theoreticians, like mathematical economists, have always been more comfortable with static than dynamic models.) Since any predictions, recommendations, etc. that a mathematical analysis can produce will likely be very sensitive to the rationale of the solution that is used, and since the big difficulties are conceptual rather than technical, it seems both possible and worthwhile to discuss salient features of the theory without recourse to heavy mathematical apparatus or overly formal arguments, and thereby perhaps make the issues involved accessible to at least some of the potential customers for the practical analyses that we wish we could carry out in a more satisfactory and convincing manner.

Two general types of "solution concept" are distinguished in game theory: *cooperative* notions, such as the core, bargaining set, von Neumann–Morgenstern stable sets, and Shapley value; and *noncooperative* notions, principally the Nash equilibrium point and its variants and elaborations, but including also the max-min solution based on "safety level" or "worst case" considerations. Cooperative notions are appropriate for situations where contracts among players are customarily adhered to and can be made legally binding; noncooperative notions where there is mistrust and no external enforcement mechanisms are available. The long-term international scene is most naturally classified as noncooperative, since there is no effective international jurisdiction

This chapter originally appeared in *Essays in Game Theory in Honor of Michael Maschler*, edited by N. Megiddo, pp. 1–15, Springer-Verlag, New York, 1994. Reprinted with permission.

in most cases, even in the short run. Adherence to major international agreements is essentially a matter of national self-interest, and to be effective in the long run such agreements must be written to be self-enforcing, i.e., so that it is to the continuing advantage of all sides to adhere to them.

Quite a bit is known about Nash noncooperative equilibria in "continuingly competitive" situations, and we shall review some of this material here. It turns out that individual self-interest in such situations can in fact dictate a kind of cooperative behavior, in many cases, sustained by the fear of "punishment" by the other players for failing to "cooperate" with the general plan—this in spite of the fact that the players have no way of legally binding themselves to carry out such punishment. The ability of the noncooperative theory to describe such arrangements and to account for their stability in a "selfish" world is an encouraging point in its favor. The price that is paid, however, is the high degree of nonuniqueness in the Nash solutions (as revealed in the two theorems described below), which removes from this theory most of its predictive power.

2 Repeated Game Models

In this section we shall review some of the known theory of a special kind of "continuingly competitive" game: that of *repeated games*. Given a finite game G in strategic form,[1] we consider an infinite game G^*, each play of which consists of an infinite repetition of plays of G. In each play of G, or "round of G^*" the players are assumed to know the outcomes of all previous rounds. The payoff for G^* may be assumed to be of the limiting average form:[2]

$$\lim_{m\to\infty} \frac{1}{m} \sum_{t=1}^{m} h_t. \tag{2.1}$$

Here h_j is the payoff for G in the j'th round of G^*. Many authors call G^* the *supergame* of G.

An alternative form of the payoff for G^* involves discounting of future payoffs at a positive discount rate:

$$\sum_{t=1}^{\infty} (1-\rho)^t h_t. \tag{2.2}$$

1. a.k.a. "normal form."

2. A technical difficulty is that this limit need not always exist; this technical difficulty has a technical solution, which we do not wish to get involved with at present.

For the time being we shall confine our discussion to the limiting average form (1), which treats the future as no less important than the present. Indeed, cumulatively, the future is all important in (1), since the contribution from any period of finite length will wash out in the long run. Nothing you actually do makes any difference; only your policies for the indefinite future have any significance. Despite this peculiarity, however, it should be remarked that for many purposes, both technical and conceptual, the limiting-average case behaves like the limit of the discounted-sum case, as the discount rate ρ in (2) goes to zero. Thus, used with care, the limiting average form can serve as an approximation to situations where a very low discount rate is appropriate. Long-term competition, almost by definition, would appear to fall into this category.

The basic theorem[3] about supergames states that a *necessary and sufficient condition for* $h = (h^1, \ldots, h^n)$ *to be the payoff vector of some Nash equilibrium point of* G^* *is that it be feasible and individually rational in* G.

Let us explain the key terms in this theorem. A "payoff vector" is simply an n-tuple of real numbers, where n is the number of players. The term "payoff vector" is used because the n coordinates signify the payoffs to the n players. By "feasible" we here mean "feasible in correlated strategies"; that is, a payoff vector is feasible if and only if it is in the convex hull of the set of payoff vectors that can be obtained by having the players play pure strategies. A payoff vector is called "individually rational" if each player receives at least his *min-max payoff*, which is the level of payoff below which he cannot be forced by the remaining players.[4] Finally, a "Nash equilibrium point," or "EP" is an n-tuple of strategies—one for each player in the game—such that each player's strategy is a best response to the $(n-1)$-tuple of the other players' strategies. In other words, no player can improve his own payoff by "defecting" to another strategy while the other players are held fixed.

To clarify the meaning of this theorem, let us see what it says about the well known "Prisoner's Dilemma." This is the two-player game whose strategic form is of the type given in the following table:

	Player II	
Player I	4, 4	0, 5
	5, 0	1, 1

The set of all feasible payoff vectors is indicated by the horizontally

3. This is a "folk-theorem"; it has never been published, but is well known to most workers in the field.

4. When there are just two players, min-max = max-min.

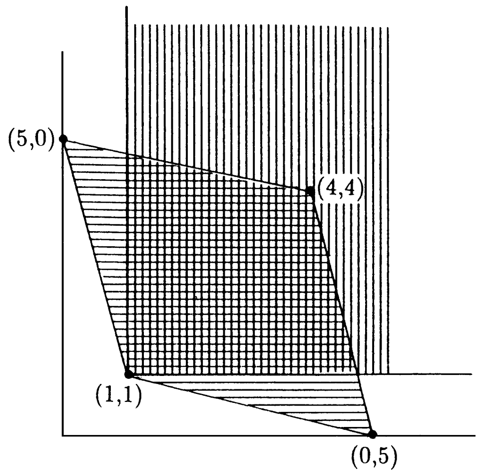

Figure 1
Feasible and individually rational payoffs in the Prisoner's Dilemma.

hatched region in Fig. 1, which is the convex hull of the four payoffs in the table. Since the minmax payoff to each player is 1, the set of all individually rational payoff vectors is indicated by the vertically hatched region. By the theorem, then, the set of payoff vectors arising from equilibrium points in the supergame is given by the *cross-hatched* region. Note in particular that the point (4, 4)—the traditional "cooperative" outcome—appears as the payoff to an equilibrium point in the supergame of the Prisoner's Dilemma. The proof of the theorem is not difficult, and as the idea of the proof is important to a proper understanding of the situation we shall take a little space to outline it here. The "necessity" part is easily established; it is intuitively clear that equilibrium is not possible if any player is below his guaranteed minimum. The more interesting and significant part of the proof is the "sufficiency."

Assume for simplicity that $n = 2$ (there are just two players). Suppose h is a feasible, individually rational payoff vector. Here we may write

$$h = \sum_{m=1}^{k} \alpha_m h_m,$$

where the α are nonnegative weights that sum to 1 and the h_m are payoff vectors corresponding to pure strategy pairs in G. Suppose first that the α_m are rational numbers and express them in the form $\alpha_m = p_m/q$, where p_m are positive integers and q is their sum. The payoff vector h can then be achieved as a limiting average in G^* by having the players play for p_1 consecutive periods an n-tuple that achieves h_1, then for p_2 consecutive periods an n-tuple that achieves h_2, and so on; after q periods, we start again from the beginning.

If the α_m are irrational, the same effect can be attained by approximating to them by rational numbers, with increasing values of q, and playing once through each approximation in turn, to yield the desired limiting average.

This procedure, however, does not yet describe a Nash equilibrium point in G^*, and in fact does not even describe a pair of supergame strategies. It only describes a particular, feasible course of play. A supergame *strategy* must describe each player's responses to all possible actions of the other player, not only when he "plays along" with a prescribed course of play, such as the one described above, but also when he "defects." This is where the requirement that h be individually rational comes in.

Since h is individually rational, we have for each player i

$$h^i \geqq \max_{\sigma} \min_{\tau} H^i(\sigma, \tau),$$

where H^i is the payoff function to player i in the game G, σ ranges over all mixed strategies of player i, and τ ranges over all mixed strategies of the other player, j. By von Neumann's minimax theorem, there is a mixed strategy $\underset{\sim}{\tau}$ of j such that for all mixed strategies σ' of i,

$$\min \max H^i(\sigma, \tau) \geqq H^i(\sigma', \underset{\sim}{\tau});$$

hence in particular

$$h^i \geqq H^i(\sigma', \underset{\sim}{\tau})$$

for all mixed strategies σ' of i. That means that by playing $\underset{\sim}{\tau}$, j can hold i down to his max-min value, and *a fortiori* to h^i.

We may now describe an EP in G^* as follows: The players start by playing to obtain an average payoff of h, as outlined above. If at any stage a player i "defects"—i.e., does not play the prescribed choice in G for that round—then starting from the next round, the other player j plays the mixed strategy $\underset{\sim}{\tau}$ forever after. This will hold i's limiting average payoff down to at most h^i, so that he will have gained nothing by his defection. Thus, h is indeed the payoff to an EP.

3 Perfect Equilibrium Points

The above line of proof has been subjected to the following criticism: Though there is no question that the strategy pair as described constitutes an equilibrium point, it is not clear under what circumstances it would ever be used. In particular, it is possible that the strategy $\underset{\sim}{\tau}$, while holding player i down to his minmax payoff, may also be very[5] disadvantageous to the player using it (or to one of the set of players participating in it, if $n > 2$). The equilibrium point dictates that $\underset{\sim}{\tau}$ will continue to be played "forever," even if i defects only once. As we pointed out at the end of the previous section $\underset{\sim}{\tau}$ is supposed to play the role of a deterrent. But an infinite unremitting repetition of $\underset{\sim}{\tau}$ seems like an unreasonable response to a single act of defection, except for the fact that—in view of (1)—any finite period of "punishment" is no punishment at all. But by the same token, a single defection is also insignificant in the limit. Thus, the threatened response may still seem unreasonable, especially when, as is often the case, it is disadvantageous or costly to the user, and hence such an unremitting repetition may not be believable as a deterrent. To have a word for these EPs, let us call them "grim."

Let us try to pinpoint the dissatisfaction with grim EPs in a slightly more general framework. The "knowledge" that j (or, more generally $N \setminus \{i\}$) will respond to a defection on the part of i by an unrelenting stream of $\underset{\sim}{\tau}$ is what keeps i from defecting; but if i does in fact defect, it may no longer be profitable for j (or $N \setminus \{i\}$) to respond with $\underset{\sim}{\tau}$. This is what makes $\underset{\sim}{\tau}$ unbelievable.

This kind of reasoning motivated a specialization of the notion of equilibrium point, first considered by R. Selten [2] and called by him a perfect equilibrium point.[6] To define this notion, we must recall more precisely the definition of a "strategy" for the player i in the supergame G^*. This is a function that tells i which pure G-strategy to choose on each round, as a function of what all the players, including i himself, did on all previous rounds. For each positive integer k, define G_k^* to be the "subgame" starting from the k'th period, i.e., after $k-1$ rounds have been played, and continuing indefinitely from that point. Thus, $G^* = G_1^*$. Each n-tuple of strategies in G^*, together with a series of actual actions on the part of all players in the first $k-1$ rounds, induces an n-tuple of strategies in G_k^*. An n-tuple $\sigma = (\sigma_1, \ldots, \sigma_n)$ of strategies in G^* is called a *perfect equilibrium point* (or PEP) if for each k and for each series of actions of the players in the first $k-1$ periods of G^*, the induced n-tuple is an EP of G_k^*.

If we set $k = 1$ we see that a PEP is in particular an EP.

It's easy to see that a grim EP is in general *not* perfect, since if player i defects on round $k-1$ it will in general not be a best response in G_k^* for the other players to "punish" him; it may even be individually irrational. It thus appears that the notion of perfect equilibrium point might hold an

5. Compare [1].

6. See also [3].

answer to the problem of the believability of deterrents. In the next three sections we shall explore this matter somewhat further. To some extent our hopes turn out to be in vain: we shall find in the next section that the payoffs associated with PEPs in G^* are the same as those for grim EPs though the grim EPs themselves are excluded. The methods may be different, so to speak, and possibly more "believable" but the upshot is the same: there is no narrowing of the class of outcomes that can be sustained in equilibrium.

However, when we modify the payoff in G^* by introducing a positive discount rate (2), which we do in Sec. 5, we find that requiring "perfection" can significantly reduce the set of equilibrium outcomes. Moreover, the concept of believability does appear to play a significant role in the description of the perfect equilibria. Thus, the notion of perfection of equilibria, though not a panacea, does appear to give us a somewhat better handle on some of the problems that we wish to model.

4 Characterization of PEPs

This section is devoted to the following theorem:

THEOREM 4.1 *The set of payoffs to perfect EPs in G^* coincides with the set of payoffs to ordinary EPs—i.e., it is the set of all feasible, individually rational payoffs in G.*

Again, to gain a good understanding of this theorem it is essential to outline the proof. As before, it is "sufficiency" that is the interesting part of the proof; the "necessity" follows from the previous theorem. We shall find the argument considerably more intricate than before.

To simplify the presentation we again assume that there are only two players. Moreover, in order to make the use of mixed strategies unnecessary, we shall assume that G is not in strategic form, with simultaneous choices by the two players, but is a game of perfect information with a single move for each player and no chance moves. Player I moves first, II is informed of I's move, and then II moves.[7] None of these assumptions are really required for the truth of the theorem, but they do simplify the proof.

Suppose h is a feasible, individually rational payoff vector of G. We shall describe a PEP with payoff h. As in Section 2, the description will be couched in terms of a tentative "agreement" on prescribed course of play. The agreement starts out as before with a sequence of choices

7. Note that the apparent asymmetry of the players disappears in the supergame: in G^* the players move alternately, each with perfect information.

which, when adhered to by both players, will lead to the desired limiting average payoff h. Let us call this "the cooperative sequence." Next, we shall specify how the players react to a defection—i.e., a departure from the cooperative sequence by one of the players. In the previous proof, the reaction was unrelenting punishment. Here, instead, the PEP strategies will specify that a defection on the part of either player be punished by a sequence of choices by the other that forces the defector's average payoff down to within ε of his max-min value, where ε is a small number that may depend on the "date" of the defection. After the defector has thus been "beaten to within an inch of his life," the punisher relents and prescribed play returns to the cooperative sequence at the point of defection.

If should be noted, however, that not only are defections from the cooperative sequence punished, but also defections from any punishing sequence (in the subgame resulting from an earlier defection) are punished. A player who "should" punish and does not do so will himself be punished. This is what provides the motivation for the punisher actually to carry out the punishment, and so keeps the EP perfect.

The situation is a little complex; in order to convince ourselves that we have actually described a PEP we shall now give a more formal treatment. Without loss of generality we may assume that the number of choices available to I on each move is the same as the number of choices available to II; call the number m. Thus, $M = \{1, \ldots, m\}$ is the set of possible choices of the players at each move. When it's Player I's turn in the n-th round of G^* he has before him the full history of previous moves; this takes the form of a sequence $(x_1, \ldots, x_{2n-2})$, where $x_i \in M$ represents the choice made on the i'th move in G^* (it is I's or II's choice according as i is odd or even). Similarly when II must move, he has before him a sequence $(x_1, \ldots, x_{2n-1})$, where again each $x_i \in M$. Let us call any finite sequence of members of M a *history*. A *strategy* for Player I [Player II] may be defined as a function from histories of even [odd] length to M. Thus, a pair of strategies is simply a function f from the set of *all* histories to M.

Now let $h = (h^I, h^{II})$ be the given feasible individually rational payoff vector. Let $(c_1, c_2, c_3, \ldots)$ be a fixed *cooperative sequence*, i.e., a sequence of moves leading to the payoff h in G^*. Let $\bar{p}$ be a G-strategy for I (i.e., a member of M) that holds II to his max-min payoff in G, and let $\bar{q}(\cdot)$ be a G-strategy for II (i.e., a function from M to M) that holds I to his max-min payoff in G.

We wish to define a strategy-pair f for G^* which is a PEP and whose associated payoff is h. The definition of f will be inductive, based on the length k of the history on which it is being defined. On a history of length 0 we define f to be c_1; this simply means that the PEP will prescribe the choice of c_1 for the first move of Player I. Suppose now that f has been defined on all histories up to length $k-1$; we wish to define it on all histories of length k. Let $(x_1, \ldots, x_k)$ be such a history. We shall say that the ℓ'th move of that history $(1 \leqq \ell \leqq k)$ is a *defection* if $x_\ell \neq f(x_1, \ldots, x_{\ell-1})$. If $(x_1, \ldots, x_k)$ contains no defection, we define $f(x_1, \ldots, x_k) = c_{k+1}$. Otherwise, suppose the *most recent* defection in $(x_1, \ldots, x_k)$ occurred at move ℓ. If ℓ and $k+1$ have the same parity—i.e., the player who is about to move is the same as the one who most recently defected—then we define $f(x_1, \ldots, x_k) = c_{k+1}$. If ℓ and $k+1$ have opposite parity, consider first the case in which ℓ is even, i.e., Player II was the last to defect. In that case k is also even, so exactly $k/2$ rounds of G^* have now been completed and it is Player I's turn to move. Consider the average payoff of Player II as measured at the end of each of the rounds $\ell/2+1, \ldots k/2$, and let $\varepsilon_\ell = 1/\ell$. If any one of these averages is $\leqq \varepsilon_\ell +$ II's max-min value in G, then we define $f(x_1, \ldots, x_k) = c_{k+1}$; otherwise, we define $f(x_1, \ldots x_k) = \bar{p}$. That means that Player I brings the offending Player II to within ε_ℓ of his max-min payoff and then returns to cooperative play.[8]

Finally, consider the case in which ℓ and $k+1$ have opposite parity and ℓ is odd, i.e., Player I was the last to defect. In that case k is odd, $(k-1)/2$ rounds of G have been completed, Player I has already made his move in the $(k+1)/2$'th round, and it is now II's turn to move. Proceeding as before, we consider I's average payoff as measured at the end of each of the periods $(\ell+1)/2, \ldots, (k-1)/2$. If any of these averages was $\leqq \varepsilon_\ell +$ the max-min value to I in G, then we define $f(x_1, \ldots, x_k) = c_{k+1}$; otherwise, we define $f(x_1, \ldots, x_k) = \bar{q}(x_k)$. As before, that means that II brings the offending Player I to within ε_ℓ of his max-min value, then returns to cooperative play. (The difference is only that, because of the asymmetry in G, II's punishing move must depend on I's move in the same "round.")

This completes the formal description of the PEP that we described informally before; the reader should be able to convince himself that it is in equilibrium, is perfect, and yields the cooperative sequence $(c_1, c_2, \ldots, c_k, \ldots)$ with limiting average payoff h.

5 Discounted Payoffs in Repeated Games: Discussion of an Example

Thus far, we have been considering only the limiting average form of payoff for repeated games, corresponding intuitively to a future discount rate of zero. We shall now try to give an idea of how positive discount

8. Note that he never has occasion to look back beyond the most recent defection; the ε_ℓ level of punishment suffices for all past transgressions.

rates can affect the behavior of EPs and PEPs by studying an apparently simple but surprisingly revealing example.

Consider the following payoff matrix for G in strategic form, the players moving simultaneously:

		Player II L	Player II R
Player I	T	0, 1	$-p, -c+1$
	B	1, 0	$-p+1, -c$

Here, p and c are positive numbers ("punishment" and "cost"); we may think of them as being rather large. Thus, II may be in a position to damage I severely, but only at a cost to himself that may perhaps be unacceptable.[9]

In the repeated G^*, we shall use the discounted payoffs

$$\sum_{t=1}^{\infty} \alpha^t h_t^I \text{ and } \sum_{t=1}^{\infty} \beta^t h_t^{II}$$

to Players I and II, respectively, where $0 < \alpha < 1$ and $0 < \beta < 1$. Sometimes we shall further assume that $\alpha \leqslant \beta$, i.e., that Player I has, if anything, a bigger discount rate (= shorter "utility horizon") than Player II.

As is easily seen, the one-shot game G has a unique EP, namely (B, L), which yields the payoff (1, 0). This means that the strategy-pair in which I always plays B and II always plays L (regardless of history) is a perfect EP of G^*, since obviously no defection, even in a subgame, can ever be profitable.

Player II, however, would naturally prefer the outcome (0, 1), corresponding to (T, R). We shall now investigate under what conditions this outcome can be sustained by an EP, or by a PEP, in the discounted repeated game. Indeed, we shall find that it can be sustained by an EP if and only if $p \geqq 1/\alpha$; and, when $\alpha \geqq \beta$, that it can be sustained by a PEP if and only if $p \geqq 1/\alpha$ and $p/c \geqq (1-\beta)/\alpha\beta$. Thus, whereas the existence of an EP is independent of the cost of the punishment to the punisher (the parameter c), the existence of a PEP is not.

Let us first consider the EP question. We claim that the following "grim" strategy-pair:

$$\begin{cases} \text{I} & \text{plays } T \text{ always} \\ \text{II} & \text{plays } L \text{ so long as I plays } T\text{, but plays } R \text{ forever} \\ & \text{if I ever plays } B \end{cases}$$

9. One could think of I and II as North Vietnam and the United States in the 1960s.

is an EP of G^*, provided that $p \geqq 1/\alpha$. Moreover, we claim that if $p < 1/\alpha$ there is no EP sustaining (1, 1). Note first that II will certainly not wish to defect, as he cannot possibly improve on the sequence of payoffs (1, 1, 1, ...). On the other hand, if I wishes to improve on his sequence (0, 0, 0, ...), his best chance is to deviate to B at some time t_0, then keep playing B forever. This yields him the payoff sequence $(0, \ldots, 0, 1, -p+1, \ldots)$, and this is worth

$$\alpha^{t_0}\left(\frac{1}{1-\alpha} - \frac{p\alpha}{1-\alpha}\right).$$

Since this is profitable to him *if and only if* $1 > p\alpha$ the truth of our claims is now evident.

For a numerical example, let $p = 2$. Then if $\alpha < .5$ there will be no EP sustaining the (0, 1) outcome, as the rewards for defecting will outweigh any possible punishment. But if $\alpha \leqslant .5$, the strategy pair given above is clearly an EP.

Nothing in this result depends on the values of c or β. Yet, intuitively, one feels that the credibility of II's "threat," with which he extracts such a favorable outcome, ought to be *very* dependent on its cost. Our next object will be to show that a PEP that sustains the (0, 1) outcome is not possible for large values of c.

First let us give an example of such a PEP. It happens that we can define it in a very simple way, making the instructions to the two players almost independent of the history:

1. In the first round, play (T, L).
2. If the choices in round $t-1$ were (T, L) or (B, R), play (T, L) in round t.
3. If the choices in round $t-1$ were (T, R) or (B, L), play (B, R) in round t.

The cooperative sequence resulting from this strategy pair is just a repetition of (T, L); this is worth $\beta/(1-\beta)$ to II and 0 to I. In checking for the PEP property, it is sufficient to look merely at deviations that occur in the first round of a typical subgame G_t^*. Suppose Player I defects when he is supposed to play T. His best possible payoff sequence from then on is $(1, -p+1, -p+1, \ldots)$, which is worth

$$\alpha^t\left(\frac{\alpha}{1-\alpha} - \frac{\alpha^2 p}{1-\alpha}\right)$$

to him. So if $p \geqq 1/\alpha$ he will not have any incentive to defect. Player II likewise will not defect when he is supposed to play L, as he cannot possibly improve on the payoff sequence $(1, 1, 1, \ldots)$. When I is supposed to play B, a defection could yield him at best the sequence $(-p, -p + 1, 0, 0, \ldots)$; this is clearly inferior for all $p \geqq 0$ to the prescribed sequence $(-p + 1, 0, 0 \ldots)$. Finally, when II is called upon to play R at the beginning of G_t^*, he will have to compare his prescribed payoff sequence $(-c, 1, 1, 1, \ldots)$ with sequences like $(0, -c, 1, 1, 1, \ldots)$, $(0, 0, -c, 1, 1, 1, \ldots)$, etc., which he can obtain by defecting for 1,2, etc. rounds, or even the sequence $(0, 0, 0, \ldots)$ which he can obtain by perpetually defecting. In the discounted sum, these "heresies" are worth

$$\beta^k(-c\beta + \beta^2/(1 - \beta)), k = t + 1, t + 2, \ldots,$$

or 0, while "orthodoxy" is worth

$$\beta^t(-c\beta + \beta^2/(1 - \beta)).$$

So if $c \leqq \beta(1 - \beta)$, II cannot gain by defecting. The given strategy pair is therefore a PEP on the assumptions that $p \geqq 1/\alpha$ and $c \leqq \beta/(1 - \beta)$, as diagrammed in Figure 2.

To wrap up our example, it is necessary to show that there are significant cases, where, because of the positive discount rate, a PEP does *not* exist. Showing nonexistence is a more difficult undertaking, because in general a PEP can be a very complex thing. In particular, while pure G-strategies have sufficed up to now, we cannot ignore the possible use of mixed strategies against a defection. In our example, if p and c are both large numbers, the threat of a small probability of using R may be

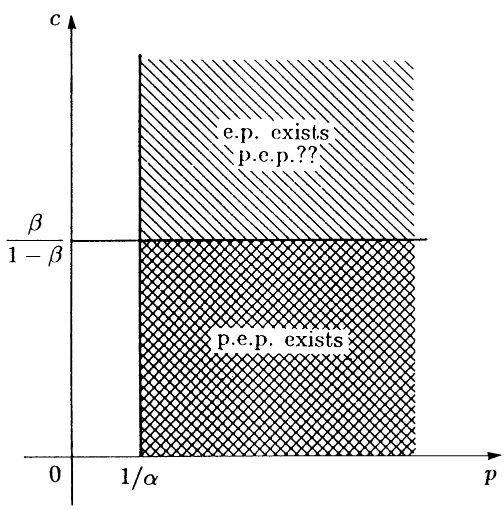

Figure 2

enough to keep I in line while holding the (expected) cost to a level that II can accept.[10] However, it would be out of place in this discussion to develop the elaborate technical apparatus of mixed strategies just for the sake of one example, whose purpose is only illustrative. Instead, we shall adopt a far simpler expedient, called "convexification in pure strategies," which is more or less equivalent to the introduction of mixed strategies.

In our example, this convexification merely means allowing Player II the option of "scaling down" his punishment by giving him a continuum of strategies in G, as follows:

		Player II	
		L	$R\lambda$ $(0 < \lambda \leqq 1)$
Player I	T	0, 1	$-\lambda p,\ -\lambda c + 1$
	B	1, 0	$-\lambda p + 1, -\lambda c$

Here, $\lambda = 1$ corresponds to the old R and $\lambda = 0$ corresponds to the old L. (However, we still indicate the latter choice by a separate column in the matrix.) Playing $R\lambda$ has much the same effect as playing a mixed strategy $\{R$ with probability λ, L with probability $1 - \lambda\}$, and it can be shown (though we shall not do it here) that if the new G^* has no PEP in *pure* strategies that sustains the cooperative sequence $((T, L), (T, L), \ldots)$, then the original G^* (with the same values of α, β, p, c) has no PEP in pure or mixed strategies that sustains that sequence.

Consider now a play of the revised G^*, with II making the sequence of moves $\Lambda = (\lambda_1, \lambda_2, \ldots, \lambda_t, \ldots)$. The total punishment received by I is then given by $P = P(\Lambda) = \sum_{t=1}^{\infty} \lambda_t \alpha^t p$, and the total cost incurred by II is $C = C(\Lambda) = \sum_{t=1}^{\infty} \lambda_t \beta^t c$. We now bring in the assumption, not used until now, that $\alpha \leqq \beta$. This implies that $\sum_{t=1}^{\infty} \lambda_t (\beta^t - \alpha^t) \geqq 0$, so that $P/C \leqq p/c$. This inequality shows that it is most efficient, in terms of the damage/cost ratio, for II to punish *immediately*; he thereby minimizes his cost for a given level of deterrence. It follows that the game has a PEP, of the type described above, whenever there is any number λ such that λp and λc satisfy the inequalities

$$\lambda p \geqq 1/\alpha, \quad \lambda c \leqq \beta/(1 - \beta).$$

This is illustrated in Fig. 3. As we can see, there is a critical ratio of c to p, namely

10. This is a realistic consideration for the world of nuclear politics and arms races, where the pressure of the nuclear deterrent is felt in every situation that creates any perceptible risk that the situation might escalate out of control.

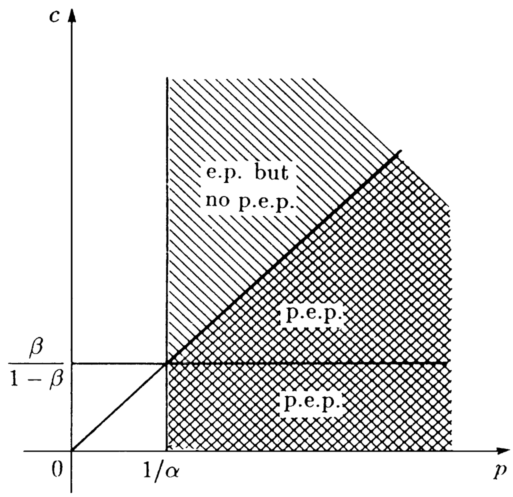

Figure 3

$$R_0 = \frac{\alpha\beta}{1-\beta},$$

above which no such λ can be found.

For a numerical example, if $\alpha = \beta = .75$ then $R_0 = 2.25$. If $p = 100$ and $c = 200$ then $c/p < R_0$, and we may, for example, choose $\lambda = .014$, giving us $\lambda c = 2.8 < \beta/(1-\beta) = 3$ and $\lambda p = 1.4 > 1/\alpha = 1.33$. So a 1.4 percent chance of II using his threat strategy R after a defection by I sustains the perfect equilibrium at (0, 1).

We can also make the converse argument. If (c, p) is *not* in the cross-hatched region indicated in Fig. 3, then there is no way for II to inflict any given amount P of punishment without incurring a cost of more than R_0P. By the foregoing, it is clear that this is above the cost that he can "afford"; in other words, he would prefer to accept his max-min payoff of 0 forever, rather than carry out the requisite threat. So a PEP cannot exist.

References

1. R. J. Aumann, "Acceptable Points in General Cooperative n-Person Games," in *Contributions to the Theory of Games IV*, Annals of Math. Study 40, Princeton University Press, 1959, pp. 287–324 [Chapter 20].

2. R. Selten, "Spieltheoretische behandlung eines oligopolmodells mit nachfrageträgheit," *Zeitschrift für die gesamte Staatswissenschaft* 12 (1965) 301–324.

3. R. Selten, "Re-examination of the perfectness concept for equilibrium points in extensive games," *International J. Game Theory* 4 (1975) 25–55.

Note added in proof Professor Jacek Krawczyk of the Victoria University of Wellington has pointed out an error on Page 405. If Player I defects when he is supposed to play T, his best possible payoff sequence from then on is $(1, -p+1, 1, -p+1, 1, \ldots)$, and not as written there. The given strategy pair is therefore a PEP on the assumptions that $p \geqslant 1 + (1/\alpha)$ and $c \leqslant \beta/(1-\beta)$, and not as written on Page 406. As a result, the neat picture presented in Figure 3 is called into question. Specifically, an EP certainly exists in the strip between the vertical lines $p = 1/\alpha$ and $p = 1 + (1/\alpha)$, but it is unclear whether a PEP exists anywhere in this strip. The exact placement of the diagonal that divides EP's from PEP's becomes similarly unclear. Nevertheless, it appears that our basic point remains valid: that for the existence of PEP's, the cost-to-punishment ratio is the determining factor, whereas for EP's, only the magnitude of the punishment matters—the cost of the punishment to the punisher is irrelevant.