Reinforcement, repeated games, and local interaction[*]

Oliver Kirchkamp[†] Rosemarie Nagel[‡]

February 6, 2002

Abstract

We investigate and compare different approaches to derive strategies from laboratory data in prisoners' dilemmas experiments. While theory suggests more cooperation in spatial structures than in spaceless ones, we find in our experiments either the opposite or no difference. In this paper we investigate to which degree learning and reinforcement explains this dependence on structure and information. Starting from a very simple model we gradually develop a setup where players use repated game strategies and choose among these strategies using a simple reinforcement rule. We then measure to which degree this model explain players' behaviour.

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

Keywords: Local interaction, experiments, prisoner's dilemma, reinforcement, repeated games.

1 Introduction

In this paper we investigate experimentally a prisoners' dilemma situation in a spatial and a spaceless model. Theoretically spatial prisoners' dilemmas have been studied by Axelrod [Axe84], Bonnhoeffer, Nowak and May [BMN93], Ellison [Ell93], Eshel, Samuelson, and Shaked [ESS98], Kirchkamp [Kir99], Lindgren and Nordahl [LN94], Nowak and May [NM92, NM93], Hegselmann [Heg94], Ely [Ely96] and several others. A brief discussion can be found in [KN00].

A large part of this literature assumes learning rules of the type "switch if better" that operate on the level of stage game strategies[1]. A smaller part of the literature studies repeated game strategies[2]. In the experimental literature only spaceless structures[3]

---

[*] We would like to express our thanks to the German DFG (through SFB 303 and SFB 504) and the Spanish DGIC Técnica PB98-1076 for supporting the project. Main ideas for this paper where conceived while the first author stayed at the Universität Bielefeld's ZiF with the research group "Making Choices". We thank Werner Güth and other members of this group for helpful comments.

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

[‡] Dep.of Economics, Universitat Pompeu Fabra; Ramon Trias Fargas, 24; E-08005 Barcelona, email: rosemarie.nagel@econ.upf.es.

[1] See [BMN93, Ell93, ESS98, NM92, NM93].

[2] See [Axe84, Kir99, LN94, Heg94].

[3] See [SMU97] and [Axe84].

are analysed with the help of repeated game strategies. Experimental studies of spatial situations[4] restrict their analysis to only stage game strategies. In the current paper we compare spatial with non-spatial behaviour, allowing for repeated game strategies.

Theoretically, “switch if better” is a compelling rule. Players compare past average payoffs of their own and of other visible players for available strategies and choose the most successful strategy in the future.

In the context of prisoners’ dilemmas and without spatial structure this behaviour eliminates cooperation. If everybody plays against everybody else, non cooperation is always the most successful strategy, and will hence be imitated by all players.

With a spatial structure, however, cooperative behaviour may survive if players follow a “switch if better” rule. Clusters of cooperative players obtain higher payoffs than clusters of non cooperative players. Since successful clusters of cooperative players are in particular visible in the vicinity of these clusters, clusters may even grow at their borders, and, since cooperation grows predominantly at the borders of cooperative clusters, clusters remain intact and successful.

Experimentally, however, this compelling theoretical property can not be replicated. Figure 1 shows for illustration the frequences of cooperation in different conditions (spatial structure, information) that we are going to describe in section 2. At this stage it is sufficient to know that “groups” have no spatial structure wheras “circles” have one. In contrast to the the theoretical prediction we find more cooperation (and not less) in the spaceless structure (group) than in the spatial one (circle), at least in the “info” condition. We find roughly the same amount of cooperation in the “no info” condition.

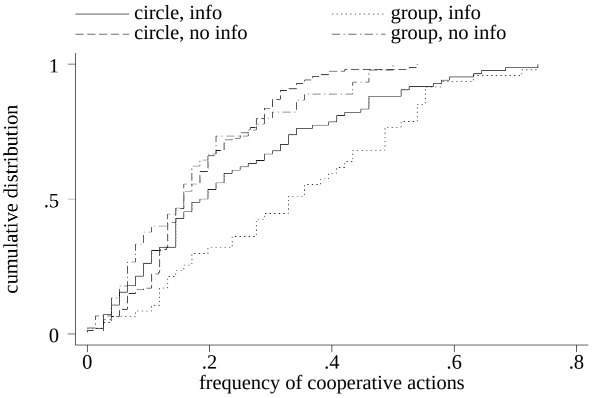
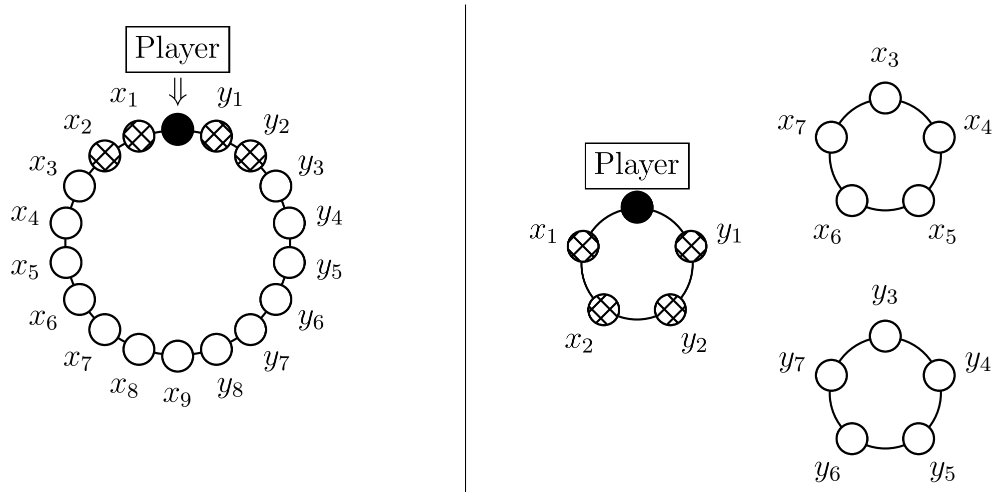


Figure 1: Frequency of cooperative actions per player

To better understand why in the experiment we get the opposite of the theoretical prediction we consider in this paper repeated game strategies that will be simple and of the following type: Cooperate if the difference between the observed payoffs between cooperation and non-cooperation is larger than a certain threshold. The threshold may be different for each player and may or may not change over time. This approach contains as

---

[4] See [KEB97, KEB98, KN00].



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

a special case the "copy if better" rule that is used throughout the theoretical literature.[5] We then measure to which degree the behaviour of players in our experiment can be explained as governed by repeated game strategies that were successful in the past, and to which degree players rely on comparison of stage game payoffs.

We will describe the experimental setup in section 2. We then introduce repeated game strategies to allow players to condition on past behaviour of their opponents. We start with a simple version with constant repeated game strategies for each player in section 3. We introduce a more elaborate model in section 4 where repeated game strategies may change over time. We relate changes in repeated game strategies to payoffs using a simple reinforcement approach in section 5. Section 6 concludes.

2 The experimental setup

In the following we outline our experimental setup. A more detailed discussion is given in [KN00]. Experiments were conducted in computerised laboratories at either Universitat Pompeu Fabra in Barcelona or at the Universität Mannheim.

We compare two structures, one that we will call 'circle', the other we will call 'group'. Circles model local interaction, groups model spaceless interaction. The structures are shown in figure 2. In each period players interact with two neighbours to the left (x_1, x_2) and two neighbours to the right (y_1, y_2). In our experiment players know that they repeatedly interact for 80 periods with the same neighbours. In each round each player

---

[5] See Axelrod [Axe84], Bonnhoeffer, Nowak and May [BMN93], Ellison [Ell93], Eshel, Samuelson, and Shaked [ESS98], Kirchkamp [Kir00], Lindgren and Nordahl [LN94], Nowak and May [NM92, NM93], Hegselmann [Heg94], Ely [Ely96].

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

has two choices: C or D.[6] Payoffs are a function of the player's own choice as well as the number of neighbours who chose C. The relation is shown in table 1.

Players also obtain information about payoffs and strategies of their neighbours during each round. We study two conditions: In one condition players obtain only information about the average payoff of the two strategies in the last round. We will call this the "no detailed information" condition. In another condition players know the payoffs of the underlying game and the actual distribution of payoffs for the two strategies in their neighbourhood. We will call this the "detailed information" condition. Even in the information condition information about payoffs of players was ordered by payoffs in each round. Thus, players only knew how their neighbourhood as a whole performed, they could not identify patterns in actions or payoffs of particular players nor could they easily infer actions or payoffs outside their neighbourhood.

In all conditions players were told that they played against the other players in the room, but that computers were networked randomly so that they never knew the identity of their neighbours in the game. After the experiment players were payed seperately and obtained between 5 and 15 Euros for an experiment that lasted for about one hour. Altogether we did 36 experiments.

3 A model with constant thresholds

Let us first assume that players follow a simple and constant repeated game strategy. Players cooperate if and only if the difference between the payoff of the two strategies C and D in the last ν periods was greater than a certain threshold τ_ν.[7] Given the payoffs of our game (see table 1) the difference $u_D - u_C$ between the two strategies must be

---

[6] A game theorist might argue that we could have obtained more information had we asked participants only for one repeated game strategy for each repeated game. This argument presupposes that the submitted repeated game strategies would also explain the players' actions if the players could choose stage game strategies on a period to period basis. However, this is only true for perfectly rational players — and not for real participants of our experiment. One of the results of this paper is that players in the experiment seem indeed to change their repeated game strategies while playing a single instance of the repeated game.

[7] The reader should note that this approach weights experience from all past ν periods equally. Alternatively one could use discounting of past experience. Our approach seems, however, sufficient to show that only the recent past ($\nu = 1$) has a substantial impact.

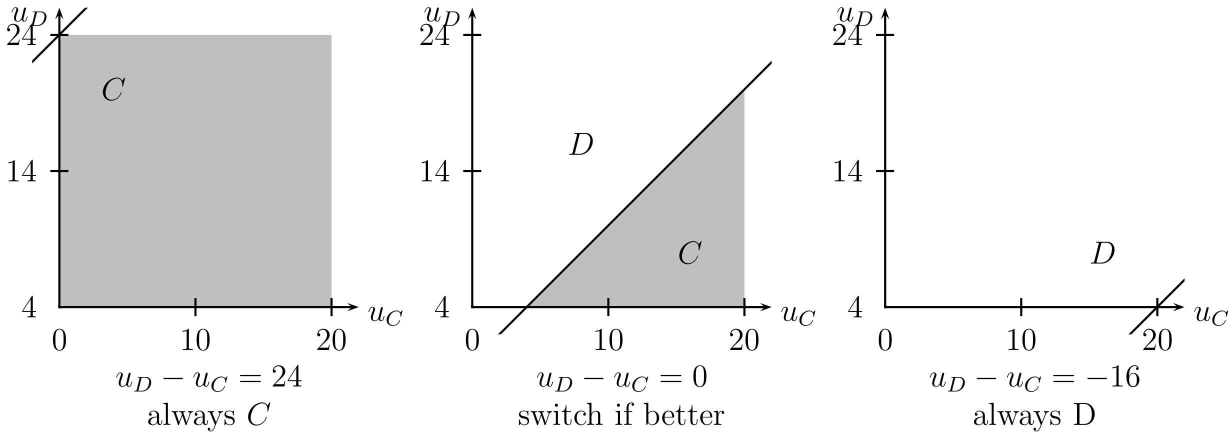
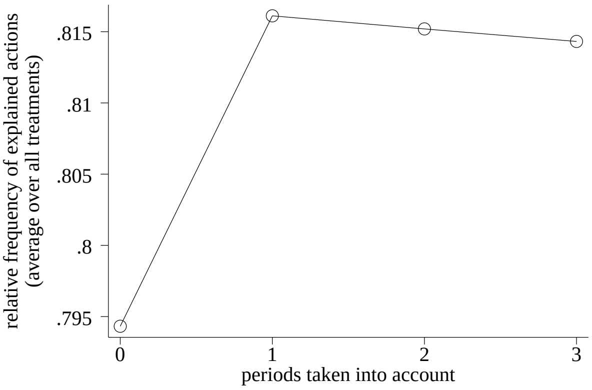


Figure 3: Threshold strategies

somewhere between -16ν (D obtains only 4 while C obtains 20) and 24ν (D obtains 24 while C gets only 0).

A player with $\tau_\nu = 24\nu$ always cooperates, a player with $\tau_\nu = 0$ always imitates the strategy with the higher payoff, and a player with $\tau_\nu = -16\nu$ never cooperates. Intermediate values of τ_ν yield an intermediate behaviour. Figure 3 illustrates this relationship. For each player separately we determine the threshold value τ_ν that maximises the number of correctly explained actions for all 80 periods. If there is no unique such value we take one randomly from the set of maximising values. We do this separately for time-spans (ν) between 0 and 3. Figure 4 shows the relation between ν and the relative frequency of correctly explained actions. With $\nu = 0$ predictions assume a very simple strategy, players either always play C or they always play D. Actually the behaviour of most players (93.01%) is best approximated with all D. This simple model explains 79.4% of all actions. Introducing information about a single previous period ($\nu = 1$) increases

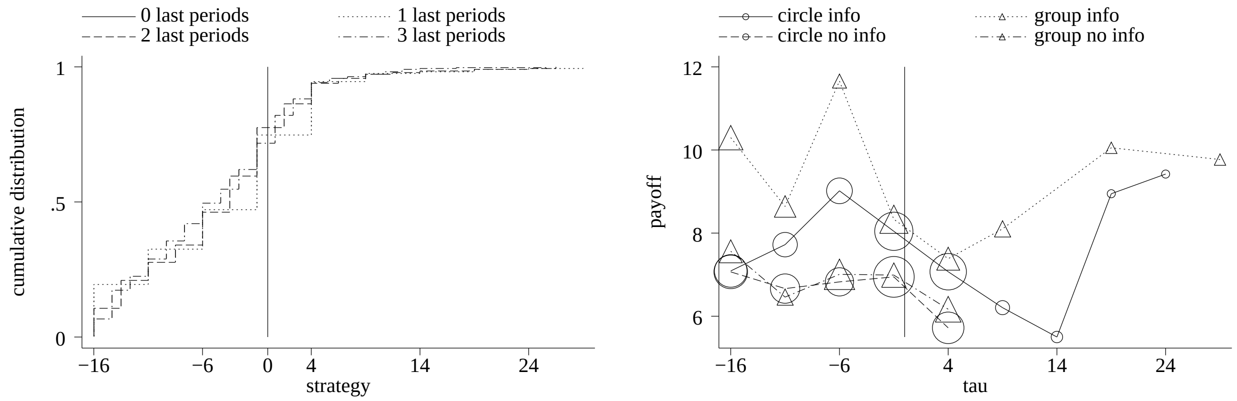


All past periods are equally weighted.

Figure 4: Relative frequency of correctly explained actions

the number of correctly predicted actions to 81.5%.[8] Introducing more periods ($\nu = 2$ or $\nu = 3$) does no longer improve the number of correctly predicted actions.[9] Apparently only one previous period has a substantial impact. Introducing more and irrelevant periods deteriorates the quality of the prediction. We will therefore restrict ourselves in the following to the case $\nu = 1$.

The left part of figure 5 shows the distribution of the threshold level τ_ν under the assumption that τ is constant for each player. Two things are worth noting: One is that the distribution of strategies τ_ν/ν does not seem to depend very much on ν as long as $\nu \geq 1$. The second, and more interesting, is that players play D significantly[10] more often than a "copy if better" strategy would recommend. Figure 6 show the distributions for each condition (circle/groups, detailed information/no detailed information) seperately. We see that the above finding does not depend on the condition.



distribution of τ for different values of ν average payoff for different τ_1

The horizontal axis shows players' strategies, normalised as τ_ν/ν.

Figure 5: Threshold levels and payoffs

Figure 7 shows how players' behaviour over time becomes increasingly consistent with this simple model. Some of the remaining unpredicted actions may be explained as experiments, but others may better be explained through a repeated game strategy that changes over time. We will therefore allow for changes of repeated game strategies over time in the next section.

4 Changing Repeated Game Strategies

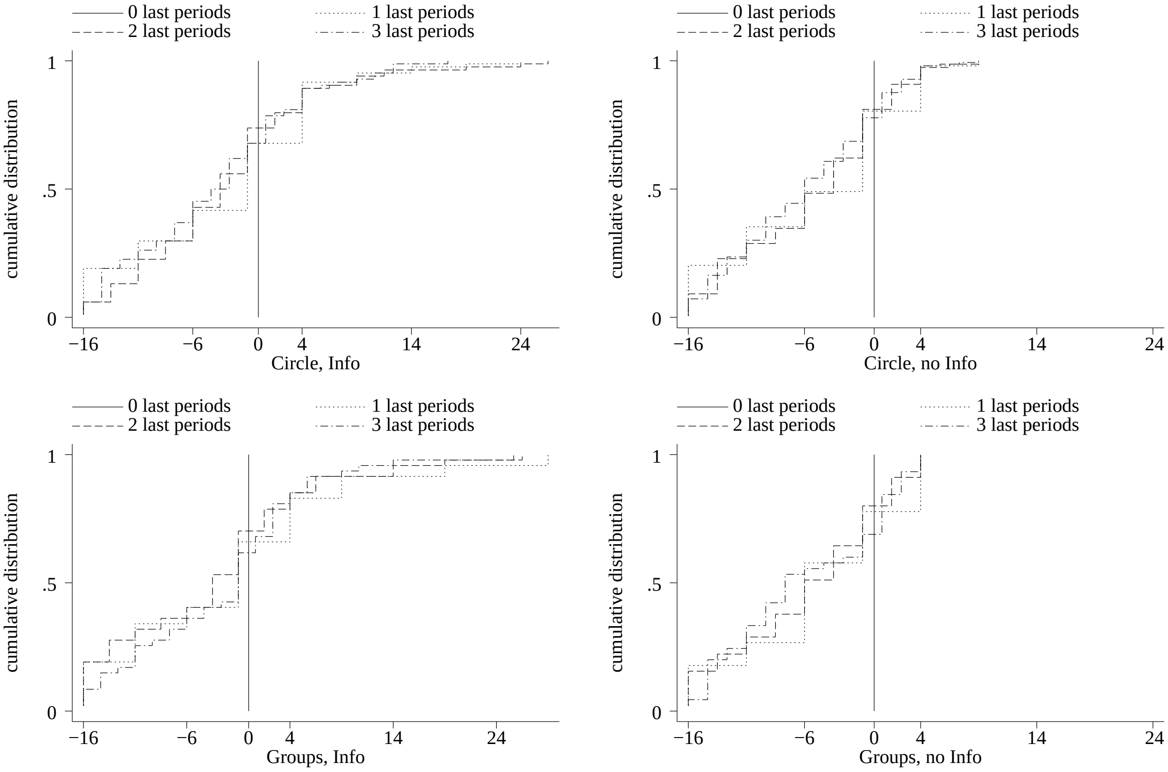
In our experiment we only observe whether a player plays C or D. The observed behaviour is consistent with several repeated game strategies. In the previous section we have

---

[8] A Wilcoxon matched-pairs signed-ranksfinds this difference to be significant ($z = 5.152, P_{>|z|} = 0.0000$).

[9] A Wilcoxon matched-pairs signed-ranks finds neither the decrease from $\nu = 1$ to $\nu = 2$ to be significant ($z = -0.476, P_{>|z|} = 0.6341$), nor the decrease from $\nu = 2$ to $\nu = 3$ ($z = -1.124, P_{>|z|} = 0.2612$)

[10] A Wilcoxon signed-rank test for $\tau_1 = 0$ yields $z = -3.892$ and $P_{>|z|} = 0.0001$



The horizontal axis shows players' strategies, normalised as τ_ν/ν.

Figure 6: Threshold levels, constant for each player, for different conditions

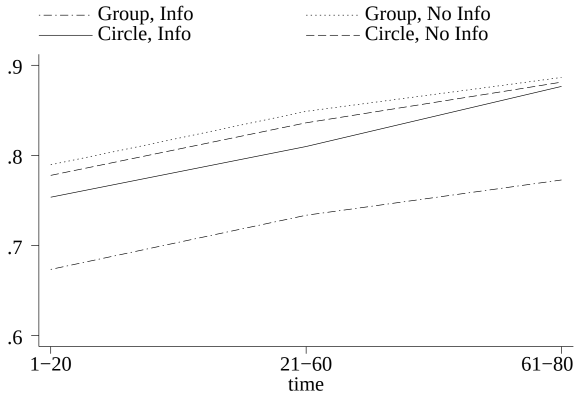


Figure 7: The number of false predictions decreases over time

introduced simple repeated game strategies that were based on threshold values τ. We have found that a fixed value of τ does not explain players' behaviour in a satisfactory way. In the current section we will allow for changes in the value of τ over time. Doing this, we have to find a solution for the problem that in each period several thresholds τ explain a player's behaviour. E.g. if the difference $u_D - u_C$ is 4 and our player plays C then all values of $\tau \in [4, 24]$ are consistent with the observed behaviour. To further identify τ we require that τ changes as little as possible. In other words, if there is a τ that explains the behaviour of a player not only at time t but also at time $t-1$ or $t+1$ we will favour this τ over another one that only explains the behaviour at time t.

Here is an example:

Period	...	$t-1$	t	$t+1$	...
Action	...	D	C	D	...
$u_D - u_C$ in previous period	...	9	4	4	...
interval of possible τ	...	$[-16, 9]$	$[4, 24]$	$[-16, 4]$	...

The example player chooses D in period $t-1$. In the previous period $u_D - u_C$ was 9. We assume that this player would also play D if $u_D - u_C$ is larger, but we do not know what this player would do for smaller values of $u_D - u_C$. So far we can restrict the range of possible values for τ to $[-16, 9]$. In the next period we see that upon a $u_D - u_C = 4$ the player chooses C. This reduces the range of possible values for τ to $[4, 9]$. In the following period τ is restricted to the unique value 4.

In this example only the value $\tau = 4$ explains all observations around t. This, however, is a lucky coincidence. With our data typically three subsequent periods do not allow to reduce the range for τ to a single value. We have to take into account more periods to determine a unique value for τ.

More formally we repeatedly apply the following algorithm:

Be $\delta(t)$ the payoff difference $u_D - u_C$ in period $t-1$. Be $I_0(t)$ the range of possible τs that is compatible with a players action in this period:

$$I_0(t) = \begin{cases} [\delta(t), u_{\max}] & \text{if the player plays } C \text{ in period } t \\ [u_{\min}, \delta(t)] & \text{if the player plays } D \text{ in period } t \end{cases} \tag{1}$$

where $u_{\max}$ and $u_{\min}$ are the maximal and minimal payoffs respectively that can be achived in the experiment. Notice that these intervals are never empty.

We distinguish the following conditions:

$$\begin{array}{rcl} a & : & I_k(t-1) \cap I_k(t) \cap I_k(t+1) \neq \emptyset \\ b & : & I_k(t-1) \cap I_k(t) \neq \emptyset \\ c & : & I_k(t) \cap I_k(t+1) \neq \emptyset \\ d & : & \min(I_k(t-1)) > \max(I_k(t)) \\ e & : & \max(I_k(t-1)) < \min(I_k(t)) \\ f & : & \min(I_k(t+1)) > \max(I_k(t)) \\ g & : & \max(I_k(t+1)) < \min(I_k(t)) \end{array}$$

We now iteratively reduce the size of the intervals using the following method:

$$
I_{k+1} = \begin{cases} I_k(t-1) \cap I_k(t) \cap I_k(t+1) & \text{if } a \\ I_k(t-1) \cap I_k(t) & \text{if } \neg a \wedge b \\ I_k(t) \cap I_k(t+1) & \text{if } \neg(a \vee b) \wedge c \\ \max(I_k(t)) & \text{if } \neg(a \vee b \vee c) \wedge d \\ \min(I_k(t)) & \text{if } \neg(a \vee b \vee c \vee d) \wedge e \\ \max(I_k(t)) & \text{if } \neg(a \vee b \vee c \vee d \vee e) \wedge f \\ \min(I_k(t)) & \text{if } \neg(a \vee b \vee c \vee d \vee e \vee f) \wedge g \\ I_k(t) & \text{otherwise} \end{cases} \tag{2}
$$

Before we discuss these conditions in more details, we should note two things:

- Once an interval consists of a singleton it will never change through repeated application of the above algorithm.
- Intervals can only become smaller, never larger. Formally $\forall_{j>k} I_j(t) \subseteq I_k(t)$. I.e. we never add something to a strategy of a player, we only make it more precise. The resulting strategy will always be compatible with what we have observed.

Condition a is the simplest and most frequent case: The ranges for τ in three subsequent periods are consistent and allow for one or possibly more values of τ. In this case we take the intersection of these ranges.

If such an intersection would be empty we try to find only two subsequent periods. We first look more into the past (b) and then more into the future (c).

If this fails as well, then neighbouring ranges for τ do not intersect at all. In our interpretation this means that we have detected a change in the conditional strategy of the player. We then assume some inertia and shrink the interval for τ_t into the direction of the neighbouring interval. We do this first for $t-1$ (conditions d and e) and then for $t+1$ (conditions f and g).

When for all players in the experiment and for all periods $I_{k+1} = I_k$ then we have reached a fixed point of the process. We will call these intervals I^*. Notice that with a finite number of observations the process always reaches a fixed point in a finite number of steps.

Will this process converge to only singletons? It is possible to show that if there is some randomness in players' behaviour which is not perfectly correlated with the behaviour of the neighbours then the probability to obtain a unique τ grows arbitrarily close to 1 when the number of observations per player (number of periods in our experiment) is only large enough[11].

Since we have a finite number of observations in our experiment we only obtain a unique τ for 99.4% of all players and periods. We dropped the remaining 0.6%.[12]

Let us start here with some summary statistics. Figure 8 shows the development of the

---

[11] To see this, one has to show that if $I_k(t)$ is not a singleton then $I_{k+l}(t)$ will be a singleton if only we find a t' such that $I_k(t) \cap I_k(t')$ is a singleton. In this case $l \leq |t'-t|$, i.e. the above process will converge to a singleton in at most $|t'-t|$ steps. To ascertain the existence of such a t' we need the assumptions of randomness in players' behaviour together with a large enough number of observations.

[12] These 0.6% are two players that consistently played D in a surprisingly cooperative neighbourhood. Instead of dropping these observations we could have replaced these observations with any (constant) element of the interval of possible τs, without affecting our results.

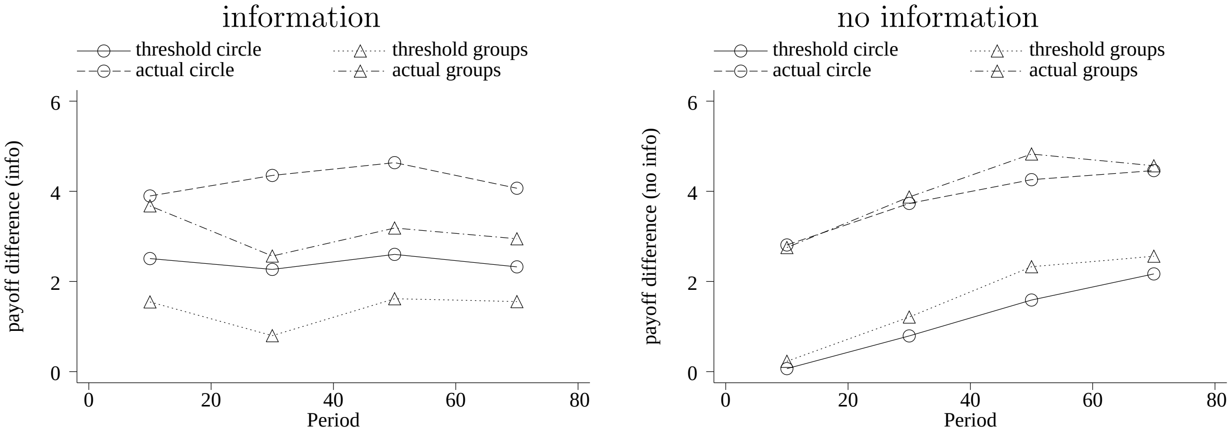


Figure 8: Threshold level and actual strategies ν

threshold level τ over time in groups and in circles separately, together with the average level of cooperation. We see that in all conditions the average threshold is lower than the actual difference $u_D - u_C$ which is in line with the small amount of cooperation (see figure 1).

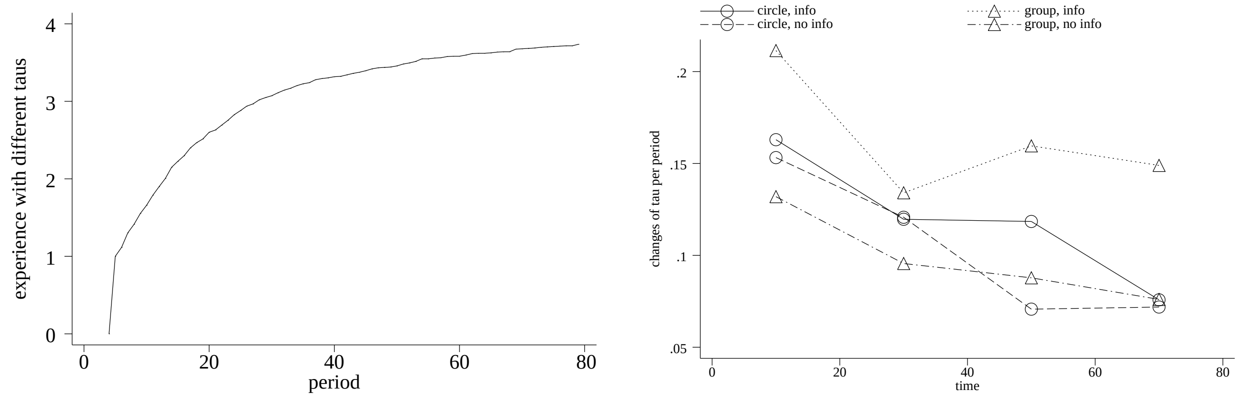
Having calculated for each player and for each period a threshold value we can explain actions through thresholds, but what can we say about how players choose and change their thresholds and if so, why?

5 A simple reinforcement model

We assume that each player in each period associates with each possible value of τ a discounted average payoff of this strategy. Reinforcement (see Erev and Roth [ER98]) suggests that players are more likely to switch to a strategy that was successful in the past.

The left part of figure 9 shows for each period the average number of different threshold levels players have experienced up to this period, the right part shows the frequency of changes in τ per period. We see that soon the average player has experience with at least three different repeated game strategies. This is less than the maximal number of strategies, but allows us to explain his choices with the help of comparisons of payoffs. To do that we concentrate on the situation when a player switches from one repeated game strategy (the 'source' strategy) to another (the 'target' strategy). The right part of figure 9 shows the frequency of such changes. While there are more changes at the beginning of the experiment players keep changing strategies until the end. What are the reasons for these changes?

One hypothesis could be that players choose different strategies based on their past success with these strategies. To investigate this hypothesis we calculate for each player,



average number of experienced τs per player changes of τ per period

Figure 9: Learning τ

period and for each repeated game strategy the discounted[13] payoff while using this strategy up to this period. When a player changes his or her strategy we call δ the difference between the discounted past payoff of the target strategy and the source strategy. The relative frequencies of δ are shown in figure 10. The diagram also includes a normal distribution as a reference.

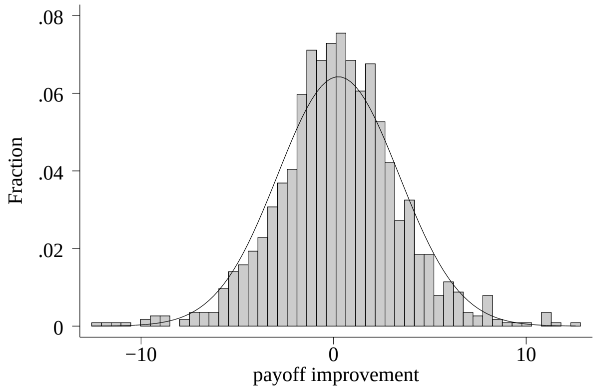


Figure 10: Difference between target and source payoff when switching strategies

Notwithstanding the variance of the distribution we see that reinforcement seems to play a role. Players rather switch to strategies that gave higher payoffs in the past. A t-test reveals that the mean of δ is significantly positive.[14]

Although players do not always choose a payoff improving repeated game strategy we will in the following section assume that they do. How much can we explain with this simple assumption. To answer this question we determine for each player and period the repeated game strategy with the highest past discounted payoff. This gives us for each

---

[13] As a discount factor we use 0.9.

[14] $t = 2.56, P_{>t} = 0.008$, the test takes into account that observations within a session may be correlated.

player and period a prediction of a stage game strategy, derived from an endogeneously determined repeated game strategy. We will call this prediction $\hat{c}_t$.

As an alternative we also calculate the hypothetical behaviour of a player who only follows a “copy if better” rule as assumed often in the literature.[15] This hypothetical behaviour gives us another prediction that we call $\bar{c}_t$.

We further allow that observed payoffs in the neighbourhood play a role. Similar to $\hat{c}_t$ and $\bar{c}_t$ we construct $\hat{c}^n_t$ and $\bar{c}^n_t$ as predictions based not on own but on average neighbours' payoff.

Finally allow for some inertia and introduce yesterday's action c_{t-1} as an explanatory variable for today's action c_t.

If players' behaviour would be consistently explained by the “copy if better” learning rule[16] only $\bar{c}_t$ and $\bar{c}^n_t$ should play a role.

Having constructed these variables we run for each condition a probit regression that explains the actual behaviour c_t as a function of c^r_t and c^n_t.

We see that player's behaviour is not only explained by “switch if better”. The coefficient `cpn` is positive and always significant. So players certainly follow their own payoffs in a way that is consistent with “switch if better”. The coefficient `cpnu`, however, is not always positive, and in the detailed information condition not significant. Should we, therefore, better explain players' behaviour as governed by a “switch if better” rule that only takes into account the players' own payoff and ignores the observed neighbours' payoff? Possibly not, since other variables also contribute significantly to explaining player's behaviour. First of all, inertia, modelled as last period's action, plays an important role and explains between 15% (circle, no detailed information) and 48% (group, detailed information) of actions.

Reinforcement of repeated game strategies also has a substantial impact. The coefficient `cpr` is always positive and significant in three out of four conditions where it contributes between 8% to 10% of actions. Only groups with no detailed information stick out.

The payoff of observed neighbours, however, does not play a great role. The coefficient `cpru` is not significant in three out of four cases and always has the wrong sign. The magnitude of the effect is small, however. Between 2% and 4% of the action is affected by `cpru`.

To summarise, we find that the simple repeated game straegies that we introduced in section 3 explain a great deal of players' behaviour. Applying a “switch if better” rule alone to these strategies, however, explains only part of players' behaviour. A simple reinforcement rule also helps us predicting what actions players choose. In both cases predominantly the players' own payoff is important. Observed neighbours' payoff is basically neglected.

---

[15] See footnote 5.

[16] see footnote 5.

group, no detailed information

```
Probit estimates                                  Number of obs =   3420
                                                  Wald chi2(5)  = 260.46
                                                  Prob > chi2   = 0.0000
Log likelihood = -1397.9574                       Pseudo R2     = 0.1198

                         (standard errors adjusted for clustering on group_id)
------------------------------------------------------------------------------
         |               Robust
      c1 |      dF/dx   Std. Err.      z    P>|z|     x-bar  [    95% C.I.   ]
---------+--------------------------------------------------------------------
     cpr |   .0030687   .0219555     0.14   0.889   .484795  -.039963  .046101
    ucpr |  -.037927    .0229514    -1.67   0.095   .589181  -.082911  .007057
    cpn*|    .1303004   .0307688     5.01   0.000    .25614   .069995  .190606
   cpnu*|    .0470498   .0291982     1.72   0.086   .073099  -.010178  .104277
      c*|    .3023972   .0367176    11.21   0.000    .17924   .230432  .374362
---------+--------------------------------------------------------------------
  obs. P |   .1754386
 pred. P |    .152746  (at x-bar)
------------------------------------------------------------------------------
(*) dF/dx is for discrete change of dummy variable from 0 to 1
    z and P>|z| are the test of the underlying coefficient being 0
```

circle, no detailed information

```
Probit estimates                                  Number of obs =   6232
                                                  Wald chi2(3)  =      .
                                                  Prob > chi2   =      .
Log likelihood = -2588.0448                       Pseudo R2     = 0.0811

                         (standard errors adjusted for clustering on group_id)
------------------------------------------------------------------------------
         |               Robust
      c1 |      dF/dx   Std. Err.      z    P>|z|     x-bar  [    95% C.I.   ]
---------+--------------------------------------------------------------------
     cpr |   .0820771   .0055242    12.55   0.000   .447288    .07125  .092904
    ucpr |  -.0263002    .008641    -3.05   0.002   .585286  -.043236 -.009364
    cpn*|    .1359608   .0291893     5.85   0.000   .269576   .078751  .193171
   cpnu*|    .0618644   .0241181     2.59   0.010   .088254   .014594  .109135
      c*|    .1483834   .0073327    29.84   0.000   .167843   .134012  .162755
---------+--------------------------------------------------------------------
  obs. P |   .1675225
 pred. P |   .1500021  (at x-bar)
------------------------------------------------------------------------------
(*) dF/dx is for discrete change of dummy variable from 0 to 1
    z and P>|z| are the test of the underlying coefficient being 0
```

group, detailed information

```
Probit estimates                                  Number of obs =   3572
                                                  Wald chi2(5)  = 519.99
                                                  Prob > chi2   = 0.0000
Log likelihood =  -1812.997                       Pseudo R2     = 0.2027

                         (standard errors adjusted for clustering on group_id)
------------------------------------------------------------------------------
         |               Robust
      c1 |      dF/dx   Std. Err.      z    P>|z|     x-bar  [    95% C.I.   ]
---------+--------------------------------------------------------------------
     cpr |   .1037308   .0292229     3.54   0.000   .573908   .046455  .161007
    ucpr |  -.027952    .0336965    -0.82   0.411   .631159  -.093996  .038092
    cpn*|    .0699172   .0337136     2.06   0.040     .3729    .00384  .135995
   cpnu*|   -.0610652    .053192    -1.09   0.275   .088186   -.16532  .043189
      c*|    .4787342   .0330236    14.49   0.000   .337626   .414009  .543459
---------+--------------------------------------------------------------------
  obs. P |   .3334267
 pred. P |   .3065019  (at x-bar)
------------------------------------------------------------------------------
(*) dF/dx is for discrete change of dummy variable from 0 to 1
    z and P>|z| are the test of the underlying coefficient being 0
```

circle, detailed information

```
Probit estimates                                  Number of obs =   6384
                                                  Wald chi2(3)  =      .
                                                  Prob > chi2   =      .
Log likelihood = -2952.1436                       Pseudo R2     = 0.1567

                         (standard errors adjusted for clustering on group_id)
------------------------------------------------------------------------------
         |               Robust
      c1 |      dF/dx   Std. Err.      z    P>|z|     x-bar  [    95% C.I.   ]
---------+--------------------------------------------------------------------
     cpr |   .0745789   .0335198     2.23   0.026   .500392   .008881  .140277
    ucpr |  -.0194182   .0299041    -0.65   0.514   .499373  -.078029  .039193
    cpn*|    .1145471   .0092291    13.27   0.000   .242794   .096459  .132636
   cpnu*|    .0426194   .0274635     1.62   0.105   .119518  -.011208  .096447
      c*|    .3781401   .0204581    19.43   0.000   .241071   .338043  .418237
---------+--------------------------------------------------------------------
  obs. P |   .2376253
 pred. P |   .2104257  (at x-bar)
------------------------------------------------------------------------------
(*) dF/dx is for discrete change of dummy variable from 0 to 1
    z and P>|z| are the test of the underlying coefficient being 0
```

The cofficients `cpr` and `cpru` stand for the predictions of the reinforcement model $\hat{c}_t$ and $\hat{c}_t^n$, the coefficients `cpn` and `cpnu` are the predictions of the "switch if better rule" $\bar{c}_t$ and $\bar{c}_t^n$. The coefficient `c` stands for last period's action and, thus, captures inertia.

Table 2: Predicting actions

6 Conclusion

In the above paper we model players' repeated game strategies with the concept of a threshold value for cooperation. The threshold is defined by the number of cooperating neighbours needed in order for a player to cooperate.

In a first step we study a model with constant thresholds. Such a model has more degrees of freedom than a (constant) stage game strategy based model and, hence, can explain more observations. We find, however, that players' behaviour can better be explained when the threshold is allowed to depend on the number of cooperative neighbours or payoffs.

We then study a simple reinforcement model and find that the repeated game strategies that we identified and that players experienced to be successful in the past are indeed more likely to be played. We observe that players change their threshold more rapidly in a local interaction structure than in a spaceless interaction structure. As a consequence a decrease of cooperation by neighbours follows an increase of threshold which leads to less cooperation on the circle than in the groups.

We then explain players' behaviour with the help of five components. One is inertia, the remaining four are repeated game strategies, driven by own payoff or observed neighbours' payoff, either following a "switch if better" strategy or following a simple reinforcement rule. We find that observed neighbours' payoff does not contribute much to a player's action, neither through a "switch if better" rule nor through a rule that is selected by reinforcement. We, hence, cast doubt on using imitation as a supporting element of cooperation in space.[17]

Player's own payoff, however, is an important factor in determining behaviour. "Switch if better"[18] certainly plays a role, but is not alone responsible for what a player does. The role of reinforcement is in the games and structures that we are studying, of about the same size.

References

[Axe84] Robert Axelrod. *The evolution of cooperation.* Basic Books, New York, 1984.

[BMN93] Sebastian Bonhoeffer, Robert M. May, and Martin A. Nowak. More spatial games. *International Journal of Bifurcation and Chaos*, 4:33–56, 1993.

[Ell93] Glenn Ellison. Learning, local interaction, and coordination. *Econometrica*, 61:1047–1071, 1993.

[Ely96] Jeffrey Christopher Ely. *Strategic Demography: The Impact of Local Interaction and Mobility on the Stability of Evolutionary Game-Theoretic Equilibria.* PhD thesis, University of California, 1996.

---

[17] See footnote 5.

[18] See footnote 1

[ER98] Ido Erev and Alvin E. Roth. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, 88(4):848–81, September 1998.

[ESS98] Ilan Eshel, Larry Samuelson, and Avner Shaked. Altruists, egoists, and hooligans in a local interaction model. *The American Economic Review*, 88:157–179, 1998.

[Heg94] Rainer Hegselmann. Zur Selbstorganisation von Solidarnetzwerken unter Ungleichen. In Karl Homann, editor, *Wirtschaftsethische Perspektiven I*, number 228/I in Schriften des Vereins für Socialpolitik, Gesellschaft für Wirtschafts- und Sozialwissenschaften, Neue Folge, pages 105–129. Duncker & Humblot, Berlin, 1994.

[KEB97] Claudia Keser, Karl-Martin Erhart, and Siegfried K. Berninghaus. Coordination games: Recent experimental results. Working Paper 97-29, SFB 504, Universität Mannheim, 1997.

[KEB98] Claudia Keser, Karl-Martin Ehrhart, and Siegfried K. Berninghaus. Coordination and local interaction: Experimental evidence. *Economics Letters*, 58(3):269–75, March 1998.

[Kir99] Oliver Kirchkamp. Simultaneous evolution of learning rules and strategies. *Journal of Economic Behavior and Organization*, 40(3):295–312, 1999. http://www.kirchkamp.de/.

[Kir00] Oliver Kirchkamp. Spatial evolution of automata in the prisoners' dilemma. *Journal of Economic Behavior and Organization*, 43(2):239–262, 2000. http://www.kirchkamp.de/.

[KN00] Oliver Kirchkamp and Rosemarie Nagel. Local and group interaction in prisoners' dilemma experiments — imitate locally, think globally. Discussion Paper 00-11, SFB 504, Universität Mannheim, 2000. http://www.kirchkamp.de/.

[LN94] Kristian Lindgreen and Mats G. Nordahl. Evolutionary dynamics of spatial games. *Physica D*, 75:292–309, 1994.

[NM92] Martin A. Nowak and Robert M. May. Evolutionary games and spatial chaos. *Nature*, 359:826–829, 1992.

[NM93] Martin A. Nowak and Robert M. May. The spatial dilemmas of evolution. *International Journal of Bifurcation and Chaos*, 3:35–78, 1993.

[SMU97] Reinhard Selten, Michael Mitzkewitz, and Gerald R. Uhlich. Duopoly strategies programmed by experienced players. *Econometrica*, 65(3):517–555, May 1997.