

MichU
DeptE.
CenREST
W
89-15

Center for Research on Economic and Social Theory
CREST Working Paper

Do People Exploit their Bargaining Power?
an Experimental Study

Ken Binmore
Peter Morgan
Avner Shaked
John Sutton

January, 1989
Number 89-15



DEPARTMENT OF ECONOMICS
University of Michigan
Ann Arbor, Michigan 48109

Do People Exploit their Bargaining Power?
An Experimental Study

by

Ken Binmore, Peter Morgan,
Avner Shaked, and John Sutton.

Current version: January, 1989

Address. Ken Binmore, Econ Dept, University of Michigan, Ann Arbor, MI 48109

Do People Exploit their Bargaining Power? An Experimental Study

Ken Binmore, Peter Morgan,
Avner Shaked, and John Sutton.

1. Introduction.

Human behavior, even in simple bargaining situations, is not well understood. Numerous rival theories compete for attention, but the data is seldom adequate to justify a rejection of one in favor of another. It is often not even clear what the significant control variables are.

This paper seeks to shed light on only one small aspect of the problem. It describes an experimental attempt to compare the predictions of two qualitatively distinct types of theory. Firstly there are the fairness/focal theories of bargaining behavior as propounded by Güth [1988], Kahneman *et al* [1986], Roth [1985], Selten [1978] and others. Secondly, there are the strategic theories of bargaining behavior, notably that of Rubinstein [1982]. Game-theoretic or strategic models treat the bargainers as rational optimizers and hence predict that the players will exploit whatever bargaining power they find at their disposal. The fairness/focal view discounts the impact of strategic considerations on behavior. Instead, the agreement on which subjects agree is seen as being determined by social norms or conventional understandings that render the agreement focal, given the circumstances under which the bargainers are working. In a bargaining context, the social norms often involve “fairness” considerations, but other features of the situation may also be important. For example, deals involving whole numbers of dollars may be salient under some circumstances.

We are grateful to the Economics and Social Research Council of the United Kingdom for generously funding this research project. We would also like to thank P. Knox and S. Chew for efficiently programming the experiments, and A. Hoolighan, A. Klin, C. Mirrlees, C. Purkhardt, and B. Thakker for their invaluable help in supervising the experiment and recruiting the subjects. We are also grateful to the Psychology Department at the London School of Economics for the use of their laboratory.

It is not easy to distinguish fairness/focal behavior from strategic behavior. Indeed, part of the message of this paper is that what people perceive as fair or focal may sometimes be a function of the strategic realities of the situation. However, we found it possible to design two simple laboratory games that, superficially at least, seem very similar from a fairness/focal viewpoint but which differ significantly in their strategic characteristics.

To summarize the results of the experiments very briefly, the subjects' behavior was biased in the direction of strategically optimal play. Under one of the two conditions, the differences in behavior between the two types of game was very marked indeed. The same turned out to be true of what the subjects asserted to be "fair" when questioned on this issue *after* playing the game.¹ That is to say, what they judged to be "fair" after experiencing actual play was biased in the direction of the outcome that would result from strategically optimal behavior in the game they had actually played.

Peter Cramton [1988] has run the same experiment using our computer programs with Yale undergraduates as his subjects. His conclusions will be reported elsewhere. We note only that they are broadly consistent with ours, although his subject population was only half the size of ours.

2. Bargaining models

The basic problem for the subjects in all the models considered is that of dividing a sum of money² that we call a "cake". If the negotiations break down, player I will receive a payment which is equivalent to a share α of the cake and player II will receive a share β of the cake, where $\alpha + \beta \leq 1$. What is "fair" in such a situation? Three possible answers to this question will be distinguished for special attention:

¹ Individuals in the same pool from which the subjects were drawn but who did *not* play the game were also surveyed on the "fairness" question. Their responses were too dispersed to be usefully reported (although there was no significant difference between what was said to be "fair" in the two types of game). Our guess is that people do not have a preconceived view of what is "fair" in these situations and hence the survey results from subjects with no experience of playing the game are largely "noise".

² We proceed throughout on the questionable assumption that utility can be identified with money.

(a) **Split-the-Difference (S-T-D)**. With this outcome, player I is assigned α , player II is assigned β and then they split the remainder of the cake equally. Another way of saying this is that the outcome is the Nash bargaining solution with the *status quo* located at (α, β) . Player I's final share is then $(1 + \alpha - \beta)/2$ and player II's share is $(1 + \beta - \alpha)/2$.

(b) **Fifty-Fifty (50:50)**. With this outcome, the breakdown payments α and β are ignored and each player simply gets half the cake.

(c) **Deal-me-Out (D-M-O)**. With this outcome, the breakdown payments are ignored and the result is 50:50, *unless* this would assign player i less than his or her breakdown payoff of γ . If so, player i gets γ and the other player gets the remaining $1 - \gamma$ of the cake.

The term deal-me-out derives from a previous paper [Binmore/Shaked/Sutton,1988] and is intended to suggest player i 's response to the proposed implementation of 50:50 when $\gamma > 1/2$. Its possible role as a "fairness" criterion was suggested by critics of this previous paper.

In all the games considered, α was taken to be very small ($\alpha = 0.04$). Two values of β were considered: a high value ($\beta = 0.64$) and a low value ($\beta = 0.36$). Figure 1 provides a convenient means of comparing the three different notions for different values of β (but with α fixed at 0.04).

To discuss *strategically optimal* play, it is necessary to be specific about the bargaining procedure to be used. We employ a procedure studied by Rubinstein [1982]. The players alternate in making proposals indefinitely until a proposal is accepted or the negotiations break down. Some incentive is necessary to encourage the players to reach an early agreement. The two classes of games considered differ in how this incentive is provided and in how breakdowns may occur.

(A) **Games with Optional Breakdown**. In these games, a player may *opt out* after refusing a proposal³ made by the opponent (and only then).

³ It matters *when* a player may opt out [Shaked,1987]

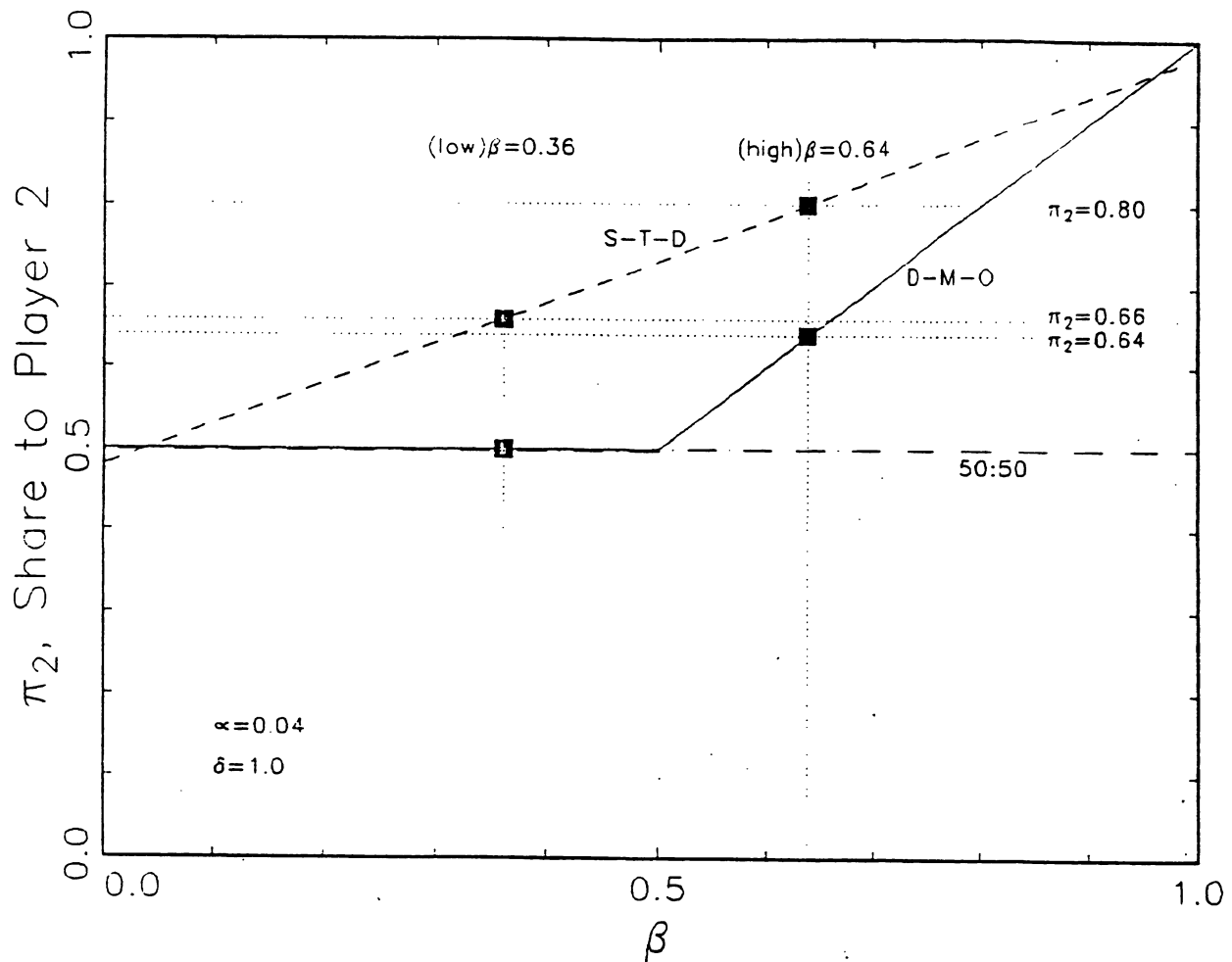


Figure 1.

The "fairness" criteria.

If a player opts out, the negotiations are deemed to have broken down, and the players receive their breakdown shares, α and β , of the current cake. The incentive for an early agreement is provided by the fact that the cake shrinks by a factor of δ ($0 < \delta < 1$) immediately before each proposal after the first.

A subgame-perfect equilibrium analysis [Binmore/Shaked/Sutton,1988] yields a unique equilibrium in which player I offers a share equal to

$\max\{\beta, \delta/(1 + \delta)\}$ to player II at time 0, and player II accepts. Notice that the equilibrium outcome converges to D-M-O as $\delta \rightarrow 1-$.

(B) **Games with Forced Breakdown.** In these games, players may not choose to opt out and the cake does not shrink. Instead, after each refusal, a random move decides whether the negotiations will be broken off, or whether they will be allowed to continue. The probability of a continuation is taken to be the same value of δ as in games of class A. The incentive for early agreement is therefore that the cake may disappear altogether if the negotiations are prolonged, leaving each player with only their breakdown payment.

A subgame-perfect equilibrium analysis [Binmore/Shaked/Sutton,1988] yields a unique equilibrium in which player I offers a share equal to $\{\delta(1 - \alpha) + \beta\}/(1 + \delta)$ to player II at time 0, and player II accepts. Notice that the equilibrium outcome converges to S-T-D as $\delta \rightarrow 1-$.

Each of these two classes of games was run under two conditions: low β and high β . This yields four different games that will be referred to as regimes 0, 1, 2, and 3 as indicated in Table I. Thus the labels 0 and 1 refer to class A games with low and high β respectively, and the labels 2 and 3 refer to class B games with low and high β respectively.

Regime 0 $\beta = 0.36$	Regime 1 $\beta = 0.64$
Class A: Optional breakdown	Class A: Optional breakdown
Regime 2 $\beta = 0.36$	Regime 3 $\beta = 0.64$
Class B: Forced breakdown	Class B: Forced breakdown

Table I.

Two points should be noted. The first is that a subgame-perfect analysis predicts future behavior even if past behavior has not been as predicted. Our earlier work [Binmore/Shaked/Sutton,1988] on class A games indicates that one should not expect instant agreement at time zero from subjects in the laboratory, but that there is reason to believe that the game-theoretic prediction of the final outcome, in terms of the cake then available, may not fare too badly. Results are therefore always reported in terms of player II's share of the cake available *at the time the game ended*. When $\delta \rightarrow 1-$, the game-theoretic prediction will always be D-M-O in class A games, and S-T-D in class B games. The second point is the more important. For a given value of β , class A games and class B games are intended to present a similar payoff profile to the subject, who may therefore be inclined to treat them in the same way in deciding what is or is not "fair" or focal. Indeed, since a subgame-perfect analysis of a class A game is *identical* to that of a class B game when $\alpha = \beta = 0$, one might expect even a strategically minded but inexperienced subject to fail to recognize the rather subtle distinction between the two classes of game. If subjects behave differently in class A games from the way they behave in class B games, one would therefore seem to have evidence in favor of the players' bargaining power being a significant factor in determining the final outcome. It is this consideration that provides the major motivation for the experimental design described in this section.

3. The Experiment.

Subjects were recruited directly from undergraduate classes in economics at the London School of Economics. The students had not studied game theory or bargaining, nor were these topics part of the curriculum for the courses they were attending. The recruiters were graduate students in Psychology who also supervised the fully automated experimental runs in the Psychology Laboratory. Subjects were informed that the experiment was "in economics" rather than "in psychology", but were not informed of the identities of the authors of this paper. As to the details of the game itself, our intention was that the subjects' information be perfect in respect of the rules of the game and the monetary payoffs to be distributed.

The main experiment ran for six days. Each day had four sessions, with each session devoted to a different regime from Table I. Each session involved four subjects, who each played ten games in all. After each game, a subject's opponent was changed. On arrival, subjects were seated in isolated booths with a minimum of interaction between them. They communicated via microcomputers linked by a local area network. First they were asked to read written instructions (appendix B), and then to operate a demonstration program that provided them with hands-on experience on how to make and respond to offers, and so on. The cake was represented on the screen by a rectangular slab. The subject made an offer by moving a "knife" up or down the cake until satisfied with the division it indicated. The monetary amounts being proposed were also displayed. Note that the demonstration program did *not* involve any sample partitions of the cake proposed by us or by the computer, since we were anxious not to create any avoidable focal points.

After running the demonstration program and asking any questions they might have,⁴ the subjects then played six games that were described as "practice games" for which no payments were made. They then played four "real games" in each of which the cake was worth £5.00. We felt this figure to be sufficiently high to provide an adequate incentive for the subject to devote some care and attention to the experiment, given that we were only asking for 30–45 minutes of their time. (At the time the experiment was run, a bland but nutritious meal could be purchased in the student cafeteria for £2.30.) Since each group of four subjects in a particular session played a total of twenty games altogether (12 for practice and 8 for real) and since each regime was in force on each of six days, we observed a total of 120 games for each of the four regimes (72 for practice and 48 for real).

To minimize on reputation effects, the subjects' bargaining partners were changed after each game. Their role in the game also varied. That is to say, half the time they occupied the role of player I (who moves first and has a breakdown payment of α) and half the time they occupied the role of player II (who moves second and

⁴ A supervisor could be summoned by pressing an appropriate key.

has a breakdown payoff of β). We attach importance to this alternation of roles in game-theoretic experiments. A rational player bases his strategic analysis of a game on the way he would play if he were in the shoes of the opponent. Alternating roles provides subjects with a painless opportunity to see things from the other player's viewpoint and hence to understand the game better. In a previous experiment [Binmore/Shaked/Sutton,1986], such role-switching proved to influence the outcome very markedly.

At the end of each session, subjects were asked to remain seated until they had completed a questionnaire (appendix C) and had been paid the total amount of money they had successfully bargained for in the four real games that each had played. They were then invited to leave one by one with a view to minimizing interaction.

Under all the four regimes of Table I, the cake was worth £5.00 in the main experiment and the parameter δ was taken to be 0.9. In all four regimes, player I's breakdown share of $\alpha = 0.04$ was therefore initially worth £0.20 in money.

Under regimes 0 and 1, the games were of class A. That is to say, the cake shrinks over time according to the discount factor δ and it is left to the players' discretion whether to force a breakdown by opting out. Under regime 0, player II's breakdown share of $\beta = 0.36$ was initially worth £1.80. Under regime 1, player II's breakdown share of $\beta = 0.64$ was initially worth £3.20.

Under regimes 2 and 3, the games were of class B. For these games, the cake does not shrink, but there is a risk of an imposed breakdown every time that an offer is refused. Our intention was that the players should believe that the game continues after a refusal with probability $\delta = 0.9$, but here we met with a difficulty in our pilot experiments. The manner in which we sought to resolve this difficulty requires some explanation.

In our initial pilot, the written instructions described the probabilistic mechanism by means of which breakdown occurred and, after each refusal, subjects saw a beautifully simulated roulette wheel turn on their screens. Nevertheless, they tended

to behave as though the possibility of a breakdown ever occurring was negligible.⁵ That is to say, they neglected to note that, although 0.9 is nearly 1, $(0.9)^n$ is small when n is sufficiently large. Such misconceptions about probabilistic matters are, of course, commonplace as laboratory phenomena.

After various attempts, we sought to evade the difficulty by telling the subjects, in their written instructions, that the maximum length for each game had already been chosen in advance, but that they were not to be told what this length was. However, they were invited to proceed on the assumption that, after each refusal, the probability of the game continuing was 0.9. The precise wording was as follows:

In each of the ten sessions, the number of proposals allowed before a breakdown is announced has been fixed in advance. But we are going to keep you guessing by not telling you what these ten numbers are. All you will know for sure is that a breakdown will occur eventually if agreement is delayed long enough. The maximum number of proposals allowed in each session may be large or it may be small, and knowing what the number turns out to be in one session will not help much in guessing what it will be in another. The numbers have been fixed so that, however many proposals there may already have been in a session, you should still reckon that there is a 90% chance of being allowed at least one more proposal. This means, for example, that it is more likely that 12 or more proposals will be allowed than 3 or less.

No subject expressed any confusion about the issue on their questionnaire. We chose the following numbers to be the maximum lengths for the ten games that each subject played:

9, 2, 11, 2, 10, 7, 7, 16, 12, 8.

The two short games were intended to bring to the subjects' attention that breakdown could indeed occur. Otherwise, our intention was that the data available to the subjects should not be such as to allow them rationally to reject the hypothesis

⁵ And, in many cases, confirmed this interpretation of their behavior by their comments on the questionnaire.

that breakdowns occur independently with probability 0.1, even if they participated in games that always ended in breakdown.⁶

4. Results

The raw data appears as appendix A. In this section, the data we believe to be relevant is summarized in six histograms (figures 3, 4 and 5). Box and whisker diagrams will be found in figure 8. We always work in terms of percentages of the cake obtained by player II. (In class A games, the cake shrinks over time. The percentage of the cake is then computed in terms of the cake available at the time the game ended.) Games that do not end in agreement are indicated by an empty box. Player II then gets his or her breakdown share β (36% under regimes 0 and 2; 64% under regimes 1 and 3).

The immediate issue is whether, in view of the similarity of their payoff profiles and the subtlety of their strategic differences, class A and class B games generate the same behavior, although a subgame-perfect analysis predicts D-M-O in the former and S-T-D in the latter. Before considering this question, some preliminary comments will be useful:

How big is epsilon?

Game theory treats players as rational optimizers, who are assumed to squeeze the last penny from a situation on the assumption that their opponents will do the same. But, in practice, one must accept that subjects will treat small enough amounts as negligible. As a rule-of-thumb, we proceed as though anything less than the price of a cup of coffee (£0.2=4% of £5.00 at the time of the experiment) is negligible. In particular, we neglect the fact that the strategically optimal outcomes with the

⁶ We do not, of course, believe that the subjects did carry out any elaborate probabilistic calculations. It is enough for our purposes if the subjects are convinced that the game will end eventually but that it is unlikely to do so immediately.

actual discount factor used ($\delta = 0.9$) differ slightly⁷ from those in the limiting case when $\delta \rightarrow 1-$.

Round number focal points.

A further source of possible distortion is the tendency of subjects to settle on deals in round numbers. Under regimes 1 and 3 (with $\beta = 64\%$), this tendency makes S-T-D attractive, since 80% of the cake is £4.00 at time 0. Under regimes 1 and 3, attention needs also to be drawn to the possible existence of a focal point at 70% which corresponds to £3.50 at time 0.

Under regimes 0 and 2 ($\beta = 36\%$), the possible existence of a round number focal point at 60% (£3.00 at time 0) needs to be taken into account.

Disagreements

In a class B game played for money, the bargaining is unilaterally terminated by the computer after a minimum of seven offers have been rejected. Note that such an outcome is *not* compatible under regime 3 with D-M-O when the latter is regarded as a “fairness” criterion, because player I then only gets α instead of his “fair” share of $1 - \beta$. Disagreements in class B games therefore convey no information relevant to this study, beyond the fact that disagreements do indeed occur even though they are not predicted. The same goes for class A games under regime 0.

However, disagreements in class A games under regime 1 are significant. A subgame-perfect analysis in such games predicts that player I will offer player II approximately $\beta = 64\%$ at time 0, and that this will be accepted. But a rational player II who is not offered the amount he deems to be rational and therefore fears that his rational counter-offer may be refused by a possibly irrational opponent has good reason to opt out immediately, thereby ensuring his subgame-perfect payoff.

⁷ For example, under regime 0, a subgame-perfect analysis predicts 47.42% for player II when player I makes the first offer, and 52.67% when player II makes the first offer. These are both approximated by the 50% predicted by D-M-O.

For this reason, we count disagreements under regime 1 as being supportive of the game-theoretic prediction (namely D-M-O) rather than dismissing them as “noise”.

Returning to the data summarized in figure 2 for regimes 1 and 3 (with $\beta = 64\%$), we observe that the difference of behavior between class A and class B games is very marked indeed. A fairness/focal theory that took no account of strategic issues would therefore receive no support from this data. Statistical analysis appears in section 7. At this point we note only that game theory predicts the observed behavior better than the alternatives listed in section 2. It is not surprising that 50:50 does not do well when player II can get 64% without the consent of his partner,⁸ but it is instructive that S-T-D predicts very much better than D-M-O in class B games, while D-M-O predicts better than S-T-D in class A games.

The position for regimes 0 and 2 (with $\beta = 36\%$) is summarized in figure 3. Here the differences between class A and class B games do not seem significant. The round number focal point at 60% (£3.00 at time 0) is perhaps responsible for producing this result, since it lies roughly midway between D-M-O (50%) and S-T-D (66%).

We therefore ran regimes 0 and 2 again, but with the £5.00 cake replaced by an \$11 cake. Subjects were told that their dollar winnings would be paid to them in pounds sterling at the then current exchange rate. Otherwise the circumstances of the experiment were identical. The point of doing this was to ensure the existence of *two* round number focal points (at \$6 and \$7) between the D-M-O prediction of 50% and the S-T-D prediction of 66%. Figure 4 shows the sharper data obtained. The differences between games of class A and class B are statistically significant (section 7). Game theory cannot be said to predict this data well, but it does better than the alternatives being considered.

⁸ Although a number of workers have observed systematic violations of individual rationality in related contexts.

Figure 2(a): Regime 1

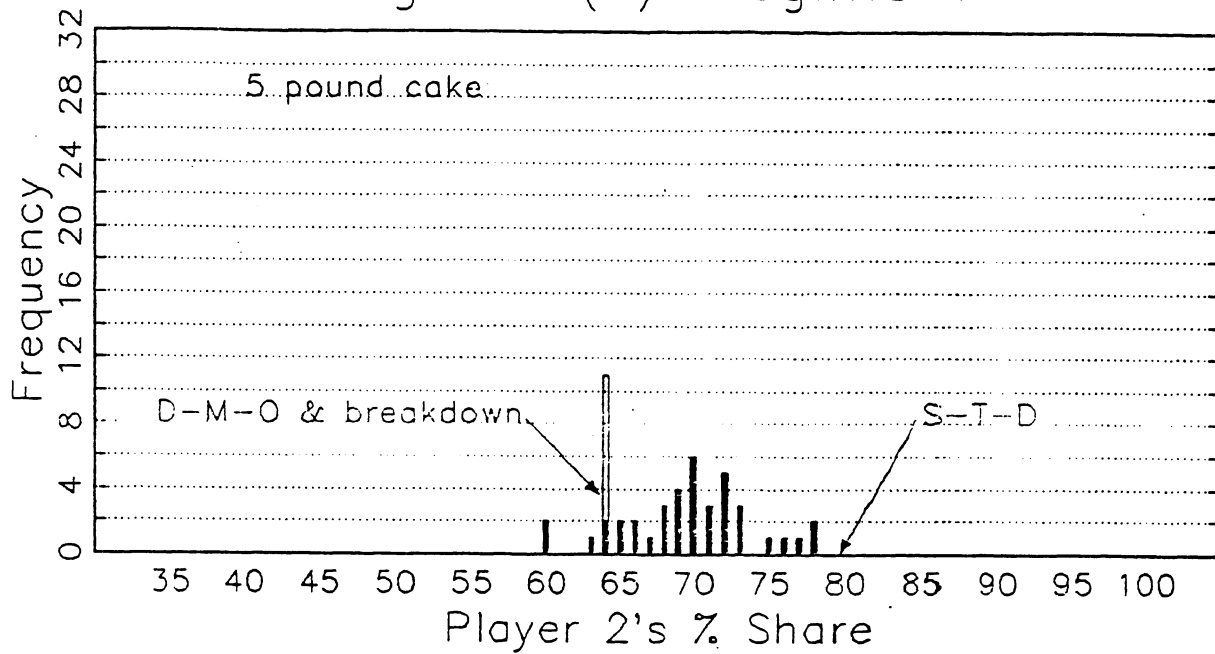


Figure 2(b): Regime 3

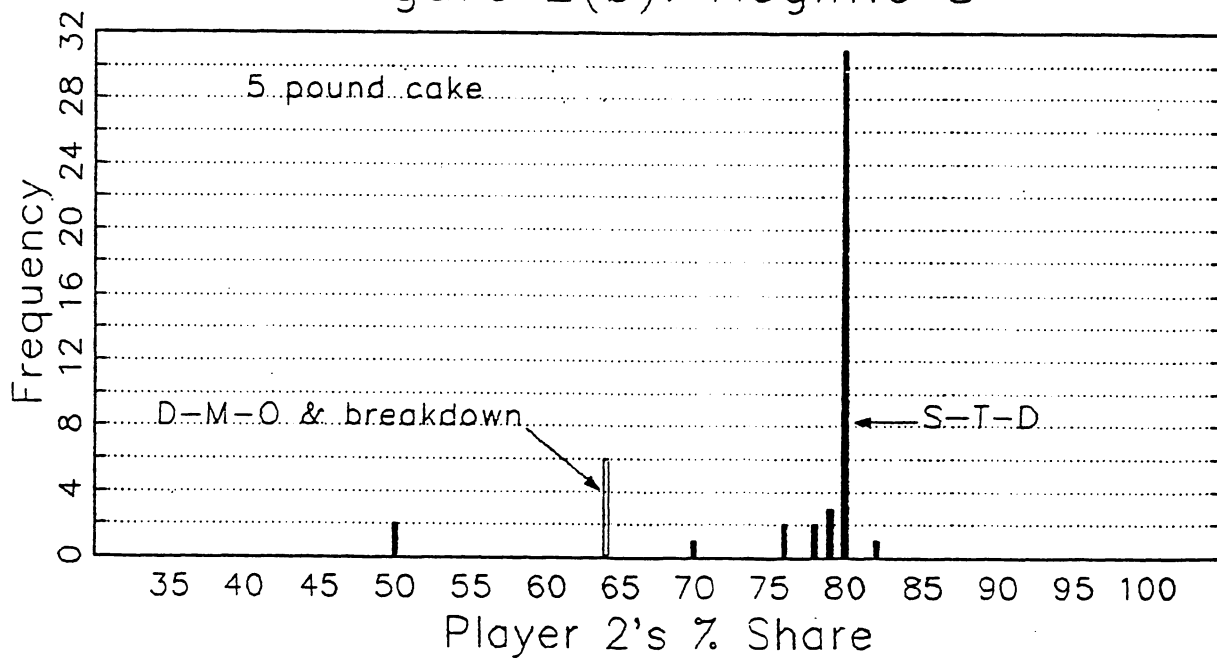


Figure 2.

This figure compares the paid games of class A and class B with high β for the £5 cake.

Figure 3(a): Regime 0

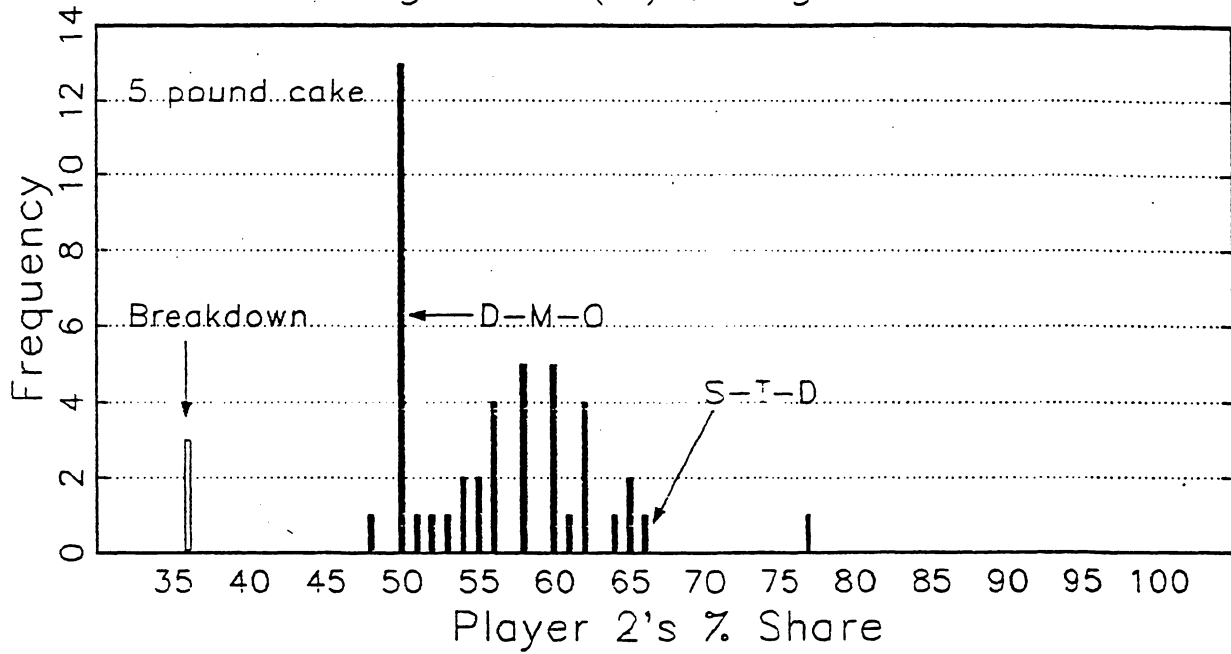


Figure 3(b): Regime 2

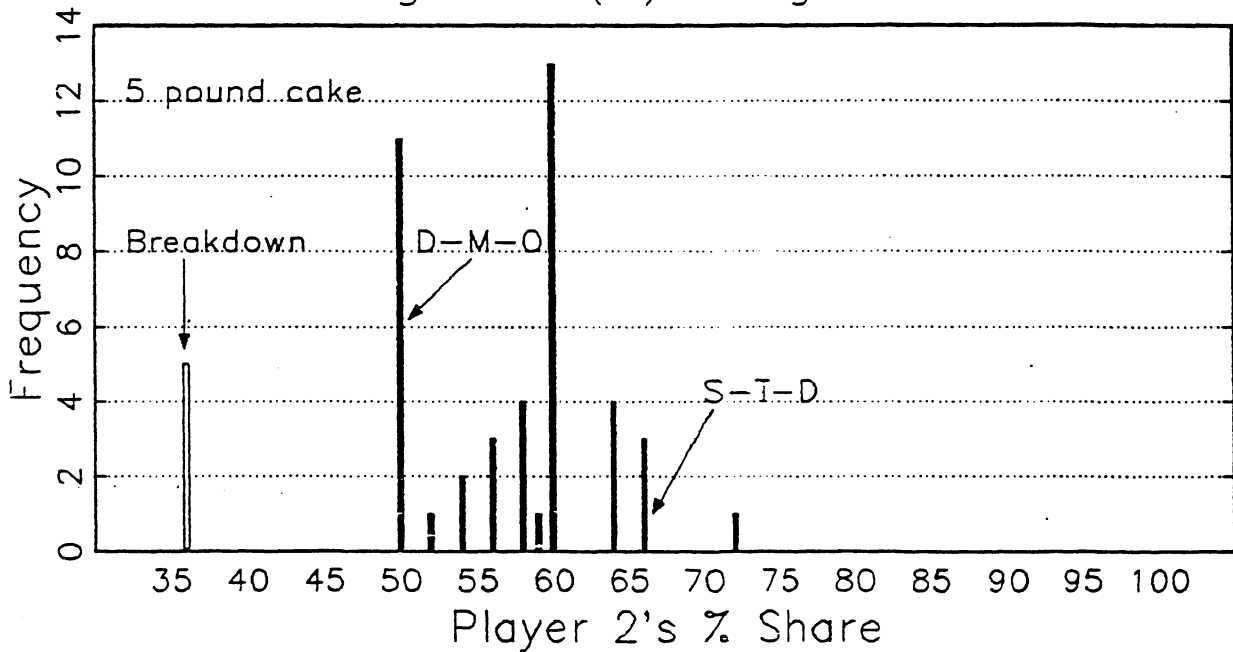


Figure 3.

This figure compares the paid games of class A and class B with low β for the £5 cake.

Figure 4(a): Regime 0

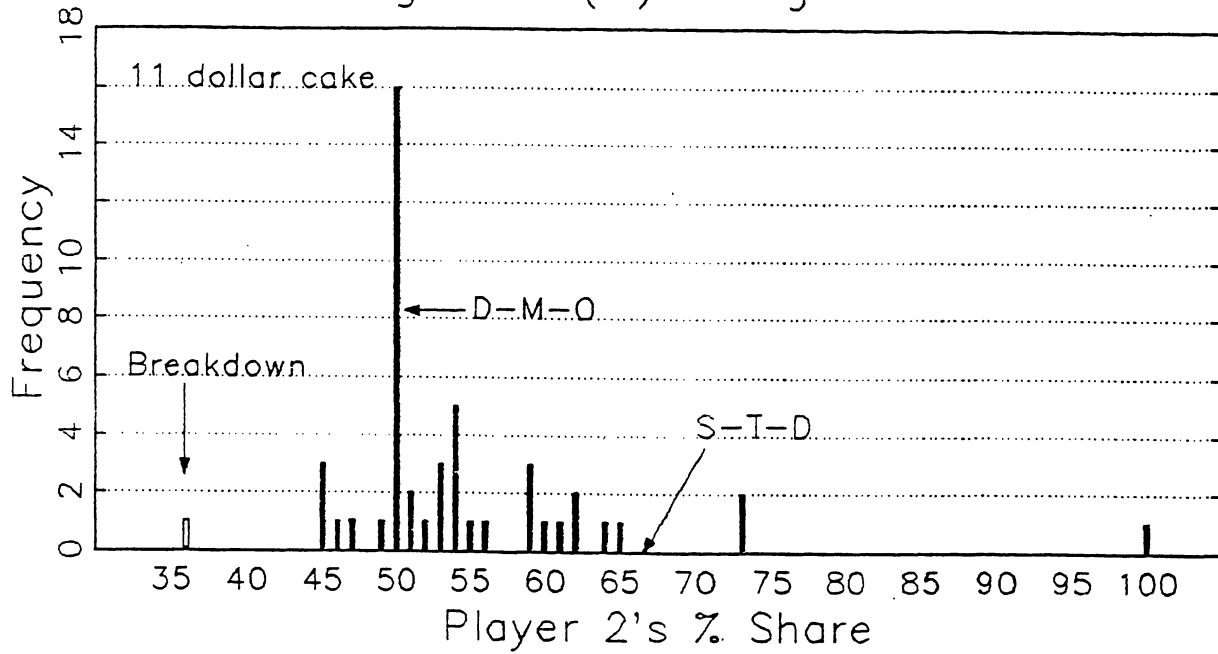


Figure 4(b): Regime 2

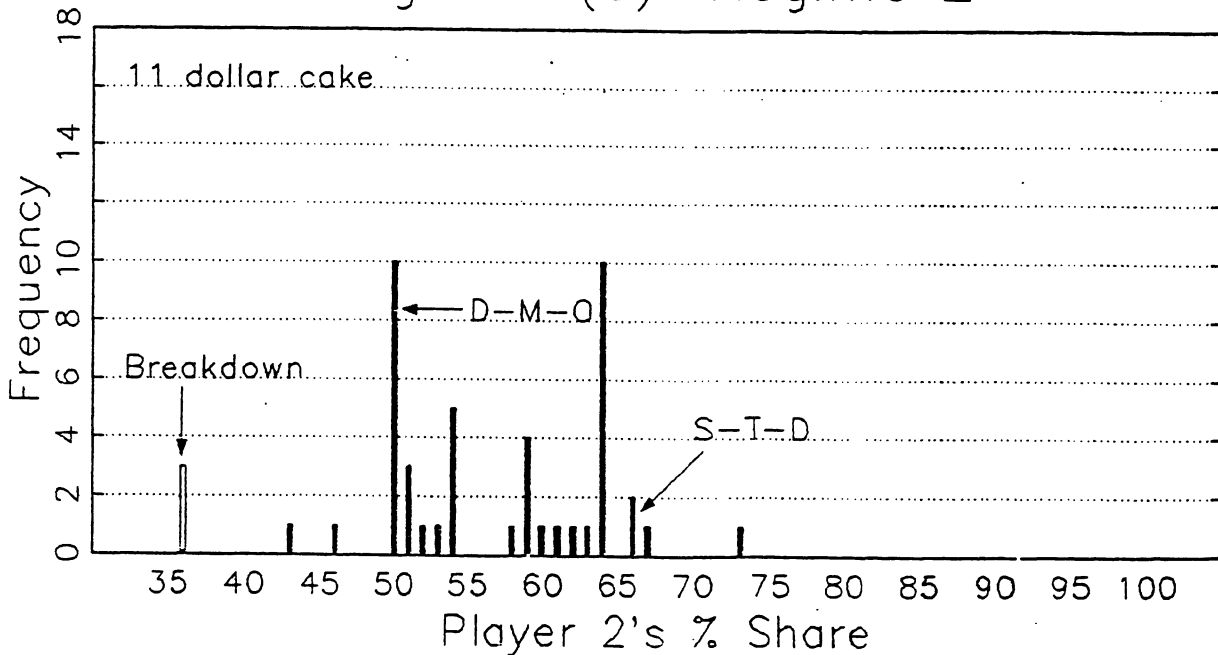


Figure 4.

This figure compares the paid games of class A and class B with low β for the \$11 cake.

5. Questionnaire.

Perhaps the most interesting results were those obtained from the subjects' answers to the questionnaire (appendix C). We discuss only⁹ the answers to questions 5, 6 and 7. On question 5, it is only necessary to observe that the subjects were unanimous that 50:50 is "fair" in this symmetric situation.

The answers to questions 6 and 7 are summarized in figures 5 and 6. Figure 5 shows the views about "fairness" expressed by all those who had experienced class A games (in which D-M-O is strategically optimal). The x -coordinate of a point¹⁰ in the figure shows what was asserted to be "fair" in a low breakdown payoff situation ($\beta = 36\%$) and the y -coordinate shows what was asserted to be fair in a high breakdown payoff situation ($\beta = 64\%$). Figure 6 similarly shows the views of all those who had experienced class B games (in which S-T-D is strategically optimal).

The difference between figures 5 and 6 is striking. (See section 7 for a statistical analysis.) Note in particular the following features:

1. In both figures, a small group (around 10%) insist that 50:50 is "fair" in spite of the asymmetries they are invited to contemplate.
2. Those who had experienced class B games (S-T-D strategically optimal) were very much more in agreement about what is "fair" than those who had experienced class A games (D-M-O strategically optimal). Results for class A games were very much more dispersed.
3. In class B games, S-T-D predicts what was asserted to be "fair" quite well, and D-M-O predicts very badly.
4. In class A games, the situation is more confused. However, D-M-O is no longer irrelevant to the data.

⁹ It is not clear to us how much weight can be given to the answers to question 8. For the record, we observe that 50% of the subjects were unambiguously of the view that it is socially acceptable to use one's bargaining power, and 17% were unambiguously of the opinion that one ought to "play fair".

¹⁰ Too much significance should not be attached to the *precise* location of points in figures 5 and 6. For example, in figure 6, most subjects simply proposed the S-T-D point. These choices have been indicated by clustering them as close to the S-T-D point as possible without overlaps. Also, subjects were not always very neat in marking their choice of a "fair division" on their questionnaires.

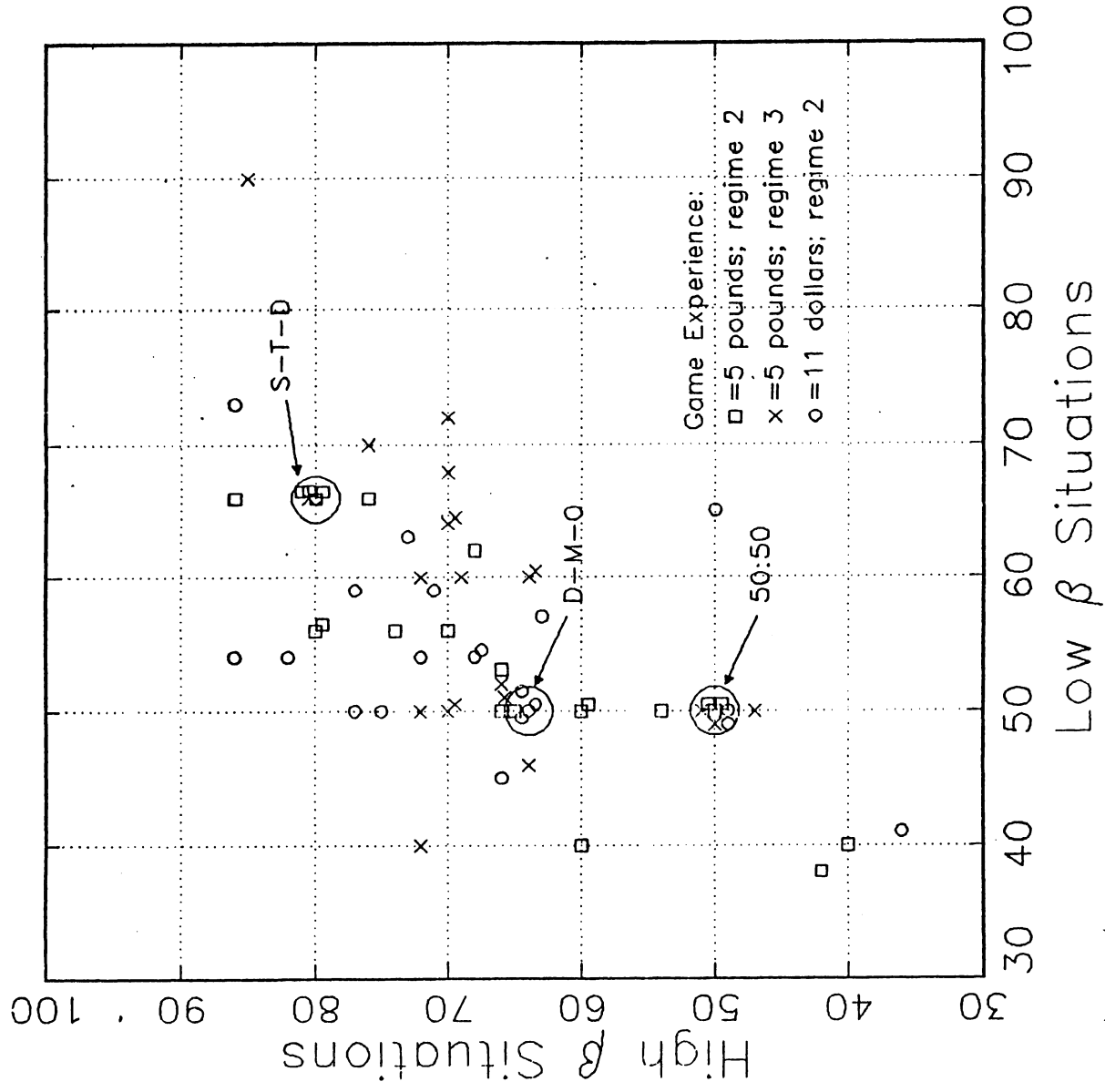


Figure 5. This figure shows the share of the cake for player II proposed as "fair" for low and high β situations by subjects who had experienced class A games.

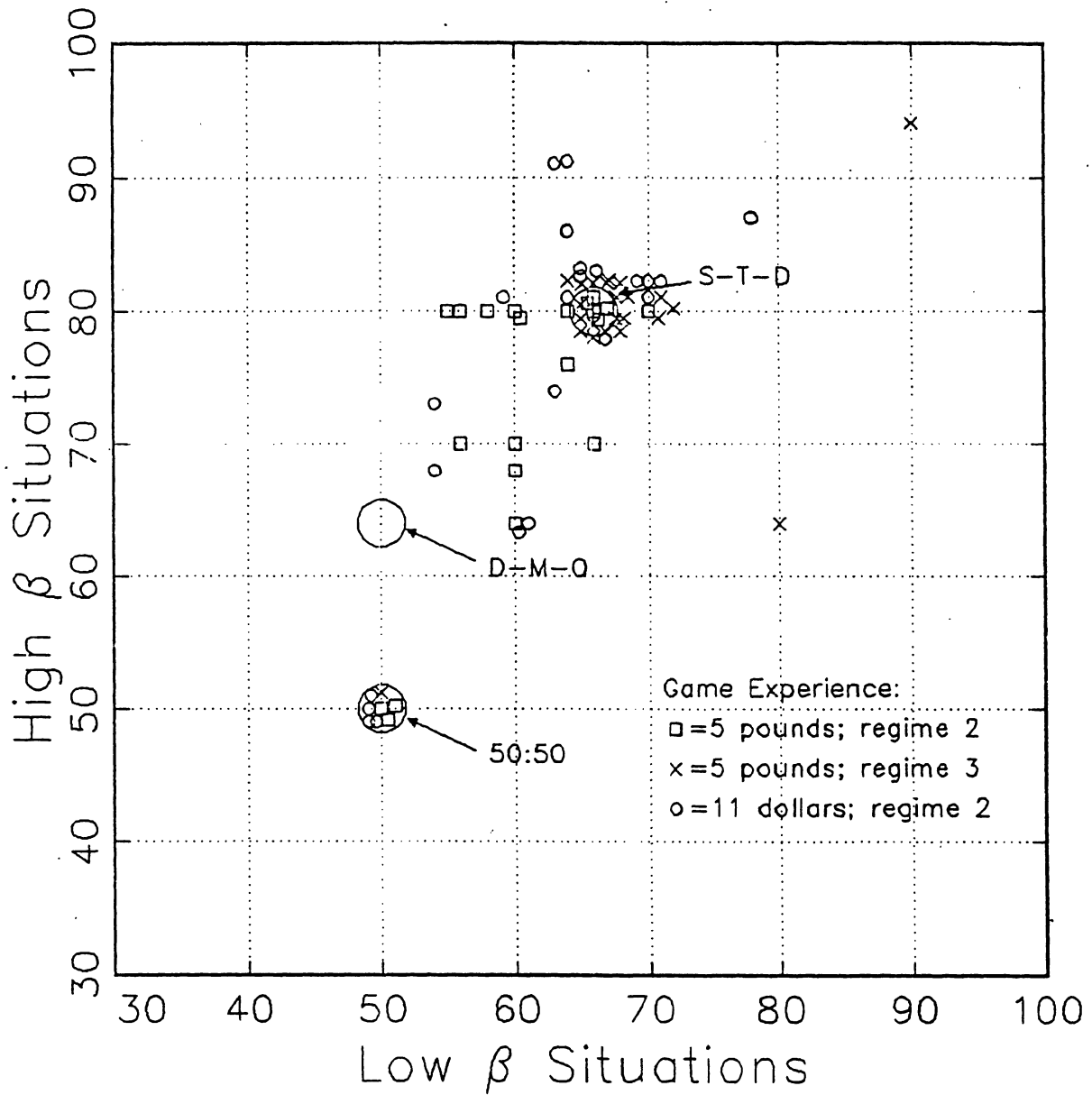


Figure 6.

This figure shows the share of the cake for player II proposed as "fair" for low and high β situations by subjects who had experienced class B games.

We do not feel that these results are conclusive, but we do feel that they make it likely that people's views about what is "fair" may be strongly influenced by their strategic experiences in situations about which they do not have established preconceptions.

6. Unlearning.

We were disappointed not to have obtained sharper results under regimes 0 and 1 (Class A games), since we had obtained sharp results in a previous study¹¹ of class A games [Binmore/Shaked/Sutton,1988], without apparent interference from round number focal points. However, in this previous study, subjects did not play repeatedly and hence had little opportunity to learn. It is therefore of interest to compare the results of our previous study with those shown in figure 7 for the first four practice games of class A (with the £5.00 cake).¹²

The results from these practice games and those from our previous study are similar, in that D-M-O predicts the data quite well in absolute terms, and overwhelmingly better than S-T-D. The drift away from this distribution will be evident from figure 3. One can tell a story for regime 1 (high β) about subjects learning that player II needs an epsilon on top of what is available from opting out if he or she is to be kept at the negotiating table. However, this would not seem to explain why player II's payoffs should improve over time under regime 0 (low β). Presumably round number focal points are somehow relevant. Evidence in support of this would seem to be provided by the differing results obtained for regime 0 with an \$11 cake (figure 4). More research will perhaps provide an explanation for what is going on here. For the moment, the only safe conclusion would seem to be that, if people are indeed "natural gamesmen",¹³ then experience in this context would appear to lead to some "unlearning" of their game-playing skills.

¹¹ The size of the cake and the values of α and β were *not* the same.

¹² We did not run regime 1 with an \$11 cake. The results from practice games under regime 0 for the \$11 cake are very similar to those for the £5 case. Only the first four practice games are reported so that the number of observations in each of the histograms of figures 2, 3, 4, and 7 is the same.

¹³ A view that has been wrongly attributed to us in the past.

7. Statistical analysis.

This section is under review.

8. Conclusion

It is *not* claimed in this paper that fairness/focal theories are mistaken. Nor is it claimed that subjects are natural gamblers. Our belief is that a more sophisticated type of theory than either of these alternatives is necessary.

It seems very likely to us that people are equipped with rules-of-thumb that they use to settle conflicts of interest that arise in real-life bargaining situations and elsewhere, and that these rules-of-thumb embody “fairness” criteria or depend in other ways on salient or focal features of the environment in which they are used. We shall follow Dawkins [1976] in referring to such rules-of-thumb as *memes*. It seems unlikely that people think very hard about these memes when using them in the real-life situations for which they are appropriate. One tends to question ingrained habits or customs only when their use generates unsatisfactory results.

At the same time, it seems to us very unlikely indeed that there are any natural gamblers at all, if by this is meant individuals who are familiar with all the theorems of game theory, even before they have been proved, and who are capable of lightning mental calculations of great complexity and who take for granted that other individuals selected at random from the population at large have the same characteristics as themselves. Nevertheless, we believe it likely that game theory *is* very relevant to real-world behavior. Given a meme which is established in a human population, one has to ask: how does it manage to survive? Why does it not get displaced by an alternative meme? The game theorist’s answer is that it survives because it is *adapted* to the environment in which it is commonly used. That is to say, it prescribes behavior that is *in equilibrium*.¹⁴ People will not usually be conscious

¹⁴ It goes without saying that this is a gross simplification. One must take into account the complexity of the meme’s environment. The more complex the environment, the more difficult it will be for better adapted memes to surface, and the longer it will take for them to become established. One must also consider the cost to individuals of implementing complicated strategies. And so on.

Figure 7(a): Regime 0

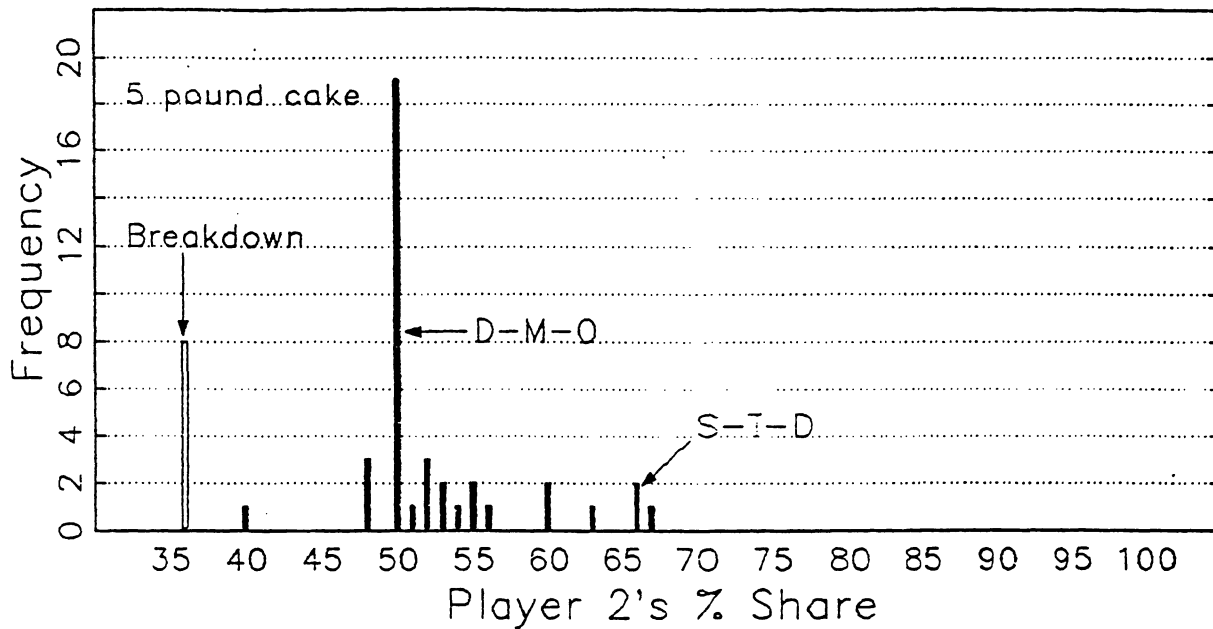


Figure 7(b): Regime 2

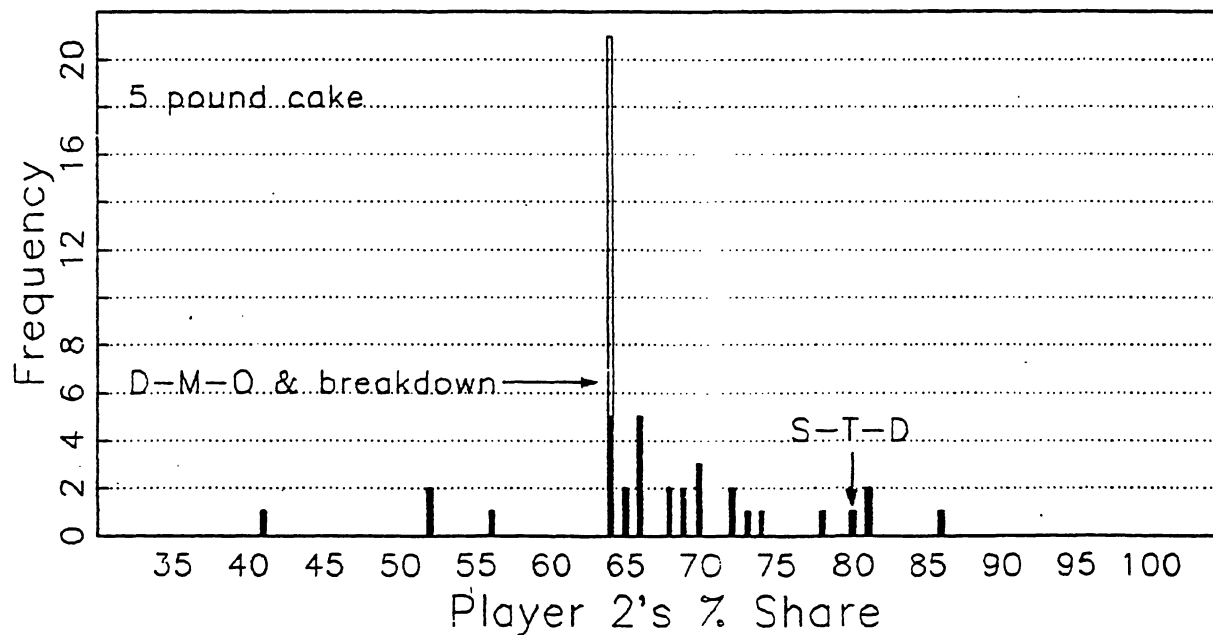


Figure 7.

This figure shows the first four *practice* games of class A and with low and high β for the £5 cake.

7. Statistical analysis.

This section is under review.

8. Conclusion

It is *not* claimed in this paper that fairness/focal theories are mistaken. Nor is it claimed that subjects are natural gamblers. Our belief is that a more sophisticated type of theory than either of these alternatives is necessary.

It seems very likely to us that people are equipped with rules-of-thumb that they use to settle conflicts of interest that arise in real-life bargaining situations and elsewhere, and that these rules-of-thumb embody "fairness" criteria or depend in other ways on salient or focal features of the environment in which they are used. We shall follow Dawkins [1976] in referring to such rules-of-thumb as *memes*. It seems unlikely that people think very hard about these memes when using them in the real-life situations for which they are appropriate. One tends to question ingrained habits or customs only when their use generates unsatisfactory results.

At the same time, it seems to us very unlikely indeed that there are any natural gamblers at all, if by this is meant individuals who are familiar with all the theorems of game theory, even before they have been proved, and who are capable of lightning mental calculations of great complexity and who take for granted that other individuals selected at random from the population at large have the same characteristics as themselves. Nevertheless, we believe it likely that game theory is very relevant to real-world behavior. Given a meme which is established in a human population, one has to ask: how does it manage to survive? Why does it not get displaced by an alternative meme? The game theorist's answer is that it survives because it is *adapted* to the environment in which it is commonly used. That is to say, it prescribes behavior that is *in equilibrium*.¹⁴ People will not usually be conscious

¹⁴ It goes without saying that this is a gross simplification. One must take into account the complexity of the meme's environment. The more complex the environment, the more difficult it will be for better adapted memes to surface, and the longer it will take for them to become established. One must also consider the cost to individuals of implementing complicated strategies. And so on.

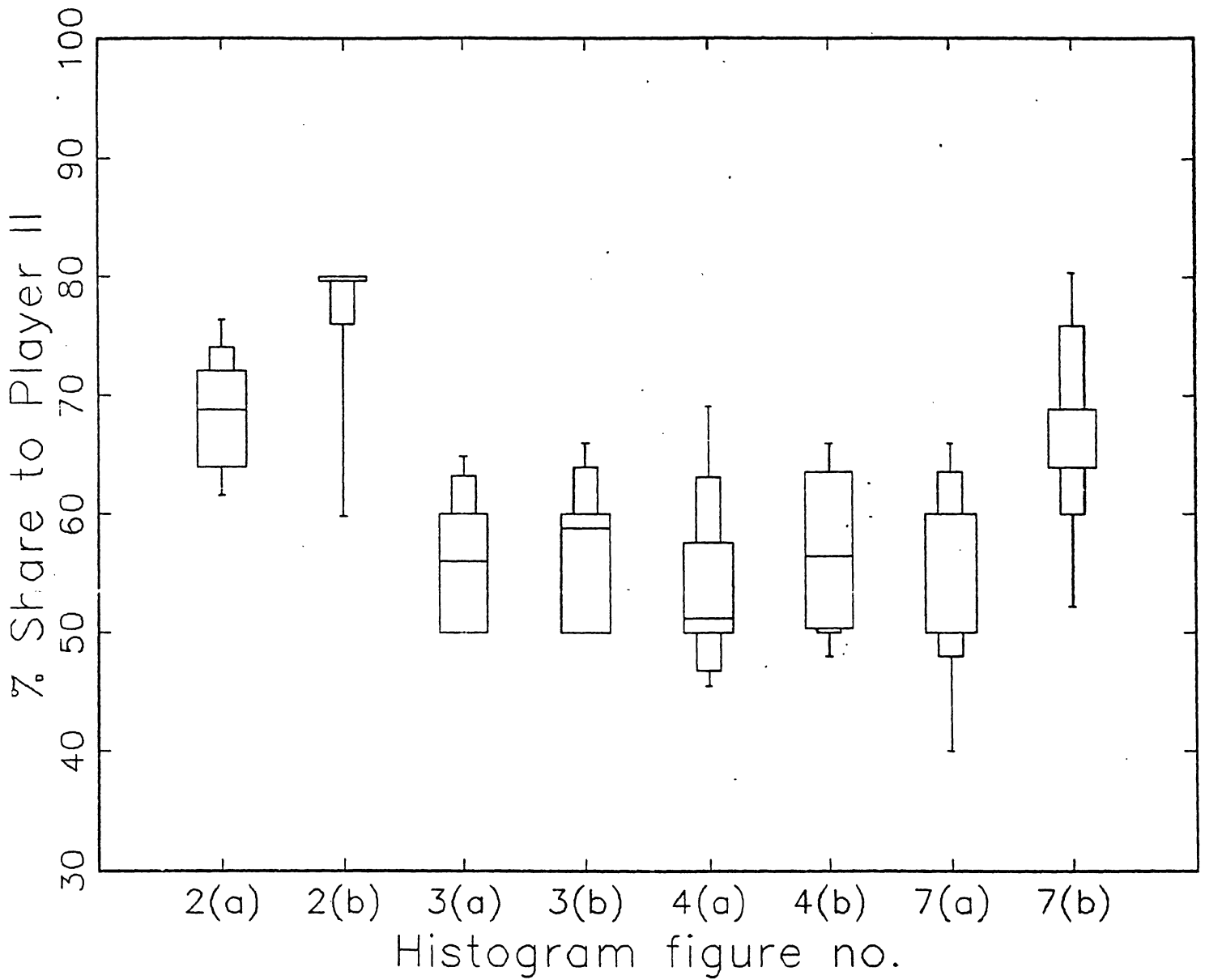


Figure 8.

This figure shows box and whisker plots for the histograms of figures 2, 3, 4, and 7. Each central box contains 25% of the mass. The flanking subsidiary boxes each contain 15% of the mass and each whisker contains 5% of the mass.

of this fact and may be quite truthful in reporting that they feel unmotivated by ,strategic considerations. But it is not necessary for individuals to know why a particular meme survives in order that it survive. Or to put the same point another way, another way, it is not necessary that they know that they are gamemen, or even that they are playing a game, for people to act *as though* they are gamesmen.

If this view is correct, at least in some circumstances, then one should expect to observe memes in operation that are triggered by hints or cues in the environment which are correlated with the strategic realities of the situation. We hope that the current paper will be seen as confirming the viability of such a standpoint rather than as just a refutation of a naive version of the fairness/focal explanation of human behavior. Subjects were put in situations for which life did not seem to have equipped them in advance with a strongly established rule-of-thumb. Behavior then evolved which correlated with the strategic situation and the subjects seemingly developed attitudes towards “fairness” that allowed them to rationalize this behavior in terms that were familiar to them. We do not doubt that it is possible to construct experiments in which such adaptation to the environment does *not* occur, even with large incentives and long time spans for learning. One might frame the experiment in such a way as to trigger a meme that it is very firmly established for use in a particular real-world context but which bears only a surface resemblance to the problem faced in the laboratory.¹⁵ Alternatively, it would not be hard to interfere with the learning process by confusing the issues facing the subjects. Indeed, we seem to have done so inadvertently under regimes 0 and 2 with the £5 cake by introducing a round number focal point at £3.00. Such experiments, however, would not and do not refute the view that we are defending here.

To summarize: in defending the relevance of game theory to actual bargaining behavior, we are not denying that fairness/focal theories are also relevant. We deny only that a theory of this type that ignores strategic considerations is likely to get

¹⁵ Invoking the meme in such pathological circumstances is, of course, precisely what one would wish to do if one’s aim were to study the mechanics of a particular established meme.

to what lies at the heart of human behavior. Most of all, we want to emphasize the importance of learning and adaptation in this context.

References

- Binmore, K., Shaked, A., and Sutton, J. (1985) "Testing Noncooperative Bargaining Theory: A Preliminary Study," *American Economic Review*, **75**, 1178–80.
- Binmore, K., Shaked, A., and Sutton, J. (1988) "A Further Test of Noncooperative Bargaining Theory: Reply," *American Economic Review*, **78**, 837–9.
- Binmore, K., Shaked, A., and Sutton, J. (1985) "An Outside Option Experiment," STICERD discussion paper TE 85/124, London School of Economics.
- Cramton, P. (1988) "Dynamic Bargaining with Outside Options," Yale School of Management.
- Güth, W. (forthcoming) "Behavioral Theory of Distributive Justice," in *Applied Behavioral Economics*, ed. Maital, S.. Brighton, Sussex: .
- Kahneman, D., Knetsch, J., and Thaler, J. (1986) "Fairness and the Assumptions of Economics," *Journal of Business*, **59**, 285–300.
- Roth, A. (1985) "Towards a Focal Point Theory of Bargaining," in *Game-Theoretic Models of Bargaining*, ed. Roth, A.. Cambridge: Cambridge University Press.
- Rubinstein, A. (1982) "Perfect Equilibrium in a Bargaining Model," *Econometrica*, **50**, 97–109.
- Selten, R. (1978) "The Equity Principle in Economic Behavior," in *Decision Theory and Social Ethics, Issues in Social Choice*, ed. Gottinger, H., and Leinfellner, W.. Dordrecht: Reifel Publishing.
- Shaked, A. (1987) "Opting Out: Bazaars versus 'Hi Tech' Markets," STICERD discussion paper TE 87/159, London School of Economics.
- Dawkins, R. (1976) *The Selfish Gene*. Oxford: Oxford University Press.

