Perfect Equilibrium in a Bargaining Model
Author(s): Ariel Rubinstein
Source: *Econometrica,* Vol. 50, No. 1 (Jan., 1982), pp. 97-109
Published by: The Econometric Society
Stable URL: http://www.jstor.org/stable/1912531
Accessed: 07/12/2010 11:13

PERFECT EQUILIBRIUM IN A BARGAINING MODEL

BY ARIEL RUBINSTEIN[1]

Two players have to reach an agreement on the partition of a pie of size 1. Each has to make in turn, a proposal as to how it should be divided. After one player has made an offer, the other must decide either to accept it, or to reject it and continue the bargaining. Several properties which the players' preferences possess are assumed. The Perfect Equilibrium Partitions (P.E.P.) are characterized in all the models satisfying these assumptions.

Specially, it is proved that when every player bears a fixed bargaining cost for each period (c_1 and c_2), then: (i) if $c_1 < c_2$ the only P.E.P. gives all the pie to 1; (ii) if $c_1 > c_2$ the only P.E.P. gives to 1 only c_2.

In the case where each player has a fixed discounting factor (δ_1 and δ_2) the only P.E.P. is $(1 - \delta_2)/(1 - \delta_1\delta_2)$.

1. INTRODUCTION

WHEN I REFER IN THIS PAPER to the *Bargaining Problem* I mean the following situation and question:

> Two individuals have before them several possible contractual agreements. Both have interests in reaching agreement but their interests are not entirely identical. What "will be" the agreed contract, assuming that both parties behave rationally?

I begin with this clarification because I would like to prevent the common confusion of the above problem with two other problems that may be asked about the bargaining situation, namely: (i) the positive question — what is the agreement reached in practice; (ii) the normative question — what is the just agreement.

Edgeworth [4] presented this problem one hundred years ago, considering it the most fundamental problem in Economics. Since then it seems to have been the source of considerable frustration for Economic theorists. Economists often talk in the following vein (beginning of Cross [3]):

> "Economists traditionally have had very little to say about pure bargaining situations in which the outcome is clearly dependent upon interactions among only a few individuals" (p. 67).

The "very little" referred to above is that the agreed contract is individual-rational and is Pareto otpimal; i.e. it is no worse than disagreement, and there is no agreement which both would prefer. However, which of the (usually numerous) contracts satisfying these conditions will be agreed? Economists tend to answer vaguely by saying that this depends on the "bargaining ability" of the parties.

Many attempts have been made in order to get to a clear cut answer to the bargaining problem. Two approaches may be distinguished in the published literature. The first is the strategic approach. The players' negotiating maneuvers are moves in a noncooperative game and the rationality assumption is expressed by investigation of the Nash equilibria. The second approach is the axiomatic method.

> "One states as axioms several properties that it would seem natural for the solution to have and then one discovers that the axioms actually determine the solution uniquely" [11, p. 129].

(For a survey of the axiomatic models of bargaining, see Roth [13].) The purpose of this approach is to bypass the difficulties inherent in the strategic approach. We make assumptions about the solution without specifying the bargaining process itself. Notice that in order to be relevant to our problem, these axioms may only either restrict the domain of the solution or be obtained from the assumption of rationality. Thus, for example, Nash's symmetry axiom can be considered as an assumption that all the differences between the players can be expressed in the set of utility pairs arising from the possible contracts and that there is no other relevant element that distinguishes between them. But, the key axiom in most axiomatizations — the "Independence of Irrelevant Alternatives" has not received a proper defense and in fact it is more suited to the normative question (see Luce and Raiffa [9] and Binmore [2]).

It was Nash himself who felt the need to complement the axiomatic approach (see [10]) with a non-cooperative game. (For a wider discussion, see Binmore [2].) In his second paper on the solution that he proposed [11], Nash proved that the solution is the limit of a sequence of equilibria of bargaining games. These models, however, are highly stylized and artificial. Among the later works, I mention here three, wherein the bargaining is represented by a multi-stage game. Ståhl [19,20] and Krelle [7] assume the existence of a known finite number of bargaining periods and their solutions are based on dynamic programming. Rice [12] uses the notion of a differential game. The bargaining period is identified with an interval, equilibrium strategies are the limits of "step-wise" strategies, and the lengths of those steps tend to zero.

In this paper I will adopt the strategic approach. I will consider the following bargaining situation: two players have to reach an agreement on the partition of a pie of size 1. Each has to make in turn, a proposal as to how it should be divided. After one party has made such an offer, the other must decide either to accept it or to reject it and continue with the bargaining. The players' preference relations are defined on the set of ordered pairs of the type (x,t) (where $0 \leqq x \leqq 1$ and t is a nonnegative integer). The pair (x,t) is interpreted as "1 receives x and 2 receives $1 - x$ at time t."

This paper is limited to the investigation of a family of models in which the preferences satisfy:

(A-1)   'pie' is desirable,

(A-2)   'time' is valuable,

(A-3)   continuity,

(A-4)   stationarity (the preference of (x, t) over $(y, t + 1)$ is independent of t),

(A-5)   the larger the portion the more 'compensation' a player needs for a delay of one period to be immaterial to him.

The two elements in which the parties may differ are the negotiating order (who has "first turn") and the preferences.

Two sub-families of models to which I will refer, are:

(i) *Fixed bargaining cost*: i's preference is derived from the function $y - c_i t$, i.e. every player bears a fixed cost for each period.

(ii) *Fixed discounting factor*: i's preference is derived from the function $y \cdot \delta_i^t$, i.e. every player has a fixed discounting factor.

So my first step has been to restrict the bargaining situation to be considered. Secondly, I will give a severe interpretation to the rationality requirement by investigating *perfect* equilibria (see Selten [17, 18]). A perfect equilibrium is one where not only the strategies chosen at the beginning of the game form an equilibrium, but also the strategies planned after all possible histories (in every subgame).

Quite surprisingly[2] this leads to the isolation of a single solution for most of the cases examined here. For example, in the fixed bargaining cost model, it turns out that if $c_1 > c_2$, 1 receives c_2 only. If $c_1 < c_2$, 1 receives all the pie. If $c_1 = c_2$, any partition of the pie from which 1 receives at least c_1 is a perfect equilibrium partition (P.E.P.). In other words, a weaker player gets almost 'nothing'; he can at most get the loss which his opponent incurs during one bargaining round. In the fixed discounting factor model there is one P.E.P., 1 obtaining $(1 - \delta_2)/(1 - \delta_1 \delta_2)$. This solution is continuous, monotonic in the discounting factors, and gives relative advantage to the player who starts the bargaining.

The work closest to that appearing here, is that of Ingolf Ståhl[3] [19, 20]. He investigates a similar bargaining situation but which has a finite and known negotiating time horizon, and in which the pie can be only partitioned discretely. Ståhl studies cases for which there exists a single P.E.P. which is independent of who has the first move.

The discussed bargaining model may be modified in numerous ways, many being only technical modifications. However I would like to point out one type of modification which I believe to be extremely interesting. A critical assumption in the model is that each player has complete information about the preference of the other. Assume on the other hand that 1 and 2 both know that 1 has a fixed bargaining cost. They both know that 2 has also a fixed bargaining cost, but only

---

[2] Especially considering that the perfect equilibrium concept has been "disappointing" when applied to the supergames, see Aumann and Shapley [1], Kurz [8], and Rubinstein [14, 15, 16].

[3] I would like to thank Professor R. Selten for referring me to Ståhl's work, after reading the first version of this paper.

2 knows its actual value. In such a situation some new aspects appear. 1 will try to conclude from 2's behavior what the true bargaining cost is, and 2 may try to cheat 1 by leading him to believe that he, 2, is "stronger" than he actually is. In such a situation one can expect that the bargaining will continue for more than one period. I hope to deal with this situation in another paper.

2. THE BARGAINING MODEL

Two players, 1 and 2, are bargaining on the partition of a pie. The pie will be partitioned only after the players reach an agreement. Each player, in turn offers a partition and his opponent may agree to the offer "Y" or reject it "N". Acceptance of the offer ends the bargaining. After rejection, the rejecting player then has to make a counter offer and so on. There are no rules which bind the players to any previous offers they have made.

Formally, let $S = [0, 1]$. A partition of the pie is identified with a number s in the unit interval by interpreting s as the proportion of the pie that 1 receives. Let s_i be the portion of the pie that player i receives in the partition s: that is $s_1 = s$ and $s_2 = 1 - s$.

Let F be the set of all sequences of functions $f = \{f^t\}_{t=1}^{\infty}$, where $f^1 \in S$, for t odd $f^t : S^{t-1} \to S$, and for t even $f^t : S^t \to \{Y, N\}$. (S^t is the set of all sequences of length t of elements in S.) F is the set of all strategies of the player who starts the bargaining. Similarly let G be the set of all strategies of the player who in the first move has to respond to the other player's offer; that is, G is the set of all sequences of functions $g = \{g^t\}_{t=1}^{\infty}$ such that, for t odd $g^t : S^t \to \{Y, N\}$ and for t even $g^t : S^{t-1} \to S$.

The following concepts are easily defined rigorously. Let $\sigma(f, g)$ be the sequence of offers in which 1 starts the bargaining and adopts $f \in F$, and 2 adopts $g \in G$. Let $T(f, g)$ be the length of $\sigma(f, g)$ (may be ∞). Let $D(f, g)$ be the last element of $\sigma(f, g)$ (if there is such an element). $D(f, g)$ is called the *partition* induced by (f, g). The outcome function of the game is defined by

$$P(f, g) = \begin{cases} (D(f, g), T(f, g)), & T(f, g) < \infty, \\ (0, \infty), & T(f, g) = \infty. \end{cases}$$

Thus, the outcome (s, t) is interpreted as the reaching of agreement s in period t, and the symbol $(0, \infty)$ indicates a perpetual disagreement.

For the analysis of the game we will have to consider the case in which the order of bargaining is revised and player 2 is the first to move. In this case a strategy for player 2 is an element of F and a strategy for player 1 is an element of G. Let us define $\sigma(g, f)$, $T(g, f)$, $D(g, f)$ and $P(g, f)$ similarly to the above for the case where player 2 starts the bargaining and adopts $f \in F$ and player 1 adopts $g \in G$.

The last component of the model is the preference of the players on the set of outcomes. I assume that player i has a preference relation (complete, reflexive, and transitive) $\succsim_i$ on the set of $S \times N \cup \{(0, \infty)\}$, where N is the set of natural numbers.

I assume that the preferences satisfy the following five assertions:

For all $r, s \in S, t, t_1, t_2 \in N$, and $i \in \{1, 2\}$:

(A-1)   if $r_i > s_i$, then $(r, t) >_i (s, t)$;

(A-2)   if $s_i > 0$ and $t_2 > t_1$, then $(s, t_1) >_i (s, t_2) >_i (0, \infty)$;

(A-3) $(r, t_1) \gtrsim_i (s, t_1 + 1)$ iff $(r, t_2) \gtrsim_i (s, t_2 + 1)$;

(A-4)   if $r_n \to r$ and $(r_n, t_1) \gtrsim_i (s, t_2)$, then $(r, t_1) \gtrsim_i (s, t_2)$;

         if $r_n \to r$ and $(r_n, t_1) \gtrsim_i (0, \infty)$, then $(r, t_1) \gtrsim_i (0, \infty)$;

(A-5)   if $(s + \epsilon, 1) \sim_i (s, 0), (\bar{s} + \bar{\epsilon}, 1) \sim_i (\bar{s}, 0)$, and $s_i < \bar{s}_i$, then $\epsilon_i \leq \bar{\epsilon}_i$.

From (A-3) we can use the notation $(r, T) \gtrsim_i (s, 0)$ and $(r, T) \lesssim_i (s, 0)$ for $(r, T + t) \gtrsim_i (s, t)$ and $(r, T + t) \lesssim_i (s, t)$, respectively.

Two families of models in which the preferences satisfy the above conditions are:

I. *Fixed bargaining costs.* Each player i has a number c_i such that $(s, t_1) \gtrsim (\bar{s}, t_2)$ iff $(s_i - c_i \cdot t_1) \geq (\bar{s}_i - c_i \cdot t_2)$.

II. *Fixed discounting factors.* Each player i has a number $0 < \delta_i \leq 1$ such that $(s, t_1) \gtrsim_i (\bar{s}, t_2)$ iff $s_i \delta_i^{t_1} \geq \bar{s}_i \delta_i^{t_2}$.

We reserve $\delta_i = 0$ for the lexicographic preference: $(s, t_1) \gtrsim_i (\bar{s}, t_2)$ if $(t_1 < t_2)$ or $(t_1 = t_2$ and $s_i \geq \bar{s}_i)$.

REMARK: In a more general framework of the model, player i would be characterized by the sequence of preferences $\{\gtrsim_i^t\}$, where $\gtrsim_i^t$ is i's preference on the outcomes assuming that the players have not reached an agreement in the first $t - 1$ periods. In fact, I assume that $\gtrsim_i^t \equiv \gtrsim_i$. This assumption precludes discussion of some interesting bargaining situation such as: (1) player i has a sequence $\{c_i^t\}$ where c_i^t is the cost to i of the bargaining in period t; (2) player i has a fixed bargaining cost and his utility is not linear.

3. PERFECT EQUILIBRIUM

The ordered pair $(\hat{f}, \hat{g}) \in F \times G$ is called a *Nash Equilibrium* if there is no $f \in F$ such that $P(f, \hat{g}) >_1 P(\hat{f}, \hat{g})$ and there is no $g \in G$ such that $P(\hat{f}, g) >_2 P(\hat{f}, \hat{g})$.

The following simple proposition indicates that even after the restriction of the bargaining problem to our model, the Nash equilibrium is a "weak" concept.

PROPOSITION: *For all $s \in S$, s is a partition induced by Nash equilibrium.*

PROOF: Let us define $\hat{f} \in F$ and $\hat{g} \in G$ as follows:

$$\text{for } t \text{ odd,} \quad \hat{f}^t \equiv s, \quad \hat{g}^t(s^1 \ldots s^t) = \begin{cases} Y, & s^t \leq s, \\ N, & s^t > s; \end{cases}$$

$$\text{for } t \text{ even,} \quad \hat{g}^t \equiv s, \quad \hat{f}^t(s^1 \ldots s^t) = \begin{cases} Y, & s^t \geq s, \\ N, & s^t < s. \end{cases}$$

Clearly, $(\hat{f}, \hat{g})$ is a Nash equilibrium and $P(\hat{f}, \hat{g}) = (s, 1)$.

The above equilibrium highlights the inadequacy of the concept of a Nash equilibrium in the current context. Assume 1 demands $s + \epsilon (\epsilon > 0)$. At this point of the game, 2 intends to insist on the original planned contract and 1 intends to agree to this offer. But if ϵ is sufficiently small so that $(s, 1) <_2 (s + \epsilon, 0)$, 2 will prefer to agree to player 1's deviation. Thus, player 1 may carry out a manipulative maneuver and offer $s + \epsilon$ in the certainty that 2 will agree to it.

In order to overcome this difficulty (see also Harsanyi [5]) I will use the concept of the Perfect Equilibrium following the definition of Selten (see [17, 18]). For ·this definition we need some additional notation. Let $s^1 \ldots s^T \in S$. Define $f | s^1 \ldots s^T$ and $g | s^1 \ldots s^T$ as the strategies derived from f and g after the offers $s^1 \ldots s^T$ have been announced and already rejected. (That is, for T odd and t odd,

$$(f | s^1 \ldots s^T)^t (r^1 \ldots r^{t-1}) = f^{T+t}(s^1 \ldots s, ^T r^1 \ldots r^{t-1}),$$

$$(g | s^1 \ldots s^T)^t (r^1 \ldots r^t) = g^{T+t}(s^1 \ldots s, ^T r^1 \ldots r^t),$$

and so on.)

Notice that if T is even it is 1's turn to propose a partition of the pie, and 2's first move is a response to 1's offer. Thus $f | s^1 \ldots s^T \in F$ and $g | s^1 \ldots s^T \in G$. If T is odd, it is 2's turn to make an offer and therefore $g | s^1 \ldots s^T \in F$ and $f | s^1 \ldots s^T \in G$. And now to the central definition which, as mentioned, follows Selten's definition of subgame perfectness [17, 18] (what may at first seem to be a slightly clumsy version of Selten's idea has been chosen to prevent the use of some additional notation which would be redundant in this paper):

DEFINITION: $(\hat{f}, \hat{g})$ is *Perfect Equilibrium* (P.E.) if for all $s^1 \ldots s^T$, if T is odd:

(P-1)   there is no $f \in F$ such that $P(\hat{f} | s^1 \ldots s^T, f) >_2 P(\hat{f} | s^1 \ldots s^T, \hat{g} | s^1 \ldots s^T)$;

(P-2)   if $\hat{g}^T(s^1 \ldots s^T) = Y$, there is no $f \in F$ such that $P(\hat{f} | s^1 \ldots s^T, f) >_2 (s^T, 0)$;

(P-3)   if $\hat{g}^T(s^1 \ldots s^T) = N, P(\hat{f} | s^1 \ldots s^T, \hat{g} | s^1 \ldots s^T) \gtrsim_2 (s^T, 0)$;

and if T is even:

(P-4)   there is no $f \in F$ such that $P(f, \hat{g} | s^1 \ldots s^T) >_1 P(\hat{f} | s^1 \ldots s^T, \hat{g} | s^1 \ldots s^T)$;

(P-5)   if $\hat{f}^T(s^1 \ldots s^T) = Y$, there is no $f \in F$ such that $P(f, \hat{g} | s^1 \ldots s^T) >_1 (s^T, 0)$;

(P-6)   if $\hat{f}^T(s^1 \ldots s^T) = N, P(\hat{f} | s^1 \ldots s^T, \hat{g} | s^1 \ldots s^T) \gtrsim_1 (s^T, 0)$.

(P-1) and (P-4) ensure that after a sequence of offers and rejections $s^1 \ldots s^T$ the player who has to continue the bargaining has no better strategy other than to follow the planned strategy. (P-2) and (P-5) ensure that a player who has planned to accept the offer s^T has no better alternative than to accept it, and (P-3) and (P-6) ensure that if a player is expected to reject an offer, it is not better for him to accept the offer.

EXAMPLE: To clarify the notation, let us show that the pair $(\hat{f}, \hat{g})$ (with $s = 0.5$) described in the proof of the above proposition is not a perfect equilibrium for fixed bargaining cost preferences with $c_1 = 0.1$ and $c_2 = 0.2$. Player 2 plans to reject a possible offer of 0.6 by player 1: that is, $\hat{g}^1(0.6) = N$. After such a rejection the players expect to agree on 0.5: that is, $P(\hat{f}|0.6, \hat{g}|0.6) = (0.5, 1)$. Player 2 prefers $(0.6, 0)$ to $(0.5, 1)$: thus, $(\hat{f}, \hat{g})$ violates condition (P-3).

REMARK: Notice that a strategy has been defined in Section 2 as a sequence of functions which is interpreted as the player's plans after every history, including histories which are not consistent with his own plans. For example, $f^3(s^1, s^2)$ is required to be defined even where $f^1 \neq s^1$ and $f^2(s^1, s^2) = Y$. The reader is directed to Selten [17, 18] and Harsanyi [5, 6] for details on the significance of the requirement.

<div style="text-align:center">4. LEMMAS</div>

In this section we have only to assume that the preferences satisfy (A-1) and (A-2). The following Lemmas establish connections between two sets: (A) the set of all P.E.P.'s in a game in which 1 starts the bargaining, that is, $\{s \in S \mid$ there is a P.E. $(f, g) \in F \times G$ such that $s = D(f, g)\}$; and (B) the set of all P.E.P.'s in a game in which 2 starts the bargaining, that is, $\{s \in S \mid$ there is a P.E. $(g, f) \in G \times F$ such that $s = D(g, f)\}$.

REMARK: In a generalized model in which the $\succsim_i^t$ are not identical the same considerations would be used to establish connections between the sets $\{A^t\}_{t=1,3,5\ldots}$ and $\{B^t\}_{t=2,4,6,\ldots}$ where $A^t(B^t)$ is the set of all P.E.P.'s in a game which starts at time t, 1(2) making the first offer.

LEMMA 1: Let $a \in A$. For all $b \in S$ such that $b > a$, there is $c \in B$ such that $(c, 1) \succsim_2 (b, 0)$.

REMARK: Lemma 1 states that for a to be in A, it has to be "protected" from the possibility that 1 will demand and achieve some better contract. Player 1 will certainly do so if there is $b \in S$ satisfying $b > a$ such that 2 would accept b if it were offered. Player 2 must therefore reject such an offer. In order that it be optimal for him to carry out this threat, player 2 has to expect to achieve a better partition in the future; that is, there must be a P.E.P. $c \in B$ in the subgame that takes place after 2's rejection such that $(c, 1)$ is preferred by 2 to $(b, 0)$.

PROOF: Let $(\hat{f}, \hat{g})$ be a P.E. such that $D(\hat{f}, \hat{g}) = a$. Let $b \in S$ and $b > a$. From (P-1), $\hat{g}^1(b) = N$ (otherwise if $f^1 = b$ then $P(f, g) = (b, 1) >_1 (a, 1) \succsim_1 (a, T(\hat{f}, \hat{g})) = P(\hat{f}, \hat{g})$ in contradiction to (P-1)). From (P-3) $P(\hat{f}|b, \hat{g}|b) \succsim_2 (b, 0)$ thus, $(D(\hat{f}|b, \hat{g}|b), T(\hat{f}|b, \hat{g}|b)) \succsim_2 (b, 0)$ and by (A-2) $(D(\hat{f}|b, \hat{g}|b), 1) \succsim_2 (b, 0)$ and therefore $D(\hat{f}|b, \hat{g}|b)$ is the desirable c.

Similarly, it is easy to prove the following lemma.

LEMMA 2: *For all $a \in B$ and for all $b \in S$ such that $b < a$, there is $c \in A$ such that $(c, 1) \gtrsim_1 (b, 0)$.*

LEMMA 3: *Let $a \in A$. Then for all b such that $(b, 1) >_2 (a, 0)$ there is $c \in A$ such that $(c, 1) \gtrsim_1 (b, 0)$.*

REMARK: Lemma 3 states that if a is a P.E.P. then 1 should have a "good reason" to reject any offer from 2 which is preferred by 2 to accepting 1's original offer. Assume that in a certain P.E., player 2 plans to agree to a in the first period (case B below). Consider b such that $(b, 1) >_2 (a, 0)$. Then, player 2 will reject a if he thinks that 1 would agree to b. Thus player 1 must threaten to reject any such offer b. In order that this threat be credible there must be a P.E. in the subgame beginning with 1's offer which yields an agreement c such that $(c, 1) \gtrsim_1 (b, 0)$. This c must be a member of A.

PROOF: Let $(\hat{f}, \hat{g})$ be a P.E. such that $D(\hat{f}, \hat{g}) = a$.

Case A: $\hat{g}^1(\hat{f}^1) = N$. Let $\hat{f}^1 = s$. Then $D(\hat{f}|s, \hat{g}|s) = a$ and $a \in B$. From (A-1) and (A-2), if $(b, 2) >_2 (a, 1)$ then $b < a$ and therefore from Lemma 2 there is $c \in A$ such that $(c, 1) \gtrsim_1 (b, 0)$.

Case B: $\hat{f}^1 = a$, $\hat{g}^1(a) = Y$. Let b satisfy $(b, 1) >_2 (a, 0)$, $\hat{f}^2(a, b) = N$, because otherwise, for any $f \in F$ satisfying $f^1 = b$, $P(\hat{f}|a, f) = (b, 1) >_2 (a, 0)$, in contradiction to (P-2). From (P-6), $P(\hat{f}|a, b, \hat{g}|a, b) \gtrsim_1 (b, 0)$. Thus $(D(\hat{f}|a, b, \hat{g}|a, b), 1) \gtrsim_1 (b, 0)$ and $D(\hat{f}|a, b, \hat{g}|a, b) \in A$.

In a similar way it is possible to prove the following lemma.

LEMMA 4: *For all $a \in B$ and for all $b \in S$ such that $(b, 1) >_1 (a, 0)$ there is $c \in B$ such that $(c, 1) \gtrsim_2 (b, 0)$.*

5. THE THEOREM

Let

$$\Delta = \left\{ (x, y) \in S \times S \,\middle|\, \begin{array}{l} y \text{ is the smallest number such that } (x, 1) \lesssim_1 (y, 0); \\ x \text{ is the largest number such that } (y, 1) \lesssim_2 (x, 0) \end{array} \right\},$$

$$\Delta_1 = \{ x \in S \,|\, \text{there is } y \in S \text{ such that } (x, y) \in \Delta \},$$

$$\Delta_2 = \{ y \in S \,|\, \text{there is } x \in S \text{ such that } (x, y) \in \Delta \}.$$

PROPOSITION 1: *If $(x, y) \in \Delta$, then $x \in A$ and $y \in B$.*

PROOF: Consider the following $(\hat{f}, \hat{g})$; for t odd

$$\hat{f}^t \equiv x, \qquad \hat{g}^t(s^1 \ldots s^t) = \begin{cases} N, & x < s^t, \\ Y, & s^t \leqq x, \end{cases}$$

and for t even

$$\hat{f}^t(s^1 \ldots s^t) = \begin{cases} N, & s^t < y, \\ Y, & y \leqq s^t, \end{cases} \qquad \hat{g}^t \equiv y.$$

It is easy to check that $(\hat{f}, \hat{g})$ is a perfect equilibrium.

PROPOSITION 2: $\Delta \neq \phi$ (and therefore A and B are not empty).

PROOF: Let

$$d_1(x) = \begin{cases} 0 \text{ if for all } y \ (y,0) >_1(x,1), \\ y \text{ if there exists } y, (y,0) \sim_1 (x,1), \end{cases}$$

and

$$d_2(y) = \begin{cases} 1 \text{ if for all } (x,0) >_2(y,1), \\ x \text{ if there exists } x, (x,0) \sim_2 (y,1). \end{cases}$$

$d_1(x)$ is the smallest y such that $(y,0) \gtrsim_1 (x,1)$ and $d_2(y)$ is the largest x such that $(x,0) \gtrsim_2 (y,1)$. Therefore

$$\Delta = \{(x,y) | y = d_1(x) \text{ and } x = d_2(y)\}.$$

It is easy to check that d_1 (and d_2) is well defined, continuous, increasing, and strictly increasing where $d_1(x) > 0$ $(d_2(y) < 1)$.

Let $D(x) = d_2(d_1(x))$. Thus, $\Delta = \{(x,y) \mid y = d_1(x)$ and $D(x) = x\}$. Notice that $D(1) \leqq 1$ and $D(0) \geqq 0$. From the continuity of D it follows that there exists x_0 such that $D(x_0) = x_0$. Thus, $(x_0, d_1(x_0)) \in \Delta$.
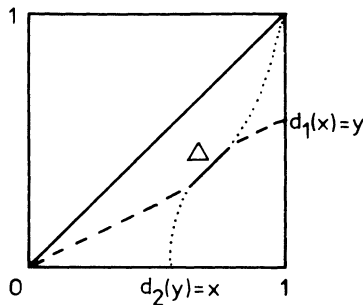

FIGURE 1

PROPOSITION 3: *The graph of Δ is a closed line segment which lies parallel to the diagonal $y = x$.*

PROOF: From the continuity of d_1 and d_2, the set Δ is closed. Notice that $x - d_1(x)$ is an increasing function. To see this, let x_0 satisfy $(0,0) \sim_1 (x_0, 1)$ (take $x_0 = 1$ if there is no x that satisfies $(0,0) \sim_1 (x, 1)$). For $x \leqq x_0, d_1(x) = 0$ and $x - d_1(x) = x$. For $x_1 > x_2 \geqq x_0$ $(d_1(x_1), 0) \sim_1 (x_1, 1)$ and $(d_1(x_2), 0) \sim_2 (x_2, 1)$. The function d_1 is an increasing function. Thus, (A-5) implies $x_1 - d_1(x_1) \geqq x_2 - d_1(x_2)$. Similarly, $d_2(y) - y$ is a decreasing function. We have to show that $x - y$ is constant for all $(x, y) \in \Delta$. Suppose that $x_2 < x_1$ and that (x_1, y_1) and (x_2, y_2) are both in Δ. Then $x_1 - d_1(x_1) \geqq x_2 - d_1(x_2)$ and $x_1 - y_1 \geqq x_2 - y_2$. Also $d_2(y_1) - y_1 \leqq d_2(y_2) - y_2$ and $x_1 - y_1 \leqq x_2 - y_2$. Thus $x_1 - y_1 = x_2 - y_2$.

PROPOSITION 4: *If $a \in A$, then $a \in \Delta_1$, and if $b \in B$, then $b \in \Delta_2$.*

PROOF: Suppose $\Delta_1 = [x_1, x_2]$ and $\Delta_2 = [y_1, y_2]$. Let $s = \sup\{a \in A\}$. Assume $x_2 < s$. Then $d_2(d_1(s)) < s$. Let $a \in A$ satisfy $r = d_2(d_1(s)) < a < s$. Let $b \in S$ satisfy $d_2^{-1}(a) > b > d_1(s)$. Then $a > d_2(b)$ and $(b, 1) >_2 (a, 0)$. From Lemma 3 there exists $c \in A$ such that $(c, 1) \gtrsim_1 (b, 0)$. Therefore there exists $c \in A$ satisfying $d_1(c) \geqq b$. The facts that d_1 is an increasing function and that $d_1(c) \geqq b > d_1(s)$ imply $c > s$ in contradiction to the definition of s.

Similarly, using Lemma 4 it is possible to show that $y_1 = \inf\{b \in B\}$. Using Lemmas 1 and 2 we get $x_1 = \inf\{a \in A\}$ and $y_2 = \sup\{b \in B\}$.

To summarize:

THEOREM: $A = \Delta_1 \neq \phi$, $B = \Delta_2 \neq \phi$. A *and* B *are closed intervals and there exists* $\epsilon \geqq 0$ *such that* $B = A - \epsilon$.
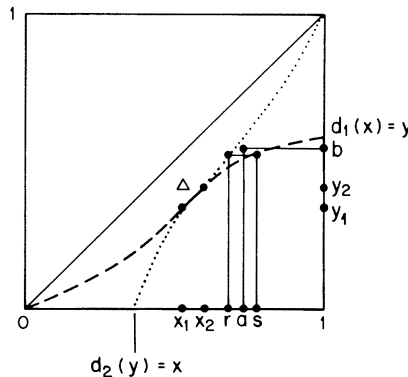

FIGURE 2

6. CONCLUSIONS

The following are applications of the theorem to the fixed bargaining cost and the fixed discounting factor models.

CONCLUSION 1: In the case where both the players have fixed bargaining costs, c_1 and c_2 (case I in the introduction):
  (1) If $c_1 > c_2, c_2$ is the only P.E.P.
  (2) If $c_1 = c_2$, every $c_1 \leqq x \leqq 1$ is a P.E.P.
  (3) If $c_1 < c_2$, 1 is the only P.E.P.

PROOF: $d_1(x) = \max\{x - c_1, 0\}$ and $d_2(y) = \min\{y + c_2, 1\}$. Thus Δ is the set of all solutions to the set of equations $y = \max\{x - c_1, 0\}$ and $x = \min\{y + c_2, 1\}$. The conclusion is implied by the three diagrams of Fig. 3 related to the cases (1) $c_1 > c_2$, (2) $c_1 = c_2$, and (3) $c_1 < c_2$.
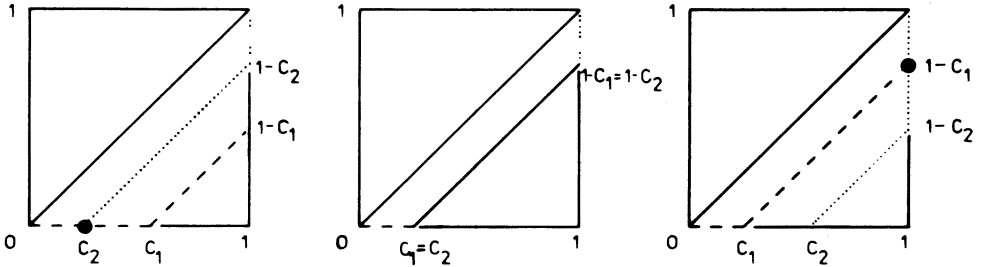

FIGURE 3

REMARK: Given a particular P.E., is agreement reached immediately after the very first offer? A positive answer results when $A \cap B = \phi$. If $(\hat{f}, \hat{g})$ is a P.E. and $T(\hat{f}, \hat{g}) > 1$, then $D(\hat{f}, \hat{g})$ is a member not only of A but also of B. Thus in almost all cases covered by Conclusions 1 and 2 (the possible exception being the case $c_1 = c_2$ in Conclusion 1) the bargaining indeed ends in the first period.

The following pair of strategies $(\hat{f}, \hat{g})$ is an example of a P.E. in the game where the players have fixed bargaining costs $c_1 = c_2 = c$. The pair $(\hat{f}, \hat{g})$ has the property that the negotiation ends at the second period: Let $\epsilon(x)$ be a non-negative function defined in the unit interval such that $\epsilon(x) \leqq \max\{0, x - c\}$. Assume $\epsilon(x)$ attains its maximum at x_0 where $x_0 > 2c$ and $\epsilon(x_0) > 2c$. Let $(\hat{f}, \hat{g})$ satisfy

$$\hat{f}^1 = x_0,$$

$$\hat{g}^1(s^1) = \begin{cases} N, & c < s^1, \\ Y, & s^1 \leqq c, \end{cases}$$

and "after," the strategies are identical to the strategies described in Proposition 1 for the partition $\epsilon(s^1)$. The partition of this P.E. is $\epsilon(x_0)$ (Figure 4).
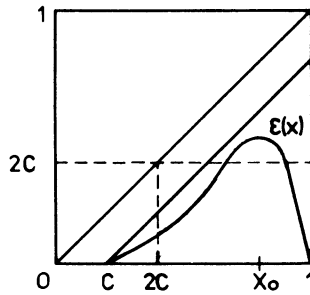
FIGURE 4

In this example the first move by player 1 serves as a signal to player 2. Player 2 interprets 1's signal s^1 as an agreement to continue with a pair of strategies that yields the partition $\epsilon(s^1)$. Not every s^1 may serve as such a signal, since 2 will agree to every partition that gives 2 more than $1 - c$. The partition $\epsilon(s^1)$ must give 2 at least $1 - s^1 + c$; therefore $\epsilon(s^1) \geqq s^1 - c$. A final restriction on ϵ is that $x_0 \geqq 2c$. Otherwise 1 would prefer to offer a partition that 2 "could not refuse" (some offer between c and $x_0 - c$). This also shows that the P.E. outcome may not be Pareto optimal; both players prefer to agree to $1 - \epsilon(x_0)$ at the beginning of the bargaining.

CONCLUSION 2: In the case where the players have fixed discounting factors — δ_1 and δ_2 (Case II in the introduction)—if at least one of the δ_i is strictly less than 1 and at least one of them is strictly positive, then the only P.E.P. is $M = (1 - \delta_2) / (1 - \delta_1\delta_2)$.

REMARK: Notice that when $\delta_2 = 0$, player 2 has no threat because the pie has no worth for him after the first period. Player 1 can exploit this to get all the pie ($M = 1$). When $\delta_1 = 0$, 1 can gain $1 - \delta_2$ only, that is, the proportion of the pie that 2 may lose if he refuses 1's offer and gets 1 in the second period. When $0 < \delta_1 = \delta_2 = \delta < 1$, 1 gets $1/1 + \delta > 1/2$. As one would expect, 1's gain from the fact that he starts the bargaining decreases as δ tends to 1.
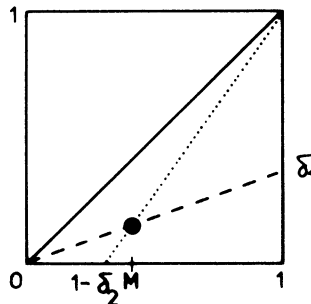

FIGURE 5

PROOF: $d_1(x) = x \cdot \delta_1$ and $d_2(y) = 1 - \delta_2 + \delta_2 \cdot y$. The conclusion follows from Figure 5 (the intersection of d_1 and d_2 is where $1 - \delta_2 + \delta_2 x \delta_1 = x$, that is where $x = M$).

The Hebrew University, Jerusalem

REFERENCES

[1] AUMANN, R. J., AND L. SHAPLEY: "Long Term Competition: A Game Theoretic Analysis," 1976, unpublished manuscript, Stanford University.

[2] BINMORE, K. G.: "Nash Bargaining Theory I," International Center for Economics and Related Disciplines, London School of Economics, Discussion Paper 80–09, 1980.

[3] CROSS, J. G.: "A Theory of the Bargaining Process," *American Economic Review*, 55(1965), 67–94.

[4] EDGEWORTH, F. Y.: "*Mathematical Psychics: An Essay on the Applications of Mathematics to the Moral Sciences,*" L.S.E. Series of Reprints of Scarce Tracts in Economics and Political Sciences, No. 10, 1932.

[5] HARSANYI, J. C.: "A Solution Theory for Non-cooperative Games and its Implications for Cooperative Games," Center for Research in Management Science, University of California, Berkeley, CP–401, 1978.

[6] HARSANYI, J. C., AND R. SELTEN: "A Generalized Nash Solution for Two Person Bargaining Games with Incomplete Information," *Management Science*, 18(1972), 80–106.

[7] KRELLE, W.: "A New Theory of Bargaining," unpublished manuscript, 1975.

[8] KURZ, M.: "Altruism as an Outcome of Social Interaction," *American Economic Review*, 68(1978), 216–222.

[9] LUCE, R. P., AND H. RAIFFA: *Games and Decisions*, New York: John Wiley and Sons, 1957.

[10] NASH, J. F.: "The Bargaining Problem," *Econometrica*, 18(1950), 155–162.

[11] ——: "Two-person Cooperative Games," *Econometrica*, 21(1953), 128–140.

[12] RICE, P.: "A Note on the Hicks Theory of Strike Bargaining," *Zeitschsrift fur die Gesamte Staatswissenschaft*, 133(1977), 236–244.

[13] ROTH, A. E.: *Axiomatic Models of Bargaining*, Lecture Notes in Economics and Mathematical Systems No. 170. Berlin: Springer-Verlag, 1979.

[14] RUBINSTEIN, A.: "Equilibrium in Supergames," Center for Research in Math. Economics and Game Theory, the Hebrew University, Jerusalem, Israel, R.M. 25, 1977.

[15] ——: "Equilibrium in Supergames with the Overtaking Criterion," *Journal of Economic Theory*, 21(1979), 1–9.

[16] ——: "Strong Perfect Equilibrium," *International Journal of Game Theory*, 9(1980), 1–12.

[17] SELTEN, R.: "Spieltheoretische Behandlung eines Oligopolmodels mit Nachfragetragheit," *Zeitschrift fur die Gesamte Staatswissenschaft*, 12(1965), 301–324 and 667–689.

[18] ——: "Re-examination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory*, 4(1975), 25–55.

[19] STÅHL, I.: *Bargaining Theory*. Stockholm: Stockholm School of Economics, 1972.

[20] ——: "An N-Person Bargaining Game in the Extensive Form," in *Mathematical Economics and Game Theory*, ed. by R. Henn and O. Moeschlin, Lecture Notes in Economics and Mathematical Systems No. 141. Berlin: Springer-Verlag, 1977.