Disagreement in Bargaining

AN EXPERIMENTAL STUDY

MICHAEL W.K. MALOUF
ALVIN E. ROTH
Department of Business Administration
University of Illinois

This study reports an experiment designed to test the predictive value of Axelrod's measure of conflict of interest. The results support the conclusion that Axelrod's measure is a good predictor of the time required to reach agreement in a given bargaining game, but that it is not a good pedictor of the frequency with which disagreements will be observed in a given game. The theoretical implications of this conclusion are discussed.

INTRODUCTION

The study of bargaining occupies an important place in economics in general and game theory in particular, with various parts of the literature having roots at least as far back as the work of Edgeworth (1881), Zeuthen (1930), Hicks (1932), and Nash (1950). The lion's share of the theoretical literature concerns the nature of the agreements which we might expect would be reached in particular kinds of bargaining situations, especially when the bargaining is conducted by suitably motivated and informed bargainers. Although the consequences of *potential* disagreement are often considered to influence the ultimate agreement, little attention has been given to what kinds of bargaining situations are likely to result in actual disagreements.

AUTHORS' NOTE: This work has been partially supported by NSF Grants SOC 78-09928 and SES 79-15356. The authors are also happy to acknowledge helpful comments and advice from Robert Axelrod, Ronald Harstad, and Keith Murnighan. Correspondence and reprint requests may be addressed to Alvin E. Roth, 350 Commerce West, University of Illinois, Urbana, IL 61801.

JOURNAL OF CONFLICT RESOLUTION, Vol. 25 No. 2, June 1981 329-348

One exception to this generalization is the work of Axelrod (1967, 1970), who proposes a measure of the "conflict of interest" in a bargaining situation. This measure is defined for bargaining games of the kind first studied by Nash (1950), and is intended to predict the likelihood of disagreement as well as the amount of other kinds of "conflictful behavior" which will be observed in such games.[1] We regard Axelrod's approach to be a promising one for the study of disagreement, and one which has perhaps received less attention than it merits. The main purpose of this article is to report some experimental results concerning Axelrod's measure, and to suggest some new directions in which the study of disagreements in bargaining might proceed.

Of course, any experimental study of behavior which seeks to test the predictive value of a mathematical theory must be designed with care and interpreted with caution. First, the experiment must be conducted under conditions which are consistent with the assumptions of the theory.[2] Second, any operational assumptions which are required to translate the predictions of the theory into observable outcomes must be explicitly formulated. Only when both these conditions are met can the results of the experiment be interpreted as a test of the theory together with its operational assumptions.

It will therefore be necessary to describe in some detail both the theoretical underpinnings of the theory to be tested, and the experimental procedures to be employed. This article will therefore be organized as follows. The second section will review Nash's (1950) theory of bargaining, together with some recent results which indicate how this theory is related to the occurrence of disagreements. The next section will review Axelrod's work, and the fourth section will review some experimental work concerned with bargaining, with particular attention to recent results which suggest the conditions under which an experimental test of Axelrod's predictions can be conducted most appropriately. The fifth section will discuss such an experimental test and its results, and the final section will consider avenues for further theoretical and experimental work.

1. Cf. Axelrod (1970): 196, 191-2. Some other approaches to the study of disagreement in bargaining are found in Crawford (1979) and Myerson (1980).

2. Conversely, experimental evidence can help establish under what circumstances the assumptions of a theory might be appropriate.

NASH'S MODEL OF BARGAINING

Following Nash (1950), we will consider two-player bargaining games defined by a pair (S, d), where d is a point in the plane, and S is a compact convex subset of the plane which contains d and at least one point x such that $x > d$. The interpretation is that S is the set of feasible expected utility payoffs to the players, any one of which can be achieved if it is agreed to by both players. If no such agreement is reached, then the disagreement point d is the result.

Nash proposed that bargaining between rational players be modelled by means of a function called a *solution*, which selects a feasible outcome for every bargaining game. That is, if we denote the class of all two-player bargaining games by B, a solution is a function $f: B \rightarrow R^2$ such that f(S, d) is an element of S. Nash further proposed that a solution should possess the following properties.

Property 1. Pareto optimality: If $f(S,d) = x$ and $y \geqslant x$, then either $y = x$ or $y \notin S$.

Property 2. Symmetry: If (S, d) is a symmetric game (i.e., if $(x_1, x_2) \in S$ implies $(x_2, x_1) \in S$ and if $d_1 = d_2$) then $f_1(S,d) = f_2(S,d)$.

Property 3. Independence of irrelevant alternatives: If (S, d) and (T, d) are bargaining games such that T contains S, and if $f(T,d) \in S$, then $f(S,d) = f(T,d)$.

Property 4. Independence of equivalent utility representations: if (S, d) and $(\hat{S}, \hat{d})$ are bargaining games such that $\hat{S} = (a_1x_1 + b_1, a_2x_2 + b_2) \mid (x_1, x_2) \in S$ and $\hat{d} = (a_1d_1 + b_1, a_2d_2 + b_2)$ where a_1, a_2, b_1 and b_2 are numbers such that a_1 and $a_2 > 0$, then $f(\hat{S}, \hat{d}) = (a_1f_1(S,d) + b_1, a_2f_2(S,d) + b_2)$.

These properties have been discussed amply elsewhere (cf. Nash, 1950; Luce and Raiffa, 1957; Harsanyi, 1977; Roth, 1977a, 1977b, 1979). Here we will simply note that only Property 4 deals at all with the assumption that the game (S, d) is defined in terms of the von Neumann-Morgenstern expected utility payoffs of the players. Before going on, it will be useful to briefly review the implications of this assumption.

Recall that an individual's utility function u is real-valued function defined on some set of alternatives A. It is a model of choice behavior, in the sense that $u(a) > u(b)$ for two alternatives a and b if and only if a is preferred to b; i.e., if and only if the individual would choose alternative a when faced with the choice between a and b. In 1944, von

Neumann and Morgenstern were the first to demonstrate conditions on an individual's preferences which are sufficient so that choice behavior over risky alternatives is the same as if the individual were maximizing the expected value of the utility function. Such a utility function is uniquely defined only up to an interval scale, which is to say that the origin (zero point) and unit of the utility function are arbitrary. Thus, if u is an expected utility function representing an individual's preferences, then another utility function v represents the same preferences if and only if $v = au + b$, where a is a positive number.

So, Property 4 states that if a game $(\hat{S}, \hat{d})$ is derived from (S, d) by transforming the utility functions of the players to equivalent representations of their preferences, than the same transformations applied to the outcome of the game (S, d) should yield the outcome selected in $(\hat{S}, \hat{d})$. Thus, it states that the solution should depend only on the preferences of the players, and not on any arbitrary features of the utility functions representing those preferences.

Note that, since the game is defined in terms of the players' utility functions (which model their choice behavior), and since the rules of the game allow each player to choose the disagreement point if wished; then any solution f which selects an outcome to which both players might potentially agree must also possess the following property.

Property 5. Individual rationality: $f(S, d) \geqslant d$.

Nash proved the following famous result.

Theorem 1: There is a unique solution which possesses Properties 1-4. It is the solution F defined by $F(S, d) = x$ such that $x \geqslant d$ and $(x_1 - d_1)(x_2 - d_2) > (y_1 - d_1)(y_2 - d_2)$ for all y in S such that $y \neq x$ and $y \geqslant d$.

Nash's solution F selects the outcome which maximizes the geometric average of the gains available to the bargainers over the set of feasible, individually rational outcomes.

Observe that any solution which is assumed to select Pareto optimal outcomes can never yield predictions which reflect the possibility of disagreement, since disagreement is never Pareto optimal for bargaining games in the class B being considered. However, the following recent result shows that the possibility of disagreement can be incorporated into the context of Nash's other assumptions (Roth, 1979).

Theorem 2: There are precisely *two* solutions which possess Properties 2-5. One is Nash's solution F, and the other is the *disagreement solution* D defined by $D(S, d) = d$ for all bargaining games (S, d).

Thus when the assumption of Pareto optimality is relaxed, the possibility of disagreement becomes compatible with Nash's solution, and disagreement is the *only* behavior compatible with that described by Nash's solution. However, since the solution F describes behavior which never leads to disagreement, while D describes behavior which always leads to disagreement, it is clear that we will have to explore different kinds of models if we wish to be able to characterize what properties of games make some games more likely to result in disagreement than others. The next section is concerned with one approach to this question.

AXELROD'S MEASURE OF "CONFLICT OF INTEREST"

Axelrod (1967, 1970) also considers bargaining games (S, d) of the kind considered in the previous section. He defines on the class B a real-valued function which we will denote by A, and calls the quantity A(S, d) the *conflict of interest* of the game (S, d). Before describing the function A, it will be convenient to define, for any game (S, d), the related game (S^+, d), where $S^+ = \{x = (x_1, x_2) \mid d \leqslant x \leqslant y \text{ for some } y \text{ in } S\}$ (see the shaded area in Figure 1).[3]

Let $\bar{x}_1 = \max\{x_1 \mid (x_1, x_2) \in S^+\}$ and $\bar{x}_2 = \max\{x_2 \mid (x_1, x_2) \in S^+\}$, and for each value of x_1 between d_1 and mean x_1, let $\phi(x_1) = \max\{x_2 \mid (x_1, x_2) \in S^+\}$. Then ϕ is the function which defines the upper boundary of S^+ (see Figure 1), and the intervals $[d_1, \bar{x}_1]$ and $[d_2, \bar{x}_2]$ define the range of feasible, individually rational demands available to players 1 and 2, respectively. Of course, not every pair of feasible demands is a feasible outcome (e.g., in general $(\bar{x}_1, \bar{x}_2)$ is not contained in S); which is to say that the rectangle $\{(x_1, x_2) \mid d_1 \leqslant x_1 \leqslant \bar{x}_1, d_2 \leqslant x_2 \leqslant \bar{x}_2\}$ of joint demands always contains S^+ and may be strictly larger than the set S^+ (see Figure 1).

Axelrod defines A(S, d) to be the proportion of the area of the joint demand rectangle which is *not* contained in S^+. This measure is invariant with respect to equivalent utility representations, in the sense that if (S, d) and $(\hat{S}, \hat{d})$ are related as in property 4, then A(S, d) = $A(\hat{S}, \hat{d})$.[4] Consequently, it will be sufficient to confine our attention to

3. Note that (S,d) and (S^+,d) share the same set of points which are both individually rational and Pareto optimal. So any bargaining process which obeys properties 1, 3, and 5 will yield the same outcome for the two games.

4. Axelrod (1970) sketches the outline of an axiomatic characterization of the function A, which requires this form of invariance with respect to equivalent utility repre-

games (S, d) normalized so that $d_1 = d_2 = 0$, and $x_1 = x_2 = 1$. For games normalized in this way, the area of the joint demand rectangle is 1, and so

$$A(S,d) = 1 - \int_0^1 \phi(x_1)dx_1.$$

That is, for games normalized in this way, A(S, d) is equal to 1 minus the area of S^+ (i.e., A(S, d) is equal to the shaded area in Figure 2).

According to this definition, the lowest possible conflict of interest occurs in a game (S, d) for which $(\bar{x}_1, \bar{x}_2)$ is a feasible outcome, since A(S, d) = 0 for such a game (cf. Figure 3a). The highest possible conflict of interest occurs in a game (S, d) in which unrestricted side payments are possible,[5] and A(S, d) = 1/2 for such a game (cf. Figure 3b).

Axelrod (1970: 45) offers the following intuitive justification for why the function A can be interpreted as measuring the conflict of interest. He points out that the greater A(S, d) becomes, "the more the region (S^+) bulges outward, the better both players can simultaneously do, and hence the less incompatible are the goals of the players. For example, if the region bulges a great deal, the players can both get nearly their best payoff, so their conflict of interest is low."

Before reporting an experimental test of this assertion in the fifth section, we will first need to consider the general question of how experiments can be designed to test theories which are stated in terms of the utility functions of the players.

EXPERIMENTAL IMPLEMENTATION OF NASH BARGAINING GAMES

Any experimental test of the predictive value of Axelrod's measure of conflict of interest is faced with many of the same problems involved

sentations. Thus, the function A is intended to depend only on the preferences of the players as expressed in their utility functions, and not on any arbitrary features of those utility functions.

5. Recall that if $(S,d) \in B$ then S must be a convex set, and so for games normalized so $d_1 = d_2 = 0$ and $\bar{x}_1 = \bar{x}_2 = 1$, the feasible set S with the smallest area is the convex hull of the points (0,0), (1,0) and (0,1), as in Figure 3b.

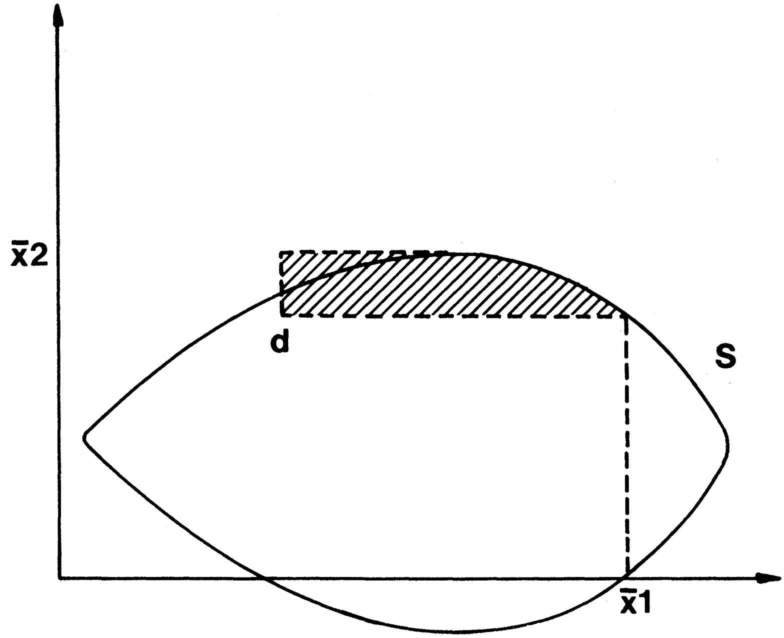

Figure 1

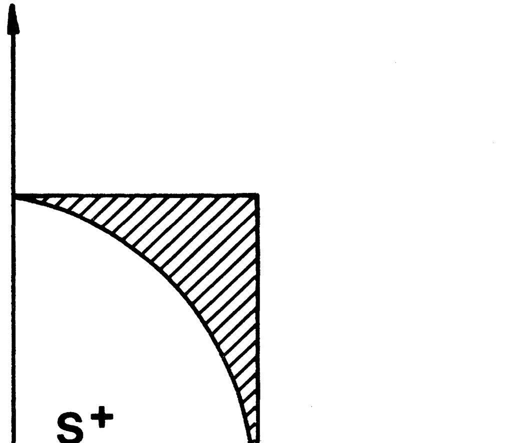

Figure 2

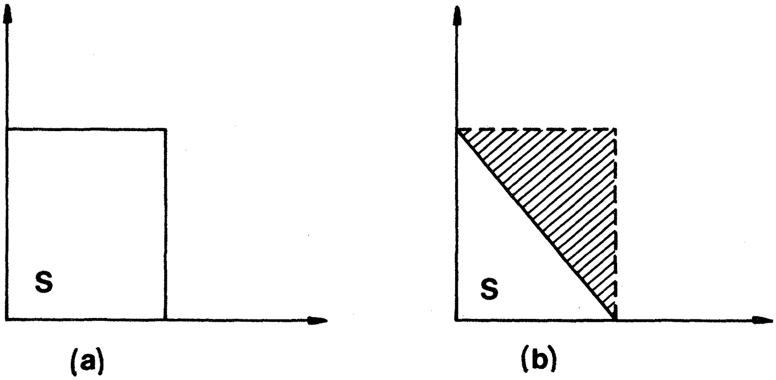

Figure 3

in testing the predictive value of Nash's solution. Both theories are stated with respect to bargaining games defined in terms of the (von Neumann-Morgenstern) expected utility functions of the players, and so they can only be directly tested in experimental situations which permit the expected utility of the bargainers for the underlying payoffs to be known. Furthermore, the assumption of both theories that both players know what game is being played means that each player must be able to determine an opponent's utility. Finally, we shall see that the assumption that the bargaining process is independent of equivalent utility representations places additional restrictions on the information which is made available to the bargainers in an experimental situation. In order to discuss how such experiments may be designed, we first need to consider what it means to determine the expected utility function of a player, over some set A of alternatives.

Consider the case in which the set A of alternatives contains elements a and c such that the player likes a strictly better than c, and for any alternative b ϵ A, the player likes a at least as well as b, and b at least as well as c. Then if u is a utility function representing this individual's preferences over the set of alternatives A, it must have the property that $u(a) \geqslant u(b) \geqslant u(c)$. Since u is defined only up to an interval scale, we may arbitrarily choose its unit and zero point, and in particular we may take $u(a) = 1$ and $u(c) = 0$. The problem of determining u(b) then becomes the problem of finding the appropriate value between 0 and 1 so that all those lotteries over alternatives that the individual prefers to b have a higher expected utility, and all those lotteries to which b is preferred have a lower expected utility. If we denote by L(p) = [pa; (1 – p)c] the lottery that with probability p yields alternative a and with probability (1 – p) yields alternative c, then the utility of participating in the lottery L(p) is its expected utility, pu(a) + (1 – p)u(c) = p. If p is the probability such that the individual is indifferent between b and L(p), then their utilities must be equal, and so, u(b) = p. Thus when we say that the utility of alternative b to a given individual is known, we mean that the probability p is known such that the individual is indifferent between having alternative b for certain or having the risky alternative L(p).

For instance, consider an individual who is faced with a choice of receiving one-half of a million dollars for certain or participating in a lottery that will yield a million dollars with a probability p and otherwise yield zero dollars. If we set the individual's utility function for zero dollars at 0 and the utility for a million dollars at 1, determining the individual's utility for one-half million dollars means determining the probability p that would leave this individual indifferent between the lottery and the one-half million dollars. Most of us would require p to be considerably greater than one-half before we would take the lottery over the assured one-half million dollars, which is to say that our utility is not linear in money, and our utility for one-half million dollars is more than halfway between our utility for zero dollars and our utility for a million dollars. In what follows, when we say that one individual knows another's utility for a given event (e.g., a particular reward), we are not requiring knowledge of any utility theory, but rather that the individual has sufficient knowledge of the other's preferences to be able to determine an equivalent lottery of the sort just described. (For a more complete discussion, see Herstein and Milnor, 1953; Krantz et al., 1971; von Neumann and Morgenstern, 1944.)

Since knowing an individual's expected utility for a given agreement is equivalent to knowing what lottery is thought as desirable as that agreement, then in a bargaining game in which the feasible agreements are the appropriate kind of lotteries, knowing the utilities of the players at a given agreement is equivalent to simply knowing the lottery they have agreed on. In our experiments, therefore, every player i was told about 2 monetary prizes: a large one l(i) and a small one s(i). In each of the games, the players bargained over the *probability* p(i) that they would receive their large prize l(i). Specifically, they bargained over how to distribute "lottery tickets" that would determine the probability that players would win their personal lotteries; i.e., a player i who received 40% of the lottery tickets would have a 40% chance of winning a large monetary prize l(i) and a 60% chance of winning a small prize s(i). In the event that no agreement was reached in the allotted time, each player i received a small prize s(i). In other words, a player would receive the large prize only if an agreement is reached on splitting the lottery tickets and if the ensuing lottery is won. Otherwise, receipt of a small prize is always assured.[6] We will refer to games of this type as *binary lottery games.*

To interpret the set of feasible outcomes of a binary lottery game in terms of each player's utility function for money, recall that if we consider each player's utility function to be normalized so that the utility for receiving the large prize is 1, and the utility for receiving the small prize is 0, then the player's utility for any lottery between these two alternatives is the probability of winning the lottery. That is, an agreement which gives player i p(i) percent of the lottery tickets also gives a utility of p(i).

Note that a change in the prizes is therefore equivalent to a change in the origin and scale of the player's utility functions. This makes it possible to use binary lottery games to experimentally investigate the circumstances under which the bargaining process is indeed independent of equivalent utility representations (property 4) resulting from a change in prizes. Such an experiment is discussed in detail in Roth and Malouf (1979). The results of that experiment support the conclusion that the information shared by the bargainers decisively influences whether the bargaining process is sensitive to a change in the prizes. Specifically, in binary lottery games in which players know only their own prizes, the observed agreements are independent of changes in the prizes of the players.[7] However, in binary lottery games in which both players know both prizes,[8] the observed agreements are *not* independent of equivalent utility representations—i.e., the observed outcomes do not obey property 4. Note that both of these information conditions meet the assumption customarily made about bargaining games, which is that the players' von Neumann-Morgenstern utility for each outcome is known.

6. This experimental design was first introduced in Roth and Malouf (1979), which also includes a review and discussion of previous experimental work in this area.

7. See, however, Roth and Malouf (1980) and Roth et al. (1980) for experimental studies concerned with equivalent utility representations arising in a different way.

8. Specifically, in games in which the size of all the prizes is common knowledge (cf. Roth and Murnighan, 1980).

As we have already noted, Axelrod's measure of conflict of interest is invariant with respect to equivalent utility representations. Thus, it is a measure intended to model conflict in situations where the bargaining process is independent of equivalent utility representations. Therefore, in order to test its predictive value under potentially favorable conditions, the experiment which follows is designed around binary games in which players are informed only of their own prizes. That is, the experiment described below is designed to test the predictive value of Axelrod's measure for games in which the bargaining process has not been observed to violate the property of independence of equivalent utility representations.

EXPERIMENTAL DESIGN

Each player played four binary lottery games in random order, against different opponents. In each game, each player had a small prize of $5.00 and a large prize of $10.00, and each game had restrictions on the manner in which the players could divide the probability of winning their large prizes. In each game, the sum of the percentage of lottery tickets which each player received could not exceed 90%.[9] In game 1 (G_1), player 1 was restricted to a maximum of 60% of the lottery tickets and player 2 was restricted to a maximum of 30%. In games 2 through 4, player 1 was restricted to a maximum of 90% of the lottery tickets, while player 2 was restricted to a maximum of 40%, 50%, and 90% in games 2, 3, and 4 (G_2, G_3, G_4), respectively. These four games are summarized in Table 1, and the set of feasible agreements for each game is depicted in Figure 4.

Note that game 1 has a unique Pareto optimal agreement, so that $A(G_1) = 0$; i.e., there is zero conflict of interest according to Axelrod's measure. Games 2, 3, and 4 have increasing conflicts of interest by Axelrod's measure; i.e., $A(G_2) = .22$, $A(G_3) = .28$, $A(G_4) = .5$. Thus, these four games span the full range of Axelrod's measure, from game 1, which with $A(G_1) = 0$ has the lowest possible conflict of interest, to game 4, which with $A(G_4) = .5$ has the highest possible conflict of interest (see Table 2).

9. In pilot experiments, a large number of (50,50) agreements were observed when players were allowed to divide 100% of the lottery tickets. In order to control for any special, cultural significance of "(50,50)", the players were restricted to divide only 90% of the lottery tickets, so that an equal division would be (45,45).

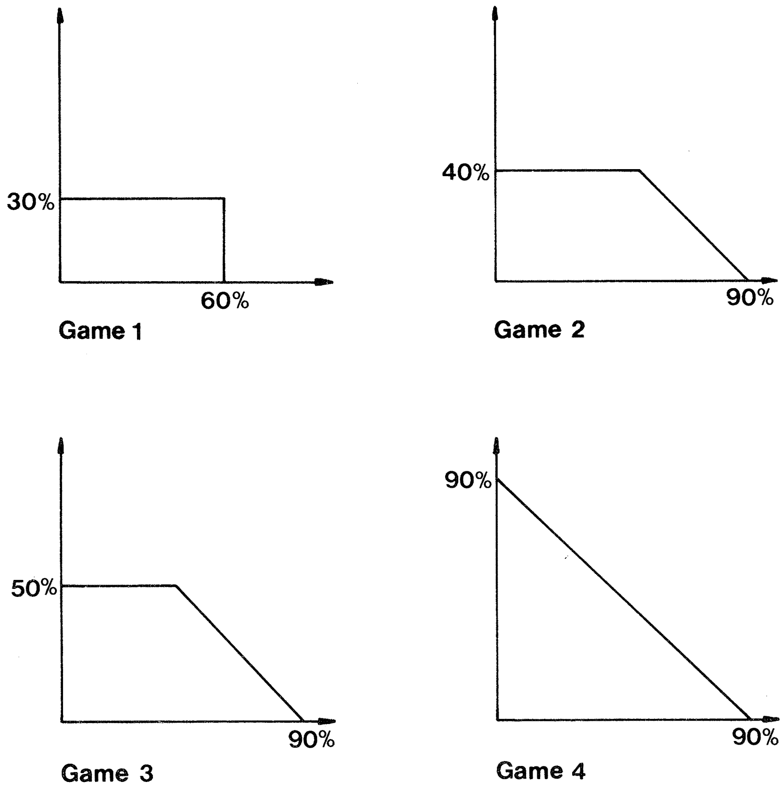

Figure 4

TABLE 1
Games

	Player's 1 Prizes		*Player's 2 Prizes*		*Maximum % Allowed*	
	Small	*Large*	*Small*	*Large*	*Player 1*	*Player 2*
Game 1	$5.00	$10.00	$5.00	$10.00	60%	30%
Game 2	$5.00	$10.00	$5.00	$10.00	90%	40%
Game 3	$5.00	$10.00	$5.00	$10.00	90%	50%
Game 4	$5.00	$10.00	$5.00	$10.00	90%	90%

NOTE: The sum of both players' percentage does not exceed 90%.

TABLE 2
Axelrod's Measure of Conflict of Interest

Games			
1	*2*	*3*	*4*
0	0.2222	0.2777	0.5

Two operational measures were selected against which to test the predictive value of the Axelrod's measure A. The first of these is the percentage of disagreements observed in each of games 1-4; i.e., the frequency with which the players failed to reach agreement in the allotted time. The second measure is the average length of time required to reach an agreement in each of games 1-4; i.e., the elapsed time from the start of negotiations until their conclusion, for every game in which agreement was reached.

METHOD

The experimental set-up employed PLATO, the computer-assisted instruction system that was developed at the University of Illinois. The attractiveness of such a system is its advanced graphic displays and interactive capability.

The participants received all instructions and conducted all communication through visually isolated terminals. The majority of the participants were taking an introductory course in business administration and were receiving credit for their participation in addition to the money they received through bargaining. Although no special knowledge or expertise were required for the experiment, pretests were run using the same subject pool to eliminate any ambiguities in the instructions.

Background information was first presented, followed by the main tools of the bargaining, which consisted of sending messages or sending proposals. A proposal was a pair of numbers: the first was the sender's share of the lottery tickets and the second was the receiver's share. The use of the computer enabled any asymmetry in the presentation to be avoided. The PLATO system also computed the expected value of each proposal and displayed the proposal on a graph of the feasible region. After being made aware of these computations, the bargainer was given the option of cancelling the proposal before its transmittal. Proposals were said to be binding on the sender; an agreement was reached whenever one of the bargainers returned a proposal identical to the one just received.

Messages were not binding. Instead, they were used to transmit any thoughts which the bargainers wanted to convey to each other. To insure anonymity, the monitor intercepted any messages that revealed the identity of the players. The monitor also intercepted messages containing information about the prizes. The intercepted message was returned to the sender with a heading indicating a reason for such action.

To verify their understanding of the procedures, the subjects were given some drills followed by a simulated bargaining session with the computer. As soon as all the participants finished this portion of the experiment, they were paired at random and the bargaining started.

At the end of eight minutes or when agreement was reached (whichever came first), the subjects were informed of the results of that game and were asked to wait until all the other bargainers were finished. For the subsequent game there were new random pairings, and the bargaining resumed. The cycle continued until all games were completed. At no point in the experiment were the players aware of what the other participants were doing, or of the identity of their opponents.

The bargaining process consisted of the exchange of messages and proposals, and participants were instructed that "your objective should be to maximize your own earnings by taking advantage of the special features of each session." All transactions were automatically recorded. After all of the games were played, a brief explanation of the purpose of the experiment was given, and the participants were offered the opportunity to type any comments, questions, and so on, and were directed to the monitor who paid them.

RESULTS

The number and percentage of disagreements are shown in Table 3, while the means and standard deviations of the bargaining time are shown in Table 4. The outcomes of all games are given in Table 5.[10]

Both games 1 and 4 have no disagreements, although Axelrod's measure A placed those two games at opposite extremes. A one-way

10. All the agreements in games 3 and 4 were Pareto optimal, 17 out of the 19 agreements in game 1 were Pareto optimal (89.5%) and so were 13 out of the 15 agreements in game 2 (86.6%).

TABLE 3
Number and Percentage of Disagreements

	Games			
	1	*2*	*3*	*4*
19 pairs	0^{a} (0%)	4^{ab} (21.1%)	7^{b} (36.8%)	0^{a} (0%)

TABLE 4
Means and Standard Deviations for Time to Reach Agreement

	Games			
	1	*2*	*3*	*4*
Excluding Disagreements	158.32^{c} (126.76)	250.73^{cd} (161.29)	215.67^{cd} (128.72)	300.00^{d} (155.24)
(total allowable bargaining time = 480 seconds)				

NOTE: Cells with common superscripts are not significantly different from one another at the 0.05 level using the Newman-Keuls test.

analysis of variance on the number of disagreements shows a significant effect for games, $F(3, 72) = 5.79$, $p < 0.02$. Axelrod's predicted distribution for conflict—assuming that the number of disagreements is proportional to A(G)[11]—was compared with the obtained distribution of the number of disagreements over games. Using the Kolmogorov-Smirnov goodness of fit test, Axelrod's model was rejected ($p < 0.01$) as a predictor of the frequency of disagreement.

A one-way analysis of variance of the time to reach agreement, excluding disagreements, reveals a significant effect for games, $F(3,61) = 3.19$, $p < 0.03$. The games are rank ordered 1, 3, 2, and 4, in order of increasing time to agreement. The results of a Newman-Keuls post hoc test are shown in Table 4. Using the Kolmogorov-Smirnov goodness of fit test, we were unable to reject, at the 5% significance level, Axelrod's model as a predictor of the time to reach agreement.[12]

11. Any alternative hypothesis that is not more than 21% from the proportionality assumption would still be rejected at the 5% level. (Axelrod formally hypothesizes only an ordinal relation.)

12. That is, the results are consistent with the hypothesis that the time to reach agreement in a game G is proportional to A(G).

DISCUSSION

Thus, the results of this experiment are consistent with the hypothesis that the quantity A(G) can serve as a predictor of the average bargaining time for agreements reached in a game G of the kind considered here. However these results lead us to reject the hypothesis that A(G) can serve as a predictor of the frequency with which disagreements will result in such games. This suggests that the characteristics of a game G which influence the frequency of disagreement may not be the same as those which influence the time required to reach a specific agreement, and that A(G) may be more closely connected to the latter characteristics than to the former.

The results of the experiment reported here permit us to speculate on the nature of these different characteristics. To do so, we need to formulate more precisely what characteristics of a game G are measured by the quantity A(G). We can do this as follows.

Suppose that the rules of the game G = (S, d) are that each player i makes a demand x_i, and the payoff to the players is (x_1, x_2) if $(x_1, x_2) \epsilon S^+$ and $d = (d_1, d_2)$ otherwise.[13] That is, the players each get what they demand if their demands (x_1, x_2) are *compatible* (i.e., if $(x_1, x_2) \epsilon S^+$), and a disagreement results if the demands are *incompatible* (i.e., if $(x_1, x_2) \not\epsilon S^+$). Then A(G) is equal to the probability of a disagreement if the players choose their demands *randomly.*

Formally, we can state this observation as follows.

Theorem 3: In a game G = (S,d), if demands x_1 and x_2 are random variables with uniform distributions of the intervals $[d_1, \bar{x}_1]$ and $[d_2, \bar{x}_2]$ respectively, then the probability that x_1 and x_2 are incompatible is equal to A(G).

Proof: Let p denote the probability that the demands (x_1, x_2) are incompatible. Then p is equal to 1 minus the probability that $(x_1, x_2) \epsilon S^+$. But $(x_1, x_2) \epsilon S^+$ if and only if $x_2 \leqslant \phi(x_1)$, where ϕ is the function which defines the Pareto optimal subset of S, as in the third section. So,

$$p = 1 - \int_{d_1}^{\bar{x}_1} \int_{d_2}^{\phi(x_1)} \frac{dx_1}{\bar{x}_1 - d_1} \frac{dx_2}{\bar{x}_2 - d_2}$$

13. These are the noncooperative bargaining rules proposed by Nash (1953).

$$= 1 - \int_{d_1}^{\bar{x}_1} \frac{\phi(x_1) - d_2}{\bar{x}_2 - d_2} \frac{dx_1}{\bar{x}_1 - d_1} = A(G).$$

Specifically, if G is normalized so that $d_1 = d_2 = 0$ and $\bar{x}_1 = \bar{x}_2 = 1$ as in the third section, then

$$p = A(G) = 1 - \int_0^1 \phi(x_1)dx_1.$$

This theorem makes explicit the notion that the quantity A(G) measures the potential "room for disagreement" in a game G, and it illuminates the experimental results obtained. Specifically, the experimental results suggest that the potential room for disagreement is a critical factor in determining the time required to reach a specific agreement in those cases when some agreement is eventually reached, but that other factors strongly influence whether an agreement will in fact be reached, or whether a disagreement will result.

Thus in game G_1, which has a unique Pareto optimal outcome, and for which $A(G_1) = 0$, the average time to reach agreement was short, and no disagreements were observed. For games G_2 and G_3, with more room for disagreement, both the time to reach agreement and the number of disagreements increase. For game G_4, which is completely symmetric but which provides the maximum room for disagreement ($A(G_4) = .5$), the average time to reach agreement was the longest of the four games, but no disagreements were observed.

This observation is consistent with the relatively low frequency of disagreements observed for completely symmetric games in other experiments,[14] and it strongly supports the conclusion that abundant room for disagreement in a game G, as measured by A(G), is not in itself sufficient to produce a high frequency of disagreements. This is not too surprising in light of the fact that we have shown A(G) to be the probability of disagreement which would result if players chose their

14. Cf. Roth and Malouf (1979), Malouf (1980). However the data from those other experiments was insufficient to give statistically significant results concerning the frequency of disagreement.

TABLE 5
Summary of Final Agreements

	Game 1			*Game 2*			*Game 3*			*Game 4*		
	Player			*Player*			*Player*			*Player*		
Group	*1*	*2*	*Time*	*1*	*2*	*Time*	*1*	*2*	*Time*	*1*	*2*	*Time*
1	30	30	261	40	40	83	45	45	173	45	45	435
	60	30	178	0	0	–	0	0	–	45	45	479
	60	30	208	50	40	33	45	45	86	45	45	297
2	60	30	420	50	40	475	50	40	343	45	45	237
	60	30	144	0	0	–	0	0	–	45	45	474
	60	30	87	50	40	479	0	0	–	45	45	151
	60	30	179	55	35	81	40	50	179	45	45	474
3	60	30	48	50	40	109	45	45	100	45	45	252
	60	30	59	50	40	131	46	44	384	45	45	118
	60	30	18	54	36	241	0	0	–	45	45	305
	60	30	39	0	0	–	0	0	–	45	45	143
4	60	30	29	50	40	162	45	45	121	45	45	68
	40	30	154	39	39	375	50	40	64	45	45	240
	60	30	382	59	31	464	0	0	–	45	45	457
	60	30	75	50	40	202	45	45	168	45	45	421
5	60	30	58	50	40	142	45	45	170	45	45	127
	60	30	217	50	40	394	0	0	–	45	45	471
	60	30	69	0	0	–	50	40	406	45	45	475
	60	30	383	50	40	390	45	45	394	0	90	76

demands randomly, since we do not expect that rational players will choose their demands in this manner.[15]

Table 5 strongly suggests that the symmetry of game G_4 played an important role in determining the outcome, since all but one[16] of the agreements reached in that game were symmetric. Thus, a theory capable of predicting the frequency of disagreement in such games will very likely have to incorporate, at least indirectly, some measure of a game's symmetry. Since all but one of the agreements in game G_4 were

15. It can be shown, however, that a rational, utility-maximizing player who believed that an opponent chose a demand randomly (as in Theorem 3) would respond by demanding the payoff given by Nash's solution (cf. Roth, 1979: 25-28).

16. The player who received 0 in the (0,90) agreement later explained that she had typed the proposal by mistake.

identical, it is somewhat surprising that so much time was needed to reach them. This serves to emphasize the significance of the room for potential disagreement, as measured by $A(G_4)$, in determining the time required to reach agreement.

REFERENCES

AXELROD, R. (1970) Conflict of Interest. Chicago: Markham.

——— (1967) "Conflict of interest: an axiomatic approach." J. of Conflict Resolution 11: 87-99.

CRAWFORD, V. P. (1979) "A theory of disagreement in bargaining." Discussion Paper 79-3. University of California, San Diego.

EDGEWORTH, F. Y. (1881) Mathematical Psychics. London: Routledge and Kegan Paul.

HARSANYI, J. C. (1977) Rational Behavior and Bargaining Equilibrium in Games and Social Situations. Cambridge, England: Cambridge Univ. Press.

HERSTEIN, I. N. and J. W. MILNOR (1953) "An axiomatic approach to measurable utility." Econometrica 21: 291-297.

HICKS, J. R. (1932) The Theory of Wages. London: Macmillan.

KRANTZ, D. H., R. D. LUCE, P. SUPPES, and A. TVERSKY (1971) Foundations of Measurement, Volume I. New York: Academic Press.

LUCE, R. D. and H. RAIFFA (1957) Games and Decisions: Introduction and Critical Survey. New York: John Wiley.

MALOUF, M.W.K. (1980) "An experimental investigation of two-person bargaining." Ph.D. dissertation, University of Illinois.

MYERSON, R. B. (1980) "Solutions for two-person bargaining problems with incomplete information." Discussion Paper 432. Northwestern University.

NASH, J. F. (1953) "Two person cooperative games." Econometrica 21: 128-140.

——— (1950) "The bargaining problem." Econometrica 28: 155-162.

ROTH, A. E. (1979) Axiomatic Models of Bargaining. New York: Springer-Verlag.

——— (1977a) "Individual rationality and Nash's solution to the bargaining problem." Mathematics of Operations Research 2, 1: 64-65.

——— (1977b) "Independence of irrelevant alternatives and solutions to Nash's bargaining problem." J. of Econ. Theory 16, 2: 247-251.

——— and M.W.K. MALOUF (1980) "Scale changes and shared information in bargaining: an experimental study." (mimeo)

——— (1979) "Game-theoretic models and the role of information in bargaining." Psych. Rev. 86: 574-594.

——— and J. K. MURNIGHAN (1980) "Strategic and sociological variables in models of bargaining." (unpublished)

ROTH, A. E. and J. K. MURNIGHAN (1980) "Common knowledge and deniability in bargaining." (unpublished)

von NEUMANN, J. MORGENSTERN, and O. MORGENSTERN (1953) Theory of Games and Economic Behavior, 3rd ed. Princeton: Princeton Univ. Press.

ZEUTHEN, F. (1930) Problems of Monopoly and Economic Warfare. London: Routledge and Kegan Paul.