legitimately expected by each party), and for the ubiquitous failure of educational institutions to recognize, much less reward, excellence in teaching. Additional understanding of students' expectations of teachers will be necessary for further exploration of the validity of this mental model, as will more sharply focused analyses of the assumptions that seem to underlie the development of policies and personnel practices on American campuses.

REFERENCES


Mid-American Review of Sociology

SOLVING' THE BARGAINING PROBLEM

James A. Schellenberg
Indiana State University


The bargaining problem is here conceived as determining a point of final agreement in bilateral bargaining situations where there is an overlap in the interests of the parties. Several formal models for describing how persons may solve the bargaining problem (in particular, those of Nash, Kalai and Smorodinsky, and Felsenthal and Diskin) are briefly reviewed, and three experiments are described which seek a comparative test of these models. Experimental results fail to provide clear support for any of these formal models, but they do lead to a more general description of how bargainers tend to arrive at cooperative agreements. This is expressed in terms of the three central considerations of (1) prominence, (2) social efficiency, and (3) equity.

INTRODUCTION

There is a substantial literature in economics and other social sciences on what has been called "the bargaining problem." This literature concerns determining the point where bargainers are likely to come together for a deal that they find mutually acceptable. The present paper is intended to discuss some experimental work which helps clarify the nature of such bargaining activity.

THE BARGAINING PROBLEM

Perhaps an example will help us see what is usually meant by the bargaining problem. A seller lists a house for sale at $91,000 but actually would be happy to sell it at $85,000. In fact, she is not sure that she will be able to get that much. A prospective buyer particularly likes the house, however, and wants to buy it. He wants to buy it so much that he would be willing to pay up to $95,000--$4,000 more than the asking price. But of course he knows that you don't offer more than the seller asks. Although he would like to get the house at as low a price as possible, he also wants to make his bid high enough that serious negotiations will begin. He decides to offer $83,000. Knowing what we know about the real interests of these house bargainers, we expect that they will in the end negotiate a successful sale. But what will be the final sale price? Somewhere between $83,000 and $91,000--but where? And what determines where this deal will be struck?

Traditional economic theory predicts with fair precision where sale prices will be in competitive markets, that is, where many buyers and sellers face one another to make trades concerning quantities of a reasonably homogeneous product. However, in cases of bilateral monopoly--where only a single buyer
faces a single seller—we have an indeterminacy problem. Assuming that their interests overlap so that there is a range of possible agreements where both parties would be better off with a deal than with no agreement, we can predict that they will come to an agreement. But precisely where this agreement will be is left indeterminate. Labor-management wage negotiations, for example, often have this characteristic feature, and this adds to the importance of the bargaining process which removes the indeterminacy. Some of this same indeterminacy carries over to situations within a market system where competitive forces are present but are limited in their application; our example of selling a home illustrates such a condition of imperfect competition. While this illustration is not quite a case of bilateral monopoly, there remains an uncertainty about where the interests of buyer and seller will come together because of the special features of their interdependence concerning this particular house.

As we move from economic exchange (where the terms of trade are stated quantitatively and in monetary terms) to forms of social exchange (where the coin of the realm may be only crudely quantified attention, approval, and influence), the indeterminacy problem remains. It is less clearly recognized because the bargaining process is less explicit. But a wide range of social situations can be seen as rough parallels of the bilateral monopoly pattern. Two sweethearts trying to decide on a wedding date, a therapist and client trying to negotiate a recognized "reality" about the client's behavior, a mother and daughter trying to decide what the daughter is to wear, or a husband and wife planning a vacation trip—in all of these examples interests may be opposed, but not so much opposed as to foreclose mutual agreement. The agreement will typically be implicitly stated rather than formally signed, and of course the terms of trade will not be clearly quantified. Nevertheless, if we look closely we will probably find the essentials of bilateral bargaining: a mutual recognition of the desirability of some agreement, initial positions which differ significantly, and movement to a final agreement which lies between the initial positions.

As an approximate definition, we may say that the bargaining problem is that of determining a point of final agreement in bilateral bargaining situations where there is an overlap in the interests of the two parties. In this sense, "the bargaining problem" becomes simply the resolution of the indeterminacy existing in the overlap of interests of two parties. However, this process cannot be neatly reduced to a single problem. There are, as Cross (1969) has pointed out, really three closely intertwined problems. There is, first of all, the problem of determining the point at which an agreement is finally made. Secondly, there is the process of making concessions through which an agreement is typically reached. And, finally, there are the more general conditions which set the stage for negotiations.

To take our example of the sale of a home, we have the problem of determining precisely what the sale price is to be. However, this in turn is closely related to the process of negotiations taking place over the sale. Finally, there is the more general set of external conditions in which negotiations take place, including such possibilities as a change in job opportunities for the buyer, unexpected difficulties in getting a bank loan, or a discovery of a leak in the roof of the house.

**SOME MODELS**

Frederik Zeuthen (1930) posed the bargaining problem in the context of labor-management wage negotiations. If the negotiations break down, there will be a conflict or "fight" in the form of a strike or a lockout. Avoidance of this conflict is assumed to be the motivation for making concessions in such negotiations. It is further assumed that both parties can make rational calculations of the costs to them of conflict, and that these costs can be compared to the net benefits which would be realized in any proposed agreement. These calculations in turn form the basis of a proneness to make concessions in bargaining.

Concessions in negotiations can be conceptualized as a series of discrete steps. At any point, we may predict which party is more likely to make a concession on the basis of the relative costs of conflict. The question, in other words, becomes: how willing is each to risk a conflict? This question is in turn evaluated in relation to one's current bargaining position and the likelihood that insistence on maintaining it may result in conflict. In Zeuthen's model it is the combination of a current position's potential value and the likelihood that maintaining it will result in conflict that determines the probability of making further concessions.

Although von Neumann and Morgenstern (1944) did not identify a unique solution for the bargaining problem, their theory of games presented a framework which has been used for most subsequent analyses in this and closely related areas. It was not long after the publication of their primary work that the theory of games was used as a framework to fashion a unique solution to the bargaining problem. Indeed, several different solutions have been proposed. Of these, the one which is generally considered the most successful is that put forward by John F. Nash.

In his initial paper titled "The Bargaining Problem," Nash (1950) identified a set of utility theory assumptions derived from the work of von Neumann and Morgenstern. He then specified several further assumptions to identify a solution. Nash's assumptions are usually summarized as consisting of the following four:

1. The assumption of **invariance with respect to utility transformations**. This holds that a solution for one bargaining game should also be the solution for any other game in which the utilities of either player (or both) are changed by a direct linear transformation of those of the original game. In other words, simply changing the utility scale values should not change the solution.
2. The assumption of **independence of irrelevant alternatives**. This holds that when a solution for one game is also a feasible outcome for a second
game contained within the first (in the sense that all feasible outcomes of the second are also included in the first), then this solution should also apply to the second game.

3. The assumption of symmetry. This holds that when the positions of the players are symmetrical, then the solution must be one which provides consideration of the available utility combinations should affect the solution.

4. The assumption of Pareto optimality. This holds that an outcome cannot be a solution when there is another feasible outcome where one of the players could do better and the other would do no worse.

With just these four assumptions, we find that only one solution to the bargaining problem can be identified. That is, for every two-person mixed-motive game, if it can be played cooperatively with Nash’s assumptions, a unique solution results. This unique solution has characteristics which can be summarized thus: it is the outcome which represents the maximum product of the gains in utilities of the two parties as they approach their mutually beneficial possibilities from their point of origin (that is, from the "conflict point," where they would find themselves if there is to be no agreement).

Several years after Nash presented his model, John Harsanyi (1956) demonstrated a surprising relationship of that model to the one presented earlier by Zeuthen. Harsanyi took Zeuthen’s basic model and added a few simplifying assumptions to make it fit with the more formal style that had emerged with game theory. Probably the most important of the assumptions added was that each party had perfect knowledge of the other’s willingness to risk ultimate conflict (that is, each knows not only its own mind, but also how eager the other is to reach an agreement). Thus fortified and refined, the Zeuthen model was shown by Harsanyi to produce a solution to the bargaining problem which is mathematically equivalent to that of Nash. It is rather surprising that these two models would have this result. Zeuthen’s is rich in psychological flavor about motivations regarding the making of concessions, while Nash’s, a paragon of mathematical simplicity and elegance, avoids any inclusion of psychological forces. Nevertheless, Harsanyi quite clearly demonstrates their mathematical equivalence, so far as a solution point for bilateral bargaining is concerned.

Let us note well the intellectual triumph which had seemed to have been achieved by the end of the 1950s. Game theory had not only produced a solution concept for all two-person zero-sum games with its minimax theorem; it had now also produced a unique solution for the two-person bargaining game, that is, for the situation in which the motives for the participants are mixtures of opposed and common interests. The solution was based on reasonable axioms, irrefutable in its proof, and the same as that which could be derived from a more process-based model, namely in Zeuthen’s earlier theory of collective bargaining processes. Even today the Zeuthen-Nash-Harsanyi solution concept is often considered the appropriate theoretical resolution for the bargaining problem. There are, however, several problems with this conclusion. One problem is the presence of still other solution concepts provided by other theoretical models. Another problem is the relatively low order of empirical support for this particular solution concept. Let us take up briefly each of these matters in turn.

Although Nash’s solution to the bargaining problem early became the favorite among game theorists, there were other solutions also put forward. One of these was first presented by Howard Raiffa, later more fully developed in its axiomatic structure by Ehud Kalai and Meir Smorodinsky, and further discussed and applied by Douglas Heckathorn. Following Heckathorn (1978), we can call this the "equal resistance" solution, for it assumes the point of final resolution to be that point among mutually favored (or "Pareto optimal") outcomes where both parties give up the same proportion of the distance between the "conflict point" (which would result if no agreement is obtained) and their "best hope."

As Kalai and Smorodinsky (1975) demonstrate, this model rests on an axiomatic structure similar to that of Nash. Only Nash’s second assumption (the independence of irrelevant alternatives) is not fully followed for this model. It is replaced with another assumption (called the "monotonicity axiom") which, though it appears less demanding, still produces a unique solution.

More recently, Dan Felsenthal and Abraham Diskin (1982) have argued for still another solution, based on what they call the "minimum utility point." Their basic argument is that the "conflict point" (where the parties would end with no agreement) is not the most reasonable base for determining where the agreement should be reached. Departing from both the Nash and the equal resistance models on this issue, they hold that the worst possible agreement should provide the minimum point for each party’s bargaining. This worst possible agreement may often be the conflict point, but sometimes it would be better for one or both of the parties, for the range of feasible agreements would be all Pareto-optimal points that for both parties are equal or better than the conflict point. If an agreement is to be reached, they reason, why not limit our attention to those points which involve feasible agreements, rather than to focus on the threat of no agreement?

In many concrete cases the same set of outcomes would be identified as the point of solution for all of these models—that of Nash and those of the equal resistance and minimum utility theories. However, though the practical differences are small, these three solution concepts (as well as several others to be found in the literature) are mathematically distinct.

Theoretical discussions about the bargaining problem have been more common than empirical tests of the various models. There have been some studies designed to test one or another of these models—more for the Nash model than the others, since it has received more and longer attention. However, even for the model most applied to empirical predictions, Nash’s, the number of well designed studies remains small. Furthermore, their results are quite varied; some studies support Nash’s predictions, while others clearly...
do not. And none of the other models have enough supporting evidence to claim a superior empirical status to that of Nash.

Given the inconclusive state of the empirical evidence for such models, we decided to undertake some further experimental work—work which would provide a comparative test of three leading mathematical models (those of Nash, Kalai and Smorodinsky, and Felsenthal and Diskin) as solutions to the bargaining problem.

THREE EXPERIMENTS

We here will briefly review three experiments designed to study some central aspects of bargaining, including solutions to the so-called bargaining problem. We will summarize the most critical aspects of this research, with the reader referred to the original report (Schellenberg 1988) for further details.

A very simple 2 x 2 game matrix (that is, a situation for interdependent choice in which two players would each have just two options available) was devised. We will be calling this our "M" matrix (after some incidental content in the verbal scenario which was first used to illustrate it). Although not precisely the same as the Prisoner's Dilemma, this matrix had a structural dilemma similar to that well-known game. When used with repeated plays and an opportunity to negotiate the choices to be planned for such a known number of plays, the M matrix has the result of predicting different results for each of the following three theoretical models: those of (a) Nash, (b) Kalai and Smorodinsky, and (c) Felsenthal and Diskin.

<table>
<thead>
<tr>
<th>Matrix M</th>
</tr>
</thead>
<tbody>
<tr>
<td>= = = = = = = = = = =</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Column</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
</tr>
<tr>
<td>C2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Row</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
</tr>
<tr>
<td>R2</td>
</tr>
</tbody>
</table>

As is shown by Matrix M (presented above), outcomes for subjects were broken down into very simple terms—identified only as positive (+), negative (-), or neutral (0). The positions of subjects are those of either the row player (whose outcomes are shown in front of each comma for each of the four possible outcomes of the matrix) or of the column player (shown after the comma for each possible case).

Situations were presented in quite abstract terms, with subjects simply encouraged to seek positive outcomes and avoid negative ones. Eventually these outcomes were translated to amounts of money the subjects received for participation in the experiment. It should be recognized that this abstract framework would not represent any particular kind of naturalistic bargaining situation. However, it was intended that some of the general tendencies implicit in all bargaining might be better explored in this way than in a less abstract mode of research.

In all of our experiments we used student volunteers, two at a time, who were placed on opposite sides of a table with a visual barrier placed between them (to prevent their knowing each other’s choices for a particular trial before these were announced by the experimenter). Subjects were given a series of pencil-and-paper exercises for over an hour, with several variations of the M matrix included in the materials presented.

Our first experiment was a rather straight-forward attempt to determine which of our three theoretical models would best predict repeated play results with the M matrix. Our results, however, failed to give consistent support for any of the three models we had set out to test. The behavior of some pairs of subjects did fit well with one of the models, but other pairs fit another model, and some fit none of the three models—but still had their own fairly consistent pattern. What we seemed to have here was a surplus of solutions—practically every pair had its own approach to a solution. But these solutions did not appear to generalize in any main direction. At least there could be no clear conclusion to favor any of the three models we had set out to test.

There was a strong temptation to end the research at this point, leaving another experiment (among many) unpublished because of its inconsistent results. But struck with the rather apparent fact that our subjects were finding solutions for their bargaining problem—their own solutions, which for most pairs had a strange internal consistency—we decided to explore further what may be behind these solutions.

In our second experiment we took a different tack. In the first experiment we had given the M matrix four sets of ten trials, following an opportunity for negotiation about how the pair of subjects might coordinate their actions for the trials. In this second experiment we gave them a single trial with the possibility of gaining (or losing) up to 10 cents, but allowed them to agree on plans for making, if they so desired, side payments. Actually, we had three conditions for presenting the M matrix to each pair of subjects in this experiment: (1) in which pairs made repeated choices without an opportunity for an advance discussion with each other; (2) in which repeated choices were made after an opportunity for negotiation, and (3) in which a single choice—but with side payments possible—was made after an opportunity for advance negotiations.

Quite different results were obtained for each of these three conditions. It was not particularly surprising that under the first condition (without an
opportunity for discussion), persons chose in terms of short-run interests. Nor was it surprising that the second condition (repeated choices after an opportunity for discussion) showed most of the same varied solutions as had been found in our first experiment. What is most noteworthy in results for this second experiment is the high consistency of results for the third (side payments possible) condition, which had not been present in the previous experiment. Here most of our pairs of subjects showed exactly the solution which would have been predicted by the Felsenthal-Diskin model. But why was this model especially successful here, but not under other conditions?

At first glance it would appear that there should be no difference between (a) presenting our M matrix ten times with an opportunity to win (or lose) up to 1 cent per trial, without side payments possible, and (b) presenting our M matrix one time with an opportunity to win (or lose) up to 10 cents, but with side payments possible. In both cases, actual play follows an opportunity for discussion and cooperative planning. In both cases the same matrix is presented. In both cases there is the theoretical possibility of gaining (or losing) up to 10 cents. However, a more thorough formal analysis may demonstrate that the actual options for possible negotiated agreements were different in the two cases. The condition in which side payments were possible was accompanied by some further possibilities for the combined distribution of rewards. Technically, there was an expansion of the "negotiation set" or "Pareto-optimal frontier"; this meant, in effect, that there was a wider range of mutually positive agreements possible in the side-payments-possible condition than under other conditions. We may call this expansion of the bargaining range an "exchange bonus." The main finding of our second experiment was the demonstration that this exchange bonus yielded a different pattern for resolution of the bargaining problem than had previously appeared.

We have mentioned that the particular pattern of results for our side-payments-possible condition was a close fit with the solution suggested by the Felsenthal-Diskin model. Was this a happy accident (a result of a correspondence of their solution with an especially prominent side payment of exactly half of the gains) or real evidence to support this particular model? To help unravel this question (as well as several other questions which had arisen), we planned a third experiment.

This experiment included presentation of the M matrix under the same main conditions as in both of the first two experiments. It also included presentation of another matrix, which we may call the "N" matrix (see next page), under all of the same conditions. The N matrix had exactly the same set of contingencies as the M matrix, only they were arranged differently, leading to a different "natural outcome" (or expected result without an agreement). The comparisons between results for the two matrices allowed us to determine the extent to which the natural outcome helped to determine the solutions that our subjects actually came up with.

Solving the Bargaining Problem

Matrix N

<table>
<thead>
<tr>
<th>Column</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
</tr>
<tr>
<td>C2</td>
</tr>
</tbody>
</table>

Row

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>-0</td>
</tr>
<tr>
<td>R2</td>
<td>+,-</td>
</tr>
</tbody>
</table>

It is impossible to give a good brief summary of all the results of our third experiment, including some especially interesting answers to questions during our debriefing procedures. However, it is fair to say that this experiment (a) tended to confirm the main findings of the first two experiments, and (b) failed to give results which would provide consistent support for any of the three models which we initially set out to support. Furthermore, the main reason that these models were unsuccessful was not to be found in their fine-tuned differences, but in a general assumption that they all had in common. All three models assumed that the solution to the bargaining problem was to be determined primarily with reference to a status-quo point. Felsenthal and Diskin identified a different status-quo point than the other two models, but all followed mathematical procedures which were anchored in a primary reference point. On the other hand, what we found our subjects emphasizing was not so much a beginning reference point as a focus upon the possibilities before them and how they might be most equitably shared. They did not pay much attention to any particular point of departure. They did not particularly let where they would end up with no agreement dictate where they actually would come together.

ANOTHER MODEL

Although we set out initially to provide a comparative test for three formal models for solving the bargaining problem, our results led us to question the adequacy of them all. The question then arises as to whether there is another model which can better fit the results we obtained. To say that there is such a model may be misleading, for certainly none exists that can do so which has either the elegance or precision of the models previously mentioned. Nevertheless, if we can accept a broader view of theoretical models to include word pictures which are not clearly defined mathematically, then it is possible that such a model can at least be given in outline form.
There are three main considerations which together appear to determine the kinds of agreements made by subjects in our experiments. These may be summarized as those of (a) prominence, (b) social efficiency, and (c) equity.

Our subjects sought solutions which would be easy to conceptualize and translate into behavior. Where such prominent solutions existed that also met standards of social efficiency and equity, they tended to be used. The even split of gains where side payments were permitted stands as a clear example of this. In cases where no prominent solution existed which met the other main considerations, subjects either worked out an agreement which was less easy to formulate and apply or else they settled on another prominent solution which was at the expense of either social efficiency or equity.

By the criterion of social efficiency, we mean that agreements tend to be sought that serve the interests of both parties. In our simple experiments, this meant that solutions should be Pareto-optimal, that joint outcomes are not satisfactory if anyone could do better without making someone else worse off. In more naturalistic bargaining settings the criterion of Pareto-optimality may be too demanding, for as Cross (1983) has pointed out, the parties may not have enough knowledge about the possible alternatives to be able to realize their best joint outcomes. Still, it is a reasonable assumption that outcomes with higher joint gains will be preferred over those where one or both parties will realize less of their interests.

We assume, finally, that persons seek agreements that they consider to be equitable. Under the conditions of our experiments, this equity was primarily expressed as simple equality. In more complex settings we would expect that considerations of relative contributions or prior status might be recognized to have a major influence on what is considered equitable, though it appears to us that the Nash and Kalai-Smorodinsky models claim a much greater generality for such an influence than is warranted. In any event, some notion of fairness needs to apply to an agreement in order for it to be fully satisfactory.

We do not assume that persons necessarily emphasize these three considerations of prominence, social efficiency, and equity in equal measure. Indeed, a good number of our experimental subjects appeared to emphasize two of these at the expense of the third. But no pairs in our experiments reached stable patterns without at least two of these considerations applying in a major way. It would be our general suggestion that bargainers generally seek solutions in which all three of these considerations apply, and that only if they are unsuccessful in this attempt is one or another of these considerations downgraded. We would never expect to find a stable agreement that is not supported by at least two of these three main considerations.

ENDNOTES

1. For a fuller discussion, see Schellenberg (1988) and Schellenberg and Druckman (1986). The present paper is based on these two earlier articles. The author is especially indebted to Daniel Druckman for his collaboration in some of this earlier work.

2. In the accompanying illustration of Matrix N, the lower right cell represents the natural outcome; while for the previously presented Matrix M, the upper left cell was the natural outcome. The only difference
between these two matrices leading to this change in the natural outcome is the exchanged locations of the ", +, +" and ", +, +, +" cells.

3. In this respect, the Felsenthal-Diskin model was closer to our results than the other two models, but other results also led us to question whether the Felsenthal-Diskin substitute for a status-quo point was not also too limiting.

REFERENCES