Discussion Paper No. 178

Induced Outcomes in Cooperative Normal-Form Games

by

Robert W. Rosenthal*

Department of Industrial Engineering and Management Sciences

Northwestern University

Evanston, Illinois

November 1975

*This research was supported by National Science Foundation Grant #SOC-740179. I have benefitted from discussions with Ehud Kalai and Donald Saari.

Abstract

For cooperative n-person games in normal form, a new model of the preplay negotiation process is introduced. In this model the players successively commit themselves irrevocably, according to a specified exogenous ordering, to coalitional strategies conditionally on the rest of the players in the coalition agreeing to play their parts of the coalitional strategy. A solution concept, the induced outcome, is defined for this model; and sufficient conditions for its existence and uniqueness are given. A modification of this concept, the ϵ-induced outcome, is found to exist under very weak assumptions.

1. Introduction

The purpose of this paper is to introduce a new class of solution concepts for n-person cooperative games in normal form. Being defined for the normal form, the concepts may be used to examine games for which the characteristic-function form does not provide an adequate description. The main feature which distinguishes these concepts from most others, however, is that they utilize the idea that players may commit themselves irrevocably to strategies in the preplay negotiation process. Such tactics may be effective when they can be accomplished prior to agreements of other players and when they can be communicated to other players. Consider

Example 1: This is an economy with three individuals, one private good, and one public good. Each individual is endowed with one unit of the private good and a utility function $u^i(x^i, y) = x^i + 2y/3$, where x^i is i's consumption of the private good and y is the sum of the units of the public good produced. Each individual may devote any fraction of his endowment of the private good to produce the same fractional number of units of the public good.

In Example 1, it is clearly in the interest of each player to commit himself irrevocably to consuming all of his endowment before either of the others can do likewise and in communicating this commitment to the others. The best joint strategy for the remaining players is then to agree to use their joint endowment for the production of the public good. This, we maintain, is the essence of the much discussed "free-rider problem". Admitting the possibility of irrevocable self-commitments makes "free-rider" and other threats accessible to game-theoretic analysis. Our model is intended to capture the viewpoint that preplay negotiations evolve under the impetus of irrevocable self-commitments or the possibilities of such commitments. Of course, if irrevocable self-commitments are not possible for some reason in preplay negotiations, then threats of the free-rider kind may not be believable and our solutions may not be applicable.

In order for irrevocable self-commitments to be possible, we shall have to hypothesize orderings on the commitment opportunities. Such assumptions are unusual in the game-theoretic literature; although the noncooperative model of monopolistic power in Kats [2], for example, reflects similar considerations. These orderings are exogenous in our model and may be interperted as representing negotiating strengths or other psychological factors. Of course, it may not always be advantageous to commit oneself early in the negotiation process; and our model sheds some light on the conditions under which it may or may not be. The possibility of allowing for irrevocable self-commitments has been discussed previously in Schelling [4] and Harsanyi [1] and considered explicitly in an arbitration model in Rosenthal [3]. In what follows, however, we shall be concerned with defining reasonable outcomes for an actual play of the games.

After some definitions and the details of the preplay negotiation process are specified in Section 2, the main solution concept, the induced outcome, is presented in Section 3. Because induced outcomes do not exist in all games, a modification of this concept is presented and discussed in Section 4. This modification is seen to possess both desirable and undesirable features.

2. Definitions and the Model

An n-person game in normal form consists of a player set $N = \{1,\ldots,n\}$; a nonempty strategy set $\sigma(S)$ for each coalition S (nonempty subset of N) satisfying $\sigma(S) \supseteq \underset{T \in P(S)}{\times} \sigma(T)$ for every partition P(S) of S;

and an ordinal utility function $u^i\colon \sigma(N) \to \mathbb{R}^1$ for each $i \in N$. We shall be considering some well-studied examples of characteristic-function games with side payments. An <u>n-person side-payment game in characteristic-function form</u> (spc game) is a pair (N,v), where N is as above and v is a real-valued function on the coalitions of N satisfying: $v(S \cup T) \geq v(S) + v(T)$ whenever $S \cap T = \emptyset$ (superadditivity). Let $\mathbb{R}^S$ denote the Euclidean space of dimension equal to $|S|$, the number of players in S, each coordinate corresponding to a player in S. The <u>canonical normal-form game for (N,v)</u> (cnf for (N,v)) is constructed as follows:

$$\sigma(S) = \{x \in \mathbb{R}^S : \sum_{i \in S} x^i \leq v(S)\} \qquad \text{for each } S \subseteq N;$$

and

$$u^i(x) = x^i \qquad \text{for each } x \in \sigma(N), \text{ each } i \in N.$$

Let $\mathcal{J}$ be a nonempty collection of coalitions of N and let $\mathcal{S}$ be a nonempty subset of $\mathcal{J}$. Let $\hat{\mathcal{O}}(\mathcal{J})$ be an ordering of the elements of $\mathcal{J}$ and let $\mathcal{O}(\mathcal{S})$ be an ordering of the elements of $\mathcal{S}$. $\hat{\mathcal{O}}$ <u>agrees</u> with $\mathcal{O}$ on $\mathcal{S}$ if the ordering of the elements of $\mathcal{S}$ under both $\hat{\mathcal{O}}$ and $\mathcal{O}$ is the same. Two coalitions S and T <u>meet</u> if $S \cap T \neq \emptyset$.

To the strategy set of each coalition of size greater than one, it will be convenient to append an extra element, "blank" or "-", to indicate no strategy. The <u>offer set</u> for each coalition S is

$$\bar{\sigma}(S) = \begin{cases} \sigma(S) & \text{if } |S| = 1 \\ \sigma(S) \cup \{-\} & \text{if } |S| > 1 . \end{cases}$$

Let M(S) be the set of coalitions which meet the coalition S. A <u>commitment</u> for S, α_S, is a pair $(\mathcal{O}(M(S)), \{\beta_T\colon T \in M(S)\})$ where $\mathcal{O}(M(S))$ is an ordering of the coalitions of M(S) and $\beta_T \in \bar{\sigma}(T)$ for each $T \in M(S)$. A commitment for S may be interpreted as follows. The players in S commit themselves to a set of offers, one from the offer set of each coalition which meets S. In addition, an ordering $\mathcal{O}(M(S))$ is established in which these offers must be considered. Suppose $i \notin S$. A <u>commitment for i given α_S</u>, $\alpha_i(\alpha_S)$, is a pair $(\hat{\mathcal{O}}(M(S \cup \{i\})), \{\gamma_T: T \in M(S \cup \{i\})\})$ where

a. $\hat{\mathcal{O}}$ agrees with $\mathcal{O}$ on $M(S)$

b. $\gamma_T \in \overline{\sigma}(T)$ for each $T \in M(S \cup \{i\})$

c. $\gamma_T = \beta_T$ if T meets S but not $\{i\}$.

d. $\gamma_T = \beta_T$ or - if T meets both S and $\{i\}$.

The commitment $\alpha_{S \cup \{i\}}$ which results when i responds to α_S with $\alpha_i(\alpha_S)$ is simply $\alpha_i(\alpha_S)$.

We are now in a position to specify the form of our preplay negotiation process with irrevocable self-commitments. An ordering of the players denoted $\langle j_1, \ldots, j_n \rangle$, is specified exogenously. Player j_1 binds himself to a commitment, say $\alpha_{\{j_1\}}$. Player j_2 now responds by accepting or rejecting (indicated by changing to a blank) each offer corresponding to $M(\{j_1, j_2\})$, proposing an offer for each coalition which meets $\{j_2\}$ but not $\{j_1\}$, and ordering all of these offers along with the rest of j_1's offers so as to agree with j_1's ordering. This process continues on through j_n's commitment. Any time a blank appears during the process it remains. Thus the first player from each coalition in the ordering determines which, if any, strategy is to be considered by that coalition. Any player from that coalition may veto the strategy when it is his turn, but no new strategy for the coalition may then be substituted (since players who have already committed themselves have no further active role to play). If all players in a coalition S agree to a strategy for that coalition, it still may not be played, however, if the strategy of some coalition which precedes S in the final ordering and which meets S has also been agreed upon by all of its members.

A commitment α_N for N generates a unique outcome in $\sigma(N)$ as follows. From the ordered offers of α_N the first nonblank (there must be at least n since $\sigma(\{i\})$ contains no blank for each $i \in N$) indicates a fully accepted offer and a strategy for the corresponding coalition. Eliminate all remaining offers in α_N for coalitions which meet this coalition. Consider the next remaining nonblank in α_N This nonblank indicates a strategy for some other coalition. Eliminate all coalitions which meet this one. Continue on with this process until a strategy has been selected for every coalition in some partition of N. We denote by $s(\alpha_N)$ the strategy combination which results in this way from α_N.

As an illustration, consider the following commitment sequence for the normal form of Example 1 relative to the ordering $\langle 1,2,3 \rangle$ in which $\sigma(S) \subseteq \mathbb{R}^S$ is the set of possible ordered private-good consumptions, all remaining private goods being used to produce the public good.

Player							
1		$\{1\}$ 1		$\{1,2\}$ -		$\{1,3\}$ -	$\{1,2,3\}$ -
2	$\{2,3\}$ (0,0)	$\{1\}$ 1		$\{1,2\}$ -	$\{2\}$ 1	$\{1,3\}$ -	$\{1,2,3\}$ -
3	$\{2,3\}$ (0,0)	$\{1\}$ 1	$\{3\}$ 1	$\{1,2\}$ -	$\{2\}$ 1	$\{1,3\}$ -	$\{1,2,3\}$ -

Here player 1 commits himself to being a free rider. Player 2 responds by offering to contribute all of his endowment to production of the public good as long as player 3 agrees to do the same. If player 3 does not agree then 2 will consume his endowment. Player 3 agrees to 2's offer. The resulting outcome in $\sigma(N)$ is (1,0,0). The corresponding utility vector is

$\left(\frac{7}{3}, \frac{4}{3}, \frac{4}{3}\right)$. Note that it makes no difference in this example which ordering player 1 selects for his offers. Players 2 and 3, however must be careful to keep the offers to $\{2,3\}$ in front of their respective offers to $\{2\}$ and $\{3\}$. Another commitment for 1 which accomplishes the same purpose is

$$1. \quad \begin{matrix} \{1,2,3\} & \{1\} & \{1,2\} & \{1,3\} \\ (1,0,0) & 1 & - & - \end{matrix}.$$

There are several aspects of this particular model of the preplay negotiations which are worth discussing. Firstly, use of this model means that a complete ordering of the players for commitment purposes is given and known to all players. As was mentioned earlier, the hypothesis that such an ordering exists represents a departure from most traditional models. Nevertheless, some such assumption seems to be necessary in order that outcomes be well-defined when more than one player may make a commitment. The game-theorist may wish to examine the commitment process under many different assumed orderings. The assumption that the ordering is known by all players may be relaxed in the following way. An initial probability distribution is given over all possible orderings. A random selection is made according to the probability distribution, but the players are not informed as to which ordering is chosen. The players are selected to make commitments according to this ordering. When a player's turn comes, he knows the order of all previous commitments and therefore has additional information about what the ordering of the uncommitted players will be. He must make his commitment based on this incomplete information. We intend to explore this extension in subsequent research, where the players will be assumed to be maximizing expected utility. In the present paper, we shall restrict ourselves to ordinal considerations.

Secondly, the assumption that each player offers at most one strategy to a coalition seems to be restrictive. In some examples it may be desirable for a player to offer a set of strategies for the subsequent players in the coalition to choose from. Such a complication also merits consideration in further research, although in the examples studied so far it does not appear to add interesting commitment alternatives.

Thirdly, we shall see in certain examples that it is not always desirable for a player to make a commitment when his position in the ordering occurs. Nevertheless, in our model a player must commit when his turn comes. This might be seen as a disadvantage of our model; but we envision one of the possible uses of our model to distinguish those games in which it is desirable from those in which it is not desirable to commit at various times. This, in turn, could reveal more about the nature of negotiations and bargaining in general.

The assumption that a player takes no further active part in the negotiations after his commitment is related to both of the last two points. This assumption considerably simplifies the task of describing irrevocable commitments.

Finally, we should take note of the fact that the normal-form game with which all of these commitments are concerned is assumed to be one of complete information. That is, each player is assumed to know every player's strategy set and every player's utility on all of $\sigma(N)$. The solution concepts to be discussed in the next two sections make strong use of this assumption. Extensions to games with incomplete information are possible but would take us beyond the purely ordinal considerations of this paper.

3. Induced Outcomes

For the last player in a commitment sequence, the decision of which offer to accept is rather uncomplicated. Since there are no more decisions to be made after his, he has nothing to consider but which available offer is best for him. Unfortunately, he may assign equal utility to the different outcomes associated with several available offers. The preceding players in the ordering, recognizing this, must take into account that they have no basis on which to predict which of these offers he will accept.

Let $T=\{j_1,\ldots, j_{n-1}\}$. $\bar{\alpha}_{j_n}(\alpha_T)$ is a rational response to α_T if

$$u^{j_n}(s(\bar{\alpha}_{j_n}(\alpha_T))) = \max_{\alpha_{j_n}(\alpha_T)} u^{j_n}(s(\alpha_{j_n}(\alpha_T)))$$

where the maximum is taken over all responses $\alpha_{j_n}(\alpha_T)$ to α_T. Of course, the maximum may not be attained in certain circumstances. We shall deal with this case presently. We denote by $r(\alpha_T)$ the set of outcomes which are associated with rational responses to α_T. Proceeding inductively, suppose that $T=\{j_1,\ldots, j_{k-1}\}$ and that $r(\alpha_{j_k}(\alpha_T))$ has been defined and is nonempty for all commitments $\alpha_{j_k}(\alpha_T)$ given α_T. Then $\bar{\alpha}_{j_k}$ is a rational response if

$$\inf_{s\in r(\alpha_{j_k}(\bar{\alpha}_T))} u^{j_k}(s) = \max_{\alpha_{j_k}(\alpha_T)} \inf_{s\in r(\alpha_{j_k}(\alpha_T))} u^{j_k}(s).$$

Similarly, if $r(\alpha_{j_1})$ is nonempty for each commitment α_{j_1} for $\{j_1\}$, then $\bar{\alpha}_{j_1}$ is a rational commitment for j_1 if

$$\inf_{s\in r(\bar{\alpha}_{j_1})} u^{j_1}(s) = \max_{\alpha_{j_1}} \inf_{s\in r(\alpha_{j_1})} u^{j_1}(s).$$

The outcome of a sequence of commitments, each of which is a rational response or commitment, is an induced outcome relative to the given ordering of the players.

To illustrate, consider the commitment sequence presented in Section 2 for Example 1. Clearly 3's only rational responses involve accepting the offer to $\{2,3\}$ before making or any other offer. Similarly, 2's only rational responses involve offering $(0,0)$ to $\{2,3\}$ before making any other offer. It is not difficult to see therefore that 1 can insure himself utility of $\frac{7}{3}$ and no more either by only offering 1 to $\{1\}$ or by only offering (1,0,0) to $\{1,2,3\}$ and 1 to $\{1\}$. In the latter case, 2 must either accept (1,0,0) or offer (0,0) to $\{2,3\}$. Clearly the only induced outcome for the ordering $\langle 1,2,3\rangle$ is (1,0,0). Similarly for the other orderings.

Unfortunately, it is easy to produce nonpathological examples in which no induced outcomes exist.

Example 2: The cnf for the spc game having $N=\{1,2\}$; $v(\{1\})=v(\{2\})=0$; $v(\{1,2\})=1$.

For the ordering $\langle 1,2\rangle$, if 1 offers (1,0) to $\{1,2\}$, 2 may rationally accept or reject. If 1 offers $(1-\varepsilon,\varepsilon)$ for $\varepsilon>0$, 2 must accept. Therefore 1 has no rational commitments. Example 2 suggests the introduction of ε-tolerances in the definition of rationality. This will be done in Section 4.

Another difficulty with the notion of induced outcome is illustrated in the following example.

Example 3: $N=\{1,2\}$; $\sigma(\{1\})=\{a,b,c,d\}$; $\sigma(\{2\})=\{e,f,g\}$; $\sigma(\{1,2\})=\sigma(\{1\})\times\sigma(\{2\})$;

$u^1(a,e)=u^1(a,f)=u^1(b,e)=u^1(c,g)=u^1(d,f)=1$;

$u^1(s)=0$ elsewhere;

$u^2(d,g)=0$;

$u^2(s)=1$ elsewhere.

For the ordering $\langle 1,2\rangle$ all outcomes except (d,g) are induced in this example, since player 1 can not assure himself of more utility than 0. Nevertheless, strategy a weakly dominates strategy b for player 1; and there seems to be no reason for him to commit to strategy b. Thus, it seems that the set of induced outcomes is too large for this example. In seeking reasonable criteria for use in reducing this set, however, we are confronted right away with some difficult judgements. Would we want to eliminate commitments by player 1 involving strategy c? Here the case is not so clear. On the one hand, there is no direct domination as with strategy b. On the other hand, there is little difference between strategies b and c from either player's point of view. What about strategy d for player 1? Again a dominates d weakly for player 1. On the other hand, 1 knows that 2 will not play g in response to d; so the situations are not actually comparable from 1's point of view.

One possible approach for the resolution of this matter is the introduction of probability assessments for each player over the commitments which a subsequent player may view with indifference. Such an approach would again involve expected utilities and take us away from the purely ordinal considerations of this paper. We shall therefore defer such considerations until a later time and continue to use the notion of induced outcome as previously defined.

The difficulties encountered in the last two examples all stem from the indifferences of players over various relevant outcomes. When indifferences are absent and some compactness and continuity assumptions hold, no such problems arise.

<u>Theorem 1:</u> Suppose that for each $S \subseteq N$, $\sigma(S)$ is a compact topological space with the property that if A is an open subset of $\sigma(S)$ and $(T_1, \ldots, T_k)$ is a partition of S, then A restricted to $\sigma(T_1) \times \ldots \times \sigma(T_k)$ is open in the product topology of $\sigma(T_1) \times \ldots \times \sigma(T_k)$. Suppose that for each $i \in N$, u^i is upper semi-continuous and one-to-one. Then, relative to every ordering of N, there is a unique induced outcome.

Theorem 1 is proved in the Appendix. We note that the hypotheses of Theorem 1 seem to restrict the strategy spaces severely. On the other hand, one-to-oneness is only crucial at certain locations in the domain. Hence, we can expect to find unique induced outcomes in a somewhat wider class of games than is treated in the theorem.

Before leaving this section we take note of an additional property of induced outcomes.

Theorem 2: Every induced outcome of a game is weakly Pareto optimal. (No outcome is strictly preferred by all players.)

Proof: Suppose not. In the commitment sequence leading to any nonoptimal induced outcome, change the offer to N in every player's commitment to the Pareto superior outcome. Now in each player's ordering of offers, make this offer to N precede all others. The result of the altered commitment sequence is now the Pareto superior outcome. Furthermore, no player can ensure a higher utility by switching to another commitment; since otherwise he could do so in the original commitment sequence. Hence, the Pareto superior outcome is induced. Hence, the nonoptimal outcome was not induced, since the first player's commitment could not have been rational.

It is easy to see that induced outcomes may fail to be strongly Pareto optimal, since an indifferent player might have to make a crucial decision.

4. ϵ-Induced Outcomes

In this section we shall introduce one extension of the concept of induced outcome. We shall see that it exists under general assumptions and that it is intuitively appealing in some simple examples. Unfortunately, it fails to possess certain desirable continuity properties. For simplicity of notation, we shall assume throughout this section that the player ordering is $\langle 1, \ldots, n \rangle$ and that T_i denotes the coalition $\{1, \ldots, i\}$ for $i = 1, \ldots, n$.

Fix $\epsilon > 0$. For player n, the set of ϵ-rational responses to $\alpha_{T_{n-1}}$ is the same as the set of rational responses whenever the set of rational responses is nonempty. When and only when the set of rational responses to $\alpha_{T_{n-1}}$ is empty,

$\bar{\alpha}_n$ is ϵ-rational if it satisfies

$$u^n(s(\bar{\alpha}_n(\alpha_{T_{n-1}}))) \geq \sup_{\alpha_n(\alpha_{T_{n-1}})} u^n(s(\alpha_n(\alpha_{T_{n-1}}))) - \epsilon.$$

Proceeding inductively, let $r_\epsilon(\alpha_i(\alpha_{T_{i-1}}))$ be the set of outcomes associated with ϵ-rational responses to $\alpha_i(\alpha_{T_{i-1}})$. Then $\bar{\alpha}_i$ is an ϵ-rational response to $\alpha_{T_{i-1}}$ if

$$\inf_{s \in r_\epsilon(\bar{\alpha}_i(\alpha_{T_{i-1}}))} u^i(s) = \max_{\alpha_i(\alpha_{T_{i-1}})} \inf_{s \in r_\epsilon(\alpha_i(\alpha_{T_{i-1}}))} u^i(s). \tag{1}$$

If the set of commitments satisfying (1) is empty, then $\bar{\alpha}_i$ is an ϵ-rational response to $\alpha_{T_{i-1}}$ if

$$\inf_{s \in r_\epsilon(\bar{\alpha}_i(\alpha_{T_{i-1}}))} u^i(s) \geq \sup_{\alpha_i(\alpha_{T_{i-1}})} \inf_{s \in r_\epsilon(\alpha_i(\alpha_{T_{i-1}}))} u^i(s) - \epsilon. \tag{2}$$

ϵ-rational commitments for player 1 are defined analogously. Note that when (1) can be satisfied by at least one response, then those responses satisfying (2) but not (1) are not ϵ-rational. Only when (1) fails for all responses is (2) resorted to.

In Example 2, the only rational (and hence ϵ-rational for all ϵ) responses for 2 to $(1-\delta,\delta)$ involve acceptance whenever $\delta > 0$. Hence the commitment

$$1. \quad \begin{matrix} \{1,2\} & \{1\} \\ (1-\delta,\delta) & 0 \end{matrix}$$

is ϵ-rational for 1, whenever $0 < \delta \leq \epsilon$. Permitting player 2 to refuse on the grounds that inequality (2) would still be satisfied by a refusal seems to us to run counter to the basic assumption that each player seeks to maximize his own utility. Only when (1) holds for no response does the ϵ-tolerance seem to us plausible, at least when ϵ can be made arbitratily small. As before, an ϵ-induced outcome results from a sequence consisting of an ϵ-rational commitment by the first player and ϵ-rational responses by the rest.

Example 4: The cnf for the spc game with

$N = \{1,2,3\}$;

$v(\{i\}) = 0, \quad i = 1,2,3$;

$v(\{1,2\}) = a, \ v(\{1,3\}) = b, \ v(\{2,3\}) = c$;

$v(\{1,2,3\}) = 1$.

From superadditivity, $0 \le a,b,c \le 1$.

From symmetry, it suffices to consider the ordering $\langle 1,2,3 \rangle$. If $c < 1$, then for any commitment by 1 consisting only of offers in which the sum of payoffs to $\{2,3\}$ does not exceed c, there is an ϵ-rational response sequence at the outcome of which 1 gets nothing. If 1 proposes (x,y,z) to $\{1,2,3\}$, where $y + z > c$, 2 and 3 have no ϵ-rational responses other than those leading to acceptance. Therefore, 1's ϵ-rational commitments must contain offers in which the total utility to $\{2,3\}$ is $c + \delta$, where $0 < \delta \le \epsilon$. The set of ϵ-induced outcomes when $0 < \epsilon \le 1-c$ is therefore

$$\{(x,y,z) \ge 0 \colon \ y + z > c, \ x \ge 1 - c - \epsilon, \ x + y + z \le 1\}.$$

Notice that the parameters a and b have no effect when the ordering is $\langle 1,2,3 \rangle$. 1's power is inversely proportional to the value of c. The solution concept treats players 2 and 3 alike. There is therefore nothing to be gained by being second rather than third in the ordering (as long as 1 behaves sensibly). Whether it pays to be first in the ordering depends on the values of the parameters and to an arbitrary decision by the player who does go first as to the split between the other players.

If $c = 1$, there exist ϵ-rational responses to all of 1's commitments such that 1 receives nothing. By offering (x,z) to $\{1,3\}$, 1 can ensure that 3 will receive more than z but not more than $z+\epsilon$ after 2's ϵ-rational response. The set of ϵ-induced outcomes for this case is

$$\{(0,y,z):\ 0 \leq z \leq b+\epsilon,\ y \geq 1-z-\epsilon,\ y+z \leq 1\} \quad \text{if } a=1,$$

and

$$\{(0,y,z):\ 0 < z \leq b+\epsilon,\ y \geq 1-z-\epsilon,\ y+z \leq 1\} \quad \text{if } a<1.$$

Being first in the ordering is not advantageous under these circumstances. The second player in the ordering appears to be the strongest, but much of his advantage relative to the third player can be offset by the commitment of player 1, which has no effect on 1's own utility.

Although the calculation of the sets of induced and ϵ-induced outcomes for the examples so far considered has been relatively simple, the reader may convince himself by trying a few examples that complications mount quickly as the number of players increases beyond three.

Under mild assumptions, we can show that the set of ϵ-induced outcomes exists.

<u>Theorem 3</u>: Suppose that u^i is bounded from above on $\sigma(N)$ for $i=1,\ldots,n$. Then the set of ϵ-induced outcomes is nonempty whenever $\epsilon>0$.

<u>Proof</u>: For every commitment $\alpha_{T_{n-1}}$, let

$$\ell(\alpha_{T_{n-1}}) = \sup_{\alpha_n(\alpha_{T_{n-1}})} u^n(s(\alpha_n(\alpha_{T_{n-1}}))) < \infty.$$

If the supremum is attained, then the set of ϵ-rational responses coincides with the nonempty set of rational ones. Otherwise, α_n is ϵ-rational relative to $\alpha_{T_{n-1}}$ if

$$u^n(s(\alpha_n(\alpha_{T_{n-1}}))) \geq \ell(\alpha_{T_{n-1}}) - \epsilon.$$

Clearly $r_\epsilon(\alpha_{T_{n-1}})$ is nonempty in either case for every $\alpha_{T_{n-1}}$. Proceeding inductively, if $r_\epsilon(\alpha_i(\alpha_{T_{i-1}}))$ is nonempty for each $\alpha_i(\alpha_{T_{i-1}})$, then

$$\sup_{\alpha_i(\alpha_{T_{i-1}})} \quad \inf_{s \in r_\epsilon(\alpha_i(\alpha_{T_{i-1}}))} u^i(s)$$

exists. If for some $\bar{\alpha}_i$,

$$\inf_{s \in \bar{\alpha}_i(\alpha_{T_{i-1}})} u^i(s) = \sup_{\alpha_i(\alpha_{T_{i-1}})} \quad \inf_{s \in r_\epsilon(\alpha_i(\alpha_{T_{i-1}}))} u^i(s),$$

it is an ϵ-rational response; and $r_\epsilon(\alpha_{T_{i-1}})$ is nonempty. Otherwise, the set of ϵ-rational responses is the nonempty set which satisfies (2). Again $r_\epsilon(\alpha_{T_{i-1}})$ is nonempty. Similarly for player 1.

If our explanation of the role of ϵ in this theory is to be convincing, a desirable property would be that at small enough levels of ϵ, the set of ϵ-induced outcomes should not behave too badly when ϵ is slightly perturbed. Unfortunately, the following examples show that the correspondence which associates to each ϵ the set of ϵ-induced outcomes may be neither upper nor lower semi-continuous near zero.

<u>Example</u> 5: $N = \{1,2\}$;

$\sigma(\{1\}) = \{(x_1,x_2): \; 0 \leq x_1 \leq 1, \; 0 \leq x_2 \leq 1\}$;

$\sigma(\{2\}) = \{a,b\}$;

$\sigma(\{1,2\}) = \sigma(\{1\}) \times \sigma(\{2\})$;

$$u^1(x,a) = -\max(x_1,x_2);$$
$$u^1(x,b) = -1;$$
$$u^2(x,a) = \max(x_1,x_2);$$
$$u^2(x,b) = \max(x_1,x_2) - |x_2-x_1|.$$

The strategy spaces in Example 5 are compact and the utility functions for both players are continuous in the strategies. Let $\{\epsilon_j\}$ be a decreasing sequence of real numbers converging to small $\epsilon > 0$. Let $x^j = (\epsilon_j - \frac{1}{2}\epsilon, \frac{1}{2}\epsilon)$. For ϵ_j sufficiently close to ϵ, (x^j,a) is ϵ_j-induced relative to the ordering $\langle 1,2\rangle$. But $\{(x^j,a)\} \to (\frac{1}{2}\epsilon, \frac{1}{2}\epsilon, a)$, which is not ϵ-induced relative to $\langle 1,2\rangle$. Hence the set of ϵ-induced outcomes in Example 5 is not upper semi-continuous at any ϵ near zero.

<u>Example</u> 6: $N = \{1,2,3\}$;

$$\sigma(\{1\}) = \sigma(\{2\}) = \{0\};\ \sigma(\{3\}) = \{z:\ 0 \le z \le 1\};$$
$$\sigma(\{1,2\}) = \{(0,0)\} \cup \{(x,y):\ x+y=1\};$$
$$\sigma(\{1,3\}) = \sigma(\{1\}) \times \sigma(\{3\});$$
$$\sigma(\{2,3\}) = \sigma(\{2\}) \times \sigma(\{3\});$$
$$\sigma(\{1,2,3\}) = \sigma(\{1,2\}) \times \sigma(\{3\}).$$
$$u^1(x,y,z) = x;$$
$$u^2(x,y,z) = y;$$
$$u^3(x,y,z) = \begin{cases} h(z) & \text{if } z \le y \\ h(y) & \text{if } z > y \end{cases} \quad \text{where } h(t) = \min(0, t\sin 1/t) \text{ for } t \ge 0.$$

The strategy spaces are compact in Example 6, and the utility functions are continuous. Note that players 1 and 2 are bargaining over one unit of utility and are not concerned with player 3. The set of ϵ-induced outcomes for their 2-player game is $\{(1-y,y):\ 0 < y \le \epsilon\}$. Note that the function $h(t)$ attains the value zero at a sequence of disjoint closed intervals converging to the point zero from above. Let $\epsilon > 0$ be the smallest value of t in one such interval. Let $\{\epsilon_j\}$ be an increasing sequence of real numbers converging to ϵ. Now $(1-\epsilon,\epsilon,\epsilon)$ is ϵ-induced; but there is no sequence of ϵ_j-induced outcomes which converges to $(1-\epsilon,\epsilon,\epsilon)$. Thus, lower semi-continuity fails for the set of ϵ-induced outcomes whenever $\epsilon > 0$ is a left endpoint of any of the above intervals. In every neighborhood of zero, therefore, lower semi-continuity fails.

Although the phenomena exhibited in the last two examples are disquieting, it remains to be seen whether they are in some sense pathological. If not, a further modification or abandonment of the ϵ-induced outcome would seem to be called for. On the other hand, the set of ϵ-induced outcomes in Examples 2 and 4 and the premeses on which the solution concept is based are appealing to us.

References

1. Harsanyi, J., "On the Rationality Postulates Underlying the Theory of Cooperative Games", Journal of Conflict Resolution, 5(1961), 179-196.

2. Kats, A., "Non-Cooperative Monopolistic Games and Monopolistic Market Games", International Journal of Game Theory, 3(1975), 251-260.

3. Rosenthal, R., "An Arbitration Model for Normal-Form Games", Mathematics of Operations Research, forthcoming.

4. Schelling, T., "An Essay on Bargaining", American Economic Review, 46(1956), 281-306.

Appendix

Proof of Theorem 1: Without loss of generality we take the player ordering to be $\langle 1,\ldots, n\rangle$. Let T_i denote the coalition $\{1,\ldots, i\}$, for $i=1,\ldots, n$. We shall first establish by backward induction that for $i=1,\ldots, n-1$, $r(\alpha_{T_i})$ is nonempty and consists of a unique outcome, denoted $\hat{r}^i(\alpha_{T_i})$ (where α_{T_i} represents any commitment for the coalition T_i); furthermore that $\hat{r}^i$ is continuous in a sense to be described. (To avoid confusion in this proof we distinguish between the various $\hat{r}^i$ which are defined on different domains. The common symbol r retains its earlier meaning. The domain of r should be obvious in each usage from the context.)

For every commitment α_{T_n}, a unique outcome is generated according to the procedure in Section 2. Define $\hat{u}^n(\alpha_{T_n}) = u^n(s(\alpha_{T_n}))$. For each fixed ordering $\mathcal{O}_N$ of the coalitions in N, we may consider the topological space $(C_{\mathcal{O}_n}, \mathcal{C}_{\mathcal{O}_n})$ of the commitments for N as the product of the offer spaces for each coalition together with the product topology. (If B is open in $\sigma(s)$ then B and $B\cup\{-\}$ are open in $\bar{\sigma}(S)$ whenever $|S|>2$.) Now $C_{\mathcal{O}_N}$ is compact; $C_{T_n} = \bigcup_{\mathcal{O}_N} C_{\mathcal{O}_N}$ is compact as a topological space, with the topology generated by the union of the open sets over all orderings $\mathcal{O}_N$; and $\hat{u}^n$ is upper semi-continuous on C_{T_n} (since $s: C_{T_N} \to \sigma(N)$ is continuous and $u^n: \sigma(N) \to \mathbb{R}^1$ is upper semi-continuous). Let $A^{n-1}(\alpha_{T_{n-1}}) = \{\alpha_{T_n} : \alpha_{T_n} = \alpha_n(\alpha_{T_{n-1}}) \text{ for some } \alpha_n\}$. Clearly, A^{n-1} is a compact-valued continous correspondence on $C_{T_{n-1}}$ (the space of commitments for T_{n-1} with topology analogous to that of C_{T_n}). Hence $\hat{u}^n$ takes on its maximum over $A^{n-1}(\alpha_{T_{n-1}})$, and each maximizer for a fixed $\alpha_{T_{n-1}}$ gives rise to the same outcome. Furthermore, the set of maximizers is a compact-valued, upper semi-continuous correspondence over $C_{T_{n-1}}$. Hence $\hat{r}^{n-1}$ is a continuous function from $C_{T_{n-1}}$ into $\sigma(N)$.

For $i = 1, \ldots, n-2$, we assume that $r(\alpha_{T_{i+1}})$ exists and consists of the unique outcome $\hat{r}^{i+1}(\alpha_{T_{i+1}})$ and that $\hat{r}^{i+1}$ is continuous as a function from $C_{T_{i+1}}$ (with topology analogous to those of $C_{T_{n-1}}$ and C_{T_n}) into $\sigma(N)$. Let $\hat{u}^{i+1}(\alpha_{T_{i+1}}) = u^{i+1}(\hat{r}^{i+1}(\alpha_{T_{i+1}}))$. $\hat{u}^{i+1}: C_{T_{i+1}} \to \mathbb{R}^1$ is then continuous. $A^i(\alpha_{T_i}) = \{\alpha_{T_{i+1}} : \alpha_{T_{i+1}} = \alpha_{i+1}(\alpha_T) \text{ for some } \alpha_{i+1}\}$ is compact, and A^i is continuous on C_{T_i}. Hence $\hat{u}^{i+1}$ takes on its maximum over $A^i(\alpha_{T_i})$, and each maximizer gives rise to the same outcome $\hat{r}(\alpha_{T_i})$ in $\sigma(N)$. The set of maximizers is compact-valued and upper semi-continuous, and $\hat{r}^i$ is continuous on C_{T_i}.

To show that the set of induced outcomes consists of a single outcome, it now suffices to apply the same argument to show that $\hat{u}^1$ takes on its maximum over C_{T_1}.