MATHEMATICS OF OPERATIONS RESEARCH
Vol. 2, No. 1, February, 1977
Printed in U.S.A.

INDIVIDUAL RATIONALITY AND NASH'S SOLUTION TO THE BARGAINING PROBLEM*

ALVIN E. ROTH

University of Illinois at Urbana-Champaign

The axiom of Pareto optimality in Nash's definition of a solution to the bargaining problem may be replaced by an axiom of individual rationality, without altering the result.

The classical two person bargaining problem, as set forth by Nash in 1950, consists of a compact convex subset S of the plane, and a point $s \in S$. A point (x_1, x_2) in S represents the von Neumann-Morgenstern utility available to each player as a result of some feasible agreement, and the set S represents the set of all feasible utility payoffs. The point $s = (s_1, s_2)$ represents the utility of the "status quo"—that is, (s_1, s_2) is the payoff to the players in the absence of any agreement.

We will only consider bargaining problems in which there is some possibility of mutual benefit; i.e., problems for which there is some $x \in S$ such that[1] $(x_1, x_2) > (s_1, s_2)$. Since the origin of each player's utility scale is arbitrary, we may without loss of generality normalize each player's utility function so that $(s_1, s_2) = (0, 0)$. Of course the units of each player's utility function may still be varied arbitrarily (see Condition 1 below). Let B denote the class of bargaining problems with status quo at the origin. We will denote elements of B by S, rather than by $(S, 0)$.

Nash defined a solution of the bargaining problem to be a function f defined on B which associates with each bargaining problem a single feasible outcome (i.e., $f(S) \in S$), and which obeys the following four conditions.[2] (We state the first condition somewhat more explicitly than Nash did.)

1. Independence of Linear Transformations. For any bargaining problem S and positive real numbers a and b, if $T = \{(ax_1, bx_2) \mid (x_1, x_2) \in S\}$, then $f(T) = (af_1(S), bf_2(S))$.

2. Independence of Irrelevant Alternatives. If S and T are bargaining problems such that T contains S, and if $f(T) \in S$, then $f(T) = f(S)$.

3. Symmetry. If S is symmetric (i.e., if $(x_1, x_2) \in S$ implies $(x_2, x_1) \in S$) then $f_1(S) = f_2(S)$.

4. Pareto Optimality. If x is a point in S and there exists a point y in S such that $(y_1, y_2) > (x_1, x_2)$, then $x \neq f(S)$.

Nash proved the following.

THEOREM [NASH]. *There is a unique solution satisfying Conditions* 1 *through* 4. *It is the function which associates with each S the element $f(S) = (x_1, x_2)$ of S such that $(x_1, x_2) \geqslant (0, 0)$ and $x_1x_2 > y_1y_2$ for all $y \in S$ such that $y \neq x$.*

The requirement that a solution be Pareto optimal may be thought of as a requirement of collective rationality, since it states that an outcome may not be chosen if there is another outcome which both players agree is preferable. In what follows, we show that the Nash solution may be obtained from a requirement of

* Received August 16, 1976; revised October 20, 1976.
AMS 1970 subject classification. Primary 90D12.
IAOR 1973 subject classification. Main: Games. Cross Reference: Decision.
Key words. Nash solution, bargaining problem, individual rationality.

[1] If x is an outcome in S, (x_1, x_2) denotes its coordinates. We write $(x_1, x_2) > (y_1, y_2)$ to mean $x_1 > y_1$ and $x_2 > y_2$. Similarly, $(x_1, x_2) \geqslant (y_1, y_2)$ means $x_1 \geqslant y_1$ and $x_2 \geqslant y_2$.

[2] See the discussion of the subject in Chapter 6 of Luce and Raiffa.

individual rationality, in place of Pareto optimality. Individual rationality is generally considered to be a more elementary requirement than collective rationality.[3]

The form[4] of individual rationality which we shall require is the following.

5. Individual Rationality. $f(S) > (0, 0)$.

The central result of this paper can now be stated.

THEOREM. *The Nash solution is the unique function defined on B which satisfies Conditions* 1, 2, 3 *and* 5.

Note that Condition 5 says that the solution will always pick a point in the nonnegative quadrant of the plane. So Condition 2 implies that we need only consider the nonnegative part of any bargaining problem. For simplicity we shall henceforth consider only sets S which are contained in the nonnegative quadrant.

It will be useful in the proof to define the *strict Pareto surface* of a set S to be the set of elements x of S for which no element y of S exists such that $y \neq x$ and $(y_1, y_2) \geqslant (x_1, x_2)$. The proof proceeds via the following.

LEMMA. *If f obeys Conditions* 1, 2 *and* 5, *then* $f(S)$ *is an element of the strict Pareto surface of* S.

PROOF. Consider a bargaining problem S (i.e., S is a compact convex subset of the nonnegative quadrant which contains the origin). Let $T = \{(x_1, x_2) \geqslant (0, 0) \mid \exists y \in S$ such that $(y_1, y_2) \geqslant (x_1, x_2)\}$. Then T contains S, and the strict Pareto surface of T is identical to the strict Pareto surface of S.

Suppose that $f(T) = z$ is not an element of the strict Pareto surface of T. Then there exists an element t of T such that $t \neq z$ and $(t_1, t_2) \geqslant (z_1, z_2)$. Let $a = z_1/t_1$ and $b = z_2/t_2$. Then $(a, b) \leqslant (1, 1)$, and $(a, b) \neq (1, 1)$.

The set $T' = \{(ax_1, bx_2) \mid (x_1, x_2) \in T\}$ is contained in T, and T' contains the point z, since $z = (at_1, bt_2)$. So Condition 2 implies $f(T') = f(T) = z$. But Condition 1 implies $f(T') = (az_1, bz_2) \neq z$. This contradiction demonstrates that z is an element of the strict Pareto surface of T. But since T and S share the same strict Pareto surface, z is an element of S. So Condition 2 implies $f(S) = f(T) = z$, and the Lemma is proved.

The proof of the Theorem is now immediate, since the Lemma implies that f obeys Condition 4, so that Nash's Theorem applies. (The result does not depend on the number of players being 2.)

[3] Note, however, that Pareto optimality neither implies nor is implied by individual rationality.

[4] The usual form of an individual rationality requirement allows $f(S) = (0, 0)$. So long as the function f is not permitted to be *identically* equal to the origin, our main theorem continues to hold. However, the lemma used in the proof, which is of interest in its own right, would no longer hold.

References

[1] Luce, R. D. and Raiffa, H. (1957). *Games and Decisions*. Wiley, New York, pp. 114–154.

[2] Nash, John F., Jr. (1950). The Bargaining Problem. *Econometrica* **18** 155–162.

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, DEPARTMENT OF BUSINESS ADMINISTRATION, 350 COMMERCE BUILDING (WEST), URBANA, ILLINOIS 61801