Journal of Economic Behavior and Organization 7 (1986) 47–70. North-Holland

END BEHAVIOR IN SEQUENCES OF FINITE PRISONER'S DILEMMA SUPERGAMES

A Learning Theory Approach

Reinhard SELTEN and Rolf STOECKER

University of Bonn, D-5300 Bonn 1, FRG

Received 7 June 1983, final version received 15 January 1985

A learning theory is proposed which models the influence of experience on end behavior in finite Prisoner's Dilemma supergames. The theory is compared with experimental results. In the experiment 35 subjects participated in 25 Prisoner's Dilemma supergames of ten periods each against anonymous opponents, changing from supergame to supergame. The typical behavior of experienced subjects involves cooperation until shortly before the end of the supergame. The theory explains shifts in the intended deviation period. On the basis of parameter estimates for each subject derived from the first 20 supergames, successful predictions could be obtained for the last five supergames.

1. Introduction

In a finite Prisoner's Dilemma supergame the same game is repeated for a fixed number of times known to both players in advance. It is well known that such games have a definite game theoretical solution which prescribes non-cooperative behavior in all periods of the supergame. However, experimental behavior does not conform to this theoretical prediction. Early experiments with finite Prisoner's Dilemma supergames [Rapoport–Dale (1966), Morehous (1966), Lave (1965)] already have shown that subjects sufficiently often choose the cooperative alternative. At first glance, the situation seems to be similar to comparable experiments where the number of periods is not known to players in advance. However, such games are more akin to the infinite Prisoner's Dilemma supergame which permits equilibrium points resulting in cooperative behavior.

More recently, experiments have been performed where subjects played the same finite Bertrand Duopoly or Prisoner's Dilemma supergames many times against changing anonymous opponents: the subjects played against this opponent within one supergame, but could not expect to meet the same opponent again in a later supergame [Stoecker (1980, 1983)]. The results show that subjects develop a pattern of behavior which may be described as tacit cooperation until shortly before the end of the supergame followed by non-cooperative choices until the end. As soon as one of the players deviates to non-cooperative behavior the other reacts with non-cooperative choices and cooperation is not established any more. This pattern of cooperation followed by an end-effect is observed in almost all supergames between experienced players. Obviously, straightforward game theoretical reasoning cannot explain experienced behavior in finite Prisoner's Dilemma supergames. One could try to account for this by the assumption that the players' utility is different from the monetary rewards. Players may for example value cooperation as such and therefore refrain from non-cooperative behavior in spite of monetary incentives. Such explanations fail to be convincing in view of the end-effect which indicates that monetary incentives are stronger than the desire to be cooperative for those who deviate to non-cooperative behavior. A more detailed discussion of this point can be found elsewhere [Selten (1978)].

An attempt to explain cooperation followed by an end-effect as the result of fully rational behavior may be based on the idea of slightly incomplete information on the other player's payoff [Kreps, Milgrom, Roberts and Wilson (1982)]. However, such theories predict the mature pattern of behavior already for inexperienced subjects. This does not agree with experimental observations. Subjects first have to learn cooperation and only afterwards do they discover the end effect. Descriptive theories cannot ignore the limited rationality of human subjects.

In this paper we shall present a learning theory approach to the explanation of end behavior in finite Prisoner's Dilemma supergames. We assume that players are motivated by monetary rewards. However, we do not assume optimizing behavior. Our theory is based on a Markov learning model where subjects change their intention to deviate from cooperation in a certain period with transition probabilities depending on experience in the last supergame.

There are obvious analogies between learning and evolution. The evolution of cooperative behavior in the infinite Prisoner's Dilemma supergame has been discussed in Axelrod's (1984) stimulating book. We shall not comment on this work in detail, since here we are concerned with the end-effect, a phenomenon which is excluded by the nature of the infinite supergame.

We shall also present the result of an experiment where each of 35 subjects participated in 25 Prisoner's Dilemma supergames of ten periods each. The data exhibit remarkable individual differences between subjects. Therefore, the parameters of the learning model are fitted separately for each subject.

If one allows for random perturbances which occasionally result in reactions which are excluded by the model, the learning theory could be viewed as roughly in agreement with the behavior of 34 of 35 subjects (one subject behaves in a rather chaotic way). The intention to deviate from cooperation can be moved forward or backwards in time or remain constant from one supergame to the next. The learning model always excludes either the forward shift or the backwards shift. In only 21 out of 585 cases could a reaction in the excluded direction be observed in the data.

A careful look at the data suggests a distinction between different groups of subjects which differ with respect to the degree of conformance between the learning model and observed behavior. In the last 13 supergames where all of the 34 subjects already had some experience with the end effect, 18 subjects never showed any response excluded by the model. However, four of these subjects had constant intentions to deviate in these supergames and therefore could be explained in a simpler way. A slightly more general model would be compatible with all responses of nine further subjects in the last 13 supergames. Each of the remaining seven subjects exhibits only one response in the direction excluded by the model in these supergames.

Statistical computations support the impression that the general ideas underlying our learning model provide a reasonable picture of observed behavior. Computer simulation based on individually estimated parameters produces results which tend to agree with the experimental observations.

2. Experimental procedure

The experiments are based on the Prisoner's Dilemma game shown in fig. 1. The payoffs shown are in German Pfennigs (one German Mark equals 100 Pfennig). In each supergame the game of fig. 1 was repeated ten times. Each subject played 25 supergames. They were told that they played against the same opponent within one supergame but against different opponents in different supergames.

Subjects were placed in separate rooms. They did not communicate with each other. The experimenters asked for each decision by intercom and announced the opponent's decision at the end of each period. Subjects kept records of previous decisions and gains.

		Player 2	
		HP	NP
Player 1	HP	60 60	−50 145
	NP	145 −50	10 10

Fig. 1. The game used in the experiment – payoffs for player 1 are shown in the upper left corner and payoffs for player 2 are shown in the lower right corner. The strategies were introduced as high price (HP, hoher Preis) and low price (NP, niedriger Preis).

The experimenters did not only ask for the subjects' decisions but also for expectations on opponents' decisions. Moreover, the subjects were required to write down reasons for each period-decision.

Subjects came to the laboratory for two afternoon sessions of four hours each. Part of the time was used for introductory explanations and for tests on altruism and risk-taking. These tests are not described here since their results will not be used in the evaluation of the experiments. The actual playing of the 25 supergames took about four hours. After some experience one period took less than a minute. It is important to point out that payoff incentives are quite high relative to such a short time span (see fig. 1).

The experimenters tried to create the impression that 26 subjects participated in each session and that they never would meet the same opponent again in a later supergame. Actually, in each session there were only 12 subjects. Unknown to the subjects the experimental design separated the 12 subjects into two groups of six. Each subject played only against changing opponents among the other five subjects in his group (see appendix A).

It was intended to have six groups of six subjects. However, in one of the second sessions one of the subjects did not come and was substituted by a fixed strategy administrated by the experimenter. This strategy prescribes cooperation until a non-cooperative choice of the opponent is observed and non-cooperative behavior from then on. This means cooperation up to the end if the opponent does not deviate in the first nine periods. There were actually three subjects who followed this policy and explicitly explained it in their written reasons.

The subjects were male and female economics and business administration students of the University of Bielefeld in their first year.

3. Experimental results

In the course of the experiment subjects learned a pattern of behavior involving cooperation followed by a non-cooperative end-effect. In order to make this statement precise we introduce the following definitions.

Definition 1. The play of a supergame is called *cooperative* if the following three conditions are satisfied:

(a) In the first m periods, where m is at least four, both players choose the cooperative alternative HP.

(b) In period $m+1$ (for $m<10$) at least one player chooses the non-cooperative alternative NP.

(c) In all periods $m+2,\ldots,10$ (if there are any) both players choose the non-cooperative alternative.

Note that this definition does not exclude the case $m=10$ where both players cooperate from the beginning to the end. Admittedly, the requirement $m \geqq 4$ is to some extent arbitrary. However, it is necessary to have some criterion in order to distinguish plays with an end-effect from plays where no cooperation has been reached at all. Moreover, in the experiment no additional case would have to be classified as cooperative if in (a) the condition $m \geqq 4$ is weakened to $m \geqq 1$.

An end-effect may also occur in plays where cooperation has been reached only after initial non-cooperation. In order to capture this possibility we adopt the following definition of an *end-effect play:*

Definition 2. An end-effect play is characterized by three conditions, (a′), (b) and (c).

(a′) Both players choose the cooperative alternative in at least four consecutive periods $k, \ldots, m$.

The conditions (b) and (c) are the same as in the definition of a cooperative play. By definition, a cooperative play is also an end-effect play.

We say that a supergame belongs to round n if it was played as the nth supergame by the subjects. Since there were 36 players [including one simulated player for rounds (9) to (25)] each round has 18 plays. Table 1 shows for every round how many plays were end-effect plays and how many of those were cooperative ones. This is indicated for each of the six groups of interacting subjects separately.

Table 1 shows that for experienced subjects most plays tend to be cooperative; however, there are some subjects who sometimes tried to gain an advantage by choosing the non-cooperative alternative in the first period hoping that the other player would not retaliate. Such behavior results mostly in end-effect plays which fail to be cooperative in the sense of the definition given above. Group 1 contains one subject who seemed to have great difficulty understanding the situation until round (21). In the first 20 rounds his behavior was highly irregular. In the last five rounds 99 percent of the plays are end-effect plays and 96 percent are cooperative plays. Appendix B gives a detailed account of the observed end-effect behavior for all subjects separately.

The learning model to be explained later contains an intended period of deviation as an internal state of the subject. In all cases where a subject deviated before the opponent or simultaneously with the opponent the intended deviation period is nothing else than the observed deviation period. However, if the opponent deviated before the subject the intended deviation period is not uniquely determined by the decisions observed in the play. This situatioh occurs in 198 out of 621 cases. In 84 of these cases the reasons written down by the subjects indicated the intended deviation period. In the

Table 1
Number of end-effect plays (EEP) and cooperative plays (CP) by rounds and subject groups.

	Group													
	I		II[a]		III		IV		V		VI		Total	
Round	EEP	CP	EEP	CP	EEP	CP	EEP	CP	EEP	CP	EEP	CP	EEP	CP
(1)									2	2	1		3	2
(2)									1	1			1	1
(3)	1				1	1	1	1	2	2	1		6	4
(4)					1	1	2	2	3	1	2	2	8	6
(5)					1	1	1	1	2		1	1	5	3
(6)					1	1	2	2	3	1	1	1	7	5
(7)			1		1	1	2	2	2	1	1	1	7	5
(8)	1		1		2	2	2	2	3	1	2	2	11	7
(9)	3	3	3	2	1	1	2	2	2	1	2	2	13	11
(10)	2	2	3	3	1	1	1	1	2	1	2	2	11	10
(11)	2	2	3	3	1	1	2	2	1		2	2	11	10
(12)	2	2	3	2	2	2	2	2	2		2	2	13	10
(13)	2	2	2	2	3	3	3	3	3	3	3	3	16	16
(14)	2	2	2	1	3	3	3	3	3	3	3	3	16	15
(15)	2	2	3	1	3	3	3	3	3	3	3	3	17	15
(16)	2	2	3	2	3	3	3	3	3	3	3	3	17	16
(17)	2	2	3	2	3	3	3	3	3	3	3	3	17	16
(18)	2	2	3	2	3	3	3	3	3	3	3	3	17	16
(19)	2	2	3	2	3	3	3	3	3	3	3	3	17	16
(20)	2	2	3	2	3	3	3	3	3	3	3	3	17	16
(21)	3	3	3	1	3	3	3	3	3	3	3	3	18	16
(22)	3	3	3	2	3	3	3	3	3	3	3	3	18	17
(23)	3	3	3	3	3	3	3	3	3	3	3	3	18	18
(24)	3	3	3	3	3	3	3	3	3	3	3	3	18	18
(25)	3	3	2	2	3	3	3	3	3	3	3	3	17	17

[a]This group contains the simulated player for rounds (9) to (25).

remaining 114 cases an estimate of the intended deviation period was based on reported expectations together with observed behavior and reasons from previous rounds.

Table 2 shows the means and standard deviation of intended deviation periods in end-effect plays for all 35 subjects who participated in rounds (13) to (25) for rounds and groups separately. In the last 12 rounds all subjects can be described as experienced in the sense that each of them had been in at least one end-effect play in an earlier round. In the computations deviation period 11 was assigned to those cases where the subject did not intend to deviate at all.

It can be seen that the end-effect has a clear tendency to shift to earlier periods in the last 13 supergames. For each of the six groups the Spearman rank correlation coefficient between the mean of the intended deviation period and the number of the supergame is negative and significant at the 0.1 percent level (two-sided) for the last 13 supergames.

Table 2

Means and standard deviation of intended deviation period in end-effect plays for rounds and groups, separately.

	Round													Spearman rank correlation coefficient[a]
	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	
Group I														
Mean	7.8	7.8	7.5	7.0	6.8	7.0	6.5	6.5	6.2	6.0	5.7	5.7	5.5	−0.99
Standard deviation	0.96	0.5	0.58	0.0	0.5	0.82	0.58	0.58	0.41	0.0	0.52	0.82	0.55	
Group II[b]														
Mean	9.2	8.7	8.5	8.5	8.3	8.5	8.2	8.0	7.7	7.8	8.0	7.7	7.5	−0.95
Standard deviation	1.3	1.5	1.2	1.2	1.4	1.2	1.5	1.6	1.6	1.6	1.6	1.9	2.4	
Group III														
Mean	10.2	10.0	10.0	9.8	10.0	9.8	9.8	9.8	9.7	9.7	9.7	9.7	9.0	−0.95
Standard deviation	0.98	1.3	1.3	1.8	1.6	1.8	1.8	1.8	2.1	2.1	2.1	2.1	2.3	
Group IV														
Mean	7.3	7.8	7.7	7.3	7.7	6.8	6.8	7.0	6.7	6.5	6.2	5.8	6.0	−0.93
Standard deviation		2.0	1.0	1.0	1.9	0.8	0.8	0.9	0.5	0.6	0.8	1.0	0.6	
Group V														
Mean	10.0	10.0	10.0	9.3	9.5	9.2	9.3	9.3	9.0	8.8	9.0	8.5	8.3	−0.95
Standard deviation	1.3	1.3	1.3	1.4	1.2	1.2	1.0	1.0	1.1	1.2	0.6	0.6	0.5	
Group VI														
Mean	10.3	10.2	10.2	10.2	9.8	9.8	9.7	9.0	8.8	8.8	8.8	8.3	8.2	−0.99
Standard deviation	0.8	0.8	1.0	1.0	1.0	0.8	0.8	0.6	0.4	0.4	0.8	0.5	0.8	
Total														
Mean	9.2	9.1	9.0	8.7	8.7	8.7	8.5	8.3	7.9	7.9	7.8	7.5	7.4	−1.00
Standard deviation	1.5	1.6	1.5	1.6	1.7	1.6	1.7	1.6	1.7	1.7	1.8	1.8	1.8	

[a] Values are rounded but rank correlations are computed for exact means.

[b] This group contains the simulated player.

Even if it is very clear from the data that there is a tendency of the end-effect to shift to earlier periods, it is not clear whether in a much longer sequence of supergames this trend would continue until finally cooperation is completely eliminated. It is also possible that the mean of the intended deviation period would have a tendency to decrease in such a way that it finally converges to a stable limit. It is interesting to note that some groups show a very strong shift to earlier periods. In round (25) the means of groups I and IV are at 5.5 and 6.0, respectively, whereas this mean is at 9.0 for group III.

4. A learning theory of end-effect behavior

Our learning model contains the intended deviation period k as the internal state of the subject. A subject is assumed to change his internal state from round t to round $t+1$ according to constant transition probabilities. Each subject is characterized by three parameters α, β and γ.

If the subject observes that in round t his opponent deviated earlier than he intended to deviate, then with probability α he will shift his intended deviation period k to $k-1$. The probability that the subject's internal state remains k in this case is $1-\alpha$.

If the subject observes that in round t his opponent deviated at the same period k as he did, then with probability β the subject's intended deviation period will be shifted to $k-1$ and with probability $1-\beta$ it will stay where it is.

If the subject observes that in round t he deviated before the opponent, then with probability γ he will shift the intended deviation period to $k+1$ provided we have $k<10$. With probability $1-\gamma$ he will not change his internal state. There is no change for $k=10$. The assumptions of the learning model are summarized by table 3.

In the explanations given above it was assumed that in round t the subject experienced an end-effect play. It is assumed that no change of intention

Table 3

Transition probabilities from round to round for the intended deviation period of a subject. k is the intended deviation period in round t.

	Intended deviation period in round $t+1$		
Subject's intended deviation period in round t	One period sooner	Unchanged	One period later
Later than his opponent	α	$1-\alpha$	
Together with his opponent	β	$1-\beta$	
Sooner than his opponent		$1-\gamma$ for $k<10$, 1 for $k=10$	γ for $k<10$, 0 for $k=10$

takes place after a round which did not result in an end-effect play. This convention is unimportant for our theoretical derivations and simulations but it has some minor significance for the interpretation of our data.

We now proceed to discuss our motivations behind the assumption of transition probabilities. A subject who has observed that his opponent deviated earlier than he himself intended to do will think that it might have been better to deviate earlier. The same is true to a lesser degree if the opponent deviated in the same period as he did. Therefore, it is reasonable to assume $\alpha \geqq \beta > 0$. In both cases there is no reason to shift the intention to deviate to later periods.

Now consider a subject who in round t deviated in a period $k < 10$ and observed that his opponent did not deviate from cooperation up to period k. He does not know exactly in which period the opponent intended to deviate. Therefore, it could have been better to deviate in a later period. We may for example look at $k = 8$. The subject does not know whether the opponent intended to deviate in period 9, 10 or not at all. In the latter two cases a deviation in period 9 would have been more advantageous. It is plausible to assume that this kind of uncertainty produces a tendency to shift the deviation periods towards the end of the supergame. Of course, for $k = 10$ there is no such uncertainty and the subject must conclude that it was right to deviate in the last period if he observed that the opponent cooperated up to the end.

In the mathematical learning models considered in the literature [see for example Restle–Greeno (1970), Bush–Mosteller (1955)] it is generally clear whether reinforcement of behavior has taken place or not. However, in a situation where a subject deviated earlier than his opponent in a period $k < 10$ he does not know whether his decision was right or wrong.

Unobserved features of the opponent's behavior prevent him from having a clear experience of success or failure. However, he knows that here is a possibility that his decision was wrong.

Our specification of the general ideas explained above contain certain simplifying assumptions. We exclude the possibility that the intended deviation period shifts by more than one period. It is, of course, easy to construct a more general learning model where shifts of two or three periods are permitted. However, the scarcity of data forces us to restrict our attention to models with as few parameters as possible.

In a situation where a subject deviated earlier than his opponent his uncertainty on the nature of his experience is the greater the earlier his deviation was. The more periods there are after the deviation until the end of the supergame, the more chances there are that the deviation was too early. Therefore, one could think of making γ dependent on k in such a way that γ increases with decreasing k. This would be a theoretically attractive modification of the model but also here the necessary increase of the number of parameters prevents us from comparing such models with the data.

5. Theoretical considerations

In the following we shall look at the consequences of our theory in an idealized situation which is not that of the experiment.

Consider a very large population of subjects where the parameters α, β and γ are the same for all subjects. With this population we imagine a fictitious experiment over a very long sequence of supergames. At each round the subjects are paired randomly.

We may ask the question how in this system the probabilities of intended deviation periods evolve. In order to describe the process which governs the evolution of these probabilities we introduce the following notations:

$$p_k^t$$

is the probability that in round t a randomly chosen subject has the intention to deviate in period k, where $k=11$ stands for the intention not to deviate at all ($k=1,\dots,11$),

$$S_k^t=\sum_{m=1}^{k} p_m^t$$

is the probability that in round t a randomly chosen subject has the intention to deviate in periods $1,\dots,k$, and

$$\alpha,\ \beta \text{ and } \gamma$$

are the parameters of table 3.

It is useful to look at the situation in a way which is similar to that of a Markov chain. We may ask the following question: what are the probabilities that a subject will intend to deviate in period $k-1$, k or $k+1$ in round $t+1$ if he intended to deviate in period k in round t? These 'transition probabilities' can be arranged in a matrix where columns correspond to intended deviation periods in round t and rows correspond to intended deviation periods in round $t+1$. A part of this matrix is shown in table 4.

With the help of table 3 it can be seen easily that the transition probabilities are in fact those shown in table 4. From what has been said up to now it is clear that the probabilities p_k^{t+1} are determined by the following equation system:

$$p_{11}^{t+1}=[(1-\alpha)S_{10}^t+(1-\beta)p_{11}^t]p_{11}^t,$$

$$p_{10}^{t+1}=[\alpha S_{10}^t+\beta p_{11}^t]p_{11}^t+[(1-\alpha)S_9^t+(1-\beta)p_{10}^t$$
$$+(1-\gamma)(1-S_{10}^t)]p_{10}^t+\gamma(1-S_9^t)p_9^t,$$

$$p_9^{t+1}=[\alpha S_9^t+\beta p_{10}^t]p_{10}^t+[(1-\alpha)S_8^t+(1-\beta)p_9^t$$

$$+(1-\gamma)(1-S_9^t)]p_9^t+\gamma(1-S_8^t)p_8^t,$$

$$p_2^{t+1}=[\alpha S_2^t+\beta p_3^t]p_3^t+[(1-\alpha)p_1^t+(1-\beta)p_2^t$$

$$+(1-\gamma)(1-S_2^t)]p_2^t+\gamma(1-p_1^t)p_1^t,$$

$$p_1^{t+1}=[\alpha p_1^t+\beta p_2^t]p_2^t+[p_1^t+(1-\gamma)(1-p_1^t)]p_1^t.$$

Table 4
Transition probabilities for a subject between rounds t and $t+1$ (explanation in the text).

	Round t				
Round $t+1$	(11)	(10)	(9)	(8)	(7)
(11)	$(1-\alpha)S_{10}^t+$ $(1-\beta)p_{11}^t$	0	0	0	0
(10)	$\alpha S_{10}^t+\beta p_{11}^t$	$(1-\alpha)S_9^t+$ $(1-\beta)p_{10}^t+$ $(1-\gamma)(1-S_{10}^t)$	$\gamma(1-S_9^t)$	0	0
(9)	0	$\alpha S_9^t+\beta p_{10}^t$	$(1-\alpha)S_8^t+$ $(1-\beta)p_9^t+$ $(1-\gamma)(1-S_9^t)$	$\gamma(1-S_8^t)$	0
(8)	0	0	$\alpha S_8^t+\beta p_9^t$	$(1-\alpha)S_7^t+$ $(1-\beta)p_8^t+$ $(1-\gamma)(1-S_8^t)$	$\gamma(1-S_7^t)$
(7)	0	0	0	$\alpha S_7^t+\beta p_8^t$	$(1-\alpha)S_6^t+$ $(1-\beta)p_7^t+$ $(1-\gamma)(1-S_7^t)$

Starting from an initial distribution $(p_1^1,\ldots,p_{11}^1)$ the probability vector $(p_1^t,\ldots,p_{11}^t)$ can be computed for every round t. We may ask the question whether this probability vector converges to a stable equilibrium distribution.

We shall not try to give a rigorous theoretical answer to the question of convergence. However, we have run a large number of numerical computations whose results show a definite pattern which will be described in the following. It must first be pointed out that the difference equations have the following property: If $p_k^t=0$ holds for $k=m,\ldots,11$ for some $t=t_0$ then the same conditions will be satisfied for every $t>t_0$. For this reason alone, the result of the simulation cannot be completely independent of the initial conditions.

However, if p_{10}^{1} and p_{11}^{1} are sufficiently high, the results of our computations do not depend on the exact initial conditions. The results obtained for $p_{11}^{1}=1$ do not change as long as the initial conditions remain in a neighbourhood of this extreme case. The size of this neighbourhood depends on the parameters but for most cases it seems to be quite large.

Table 5 shows numerical results for selected parameter combinations. All these computations have been run starting from the initial condition $p_{11}^{1}=1$. Our experimental results suggest that subjects learn to cooperate before they learn to show any end-effect. Therefore, the assumption $p_{11}^{1}=1$ is quite reasonable.

All the computations with $p_{11}^{1}=1$ converged to a stationary distribution which was always mostly concentrated either at the end or at the beginning of the supergame. In table 5 either the first three or the last three periods obtained at least 97 percent of the total mass of the probability.

The parameter combinations of table 5 are arranged in groups with constant β and γ and increasing α. If α is small in comparison to $\gamma-\beta$ the distribution is mostly concentrated near the end of the supergame. With increasing α this concentration becomes less pronounced until a critical value of α is reached beyond which the stationary distribution is mostly concentrated at the beginning of the supergame. As can be seen in table 5 the critical value for α is a little below $\gamma-\beta$. It can be checked analytically without much difficulty that for $\alpha+\beta=\gamma$ the distribution $p_1=p_2=0.5$ is stationary. In fact, in cases with $\alpha+\beta=\gamma$ the process converges to this distribution.

The results of these computations suggest an abrupt change of the stationary distribution at the critical value of α. In table 5 the critical values of α are enclosed by intervals of the length of 10^{-3}. It can be seen that within this small interval the stationary distribution reached by the process changes drastically. The change is somewhat less pronounced if the interval is narrowed down to the length of 10^{-7} but even there p_5 and p_6 are practically 0 before and after the change from a concentration at the end to a concentration at the beginning.

In the experiments a group of interacting subjects had only six members and the parameter values varied considerably from subject to subject. Moreover, the experimental pairings of subjects are not random, but follow a repetitive scheme (see appendix A). Nevertheless, the model applied to the experimental situation can be looked upon as a Markov chain with a suitably defined state space. The highest among the intended deviation periods of the subjects cannot increase from one supergame to the next, but if α and β are positive and λ is smaller than one for all subjects, then there always is a positive probability that the highest intended deviation time will

Table 5
Stable probability distributions over intended deviation periods for selected parameter combinations.

α	β	γ	p_1	p_2	p_3	…	p_8	p_9	p_{10}
0.100	0.1	0.1	1.000	—	—		—	—	—
0.100	0.1	0.4	—	—	—		0.017	0.250	0.733
0.200	0.1	0.4	—	—	—		0.038	0.346	0.615
0.264	0.1	0.4	—	—	—		0.135	0.520	0.340
0.265	0.1	0.4	0.119	0.509	0.371		—	—	—
0.300	0.1	0.4	0.500	0.500	—		—	—	—
0.300	0.1	0.5	—	—	—		0.034	0.356	0.610
0.355	0.1	0.5	—	—	—		0.105	0.520	0.373
0.356	0.1	0.5	0.098	0.512	0.390		—	—	—
0.400	0.1	0.5	0.500	0.500	—		—	—	—
0.500	0.1	0.5	0.990	0.010	—		—	—	—
0.200	0.2	0.5	—	—	—		0.086	0.400	0.511
0.300	0.2	0.5	0.500	0.500	—		—	—	—
0.400	0.2	0.5	0.667	0.333	—		—	—	—
0.100	0.1	0.6	—	—	—		0.005	0.167	0.829
0.200	0.1	0.6	—	—	—		0.007	0.201	0.791
0.300	0.1	0.6	—	—	—		0.013	0.256	0.731
0.400	0.1	0.6	—	—	—		0.030	0.364	0.606
0.449	0.1	0.6	—	—	—		0.095	0.533	0.370
0.450	0.1	0.6	0.087	0.522	0.391		—	—	—
0.200	0.2	0.6	—	—	—		0.044	0.333	0.622
0.300	0.2	0.6	—	—	—		0.083	0.417	0.498
0.353	0.2	0.6	—	—	—		0.188	0.516	0.282
0.354	0.2	0.6	0.151	0.500	0.349		—	—	—
0.400	0.2	0.6	0.500	0.500	—		—	—	—
0.500	0.2	0.6	0.667	0.333	—		—	—	—
0.600	0.2	0.6	0.993	0.007	—		—	—	—
0.100	0.1	0.7	—	—	—		0.003	0.143	0.854
0.200	0.1	0.7	—	—	—		0.004	0.167	0.828
0.300	0.1	0.7	—	—	—		0.006	0.203	0.791
0.400	0.1	0.7	—	—	—		0.011	0.259	0.730
0.500	0.1	0.7	—	—	—		0.028	0.370	0.602
0.543	0.1	0.7	—	—	—		0.073	0.517	0.409
0.544	0.1	0.7	0.072	0.515	0.413		—	—	—
0.600	0.1	0.7	0.500	0.500	—		—	—	—
0.700	0.1	0.7	0.990	0.010	—		—	—	—
0.440	0.2	0.7	—	—	—		0.164	0.525	0.302
0.441	0.2	0.7	0.133	0.506	0.361		—	—	—
0.529	0.2	0.8	—	—	—		0.144	0.530	0.320
0.530	0.2	0.8	0.119	0.509	0.371		—	—	—
0.445	0.3	0.8	—	—	—		0.213	0.510	0.257
0.446	0.3	0.8	0.165	0.495	0.340		—	—	—
0.538	0.4	0.9	—	—	—		0.227	0.504	0.244
0.539	0.4	0.9	0.173	0.492	0.335		—	—	—
0.100	0.1	1.0	—	—	—		—	0.100	0.900

decrease by one. Therefore, with probability one the highest deviation time will finally decrease to one in an infinite sequence of supergames. This means that in the long run behavior converges to complete non-cooperation.[1]

Even if the possibility of convergence to a stationary distribution exhibiting a stable end-effect is excluded by our model, if it is applied to a finite population of subjects, the computations for the idealized situation with an infinite population are not without interest. They suggest that in the finite situation convergence to non-cooperation may be very slow if the parameter values for α and β are relatively small and those for λ are relatively large. Moreover, in the light of the computations for the infinite case one must consider the possibility that the conclusion on convergence to non-cooperation is not robust with respect to slight misspecifications of the learning model. Suppose that the probabilities for the excluded transitions are not really zero, but only relatively small. It is reasonable to expect that under this condition a stationary distribution exhibiting a stable end-effect might be obtained for suitable parameter combinations in the finite case.

However, it can be expected that the results obtained for the large group case with equal parameters are indicative for what can be expected to happen in the experimental situation if the model is correct.

6. Subject differences

After the theoretical considerations of the last section we shall now turn our attention to some important features of our experimental results. The behavioral assumptions of our model do not fit all subjects equally well.

There are several deviations from the theoretical behavior which may occur. Some subjects occasionally change the intended deviation period by more than one step from one round to the next. Even if this is not a deviation from the spirit of our model, it is a deviation from the specification which had to be used in view of the scarcity of observations. A more serious deviation which occurred only rarely is a shift of the intended deviation period in the wrong direction. Some subjects do not show any reaction excluded by the model but they have a constant intended deviation period. An intended deviation period which does not change over time can be explained in simpler and possibly more adequate ways than by our model.

Table 6 distinguishes several groups of subjects according to the conformance of their behavior to the model in the last 13 rounds. We restricted this evaluation to the second half of the experiment since there almost all subjects had learned to cooperate. Only a subject who has learned to cooperate can experience an end-effect play.

It can be seen that only 20 percent of all subjects show a shift in the

[1] We are grateful to an anonymous referee who directed our attention to this point.

Table 6
Grouping of subjects by conformance of behavior to the model in the last 13 rounds. Each subject is listed only once.

Subject category	Number of subjects	Number of deviations	Total number of cases
No deviations from the model and varying intended deviation period	14	—	158
Constant intended deviation period[a]	4	—	48
Shifts of more than one step but no other deviations from the model	9	14	108
Shifts in the wrong direction[b]	7	7	81
Failure to learn cooperation[c]	1	—	—

[a] Three of these subjects never intended to deviate as a matter of principle.

[b] Two of these subjects also showed jumps of more than one step.

[c] Unlike all other subjects this subject did not learn to cooperate in the first half of the experiment. He began to experience end-effect plays only in the last five rounds.

wrong direction. For each of these seven subjects such a shift occurs only once.

Three subjects always had the intention to cooperate until the last period. The protocols written by these subjects show that they did this on principle. Obviously, the learning model does not adequately describe the motivations of these subjects even if it formally fits their behavior. One subject always intended to deviate in period (8). He thought that this is the optimal deviation period. Since this opinion was based on experience rather than theoretical reasoning, his behavior may be adequately explained by the model.

Up to occasional deviations, a learning theory approach like that of our model seems to offer a plausible explanation for the behavior of the vast majority of subjects. A fundamentally different theory may be required for those three subjects who never intended to deviate in rounds (13) to (25) as a matter of principle. The learning model cannot be compared with the behavior of the subject who failed to learn to cooperate in rounds (1) to (20). With these exceptions the learning model can be proposed as an idealized picture of observed behavior. The next section will try to throw further light on the extent to which the data agree with the learning model.

7. Parameter estimates

The observations for rounds (1) to (20) have been used in order to obtain parameter estimates $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$ of α, β and γ, respectively, for all subjects with the exception of subject 1 who failed to learn to cooperate in rounds (1) to (20). On the basis of these parameter estimates Monte Carlo simulations have been run in order to generate predictions for rounds (21) to (25) which can be compared with the data. The Monte Carlo simulations will be discussed in section 9.

As far as possible relative frequencies of transitions have been taken as parameter estimates. In the determination of relative frequencies shifts of more than one step in the right direction have been counted as if they were shifts of one step. Shifts in the wrong direction have been counted as if they were cases of unchanged deviation periods.

The parameter estimates are shown in table 7. In the three cases indicated by the superscript 'a', relative frequencies were not available due to lack of observations and estimates had to be obtained in another way.

It is plausible to assume $\alpha \geqq \beta$ since there is more reason for a shift to an earlier deviation if the opponent has deviated earlier than in the case that he has deviated at the same time. In fact, in 26 of the 31 cases where relative frequencies estimates $\hat{\alpha}$ and $\hat{\beta}$ are available, the inequality $\hat{\alpha} \geqq \hat{\beta}$ is satisfied. Therefore, it seems to be reasonable to take the following inequality as a point of departure:

$$0 \leqq \hat{\beta} \leqq \hat{\alpha} \leqq 1.$$

Accordingly, an auxiliary estimate $\hat{\beta} \leqq \hat{\alpha}/2$ is formed at the midpoint of the relevant interval delineated by this inequality if a relative frequency estimate is available for α but not for β. Analogously, an auxiliary estimate $\hat{\alpha} = (\hat{\beta} + 1)/2$ is formed if a relative frequency estimate is available for β but not for α.

It can be seen that the estimates in table 7 vary considerably from subject to subject. This is also true for the 14 subjects whose behavior completely conforms to the model. For these 14 subjects a second set of parameter estimates has been obtained in the same way on the basis of the data from rounds (1) to (25). These estimates will be used for a comparison of the learning model with a simple alternative hypothesis to be explained in the next section.

8. Comparison with a simple alternative hypothesis

In the following we want to look at the question whether our model provides a better explanation of the data than a simple alternative hypothesis based on the assumption that no learning takes place at all. We compare the learning model with the simplest alternative theory of this kind. In the alternative hypothesis each subject is assumed to have a probability distribution over his intended deviation period which does not vary over time. The intended deviation period of each round is assumed to be stochastically independent from those of other rounds.

Table 7

Parameter estimates based on rounds (1) to (20). Subjects are grouped according to the categories of table 6, in the same order.

Subject	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$
2	1.00	0.67	0.50
3	1.00	1.00	0.17
6	0.50	0.30	0.25
7	1.00	0.33	0.20
9	0.60	0.00	0.00
11	0.33	0.40	0.50
12	0.50	0.00	0.00
13	1.00[a]	1.00	0.00
18	0.22	0.00	0.50
22	0.67	0.40	0.50
28	0.25	0.00	0.00
30	0.25	0.00	0.50
31	0.57	0.33	0.57
34	1.00	0.50	0.50
15	0.00	0.00	0.00
17	0.00	0.00	0.00
21	0.00	0.20	0.50
25	1.00	0.00	0.00
16	0.43	0.00	0.67
19	0.00	0.00	0.67
20	0.75	0.50	0.14
24	0.00	0.50	0.09
27	0.00	0.13	0.00
29	0.00	0.14	1.00
32	0.38	0.00	0.00
33	0.14	0.00	0.33
35	1.00	1.00	0.75
4	0.33	0.20	0.00
5	0.50	0.00	0.00
8	1.00	0.75	0.33
14	0.17	0.08[a]	1.00
23	1.00	0.50[a]	0.18
26	0.50	0.50	0.33
36	1.00	0.50	0.50

[a] No relative frequency estimate available; auxiliary estimate according to $\hat{\alpha}=(\hat{\beta}+1)/2$ or $\hat{\beta}=\hat{\alpha}/2$, respectively.

The comparison will be restricted to those 14 subjects which never showed a reaction excluded by the learning model in the last 13 rounds and also had varying intended deviation periods in these rounds. For each of these subjects the probabilities for the actually observed intended deviation periods have been computed under the assumption of the model and under the alternative hypothesis. The parameters α, β and γ have been estimated on the basis of all 25 rounds. In the computation of the probabilities for the learning model the behavior of the other subjects in the same group had been taken as given. The probabilities therefore are conditional on the behavior of the other players. For each of the 14 subjects a conditional likelihood ratio has been formed as the quotient of the probability generated by the model divided by the probability generated by the alternative hypothesis. The conditional likelihood ratios are shown in table 8.

Table 8
Conditional likelihood ratios for the 14 subjects in the first group of table 6.

Subject no.	Ratio	Subject no.	Ratio
2	5.2	13	8830.0
3	131.0	18	540.0
6	23796.0	22	1628.0
7	2421.0	28	4.5
9	34.0	30	41086.0
11	660.0	31	2.0
12	1077.0	34	13.0

It can be seen that all 14 of the conditional likelihood ratios are greater than one; most of them are quite high.

The results of table 8 support the assumption that learning is an important factor in the choice of the intended deviation period. It would not make much sense to extend the method to the other subjects. Of course, a subject with a constant intended deviation period is better explained by the alternative hypothesis. The probabilities generated by the learning model for subjects who do not conform to it, are always 0, even in cases where there is only one isolated deviation. However, it is important to exclude the possibility that even for those subjects whose behavior does not violate the restrictions of the model the simpler alternative hypothesis yields a better explanation.

It would have been desirable to compute likelihood ratios for whole groups of interacting subjects rather than for individuals. Unfortunately, every group had at least one member not among those subjects to which the comparison was restricted.

Even if it must be admitted that our method of comparison is not entirely satisfactory the results confirm the impression that the learning model captures important aspects of the dynamics of end-effect behavior.

9. Monte Carlo simulations

The Monte Carlo simulations which already have been mentioned in the section on theoretical considerations serve the purpose to examine the predictive potential of the learning model. Therefore, in table 7 the parameters α, β and γ have been estimated individually for the subjects on the basis of observed behavior in the first 20 rounds. With these parameters. the last five rounds have been simulated starting from the observed values of intended deviation periods in round (20) as initial conditions. The pairing of the subjects followed the schedule of appendix A. The simulations only cover five of the six groups. The first group had one member who did not learn to cooperate before the last five rounds (see table 6). Therefore, for this subject no parameter estimates could be computed on the basis of the first 20 rounds.

The size of the end-effect is best described by the 'intended deviation time' which is defined as 11 minus the intended deviation period.

For each of the five groups fig. 2 shows the means of the intended deviation times over the six subjects for each of the last five rounds. These means are indicated both for the actual experiment and for the eight Monte-Carlo simulations.

It can be seen that the actual observed means are not too dissimilar from those generated by the Monte-Carlo simulations. It must be pointed out, however, that some shifts of more than one period occurred in the last five rounds. There was, for example, one subject in group III who shifted his intended deviation period from 11 to 6 from round (24) to round (25). Of course, the Monte-Carlo simulations cannot reproduce the effects of such jumps. This explains the special features in the drawing for groups III and V.

A meaningful statistical comparison of the simulations and the observations must be based on some features of the simulations which do not vary too much from realization to realization. It is plausible to conjecture that the rank order of the cumulative shifts of intended deviation periods over the last five rounds satisfies this criterion. The cumulative shift is the difference of the intended deviation periods in round (25) and in round (20). For each simulation run we obtain a rank order of these shifts over the six subjects of the group. In this way, the eight simulation runs for each group yield eight rank orders. Kendall's concordance coefficient W has been computed for the eight rankings in each of the five groups, separately. All five concordance coefficients are significant on the 0.01 level. This supports the conjecture that the rank order of cumulative shifts is a variable which can be predicted with some reliability if the model is correct. The predicted mean rank order has been computed by the sum of ranks following Kendall's proposal [Siegel (1957), Kendall (1948)].

For each of the five groups we have correlated the mean rank order of

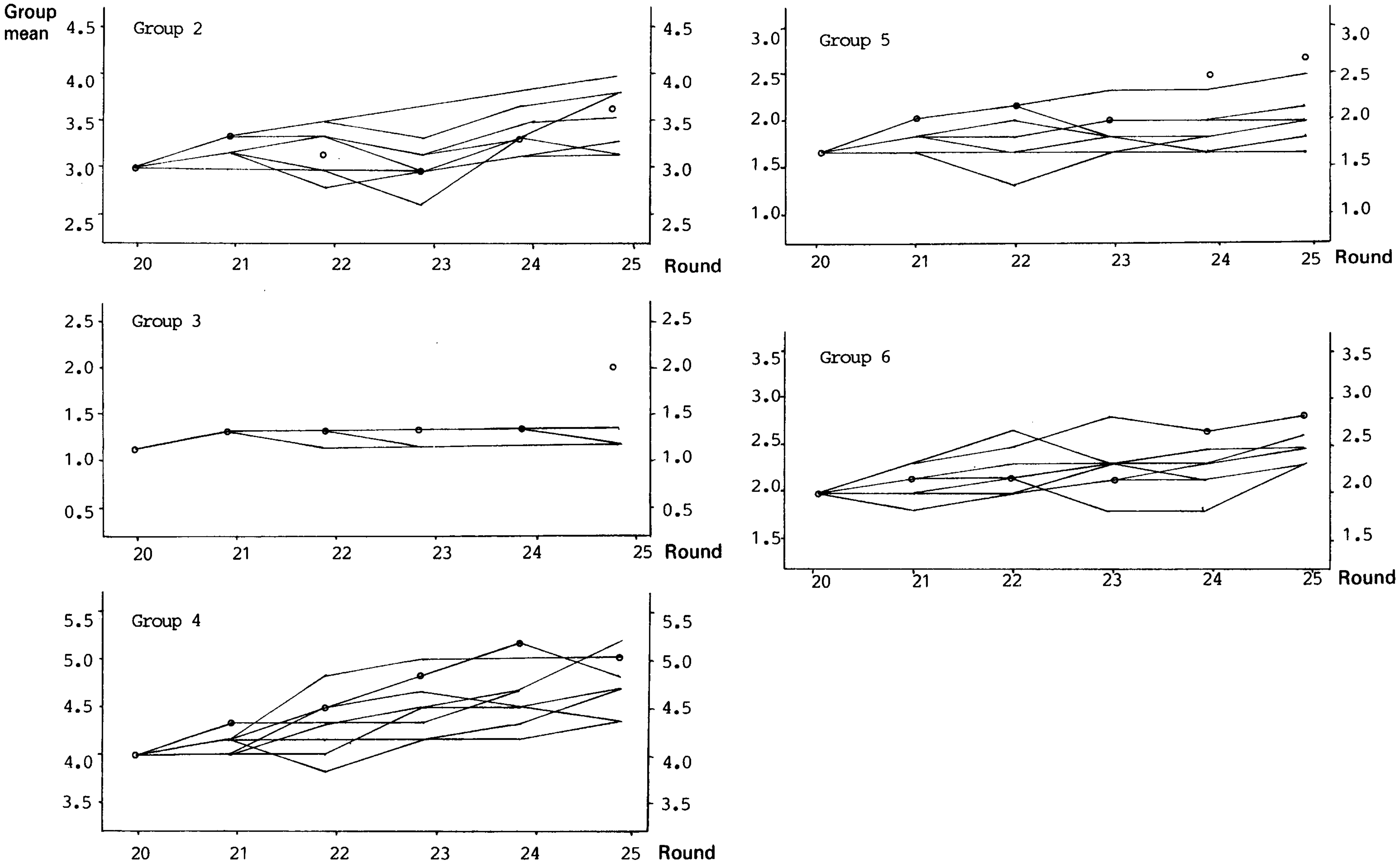

Fig. 2. Group means of intended deviation times in the experiment (o) and in eight Monte Carlo simulations.

cumulative shifts derived from the eight simulation runs with the rank order of cumulative shifts observed in the experiment. The Spearman Rank correlation coefficients are 0.880 ($p<0.05$), 0.548, 0.956 ($p<0.01$), 0.462 and 0.926 ($p<0.05$) for groups II, ..., VI, respectively.

The cumulative shift between round (20) and round (25) rather than the intended deviation period of round (25) has been chosen as the basis of the comparison between simulations and observations since the latter variable could reflect the initial conditions of round (20) more than the effects of the parameter values. On the other hand, the cumulative shift is a measure which can be expected to be more closely connected to the dynamics of the learning process.

If the learning model had no predictive value one would expect positive and negative rank correlation coefficients between predicted and observed rank orders of cumulative shifts with equal probability. The binomial test rejects this null hypothesis on the 0.05 level (one-sided). Moreover, three of the five-rank correlation coefficients are significant at the 0.05 level.

The result of the comparison of predicted and observed rank orders of cumulative shifts support the learning model as an idealized picture of end-effect behavior in repeated Prisoners' Dilemma supergames.

Appendix A

The six groups of interacting subjects were composed as shown in table A.1. Within each group of six interacting subjects the pairings were determined according to the scheme in table A.2. The same pattern was repeated in rounds (6) to (10), (11) to (15), (16) to (20) and (21) to (25).

For group II the numbers 1, 3, 5, 7, 9, 11 have to be replaced by 2, 4, 6, 8, 10, 12, in that order. The pairings within the other groups are obtained analogously.

Table A.1
Composition of groups I to VI.

Group	Subjects
I	1, 3, 5, 7, 9, 11
II	2, 4, 6, 8, 10, 12
III	13, 15, 17, 19, 21, 23
IV	14, 16, 18, 20, 22, 24
V	25, 27, 29, 31, 33, 35
VI	26, 28, 30, 32, 34, 36

Table A.2
Pairings in group I for rounds (1) to (5).

Round	Pair 1	Pair 2	Pair 3
(1)	1, 3	5, 7	9, 11
(2)	1, 5	3, 11	7, 9
(3)	1, 7	3, 9	5, 11
(4)	1, 9	3, 5	7, 11
(5)	1, 11	3, 7	5, 9

Appendix B

Table B.1

Subjects' intended deviation period and the observed behavior of the opponent in end-effect plays. Subjects are ordered according to interacting groups (see appendix A).[a]

	Round																								
Subject	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)
1. intended									0												7	6	6	7	6
observed									10												6	6	5	6	5
3. intended									9	8		7	7	7	7		7	6	6	7	6	6	6	6	5
observed									9	7		>	>	>	>		7	>	>	6	>	6	6	5	5
5. intended			9					9	9	9	9		9	8	8	7		8	7	6	6	6	6	5	5
observed			9					9	9	>	8		7	7	7	7		7	6	>	6	6	5	>	>
7. intended									7	7	8	8		8	8	7	6		6	6	6	6	5	5	5
observed									>	>	>	8		8	7	7	>		>	>	6	6	>	5	5
9. intended									10	10	9	8	8		7	7	7	7		7	6	6	6	6	6
observed									0	9	8	8	7		>	7	6	6		6	6	6	6	>	5
11. intended			9					9	8		8	8	7	8		7	7	7	7		6	6	5	5	5
observed			9					9	>	>	>	>	>	8		7	7	6	6		6	6	>	5	5
2. intended							8	8	10	10	9	8	9	8	8	8	7	8	7	7	7	7	7	7	7
observed							8	>	0	9	9	>	9	>	8	8	>	8	7	>	7	7	>	>	6
4. intended									10	10	9	8	8	8	8	8	8	8	8	8	7	7	8	8	
observed									9	9	9	>	>	8	8	8	8	>	7	7	7	7	>	6	
6. intended							8		9	10	10	9		8	8	8	8	8	7	7	7	7	7	6	6
observed							8		>	0	9	8		8	>	7	7	8	>	>	7	7	7	>	>
8. intended								9	8	9	9	9	9		8	7	8	8	8	7	7	8	8	8	
observed								8	>	>	>	>	9		8	>	>	8	8	>	7	>	7	6	
10. intended	Simulated player for rounds (8) to (25)																								
observed																									
12. intended									9	9	9	9			8	8	8	8	8	8	7	7	7	6	6
observed									8	>	>	8			8	>	8	8	8	7	>	7	7	>	>
13. intended													9	8	8	8	8	8	8	8	7	7	7	7	7
observed													9	>	>	>	>	>	>	7	>	>	>	>	>
15. intended				0			0	0	0			0	0	0	0	0	0	0	0	0	0	0	0	0	0
observed				0			9	10	0			10	0	0	10	8	8	0	0	0	7	7	0	0	6
17. intended			0	0	0			0	0	0			0	0	0	0	0	0	0	0	0	0	0	0	0
observed			9	0	0			9	0	10			10	0	0	0	8	7	0	0	0	7	7	0	0
19. intended												9	9	9	10	0	0	0	0	0	0	0	0	0	6
observed												>	9	>	0	0	0	8	7	0	0	0	7	7	>
21. intended					0	10		10		10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
observed					0	9		0		0	10	9	0	8	0	7	0	0	8	0	7	0	0	7	0
23. intended			9			9	9	9			10	10	10	10	9	7	8	7	7	7	7	7	7	7	8
observed			>			>	>	>			0	0	0	9	8	>	>	>	>	>	>	>	>	>	7

14. intended													8	9	9	9	θ	θ	θ	8	7	7	7	7	6
observed													>	7	7	7	8	6	6	7	7	7	6	5	>
16. intended				9	9		9	8	8	8		8	8	5	6	7	8	7	7	8	7	6	6	5	6
observed				>	7		7	8	>	6		6	6	>	>	>	7	6	7	6	7	6	6	>	6
18. intended			5	θ		θ		θ	θ		6		8	8	8	8	8	8	8	7	7	7	7	7	6
observed			>	9		6		7	8		6		7	5	8	6	>	7	7	6	6	7	5	5	5
20. intended				θ	6	6	7		7	6	6	5	θ	θ	8	6	6	6	6	6	6	6	6	6	6
observed				7	>	>	>		7	>	6	5	8	7	6	>	6	>	>	>	>	>	>	5	6
22. intended						θ	θ	8			θ	5	6	7	8	7	6	6	6	6	6	7	6	5	5
observed						7	7	8			6	5	>	>	8	7	6	>	>	>	>	6	6	>	>
24. intended			7	7		7	7	7	7		6	6	7	7	7	7	7	7	7	7	7	6	5	5	7
observed			5	7		>	>	>	7		>	>	>	>	>	7	>	>	6	>	6	6	>	>	6
25. intended	9		9	8		8		8	8				8	8	8	8	8	8	8	8	8	8	8	8	8
observed	9		7	>		>		>	8				>	>	>	>	>	>	>	>	>	>	>	>	8
27. intended	9		9	9	9	9	9	9	9	8		9	θ	θ	θ	9	9	9	9	9	9	9	9	9	9
observed	9		>	9	9	8	>	8	9	8		8	10	θ	9	8	9	9	>	9	8	8	>	8	9
29. intended				9	9	8		9	9	9	9		θ	θ	θ	θ	θ	θ	θ	θ	θ	θ	9	8	8
observed				9	9	>		9	9	8	7		θ	θ	θ	9	8	8	9	9	9	8	9	>	8
31. intended		9	7	9	9	9	8	9		8	7	8	9	10	9	9	9	10	10	9	9	9	9	9	9
observed		>	>	8	9	8	>	8		8	>	8	8	9	>	>	>	8	9	9	>	8	8	8	9
33. intended	10	10	10	10	9	9	9	8	8	8		8	10	θ	θ	θ	θ	9	9	9	9	8	10	9	8
observed	9	9	9	8	9	8	8	>	8	>		8	θ	8	θ	8	9	9	8	>	8	>	9	8	8
35. intended	9			8		8	10	9				8	θ	9	10	8	9	8	9	10	8	8	9	8	8
observed	>			>		>	9	9				>	θ	>	8	>	9	>	>	8	>	>	9	>	8
26. intended													θ	10	10	θ	10	10	10	10	9	9	10	9	8
observed													θ	10	θ	10	θ	θ	9	8	9	9	9	8	8
28. intended				9	9		9	9	9	9		9	10	9	9	9	9	9	9	9	9	9	9	9	9
observed				9	9		>	7	9	7		8	9	>	>	>	9	9	>	9	9	8	9	8	8
30. intended	9		9	9		9		9	9	9	9		10	θ	θ	θ	θ	10	10	9	9	9	8	8	9
observed	>		8	9		9		8	9	7	7		θ	9	9	θ	10	10	9	9	9	9	8	>	8
32. intended	10			10	9	9			9	7	7	7	θ	θ	θ	θ	θ	θ	θ	9	9	9	9	9	7
observed	9			9	9	9			8	>	>	6	θ	10	9	θ	9	10	9	9	9	9	>	8	>
34. intended								6		7	9	6	9	10	9	9	9	9	9	9	9	9	9	8	8
observed								>		>	8	>	>	10	>	>	>	9	>	9	8	9	9	>	>
36. intended			8	9			10	8	8		8	8	θ	10	θ	10	9	10	9	8	8	8	8	8	8
observed			>	>			9	>	>		>	>	10	θ	10	9	9	10	>	>	>	>	8	>	8

[a]The entries are as follows: 'intended': intended deviation period; 'observed': opponent's deviation period as observed by the subject; >: the subject deviated before the opponent; θ: in an 'intended'-row: the subject did not intend to deviate, and in an 'observed'-row: the opponent cooperated up to the end.

References

Axelrod, R., 1984, The evolution of cooperation (Basic Books, New York).

Bush, R.R. and F. Mosteller, 1955, Stochastic models for learning (Wiley, New York).

Kendall, M.G., 1948, Rank correlation methods (Griffin, London).

Kreps, D. and R. Wilson, 1982, Reputation and imperfect information, Journal of Economic Theory 27, 253–279.

Kreps, D., P. Milgrom, J. Roberts and R. Wilson, 1982, Rational cooperation in the finitely repeated Prisoner's Dilemma, Journal of Economic Theory 27, 245–252.

Lave, L.B., 1965, Factors affecting cooperation in the Prisoner's Dilemma, Behavioral Science 10, 26–38.

Milgrom, P. and J. Roberts, 1982, Predation, reputation and entry deterrence, Journal of Economic Theory 27, 280–312.

Morehous, L.G., 1967, One-play, two-play, five-play and ten-play runs of Prisoner's Dilemma, Journal of Conflict Resolution 11, 354–362.

Rapoport, A. and Ph.S. Dale, 1967, The 'end' and 'start' effects in interated Prisoner's Dilemma, Journal of Conflict Resolution 11, 354–462.

Restle, F. and J.G. Greeno, 1970, Introduction to mathematical psychology (Addison-Wesley, Reading, MA).

Selten, R., 1978, The chain store paradox, Theory and Decision 9, 127–159.

Siegel, S., 1956, Nonparametric statistics for the behavioral sciences (McGraw-Hill, New York).

Stoecker, R., 1980, Experimentelle Untersuchung des Entscheidungsverhaltens im Bertrand–Oligopol (Pfeffer, Bielefeld).

Stoecker, R., 1983, Das elernte Schlußverhalten — eine experimentelle Untersuchung, Zeitschrift für die gesamte Staatswissenschaft, 100–121.