Exclusion and Cooperation in Social Network Experiments*

Arno Riedl[‡] and Aljaž Ule[§]

Preliminary version: February, 2002

Abstract

This study examines the evolution of social networks within groups of six subjects playing a prisoners dilemma game. In each round subjects cannot discriminate in their action choice but can exclude others from their social environment. Four treatments are considered, varying the cost of exclusion and the information flow through the network. Observed cooperation levels strongly depend on the treatment conditions but are always significantly higher, reaching 93 percent, compared with the control treatment with predetermined network structure. In addition, the evolution of social structure is analyzed in parallel to the dynamics of action choices. It appears that when cooperators stay connected in cliques separated from defectors, cooperation remains high until the last rounds.

JEL Classification Number: C72, C91, C92, D83, H41
Keywords: Networks, cooperation, Prisoner's Dilemma, partner selection, social exclusion, information, experiments

*This is a very preliminary version prepared for the ESEM 2002. Please do not cite or quote withouit permission of the authors. This paper is part of the EU-TMR research network ENDEAR (FMRX-CT98-0238) and the research project on "Strategic Bargaining and Coalition Formation" financed by the Oesterreichische Nationalbank (project number: 6933). We are very indebted to Jos Theelen who wrote the experimental software. As always the critical comments and remarks of our colleagues at CREED have been very helpful in designing the experiment. The usual disclaimer applies.

[‡]CREED, Faculty of Economics and Econometrics, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands.

[‡]Tinbergen Institute and CREED, Faculty of Economics and Econometrics, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands.

1 Introduction

This paper experimentally investigates the dynamics of cooperation between agents structured in endogenously developed social networks. Both the endogenous choice of partners and the structure of the social network have been previously proposed as determinants for the levels of cooperation in social dilemmas. However related these two factors may be, the focus of recent research has been to consider them in isolation. Freedom to refuse participation in the social dilemma appeared to improve cooperation but the evidence on the effect of the social structure remains mixed and scarce.

In this paper we combine both approaches and consider the simultaneous evolution of social network and cooperation. The motivation for our approach is our belief that social structure is a consequence of behavior rather than its determinant. People tend to manipulate social relationships on the basis of observed past behavior; they form social ties with other cooperators and exclude those that cheat. As the number of cooperators increases, the threat of social exclusion gains credibility and even the most rigid defectors switch to cooperation. The dynamics of behavior is thus a consequence of both the actual neighborhood structure, and of its potential variability.

Our approach builds on several recent developments in social choice theory but is ultimately based on the insights on cooperation in iterated Prisoners dilemma.[1] Repeated games, it has been argued, permit for complex strategies inducing reputation, retaliation, learning, etc. This facilitates identification of defectors and their punishment or evasion, which potentially benefits cooperators. We are interested in the role of social exclusion as a mechanism to both retaliate and prevent from future defection. Ostracism is the most radical form of social exclusion and can be very successful in promoting cooperation (Hirshleifer and Rasmusen, 1989). It however requires that exclusion forms a strong social norm between cooperators. To detect if and how such a norm evolves we permit agents to shape individually their own exclusion practices. A paradigm for this set-up is the option to refuse to play in the Prisoners dilemma game. Outside option, in short, has been experimentally proven effective in promoting cooperative relationships, both for one-shot games by Orbell and Dawes, 1993, and for repeated games by Hauk, 1999. See also Fehr and Gachter, 2000 for further experimental evidence on retaliation based cooperation.

Individual choices of interaction partners, comprehended from the global perspective, form a social network between the agents. This social structure can however itself affect the evolution of conventions and social norms (Ellison, 1993, Blume, 1993).[2] Initial experimental investigations of economic action in structured population were reported by Berninghaus and Schwalbe, 1996, Kirchkamp and Nagel, 2000, and Cassar, 2002. Kirchkamp and Nagel compared cooperation on different network structures with equal neighborhood size and reported lower contributions on the (segregated) circles than in the (fully connected) groups. Cassar varied neighborhood sizes and observed that cooperation increases with segregation, but remains difficult to achieve. Both studies report cooperation levels below 50%.

Combining these research programs, Jackson and Watts, 2000 and Goyal and Vega-Redondo, 2000 proposed models of co-evolution of social structure and

[1]Discussion on cooperation in finitely repeated Prisoners dilemma has long been initiated by Axelrod, 1984 and Kreps et al., 1982. See Axelrod, 2000 for an overview of recent developments.

[2]For recent update see Eshel et al., 1998, Chwe, 2000 and Morris, 2000.

coordination. The insights from these papers remain restricted to theory[3]. The co-evolution of social structure and cooperation in this paper is investigated experimentally.

We consider groups of six players playing Prisoners dilemma games with their chosen neighbors. In between each of the 60 rounds they can modify personal neighborhoods and the action in the game. They cannot however discriminate in the action choice between their neighbors. We investigate four experimental treatments, which differ in the cost of social exclusion and in the degree of information shared between players. We compare the experimental treatments to a control treatment with fixed neighborhoods and conclude that free partner choice increases cooperation. The magnitude of this increase as well as the network structures do vary between treatments and we identify conditions that promote group welfare. In particular, when exclusion is cheap and information is maximally shared, majority of groups in our experiment establish unanimous cooperation in the full network in all but the initial and the final periods.

2 The Network Game

The network game is a multi-player generalization of a Prisoners Dilemma (PD) game with option of not playing the game (PD game with outside option, Orbell and Dawes, 1993). In the basic PD game, two players simultaneously choose between two actions: "cooperate" and "defect". Depending on their chosen pair of actions their respective payoffs are given by the following bimatrix, where $d > a > b > c$:

	cooperate	*defect*
cooperate	a, a	c, f
defect	f, c	b, b

A PD with outside option is an extension of the basic game in which a third action can be taken by each of the two players, namely, to refuse to play the game. In case one of them chooses not to play the game both players receive v, the value of the outside option, otherwise they receive payoffs according to their action choices.

The "prisoner" metaphor dominates the definition of the basic Prisoners Dilemma. However, people are in general not forced to engage in risky social dilemmas, and the high contribution rates observed in PD games could be related to the players' decision to engage. In our opinion, the outside option suggests closer resemblance of a general social situation, which motivates its generalization in models of evolution of social behavior. We consider the following social dilemma based on an endogenous choice of partners. In the network game a group of N players simultaneously make two decisions, (1) neighbor choice and (2) action choice.

(1) Each player proposes links between himself and other players. Linking is costless. A player can propose a link to any other player in his group and there is no limitation on the number of proposed links. A link between two players is however established only when it is mutually proposed and in this case we say the players are neighbors.

[3] Corbae and Duffy, 2000 announced preliminary experimental tests.

(2) Each player chooses one action for a PD game. One game is played by each pair of linked players. Players, however, cannot discriminate in their action choices. They can either cooperate with all their neighbors or defect with all of them. A player receives v, the value of the outside option, for each player he is not linked to.

Formally, let $\mathcal{N} = \{1, 2, ..., N\}$ be the set of players, let P_i be the set of players to whom a player i proposed a link, and let s_i be his action. A link between players i and j is established only when $i \in P_j$ and $j \in P_i$. Let $\mathcal{L}$ be the set of all established links:

$$\mathcal{L} = \{ij \mid i \in P_j \text{ and } j \in P_i\},$$

Let L_i be a set of players with which a player i has established a link. We call players in L_i the *neighbors* of player i and the set L_i his *neighborhood*. We also say player i *refused* a link with player j, when $i \in P_j$ and $j \notin P_i$. The number of links a player i proposes, $|P_i|$, is his *outdegree* and the number of other players that proposed a link to i, $|\{j \mid i \in P_j\}|$, is his *indegree*.

Player i plays a Prisoners dilemma game with all his neighbors and his payoff is

$$\pi_i(s_i, s_{-i}) = \sum_{j \in L_i} u(s_i, s_j) + \sum_{j \notin L_i} v,$$

where s_{-i} denotes the strategies chosen by remaining players $\mathcal{N} \setminus i$, $u(s_i, s_j)$ gives the payoff for choosing action s_i when opponent chooses s_j in the basic Prisoners dilemma game, and v is the value of the outside option.

When the outside option value is higher than the mutual defection payoff, $v > b$, all Nash equilibria predict an *empty network*, in which no links are established. When the outside option value is lower than mutual defection payoff, $v < b$, the unique strict Nash equilibrium predicts that all players defect and link to all other players thus establishing the *full network*[4]. In particular, proposing a link and cooperating is never part of an equilibrium strategy, for any value of v. The network game can be viewed as a linear public game with endogenous group choice.

Another N-player generalization of the Prisoners dilemma with outside option has been studied by Orbell and Dawes, 1993 and in its repeated form by Hauk, 1999 and Hauk and Nagel, 2001. Orbell and Dawes studied one-shot PD games with outside option value $v > b$. They report increased cooperation between players who chose to play the game and attribute this to the "false consensus" effect[5]. Hauk confirms this observation in the repeated game but reports decreased cooperation for repeated PD game with outside option value $c < v < b$.

[4]In case of $v < b$ there are other Nash equilibria in which some links are not established. Namely, proposing no links between two players can be a part of a (non-strict) equilibrium. Only when player has no links established he is indifferent between cooperating and defecting. As proposing all links and defecting dominates all other strategies, we consider only the strict Nash equilibrium in the one-shot game.

[5]When $v > b$ it is not rational to link to a defector. According to the "false consensus" the cooperators expect more cooperation than defectors do. Therefore they are more likely to play the game.

The players in these studies could propose their neighbors as in (1) but could discriminate in action choices between the neighbors. In our opinion this formulates a collection of independent Prisoners dilemma games with outside option between all possible pairs of players in the group. In contrast to this setting the players in our network game do endogenously determine their neighborhood but cannot discriminate in action choices between their neighbors. The motivation for this restriction is to allow for network effects, such as the evolution of norms and the implicit interference between players that are not direct neighbors. An agent in our setting might choose her action only as a reaction to behavior of some of her neighbors, yet her decision affects all her neighbors, even if they are not linked among themselves. In an extreme example, a defective action by one agent triggers defection by her neighbors, which in turn triggers defection by the neighbors of her neighbors, until defective behavior spreads to the whole group. To protect cooperation, though, the agents might remove links with defectors. Such a setting therefore allows us to study the evolution of social norms in parallel to the dynamics of the network structure. This however requires that the network game be played repeatedly by the same set of players.

Repeated network game. Repeated games potentially allow for more complex strategies, such as reputation building and retaliation. In a one-shot network game it is never rational to cooperate. In the repeated network game on the contrary, cooperating with the neighbor may be individually rational. The following is the basic example extended to the network game (see e.g. Kreps et al., 1982). Assume that a player believes there is a probability that her own defection induces her neighbors to remove the links in subsequent rounds. If the player believes this probability to be high enough, her cooperation is the best response.

Consider therefore the network game with the same players that is repeated $m > 1$ times. At the end of each round, the information about the history of play is given to the players according to an information condition. For each player, the information condition defines which past decisions of which other players she observes. The extensive form of the repeated network game thus depends on the corresponding information partitions at each decision node. The subgame perfect equilibria (SPE), however, are independent of the information condition. In particular, as in the one-shot case no links are ever established in an SPE, when $v > b$. In the case when $v < b$, the unique strict SPE predicts that players in all rounds establish a full network and unanimously defect[6]. Regardless of the outside option value and the information condition, proposing a link and cooperating is never part of a subgame perfect equilibrium.

Cooperation does, however, appear in the non-subgame perfect Nash equilibria, which exist only when $v < b$. The weakly dominated option to refuse a link adds in this case a punishment possibility that is harder than defection itself. But the threat of exclusion is non-credible, and can be present only off the equilibrium path, as for example in the strategy:

> "Cooperate until round $m - 1$ and defect in round m. Punish all deviations from this sequence by exclusion for one round, otherwise propose all links",

[6] Other non-strict SPE exist in case of $v < b$, predicting some links simultaneously not proposed, see footnotes above.

	cooperate	defect
cooperate	$50, 50$	$10, 70$
defect	$70, 10$	$30, 30$

Figure 1 Payoff matrix of the elementary PD game.

which constitutes the symmetric Nash equilibrium when $v < (a + b - d)$.

We shall distinguish between the level of cooperation and the level of cooperative play. The *level of cooperation* is the normalized number of players who chose to cooperate, regardless of their linking decisions. If all players in the group cooperate we say that the level of cooperation is 1, if all defect this level is 0. The *level of cooperative play* measures the rate of cooperation in the games that were actually played: it is calculated as the number of cooperative actions over all links, normalized by double the number of established links:

$$\textit{level of coop. play} = \frac{1}{2\,|\mathcal{L}|} \sum_{ij \in \mathcal{L}} (\delta_i + \delta_j),$$

where δ_i is one if player i cooperates and zero otherwise and $|\mathcal{L}|$ is the number of established links. This measure is not affected by the actions of players that are not linked. Further, the weight of a player's action increases with the number of games he played (his links). The level of cooperative play is 1 if all linked players cooperate and 0 if all linked players defect.

The *density* of the network gives the number of established links normalized by 15, the number of all possible links between 6 people. Density is thus 1 when the network is full and 0 when the network is empty.

3 Experimental design

Our design consists of four experimental treatments and one control treatment. In all five treatments the elementary Prisoners Dilemma game is the same. In the four experimental treatments groups of $N = 6$ subjects play 60 rounds of the network game with the elementary PD game given in Figure 1. In the control treatment groups of $N = 6$ subjects play 60 rounds of a multi-player generalization of PD game without the outside option: neighborhoods of all six subjects are in each turn exogenously fixed to create the full network, i.e. $P_i = \mathcal{N} \setminus i$, for each i.

The experimental treatments I-IV differ from the control treatment V only in that players have the freedom to exclude other others. This allows us to investigate the consequences of the endogenous choice of neighbors for cooperation levels and group welfare. We consider two values of the outside option:

1. high, v = 40, and
2. low, v = 0.

information:	maximal	minimal
$v = 0$	I	II
$v = 40$	III	IV

Figure 2 Information condition and outside option value for experimental treatments I - IV.

In addition, we vary the information about the history of play that is shared among the subjects. We consider two information conditions:

1. Maximal information: in each round the subjects observe all past choices of all other subjects in their group. This includes both choices of neighbors and action choices.
2. Minimal information: the subjects in each round observe only the action choices of their neighbors.

Using four experimental treatments we adopt a two by two design as is shown in Table 2.

If there were no relationship between the choice of neighbors and the choice of actions, the behavior in the treatments with the low outside option (I and II) would coincide with the behavior in the control treatment. Then a full network would always emerge, and varying the information condition would have no effect. In the treatments with the high outside option (III and IV) the unique equilibrium predicts an empty network, again regardless of the information condition. We thus take as our initial hypotheses:

H0.1a No difference in cooperation levels between treatments I, II, and V.

H0.1b Full network in treatments I and II.

H0.2 Empty network in treatments III and IV.

In all experimental treatments it is rational to propose links to potential cooperators. When one believes there is a large enough chance that another player will cooperate, linking to that player payoff-dominates not linking. Proposing to potential defectors and other strategic considerations depend, however, on the value of the outside option. We therefore discuss treatments I and II separately from treatments III and IV.

3.1 Treatments with low outside option

Linking to potential defectors is individually rational from the one-shot perspective: playing the PD game always payoff-dominates the low outside option. In a repeated network game it may, however, be rational to exclude a defector if one believes that this will induce the defector to propose a link and cooperate in subsequent rounds. As discussed above, there exist Nash equilibria that prescribe cooperation and exclusion of defectors. If experimental subjects in treatments I and II play these equilibrium strategies, we should observe an increase in cooperation with respect to our control treatment.

H1.1a The indegree of cooperators is higher than indegree of defectors. Cooperation levels in treatments I and II are higher than in treatment V.

Punishment by exclusion is costly and there is a free-riding incentive among cooperators not to exclude a defector if others do. Higher levels of exclusion might be necessary to increase the credibility of punishment (see Axelrod, 1986 for an elaboration). For example, cooperators who did not exclude defectors might find themselves excluded by cooperators who did exclude defectors. This threat is, however, even less credible than the exclusion of defectors, and requires that the linking choices are public information. We should detect differences between treatments I and II, if these higher forms of exclusion do play a role. In general, the minimal information condition limits the evolution of common social norms, as subjects cannot observe any linking choices, and when excluded not even action choices. Loss of information parallels the loss of links and cooperators might not exclude. Subjects in Treatment I, on the other hand, observe exclusion practices of all other subjects and can adjust their behavior, e.g. following their best response or imitate the most prominent behavior in the group.

H2.1 Cooperation levels and indegrees to defectors differ between Treatment I and II.

The density of links depends both on the number of defectors and on the propensity of cooperators to exclude them. In general, if there are many cooperators or many defectors, the density of the network increases.

H1.1b The network density increases with the homogeneity of action choices.

3.2 Treatments with high outside option

In treatments III and IV it is rational to link to a cooperator but not to a defector. The threat of punishment of defectors by exclusion is thus credible and cooperators, if any, are more likely to establish links than defectors. This in turn increases cooperative choices in the games that are actually played.

H1.2a The levels of cooperative play in treatments III and IV are higher than in treatment V.

In a repeated network game the subjects update at each round their beliefs on the distribution of linking and action choices in the group. The updating quality depends on the quality of information the subject obtains each round. In treatment III all subjects observe cooperative play and can respond by proposing a link. This induces cooperative action by the proposers of links, as signalling defection seriously reduces others' incentives to establish these links. One cooperative play signalling intentions to cooperate in a game might therefore lead to complete cooperation in the full network.

Session	1	2	3	4	5	6	7	8	9	10	11
Treatment	I	I	II	II	III	III	IV	IV	V	V	IV
number of groups	3	3	3	3	3	4	2	3	2	3	3

Figure 3 Distribution of treatments and number of independent groups over sessions.

In contrast, in treatment IV the quality of information depends on the size of the neighborhoods. Subjects whose prior beliefs lead them to behave according to the subgame perfect equilibrium will never link and therefore never update their prior beliefs in Treatment IV. Defectors, once excluded, cannot signal cooperative intentions and cooperators might get separated into small cliques.

H2.2 The level of cooperation is higher in treatment III than in treatment IV. The network density is higher in treatment III than in treatment IV.

H1.2b The network density increases with the cooperation level.

4 Experimental procedures

All the experiments were conducted in the CREED laboratory at the University of Amsterdam. Subjects were students of various disciplines but mainly economics undergraduates. They were recruited through notices on bulletin boards in the faculties and through email announcements. Each subject participated in only one session and none had participated in a similar experiment in CREED before. The experiments were computerized using the software developed in CREED as a general tool for network experiments. To ensure anonymity at least twelve subjects were recruited for every session, and were randomly divided into at least two independent groups. See Table 3 for the distribution of treatments per sessions.

Instructions were read aloud and each participant received a written copy. Before the experiment started, each participant had to answer a number of questions at her computer terminal, specifically designed to test for understanding of the network game, the payoff procedure and the computer screen. There was also a single practice round at the end of which no information about the choices of other subjects was given. No communication other than through decision-making was permitted. To increase anonymity during experiment each subject referred to herself as "Me" and to the other five members of her group with the letters "A", "B", "C", "D", and "E". The letters referred to the same subject throughout all rounds, and this was common knowledge. The network game was neutrally framed: subjects collected points by forming "connections" and choosing between the actions "green" and "blue".

Care was taken to minimize differences in the instructions for different treatments. In all the treatments the subjects were informed about the size of their group, procedures to propose and establish links, the payoff function, the information condition, and the number of rounds[7]. Each session took approximately two hours.

[7]Please contact authors for a copy of instructions and software presentation.

Treatment	Payoff in SPE	Min payment	Max payment	Average payment
I	18	11.5	29.8	21.5
II	18	10.1	29.8	22.7
III	24	24.7	30	28.8
IV	24	23.5	27.8	25.4
V	18	16.8	27.5	22.2

Figure 4 Payoff per subject in the subgame perfect equilibrium. Actual minimum, maximum, and average payoffs over the treatments.

4.1 Payment

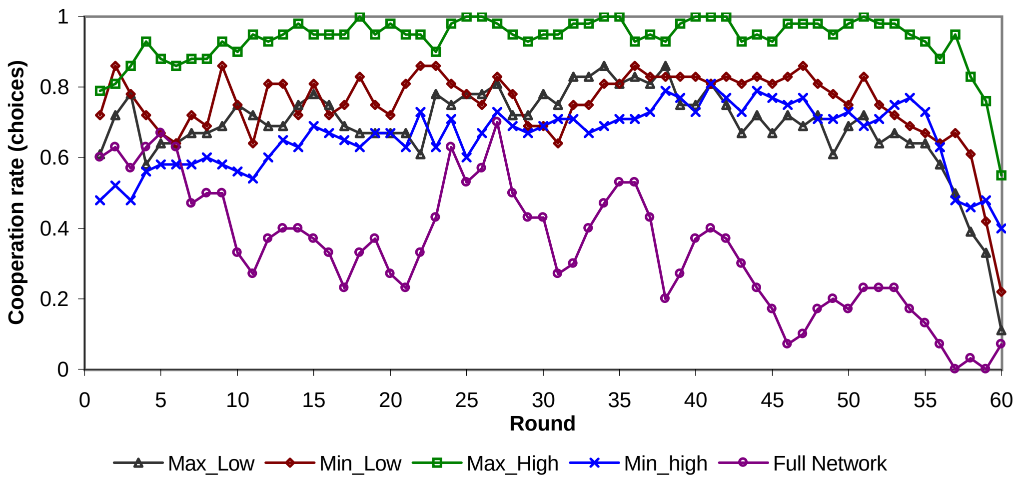
At the end of the session each subject was individually paid her own payoff in cash. The payoff of a subject was determined by the sum of the points earned over all the rounds. The conversion rate was 1 euro for 500 points. In all treatments the joint group earning was maximized when all subjects in all rounds cooperated and established the full network. In this case the payoff per subject equals 30 euro. The payoff per subject when all subjects played the subgame perfect equilibrium differs between treatments and is given in Table 4, together with the minimum, maximum, and average of the actually earned payoffs.

5 Experimental Results

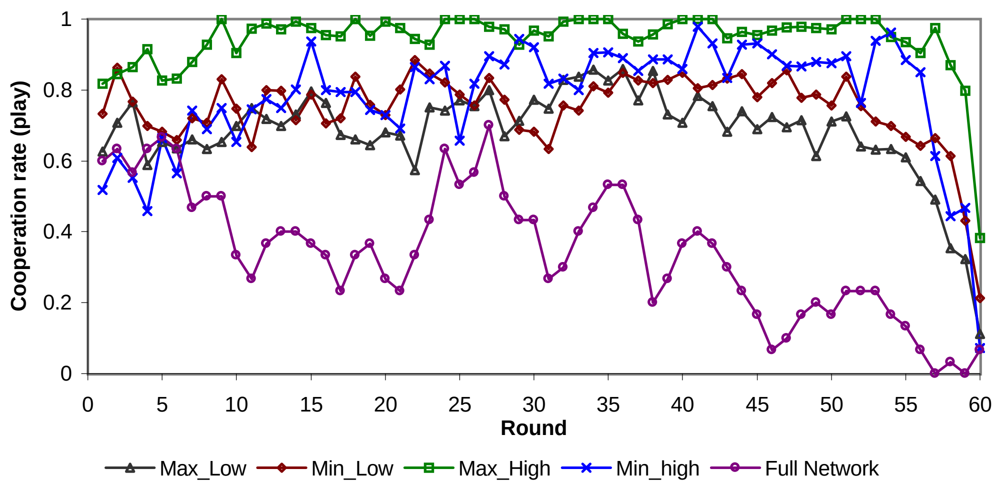
We start with analyzing the action decisions (i.e. *cooperate* or *defect*) in the different treatments. As expected in none of the treatments subgame perfect equilibrium play, i.e. always defect, is observed right from the beginning. Actually, in all treatments subjects cooperation rates have been considerably above zero and very similar. However, there are remarkable differences in the evolution of action choices in particular between the control treatmen T-V and the four experimental treatments T-I to T-IV.

Result 1. Cooperation in endogenous and exogenous networks
(i) In the first period action behavior is not subgame perfect. The cooperation rates are around or above 50 percent in all five treatments. (ii) During later rounds cooperation rates converge to complete free riding only in T-V. Cooperation rates in the four treatments with endogenous link formation are significantly higher than in T-V and close to full cooperation in the later rounds of the game. Furthermore, except for some last round(s) effect, in all four experimental treatments there is no tendency towards full defection.

A first support for this result comes from Figure 5(a)-(b) and Figure 6(a)-(b). Figure 5 depicts, for each treatment, the evolution of cooperative choices (a) and cooperative plays (b) for all sixty rounds. It is obvious from the figure that in the first round there is a substantial amount of cooperation in all treatments (see also Table 1). What these figures also show is that the dynamics of cooperative choices differ significantly between the four experimental treatments T-I to T-IV and the control treatment T-V.



(a) Cooperative choices



(b) Cooperative plays

Figure 5. Evolution of cooperation rates (group averages)

In particular, Figure 6(a)-(b) show nicely that in T-V there is a strong trend towards complete defection. Interestingly, there are frequent tries to return to more cooperation (see the 'hills' in Figure 5(a)-(b)) which are, however, not successful. A completely different dynamics can be observed in the other four treatments. Although the levels of cooperation are not the same in the four experimental treatments (see also below) the dynamics are very similar. In particular, all four treatments exhibit very high and stable cooperation rates (choices as well as plays) across rounds. The cooperation rates steadily increase till they reach a maximum in rounds 41 to 50 (except for T-I where the maximum is reached in rounds 31-40). Thereafter - due to end round(s) effect - the cooperation rates are decreasing again. Note, however, that they are with the exception of the very last round always considerable above the cooperation rates of T-V, in all four experimental treatments.

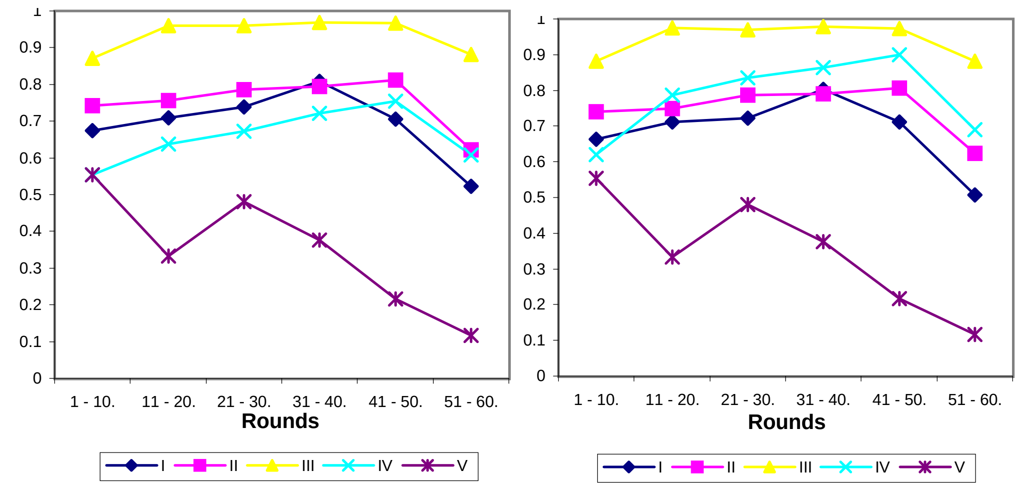
	Cooperative choices:			Cooperative plays:		
Treatment	first round	all rounds	rounds 1-55	first round	all rounds	rounds 1-55
I	0.611	0.693	0.721	0.626	0.686	0.715
		(0.188)	(0.140)		(0.198)	(0.150)
II	0.722	0.752	0.774	0.734	0.750	0.771
		(0.151)	(0.111)		(0.150)	(0.110)
III	0.786	0.934	0.947	0.817	0.93	0.957
		(0.113)	(0.071)		(0.134)	(0.079)
IV	0.479	0.658	0.673	0.519	0.786	0.809
		(0.159)	(0.140)		(0.247)	(0.206)
V	0.600	0.346	0.375	0.600	0.346	0.375
		(0.269)	(0.262)		(0.269)	(0.262)

Table 1 – Cooperation rates (group averages)

To get further support for this result we have run non-parametric Mann-Whitney tests for the first round, rounds 1-10, 11-20, 21-30, 31-40, 41-50, 51-60, and the last round, with group averages in each treatment as unit of observation. These statistics corroborate the observations drawn from the figures. In the first round only the actions taken by subjects in T-III differ significantly from those taken in T-V ($p = 0.090$ for "choices", $p = 0.042$ for "plays"; all test statistics are one-sided). In the later rounds, however, the cooperation is higher in all four treatments T-I to T-IV than in T-V. In particular, in periods 31-40, 41-50, and 51-60 the differences are significant at least at the 5 percent significance level and mostly (in nine out of 12 cases) at the 2.5 or 1 percent level. This holds for both measures of cooperation. Interestingly, even in the very last round cooperative choices in T-III and T-IV are significantly above those in T-V ($p < 0.005$, one-sided). Cooperative plays in the last round are significantly larger in T-III than in T-V ($p < 0.025$, one-sided). The same results are found when we look at the cooperation levels across all rounds or (to control for the end round(s) effect) across rounds 1-55 (see also Table 1. A Mann-Whitney test shows that cooperative choices and cooperative plays in T-I to T-IV are significantly above those in T-V (at least at the 5 percent level; mostly at the 2.5 percent level).

Figure 5 not only shows that the cooperation rates are higher in our experimental treatments but also points to the possibility that behavior in these treatments is much more stable than in the control treatment. To find statistical support for this impression we calculated the standard deviation of cooperative choices and plays across rounds for each group. Table 1 shows in parentheses the averages of these standard deviations for each treatment for rounds 1-60 and rounds 1-55. For cooperative choices a Mann-Whitney test shows that in comparison to T-V the standard deviations in T-I to T-IV are significantly smaller ($p < 0.01$ for all pair-wise comparisons, one-sided test) when taking all rounds into account and also when taking all but the last five rounds into account. For the cooperative plays the results are less pronounced but qualitatively the same as for the choices ($p < 0.025$ for all but the comparison between T-IV and T-V for which it is only marginally significant when taking all but the last five rounds into account; one-sided tests). We summarize in the following

Result 2. Stability of cooperative behavior
Compared to T-V the action choices in the treatments with endogenous network formation exhibit significantly more stability.



(a) Cooperative choices

(b) Cooperative plays

Figure 6. Cooperation rates (10 round averages)

Together with Result 1 this result undoubtfully shows that the possibility to form the network endogenously, i.e. to choose with whom to play, serves as a powerful mean for promoting *and* stabilizing cooperation at a rather high level. However, the figures also indicate that there are some differences between our four experimental treatments with respect to cooperation levels and also concerning network formation. In the following we shall elaborate more on these issues.

Result 3.
In T-III significantly more cooperation is observed than in all other treatments. Cooperation rates in the other three experimental treatments do not differ significantly.

Support for this result comes again from Figures 5 and 6. From them it is obvious that cooperation rates (choices as well as plays) are highest in the treatment with full information and the high outside option, T-III. As can be seen for many rounds even full cooperation is observed. Furthermore, Figure 7 shows the distribution of the number of cooperators in a group for the different treatments. It impressively shows that by far most frequently all players in a group cooperate when they are in T-III. These visual impressions are corroborated by a Mann-Whitney test applied to cooperative choices and plays for all rounds and for rounds 1-55. In pairwise comparisons we find that the differences are significant at least at the 5 percent level (one-sided test) when looking at all but the last five rounds. When taking all rounds into account the differences are significant at least at the 2.5 percent level when compared to T-I and T-IV and marginally significant when compared to T-II (all tests one-sided). A pairwise comparison of cooperation rates of the other three experimental treatments does not show any significant differences ($p > 0.12$ in all cases, one-side tests).

Another important question is how dense the endogenously formed networks are. Recall that, according to standard game theory, in the subgame perfect equilibrium the full network should be formed in T-I and T-II whereas the empty network should

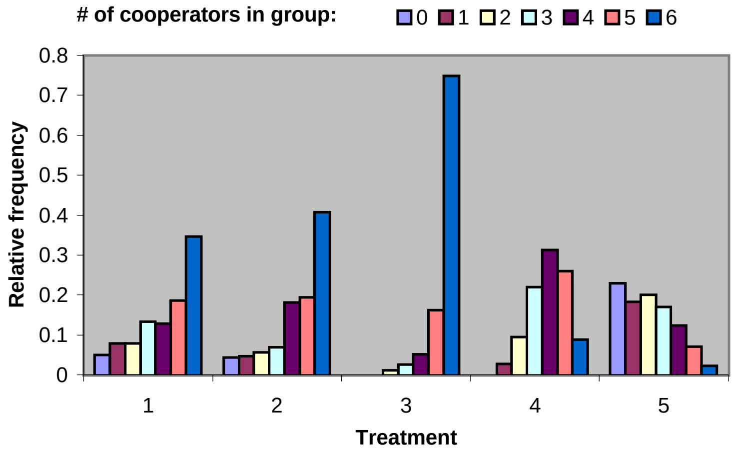


FIGURE 7. DISTRIBUTION OF NUMBER OF COOPERATORS PER TREATMENT

occur in treatments T-III and T-IV. In addition, in some non-subgame perfect Nash equilibria non-full networks may be observed in the treatments with low outside options (T-I and T-II).

Result 4. NETWORK DENSITIES
In the two experimental treatments with the low outside option (T-I and T-II) the network density is rather high (around 0.8 over all rounds). In stark contrast to standard theory the network density is also high in the treatment with high outside option and full information (T-III). In T-IV, however, the network density is relatively low (around 0.3 over all rounds).

Support for this result comes from Figure 8. The figure shows the evolution of the network density over time. As can be seen average network density in the two treatments with the low outside option is constantly around 0.8. Over all rounds the average density is 0.82 in T-I and 0.84 in T-II.This tells us that on average 12.3 and 12.6, respectively, links out of 15 possible links are formed. Surprisingly, a similar figure can be observed for treatment T-III (high outside option and full information). There the average density over all rounds is with 0.86 (12.9 links) even the highest among all treatments. In T-IV, the treatment with also a high outside option but minimal information, however, an average network density of only 0.35 (5.3 links) is observed. The large and significant ($p < 0.01$, two-side Mann-Whitney test) difference in network density between between T-III and T-IV indicates the importance of information about previous play in case of high outside options. A further indication for the importance of information of past play comes from the fact that the densities in T-III and T-IV start out at a similar level (0.51 and 0.57, respectively), but develop in completely opposite directions. While network density in T-III increases steadily it drops strongly in the beginning in T-IV and stays then at the low level. This finding also verifies Hypothesis H2.2.

Having observed the high cooperation rates in all experimental treatments as well as different network densities it is natural to ask for the driving force behind this result. The most natural candidate is (fear from) social exclusion. To get a first impression we look at the proposed links by cooperators C_t and defectors D_t in round t to cooperators

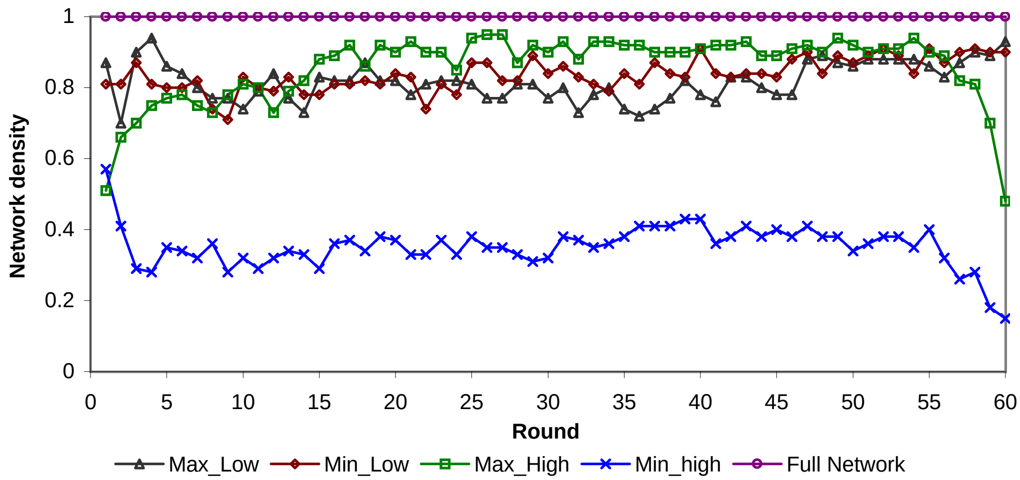


FIGURE 8. EVOLUTION OF NETWORK DENSITY PER TREATMENT

C_{t-1} and defectors D_{t-1}. Figure 9 depicts these proposed links. The figures show that in all treatments cooperators are relatively little likely to link to previous period defectors. In T-I in only 46 percent of all possible cases a cooperator proposed a link to a previous period defector. In T-II this is in 64 percent the case. Note that those who do not propose a link forego for sure 10 points (the 'sucker' payoff) in these treatments. Hence, cooperators to not hesitate to *costly* exclude defectors. In T-III and T-IV cooperators are even less likely to link to previous round defectors. Only in 23 percent and in 24 percent of all possible cases cooperators propose a link to previous period defectors. That the number of proposed links to defectors are smaller in these treatments than in the other two is intuitively clear. In T-III and T-IV a player foregoes money by not proposing a link only if the other person would also cooperate, otherwise it is better to stay out. This is of course also true for defectors.While in T-I and T-II defectors propose almost always links to both, previous period cooperators and defectors, these numbers are much smaller in T-III and T-IV. We summarize in the following

Result 5. PROPOSED LINKS
In all four treatments cooperators are much more likely to propose links to previous period cooperators than to previous period defectors. Cooperators do not hesitate to exclude previous period defectors even if this is costly for sure (T-I and T-II). This tendency to refuse links to defectors is reinforced in the T-III and T-IV.
For low outside options defectors try to link to everybody. In the treatments with high outside options they are more likely to link to previous period cooperators than defectors.

The above result tells us that intending cooperators are more likely to link to past cooperators than to past defectors. This does not ensure, however, that the network structure is such that cooperators are mainly the neighbors of cooperators. There is still the possibility that past cooperators change to actual defectors and vice versa. In particular it is interesting to examine to what extent cooperators were able to exclude defectors and how this differs across treatments.

Result 6.ESTABLISHED LINKS *In all treatments there is a tendency that cooperators and defectors stay isolated in small cliques. This isolation takes plays more frequently* *when the outside option is high (T-III and T-IV). In particular, in T-IV isolation of cooperators is almost complete.*

Support for this result comes from Figure 10. It depicts the average frequencies of established links (in a period t) between different types of players for our four experimental treatments. The diameters of the circles around C_t (cooperator in round t) and D_t (defector in t) are proportional to the average frequency of cooperators and defectors, respectively, in a round. For T-I and T-II (low outside option) the figure shows that cooperators are almost always neighbors of cooperators (in 85 and 87 percent of all possible cases in T-I and T-II, respectively) and defectors are almost always neighbors of defectors (96 and 89 percent T-I and T-II, respectively). A link between cooperators and defectors, however, is established much less likely. In T-I only in 67 percent of all possible cases a defector is a neighbor of a cooperator. In T-II this is the case in 74 percent of all cases. This shows that cooperators are not only trying to exclude defectors (see 5 but also that they are (at least to a certain) successful in doing so. If the outside option is high exclusion of defectors is cheaper and therefore more easily done. This is reflected by the low frequencies of established links between cooperators and defectors in T-III and T-IV. In particular, in T-IV when no information is provided defectors are almost completely excluded. This also shows that the low density of the networks in this treatment is almost exclusively due to the exclusion of defectors.

Result 7 *(i) In T-I and T-II homogeneity of action is (weakly) significantly positively correlated with network density.*
(ii) In T-III and T-IV cooperation is significantly positively correlated with network density.

To support part (i) of this result we calculate a measure for action homogeneity. For that we take the absolute distance of actual cooperative choice (in a group over all rounds) from 0.5. A cooperation rate of 0.5 means that half of the subjects cooperate and half of the subjects defect, which defines the least homogeneous group. The Spearman rank order correlation coefficient is 0.77 ($p = 0.036$, one-sided) for T-I and 0.58 ($p = 0.114$, one sided). In T-III and T-IV the Spearman rank order correlation coefficient is 0.72 ($p = 0.035$, one-sided) and 0.91 ($p = 0.001$, one-sided). The result gives support to our hypotheses H1.3 and H2.3.

6 Discussion and Conclusion

In this paper we investigate experimentally the dynamics of cooperation in endogenous social structures. We implement an environment where agents are free to choose their neighborhoods within which they play Prisoners Dilemma games repeatedly. In this set-up we can examine the simultaneous evolution of cooperative behavior and of the social structure. We report on four experimental treatments, which differ in the cost of social exclusion and in the flow of information between the agents. We compare the outcomes of these to a control treatment where the network structure is exogenously imposed. The following are our main conclusions: (i) Free choice of social links drastically increases cooperation. In all four experimental treatments cooperation is significantly higher compared to the control treatment. Cooperation levels in the

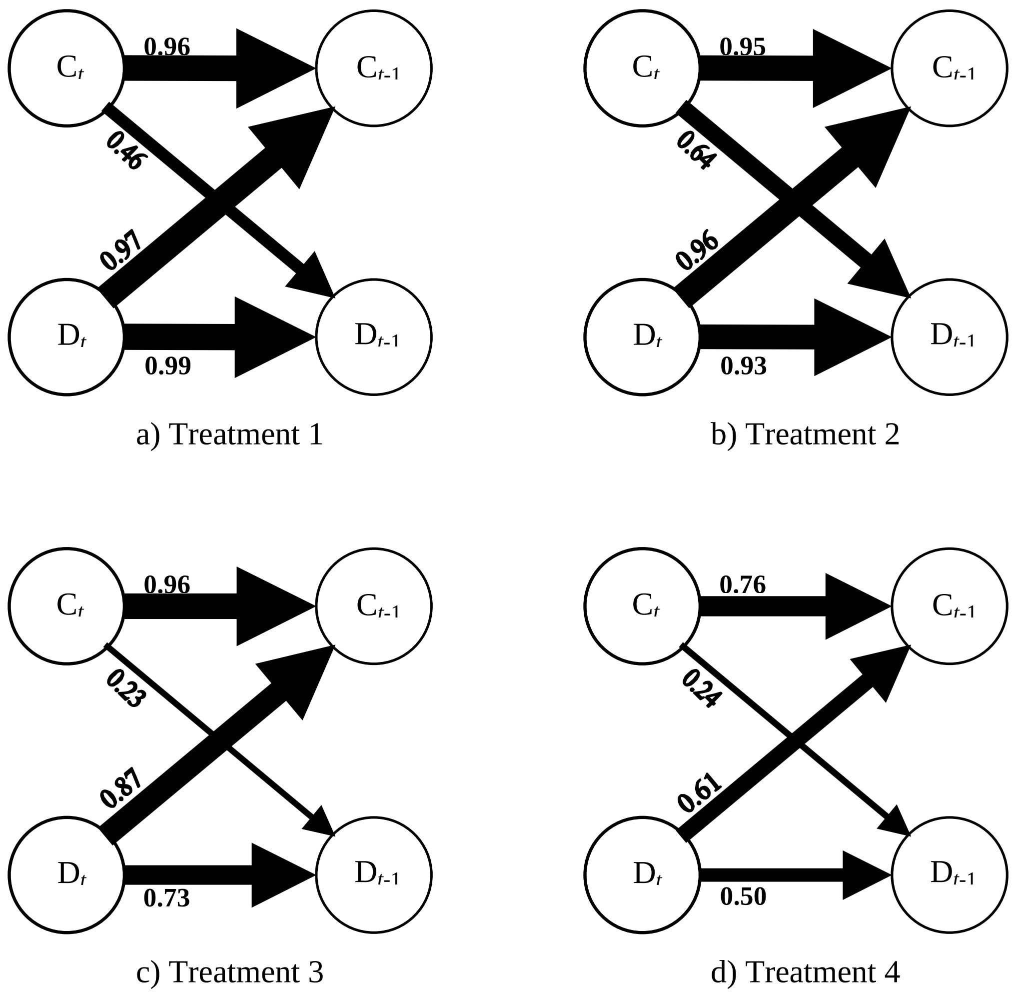


Figure 9. Proposed links by cooperators and defectors to previous round cooperators and defectors

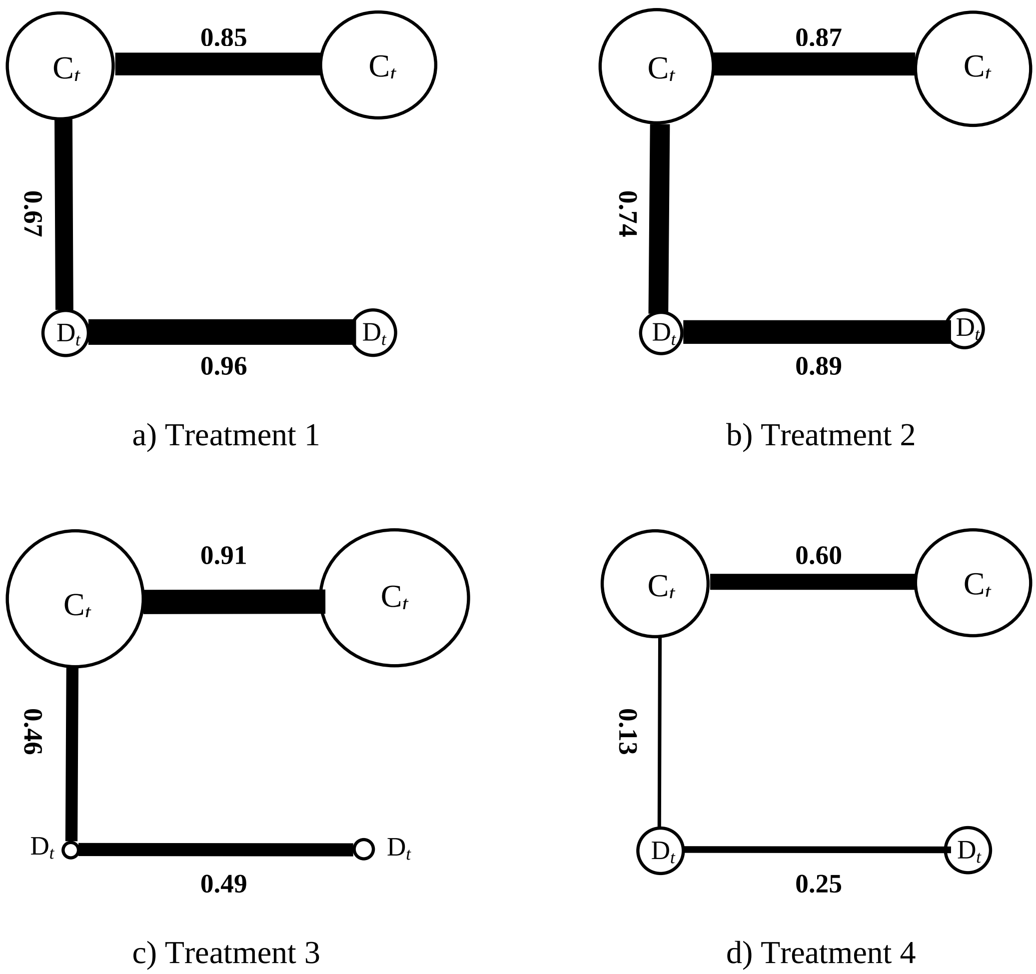


FIGURE 10. FREQUENCIES OF ESTABLISHED LINKS BETWEEN DIFFERENT TYPES OF PLAYERS

experimental treatments appear to be constant or are even increasing. (ii) Contrary to the subgame perfect predictions, subjects use the option of not playing with some of their neighbors and costly exclude those who tend to defect. In particular, when exclusion is relatively cheap our subjects keep forming links and cooperate. In all experimental treatments cooperators that refuse linking to defectors account for most of social exclusion. (iii) Cooperation however appears to be affected by both the cost of exclusion and the information flow. When exclusion is costly, subjects reach high, but rarely full cooperation, independently of the information flow. When exclusion is cheap, however, information about past behavior appears a major determinant of behavior. Under maximal information, cooperation levels reach 93 percent and five out of seven groups unanimously cooperate and fully connect for at least 30 consecutive periods. In contrast, when information is minimal, only 66 percent of the actions are cooperative. This is attributed to different belief updating potential, as defectors under the minimal information, being isolated, never realize that there exist cooperative cliques and therefore never adjust their behavior. (iv) This accounts also for the large differences in the network structures between treatments with maximal and those with minimal information, when exclusion is cheap. The groups under maximal information completely integrate while groups under minimal information disintegrate into cooperative cliques and isolated defectors. Such a difference, however, was not observed in treatments with costly exclusion.

References

Axelrod, R. (1984). *The Evolution of Cooperation*. New York: Basic Books.

Axelrod, R. (1986). An evolutionary approach to norms. *American Political Science Review*, 80:1095–1111.

Axelrod, R. (2000). On six advances in cooperation theory. *Analyse & Kritik*, 22:130–151.

Berninghaus, S. K. and Schwalbe, U. (1996). Conventions, local interaction, and automata networks. *Journal of Evolutionary Economics*, 6:297–312.

Blume, L. E. (1993). The statistical mechanics of strategic interaction. *Games and economic behaviour*, 5(3):387–424.

Cassar, A. (2002). Games in networks. mimeo.

Chwe, M. S. (2000). Communication and coordination in social networks. *Review of economic studies*, 67:1–16.

Corbae, D. and Duffy, J. (2000). Experiments with network economies. mimeo.

Ellison, G. (1993). Learning, local interaction, and coordination. *Econometrica*, 61(5):1047–1072.

Eshel, I., Samuelson, L., and Shaked, A. (1998). Altruists, egoists and hooligans in a local interaction model. *Journal of economic theory*, 88:157–179.

Fehr, E. and Gachter, S. (2000). Cooperation and punishment in public goods experiments. *The American Economic Review*, 90(4):980–994.

Goyal, S. and Vega-Redondo, F. (2000). Learning, network formation and coordination. mimeo.

Hauk, E. (1999). Multiple prisoner's dilemma games with(out) an outside option: An experimental study. mimeo.

Hauk, E. and Nagel, R. (2001). Choice of partners in multiple two-person prisoner's dilemma games: An experimental study. *Journal of conflict resolution*, 45 (6):770–793.

Hirshleifer, D. and Rasmusen, E. (1989). Cooperation in a repeated prisoners' dilemma with ostracism. *Journal of Economic Behavior and Organization*, 12(1):87–106.

Jackson, M. O. and Watts, A. (2000). On the formation of interaction networks in social coordination games. To appear: Games and Economic Behavior.

Kirchkamp, O. and Nagel, R. (2000). Local and group interaction in prisoners' dilemma experiments - imitate locally, think globally. mimeo.

Kreps, D. M., Milgrom, P., Roberts, J., and Wilson, R. (1982). Rational cooperation in the finitely repeated prisoners dilemma. *Journal of Economic Theory*, 27:245–252.

Morris, S. (2000). Contagion. *Review of Economic Studies*, 67:57–78.

Orbell, J. M. and Dawes, R. M. (1993). Social welfare , cooperators' advantage, and the option of not playing the game. *American Sociological Review*, 58:787–800.