Presidential Address

Features of experimentally observed bounded rationality

Reinhard Selten*

Laboratorium für experimentelle Wirtschaftsforschung, Universität Bonn, Adenauerallee 24-42, D-53111 Bonn, Germany

Abstract

On the basis of experimental evidence reported in the literature the paper draws conclusions about the bounded rationality exhibited by human economic behaviour. Among the topics discussed are presentation effects caused by superficial analysis, strategic reasoning and strategy construction based on reciprocity and fairness, avoidance of circular concepts in step by step strategic reasoning, ex-post rationality and learning direction theory, presence of both adaptive and analytic approaches to repeated decision tasks, and the absence of quantitative expectations and optimization in typical repeated game strategies. © 1998 Elsevier Science B.V. All rights reserved.

JEL classification: C70; C90; D63; D83

Keywords: Bounded rationality; Experiments; Strategic interaction

1. Introduction

In this talk the term bounded rationality is used in the tradition of H.A. Simon (1957). Bounded rationality is understood as rationality exhibited by actual human economic behaviour.

*Tel.: +49-228-73-9190; fax: +49-228-73-9193; e-mail: selten@lab.econ1.uni-bonn.de.
The picture of rational decision making underlying most of contemporary
economic theory is far away from observed behaviour. It is therefore necessary
to develop theories of bounded rationality. Experimental results support some
theories of limited range. However, an empirical-based general theory remains
a task of the future.

Behaviour cannot be invented in the armchair. It has to be observed. There-
fore, the development of theories of bounded rationality needs an empirical
basis. Laboratory experimentation is an important source of empirical evidence.
Of course, also field data are important, but they are more difficult to obtain and
harder to interpret.

Our knowledge of boundedly rational economic behaviour is still very lim-
ited, but nevertheless experimental research has already supplied many insights.
Only some of them will be presented here. The talk concentrates on selected
topics of importance for strategic interaction. Wide areas of successful behav-
ioural research will not be touched, e.g. risky choice, probability judgment, and
decimal prominence, to name only a few among many.

2. Three roots of behaviour

Human economic behaviour has a complex structure. At least three different
kinds of mental processes interact:

1. Motivation – the driving force.

This is probably a somewhat simplistic description of what we might call the
roots of behaviour, but nevertheless a useful one. Motivation is concerned with
what behaviour aims at and how a multitude of fears and desires combine to
determine human action. The word *adaptation* is used in order to describe the
kind of learning behaviour which is already observed in animals with relatively
primitive central nervous systems to whom one would not ascribe any powers of
reasoning. Human behaviour is not just due to the interplay of motivation and
adaptation but also heavily influenced by cognition. The word *cognition* stands
for all the reasoning processes in the human mind, regardless of whether they are
fully conscious or not.

Much of recent experimental research deals with questions of motivation,
e.g. the influence of fairness and reciprocity. *Reciprocity* can roughly be de-
scribed by the saying *I do unto you as you do unto me*. Here we are not primarily
concerned with motivational issues, but, as we shall see, we cannot completely
avoid them because motivation and bounded rationality are not completely separable.
Adaptive explanations of behaviour have proved to be quite successful. The results obtained by Roth and Erev (1995) are impressive. They succeeded in explaining behaviour in various repeated games with the help of a simple reinforcement model. In this talk the emphasis will be more on cognition than adaptation. More detailed comments on these interesting results would be beyond the limited scope of this talk.

Reciprocity means that there is a tendency to react with friendliness to friendly acts and with hostility to hostile acts. This requires an interpretation of acts of others as friendly, hostile, or neutral. Here boundedly rational cognition enters the picture. Whether an act is perceived as friendly, neutral, or hostile depends on boundedly rational reasoning processes.

New descriptive theories based on experimental results indicate that in many interactive situations subjects are not clearly motivated by reciprocity in the sense explained above (Bolton and Ockenfels, 1997; Fehr and Schmidt, 1997). According to these theories, motivation can be described by a utility function depending on own payoff and deviation from fairness. Here, too, cognition comes into play. What is perceived to be fair in a specific situation depends on boundedly rational reasoning processes.

Adaptation may involve an element of cognition, too. More about this will be said in connection with learning direction theory (Selten and Stoecker, 1986; Selten and Buchta, 1994). In this theory the interpretation of experiences in the light of qualitative causal beliefs guides the direction of adaptation.

3. A presentation effect (framing) in a duopoly experiment

The response to a decision task often depends on how it is presented to the subjects. It may be important whether the results of a medical treatment are described in terms of dead or of surviving patients (Tversky and Kahneman, 1981). Today it has become customary to refer to such effects as the results of framing. Actually, the phenomenon has been observed long before this term was coined. In our study (Selten and Berg, 1970) we referred to it as a presentation effect. Without going too much into detail the effect will be described in the following.

The paper by Selten and Berg (1970) reports on duopoly experiments in continuous time with face to face interaction. The underlying model was an asymmetric Cournot duopoly with linear demand and quadratic costs. Each player started with some initial assets $A$. The final payoff $P$ was the sum of $A$ and the profit $G$ obtained in the course of the game.

In the experiments the presentation was changed systematically as indicated by Table 1. In the transition from one presentation to another the initial assets of a player was changed by adding a constant $C$. This change was exactly compensated by an adjustment of fixed costs per time. Money payoffs were
proportional to final assets. Different constants were used for both players. The way in which payoffs depend on behaviour is not changed from one presentation to the other. Nevertheless, the presentation shift parameters had a strong influence on behaviour.

In the experiments, most pairs of subjects, sooner or later, arrived at one of two modes of cooperation. Sometimes the players agreed to aim at high joint profit (not always the maximal one) and to divide their money evenly after having received their final payoffs. Such agreements were not protected by the experimenter. Subjects would not always trust each other to stick to such promises.

Another mode of cooperation involved alternating between several quantity combinations, mostly Pareto-optimal ones, in such a way that profits were approximately equal. The shift parameters change the relationship between quantity combinations and profits. Since the shift parameters were different for both players cooperation at approximately equal profits became more favourable for one player and less favourable for the other by a change of the presentation. This leads to a presentation effect.

It seems to me that the presentation effect described here is due to a feature of boundedly rational reasoning which may be called superficial analysis. A decision situation is first approached in a superficial way, since a reasoning process has to begin somewhere and deep properties of the problem cannot be seen already at its start. In the duopoly experiment subjects looked at both profit tables and saw their regions of positive and negative profits. Superficial analysis suggests that zero profits should be a lower bound for a player’s aspiration level. Starting from there it is natural to agree on equal profits. Since in this way a satisfactory solution to the problem where to cooperate is easily found, the analysis is not deepened to the point where it becomes clear that the distinction between negative and positive profits is not as relevant as it may seem at first glance.

A deeper analysis would involve looking at the game-theoretic properties of the situation, like the Cournot–Nash equilibrium or cooperative solutions of the underlying model. However, the reasoning process does not progress to a point at which the veil of the presentation is lifted.
4. Presentation effect in prisoner’s dilemma supergames

Pruitt (1970) compared two different presentations of a prisoner’s dilemma game repeated over many periods. The usual presentation describes the game as a bimatrix as in Table 2.

The game of Table 2 can also be presented in a *decomposed form* as in Table 3.

The decomposed form presentation explains the game as a situation in which each of both players simultaneously chooses between two possibilities \(A\) and \(B\). A player who chooses \(A\) takes zero for himself and gives 3 to the other player. Similarly, the choice \(B\) consists of taking one for oneself and giving zero to the other player. The payoff of a player is the sum of what he takes for himself and what the other player gives him. Thus, if player 1 chooses \(A\) and player 2 selects \(B\), player 1 takes nothing for himself and receives nothing from the other player; his payoff is zero. Player 2 takes 1 and receives 3 from player 1. Player 2 has payoff 4. In this way, one can easily check that Tables 2 and 3 show different presentations of the same game.

Table 2
A prisoner’s dilemma game. Payoffs of player 1, the row player, are shown in the upper left corner and payoffs of player 2, the column player, are shown in the lower right corner

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3
The game of Table 2 in decomposed form

<table>
<thead>
<tr>
<th></th>
<th>For me</th>
<th>For him</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Game theoretically it should make no difference whether the game is presented in bimatrix form or in decomposed form. However, as Pruitt has shown, one receives dramatically more cooperation in the decomposed form presentation. This is a very important result which cannot be explained by any theory which exclusively focuses on the relationship between behaviour and payoff.

My interpretation of Pruitt’s findings is based on the idea of superficial analysis which already has been explained in Section 3. The decomposed form makes it easy to recognize choice $A$ as a friendly act. In the bimatrix form a choice of $A$ by the opponent might be interpreted as motivated by the aim of obtaining 3 which is a relatively high payoff. In the decomposed form $A$ is a much clearer signal of cooperativeness. A superficial look at the decomposed form immediately suggests that it is generous to choose $A$ and egoistic to choose $B$. This is not the case for the bimatrix form.

As explained in Section 2, reciprocity means *I do unto you as you do unto me*. For the point I want to make here it is not important whether subjects are motivated in this way or rather by a utility function which depends on the own payoff and the deviation of the outcome from fairness. In symmetric games like the prisoner’s dilemma it is very clear what fairness means: Both players should get the same payoff. Even if *I do unto you as you do unto me* is not an original motivation, it may very well be a strategy used in order to bring fair cooperation about. In the case of the decomposed form, a supergame strategy like tit-for-tat may much more easily be discovered by one player and recognized by the other one as an attempt to establish fair cooperation.

5. Strategic reasoning without circular concepts

In this section we shall look at the theory of equal division payoff bounds (Selten, 1987) for zero-normalized three person games in characteristic function form. For the sake of simplicity, attention will be restricted to games without the grand coalition, i.e. the coalition of all three players. Only the two-person coalitions 12, 13, 23 are permitted in such games (The notation $ij$ indicates that the players $i$ and $j$ are members of the coalition). The term *zero-normalized* means that players outside a two-person coalition receive zero payoffs.

A characteristic function of such a game can be described by three numbers $a$, $b$, and $c$ which denote the values of the coalitions 12, 13, and 23, respectively. In Fig. 1 the situation is described symbolically by a triangle the corners of which present the players and the sides of which stand for the possible two-person coalitions. At these sides the corresponding coalition values are indicated. On the right-hand side the figure shows a numerical example.

If two players form a coalition they can divide its value among themselves in any way they want. Only one two-person coalition can be formed. The player
not belonging to it receives zero. It is also possible that no coalition is formed at all, in which case all players receive zero.

5.1. Quotas

In order to keep our explanations as simple as possible we further restrict our attention by the inequalities shown in Fig. 1. The first inequality excludes negative coalition payoffs and symmetries among players, but, apart from that, entails no restriction of generality, since the order of the coalition payoffs can be achieved by a suitable numbering of the players. The second inequality is the so-called triangular inequality which secures the existence of positive numbers $q_1$, $q_2$, $q_3$ with the property that for any two different players $i$ and $j$ the sum of $q_i$ and $q_j$ is the coalition value for $ij$:

$$q_1 + q_2 = a, \quad q_1 + q_3 = b, \quad q_2 + q_3 = c.$$ 

The number $q_i$ is called the quota of player $i$. In the numerical example of Fig. 1 the quotas of the players 1, 2, and 3 are 50, 30, and 20, respectively. A division of the value of a two-person coalition $ij$ according to the quotas of its members is called a quota agreement.

A quota agreement has a certain stability property. If such a coalition agreement is considered, one of both members, say player $i$, cannot propose a coalition $ik$ to the third player $k$ which gives player $i$ more than $q_i$, say $q_i + \varepsilon$, and $q_k - \varepsilon$ to player $k$ without giving player $j$ the opportunity to make a better offer to $k$ by proposing a quota agreement for the coalition $jk$. In this way player $j$ can counter $i$'s offer to $k$ in a way which protects his payoff proposed for the coalition $ij$.

For the three-person games considered here the stability property outlined above supports the idea that rational players should end up with a quota agreement. In fact, an important cooperative solution concept, the bargaining set (Aumann and Maschler, 1964) leads to the quota agreements in our case. The bargaining set is the only classical cooperative solution concept, which has been compared to experimental data with at least some success (Maschler, 1978; Kahan and Rapoport, 1984).
The quota concept is circular in the sense that none of the quotas is defined independently of the others. In order to find the quotas, one has to solve a simultaneous equation system. This kind of circularity is typical for rational game theory. However, subjects in the laboratory usually do not compute quotas. They seem to avoid circular concepts in their strategic reasoning.

5.2. Equal division payoff bounds

The theory of equal division payoff bounds (Selten, 1987) describes a boundedly rational reasoning process which arrives at lower bounds \( s_1, s_2, \) and \( s_3 \) for payoffs of players 1, 2, and 3 respectively, within a two-person coalition. These numbers are called equal division payoff bounds. We shall now explain the theory for the restricted class of games considered here. The reasoning process starts with the observation that player 1 has better coalition possibilities than player 2, and player 2 has better coalition possibilities than player 3. We express this by saying that the order of strength is

\[ 1 \succ 2 \succ 3. \]

Here the symbol \( \succ \) stands for ‘stronger’. The continuation of the reasoning process is based on the principle that the stronger member in a two-person coalition should at least get his or her equal share of the coalition value. Player 1 is strongest in 12 and player 2 is strongest in 23. This leads to the lower bounds \( s_1 \) and \( s_2 \) for players 1 and 2, respectively,

\[ s_1 = \frac{a}{2}, \quad s_2 = \frac{c}{2}. \]

From these lower bounds and upper bounds \( h_1 \) and \( h_2 \) for the payoffs of 1 and 2, respectively, in 12 are derived:

\[ h_1 = a - s_2, \quad h_2 = a - s_1. \]

Player 3 is in a difficult situation since the coalition 12 with the highest equal share is the most attractive one. Moreover, in no two-person coalition player 3 is the stronger member. Therefore, for player 3, a lower bound cannot be derived in the same way as for players 1 and 2. In order to have a chance to be in the final coalition, player 3 may have to be willing to give to both players 1 and 2 the upper bounds \( h_1 \) and \( h_2 \) they can obtain in 12. This leaves the minimum of \( b - h_1 \) and \( c - h_2 \) for player 3. However, player 3 also must get at least zero. This leads to the lower bound

\[ s_3 = \max[0, \min(b - h_1, c - h_2)]. \]

We call \( s_3 \) player 3’s competitive bound. For the numerical examples of Fig. 1 we obtain \( s_1 = 40, \ s_2 = 25, \) and \( s_3 = 10. \) The theory predicts that a two-person
coalition $ij$ will be formed in which both members receive at least their equal division payoff bounds $s_i$ and $s_j$, respectively.

5.3. Experimental evidence

Actually, the theory of equal division payoff bounds is somewhat modified by taking account of rounding effects in comparison to experimental data. Since usually experiments restrict payoffs agreed upon to integer multiples of a smallest money unit, it is also postulated that in a two-person coalition a player should not get zero but at least one smallest money unit. These modifications lead to the theory of rounded equal division payoff bounds.

Also the bargaining set needs to be modified in order to take rounding effects into account. Further modifications based on power transformations (Maschler, 1978) have to be introduced in order to improve predictive power. The comparison of the most favourable version of the bargaining set and the theory of rounded equal division payoff bounds can be based on a measure of predictive success which can vary between $-1$ and $+1$ (Selten and Krischker, 1982; Selten, 1987). An investigation by Uhlich (1989) of zero-normalized three-person games with and without the grand coalition obtains measures of predictive success of 0.19 for the most favourable version of the bargaining set and 0.64 for the rounded equal division payoff bounds. This investigation was based on 25 independent experimental studies by various authors. Clearly, rounded division payoff bounds are in much better agreement with the data.

5.4. Step-by-step reasoning without circular concepts

Cooperative solution concepts in game theory typically start with an abstract definition which has the same kind of circularity as the quota concept. It is then a mathematical problem whether something exists which satisfies the requirement and how a solution can be found in any particular case. The structure of the theory of equal division payoff bounds is quite different. It does not involve any circular concepts. It directly specifies a procedure by which the solution is found. The fact that the theory of equal division payoff bounds is empirically much more successful than alternative theories developed in classical game theory strongly suggests that the natural way of looking at game situations of this kind is not based on circular concepts, but rather on a step-by-step reasoning procedure.

6. Learning direction theory

Learning direction theory has its origin in an experimental investigation by Selten and Stoecker (1986) about the end effect in prisoners dilemma supergames.
Since then it has been successfully applied in many different experimental contexts, mainly by my associates and myself, but also by others (Mitzkewitz and Nagel, 1993; Selten and Buchta, 1994; Kuon, 1994; Ryll, 1996; Nagel, 1996; Kagel and Levin, 1996; Berninghaus and Ehrhart, 1996; Nagel and Tang, 1997; Cason and Friedman, forthcoming; Selten et al., in preparation).

For a class of repeated decision tasks, learning direction theory describes an influence of cognition on adaptation. It is not a full-fledged learning theory, but rather a statement about an aspect of learning in many situations. It is a qualitative theory which makes only weak predictions. However, these predictions seem to be surprisingly robust and reliable.

Consider the example of an archer who wants to hit the trunk of a tree by bow and arrow. If the arrow misses the tree on the left-hand side, the archer will be inclined to aim more to the right. Similarly, a miss on the right-hand side will result in a tendency to aim more to the left. Of course, the archer may also not change his aim at all, because he may think of the miss as caused by temporary exogenous influences like a gust of wind. However, if the aim is changed, then it will have a tendency to be changed in the indicated direction.

Learning direction theory is applicable to repeated decision tasks in which a parameter $p_t$ has to be chosen in a sequence of periods $t = 1, \ldots, T$ and in which feedback information after each period permits causal inferences about what might have been better last time. The adjustment tends to follow a principle of ex-post rationality:

$$p_t \leq p_{t-1} \quad \text{if } p < p_{t-1} \text{ could have been better, but not } p > p_{t-1},$$

$$p_t \geq p_{t-1} \quad \text{if } p > p_{t-1} \text{ could have been better, but not } p < p_{t-1}.$$

Ex-post rationality does not mean that the parameter must be changed, but only that the parameter is changed in the indicated direction if it is changed at all.
Learning direction theory does not predict that every single change of the parameter will obey ex-post rationality, but only that observed changes will be more than randomly expected in the indicated direction.

As has been said before, learning direction theory does not claim to be a full fledged learning theory, but rather the description of an important aspect of behaviour in a class of repeated decision tasks. It is not excluded that behaviour is also subject to other influences. However, it is assumed that ex-post rationality has a strong impact which results in a recognisable tendency to move more than randomly expected in the right direction.

A situation with repeated private value first price sealed bid auctions may serve as an example. In an auction of this kind all bidders simultaneously and independently submit a bid for an object to be sold. The highest bid is the price and the object is sold to the highest bidder or, if there are several highest bidders, to one of them chosen randomly with equal probabilities. Let $b$ be last period’s bid and $v$ be last period’s private value of a participant; moreover let $p$ be last period’s price. The participant may find himself in one of the following three experience conditions:

<table>
<thead>
<tr>
<th>Experience condition</th>
<th>Causal inference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Successful bid:</td>
<td>$b = p$ Bid possibly too high</td>
</tr>
<tr>
<td>Lost opportunity:</td>
<td>$b &lt; p &lt; v$ Bid too low</td>
</tr>
<tr>
<td>Outpriced value:</td>
<td>$p &gt; v$ None</td>
</tr>
</tbody>
</table>

In the case of a successful bid the participant has received the object. The participant then can infer that a lower price might have been sufficient. This will tend to make his bidding behaviour less aggressive. In the lost opportunity case the price could have been profitably overbid. This will tend to make bidding behaviour more aggressive. In the outpriced value condition there was no chance for profitably buying the object. Therefore, no influence in either direction is expected in this case.

7. Winner’s curse

In this section we shall look at an experimentally observed phenomenon which will later be explained with the help of learning direction theory. The winner’s curse was first described by Capen et al. (1971) in connection with oil-field auctions: on average, the winner of the auction who obtains the object sold by the highest bid makes a loss.

The phenomenon occurs in so-called common value auctions in which the object has an objective value not known to the bidders at the time of the bid. The highest bidder is very likely to be one of those who overestimate the value of the object. The higher the overestimate, the higher the bid. The phenomenon results,
if the estimated value is not sufficiently reduced in the computation of the bid in order to take account of the danger of overestimating.

In the following we shall look at a very simple example which involves only one bidder: $A$ is willing to sell an object to $B$. The value $v$ of the object for $A$ is uniformly distributed over the interval $0 \leq v \leq 100$. The value of the object for $B$ is $1.5v$. Player $A$ knows $v$, but $B$ knows the distribution only. $B$ names a price $x$ and $A$ either accepts or rejects the bid. This game is repeated over a sequence of periods. A new $v$ is drawn independently of the past for each period. After each period, $v$ is made known to $B$, regardless of whether the object was sold to him or not.

Bazerman and Samuelson (1983) were the first to perform experiments on this repeated decision task. In these experiments subjects are in the role of $B$. Player $A$ is simulated by the computer. The subjects are told that $A$ sells if $v \leq x$ and rejects the bid otherwise.

At least in the beginning many subjects are guided by the unconditional expectation $E(v) = 50$ of $v$ and choose their bids in the vicinity of this value. However, this approach to the task is erroneous. $A$ sells only in the case $v \leq x$; therefore only the conditional expectation for this case $E(v|v \leq x)$ is relevant. This expectation is $0.5x$. Therefore, the conditional expectation for $B$ is $0.75x$. This has the consequence of a conditional loss expectation of $0.25x$. It follows that $x = 0$ is the only optimal bid.

Fig. 2 shows the average bids in an experiment by Ball et al. (1991) over 20 periods. 33 subjects took part in the experiment. According to the criteria of Ball et al. (1991) only two of them avoided the winner’s curse. It can be seen that the average bid remains on roughly the same high level. There is no tendency towards the optimum bid.
The argument that subjects choose a high bid mistakenly guided by the unconditional expectation 50 explains why average bids are high at the beginning. However, it is not clear why subjects should not learn to make lower bids by the experience of average losses. As we shall see in the next section, learning direction theory offers an explanation.

8. Winner’s curse and learning direction theory

This section is based on results obtained by Selten et al. (in preparation). Our winner’s curse experiments are very similar to those described in Section 7 with the following modifications: The experiment was extended over a hundred periods. \( v \) was uniformly distributed over a range of integer values \( u, \ldots, 99 \). Altogether 54 subjects, 18 for each of the cases \( u = 1, 11 \) and 21, took part in the experiment. As before \( v \) was the value for \( A \) and \( 1.5v \) was the value for \( B \). The subjects had a start capital of 250. Moreover, they received a fixed period income of 20 in addition to what they earned in the decision task. This served to avoid bankruptcy of subjects falling into the trap of the winner’s curse.

In the case of the winner’s curse ex-post rationality means that the bid is changed in the direction of last period’s value, if it is changed at all. If the bid was greater than the value then a lower bid would have been more profitable. If the bid was smaller than the value then no profit was made but a profit could have been made by a higher bid.

The tendency towards ex-post rationality was clearly present in the data. A comparison with an alternative random theory showed that subjects changed their bids more frequently in the right direction than randomly expected. On the basis of subjects taken as independent observations the null hypothesis can be rejected on the 5\% significance level (one-tailed) for \( u = 1, 11 \) and 21 separately.

8.1. Analytical and adaptive approaches to the decision task

The overall conformance to learning direction theory hides the fact that different subjects behave differently. This becomes clear by a classification of subjects by modal bids. One can distinguish between analytical and adaptive approaches to the decision task and exceptional ones which fit neither of both categories. Analytical approaches are based on explicit computations which, however, do not necessarily involve optimisation. The adaptive approach is taken by subjects who adjust to experience in an uncaring way, more or less, guided by ex-post rationality. The exceptional approaches were taken by two subjects referred to as gamblers whose most frequent bid was 99 and by two non-participants who did not make bids in the range of possible values in spite of the fact that bids with guaranteed riskless profit opportunities were available.
The numbers of subjects in the different categories are as follows:

**Analytical:** 10 optimizers, 3 loss-avoiders, 8 asset conservers.

**Adaptive:** 29 adapters.

**Exceptional:** 2 gamblers, 2 non-participants.

The optimizers were subjects whose modal bids were near to those which optimize expected value. The three *loss-avoiders* most frequently chose bids near to the highest one which excludes losses. The *asset conservers* had modal bids near to the highest bid at which losses up to 20 are possible but not higher ones. Since subjects receive a fixed additional period income of 20, excluding losses of more than 20 prevented decreases of assets from one period to the next.

Instead of optimising, loss avoiders and asset conservers limit their losses by aspiration levels naturally suggested by the situation. In the case of the asset conservers the loss limit 20 is due to superficial analysis. One could, for example, decrease the fixed additional income by \( k \) units and compensate this change by an increase of initial assets by \( 100k \). This manipulation is likely to result in a presentation effect, similar to that discussed in Section 3.

It is interesting to see that a sizeable minority of subjects took an analytical approach to the decision task. Optimisation is only one of these approaches. The difference between analytical and adaptive approaches seems to be very important for many decision tasks.

One of the reasons why we extended the experiment over 100 periods was the intention to give learning a better chance to approach the optimum than in a relatively short sequence of periods. Only for the adapters it is meaningful to ask the question whether there is a tendency of learning to lead towards convergence to the optimum. No such tendency can be seen in the data. The difference of the bid in period 90 and the bid in period 10 is positive for 13 adapters, zero for 3 of them, and negative for 13 of them.

A statistical test shows that adapters conform more closely to learning direction theory than other subjects (1% significance level, one-tailed).

### 9. Measure for measure

This section will describe an investigation by Selten et al. (1997) in which the strategy method was applied to a 20 times repeated asymmetric quantity variation duopoly with numerically specified linear costs and demand. The strategy method is an experimental procedure in which subjects have to write strategies for a game involving a sequence of decisions. In our case the strategies took the form of computer programs specifying decisions for every possible history.

Game playing experiments with decisions made as they arrive in the course of play provide much less information about the strategic reasoning of the subjects
than studies employing the strategy method. Therefore, the strategy method is an important tool for the exploration of bounded rationality.

In our case, 24 participants first played the 20 period supergame three times against changing anonymous opponents. Then they programmed strategies, each of them both for player 1 and for player 2. Then, in a computer tournament, each strategy was matched against all others on the other side. After the tournament each participant received printouts of all tournament games involving his strategy program. In the light of this information the subjects could change their strategies. The new strategies were then again matched in a computer tournament with the same feedback information as before. Finally, a third strategy programming round followed with strategy matching in a last tournament.

The experiment was run in the form of a student seminar and took a whole semester. The participants did not receive money payoffs, but were motivated by grades mostly based on success in the final tournament. The participants also had to give a written account of the reasoning behind their final strategies.

9.1. Typical strategies

A typical strategy structure emerged in the final tournament. The typical approach to the strategic problem involves two questions answered one after the other:

1. What is my cooperative goal?
   The typical answer takes the form of an ideal point, a quantity pair for both players determined on the basis of fairness criteria.
2. How do I reach cooperation at my cooperative goal?
   The typical answer is a measure for measure policy which reacts to movements towards or away from one’s ideal point by own movements of a similar size in the same direction.

Actually, a typical measure for measure policy is a little more complex. The response is bounded by the ideal point quantity from below and by the Cournot output from above.

Obviously, a measure for measure policy is based on a principle of reciprocity. If the opponent lowers her output and thereby brings it nearer to my ideal point, I perceive this as a friendly action which I reciprocate. Similarly, if she increases her quantity and thereby moves away from the ideal point, I also reciprocate by increasing my output.

The typical strategy has a phase structure. In an initial phase of up to 4 periods, a fixed decreasing sequence of outputs is specified. Then in a main phase a measure for measure policy is applied. Finally, in an end phase of up to 4 periods, behaviour becomes non-cooperative. In most cases Cournot outputs are specified for the end phase.
The decreasing quantities in the initial phase signal a cooperative intention. The measure for measure policy then tries to guide the behaviour of the other player towards one's own cooperative goal. Finally, the attempt to cooperate is abolished in the end phase.

The ideal points are based on fairness criteria like equal profits or equal surplus of profits over Cournot profits or profits proportional to Cournot profits, etc. Two players who try to achieve cooperation at different ideal points may still cooperate if the ideal points are compatible in a certain sense. Otherwise cooperation is not achieved.

9.2. No quantitative prediction and no optimization

The typical strategic approach neither involves quantitative predictions of the other player's behaviour nor any attempt to optimise against such predictions. This is very different from what we find in game theory and traditional oligopoly theory. Virtually, all theories proposed in the literature postulate optimisation against quantitative expectations about the behaviour of the others. Nevertheless, the typical approach to the strategic problem taken by our subjects is by no means unreasonable. It embodies its own rationality which of course is bounded one.

A Bayesian approach would force a strategy programmer to form an a priori distribution over all possible opponent strategies and to solve a dynamic programming problem on this basis. The subjects typically do nothing even remotely similar to that. Instead of passively adapting to quantitative expectations about the others, they actively try to exert an influence on the opponent's behaviour.

The reciprocity embodied in measure for measure strategies seems to be instrumental rather than directly related to motivational factors. The fact that in the end phase cooperative intentions are abolished points in this direction.

Measure for measure is similar to tit-for-tat. In fact, the tit-for-tat strategy which did so well in the Axelrod (1984) tournaments can be looked upon as a measure for measure policy for the repeated prisoner's dilemma. Since the prisoner's dilemma is symmetric there is virtually only one reasonable cooperative goal which can be pursued, namely the selection of cooperative actions by both. The problem of choosing an ideal point does not arise. In this respect measure for measure is more general than tit-for-tat.

9.3. Typicity and success

The nearness of a particular strategy to the typical approach can be measured by a typicity index, the definition of which will not be discussed here. Typicity and success in the final tournament are connected by a rank correlation of + 0.616. This is significant at the 1%- level (two-tailed). A strategy tends to be
the more successful the more typical it is. The typical approach to the strategic problem seems to be not only reasonable, but also recommendable.

9.4. The depth of memory

Consider a game played in discrete time. Wherever a player has to make a decision a strategy specifies one of his possible choices (or a probability distribution over possible choices) depending on the past history of play. We say that a strategy has memory depth \( k \), if its specifications sometimes depend on the last \( k \) periods played but never on periods earlier than that. The memory depth of the typical strategies described above is 2. The specifications of a strategy of depth \( k \) may not only depend on what has been played in the last \( k \) periods, but also on the number \( t \) of the period, for which the decision is made. In our case the phase structure introduces a dependence on \( t \). The memory depth of only 2 is a remarkable simplicity property of the typical strategies.

10. Public good game strategies

In a study by Claudia Keser (1997) the strategy method was applied to a 25 times repeated public goods game with the following structure: Each of 4 players \( i = 1, \ldots, 4 \) can spend a total amount of 20 for a private activity and the production of a public good. The amount contributed to the public good by player \( i \) is \( b_i \). The amount spent for the private activity is the remainder \( a_i = 20 - b_i \). The amounts \( a_i \) and \( b_i \) must be integer numbers. Player \( i \)'s payoff \( U_i \) is as follows:

\[
U_i = 41a_i - a_i^2 + 15 \sum_{i=1}^{4} b_i.
\]

The first part, depending on \( a_i \), is the payoff derived from the private activity and the second part is the utility, due to the public good. It can be seen easily that \( b_i = 7 \) is a dominant strategy for this game.

Fifty economists from 13 European countries submitted strategies in the form of flow charts for 3 consecutive tournaments. A very interesting typical structure of the strategies for the final tournament was observed. In the following this structure will be described.

As in the case of the 20 times repeated duopoly discussed in the last section, a typical strategy prescribes different behaviour for three phases of the game. Fixed contributions are specified for the initial phase with \( b_i = 20 \) in period 1. In an end phase extending over up to 6 periods, the dominant strategy \( b_i = 7 \) is played. The behaviour in the main phase between the initial and the end phase can be described as reciprocation. \( b_i \) is chosen according to a formula of the
following type:

\[ b_i = \overline{b}_- + x. \]

Here \( \overline{b}_- \) is a rounded last period’s contribution average either over all 4 players or the other 3 players. The constant \( x \) added to the average contribution is mostly \(-1, 0 \) or \(+1\). Positive values are more frequent than negative values. In most cases \( x \) does not vary. Typically, the depth of memory is one.

It is not unreasonable to choose a positive value of \( x \). This helps to keep average contributions high, even if some people occasionally try to take advantage of the cooperativeness of others.

Thirty of the 50 final strategies make use of the type of formula shown above in the main phase. It is interesting to see that here, again, reciprocity can be observed. As in the measure for measure case, reciprocity seems to be instrumental rather than directly due to motivation. In the public good game we also observe an end effect.

There is an obvious similarity to measure for measure and tit-for-tat. Since the underlying game is symmetric the formation of a cooperative goal is easy. The most efficient symmetric cooperation is obtained if all players choose \( b_i = 20 \). The behaviour prescribed by the type of reciprocation formula shown above is similar to that specified by a measure for measure policy. There seems to be a common underlying general principle of *reciprocity guided strategic interaction in repeated games*.

Claudia Keser (1997) also looked at the connection between typicity and success. She obtained a rank correlation of \(+0.60\) (significant on the 1%-level, two-tailed). Here, too, the typical behaviour seems to be recommendable.

### 11. Strategies for incomplete information bargaining

Bettina Kuon (1994) has applied the strategy method to an incomplete information bargaining situation. The underlying game has the following structure: two players 1 and 2 bargain about the division of 100. Each of them can be of one of two types LO (low) or HI (high) with equal probability. LO has a zero conflict payoff. The conflict payoff of HI is a positive constant \( a \). Information is incomplete in the sense that each player knows only his type. He also knows that on both sides both types have probability \( 1/2 \). Bargaining followed a procedure of alternating bids. A randomly selected player would be the first one to act. Then after each period the initiative would shift to the other player. A player whose turn it was could either accept the last offer of the other player, or make a bid, or break off. In the case of break off conflict results. Payoffs were discounted at a rate of 1% for every bidding period after the first one.

32 participants took part in 3 tournaments. They had to write strategies for both LO and HI for 3 values of \( a \), namely \( a = 30, a = 45, \) and \( a = 60. \)
Table 4

The typical structure of final strategies in the incomplete information bargaining study by Kuon (1994). A horizontal arrow indicates a constant sequence and the downward arrow means that the sequence decreases at least in some periods. The numbers indicate the initial values of the sequences.

<table>
<thead>
<tr>
<th>$a$</th>
<th>Type</th>
<th>Demand sequence</th>
<th>Acceptance level sequence</th>
<th>Break off period</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>LO</td>
<td>50 →</td>
<td>50 →</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>HI</td>
<td>50 →</td>
<td>50 →</td>
<td>Fixed</td>
</tr>
<tr>
<td>45</td>
<td>LO</td>
<td>50 →</td>
<td>50 → 50 (\searrow)</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>HI</td>
<td>50 →</td>
<td>50 →</td>
<td>Fixed</td>
</tr>
<tr>
<td>60</td>
<td>LO</td>
<td>70, 80 (\searrow)</td>
<td>50 (\searrow) 50 (\searrow)</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>HI</td>
<td>70, 80 (\searrow)</td>
<td>60, . . . , 70 (\searrow)</td>
<td>Depends on opponent’s proposal</td>
</tr>
</tbody>
</table>

The typical strategy structure emerging in the last tournament is shown in Table 4.

A typical strategy has three parts. A demand sequence, an acceptance level sequence, and a break off rule. The payoff demand in period $t$ depends on $t$ only and not on other aspects of past history. Another sequence also depending only on $t$ specifies acceptance levels. The opponent’s last proposal is accepted if it offers a payoff at least as high as the current acceptance level.

A typical strategy never specifies break off for type LO. This is reasonable since LO cannot loose anything by continuing bargaining. For $a = 30$ and 45 break off is typically specified for a fixed period. For $a = 60$, however, the break off decision depends not only on $t$ but also on the opponent’s last proposal. Typically the memory depth is one.

For $a = 30$ and 45 typical strategies mostly specify demand and acceptance level sequences constant at 50. Only type LO in the case $a = 45$ also has acceptance level sequences decreasing in some periods. In the last tournament, agreement for $a = 30$ and 45 is predominantly reached at the equal split of 100. Conflict is very rare. In the face of the incompleteness of the information the equal split of 100 seems to be the only sensible fairness norm for $a = 30$ and even for $a = 45$.

In the case $a = 60$ agreement at the equal split of 100 cannot be reached, if one of the players is of type HI. In the following we shall concentrate on this case. Typically, a type HI has a demand sequence beginning with values between 70 and 80 and then staying constant or going down in some of the periods. The demand behaviour of LO is similar but the typical demand of types LO do not stay constant. The acceptance level of type HI typically stays constant at a level between 60 and 70. Some types LO have acceptance levels constant at 50 whereas others start with 50 but decrease their acceptance level later.
Conflict must be reached if both players are of type HI. Conflict also often is reached, if the combination of types is HI and LO. In this case, however, two typical strategies may reach agreement, if the acceptance level sequence of type LO is sufficiently decreasing.

In the case of the HI–LO combination, agreement by typical strategies is usually reached by sudden acceptance. Since the acceptance level sequence of a type LO is much below the demand sequence, type LO finally accepts a payoff offer markedly below his demand. This is meant by sudden acceptance. Kuon (1994) also observed this phenomenon in spontaneous play.

It cannot be denied that spontaneous play may be very different from the strategic interaction of programmed strategies. However, as in the case of the sudden acceptance phenomenon the comparison of both, where it has been made, also shows elements in common. Possible differences between spontaneous play and programmed strategies are not a valid argument against the strategy method. Both spontaneous behaviour and carefully devised strategies are of interest for behavioural decision and game theory.

Interestingly, the typical strategies do not show any reciprocity in the sense of a reaction to a concession of the opponent by a similar own concession. Demand sequences are prespecified. This lack of response to the opponent’s behaviour is maybe related to the special bargaining situation examined here. A large concession is possibly seen as a sign of weakness indicating that the concession maker is of type LO and therefore is not reciprocated. Some experimental results (Tietz and Weber, 1972; Kuon and Uhlich, 1993; Hennig-Schmidt, 1996) convey the impression that reciprocity is not necessarily completely absent in concession making; whether this is the case or not, probably depends on the bargaining situation.

12. Features of bounded rationality suggested by experiments discussed here

In the following the conclusions drawn from the experimental evidence presented here will be summarized. Of course, the statements of this section are interpretations of observed experimental results which may be seen differently by other researchers. As soon as one moves from the description of mere fact in the direction of a theoretical appraisal, something more comes in than a mere description of statistical relationships.

12.1. Superficial analysis

It has been argued that decision processes start with a superficial analysis of the situation. Later, further reasoning may or may not deepen the analysis to some extent. This explains why often presentation effects are observed as in the examples discussed in Sections 3 and 4. A superficial look at the situation does
not necessarily reveal what is really relevant and a satisfactory solution of the problem may be found before this becomes clear.

12.2. Strategic reasoning and strategy construction based on reciprocity and fairness

As we have seen, fairness and reciprocity are very important for strategic reasoning. This is maybe only partially due to motivational factors. Fairness standards may serve as focal points in bargaining and strategic interaction. Sometimes reasoning does not aim at fairness, but fairness standards, nevertheless, have an important benchmark role. Thus, in the case of equal division payoff bounds (Section 5) equity considerations determine a lower bound for the payoff of the more powerful party. Reciprocation can also be seen as a strategic device to achieve cooperation at a cooperative goal based on a fairness standard. This is suggested by the strategy studies discussed in Sections 9 and 10. There we have also expressed the idea that what is observed in these studies may be due to a general principle of reciprocity guided strategic interaction in repeated games.

12.3. Avoidance of circular concepts in step-by-step strategic reasoning

Basic concepts in game theory are often circular in the sense that they are based on definitions by implicit properties. We have seen this for the example of quotas for three-person coalition games in Section 5. Boundedly rational strategic reasoning seems to avoid circular concepts. It directly results in a procedure by which a problem solution is found. Each step of the procedure is simple, even if many case distinctions by simple criteria may have to be made.

12.4. Ex-post rationality as a strong influence on adaptation

Bayesian rationality is ex-ante in the sense that it looks to the future and optimises on the basis of current beliefs. Contrary to this, boundedly rational behaviour is often based on ex-post rationality. Experiences are interpreted in the light of qualititative beliefs and behaviour is adjusted accordingly without any attempt at forward looking optimization. This is shown by the empirical success of learning direction theory (Sections 6–8).

12.5. Presence of both adaptive and analytical approaches to repeated decision tasks

Repeated decision tasks may be approached differently by different subjects. Some take an analytical approach which computes the decision on the basis of numerical information present in the situation. Analytical approaches are not necessarily attempts at optimization. They may be based on natural aspiration
levels for the avoidance of risk, like loss avoidance or asset conservation in the winner’s curse task of Section 8.

Other subjects are adapters in the sense that their decisions are based on uncalculated adaptation which, however, may be influenced more or less by ex-post rationality. Especially, in situations where analytical approaches are difficult to find, one must expect that this is the more frequent mode of behaviour.

12.6. No quantitative expectations and no optimization in typical repeated game strategies

Consider a two-person game in normal form with a unique equilibrium point and Pareto-optimal payoff combinations which are better for all players. Suppose that such a game is repeated many times. The typical approach to a situation of this kind seems to involve neither quantitative expectations nor optimization. Instead of this, cooperative goals are formed on the basis of fairness criteria and efficiency considerations. These goals are pursued with the help of measure for measure strategies or similar methods of reciprocation (Sections 9 and 10).

12.7. Low memory depth

The three applications of the strategy method discussed in Sections 9–11 have something in common. The typical strategies have a low depth of memory (2 in the duopoly study of Sections 9 and 1 in both other cases). The same is true for other investigations not reported here (e.g. Keser, 1992). Only sparing use of history dependence seems to be made by typical strategies for dynamic game situations. However, the prescriptions of typical strategies are not stationary in the sense that behaviour exclusively depends on what happened in the memory span. Behaviour is also directly time dependent.

References


