Discussion Paper No. 121

An Arbitration Model for Normal-Form Games

by

Robert W. Rosenthal*

Revised September, 1975

*This work was supported by National Science Foundation grant # SOC-7401790 to Northwestern University. I am grateful to Ehud Kalai and Nimrod Megiddo for key observations concerning the Theorem in Section 3 and to Robert Aumann for several improvements in the paper.

An Arbitration Model for Normal-Form Games

by

Robert W. Rosenthal

Abstract: An arbitration model for cooperative two-person normal-form games is suggested in which each player's strength is measured by what he can obtain through committing himself to a course of action before his opponent does. This approach differs from earlier models in which threats were evaluated on the basis of their relative effects on the two players. Possibilities for extensions to the n-player case are included.

Received Revision received

AMS 1970 Subject Classifications: Primary: 90D12. Secondary: none.

IAOR 1973 Subject Classifications: Main: Games. Cross references: none.

Key Words and Phrases: game theory; bargaining; arbitration; irrevocable commitments

Discussion Paper No. 121

An Arbitration Model for Normal-Form Games

by

Robert W. Rosenthal*

Revised September, 1975

*This work was supported by National Science Foundation grant # SOC-7401790 to Northwestern University. I am grateful to Ehud Kalai and Nimrod Megiddo for key observations concerning the Theorem in Section 3 and to Robert Aumann for several improvements in the paper.

1. Introduction

The purpose of this paper is to develop a new model for the arbitration of cooperative games in normal form. We take it for granted that the purpose of arbitration is to select a single Pareto-optimal outcome which, in some sense, reflects the relative strengths of the players in the underlying cooperative game. The novel feature of our approach is the method by which these strengths are measured.

For two-person games, Zeuthen proposed in [14] and Nash proposed and axiomatized in [6] the same solution for the special case of bargaining with a fixed threat point. Nash later extended this in [7] to two-person games in normal form. The Zeuthen-Nash bargaining model and Nash's extended bargaining model have been the major focus of interest in the two person case. (Since bargainers presumably take advantage of their relative strengths, we assume that arbitrators could make use of bargaining solutions.) In [9] Raiffa introduced a different solution for the fixed-threat case which has recently been axiomatized by Kalai and Smorodinsky [4].

For n-person games the major arbitration-type solution concepts are all derived from the Shapley value, proposed and axiomatized in [12] for n-person characteristic-function games with side payments. This has been extended to normal-form games with side payments by Harsanyi [2] and Selten [11]. The Harsanyi-Selten concept also generalizes Nash's extended bargaining solution. A further extension to characteristic-function and normal-form games without side payments has been indicated by Shapley in [13]. (Although the references in the above discussion do not exhaust the important work done on models of arbitration and bargaining, they should be sufficient for purposes of comparison with what we will propose. Some other arbitration schemes are described, for example, in [5] and [9].)

Differing from all of these, our approach focuses on what each player can gain by committing himself to a course of action before his opponent does. It is perhaps most easily seen in contrast with Nash's extended bargaining model for the case of two-person games with side payments. For this reason we begin by developing this case in Section 2, although our solution coincides with the Shapley value of the characteristic-function form in this setting. In Section 3, we extend our idea to the two-person no-side-payment case. In Section 4 we indicate the relationship with the Raiffa-Kalai-Smorodinsky (RKS) model and the Zeuthen-Nash model for the special case of fixed threats. In Section 5, extensions to n-person games are proposed.

2. Two-Person Games With Side Payments

In this section we begin with the problem of how to arbitrate a cooperative two-person game in normal form when unrestricted side payments in a transferable utility medium are possible. For this class of games, Nash's extended bargaining model involves constructing a zero-sum "threat game" from the given normal-form game by replacing the payoff pairs (a_{ij}, b_{ij}) for each pure strategy combination (i,j) by $(a_{ij} - b_{ij})$. The extended bargaining solution to the cooperative game is then that Pareto optimal point (x,y) for which $x - y$ equals the equilibrium value of the threat game. Usage of equilibria in the threat game, however, reveals a particular point of view about which threats are believable; namely, that a threat is believable or not according to the relative severity of damage which is done to the opponent vs. damage to the threatener. "This will hurt you more than me" is acceptable reasoning.

Consider the following three examples, all normal-form games played cooperatively and with side payments.

Example 1:

10,0	-10,-30

Example 2:

10,0	-20,-30

Example 3:

10,0	-30,-30

The extended bargaining solutions are, respectively, (10,0), (10,0) and (5,5). In the first two examples, player 2's threat to play his second strategy is ignored by the solution; while in the third example, it is evidently a reasonable enough threat to warrant a side payment of 5. To us the relative damage kind of argument and the consequent solution seem to be somewhat arbitrary.

It seems to us that whether or not player 2's threat in each of the above games is believable depends upon whether or not he can convince player 1 that he is irrevocably committed to the threat before player 1's last chance to agree to a side payment. If player 2 desires only to maximize his own utility and if he has left himself the choice, then when it comes time actually to play the game, he will not use his threat. Thus it seems to us that threats are inextricably tied up with the idea of self-commitment to various strategies. In example 1, if player 2 is able to commit himself irrevocably to the following policy "play strategy 2 unless player 1 agrees to a side payment of $(20 - \epsilon)$ utility units (where ϵ is a small positive number)", then player 1's only rational action is to agree. The outcome is then $(-10 + \epsilon, 20 - \epsilon)$.

The concept of irrevocable self-commitment has been discussed in the bargaining literature (see, for example, [3] and [10]) but never previously incorporated into a formal model to our knowledge. In other game-theoretic settings it has been the subject of some controversy (see, for example, [1]). We shall avoid the controversy for the present by stating that our proposed solution is meant to apply to situations in which irrevocable self-commitments are possible, but the arbitrator is unable to know which player is in a better position to commit first. For situations in which this is not a valid assumption, our solution is certainly based on somewhat arbitrary considerations. (With unspecified preplay negotiation rules, however, some arbitrariness is probably unavoidable.)

We propose that the arbitrator consider the commitment statement for each player which is best for him along with the opponent's best response and then treat the resulting two points symmetrically; i.e., split the difference. (This assumes an absence of any information for the arbitrator about which player is in a better position to commit first. In practice, such information may be available and perhaps should be taken into account, as should any other unmodelled information.) In the examples, 1's best alternative is to commit himself to no side payment, thus insuring the outcome (10,0). Good alternatives for 2 are to commit himself to his second strategy unless 1 agrees to pay $(20-\epsilon)$, $(30-\epsilon)$, and $(40-\epsilon)$, respectively (where ϵ is a small positive number). Treating the commitment points symmetrically and letting ϵ to go to zero, the arbitrated outcomes in the three cases are (0,10), (-5,15), and (-10,20), respectively.

More generally, let (a_{ij}, b_{ij}) be the respective payoffs (in transferable von Neumann-Morgenstern utility) to 1 and 2 for strategy combination $(i,j) \in I \times J$, where I and J are the finite pure strategy sets for 1 and 2, respectively. Since unrestricted side payments are assumed, the Pareto optimal surface for this game is $\{(x,y) : x+y = \max_{(i,j) \in I \times J} (a_{ij} + b_{ij})\}$. Let $I^* = \{p = (p_1, \ldots, p_{|I|}) : p \geq 0, \sum_{i=1}^{|I|} p_i = 1\}$ be the set of mixed strategies for player 1, where $|I|$ and $|J|$ are the cardinalities of I and J, respectively. If 1 were to play mixed strategy p, the payoff to 2's best response would be $\max_{j \in J} \sum_{i=1}^{I} p_i b_{ij}$. 1 can therefore guarantee that 2 receives no more than $\min_{p \in I^*} \max_{j \in J} \sum_{i=1}^{|I|} p_i b_{ij}$. This is simply the value of the zero-sum game defined by the matrix $B^T = (b_{ij})^T$. By committing first, player 1 can therefore effect the payoff (x,y) in the cooperative game such that

$$x + y = \max_{(i,j) \in I \times J} (a_{ij} + b_{ij})$$

$$y = \text{val}\ (B^T) + \epsilon,$$

where ϵ is any positive number (nonnegative if 2's best reply to 1's threatened strategy is Pareto optimal). Letting ϵ go to zero we obtain player 1's "commitment point". Reasoning analogously for player 2 and averaging, the arbitrated outcome of the cooperative game is (x,y) where (setting $A = (a_{ij})$)

$$x + y = \max_{(i,j) \in I \times J} (a_{ij} + b_{ij})$$

$$x - y = \text{val}\ (A) - \text{val}\ (B^T)$$

The Nash extended bargaining solution is, by contrast, the point (x,y) where:

$$x + y = \max_{(i,j) \in I \times J} (a_{ij} + b_{ij})$$

$$x - y = \text{val}\ (A - B)$$

Before leaving this section, let us consider another example.

Example 4:

1,4	-1,-4
-4,-1	4,1

In Example 4, Nash's extended bargaining solution is (1,4), whereas our solution is $\left(\frac{5}{2},\frac{5}{2}\right)$. It is argued in [5] that in this game, player 2 is in a stronger bargaining position than player 1 and that the symmetric Pareto optimum (arrived at there from other considerations) is therefore inappropriate. In that argument it is player 2's threat to play his first pure strategy which is so strong. In our minds, however, player 1's threat to play his second pure strategy is equally viable as long as commitments are possible.

3. Two-Person Games Without Side Payments

For two-person games with no side payments possible, the feasible set of utility pairs, achievable by correlating strategies in the given normal-form game, is the convex hull of the set of payoffs in the normal-form game. The solution proposed in Section 2 for the side-payment case is not applicable for the no-side-payment case for two reasons: 1) the commitment points assuming side payments may not be attainable without side payments, and 2) the average of two Pareto optimal commitment points may no longer be Pareto optimal. We shall extend the notion of commitment point to the no-side-payment case and then provide a natural method for selecting an arbitrated Pareto optimal outcome on the basis of the commitment points.

The rationale behind commitment points remains the same. Whatever strategy player 1 conditionally commits himself to, player 2 can consider his best response. Thus player 1 can hold player 2 to val (B^T). If C is the convex hull of $\{(a_{ij},b_{ij}) : i \in I,\ j \in J\}$, then 1's commitment point is $(\bar{x}_1,\bar{y}_1)$ where:

$$\bar{x}_1 = \max x$$

$$\text{subject to } (x,y) \in C$$

$$y \geq \text{val } (B^T)$$

$$\bar{y}_1 \in Y_1 = \{y : (\bar{x}_1, y) \in C\}$$

The set Y_1 consists of a single point whenever the weak Pareto boundary of C contains no vertical line segments. (A point $(x,y) \in C$ is weakly Pareto optimal if $\nexists (u,v) \in C$ with $u > x$ and $v > y$. (x,y) is strongly Pareto optimal in C if $(u,v) \in C$, $(u,v) \geq (x,y)$ implies $(u,v) = (x,y)$.) In this case, 1's commitment point is uniquely determined. Otherwise, player 1 (who is in control) is indifferent over $\{(\overline{x}_1, y) : y \in Y_1\}$ and 1's commitment point may not be uniquely determined (see Figure 1).

Player 2's commitment point is similarly defined as $(\overline{x}_2, \overline{y}_2)$ where:

$$\overline{y}_2 = \max y$$

$$\text{subject to } (x,y) \in C$$

$$x \geq \text{val }(A)$$

$$\overline{x}_2 \in X_2 = \{x : (x, \overline{y}_2) \in C\}$$

Notice that both commitment points are continuous functions of the pair (A,B) except at some games in which the weak Pareto boundary of C does not coincide with the strong Pareto boundary. Since the solution we shall propose depends continuously on the commitment points, this solution may in turn be discontinuous in (A,B) at such games.

The problem remains for the arbitrator to select a single Pareto optimal point reflecting the strengths of the respective commitment points. We propose the following function F of that part P of the weak Pareto boundary lying between the commitment points: if L is the line segment joining $(\overline{x}_2, \overline{y}_1)$ to $(\overline{x}_1, \overline{y}_2)$ then F chooses the unique (whenever $\overline{x}_1 > \overline{x}_2$ and $\overline{y}_2 > \overline{y}_1$) point common to L and P (see Figure 2). (If $\overline{x}_1 = \overline{x}_2$ or $\overline{y}_1 = \overline{y}_2$, there is a unique outcome which is strongly Pareto optimal. This is the value of F in such cases. Notice that $\overline{x}_1 \geq \overline{x}_2$ and $\overline{y}_2 \geq \overline{y}_1$ always hold.)

Before proceeding to justify this choice of F, we note the following

1. If P consists of a single line segment with slope strictly between zero and $-\infty$, then F selects the midpoint of the line segment. Thus F generalizes our solution for the side-payment case.

2. The domain of F may be extended to include any closed connected subset of the weak Pareto boundary of any compact convex set in $\mathbb{R}^2$. The set of all such sets will be referred to as dom F.

3. F is continuous on dom F (for example, with the Hausdorff distance on closed subsets of $\mathbb{R}^2$). The solution is therefore continuous in the game matrices except at those games mentioned previously.

We shall now attempt to justify F by showing that it is the only function from dom F to $\mathbb{R}^2$ which satisfies the following axioms. (Of course, by restricting matters to dom F, a considerable simplification of the problem has already been made.)

A1: (Pareto optimality) F(P) is strongly Pareto optimal in P.

A2: (symmetry) Suppose P is symmetric (i.e., $(x,y) \in P \Leftrightarrow (y,x) \in P$), then the components of F(P) are equal.

A3: (independence of affine utility transformations) If $P' = \{(\alpha x + \beta, \gamma y + \delta) : (x,y) \in P\}$ and if $F(P) = (x^0, y^0)$, then $F(P') = (\alpha x^0 + \beta, \gamma y^0 + \delta)$.

A4: (monotonicity) Suppose that P and P' are elements of dom F with identical endpoints, and that coP (the convex hull of P) contains coP', then $F(P) \geq F(P')$.

Theorem: F is the only function from dom F to $\mathbb{R}^2$ which satisfies A1-A4.

Proof: That F satisfies A1-A4 is immediate. Let P be any set in dom F with more than one strongly Pareto optimal point. By A3, we may assume that the commitment points of P are at (1,0) and (0,1), respectively. In this case, F selects the symmetric point p in P. Let $P' \in$ dom F be the union of the line segments [(0,1),p] and [p,(1,0)]. By A2, our function must map P' into p. $\text{co}P \supseteq \text{co}P'$; hence our function must map P into that part of P which is at least as large as p. Thus our function must be F.

The axioms, theorem, and proof are all closely related to their analogues in [4].

We shall not go into the details of Nash's extended bargaining model for the no-side-payment case (see [8] for a clear discussion). As in the side-payment case, those threats which in some sense harm the opponent more than the threatener are viable for the threatener in Nash's solution. This is in sharp contrast to our solution in which the harm to the threatener is never taken into account. Our model is therefore responsive to what may seem to be extremely irrational threats. On the other hand, our arbitrator need never consider whether it makes sense to carry out any threats at all, only commitments. Another difference between the two models is that Nash's solution is relatively more sensitive to small changes in the Pareto curve near the solution and completely insensitive to small changes in the Pareto curve near the commitment points.

4. Fixed-Threat Games

Let C be a compact, convex set in $\mathbb{R}^2$ and let $q=(q_1,q_2)$ be a point in C. The fixed threat bargaining game (C,q) may be simply described as follows: if the two players can agree on a utility pair in C, then they receive that point; if they cannot agree, they receive utility pair q. By reasoning as in the previous sections, we can define commitment points $(\bar{x}_1,\bar{y}_1)$, $(\bar{x}_2,\bar{y}_2)$ as follows:

$$\bar{x}_1 = \max x \quad \text{subject to } (x,y) \in C, \ (x,y) \geq q$$

$$\bar{y}_1 \in Y_1 = \{y : (\bar{x}_1,y) \in C\},$$

$$\bar{y}_2 = \max y \quad \text{subject to } (x,y) \in C, \ (x,y) \geq q$$

$$\bar{x}_2 \in X_2 = \{x : (x,\bar{y}_2) \in C\},$$

where Y_1 and X_2 are singletons except possibly when the weak and strong Pareto boundaries of C are not identical. F is then defined as in Section 3.

The solution just defined bears a close relationship with the RKS solution for fixed-threat games. Instead of selecting the point which P and L have in common, the RKS solution ψ selects the point which P and L' (the line segment $[(q_1,q_2), (\bar{x}_1,\bar{y}_2)]$ have in common (see Figure 3). ψ has been shown to be the only solution which satisfies certain reasonable axioms which are closely related to A1 - A4. Our solution agrees with ψ whenever the respective commitment points lie on the horizontal and vertical lines, respectively, through q. When neither of these is true, our solution is insensitive to movements in q; while ψ is sensitive to such movements.

The Nash-Zeuthen solution, by way of contrast, selects the point η in P at which there exists a line supporting C with slope equal to the negative of the slope of the line from q to η. η has been shown to be the only solution which satisfies another set of reasonable axioms. Unlike ψ and F, η depends on the local behavior of P and is insensitive to changes in commitment points.

Both ψ and η select unique outcomes for all fixed threat bargaining games and are continuous functions of C and q. Our solution, once again, is either not uniquely defined for certain games or, if a somewhat arbitrary choice is made from Y_1 or X_2, is discontinuous.

5. n-Person Games

In this section we will indicate briefly how one might extend our two-person solution to an n-person normal-form game Γ. No attempt will be made to justify this extension except by analogy to existing models.

First let Γ be a side payment game. We begin by recalling the Harsanyi-Selten value [2,11]. Given a coalition S, define a 2-player side payment game Γ_S in normal form as follows: The players are S and its complement $N\backslash S$, the pure strategies of S (or $N\backslash S$) are the $|S|$-tuples (or $|N\backslash S|$-tuples) of pure strategies of its members, and the payoff to S (or $N\backslash S$) is the total payoff to its members. Let $v'(s)$ be the payoff to S when the Nash arbitration scheme is used in Γ_S. Then the Harsanyi-Selten value of Γ is the Shapley value $\varphi v'$ of the characteristic function v'.

Rather than applying the Nash arbitration scheme to Γ_S, we may apply our scheme. Let v'' be the resulting arbitrated payoff to S, and let $\varphi v''$ be the corresponding Shapley value. $\varphi v''$ is the arbitrated solution we suggest for Γ. Interestingly, this arbitrated solution coincides with the Shapley value φv when v is the von Neumann-Morgenstern characteristic function, i.e. v(S) is the maxmin of the payoff to S in Γ_S. To see this, note that

$$v'(S) = \frac{1}{2}(v(S) + v(N) - v(N\backslash S)).$$

For games without side payments, we can use the method of [13] to show that for some choices of weight-vectors λ, the arbitrated value $\varphi v''_\lambda$ obtained by weighting each player i's utility by λ_i, allowing side payments, and then using the above arbitration scheme, results in a payoff attainable without side payments in Γ. Such outcomes become candidates for the arbitrated solution to Γ.

References

1. AUMANN, R., and MASCHLER, M., "Some Thoughts on the Minimax Principle", Management Science Vol. 18 (1972), pp. 54-63.

2. HARSANYI, J., "A Bargaining Model for the Cooperative n-Person Game", Annals of Mathematics Studies 40 (Tucker and Luce, eds.) Princeton University Press, 1959, pp. 325-355.

3. HARSANYI, J., "On the Rationality Postulates Underlying the Theory of Cooperative Games", Journal of Conflict Resolution Vol. 5 (1961), pp. 179-196.

4. KALAI, E., and SMORODINSKY, M., "Other Solutions to Nash's Bargaining Problem", Econometrica Vol. 43 (1975), pp. 513-518.

5. LUCE, D., and RAIFFA, H., Games and Decisions, John Wiley and Sons, New York, 1957, Chapter 6.

6. NASH, J., "The Bargaining Problem", Econometrica Vol. 18 (1950), pp. 155-162.

7. NASH, J., "Two-Person Cooperative Games", Econometrica Vol. 21 (1953), pp. 128-140.

8. OWEN, G., Game Theory, W.B. Saunders Co., Philadelphia, 1968, Chapter 7.

9. RAIFFA, H., "Arbitration Schemes for Generalized Two-Person Games", Annals of Mathematics Studies 28 (Kuhn and Tucker, eds.), Princeton University Press, 1953, pp. 361-387.

10. SCHELLING, T., "An Essay on Bargaining", American Economic Review, Vol. 46 (1956), pp. 281-306.

11. SELTEN, R., "Valuation of n-Person Games", Annals of Mathematics Studies 52 (Dresher, Shapley, and Tucker, eds.) Princeton University Press, 1964, pp. 577-626.

12. SHAPLEY, L., "A Value for n-Person Games", Annals of Mathematics Studies 28 (Kuhn and Tucker, eds.), Princeton University Press, 1953, pp. 307-317.

13. SHAPLEY, L., "Utility Comparison and the Theory of Games", La Decision, Editions du Centre National de la Recherche Scientifique, Paris, 1969, pp. 251-262.

14. ZEUTHEN, F., Problems of Monopoly and Economic Warfare, G. Routledge and and Sons, London, 1930.

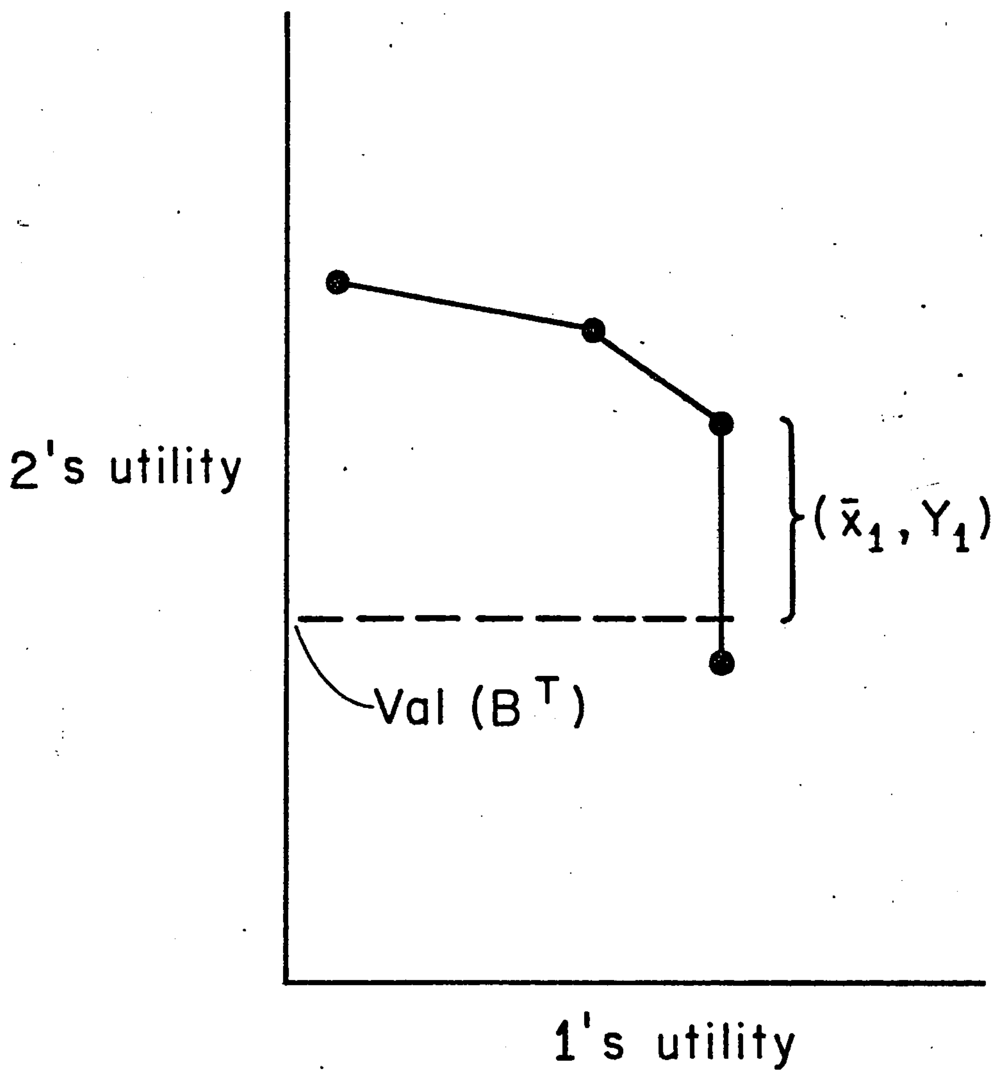

Figure 1

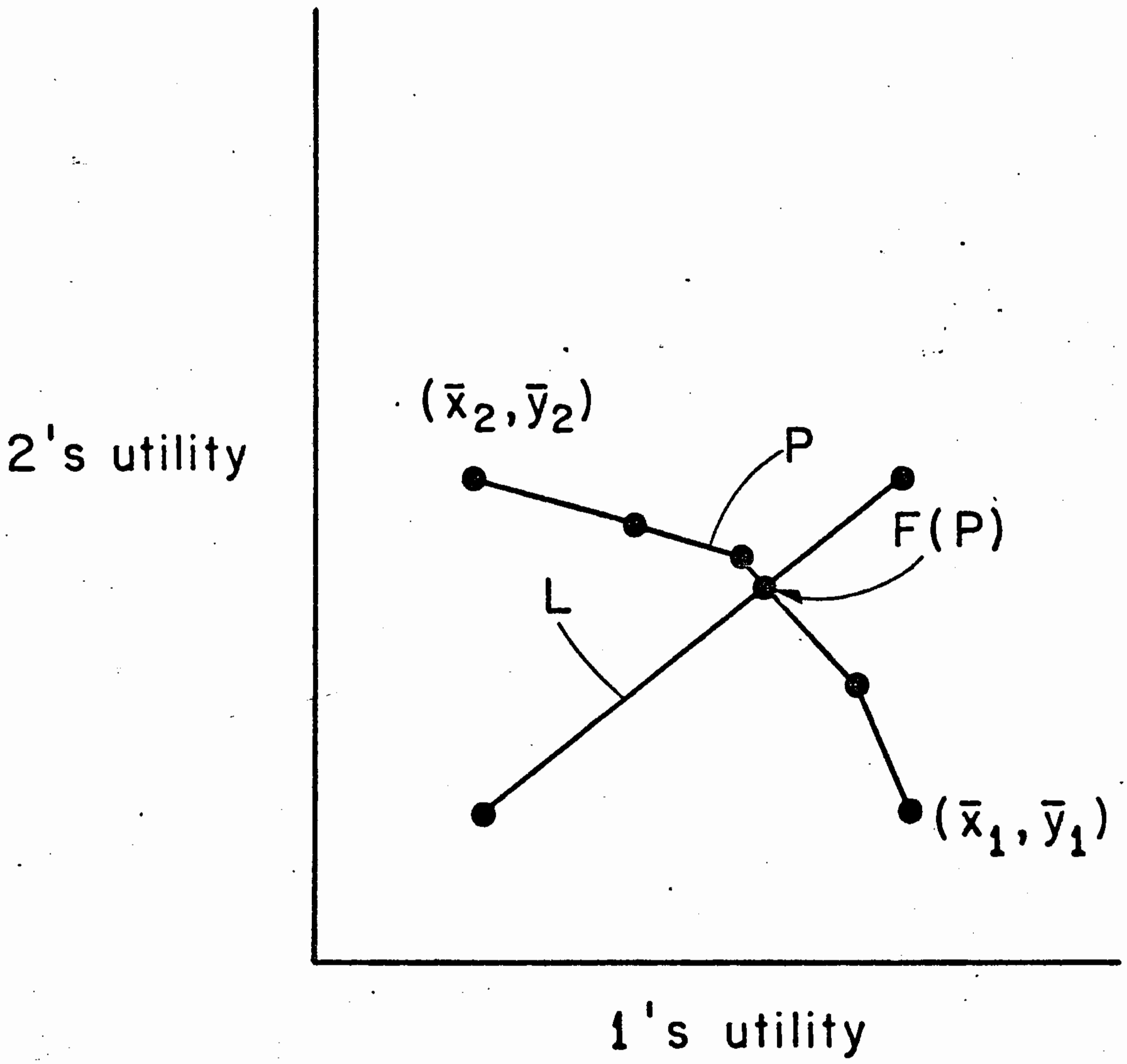

Figure 2

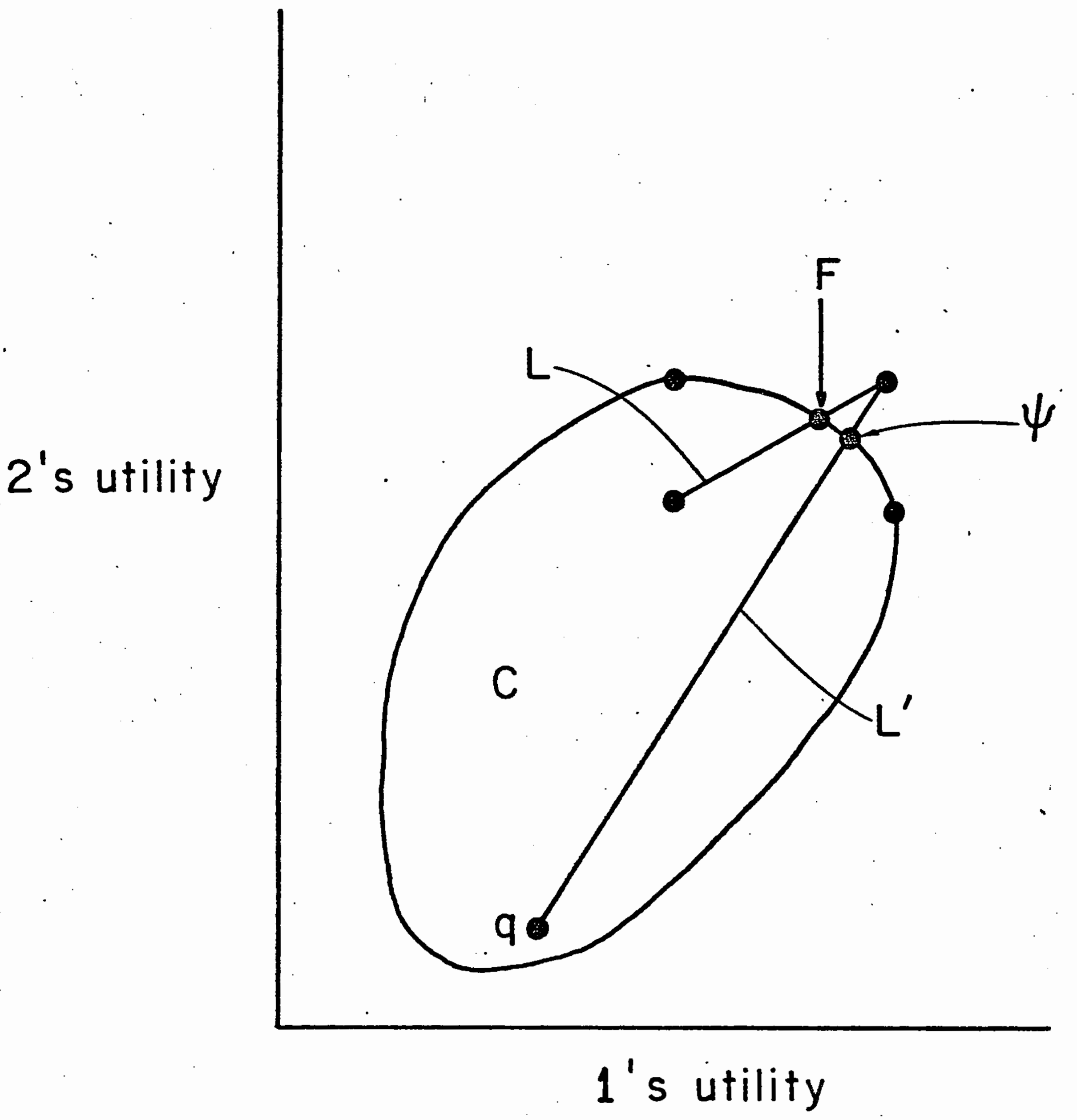

Figure 3