Disentangling reinforcement learning and imitation in network experiments*

Oliver Kirchkamp† Rosemarie Nagel‡

September 15, 2003

Abstract

In this study we disentangle imitation, reinforcement, and reciprocity in repeated prisoners' dilemmas experiments. We compare a simple situation in which players interact only with their neighbours (local interaction) with one where players interact with all members of the population (group interaction). We observe choices under different information conditions and estimate parameters of a learning model. We find that imitation, while assumed to be a driving force in many models of spatial evolution, is often a negligible factor in the experiment. Behaviour is predominantly driven by reinforcement learning.

JEL-Classification: C72, C92, D74, D83, H41, R12

Keywords: Imitation, reinforcement, learning, local interaction, heterogeneity of environment, experiments, prisoners' dilemma.

*We would like to express our thanks to the German DFG (through SFB 303 and SFB 504) and the Spanish DGIC Técnica PB95-0983 and PB98-1076 for supporting the project. We are grateful to Antonio Cabrales, Jörg Oechsler, Avner Shaked, Nick Vriend and two anonymous referees for their helpful comments.

†University of Mannheim, SFB 504, L 13, 15, D-68131 Mannheim, email: oliver@kirchkamp.de.

‡Dep.of Economics, Universitat Pompeu Fabra, 132, Balmes, E-08008 Barcelona, email: rosemarie.nagel@upf.edu.

1 Introduction

In this paper we use experiments to disentangle imitation and reinforcement behaviour. This research is motivated by an approach frequently followed in evolutionary game theory: if agents can learn from their own experience as well as from other players' experience, then evolutionary game theory traditionally assumes that players weight information from both sources equally. As an example for this assumption Axelrod (1984, p. 158ff) discusses the evolution of a network of cooperators and defectors in a prisoners' dilemma and assumes that players use a copy-best rule, i.e. they choose the strategy with the highest payoff in the past, regardless whether this payoff was obtained by the learning players or by their neighbours. Nowak and May (1992), Eshel, Samuelson, and Shaked (1998) and several other articles[1] follow this approach. The copy-best rule is interesting since it helps us to explain how cooperation emerges through imitation in networks.

A rule like copy-best involves a specific mixture of imitation and learning from own experience. We want to find out whether this specific mixture is a good description of behaviour. A rule like copy-best is simple and reasonable if we interpret the evolutionary dynamics in a biological context where successful species displace less successful ones. Also in an economic context where successful firms invade the markets of less successful ones we may treat both sources of information equally. Neither biological resources nor markets have the cognitive capabilities to make a distinction between the success of the incumbent species or firm and the success of the invading species or firm.

Learning agents, however, may be able to distinguish between their own success and the success of their neighbours. Whether they should do so depends on the heterogeneity of the environment. Kirchkamp (1999) shows that if agents and neighbours are in the same environment a neighbours' experience is as good as an agent's own experience — there is no reason to value information differently. In a heterogeneous environment, however, the experience of a neighbour may be specific to a situation that is different from the agent's situation. In such an environment agents should learn relatively more from their own experience and relatively less from the experience of other players.

In this paper we investigate whether human players indeed weight own and neighbours' information equally in homogeneous environments and differently in heterogeneous environments. To control the degree of homogeneity, we compare two structures: In one structure agents are located on a circle and interact in overlapping neighbourhoods. This is what we call *local interaction* or a *spatial structure*. In such a structure players' environments are not entirely identical. Players may learn from their neighbours, however,

[1] See also Nowak and May (1993), Bonhoeffer, May, and Nowak (1993), Lindgreen and Nordahl (1994), Kirchkamp (2000).

their neighbours' success might be due to opponents that are not part of the interaction neighbourhood of the learning players. In the other structure agents are in a group and each agent is equally likely to interact with every other agent. This is what we call *group interaction* or *spaceless structure*. In this structure all agents face the same interaction partners.

From several other experiments we know that players learn from their own experience and that they also imitate. A classic study that describes how players learn from their own experience is Erev and Roth (1998). From this literature we know that a reinforcement model describes a learning process fairly well in many different situations. Pingle and Day (1996) find that participants of their experiments also imitate choices to economise decision cost. Offerman and Sonnemans (1998) observe that players imitate beliefs of other players if these are available. However, Offerman and Sonnemans (1998, p. 571) also suggests that *own experience* might be "more important" for the adaptation of beliefs than *imitation*.

What kind of framework should we use in order to study learning and imitation? We have chosen a very simple framework, a prisoners' dilemma. This is, however, not the only possible choice. Some recent studies of imitation behaviour use the context of an oligopoly. The oligopoly framework is particularly interesting in the context of learning and imitation since learning and imitation may affect the equilibrium process. Vega-Redondo (1997) presents a theoretical analysis of a Cournot oligopoly and finds that an imitation based evolutionary process converges to the Walras equilibrium which is far away from the Cournot-Nash equilibrium and which is also more competitive. Huck, Normann, and Oechssler (1999) and Offerman, Potters, and Sonnemans (2002) use experiments to find that players do imitate and do indeed tend to converge to the Walras equilibrium in oligopolies if information about other players is available. Selten and Ostmann (2001) develop the theoretical concept of an imitation equilibrium which is studied in Selten and Apesteguia (2002) with the help of an experiment based on an oligopoly with spatial competition. Selten and Apesteguia find that, indeed, features of the imitation equilibrium describe parts of actual behaviour better than the Cournot Nash concepts. In an experiment by Bosch-Domènech and Vriend (2003), however, play remains close to the Cournot Nash equilibrium and does not converge to the Walras equilibrium.

What we learn from these experiments is that imitation may play a role in oligopolies. These experiments also help to distinguish among different equilibrium concepts in oligopoly models. However, these experiments also show that it is not easy to disentangle imitation of others from learning from own experience with the help of an oligopoly experiment. The reason is the large strategy space. Players can and will choose many different strategies among a large number of possible quantities. Hence, often players will choose new quantities that have not been tried before. How can we interpret the choice of new quantities as imitation or learning from own experience? Perhaps the chosen strategy was close to one or more successful strategies used by other players or used by the learning player, but how close must a choice be to be qualified as imitation? With so many candidate strategies one would need additional and hard to justify assumptions to relate players' choices to past strategies.[2]

To avoid this problem we use a simpler setting. With the prisoners' dilemma we study a game with only two strategies. This game is conceptually close to an oligopoly game, still, with only two strategies it is technically easier to interpret choices as learning. A prisoners' dilemma is not only interesting because it describes the well known dilemma situation. What is useful here are two other properties: firstly, learning and myopic optimisation may call for very different actions in this game, and, secondly, theoretical analysis shows that the interaction structure may crucially determine the behaviour of a population. If players copy successful strategies from their neighbours, cooperation may be a stable outcome in prisoners' dilemma games in a locally structured population, but can not be stable in a population without such a structure (see footnote 1).

Heterogeneity of the environment should theoretically influence imitation behaviour (Kirchkamp 1999). We will, hence, compare two setups, a homogeneous environment where all players face the same interaction neighbourhood and a heterogeneous environment with overlapping neighbourhoods.

Experiments where players are linked through a network and, thus, are in a heterogeneous situation have been done with coordination games, market games and prisoners' dilemma games. Kosfeld (2003) provides an exhaustive summary of networks experiments. Close to our study are those of Keser, Ehrhart, and Berninghaus (1998), Cassar (2002), and Selten and Apesteguia (2002).

Keser, Ehrhart, and Berninghaus (1998) study how the structure of the network af-

[2] A solution for a related problem is used by Huck, Normann, and Oechssler (2000). In each round the authors determine a best-reply quantity and an imitation quantity. Then they count the number of choices that are within an interval around these two quantities. A similar procedure could be used to distinguish between imitating others' and imitating own experience. The problem of Huck, Normann, and Oechssler is, however, easier to solve. In each round there is always a best response and always an imitation quantity. Since we want to distinguish between imitating others and learning from own experience we have to deal with the problem that in each round the best strategy is either used by the learning player or used by the other players. In each round there is one explanatory observation missing. To impute this missing observation one could assume that players create two quadratic models, one for own experience and the other for others' experience and then continue with Huck, Normann, and Oechssler's procedure. Reducing the strategy space, which is what we are going to do in the next section, might be simpler.

fects selection of Pareto and risk dominant equilibria in coordination games. To answer our question, however, coordination games are not ideally suited. In coordination games we can not distinguish between a player who chooses a strategy as a result of imitating successful neighbours, and a player who chooses a strategy as a result of myopic optimisation. Both motives call for the same action. Since we want to learn more about imitation we have to study a different game.

Cassar (2002) studies coordination games and prisoners' dilemmas. In her experiments with prisoners' dilemmas she finds how perturbations in the structure of a spatial network affects choices. She compares three structures, a local one, a slightly perturbed one (what she calls a small world) and a random network. She finds an interesting non-monotonicity: The slightly perturbed network yields the smallest amount of cooperation.

Selten and Apesteguia (2002) study an oligopoly with a spatially differentiated product. They are, however, not interested in the relation between learning from own experience versus imitation. They do not measure this relationship and they do not vary the heterogeneity of their environment. What they find is that imitation seems to be a relevant factor. What we want to find in this article is how relevant this factor is, as compared to learning from own experience.

While we use space here to model similarity of situations and to allow studying the evolution of strategies, space is also crucial in many economic situations. Restaurants or shops along a street do not compete with the same strength with all other restaurants or shops on that street. Strategic interaction and imitation is more important among producers of similar products. Should we, therefore, find more tacit collusion in industries where product space or geographic space is relevant for interaction?

In our experiment groups of players repeatedly play prisoners' dilemmas either within a locally structured neighbourhood (a circle with overlapping neighbourhoods) or within an unstructured (spaceless) group. Players receive information about their neighbours and their own payoffs. We find that players learn from their own experience. Success of their neighbours, however, does not seem to play a large role. This holds for both structures: the spatial as well as the spaceless one. As a consequence we do not find the higher levels of cooperation in the spatial structure predicted by the theoretical literature under the assumption of learning from neighbours (see footnote 1). Various modifications of our setup do not change this result.

In section 2, we briefly summarise a theoretical argument that is based on imitation and that suggests more cooperation in a spatial world than in a non-spatial world. We will describe the first experimental setup without information about the payoff matrix in section 3. In section 4 we come to our experimental results. We will study stage game behaviour and learning behaviour. In section 4.2 we study a structure where a permanent cluster of computerised cooperators facilitates the imitation of successful cooperation. Section 4.3 studies the effect of introducing information not only about realised payoffs but also about the payoff matrix. Section 5 concludes.

2 A simple model based on copy-best

In this section we will sketch a simple and common evolutionary learning process based on copy-best[3] which suggests more cooperation in a spatial environment and less in a non-spatial one.

Let us assume that players play a prisoners' dilemma in a neighbourhood of five as described in table 1. Players can only use the same strategy against all four neighbours/group members. Playing C contributes 5 points to the payoff of each neighbour, playing D contributes nothing to the others but adds always 4 points to the own payoff.

Obviously, in a non-spatial (group) setting with myopic imitation, or replicator dynamics, non-cooperation is always more successful than cooperation. Hence, in a non-spatial setting, cooperation always dies out. In the upper part of figure 1 we give an example. We simulate a group of five players who always play the strategy with the highest average payoff in their neighbourhood (*copy best average*). With a small probability (1% in this example) players 'mutate' and choose the other strategy.

If average payoffs for C and D in period t are called u_t^C and u_t^D, respectively, if all strategies are used in period t, and $u_t^C \neq u_t^D$ then we can express the probability to play c tomorrow as follows

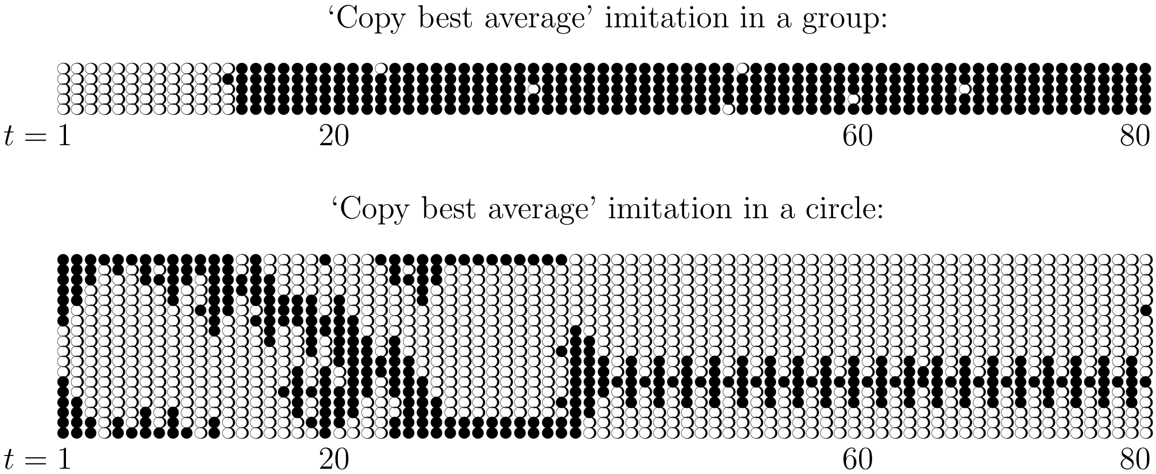
$$P(c_{t+1}) = \frac{1}{2} + \left(\frac{1}{2} - \epsilon\right) \operatorname{sign}(u_t^C - u_t^D) \tag{1}$$

where ϵ is the mutation rate. If $u_t^C = u_t^D$ or one strategy was not used in period t then players repeat their choice.

In the example we start with 5 cooperating players who choose cooperation until the first mutant arrives. This happens in our example in period 13 where one player mutates and plays D. Being very successful, this player is imitated by all neighbours and from period 14 on everybody plays D. Further mutants that appear in later periods do not lead the group back to cooperation.[4]

[3] Similar processes are used e.g. in Nowak and May (1992, 1993), Bonnhoeffer, Nowak, and May (1993), Lindgren and Nordahl (1994), Eshel, Samuelson, and Shaked (1998), Kirchkamp (2000).

[4] The only way to move a population where everybody plays D back to cooperation is a simultaneous mutation of all five players. With independent mutations this is not very likely. And even if it happens,



❍ $=C$, ● $=D$. Time is shown on the horizontal axis, different players are shown on the vertical axis. The first mutant D makes cooperation disappear completely in groups. Cooperation in circles, however, persists despite mutant Ds.
(The imitation rule is ‘copy best average payoff’, the mutation rate is 1%, the imitation and interaction radius is 2, as in the experiment. Simulations starts with 5 cooperators in the first period.)

FIGURE 1: Simulated learning.

Own payoff:					
own	number of	[neighbours / group members]			choosing C
action	0	1	2	3	4
C	0	5	10	15	20
D	4	9	14	19	24

TABLE 1: Payoff Matrix

Player						…	1	2	…	…		
Neighbourhood of Player 2						-	-	⇓	-	-		
Action:	…	D	C	C	C	C	C	D	D	D	D	D	…
# of other Cs in the neighbourhood	…	2	2	3	4	3	2	2	1	0	0	0	
Own payoff	…	14	10	15	20	15	10	14	9	4	4	4	
Average payoff of C		12.5	15	15	14	15	15	12.5	10	—	—	—	
payoff of D in the neighbourhood		9	11.5	14	—	14	11.5	9	7.75	7	5.25	4	

TABLE 2: Example of a neighbourhood of Cs and Ds

In a spatial setting and with similar imitation dynamics (see footnote 1) however, cooperation is protected through space and may, hence, survive.[5] Let us assume that player 2 from table 2 knows his own payoff from playing D, which is 14, but also the payoff from his two D-playing neighbours, 9 and 4. The average payoff of playing D is, hence, 9. The two C-playing neighbours of this player have a payoff of 15 and 10, on average, hence, 12.5. If player 2 copies the strategy with the highest average payoff then player 2 will choose C in the next period — thus, cooperation will grow.[6]

In our example (see the bottom part of figure 1) cooperation grows from the initial configuration of only five Cs and is not much affected by mutants.

In describing the above dynamics we used the rule 'copy best average payoff' (see the literature given in footnote 3). We should note that this learning rule does not distinguish between a players' own experience and his neighbours' experience. This is expressed in the following hypothesis:

Hypothesis SYM_{NB} A player learns as much from his neighbours' experience as from his own.

We, furthermore, assumed that players would learn from payoffs of C and D in the same way, i.e. an increase in the observed payoff of C would increase a player's inclination to play C in the same way as a similar decrease in the observed payoff of D.

Hypothesis SYM_{STR} Players learn from C and D in the same way.

It is, however, not obvious, that hypothesis SYM_{NB} and SYM_{STR} *should* hold. In a spatial structure players' environments are not identical. Making no distinction between own experience and one's neighbours' experience may, hence, be suboptimal.[7] We summarise this in the following hypothesis:

Hypothesis ASYM_{NB} Players learn relatively more from their own experience and less from their neighbours' experience the more local their interaction structure is.

If hypothesis SYM_{NB} and SYM_{STR} hold, then we should, following the argument sketched in section 2 and discussed in detail in the literature (see footnote 3), expect the following:

cooperation will not last for long since the first single mutant leads the population back to D. As a result the population will spend most of the time in a state where most of them play D.

[5]With myopic optimisation (Ellison 1993) players would obviously never cooperate.

[6]Once the cluster of Ds becomes small the payoff of the remaining Ds grows and the process stops or enters a cycle. With standard imitation processes stable equilibria are often reached when clusters of successful Cs are separated by small clusters of equally successful Ds.

[7]See Kirchkamp (1999).

Hypothesis COOP We find more cooperation in populations with a spatial structure than in populations without such a structure.

If, however, learning is not symmetric and instead ASYM_{NB} holds, the forces of imitation are weaker. Imitation is, as we have seen in the example above, a major driving force behind the survival of cooperation in a spatially structured population. A player who looks only at his own payoff in a prisoners' dilemma will quickly learn that defection gives a higher payoff — regardless whether this player is learning in a spatial or a spaceless structure. We might then find the following:

Hypothesis NO-COOP Levels of cooperation are *not* higher in a spatial structure.

3 The experimental setup

In this paper we describe results from five different treatments which are based on 35 sessions run in Barcelona and Mannheim, involving 339 participants.[8]. A list of these sessions is given in appendix A

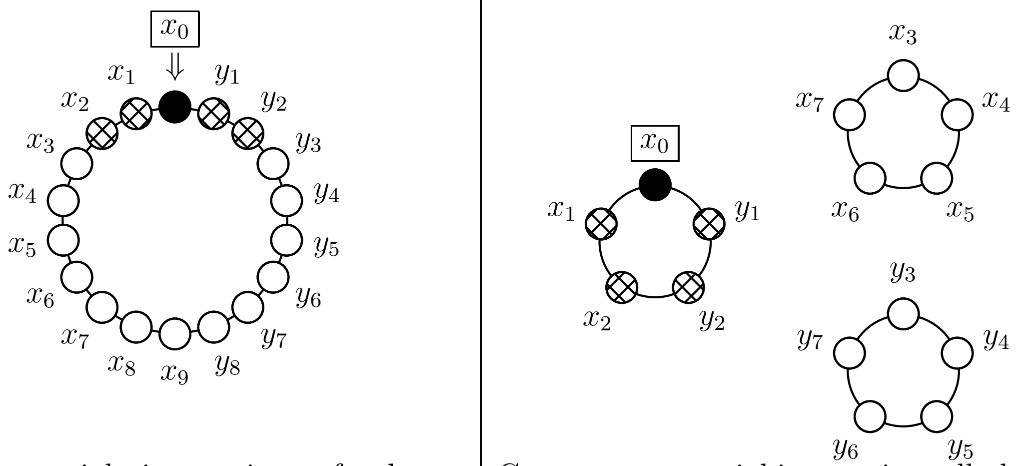
In the current section we will give a description of the first two treatments. One of them will be called a 'circle' treatment, the other 'group' treatment. We ran 4 sessions on a circle and 10 in groups. The remaining three treatments are modifications that are described in sections 4.2 and 4.3 below.

- In each session of the circle treatment we study a spatial structure of 18 players.[9] Participants are randomly seated in front of computer terminals that are networked to create a neighbourhood structure (see left part of figure 2). Each player interacts in each round with two neighbours to the left and two neighbours to the right. Player x_0 in the figure is in interaction with x_1, x_2, and y_1, y_2. Player x_2 is in interaction with x_3, x_4, and x_1, x_0. Players are aware of this structure and observe average payoffs of the strategies used by their four interaction neighbours. We ran four sessions of this treatment.

- In the group treatment we study groups consisting of five players each. Each member of a group interacts in every round with all members of the group (see right part of figure 2). Players are again aware of this structure and can average observe payoffs of the strategies used in their group. We ran ten sessions of this treatment.

[8]Students of the UPF in Barcelona and Universität Mannheim respectively.

[9]In one of the treatments only 14 players showed up for the experiment. We pool the data from this experiment with the others.



Circle: spatial interaction of players through overlapping neighbourhoods

Groups: non-spatial interaction, all players are either in the same neighbourhood, or do not interact at all.

FIGURE 2: Neighbourhoods

Thus, both in the group and in the circle treatment the number of interaction partners is four. In the group sessions we invited 15 players that were randomly divided into groups of five. We conducted three sessions, thus involving nine independent groups.

During any session players always interact with the same neighbours. Sessions last for 80 periods. In each period participants play a prisoners' dilemma against all members of their neighbourhood/group as described in table 1.

A critical issue is the information that we provide in each period to participants. To be faithful to the copy-best setup in section 2 we could give detailed information about payoffs and choices off all neighbours to participants. We will study such a setup in section 4.3.2. However, this setup has a drawback. The structure of the prisoners' dilemma game becomes obvious and strategic considerations replace learning.

We therefore start with a setup where we give only average information about payoffs. Also the payoff matrix (table 1) is not known to participants. To allow players to learn from their own experience we have to give them information about their own payoff u^{own}. In addition we could give them information about average payoffs of the two strategies in their neighbourhood u^C and u^D. This information would include their own success. Participants could apply equation (1) and use the copy best rule if they wanted to. We fear, however, that in this setup participants are confused by obtaining information about their own payoff u_{own} next to information about average payoff that again includes their own payoff. We believe that participants will better understand the experiment if we separate payoff information into own payoff u^{own} and payoff of the neighbours $u^{c,\text{other}}$

History			
Round	Your action and gains are	in your neighbourhood the average payoff was with… C	D
…	……	…	…
…	D 14	12.5	9
…	……	…	…

In the experiment strategies were called A and B. In some sessions A was the cooperative strategy, in others B. Payoffs of Cs are shown in a box, payoffs of Ds are shown in gray. In the experiment we use the colours red and blue.

Table 3: Representation of payoffs in the 'less-information' treatment

and $u^{d,\text{other}}$. This is a small change but we should note that now participants can only approximatively implement copy-best. In section 4.1.4 we will explain in more detail how they could do this.

Own payoffs and actions and the average payoff for their neighbours' actions C and D are shown as in table 3. This takes place in circles and groups in the same way. Players could change their strategy from period to period, but they always had to choose a single strategy for all their neighbours.

4 Results

4.1 Results from the baseline treatments

We will first study stage game behaviour and find that in contrast to the simple imitation dynamics discussed in section 2 and summarised in hypothesis COOP there is not more cooperation in space (in circles) than without space (in groups). Then we relate this observation to learning. We will see that, contradicting hypothesis SYM_{NB}, imitation is a weak force. Players' behaviour is much more driven by their own experience (learning through reinforcement) than by their neighbours' experience which is in line with hypothesis ASYM_{NB}. The players' actions over time are shown in appendix B.1 and B.2.

4.1.1 Stage game behaviour

In figure 3 we show the relative frequency of cooperation in circles and groups. Levels of cooperation decrease over time and are about the same in groups and in circles. In circles the average relative frequency of cooperation over all 80 periods is 0.176, in groups

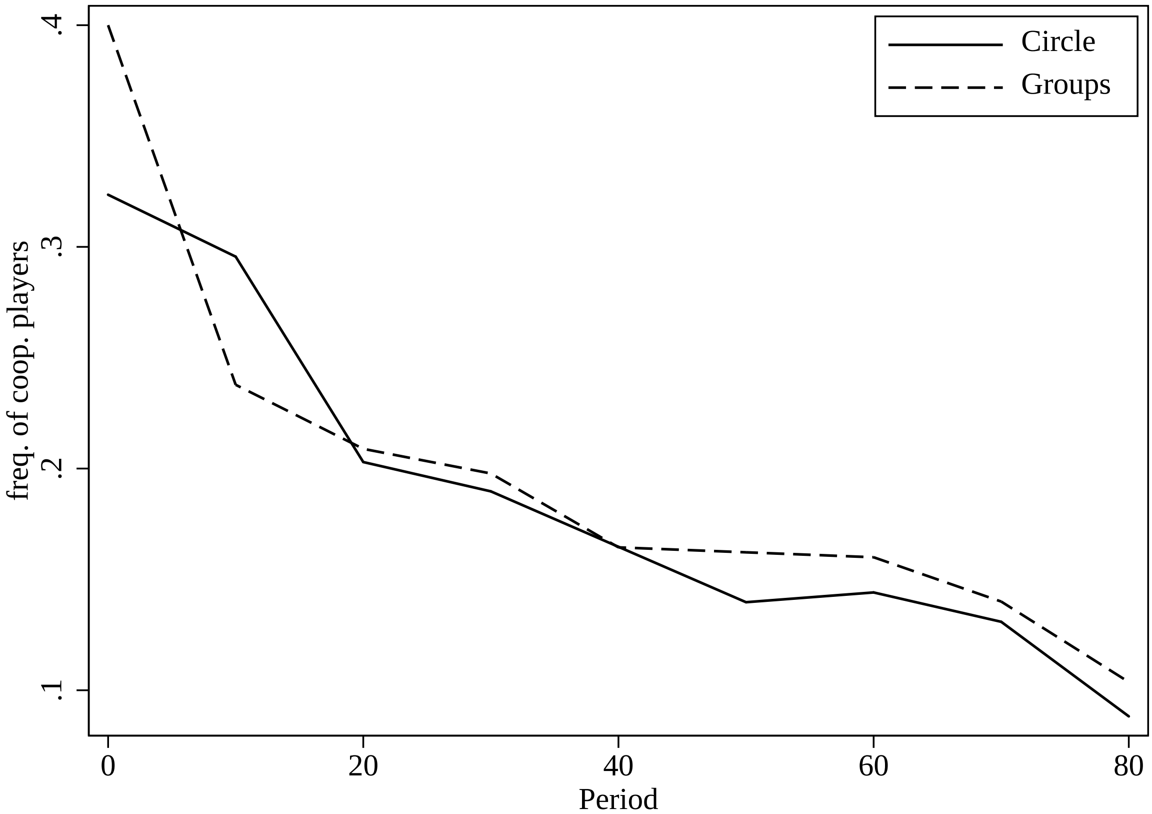


FIGURE 3: Frequency of cooperative players in circles and groups over time

the level is 0.187. Neither a t-test[10] ($t = -0.47$, $P_{>|t|} = 0.646$, allowing for correlations within sessions) nor a two-sample Wilcoxon rank-sum test ($z = 0.820$, $P_{>|z|} = 0.4120$) find a significant difference between groups and circles. They are similar to what is found in other non-spatial experiments.[11] Hence, we do not find support for hypothesis COOP. Hypothesis NO-COOP is, however, consistent with our observation.

This observation is also supported by Cassar (2002) who finds similar decline in cooperation in circles and also in perturbed circles where some connections between neighbours are broken and replaced by connections to distant players.

4.1.2 A simple learning model

In this section we start investigating hypothesis SYM_{NB} and ASYM_{NB} and study players' learning and imitation behaviour. Since we can not directly observe the learning process but only its outcomes, i.e. players' choices, we have to use a statistical model of the learning process. The logit model is perhaps the most common model that allows us to describe discrete choices between two alternatives, here C and D. Own payoffs from C and D will be called $u^{c,\text{own}}$ and $u^{d,\text{own}}$, respectively. If a player does not cooperate in a given period t the value of $u_t^{c,\text{own}}$ can not directly be determined. In this case we recursively use $u_t^{c,\text{own}} := u_{t-1}^{c,\text{own}}$ until we reach a period where the player actually cooperated. In a similar way we define $u^{c,\text{other}}$ and $u^{d,\text{other}}$. We define recursively $u_t^{s,i} := u_{t-1}^{s,i}$ for $s \in \{C, D\}$ and $i \in \{\text{own}, \text{other}\}$.

In line with equation (1) we use differences in payoffs of C and D as explanatory variables of our model. $\Delta_t^{\text{own}} := u_t^{c,\text{own}} - u_t^{d,\text{own}}$ is the difference between payoff from cooperation and payoff from non cooperation as experienced by the player in period t. $\Delta_t^{\text{other}} := u_t^{c,\text{other}} - u_t^{d,\text{other}}$ is the difference between payoff from cooperation and payoff from non cooperation as experienced by player's neighbours in period t. To allow for some inertia we include the current choice c_t which we code as 1 if the player cooperates today, and 0 otherwise. In section 4.1.6 we will study a richer model with more parameters which

[10] When calculating levels of standard deviations and levels of significance we have to take into account that observations within any session may be correlated. We can, however, assume that covariances of observations from different sessions are zero. Covariances of observations from the same session are replaced by the appropriate product of the residuals (Rogers 1993). We will use this approach throughout the paper to calculate standard errors.

[11] Bonacich et. al. (1976) studied cooperation within groups of 3, 6, and 9 players in a game where cooperation is less attractive than in our game. They found levels of about 30% of cooperation in groups, which is close to the initial levels results in our experiment.

Fox and Guyer (1977) used a non-linear payoff scheme where sometimes cooperation was more attractive than in our game. They found more cooperation (around 50%) in a game with groups of 3 and 12 players.

allows players to treat C and D in an asymmetric way. Let us first estimate

$$P(c_{t+1}) = \mathcal{L}(\beta_0 + \beta_c c_t + \sum_{i \in \{\text{own},\text{other}\}} \beta^i \Delta_t^i) \tag{2}$$

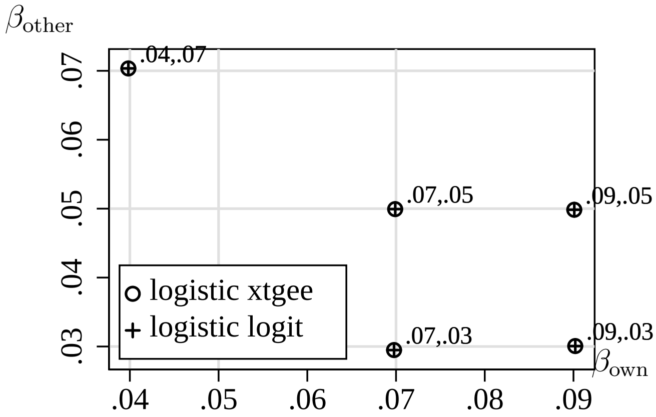
where $\mathcal{L}(x) = e^x/(1+e^x)$, c_{t+1} is 1 if a player cooperates tomorrow, and 0 otherwise. The factor β_{own} captures, hence, reinforcement, β_{other} measures the amount of imitation, β_c measures inertia, and β_0 a general inclination to play C. The two factors, β_{own} and β_{other}, hence, allow us to measure different degrees of imitative behaviour. The relationship to the copy-best rule will be investigated in the next section.

4.1.3 How to estimate learning behaviour

When estimating the above model we have to take into account correlations within variables. The dependent variable c_{t+1} influences payoffs in the next period and, hence, the explanatory variables $\Delta_{t+1}^{\text{own}}$ and $\Delta_{t+1}^{\text{other}}$. The AR(1) process can be estimated with the help of a GEE population-averaged model.[12] Before we come to the results of our experiments we want to do two things. We want to convince the reader that the GEE estimator is unbiased. Furthermore we show the results of the estimator when the population does not follow the model described in equation (2) but instead a copy-best imitation process as assumed by Axelrod (1984), Eshel, Samuelson, and Shaked (1998), Nowak and May (1992), and others.

In order to find out whether the GEE estimator is indeed unbiased in our context we did a Monte Carlo study which was based on our laboratory setting. A circle of 18 players starts from a random configuration where each player chose with probability 0.36 to cooperate. This corresponds to the initial behaviour in our experiments. In the following 80 periods each player chooses C with a probability given by equation (2). Based on six of these simulations we run a GEE estimate as described above and also a simple logit estimate. We repeat this procedure 100 times and take averages. These averages are shown in figure 4 for five different parameter vectors. We see that for both methods, GEE and logit, the estimates coincide with the true coefficients — the estimator is unbiased. The number of periods (80) seems to be sufficiently large to counter the possible bias of the AR(1) process. In the following we will present results from GEE and logit estimates in our graphs. For simplicity we will present only the GEE estimates on our tables.

[12]See Liang and Zeger (1986). We use as a link function the logistic function and specify c_{t+1} to be binomially distributed.



The behaviour of a simulated populations is based on one of five different parameter vectors: $(\beta_0, \beta_c, \beta^{\text{own}}, \beta^{\text{other}}) \in \{(-1, 0, .09, .03), (-1, 0, .07, .03), (-1, 0, .09, .05), (-1, 0, .07, .05), (-.6, 2.6, .04, .07)\}$. The figure shows the true value of the parameter and the estimated value of the parameter. The average GEE estimates are shown as circles, the average logit estimates are shown as pluses. Next to the averages we show the true coefficients $\beta_{\text{own}}, \beta_{\text{other}}$ of the simulation.

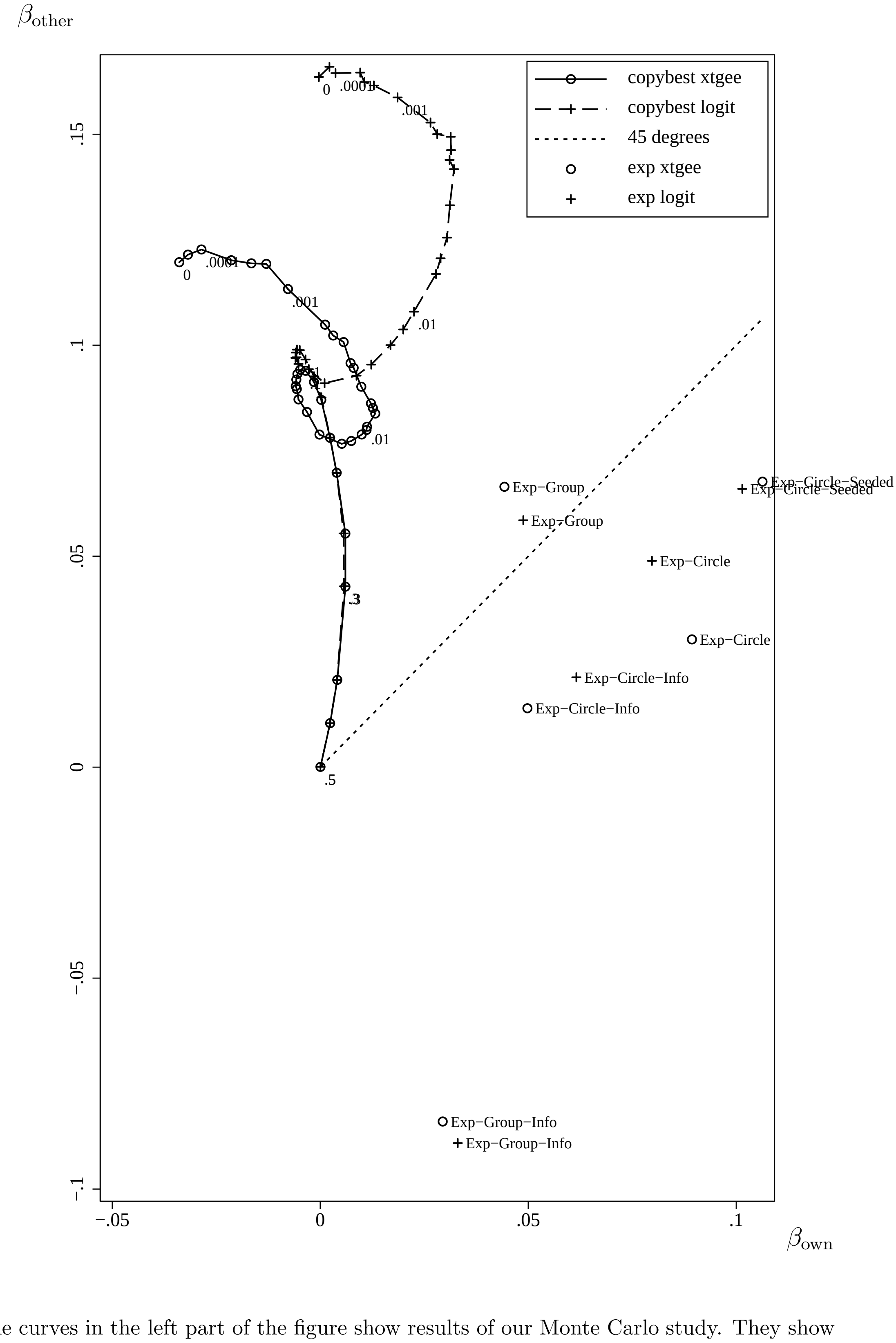
FIGURE 4: Results of Monte Carlo study — players using a logistic function

4.1.4 Where is copy-best in our learning model

As said above, we are interested in a comparison with the more theoretical literature (Axelrod (1984), Eshel, Samuelson, and Shaked (1998), Nowak and May (1992), and others). This literature assumes that players use a copy-best imitation mechanism when they update their strategy. In each period they determine the strategy with the highest payoff in their neighbourhood and follow this strategy in the next period. If payoffs of C and D are the same they stick to their current strategy. What would a GEE or a logit modes estimate if confronted with such a behaviour? To answer this question we use again a Monte Carlo study. We simulate the same situation as in the experiment, a circle of 18 players who play for 80 periods. Players follow the learning rule that is used by Eshel, Samuelson, and Shaked (1998).

In order to narrow down the properties of this process Eshel, Samuelson, and Shaked (1998), Kirchkamp (2000) Nowak, Sigmund, and El-Sedy (1993), and others use mutations, i.e. they assume that with a small probability p players make a mistake and choose the opposite strategy. For each mutation rate we simulate 1000 groups of six circles and present average estimates of the coefficients of equation (2) in figure 5.

The two curves in the left part show the result of our Monte Carlo study. Each point in each curve corresponds to one mutation rate. Mutation rates are shown next to the curve.



The curves in the left part of the figure show results of our Monte Carlo study. They show how a GEE or a logit estimate perceives copy-best behaviour (mutation rates are shown next to the curve).

For comparison our experimental results are shown in the same figure (Exp-Circle... and Exp-Group...). Most of the experiments can be found below the 45° line.

FIGURE 5: Copy best versus experimental results

coeff. from eq. (2)	Learning own and others' payoff in circles					
	β	σ	t	$P_{>\|t\|}$	95% conf. interval	
β_c	$-.1527977$	.1097514	-1.39	0.164	$-.3679064$	.062311
β^{own}	.0893514	.0089972	9.93	0.000	.0717171	.1069857
β^{other}	.0302698	.0121571	2.49	0.013	.0064423	.0540972
β_0	-1.079484	.067418	-16.01	0.000	-1.211621	$-.9473476$
coeff. from eq. (2)	Learning from own and others' payoff in groups					
	β	σ	t	$P_{>\|t\|}$	95% conf. interval	
β_c	2.585381	.1193796	21.66	0.000	2.351402	2.819361
β^{own}	.0442708	.0097167	4.56	0.000	.0252265	.0633152
β^{other}	.0664513	.0192851	3.45	0.001	.0286532	.1042494
β_0	-1.622957	.1131753	-14.34	0.000	-1.844777	-1.401138

TABLE 4: GEE population-averaged estimation of equation (2)

A mutation rate of $p = 0.5$ corresponds to random behaviour — half of the time players choose the right strategy, and half of the time they choose the wrong strategy. Thus, both coefficients β_{own} and β_{other} are estimated to be zero. For smaller mutation rates behaviour is more structured. We see that for all mutation rates the GEE and the logit estimate are above the 45° line, i.e. $\beta_{\text{other}} > \beta_{\text{own}}$. The intuition is that when calculating the average payoff of a strategy all players are treated equally. The payoff experience of the learning player has a smaller impact than the experiences of the four neighbours.

Finding $\beta_{\text{other}} > \beta_{\text{own}}$ is, hence, what we should expect in a world where players use a copy-best rule, i.e. where hypotheses SYM_{NB} and SYM_{STR} hold. Finding $\beta_{\text{other}} < \beta_{\text{own}}$ would be evidence for ASYM_{NB}.

We did similar simulations for groups. As we have already seen in section 2 cooperation dies out quickly in the group setting. Therefore estimated coefficients are independently of the mutation rate very close to zero. In figure 5 estimation results could not be visibly distinguished from the origin.

4.1.5 Learning in the experiment

Anticipating briefly our result figure 5 includes also an overview of our experiments. Average coefficients for β^{own} and β^{other} are shown next to a label of the experiment. Regardless whether we look at the GEE or the logit estimate, most of our experimental results are below the 45-degrees-line, i.e. $\beta^{\text{own}} > \beta^{\text{other}}$. This is in stark contrast to the copy best behaviour described above.

The experiments that we described so far are the ones labelled "Exp-Circle" and "Exp-Group" in figure 5. Detailed estimation results are shown in table 4. The value of β_{own} is significantly larger than β_{other} ($\chi^2(1) = 10.49$, $P_{>\chi^2} = 0.0012$). This means that

coeff. from eq. (3)	Learning from C and D in circles					
	β	σ	t	$P_{>\|t\|}$	95% conf. interval	
β_c	−.3078639	.16064	−1.92	0.055	−.6227125	.0069846
$\beta^{c,\text{own}}$	.0909994	.0108732	8.37	0.000	.0696884	.1123104
$\beta^{d,\text{own}}$	−.0849679	.015454	−5.50	0.000	−.1152572	−.0546786
$\beta^{c,\text{other}}$	.0462106	.0139529	3.31	0.001	.0188635	.0735577
$\beta^{d,\text{other}}$	−.0134582	.0192408	−0.70	0.484	−.0511695	.024253
β_0	−1.257075	.0987834	−12.73	0.000	−1.450687	−1.063463
coeff. from eq. (3)	Learning from C and D in groups					
	β	σ	t	$P_{>\|t\|}$	95% conf. interval	
β_c	2.657546	.1513503	17.56	0.000	2.360905	2.954187
$\beta^{c,\text{own}}$	.0725666	.0131783	5.51	0.000	.0467377	.0983955
$\beta^{d,\text{own}}$	.0152466	.0168734	0.90	0.366	−.0178246	.0483179
$\beta^{c,\text{other}}$	.1177859	.022863	5.15	0.000	.0729753	.1625966
$\beta^{d,\text{other}}$	−.1191031	.024038	−4.95	0.000	−.1662168	−.0719894
β_0	−1.822549	.1188607	−15.33	0.000	−2.055512	−1.589586

TABLE 5: GEE population-averaged estimation of equation (3)

there is more reinforcement than imitation. In groups $\beta_{\text{own}} < \beta_{\text{other}}$ but the difference is not significant ($\chi^2(1) = 0.82$, $P_{>\chi^2} = 0.3663$). Players living in a spatial structure are less sensitive to their neighbour's payoffs than one in a non-spatial one. This is consistent with hypothesis ASYM$_{\text{NB}}$. In a spatial structure a neighbour's success with a strategy may be due to this neighbour's neighbourhood and might not apply to the learning players. We can, hence, reject hypothesis SYM$_{\text{NB}}$ in the spatial structure (in circles) but not in groups.

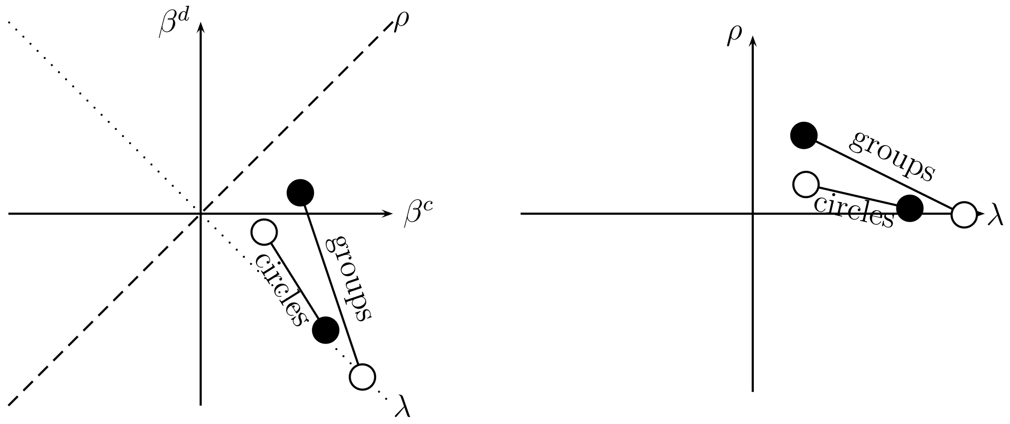
Remember that in section 2 we explained that survival of cooperation in a spatial structure crucially depends on imitation of neighbours. Finding only a small amount of cooperation in circles in section 4.1.1 should, hence, not come as a surprise, given that imitation plays only a limited role ($\beta_{\text{other}} < \beta_{\text{own}}$).

4.1.6 Learning and reciprocity

When we estimated equation 2 we made the simplifying assumption that players are equally sensitive to payoffs from the two strategies C and D. Equation 3 describes an approach which allows for different sensitivities.

$$P(c_{t+1}) = \mathcal{L}(\beta_0 + \beta_c c_t + \sum_{\substack{s\in\{C,D\}\\ i\in\{\text{own},\text{other}\}}} \beta^{s,i} u_t^{s,i}) \tag{3}$$

Results are shown in table 5 and are again in line with hypothesis ASYM$_{\text{NB}}$. The left



$(\beta^{c,\text{own}}, \beta^{d,\text{own}})$ (see table 5) are displayed as $\bullet$, $(\beta^{c,\text{other}}, \beta^{d,\text{other}})$ are displayed as $\circ$. The left part of the figure shows the original coefficient, the right part uses the transformation $\lambda^i := \beta^{c,i} - \beta^{d,i}$ and $\rho^i := \beta^{c,i} + \beta^{d,i}$ (see table 6).

FIGURE 6: Estimation of equation (3) for the baseline treatment

part of figure 6 shows the estimated coefficients graphically.

To make the interpretation of the coefficients easier we use a simple translation. We will call a player who is characterised by $\beta^c = -\beta^d$ a *learning* player. Such a player acts according to hypothesis SYM_{STR}. Estimates of β^c and β^d for such a player should be found on the dotted line in the left part of figure 6.

Player might, however, not only learn but might understand the prisoners' dilemma nature of the game and behave in a reciprocating way. We will call a player *reciprocating* if the player chooses C more frequently when payoffs of either C or D are high, since high payoffs indicate the presence of other cooperators. Such a player is characterised by points on the dashed line in the left part of figure 6.

As a measure for learning we take, hence, $\lambda^i := \beta^{c,i} - \beta^{d,i}$ for $i \in \{\text{own}, \text{other}\}$. As a measure for reciprocity we take $\rho^i := \beta^{c,i} + \beta^{d,i}$ for $i \in \{\text{own}, \text{other}\}$. If a player learns as assumed in hypothesis SYM_{STR} then $0 < \beta^{c,i} = -\beta^{d,i}$ or, in other words, $\lambda^i > 0$ and $\rho^i = 0$ (this was the implicit assumption when we estimated equation (2)). If, however, a player is only reciprocating $0 < \beta^{c,i} = \beta^{d,i}$ or $\lambda^i = 0$ and $\rho^i > 0$. Characteristics of these expressions are shown in table 6 and in the right part of figure 6.

We first test the simplifying assumption that we made above in section 4.1.2. When we estimated equation (2) we implicitly assumed hypothesis SYM_{STR}, i.e. players are equally sensitive to payoffs from the two strategies C and D. Then we should have in equation (3) $\forall i \in \{\text{own}, \text{other}\} : \beta^{c,i} = -\beta^{d,i}$, or, in our transformation, we should have $\rho = 0$. In the left part of figure 6 estimates should be on the dotted line. We use a Wald test to jointly test $\forall i \in \{\text{own}, \text{other}\} : \rho^i = 0$ and find for circles ($\chi^2(2) = 6.01$, $P_{>\chi^2} = 0.0496$) and for

coeff. from eq. (3)	Learning and reciprocity in circles					
	β	σ	z	$P_{>\lvert z\rvert}$	95% conf. interval	
λ^{own}	.1759673	.018847	9.34	0.000	.1390278	.2129068
ρ^{own}	.0060315	.0189445	0.32	0.750	−.031099	.0431621
λ^{other}	.0596688	.026113	2.29	0.022	.0084883	.1108494
ρ^{other}	.0327524	.0211634	1.55	0.122	−.0087272	.0742319
coeff. from eq. (3)	Learning and reciprocity in groups					
	β	σ	z	$P_{>\lvert z\rvert}$	95% conf. interval	
λ^{own}	.05732	.0199049	2.88	0.004	.0183071	.0963329
ρ^{own}	.0878133	.0228156	3.85	0.000	.0430956	.132531
λ^{other}	.2368891	.042429	5.58	0.000	.1537298	.3200483
ρ^{other}	−.0013172	.0200217	−0.07	0.948	−.0405591	.0379247

TABLE 6: Learning λ and reciprocity ρ as estimated in the GEE estimation of equation (3) using the transformation $\lambda^i := \beta^{c,i} - \beta^{d,i}$ and $\rho^i := \beta^{c,i} + \beta^{d,i}$

groups ($\chi^2(2) = 27.82$, $P_{>\chi^2} = 0.0000$) a significant amount of asymmetric learning. So, while the simplifying approach from section 4.1.2 may help us gain a first insight, it seems justified to abandon hypothesis SYM_{STR} and to introduce ρ as an additional dimension, allowing for differences in learning from C and D.

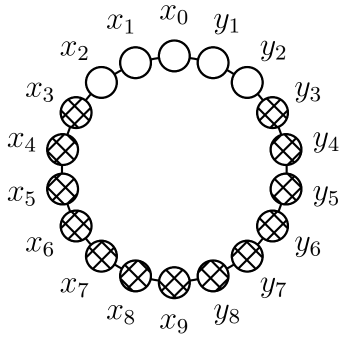
Let us next compare hypotheses SYM_{NB} and ASYM_{NB}. Following hypothesis ASYM_{NB} we should expect relatively more reinforcement and less imitation in spatial structures (in circles). In groups we should expect the opposite. Our data confirms this hypothesis. In circles we find $\lambda^{\text{own}} > \lambda^{\text{other}}$ ($\chi^2(2) = 8.60$, $P_{>\chi^2} = 0.0034$) while in groups $\lambda^{\text{own}} < \lambda^{\text{other}}$ ($\chi^2(2) = 10.86$, $P_{>\chi^2} = 0.0010$).

This explains why we do not find more cooperation in circles. Players put relatively more weight on their own experience the more spatial a structure becomes. As a result the mechanism that would otherwise support growth of cooperation in a spatial structure ceases to work.

4.2 A treatment with some computerised players

In section 2 we explained how imitation of successful neighbours supports cooperation in a spatial environment. This argument relies on the assumption of an initial cluster of cooperators of sufficient size — with our payoffs we need at least five neighbouring cooperators. But how does such a cluster appear? An evolutionary game theorist might argue that we only have to wait long enough until such a cluster appears with a mutation. Sessions in our experiment, however, last only for a limited number of periods, and if the cooperative cluster does not appear during this time cooperation might never get started.

To give cooperation in circles the best possible conditions we therefore introduced a



The five white dots indicate the position of computerised players that always play C. The remaining dots indicate the position of the human players.

Figure 7: The structure of circles with some computerised players

cluster of five computerised players into the circle. In figure 7 players x_2, x_1, x_0, y_1, y_2 are played by the computer and cooperate in every period.[13] The remaining players are human which obtain the same information as in the above treatment (section 3). Players x_3, x_4, y_3, y_4 do not know that their neighbours are computers. The detailed behaviour of the human players is shown in appendix B.3.

4.2.1 Stage game behaviour in circles with some computerised players

Figure 8 shows the frequency of cooperation depending on the distance to the computerised players. Players with a smaller distance to the computerised players cooperate significantly more.[14] The four players which are closest to the computerised players and who obtain information about payoffs of these players (x_3, x_4, y_3, y_4 in figure 7) cooperate more frequently.[15] The average frequency of cooperation in the circle with some computerised players is slightly, but not significantly, higher than in the baseline treatment.[16] If we drop players x_3, x_4, y_3, y_4 the average frequency of cooperation in circles with computerised players is even slightly (but not significantly) lower than in the baseline treatment.

[13] Participants were told that they would play a game with 18 agents sitting round a circle. They could see that only 13 players were present in the laboratory but in our experiment no participant missed the other five.

We ran four sessions with 13 players and 5 computerised agents. We conducted two more sessions where only 10 players showed up (again with 5 computerised agents). We pool the data from these two sessions with the other four.

[14] A Cuzick-Altman test finds $z = 2.55$, $P_{>|z|} = 0.01$.

[15] a t-test finds $t = -2.68$, $P_{>|t|} = 0.044$, a one sample Wilcoxon signed-rank test finds $z = 2.201$, $P_{>|z|} = 0.0277$.

[16] A t-test finds $t = 1.06$, $P_{>|t|} = 0.303$, a two-sample Wilcoxon rank-sum test finds $z = 0.702$, $P_{>|z|} = 0.4829$.

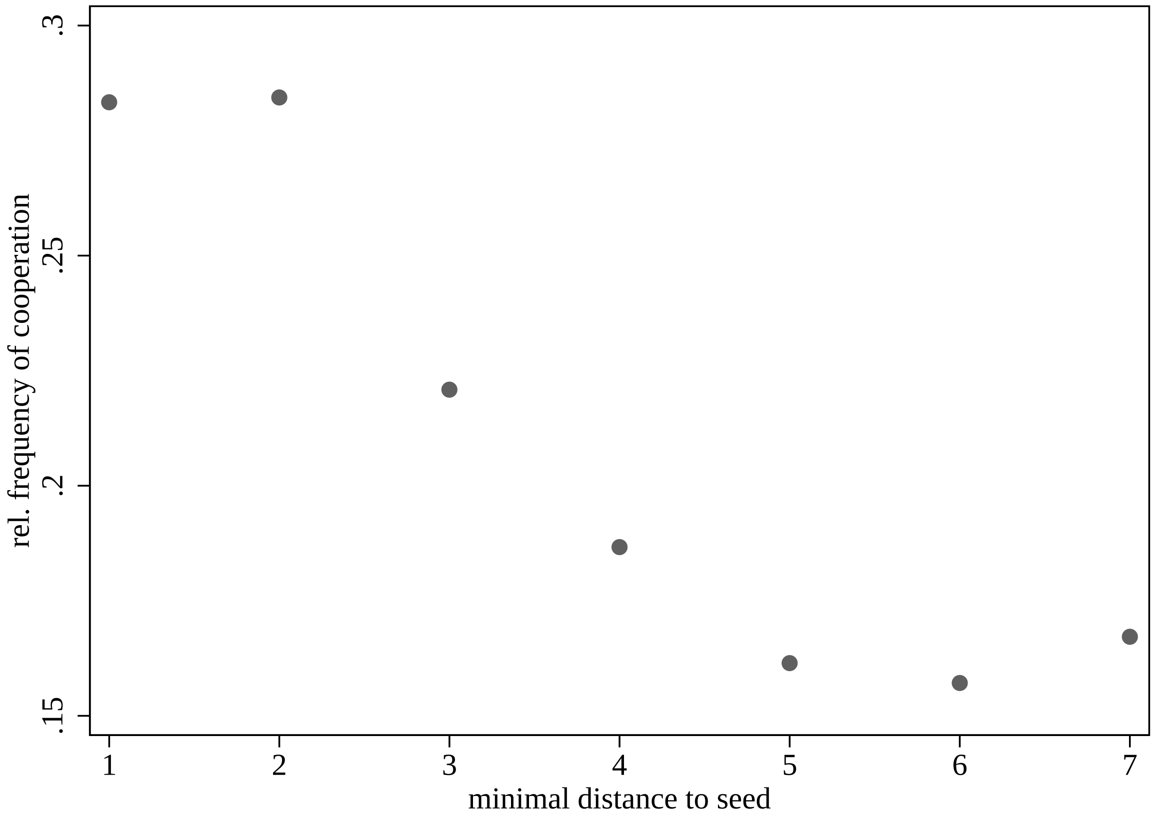
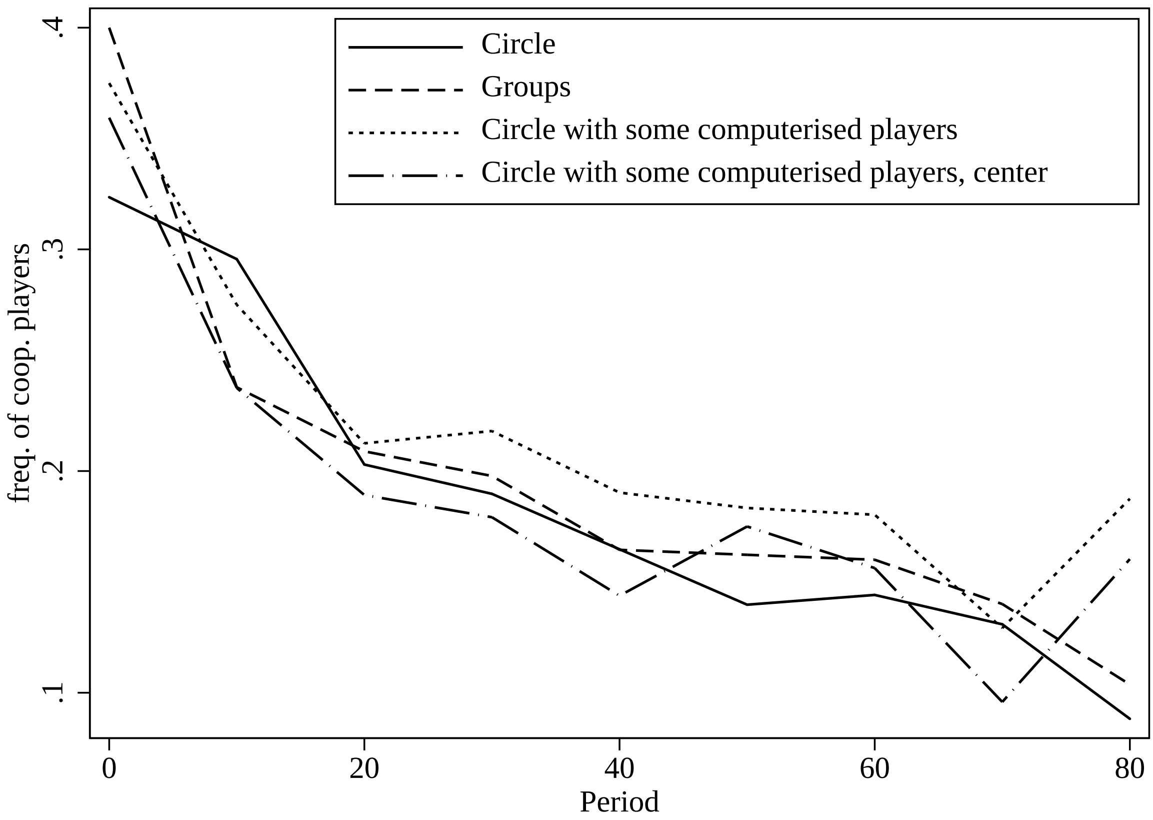


FIGURE 8: Cooperation in circles with some computerised players depending on the distance to the computerised players



The line "circle with some computerised players" shows the relative frequency of all human players, the line "circle with some computerised players, center" excludes players x_3, x_4, y_3, y_4 who set next to the computerised neighbours.

Figure 9: Cooperation in circles with some computerised players

We see, hence, that there is some imitation going on, but, as we will see in the next paragraph, this is not enough to increase overall cooperation substantially.

Figure 9 shows the development of cooperation in the baseline treatment and in circles with computerised players. For the circles with some computerised players we show two lines. The upper one shows all participants, including those that have immediate computerised neighbours. The latter cooperate more than those who are farther away from the cluster of cooperators. When we exclude them, we obtain the lower line. If we compare the average frequency of cooperation in circles with some computerised players with the one in groups we find no significant difference (a t-test finds $t = 0.68$, $P_{>|t|} = 0.507$, a two-sample Wilcoxon rank-sum test finds $z = 0.589$, $P_{>|z|} = 0.5557$).

coeff. from eq. (3)	Learning in circles with some computerised players					
	β	σ	z	$P_{>\lvert z\rvert}$	95% conf. interval	
β_c	−.2013955	.0923885	−2.18	0.029	−.3824737	−.0203173
β^{own}	.1071798	.0089123	12.03	0.000	.0897119	.1246476
β^{other}	.0679119	.0080964	8.39	0.000	.0520432	.0837806
β_0	−.7267049	.0546413	−13.30	0.000	−.8337999	−.6196098

TABLE 7: GEE population-averaged estimation of equation (2) for circles with some computerised players

coeff. from eq. (3)	Learning from C and D in circles with some computerised players					
	β	σ	z	$P_{>\lvert z\rvert}$	95% conf. interval	
β_c	−.6312252	.128256	−4.92	0.000	−.8826023	−.379848
$\beta^{c,\text{own}}$	.1071357	.0111059	9.65	0.000	.0853685	.1289029
$\beta^{d,\text{own}}$	−.1061983	.0124075	−8.56	0.000	−.1305165	−.0818801
$\beta^{c,\text{other}}$	.0844496	.0114302	7.39	0.000	.0620468	.1068523
$\beta^{d,\text{other}}$	−.0336295	.0134328	−2.50	0.012	−.0599573	−.0073016
β_0	−1.074731	.1057231	−10.17	0.000	−1.281945	−.8675178

TABLE 8: GEE population-averaged estimation of equation (3) for circles with some computerised players

To summarise: even when we give players in circles the best possible starting conditions we do not find support for hypothesis COOP — players still do not cooperate more in circles than in groups.

4.2.2 Learning in circles with some computerised players

Theoretically we do not see any reason why learning behaviour in the treatment with some computerised players should differ from learning in the baseline treatment. This is confirmed by our estimations. Tables 7, 8, 9, and figure 10 show GEE estimates for circles with some computerised players similar to tables 4, 5, 6 for circles from the baseline treatment. Also in circles with some computerised players we find $\lambda^{\text{own}} > \lambda^{\text{other}}$ ($\chi^2(2) = 11.82$, $P_{>\chi^2} = 0.0006$).

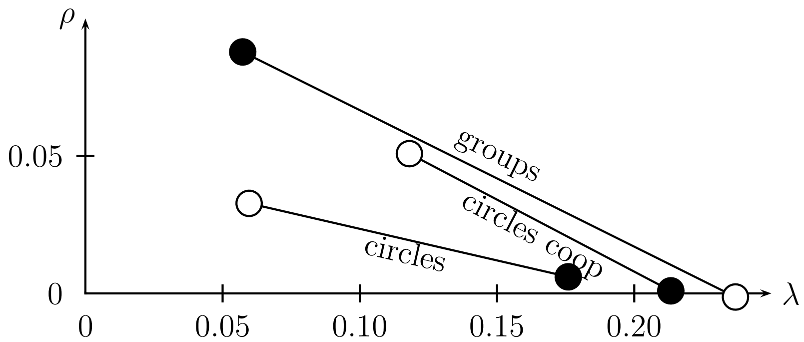
4.3 A treatment with information about the payoff matrix

The treatments described in sections 4.1 and 4.2 where designed to disentangle imitation from learning from own experience (reinforcement). Our focus was on learning. Indeed, in the estimations of equation (3) learning was significantly stronger than reciprocity.[17]

[17] A joint test of $\forall i \in \{\text{own}, \text{other}\} : \lambda^i = \rho^i$ yields in circles $\chi^2(2) = 75.60$, $P_{>\chi^2} = 0.0000$ and in groups $\chi^2(2) = 35.51$, $P_{>\chi^2} = 0.0000$.

coeff. from eq. (3)	Learning and reciprocity in circles with some computerised players					
	β	σ	z	$P_{>\|z\|}$	95% conf. interval	
λ^{own}	.213334	.0178997	11.92	0.000	.1782511	.2484168
ρ^{own}	.0009374	.0153027	0.06	0.951	$-.0290554$	.0309302
λ^{other}	.118079	.0163523	7.22	0.000	.086029	.150129
ρ^{other}	.0508201	.0188356	2.70	0.007	.013903	.0877373

TABLE 9: Learning λ and reciprocity ρ as estimated in the GEE estimation of equation (3) for circles with some computerised players



In addition to the baseline treatment that is shown in figure 6 we show here (larger than in figure 6) estimates for circles with computerised cooperators (circles coop). $(\beta^{c,\text{own}}, \beta^{d,\text{own}})$ are displayed as $\bullet$, $(\beta^{c,\text{other}}, \beta^{d,\text{other}})$ are displayed as $\circ$. We use the transformation $\lambda^i := \beta^{c,i} - \beta^{d,i}$ and $\rho^i := \beta^{c,i} + \beta^{d,i}$. See also table 9 for circles with computerised cooperators and table 6 for the baseline treatment.

FIGURE 10: Estimation of equation (3) with computerised cooperators

History						
Round	Your strategy and gains are		your neighbours received			
...	...	...	...	...	...	...
...	C	10	20	15	14	9

The table shows payoff information as seen by player 1 from table 2

TABLE 10: Example of payoff representation in the detailed information treatment

Still, reciprocity might have some impact. In the discussion of tables 6 and 9 and in figure 6 we found that reciprocity is small in circles but larger in groups. One reason for this difference might be that in groups participants of the experiment understand the prisoners' dilemma nature of the game more easily despite the fact that participants do not have access to the payoff matrix of the game. Having understood that a game is a prisoners' dilemma allows players to analyse the game strategically and to rely less on imitation or reinforcement.

In the current section we want to better understand the influence of reciprocity. We study a treatment where players know the payoff matrix of the game as shown in table 1, i.e. they are able to see that they are playing a prisoners' dilemma. Bosch-Domènech and Vriend (2003) show in a Cournot game that the amount of imitation is not affected by the available information. We will see below that in our setup the available information affects the degree of imitation considerably.

We will call the treatment the *treatment with detailed information*. Together with the payoff matrix we present the information during the course of the session as shown in table 10. Players do not see average payoffs, as in sections 4.1 and 4.2 but payoffs of each individual player. We actually run this treatment before the other treatments mentioned above. At this time we wanted to present payoffs in the different rounds (table 10) in a way that is similar to the presentation of payoffs of the game (table 1). This means that payoffs in the treatment with detailed information are presented not in exactly the same way as they are presented in the other treatments. However, we think that the difference is only small and should not affect the results.

Consider player 1 from table 2 who has two neighbours with action C and two other neighbours with D. Information about payoffs in this round is presented as shown in table 10. Player 1's own payoff is shown as 10, and displayed next to the player's own action C. The player has two neighbours with action C and payoffs 20 and 15 respectively. The two other neighbours choose action D and receive payoffs 14 and 9. Payoffs obtained with either C or D are displayed in different colours in the experiment. The payoffs are

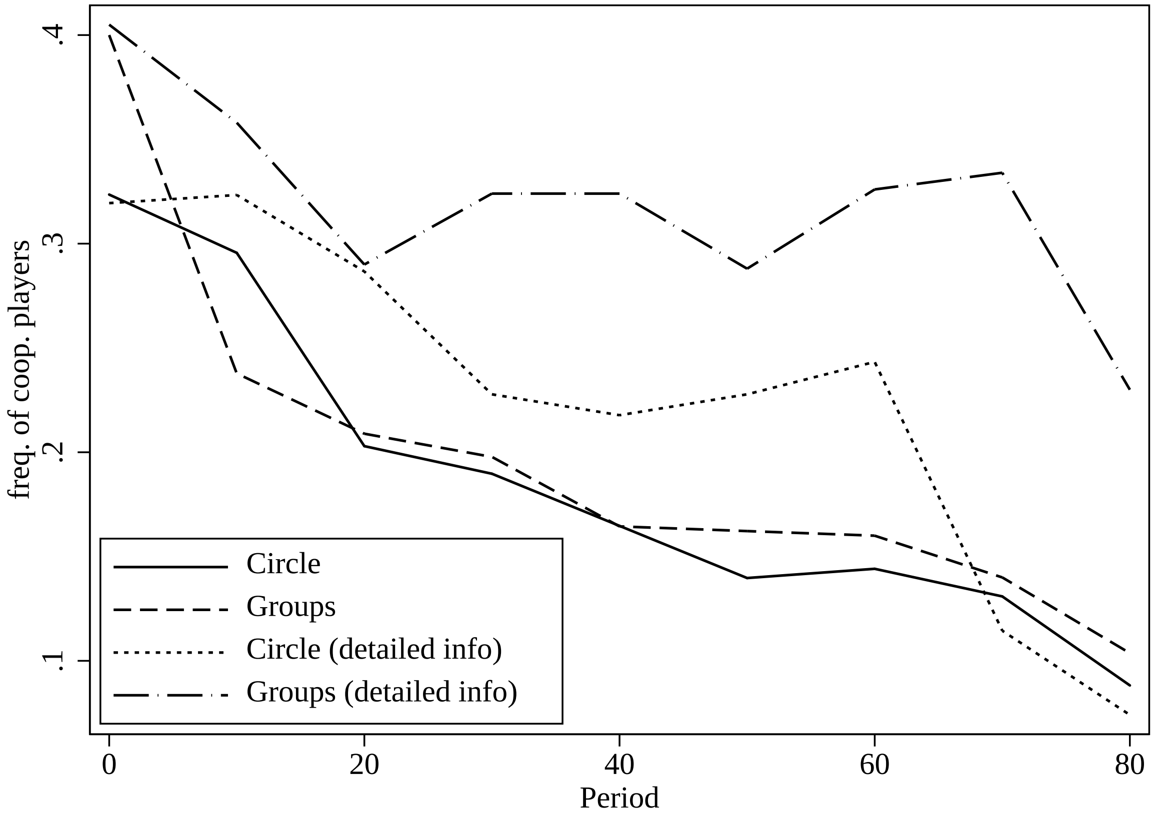


FIGURE 11: Frequency of cooperative players over time in the detailed information and in the baseline treatment

shown in the rightmost column and ordered from highest to lowest. Thus, it is not obvious to the player *which* of the player's neighbours has chosen a certain action and received a certain payoff.

4.3.1 Stage game behaviour in the detailed information treatment

In figure 11 we show the relative frequency of cooperation in the detailed information treatment as a solid line for circles and as a dotted line for groups. For comparison the figure also shows results for the baseline treatment with dashed lines.

With detailed information we find significantly more cooperation than without detailed information in circles and also in groups.[18] However, the increase in the frequency of cooperation is not the same in the two structures. While without detailed information

[18]For circles we find in a t-test $t = 2.95$, $P_{>|t|} = 0.018$, in two-sample Wilcoxon rank-sum test we find $z = -1.960$, $P_{>|z|} = 0.0500$. For groups we find in a t-test $t = 3.65$, $P_{>|t|} = 0.002$, in two-sample Wilcoxon rank-sum test we find $z = -2.694$, $P_{>|z|} = 0.0071$.

coeff. from eq. (3)	Learning own and others' payoff in circles					
	β	σ	z	$P_{>\|z\|}$	95% conf. interval	
β_c	3.019047	.0658539	45.84	0.000	2.889976	3.148119
β^{own}	.0497975	.0061767	8.06	0.000	.0376913	.0619036
β^{other}	.0139596	.0079059	1.77	0.077	−.0015357	.0294549
β_0	−1.888248	.0564207	−33.47	0.000	−1.99883	−1.777665
coeff. from eq. (3)	Learning from own and others' payoff in groups					
	β	σ	z	$P_{>\|z\|}$	95% conf. interval	
β_c	2.560402	.0862198	29.70	0.000	2.391414	2.72939
β^{own}	.0294385	.0065031	4.53	0.000	.0166926	.0421844
β^{other}	−.0839884	.015045	−5.58	0.000	−.113476	−.0545007
β_0	−2.239785	.1129711	−19.83	0.000	−2.461205	−2.018366

TABLE 11: GEE population-averaged estimation of equation (2) when detailed information is given

in section 4.1.1 we did not find a significant difference between circles and groups we find with detailed information more cooperation in groups than in circles[19], i.e. an even stronger contradiction of hypothesis COOP than what we found above. A reason might be that in this treatment learning, which supports COOP in circles, is relatively less important and reciprocity becomes more important.

4.3.2 Learning and reciprocity in the detailed information treatment

Similar to the estimations in sections 4.1.5 and 4.2.2 we estimate again equations (2) and (3). Results are shown in tables 11, 12, 13 and in figure 12. Let us skip tables 11 and 12 and concentrate on table 13 and figure 12. Not only in the baseline treatment, also with detailed information players learn more from their own experience in circles.[20] In contrast to the baseline treatment players learn more from their own experience in groups too.[21]

In the treatment without detailed information, most reciprocity terms ρ were not significantly different from zero. Now, in the treatment with detailed information, they are.[22] If reciprocity plays a relatively larger role in this treatment, hypothesis COOP, which is based on learning, does not hold.

In table 13 and figure 12 we also find two coefficients with unexpectedly negative signs. In groups with detailed information ρ^{own} and λ^{other} are both negative. Technically

[19]In a t-test we find $t = 2.89$, $P_{>t} = 0.006$, in two-sample Wilcoxon rank-sum test we find $z = 1.715$, $P_{>z} = 0.0432$.

[20]Testing $\lambda^{\text{own}} > \lambda^{\text{other}}$ yields a $\chi^2 = 18.45$, $P_{>\chi^2} = 0.0000$.

[21]Testing $\lambda^{\text{own}} > \lambda^{\text{other}}$ yields a $\chi^2 = 27.80$, $P_{>\chi^2} = 0.0000$ in groups.

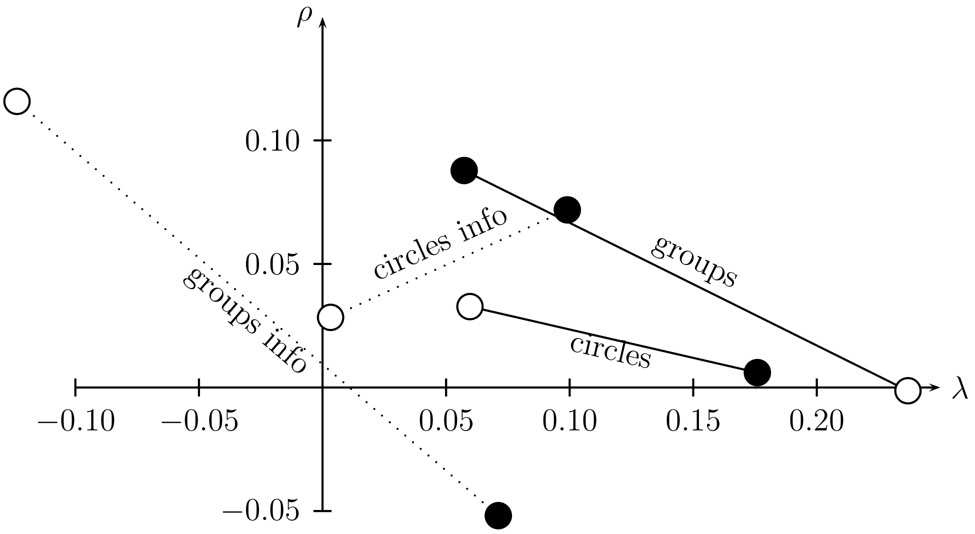
[22]$\chi^2 = 3.38$, $P_{>\chi^2} = 0.0659$ in circles, $\chi^2 = 32.67$, $P_{>\chi^2} = 0.0000$ in groups.

coeff. from eq. (3)	Learning from C and D in circles with detailed information					
	β	σ	z	$P_{>\|z\|}$	95% conf. interval	
β_c	2.558246	.0968281	26.42	0.000	2.368467	2.748026
$\beta^{c,\text{own}}$	.0854751	.0079583	10.74	0.000	.0698772	.1010731
$\beta^{d,\text{own}}$	−.0135855	.0085466	−1.59	0.112	−.0303365	.0031655
$\beta^{c,\text{other}}$	.0158278	.0090937	1.74	0.082	−.0019956	.0336512
$\beta^{d,\text{other}}$	.0125569	.0130059	0.97	0.334	−.0129342	.0380479
β_0	−2.465581	.0916312	−26.91	0.000	−2.645175	−2.285987
coeff. from eq. (3)	Learning from C and D in groups with detailed information					
	β	σ	z	$P_{>\|z\|}$	95% conf. interval	
β_c	1.662233	.1150202	14.45	0.000	1.436797	1.887668
$\beta^{c,\text{own}}$	.009631	.0094656	1.02	0.309	−.0089213	.0281833
$\beta^{d,\text{own}}$	−.061498	.0111853	−5.50	0.000	−.0834207	−.0395752
$\beta^{c,\text{other}}$	−.0039144	.0175425	−0.22	0.823	−.038297	.0304682
$\beta^{d,\text{other}}$	.1197077	.0173654	6.89	0.000	.0856721	.1537434
β_0	−2.278528	.1198936	−19.00	0.000	−2.513516	−2.043541

TABLE 12: GEE population-averaged estimation of equation (3) with detailed information

coeff. from eq. (3)	Learning and reciprocity in circles with detailed information					
	β	σ	z	$P_{>\|z\|}$	95% conf. interval	
λ^{own}	.0990606	.0123706	8.01	0.000	.0748148	.1233065
ρ^{own}	.0718897	.0109419	6.57	0.000	.0504439	.0933355
λ^{other}	.003271	.0168562	0.19	0.846	−.0297666	.0363085
ρ^{other}	.0283847	.0148177	1.92	0.055	−.0006576	.0574269
coeff. from eq. (3)	Learning and reciprocity in groups with detailed information					
	β	σ	z	$P_{>\|z\|}$	95% conf. interval	
λ^{own}	.071129	.013689	5.20	0.000	.044299	.0979589
ρ^{own}	−.051867	.0155572	−3.33	0.001	−.0823585	−.0213754
λ^{other}	−.1236221	.0312888	−3.95	0.000	−.184947	−.0622973
ρ^{other}	.1157934	.0154792	7.48	0.000	.0854546	.1461321

TABLE 13: Learning λ and reciprocity ρ as estimated in the GEE estimation of equation (3) with detailed information



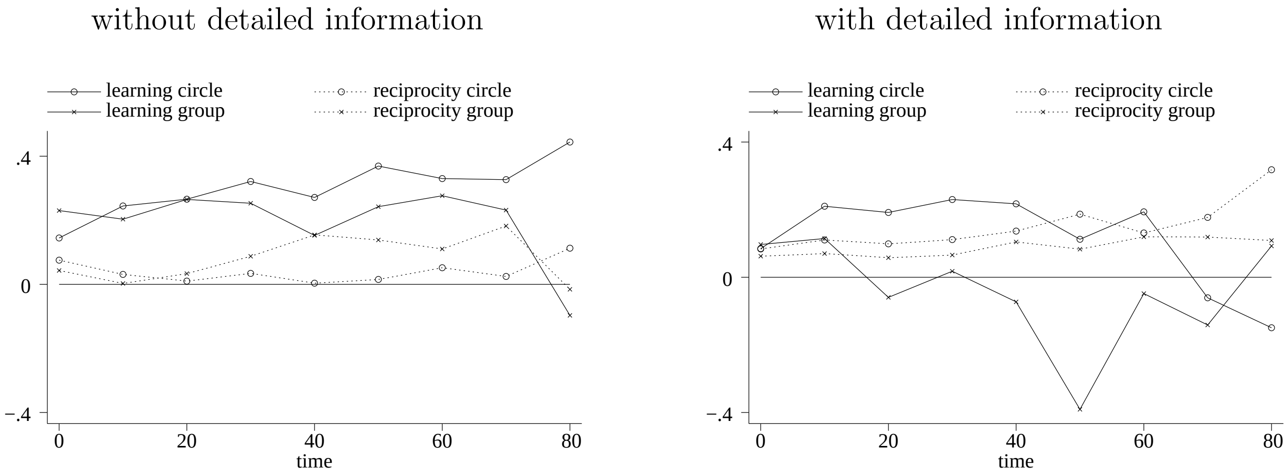
In addition to the baseline treatment that is already shown in figure 6 (table 6) we show here also with dotted lines the estimation for the treatment with detailed information (table 13). $(\beta^{c,\text{own}}, \beta^{d,\text{own}})$ are displayed as $\bullet$, $(\beta^{c,\text{other}}, \beta^{d,\text{other}})$ are displayed as $\circ$. We use the transformation $\lambda^i := \beta^{c,i} - \beta^{d,i}$ and $\rho^i := \beta^{c,i} + \beta^{d,i}$.

FIGURE 12: Estimation of equation (3) for the treatment with detailed information

the negative λ_{other} in groups results from a large value of $\beta^{d,\text{other}}$, i.e. players cooperate more when average payoffs of Ds are large. Stronger reciprocity might be a reason, which should, a priori, affect $\beta^{c,\text{other}}$ in the same way. However, since there are more Ds than Cs in a neighbourhood, Ds average payoffs might be considered more reliable information and, therefore, $\beta^{d,\text{other}}$ might be larger. Another reason might be that participants do not seem to react linearly in this treatment. If a certain level of cooperation is exceeded players stop reciprocating. Whatever the reasons for these coefficients are, in any case we should be careful not to over-stretch the interpretation of our simple learning model in the group case with detailed information.

4.4 Learning how to learn

In the discussion in the previous sections we always assumed that learning and reciprocity were constant over time. In figure 13 we see that changes over time do not follow an obvious pattern. The figure shows results of estimating the GEE population-averaged model of equation (3) for subsets of 10 adjoining periods of all experiments without detailed information. To simplify the figure we show $\sum_{i\in\{\text{own,other}\}} \lambda^i$ as an indicator for learning and $\sum_{i\in\{\text{own,other}\}} \rho^i$. All major results that we found above seem to hold during the whole experiment. Trends, if they can be found at all, are weak and not significant.



The figure show $\sum_{i\in\{\text{own,other}\}} \lambda^i$ as a measure for learning and $\sum_{i\in\{\text{own,other}\}} \rho^i$ as a measure for reciprocity.

FIGURE 13: Learning and reciprocity over time

5 Conclusion

In this paper we have tried to find out how players learn and how their learning behaviour depends on the heterogeneity of their environment. In particular we compared actual learning behaviour with copy-best learning, which is a common model of learning in the literature. We have seen that copy-best learning is characterised by a large amount of imitation and a smaller amount of reinforcement. Learning of human players, in particular in heterogeneous structures, does not fit this description. Players do imitate sometimes, however, learning from own experience has a stronger influence.

When players are in a homogeneous environment (as they are in our group treatment) then imitation plays a relatively larger role as compared to a heterogeneous environment (as in our circle treatment). This is interesting for the literature that builds upon imitation in local interaction models to explain cooperation. This literature explains very elegantly how local interaction supports cooperation in an evolutionary model. The argument, however, depends substantially on imitation. If, as we find in our experiments, there is only a small amount of imitation in particular in spatial settings, cooperation breaks down.

Also the available information affects the amount of imitation in an intuitive way. The more information is available the less players rely on imitation and the more reciprocity plays a role. Given that in other games information is not affected by information (Bosch-

Domènech and Vriend 2003) complexity of the game might be a moderating factor. In the fairly complex game of Bosch and Vriend players might disregard information altogether, always relying on imitation. In simpler games, like the prisoners' dilemma, information, if available, may be helpful and displaces imitation.

There are other questions that we had to leave aside. The development of learning over time (see figure 13) should be further explored. Also, disentangling of the parameters of our regression into learning and reciprocity effects was helpful in the analysis but lead to sometimes unexpected coefficients. Given the sheer number of coefficients that we estimate this may be hardly surprising, still, we feel that more effects than learning through reinforcement, imitation and reciprocity might be at work here.

References

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

Bonacich, P., G. H. Shure, J. P. Kahan, and R. J. Meeker, 1976, Cooperation and Group Size in the N-Person Prisoners' Dilemma, *Journal of Conflict Resolution*, 20(4), 687–706.

Bonhoeffer, S., R. M. May, and M. A. Nowak, 1993, More Spatial Games, *International Journal of Bifurcation and Chaos*, 4, 33–56.

Bosch-Domènech, A., and N. J. Vriend, 2003, Imitation of successful behavior in cournot markets, *The Economic Journal*, pp. 495–.

Cassar, A., 2002, Coordination and Cooperation in Local, Random and Small World Networks: Experimental Evidence, in *Proceedings of the 2002 North American Summer Meetings of the Econometric Society: Game Theory*, ed. by D. K. Levine, W. Zame, L. Ausubel, P.-A. Chiappori, B. Ellickson, A. Rubinstein, and L. Samuelson.

Ellison, G., 1993, Learning, Local Interaction, and Coordination, *Econometrica*, 61, 1047–1071.

Erev, I., and A. E. Roth, 1998, Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria, *American Economic Review*, 88(4), 848–81.

Eshel, I., L. Samuelson, and A. Shaked, 1998, Altruists, Egoists, and Hooligans in a Local Interaction Model, *The American Economic Review*, 88, 157–179.

Fox, J., and M. Guyer, 1977, Group Size and Other's Strategy in an N-Person Game, *Journal of Conflict Resolution*, 21(2), 323–338.

Huck, S., H.-T. Normann, and J. Oechssler, 1999, Learning in Cournot Oligopoly — An Experiment, *The Economic Journal*, 109, C80–C95.

——— , 2000, Does information about competitors' actions increase or decrease competition in experimental oligopoly markets, *International Journal of Industrial Organization*, 18, 39–57.

Keser, C., K.-M. Ehrhart, and S. K. Berninghaus, 1998, Coordination and Local Interaction: Experimental Evidence, *Economics Letters*, 58(3), 269–75.

Kirchkamp, O., 1999, Simultaneous Evolution of Learning Rules and Strategies, *Journal of Economic Behavior and Organization*, 40(3), 295–312, http://www.kirchkamp.de/.

——— , 2000, Spatial Evolution of Automata in the Prisoners' Dilemma, *Journal of Economic Behavior and Organization*, 43(2), 239–262, http://www.kirchkamp.de/.

Kosfeld, M., 2003, Network Experiments, Discussion paper, University of Zurich, Institute for Empirical Research in Economics, Blümlisalpstrasse 10, CH-8006 Zürich, email: kosfeld@iew.unizh.ch.

Liang, K.-Y., and S. L. Zeger, 1986, Longitudinal data analysis using generalised linear models, *Biometrica*, 73, 13–22.

Lindgreen, K., and M. G. Nordahl, 1994, Evolutionary dynamics of spatial games, *Physica D*, 75, 292–309.

Nowak, M., K. Sigmund, and E. El-Sedy, 1993, Automata, Repeated Games and Noise, Mimeo, University of Oxford.

Nowak, M. A., and R. M. May, 1992, Evolutionary Games and Spatial Chaos, *Nature*, 359, 826–829.

——— , 1993, The Spatial Dilemmas of Evolution, *International Journal of Bifurcation and Chaos*, 3, 35–78.

Offerman, T., J. Potters, and J. Sonnemans, 2002, Imitation and Belief Learning in an Oligopoly Experiment, *Review of Economic Studies*, 69, 973–997.

Offerman, T., and J. Sonnemans, 1998, Learning by experience and learning by imitating successful others, *Journal of Economic Behavior and Organization*, 34, 559–575.

Pingle, M., and R. H. Day, 1996, Modes of Economizing Behavior: Experimental Evidence, *Journal of Economic Behavior and Organization*, 29(2), 191–209.

Rogers, W. H., 1993, Regression standard errors in clustered samples, in *Stata Technical Bulletin*, vol. 13, pp. 19–23. Stata, Reprinted in Stata Technical Bulletins, vol. 3, 88-94.

Selten, R., and J. Apesteguia, 2002, Experimentally Observed Imitation and Cooperation in Price Competition on the Circle, Discussion Paper bgse19_2002, Bonn Econ, University of Bonn, Germany.

Selten, R., and A. Ostmann, 2001, Imitation Equilibrium, *Homo Oeconomicus*, 43, 111–149.

Vega-Redondo, F., 1997, The evolution of Walrasian behavior, *Econometrica*, 65, 375–384.

A List of Sessions

Overview:

Number of sessions in different treatments			
	information...		
	detailed	less	
			5 computerised cooperators
group	9	10	
circle	5	5	6

Parameters of each session:

	structure	information	computerised cooperators	number of players
1.	Group	baseline info	0	5
2.	Group	baseline info	0	5
3.	Group	baseline info	0	5
4.	Group	baseline info	0	5
5.	Group	baseline info	0	5
6.	Group	baseline info	0	5
7.	Group	baseline info	0	5
8.	Group	baseline info	0	5
9.	Group	baseline info	0	5
10.	Group	detailed info	0	5
11.	Group	detailed info	0	5
12.	Group	detailed info	0	5

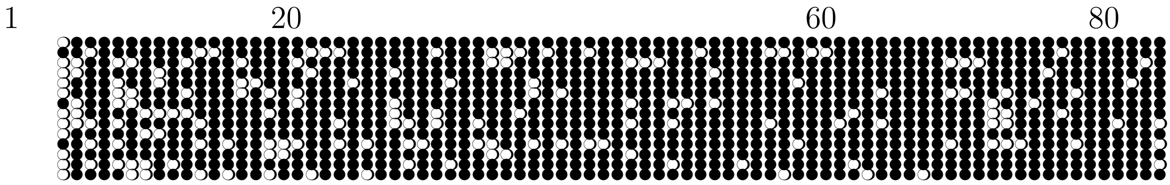
continued on next page

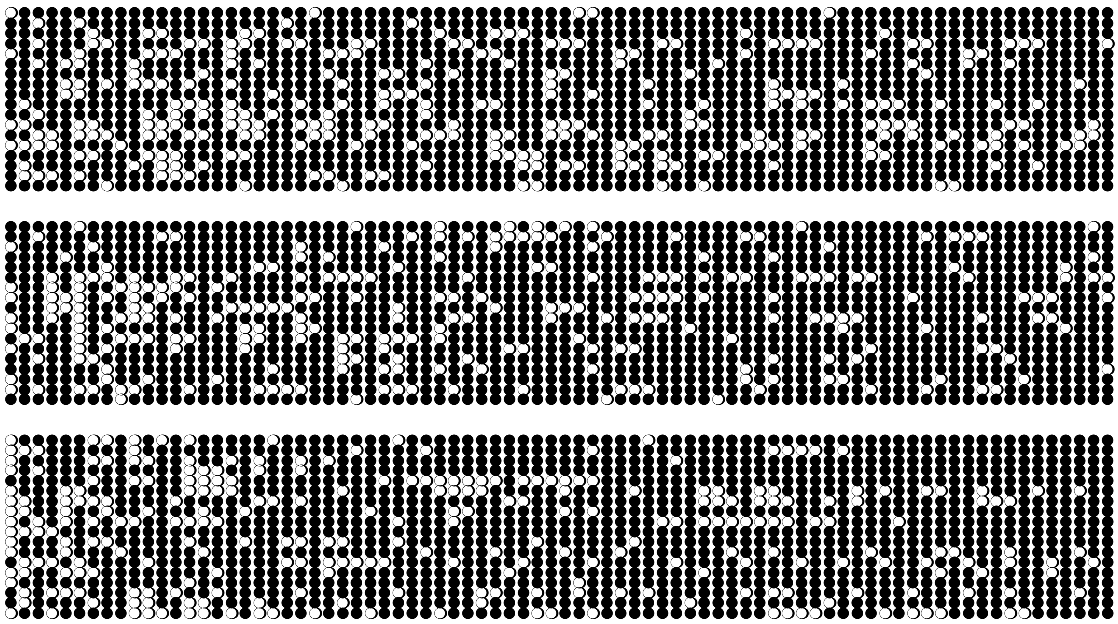
				continued from previous page
	structure	information	computerised cooperators	number of players
13.	Group	detailed info	0	5
14.	Group	detailed info	0	5
15.	Group	detailed info	0	5
16.	Group	detailed info	0	5
17.	Group	detailed info	0	5
18.	Group	detailed info	0	5
19.	Group	detailed info	0	5
20.	Circle	baseline info	0	14
21.	Circle	baseline info	0	18
22.	Circle	baseline info	0	18
23.	Circle	baseline info	0	18
24.	Circle	baseline info	5	13
25.	Circle	baseline info	5	10
26.	Circle	baseline info	5	13
27.	Circle	baseline info	5	10
28.	Circle	baseline info	5	13
29.	Circle	baseline info	5	13
30.	Circle	detailed info	0	18
31.	Circle	detailed info	0	18
32.	Circle	detailed info	0	18
33.	Circle	detailed info	0	18
34.	Circle	detailed info	0	18

B Raw data

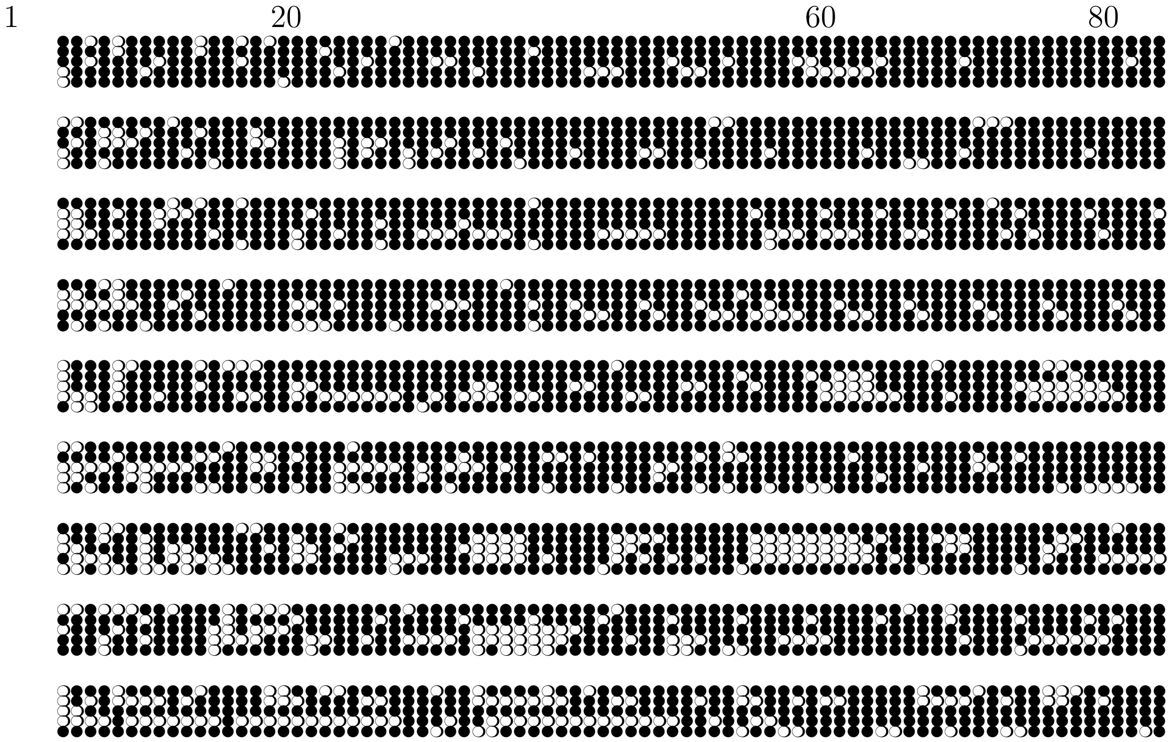
In the following graphs each line represents the actions of a player from period 1 to period 80. Cooperation is shown as ❍, non cooperation as ●. Neighbouring lines correspond to neighbouring players in the experiment. In all treatments without computerised cooperators (sections B.4 to B.2) the last line of each block of lines is in circles always a neighbour of the first line of the same block. In these sections the display of circles is always rotated such that least cooperative players are found in the first and the last lines.

B.1 Circle treatment



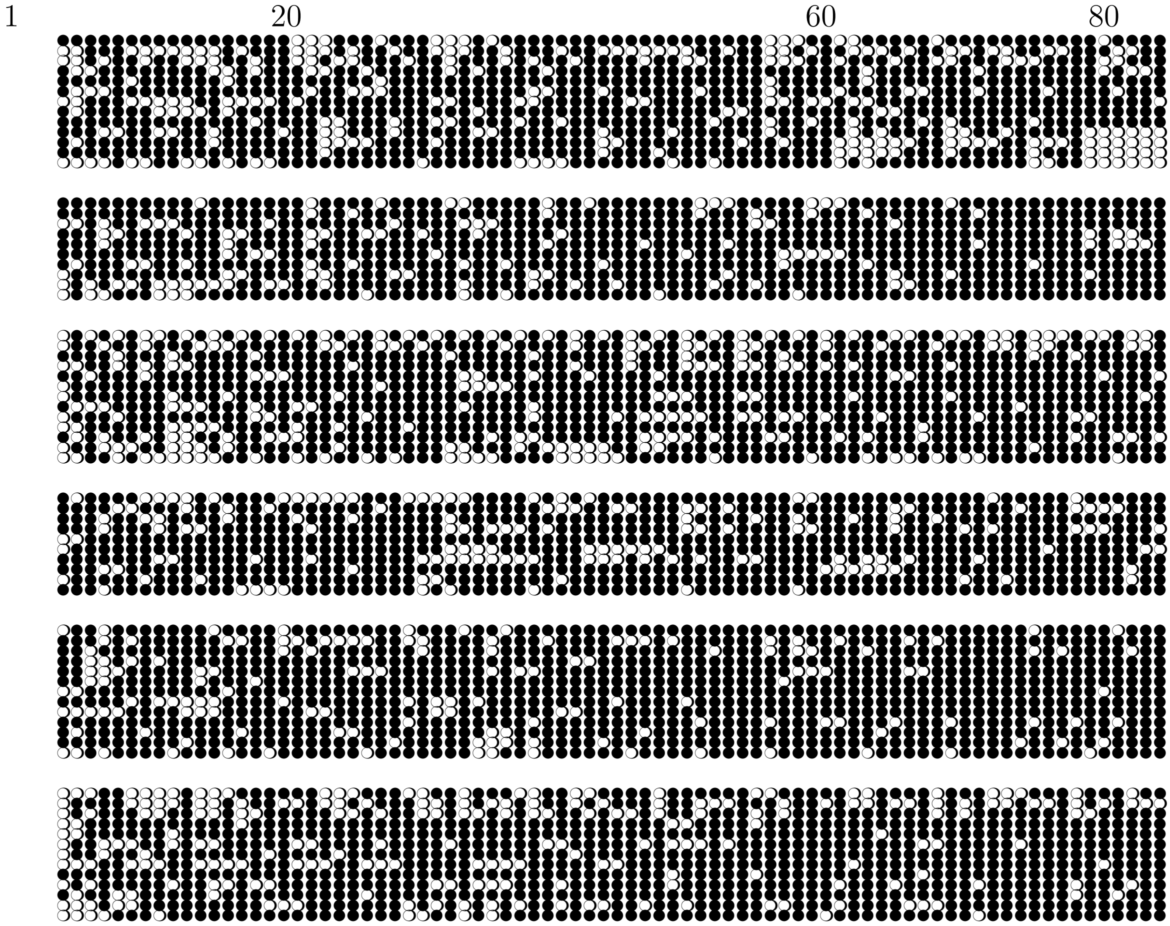


B.2 Group treatment

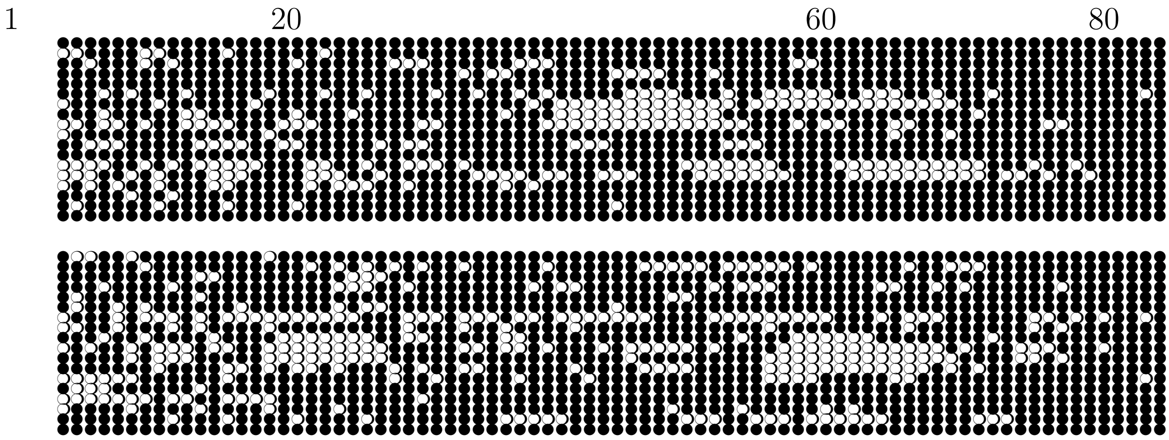


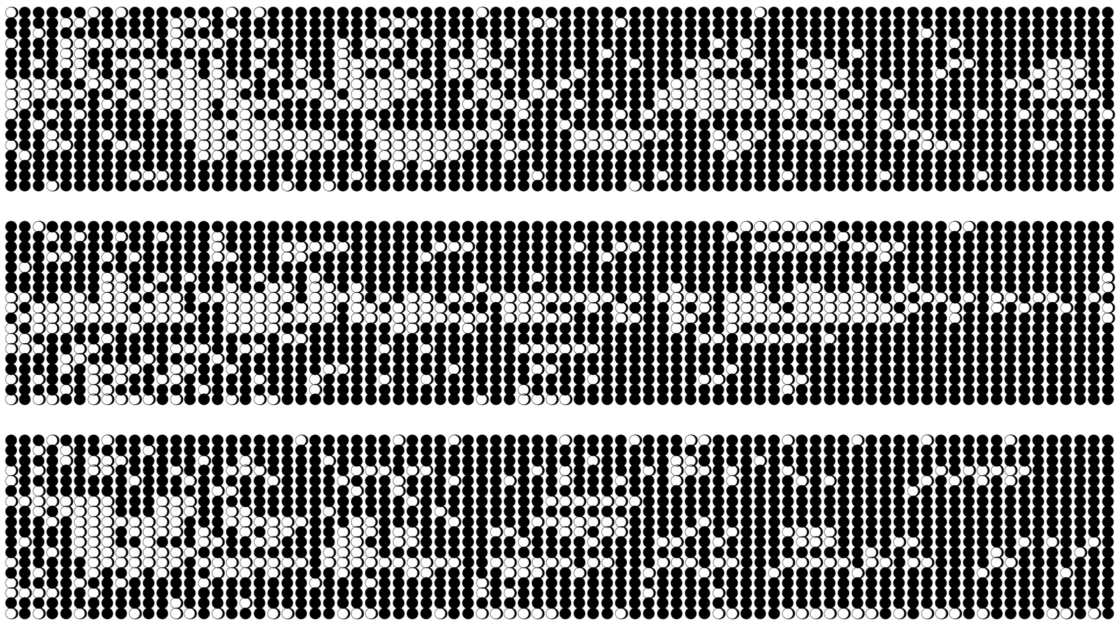
B.3 Circles with some computerised players

In the display of the circles the five computerised cooperators are not displayed. Their location is on top of the first line and below the last line of each block. The two top lines and the two bottom lines of each block are, hence, immediate neighbours of computerised cooperators.

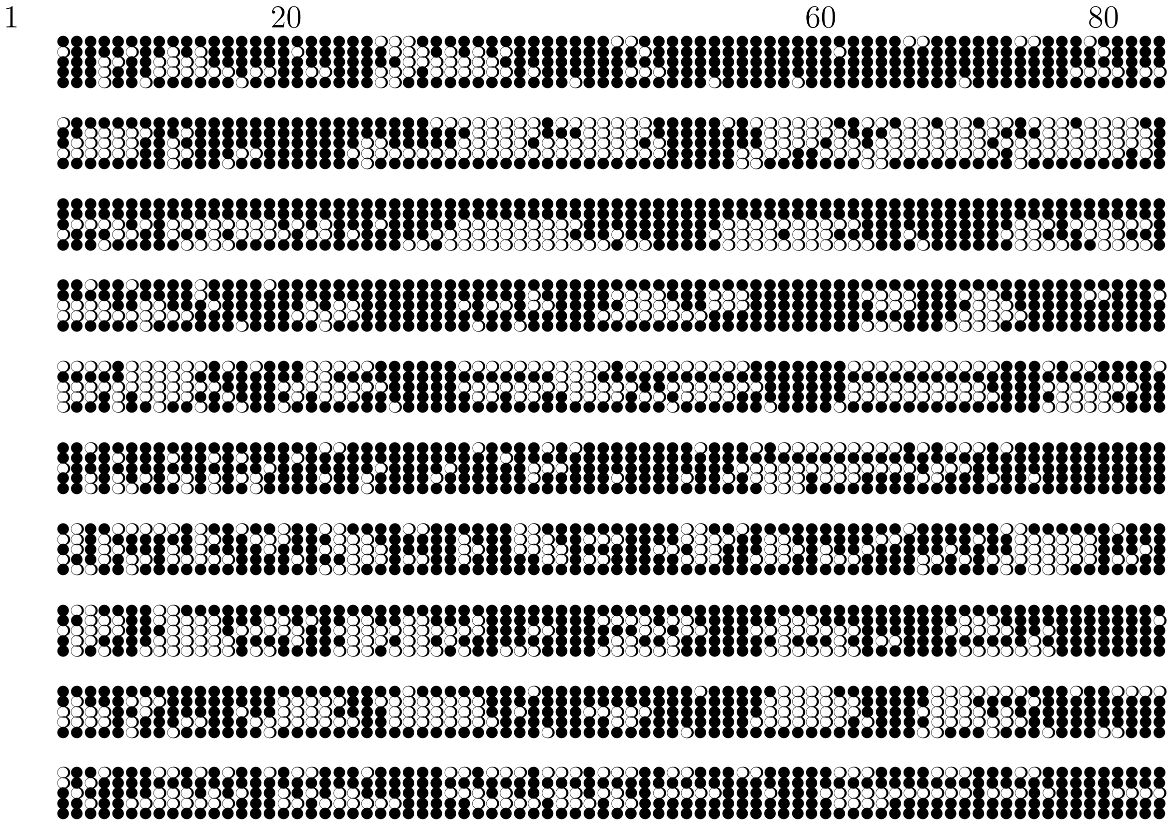


B.4 Circle treatment with detailed information





B.5 Group treatment with detailed information



C Instructions

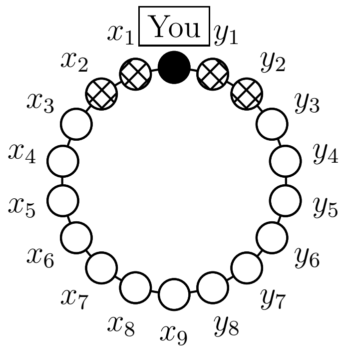
C.1 Instructions for the circle treatment without detailed information

Please sit down and read the following instructions. It is important that you read them attentively. A good understanding of the game is a prerequisite of your success.

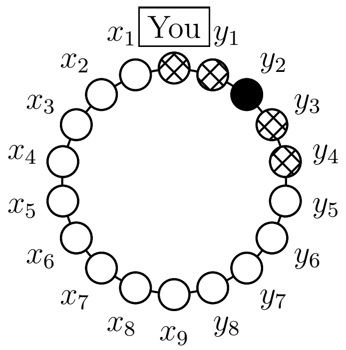
After having read the instructions you will continue with a little quiz on the computer screen. There you will be asked questions that will be easy to answer once you have read the instructions.

You may take notes but you may not talk to each other.

The structure of the neighbourhood



Your gain depends on your decision and on the decision of your two neighbours to the left and your two neighbours to the right. These four neighbours remain the same during the course of the experiment. You are connected through the computer with these neighbours. We will not tell who these neighbours are. Similarly your neighbours will not be told who you are.
In the diagram on the right side your four neighbours are shown cross-hatched.



Also your neighbours have neighbours. E.g. the neighbours of y_2 are players y_4, y_3, y_1 and you.

Rounds

In this experiment you play several rounds. In each round you take a decision. Depending on your decision and on the decision of your neighbours you receive points that will be converted to € at the end of the experiment.

Decision

In each round you choose among two decisions. You choose A or B. Your gain depends on what you have chosen and on how many of your neighbours have chosen A or B.

This relation between choices and gains is the same for all participants.

If you choose e.g. A, and all your neighbours choose B than you receive the same number of points as any other person who chooses A while the neighbours of this person all choose B. All players choose simultaneously, without knowing the decision of the others.

When all players have made their decision we continue with the next round.

Information after each round

In each round your receive information about your gain. Additionally you receive information about the decision of your neighbours and their gain.

Round	Your Decision	Your Gain	Average gain with A in your neighbourhood	Average gain with B in your neighbourhood
...	...	...	...	...

In each row you obtain information about one round. You find your decision and your gain in the second and the third column.

In the two columns to the right you find the average gain of all your neighbours who chose A and the average gain for those who chose B. The average gain is the sum of gains of all neighbours who made a decision divided into the number of these neighbours. Your own gain is included when calculating average gains.

If nobody in your neighbourhood has chosen A or B this columns will be marked with "—".

Quiz

Please answer now the questions from the quiz on the computer screen. If you are unsure how to answer a question, please consult your instructions.

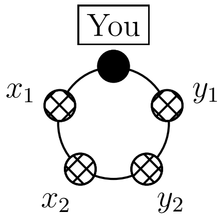
C.2 Instructions for the group treatment without detailed information

Please sit down and read the following instructions. It is important that you read them attentively. A good understanding of the game is a prerequisite of your success.

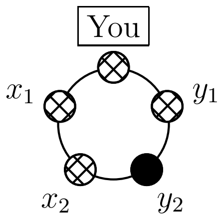
After having read the instructions you will continue with a little quiz on the computer screen. There you will be asked questions that will be easy to answer once you have read the instructions.

You may take notes but you may not talk to each other.

The structure of the neighbourhood



Your gain depends on your decision and on the decision of your two neighbours to the left and your two neighbours to the right. These four neighbours remain the same during the course of the experiment. You are connected through the computer with these neighbours. We will not tell who these neighbours are. Similarly your neighbours will not be told who you are.
In the diagram on the right side your four neighbours are shown cross-hatched.



Also your neighbours have neighbours. E.g. the neighbours of y_2 are players x_1, x_2 , y_1 and you.

Rounds

In this experiment you play several rounds. In each round you take a decision. Depending on your decision and on the decision of your neighbours you receive points that will be converted to € at the end of the experiment.

Decision

In each round you choose among two decisions. You choose A or B. Your gain depends on what you have chosen and on how many of your neighbours have chosen A or B.

This relation between choices and gains is the same for all participants.

If you choose e.g. A, and all your neighbours choose B than you receive the same number of points as any other person who chooses A while the neighbours of this person all choose B. All players choose simultaneously, without knowing the decision of the others.

When all players have made their decision we continue with the next round.

Information after each round

In each round your receive information about your gain. Additionally you receive information about the decision of your neighbours and their gain.

Round	Your Decision	Your Gain	Average gain with A in your neighbourhood	Average gain with B in your neighbourhood
...	...	...	...	...

In each row you obtain information about one round. You find your decision and your gain in the second and the third column.

In the two columns to the right you find the average gain of all your neighbours who chose A and the average gain for those who chose B. The average gain is the sum of gains of all neighbours who made a decision divided into the number of these neighbours. Your own gain is included when calculating average gains.

If nobody in your neighbourhood has chosen A or B this columns will be marked with "—".

Quiz

Please answer now the questions from the quiz on the computer screen. If you are unsure how to answer a question, please consult your instructions.

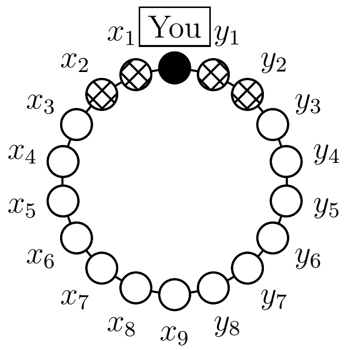
C.3 Instructions for the circle treatment with detailed information

Please sit down and read the following instructions. It is important that you read them attentively. A good understanding of the game is a prerequisite of your success.

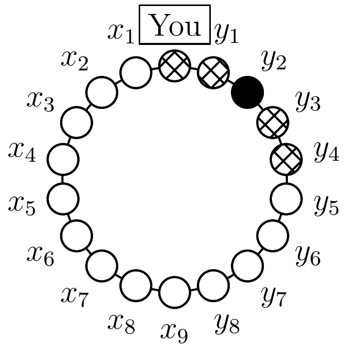
After having read the instructions you will continue with a little quiz on the computer screen. There you will be asked questions that will be easy to answer once you have read the instructions.

You may take notes but you may not talk to each other.

The structure of the neighbourhood



Your gain depends on your decision and on the decision of your two neighbours to the left and your two neighbours to the right. These four neighbours remain the same during the course of the experiment. You are connected through the computer with these neighbours. We will not tell who these neighbours are. Similarly your neighbours will not be told who you are.
In the diagram on the right side your four neighbours are shown cross-hatched.



Also your neighbours have neighbours. E.g. the neighbours of y_2 are players y_4, y_3, y_1 and you.

Rounds

In this experiment you play several rounds. In each round you take a decision. Depending on your decision and on the decision of your neighbours you receive points that will be converted to € at the end of the experiment.

Decision

In each round you choose among two decisions. You choose A or B. Your gain depends on what you have chosen and on how many of your neighbours have chosen A or B.

This relation between choices and gains is the same for all participants.

It will be shown on the screen in the form of a table.

	Your neighbours play...
You play A You play B	...Your gain ...

All players choose simultaneously, without knowing the decision of the others.

When all players have made their decision we continue with the next round.

Information after each round

In each round your receive information about your gain. Additionally you receive information about the decision of your neighbours and their gain.

Round	Your Decision	Your Gain	Decisions and gain in your neighbourhood, ordered by gain
...	...	...	...

In each row you obtain information about one round. You find your decision and your gain in the second and the third column.

On the right side we show for each of your neighbours the decision of the neighbour and the obtained gain. The ordering of neighbours in this column depends on the gain in this period. First comes the neighbour with the highest gain, then the one whose gain was second, etc.. This implies that in each period a different person can be the first in the right column.

Quiz

Please answer now the questions from the quiz on the computer screen. If you are unsure how to answer a question, please consult your instructions.

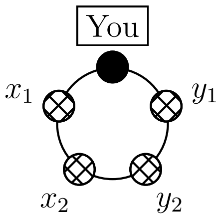
C.4 Instructions for the group treatment with detailed information

Please sit down and read the following instructions. It is important that you read them attentively. A good understanding of the game is a prerequisite of your success.

After having read the instructions you will continue with a little quiz on the computer screen. There you will be asked questions that will be easy to answer once you have read the instructions.

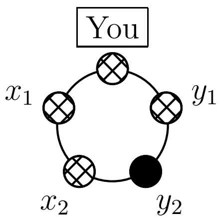
You may take notes but you may not talk to each other.

The structure of the neighbourhood



Your gain depends on your decision and on the decision of your two neighbours to the left and your two neighbours to the right. These four neighbours remain the same during the course of the experiment. You are connected through the computer with these neighbours. We will not tell who these neighbours are. Similarly your neighbours will not be told who you are.

In the diagram on the right side your four neighbours are shown cross-hatched.



Also your neighbours have neighbours. E.g. the neighbours of y_2 are players x_1, x_2 , y_1 and you.

Rounds

In this experiment you play several rounds. In each round you take a decision. Depending on your decision and on the decision of your neighbours you receive points that will be converted to € at the end of the experiment.

Decision

In each round you choose among two decisions. You choose A or B. Your gain depends on what you have chosen and on how many of your neighbours have chosen A or B.

This relation between choices and gains is the same for all participants.

It will be shown on the screen in the form of a table.

	Your neighbours play...
You play A You play B	... Your gain ...

All players choose simultaneously, without knowing the decision of the others.

When all players have made their decision we continue with the next round.

Information after each round

In each round your receive information about your gain. Additionally you receive information about the decision of your neighbours and their gain.

Round	Your Decision	Your Gain	Decisions and gain in your neighbourhood, ordered by gain
...	...	...	...

In each row you obtain information about one round. You find your decision and your gain in the second and the third column.

On the right side we show for each of your neighbours the decision of the neighbour and the obtained gain. The ordering of neighbours in this column depends on the gain in this period. First comes the neighbour with the highest gain, then the one whose gain was second, etc.. This implies that in each period a different person can be the first in the right column.

Quiz

Please answer now the questions from the quiz on the computer screen. If you are unsure how to answer a question, please consult your instructions.