Social status and group norms: Indirect reciprocity in a helping experiment

Seinen, I.; Schram, A.

Publication date
2001

Social Status and Group Norms: Indirect Reciprocity in a Helping Experiment

Ingrid Seinen* and Arthur Schram*

CREED
Department of Economics and Econometrics
University of Amsterdam
Roetersstraat 11
1018 WB Amsterdam
The Netherlands

e-mail:
* *IngridS@fee.uva.nl*
* *ArthurS@fee.uva.nl*

Manuscript, submitted

This version: 24 September 2001

Abstract

Experimental evidence is provided showing that indirect reciprocity is important in economic decision making and in the development of group norms. We study a so called repeated helping game with random pairing in large groups. Pairs consist of a donor and a recipient. Donors decide whether to help the individuals they are matched with against a certain cost or not to help, enduring no costs. We observe that many decision makers respond to the information we give them about previous decisions of the recipients, even though this information is based on transactions with third parties. (*JEL* C92)

I Introduction

Among economists, there is a growing awareness that both trust and reciprocity play important roles in many economic interactions. Because many activities take place sequentially, where one actor endures costs before obtaining the benefits, it is often a matter of trust that the future benefits will indeed be delivered by another actor. In theory, this future delivery can be governed by formal contracts. In practice, information asymmetries and prohibitive costs of perfect monitoring often make the use of enforceable contracts impossible. One has to rely on implicit contracts [Arrow, 1974], the enforcement of which is based on reciprocity.

We define reciprocity as conditional behavior where kind acts are rewarded and hostile acts are punished, even when this is costly. The distinguishing feature of reciprocity is that it is not based on explicit incentive schemes. In other words, the individual decision to reward or punish acts of others is not governed by her or his direct self interest. Note that we explicitly define reciprocity as behavior and not as a motivation or a preference. The motivation underlying reciprocal behavior can be related to other regarding preferences such as (reciprocal) fairness [*e.g.* Rabin, 1993, Bolton and Ockenfels, 2000, Fehr and Schmidt, 2000] or (reciprocal) altruism [Trivers, 1971, Levine, 1998] or to the quest for efficiency gains through cooperation [Brandts and Schram, 2000]. For an overview of this type of motivations, see Schram [2000].

In this paper we do not intend to contribute to the discussions on the motivations underlying other regarding behavior in general or reciprocity in particular. Instead, we focus on the form that reciprocal behavior takes and on the consequences it has for the (repeated) interaction between individuals in a group. In the concluding discussion we will briefly explore the extent to which our results enable us to draw conclusions about the motivations underlying the observed indirect reciprocity.

A good example of an (experimental) setup used to study implicit contracts and the role of reciprocity is the investigation of the relationship between an employer and an employee by Fehr et al. [1997, 1998]. In these experiments, a wage level (and corresponding effort level) is first determined in a contract. Once this has been done, the employee has to decide whether to supply the work effort agreed upon or to shirk. In some experiments this is followed by an opportunity for the employer to punish or reward the employee. This type of experiment has been undertaken in various institutional settings. A major conclusion is that "... if both parties in a trade have the opportunity to reciprocate, reciprocal motivations have a robust and very powerful impact on the enforcement of contracts" [Fehr et al., 1997, p. 836]. [1]

There are many other experimental studies in which reciprocal behavior is observed [*e.g.* Fehr and Gächter, 1998]. It can be either negative [punishing uncooperative actions; *e.g.* Güth et al., 1982] or positive [rewarding cooperative actions; *e.g.* Fehr and Gächter, 1996]. The games in which reciprocity has been observed include public goods games [Brandts and Schram, 2000], the prisoners' dilemma [Andreoni and Miller, 1993, Cooper et al., 1996], centipede games [McKelvey and Palfrey, 1992], an investment game [Berg et al., 1995] and a gift exchange game [Gächter and Falk, 1997]. Most experimental studies are firm in their conclusions that some type of reciprocal behavior can be observed in the laboratory. In addition, reciprocity often seems to be a stable outcome, in which frequency and strength do not decline over time [Roth et al., 1991, Fehr et al., 1993, 1998, Cooper et al., 1996, Gächter and Falk, 1997].

Once the occurrence of reciprocity has been established, there is also room for strategic reputation building by individuals. Even those who would otherwise not act cooperatively, might do so in order to increase the probability of being reciprocated. Hence, reciprocity provides an explanation for cooperative behavior of individuals, for whom it is not in their (short-term) interest to cooperate. As a consequence, it provides necessary conditions for cooperative behavior to be stable in the long run and is therefore seen as an important

[1] Note that Fehr et al. refer to reciprocal *motivations*. Their conclusions are based on observed behavior, however.

mechanism in the evolution of cooperation in human societies [Trivers, 1971, Axelrod, 1984, Ridley, 1996, Binmore, 1998].

To a large extent the literature has focused on *direct* reciprocity, *i.e.*, a reaction to the acts of individuals one has interacted with before. There has been much less attention for *indirect* reciprocity, where a cooperative action is reciprocated by a third actor, not involved in the original exchange. The biologist Alexander [1987] argues that indirect reciprocity plays a central role in human societies. In his view, the link between actors is made through 'reputation' or 'social status'. Individuals in society are continuously being evaluated and reassessed with respect to how 'cooperative' they are. This gives them a reputation that may be used by others when deciding on how cooperatively to act towards them. We will discuss the literature on indirect reciprocity more extensively in the next section.

In this paper, we present experimental evidence showing that indirect reciprocity is an important phenomenon in the laboratory. The hypothesis tested is that people behave cooperatively to people who were cooperative towards others. Information about choices made by someone else does not require direct interaction, but can be obtained by observation. Hence, (providing) information about former behavior of the partner in a game is a way to control levels of cooperation. We use a 'repeated helping game' with random pairing in large groups, in which individuals either decide whether or not to help the subject they are matched with or are on the receiving end of this interaction. We observe that many decision makers respond to the information we give them about the 'social status' of the person they are matched with. This occurs even though they realize that this status is based on transactions with third parties and not on previous interactions with themselves. We use this result to study the consequences of indirect reciprocity for the dynamics of decision making in a group and the development of group norms.

The paper continues as follows. The following section discusses the theoretical literature on indirect reciprocity. Section III presents our experimental design and procedures. The results are presented in section IV, where we distinguish general results and a more detailed study of individual strategies in our experiments. The consequences for group dynamics and group norms are discussed in section V. Section VI summarizes and concludes.

II Models of indirect reciprocity

Since Axelrod ran his famous computer tournaments [Axelrod and Hamilton, 1981, Axelrod, 1984], the most commonly used framework for analyzing reciprocity has been the repeated Prisoner's Dilemma (PD) game. The tournament consisted of a sequence of repeated PD games between all possible pairs. The winning strategy turned out to be Tit for Tat (TFT), that is, start cooperatively in the first round and imitate (reciprocate) in all subsequent rounds. [2]

In a slightly different setup, introduced by Trivers [1971], two players act

[2] Although TFT was the winning strategy in Axelrod's tournaments, it is not an evolutionary stable strategy [Selten and Hammerstein, 1984] and not the winning strategy in all environments [Boyd and Lorberbaum, 1987, Foster and Young, 1990, Nowak and Sigmund, 1994]. However, reciprocal strategies, like TFT, perform better than 'all-defect' strategies in many theoretical and experimental environments and thus have at least some explanatory power in the evolution of cooperation.

in an *alternating* sequence of moves. A player makes either an altruistic or a non-altruistic choice. In this setup, conditional altruism is protected against exploitation, and both partners have an incentive to cooperate. This repeated alternating PD has been studied less intensively than the simultaneous PD, but it has been shown to have a cooperative evolutionary stable equilibrium [Nowak and Sigmund, 1993, Leimar, 1997].

In the literature, reciprocity is generally limited to direct reciprocity, requiring repeated encounters between the same two individuals. However, many authors have stressed that reciprocity does not need to be restricted to two individuals [Trivers, 1971, Sugden, 1986, Alexander, 1987, Binmore, 1992].

According to Alexander, who introduced the term indirect reciprocity, individuals use not only information from their own experience, but also react to interactions they observe between other individuals within their group.

> "In indirect reciprocity the return is expected from someone other than the recipient of the beneficence. This return may come from essentially any individual or collection of individuals in the group. Indirect reciprocity involves reputation and status, and results in everyone in a social group continually being assessed and reassessed by interactants, past and potential, on the basis of their interactions with others" [Alexander, 1987, p. 85].

Moreover, he calls indirect reciprocity the evolutionary basis of moral systems, which prescribe cooperative behavior.

An early experimental study including indirect reciprocity is Kahneman et al. [1986]. They study behavior in a reduced ultimatum game, where subjects knew the choices of their partners in a previous game with a third party. They find that subjects are more willing to split \$10 evenly with someone who was 'fair' in the past than to split \$12 with an 'unfair' allocator.

A first attempt to model reciprocity in larger groups was within small cyclical networks: individual A helps B, who helps C, who helps D, who in return helps A [Boyd and Richerson, 1989]. This type of indirect reciprocity has been studied experimentally in pension games by Heijden [1996]. However, it requires very strict interactions and is not based on reputation.

With the strategy *Observer Tit for Tat* Pollock and Dugatkin [1992] implement a notion of indirect reciprocity based on reputation. They study a sequence of repeated PD games with different pairs in which they allow players to occasionally observe a co-player's behavior with another partner before starting the repeated interaction. If the defection was observed in the last game with the other partner, Observer Tit For Tat prescribes defecting in the first round. This strategy outperformed the usual Tit for Tat and could even coexist with a subpopulation of defectors under certain circumstances.

The first authors who fully recognized the importance of Alexander's indirect reciprocity theory for the evolution of cooperation were Nowak and Sigmund [1998a,b]. In their main model, they use image scores, which are integer values in the range from -5 to 5, to describe the 'past level of cooperation' of individuals. Wedekind [1998] uses Alexander's term 'social status' when referring to this score. All individuals have their own score, starting at zero, but changing with decisions made in the game.

The game is played for several rounds. In each round two individuals are randomly chosen, one recipient and one donor. The donor decides whether or

not to give an amount b to the recipient at a cost $c < b$. When the donor gives to the recipient ('helps'), her or his score is increased by one point, otherwise (s)he loses one point. The decision whether or not to help may be based on the score of the recipient. Every individual is assumed to use a so called 'strategy k', where k is an integer value between -5 and 6. Using this strategy, donors cooperate if and only if the image score of the recipient is at least k. Strategy $k = -5$ thus describes unconditional cooperation, and strategy $k = +6$ describes unconditional defection.

The game is studied through simulation. To start, k-values are randomly distributed across the individuals. After 125 rounds have been played with alternating pairs, a new generation starts, with a distribution of k-values proportional to their payoffs in the previous generation. Nowak and Sigmund find that the whole population consists of strategy $k = 0$ after 166 generations, which is the most discriminating cooperative strategy.

Cooperative regimes also evolve when mutations are added, when interactions are only observed by some of the group members, or when individuals also care about their own score. An important requirement is that a sufficient proportion of the individuals is conditionally cooperative, *i.e.*, uses a strategy $k < 5$. If the entire population consists of self-regarding strategies ($k = 6$), no cooperative society can evolve.

An essential part of the Nowak and Sigmund studies is that they use replicator dynamics. The strategies are cut-point strategies based upon the scores of the recipients and/or on their own scores. Under these assumptions they are able to show that cooperation can evolve when at least some of the strategies are based on the score of the recipient.

It is this Nowak and Sigmund model of indirect reciprocity that we test experimentally in this paper[3]. From earlier experiments we know that direct reciprocity leads to high levels of cooperation in interaction between fixed partners. People behave nicer to people who were nice to them. The new hypothesis tested in this paper is that people also behave nicer to people who were nice to others. [4]

III Experimental procedures and design

The experiment consisted of 6 sessions at the CREED laboratory of the University of Amsterdam. Subjects voluntarily signed up after public announcement. Participants were students from various departments including economics, social sciences, law, chemistry and computer sciences. In each session 28 subjects participated simultaneously. They were randomly assigned to seats in the laboratory, which are separated by partitions. No communication was allowed. Instructions were computerized and could be read at one's own pace (see the appendix for an English translation). A quiz was used to ensure that subjects

[3] In this model, various strategies can be supported in equilibrium under the assumption of full rationality [*e.g.* Kandori, 1992, Ellison, 1994]. We are not interested in testing any specific equilibrium prediction, but in studying the evolution of cooperative behavior in our experiments.

[4] A first experimental test of the Nowak and Sigmund model was developed independently of our study by Wedekind and Milinski [2000]. In a game of only 6 rounds, they find evidence in support of indirect reciprocity. Our experiments allow for a much more extensive analysis. In a completely different setting, Dufwenberg et al. [2000] find evidence of indirect reciprocity in an investment game.

understood the instructions. When all subjects had finished reading the instructions the experiment started. The experiment itself was also computerized and consisted of at least 90 rounds. Thereafter, any additional round was started with a common knowledge probability of 90%. The duration of one session was about 75 minutes. There was a show up fee of either 20 or 30 guilders, depending on the specific parameters in the session concerned (at the time of the experiments, 1 guilder $\approx$ \$0.50). In addition, the subjects were told that they would be paid the earnings (which could also be negative) of 20 rounds that would be randomly chosen at the end of the experiment. Therefore, at any point, subjects did not know exactly how much they had earned. This design was used to avoid changes in behavior as a result of income effects. [5] On average subjects earned 35.37 guilders. At the end of the experiment, subjects were asked to fill in a questionnaire concerning personal background and motivation.

In each round, pairs of subjects were chosen, consisting of one donor and one recipient. Both the roles and the pairing were determined completely at random. Subjects were told that they could be randomly matched with any of the other 27 participants in the laboratory. In fact, however, the randomization was done in two separate groups of 14 subjects in order to increase the number of independent observations.

Our setting is a 'helping' experiment. In any given round, only the donor has to make a decision. (S)he has the choice of either 'helping' the recipient at a cost c, in which case the recipient receives a benefit $b > c$, or 'passing', in which case both individuals receive 0. After the decisions are made the recipients are informed about the decision of the donor they were matched with. They do not receive any information other than this outcome.

The choices were presented to the subjects in a payoff table with yellow and blue representing the choices 'help' and 'pass', respectively. The payoff values for the case where the donor chose yellow (help) are presented in table I. Note that the benefit b for the recipient was always 250 cents, whereas the cost varied across sessions. Two values were used: $c = 50$ cents in the low-cost (LC) condition and $c = 150$ cents in the high-cost (HC) condition.

Besides the costs, we also varied the amount of information given to the donor. In a baseline condition without information, donors were not told anything about the previous choices of the recipient. This is referred to as the No-Information (N) condition. This baseline was only used in the HC condition. In the Information (I) condition the notion of social status was implemented in the following way. Before making their choice, donors were given information about the previous 6 decisions made by the recipient when (s)he had been a donor. This information was summarized in two numbers, the number of yellow (help) choices and the number of blue (pass) choices. Only the number in the last 6 choices was given, not the order. All in all, we ran three treatments: LCI; HCI; and HCN (*cf.* table I). Because we organized two sessions each and distinguished two groups of 14 in each session as discussed above, we have four (statistically independent) groups per treatment. Within each group we have 14 subjects[6], each making about 45 to 50 decisions.

Note that information about 6 decisions (on average 12 rounds) reflects a limited memory, as often used in evolutionary game theoretical models [Young,

[5] Of course, the implicit assumption here is that the difference in show up fee does not cause an income effect. We did not find any evidence of such an income effect.

[6] In two groups of HCN we had only 10 subjects due to no-shows.

Table I: Parameters used

Treatment	Benefit b for recipient	Cost c for donor	# of previous choices of recipient shown to donor	Show-up fee
LCI	250	50	6	2000
HCI	250	150	6	3000
HCN	250	150	0	3000

Note: The numbers referring to costs and benefits represent the payoffs (in Dutch cents) in case the donor chooses yellow (help). The payoff for both the recipient and the donor is zero in case the donor chooses blue (pass).

1998]. This gives subjects an opportunity to 'clean' their record. We chose a limited memory in order to decrease the influence of early rounds and to keep both the impact of one single decision and the information level constant.

In all sessions individuals were given a summary of their own last 6 choices and they were shown their own outcomes of all previous rounds. Both were displayed on the screen continuously. No summary statistics or information about the decisions of others (except the recipient) were provided. We did not provide information about the donor to the recipient in order to avoid direct reciprocity towards subjects with the same score and to minimize the information flow about the strategies of others. In this setup, subjects can learn about what others do, because they interact with someone else every round, but they cannot infer how a donors decision is related to her or his own score.

IV Results

The results are presented in two subsections. First, we present a general overview and analysis of helpful behavior, both over rounds and across treatments. In addition, we provide some data reflecting the relationship between individual choice and the social status of the receiver and donor. Then, we elaborate on this relationship by analyzing individual strategies in more depth. These strategies may depend on the social status in various ways.

1 General Results

We find a large proportion choices to be cooperative (helping) in all sessions. The percentage (in the first 90 rounds) is 86 % in LCI, 70 % in HCI and 22 % in HCN. Figure I presents the fraction of helpful choices per round (up to round 90) for the three treatments.

This figure shows various things. First of all, there is a clear order in helping. The highest fraction of helpful choices is observed in LCI and the lowest in HCN. Second, LCI and HCI show some signs of an end effect, with helpful choices starting to drop around round 80. Nevertheless, *help* is still being chosen in more than 50 % of the cases (except for the final two rounds in HCI). No end effect is observed in HCN.

Besides differences in the average fraction of helpful choices across treat-

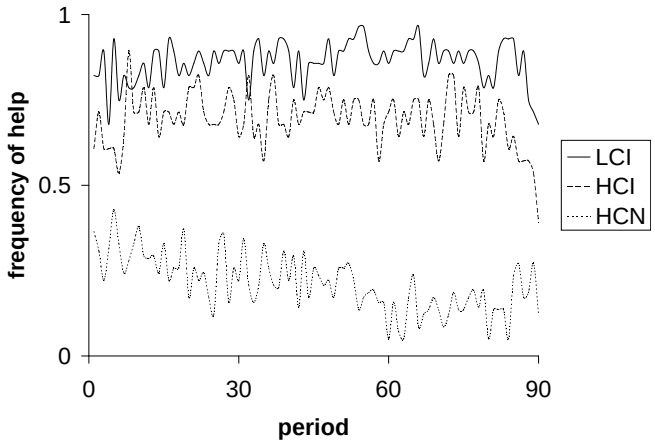


Figure I: Average fraction of helpful choices per treatment.

Table II: Helpful choices per group.

Treatment	groups				total
LCI	0.74	0.89	0.90	0.92	0.86
HCI	0.49	0.66	0.77	0.86	0.70
HCN	0.18	0.19	0.22	0.27	0.22

Note: The numbers in the cells are the fraction of helpful choices in the first 90 rounds. Groups are shown in increasing order of helpful choices from left to right.

ments, there are also important differences across groups within a treatment. Figure II shows the 7-round moving average of choices of each of the four groups of 14 subjects per treatment.

Figure II shows that different groups have distinct dynamics. This is especially clear in HCI. Note that this makes it unreasonable to assume statistical independence of choices within a group. To undertake testing, we therefore need to summarize group statistics. Table II presents the fraction of helpful choices per group for each treatment.

Using the four averages per treatment as four independent observations, we conducted a Mann–Whitney test for a pair wise comparison across treatments. The results show that the differences LCI-HCN and HCI-HCN are significant at the 5 %-level and the difference LCI-HCI is significant at the 10 % level. Therefore, (even with such a conservative test) we find treatment effects, where both the level of costs and the provision of information affects choices.

Next, we turn to the influence of the social status on the choices made by donors. Recall that this is reflected in the information we provide to the donor about the six previous choices of the recipient. Because the number of helpful choices in LCI does not show enough variation for a fruitful analysis, we will not analyze this treatment.

Social status may matter in two ways. First we will have a look at the influ-

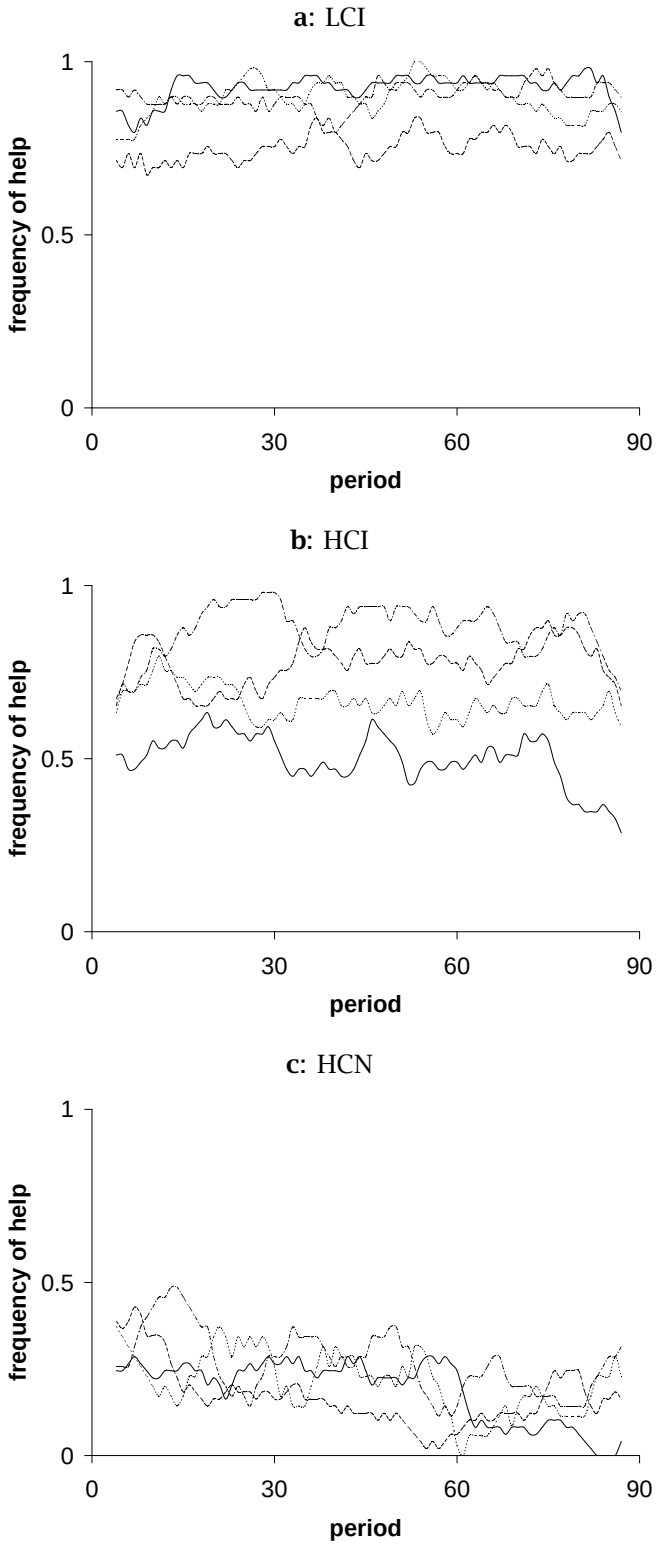


Figure II: Moving average of helping fraction per group.

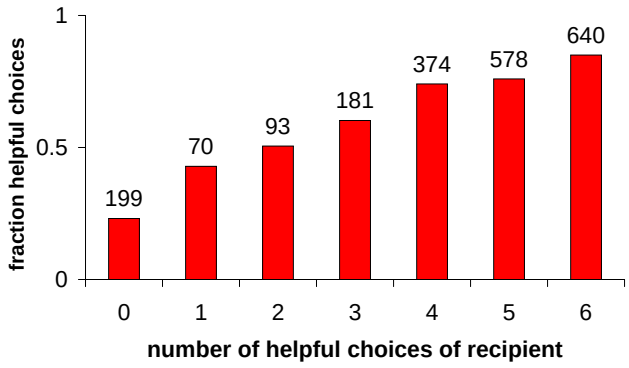


Figure III: The fraction of helpful choices in the first 90 rounds is shown, dependent on the previous 6 decisions of the recipients. (Early) rounds where either the recipient or the donor had made less than 6 previous decisions are not included. The number of observations per category is shown above the respective bars.

ence that the social status of recipients has on the decision making of donors.

Figure III shows the fraction of helpful choices as a function of the recipient's social status in HCI (note that this is not observed in HCN). Social status may vary across 7 categories: from 0 (zero help and six pass) to 6 (six help and zero pass) in the previous six choices. The figure shows that subjects with a high score are almost always helped (reciprocated) whereas individuals who never choose to help are helped about 25% of the time. Apparently it matters what the social status of the recipient is. This is a first indication of indirect reciprocity in subjects' behavior. Note that this behavior is not likely to be strategically aimed at direct reciprocity. Subjects believe that the probability that they will meet their partner again in any given round is $1/28$. This probability is much too low to motivate any kind of decisions in the hope of receiving a direct reciprocal response from the same partner.

Given the numbers underlying figure III, one can estimate the expected return for various levels of social status. It turns out that the highest return (85 cents per round) is to be expected from keeping a score of 4 out of 6 and the lowest (58 cents per round) from keeping a score of 0 (unconditional defection).

To test the influence of indirect reciprocity, we conducted a Page test [Siegel and Castellan, 1988] for ordered alternatives on the rankings from 1 to 7 for each of 7 categories in the 4 groups in HCI. The rankings were based on the mean fraction of helpful choices towards recipients with a score 0 to 6. The null hypothesis was that the average rank in each of the categories were the same. The alternative hypothesis is that the average rank increases from category 1 to 7 (all differences are inequalities and at least 1 difference between 2 successive categories is a strict inequality). The relation between the score of the recipient and the fraction of helpful choices turned out to be significant at the 0.001 level.

If social status of the recipient is important in deciding whether to help or not, then a donor might consider the effect that her or his own score will have on

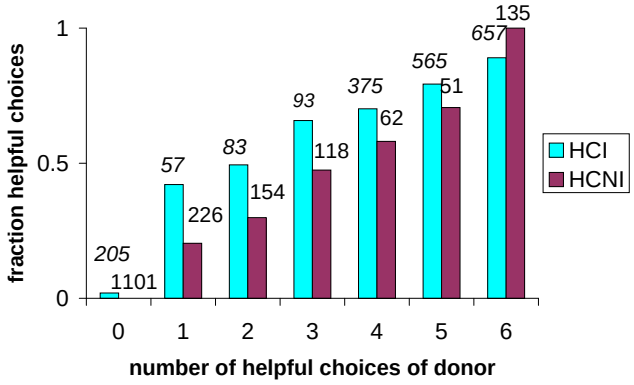


Figure IV: The fraction of helpful choices in the first 90 rounds is shown, dependent on the previous 6 decisions of the donor. (Early) rounds were either the recipient or the donor had made less than 6 previous decisions are not included. The number of observations per category is shown above the respective bars.

future donors when (s)he her- or himself will be recipient. For HCN this should not matter, because the donor knows that no information will be passed on to others. This is reflected in the data presented in table II. The average level of helping is much higher in case the information is passed on. Figure IV shows the fraction of helpful choices as a function of the own social status. Social status may again vary from 0 (zero help and 6 pass) to 6 (six help and zero pass). The figure for HCN shows a uniformly increasing fraction. This simply reflects some stability in individual choices: people who chose *help* more often at the previous 6 opportunities are more likely to do so now.

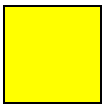
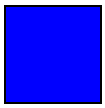
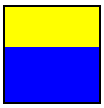
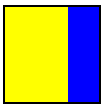
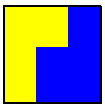
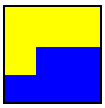
The own social status plays a much more important role in the decisions of donors in HCI than in HCN. Besides from the total fraction of helpful choices, this is also apparent from figure IV. Subjects are much more likely to help if they know that their score will be passed on.

The numbers underlying figure IV provide an indication of the dynamics of helping behavior during the experiment. For example, if the fraction of helpful choices with a score of 1 is higher than 1/6, then there is a drift towards higher fractions. Figure IV reflects a drift towards lower fractions for HCN, except by those who chose to help at all previous 6 opportunities (score of 6). HCI shows a tendency towards behavior where donors choose to help in 4 or 5 out of 6 rounds. Recall that we concluded from the analysis of figure III that the highest expected payoffs are obtained with a social status of 4. Hence, the drift in case information is passed on is to the score that maximizes the expected gains.

2 Individual strategies

In this section, we will study individual strategies in more detail. We will allow these strategies to depend on the own social status and/or the social status

Table III: The individual strategies we distinguish: For each of the six strategies, the x-axis shows the own social status and the y-axis the social status of the recipient. Light grey denotes helping and dark grey denotes passing. Strategy 5 also includes strategies with a minimum own score, but no maximum own score (always *help* when the own score is equal to or below k_i; also *help* when the own score is above k_i and the recipient's score is above k_j). Strategy 6 also includes strategies for which no score of the recipient is high enough to ensure help (*help* when the own score is below k_i *and* the recipient's score is at least k_j).

Strategy	1	2	3	4	5	6
Description	unconditional		conditional cooperation with helping decision based on k('s) of:			
	cooperation	defection	recipient	donor	both recipient and donor	
k recipient / k donor						
group 1	3	3	2	0	3	3
group 2	3	2	4	2	2	1
group 3	3	1	4	1	2	3
group 4	11	0	0	0	3	0

of the recipient. [7] Again, we focus on the HCI sessions, because LCI does not provide enough variation in strategies and HCN does not provide donors with information about the score of the recipient. We describe strategies using a figure where the number of own previous helpful choices (out of 6) is given on the horizontal axis and the number of previous helpful choices by the other is given on the vertical axis. Table III shows the strategies we distinguish.

In this table, light gray areas refer to helping and black areas to passing. Hence, strategy 1 shows a strategy of unconditional cooperation (helping independent of the own or the other's social status) and strategy 2 describes unconditional defection (never helping). Strategy 3 reflects a cut-point strategy of cooperation if and only if the other has helped enough recipients in the past, whereas strategy 4 reflects a cut point with respect to the own social status (keeping one's own score at some specified level). Finally, strategies 5 and 6 use both the own score and the score of the recipient. In strategy 5, the donor ensures an own score at a certain level but the willingness to help beyond that depends on the score of the recipient. In strategy 6, the donor never helps a recipient whose score is too low, but is willing to help others, depending on the scores of both. The numbers in table III refer to the number of participants in the various HCI groups estimated to use the respective strategy. This was simply determined by the number of choices that could be explained by the various strategies in the first 60 rounds, excluding the rounds in which less than 6 decisions had been made. If two strategies could explain an equal number of choices, the more 'simple' one (*i.e.* with the fewest parame-

[7] An obvious naive strategy would be to consider the effect of the outcome in the previous round. However, this is only found to be significant in 4 individuals and therefore ignored in this analysis.

Table IV: Predictive power of strategies

Group	Explained choices rounds < 61	Predicted choices rounds 61–80	Baseline predictions
1	0.92	0.83	0.63
2	0.90	0.81	0.85
3	0.89	0.85	0.81
4	0.93	0.88	0.86
Total	0.92	0.84	0.79

Note: The third column presents the fraction of correct predictions for rounds 61–80 based on the strategies calibrated on rounds ≤60. The fourth column shows the fraction of correct predictions based on strategies 1 and 2: if *help* (*pass*) was chosen in the majority of cases in rounds 1–60, the prediction for 61–80 is to always choose help (pass).

ters) was chosen. [8] The strategies estimated involve unconditional cooperation (36%); unconditional defection (11%); conditional cooperation based on the recipient's social status (18%); conditional cooperation based on the own social status (5%); conditional cooperation primarily based on the own social status, but also on the recipient's social status (18%) and conditional cooperation primarily based on the recipient's social status, but also on the own social status (12%). These individual strategies were used to predict individual choices in rounds 61–80. Table IV presents the results of this exercise.

The fraction of choices in rounds 1–60 that can be explained when allowing for cut-point strategies (3–6) is significantly higher than the fraction explained when restricting strategies to choosing *help* or *pass* only (1 or 2). The total improvement relative to this benchmark is 11%-points. [9]

Table IV shows that the strategies can predict out of sample quite accurately. On average, the choice made is predicted 84% of the time in rounds 61–80. As a benchmark, we again restricted the strategy set to 1 and 2. In this benchmark, a donor is predicted to always help (pass) in rounds 61–80 if the observed fraction of *help* in rounds before 61 is larger (smaller) than 0.5. The improvement from the benchmark to the full strategy analysis is 5%-points.

Unfortunately, this analysis of individual strategies did not always provide unique scoring rules for individuals. In some cases, different scores could explain observed behavior equally well, mainly because not all individuals were confronted with every score. [10] For the 27 subjects (in four groups) estimated to use strategy 3, 5 or 6, we estimate the average cut point for the recipient's

[8] Strategies 1 and 2 are considered to be more simple than 3 and 4; 3 and 4 are considered to be more simple than 5 and 6.

[9] To determine the extent to which this improvement can simply be attributed to increasing the number of parameters, we used a randomization procedure to determine the fraction of choices that can be explained by adding three parameters, dividing the strategy space in 6 sections. Reshuffling *help* and *pass* per individual across rounds and calculating best strategies 10000 times shows an increase of 4%-points at the 0.001 level. Hence, the increase in explanatory power we observe is not simply a consequence of adding parameters.

[10] Consider an individual who always helped recipients with a score of 4 or higher and never helped recipients with a score of 1 or lower. If (s)he never met a subject with a score of 2 or 3, his or her cut point can be 2, 3 or 4.

status to lie in the interval {(1.7–2.4), (2.1–2.7), (2.3–3.2) and (4.0–5.0)} respectively (groups are ordered in fraction of helpful choices). For the predictions of choices in rounds 61–80 we used the maximum of this interval. The scoring rules of the own social status show less ambiguity. The average cut point for the own social status of the 20 subjects with strategy 4, 5 or 6, is estimated as {3.5–3.8, 4, 4.8, 5.3}. [11] Note that both cut points seem to increase with the level of observed cooperation in the group. This is elaborated in section V.

V GROUP DYNAMICS AND GROUP NORMS

In the previous section we analyzed individual strategies and showed that many strategies can be characterized by minimum score rules (cut-point strategies) for the recipient and/or the donor. The minimum score for the recipient can be seen as a norm that (s)he has to satisfy. We saw that the estimated norms differ across groups. In this section we will focus on the development of these norms across rounds and their relation to group composition. We will argue that individual norms are at least partly determined by group composition, developing similarly within, but distinct across groups. This leads to the emergence of 'group norms'. We use the term 'group norms' to indicate a correlation of individual norms within a group. These individual norms refer to the social status of the recipient that is demanded to help her or him. A consequence of distinct group norms across groups is that a specific social status may be sufficient to induce helpful behavior in some groups but not in others.

The possibilities for learning are quite limited in our setting. Experimenting with one's own score is possible, but time consuming and the outcome will differ across subjects, because every individual is matched with a different set of others. Subjects do not know how often others are helped, and recipients do not know what type of donor (high status or low status) they are matched with. Therefore they cannot imitate behavioral strategies of (successful) others. The only thing they can try to copy is the score of recipients or some measure (*e.g.* mean or modus) of the recipients they are matched with.

From the information they receive about the scores of recipients, subjects can estimate a distribution of scores within their group. They will learn what score can be considered to be low and what score can be considered to be high. If they are more helpful to subjects with a high social status than to subjects with a low status (*e.g.*, if they use a strategy type 3, 5, or 6), their own score will be affected by the relative frequency of encounters with high and low status recipients. Taking this into account, subjects using a norm for the recipient's social status can determine a maximum norm k_i they can use for this status, that will allow them to keep their own score $s_i \geq k_i$. In other words, subjects using strategies 3, 5 or 6 can learn to adjust their cut point in a way that their own score adheres to it. It is this type of learning, finding a norm that is consistent with the own social status, that we think is important in synchronizing norms within a group, *i.e.*, creating group norms. Our data provide some evidence that this type of learning is taking place.

[11] For two individuals with strategy 6 the own minimum score is not well defined, because the interval between the two cut points of the recipient's score is very small. There were no observations in this interval in rounds 61–80.

As shown above, the groups within the HCI treatment have distinct dynamics (figure II) and vary in the average fraction of helpful choices (table II). To investigate whether these differences across groups exist from the beginning or develop during the experiment, we use two tests. For all rounds aggregated in groups of 10, we test (i) whether the fraction of helpful choices differs across groups and (ii) whether the fraction of helping per individual is different across groups. Hence, to test (i) the unit of observation is the choice whereas in (ii) it is the individual.

Using a χ^2 test, we find significant differences in the fraction of *help* across groups in each 10-round block (test i), although the difference is only marginally significant in the first block of 10 rounds (first 10 rounds $\chi^2 = 7.13, p = 0.068$, later rounds $\chi^2 > 14, p \leq 0.003$). When comparing the helping fraction per individual during 10 rounds using a Kruskal–Wallis test (ii), groups appear to be different from the third 10-round block onwards (first 20 rounds $p > 0.2$; later rounds $p < 0.05$).

The difference between the two test results is an indication that differences across groups in early rounds (< 20) are concentrated in a few individuals, whereas in later rounds a larger group of individuals differ in their frequency of *help*. This could imply that the interaction between subjects changes individual behavior and therefore influences the dynamics within groups.

Our results from the individual strategy analysis are in agreement with these findings. Of the 56 subjects in HCI, 27 were found to use a strategy based on the score of the recipient (strategy 3, 5 or 6). These subjects behave less helpful towards subjects with a low status then towards subjects with a higher status. As a consequence, in a group with more defectors (strategy 2), these conditional cooperators will have to punish more frequently and therefore obtain a lower social status themselves. We might therefore expect a negative correlation between the helping frequency of conditional cooperators in a group and the number of unconditional defectors (strategy 2). This is indeed the case: for the 4 groups in order of helping frequency, the fraction of helpful choices by conditional cooperators (strategies 3, 5 and 6) are {0.59, 0.74, 0.79, 0.89} and the numbers of unconditional defectors were found to be: $n_1 = 3, n_2 = 2, n_3 = 1, n_4 = 0$. Recall that the norms estimated in the previous section were also shown to be increasing in the frequency of help. If the latter were not the case, the lower level of helping when there are more unconditional defectors could make the conditional cooperators fall below there own norms. In turn, this would make them punish each other, and cause a reduction in helpful choices towards zero. Note that this is not observed.

To study the relationship between behavior and group norms in more detail we use data from the post-experimental questionnaire in which 68% reported having used a minimum score rule for the recipient. Note that in order to test for changes in behavior we are forced to use the cut points reported in the questionnaire as opposed to the ones estimated from the observed choices. One cannot test behavior using a categorization based on that behavior. The average cut points $\bar{k}$ reported in the four groups are: $\bar{k}_1 = 2.4, \bar{k}_2 = 3.4, \bar{k}_3 = 3.8, \bar{k}_4 = 4.3$. First note that these cut points are similar to the cut points found in the individual strategy analysis, where $\bar{k}_1 = 2.4, \bar{k}_2 = 2.7, \bar{k}_3 = 3.2, \bar{k}_4 = 5$, as reported in section 2. (We use the upper limit of the estimated interval, because this is more closely related to the question in the questionnaire)

Table V: Donors reported norm k_r, compared with social status s_r of recipients they were matched with.

rounds	$s_r > k_r + 1$	$s_r = k_r + 1$	$s_r = k_r$	$s_r = k_r - 1$	$s_r < k_r - 1$
	a: fraction of *help*				
≤ 20	0.82	0.77	0.76	0.90	0.53
21–40	0.90	0.84	0.88	0.74	0.29
41–70	0.88	0.90	0.85	0.59	0.22
71–90	0.87	0.86	0.64	0.50	0.14
	b: fraction recipients				
Total	0.36	0.28	0.17	0.07	0.13

Second, it is noteworthy that 70% of the conditional cooperators report using the mode cut point in their group: $\hat{k}_1 = 2, \hat{k}_2 = 3, \hat{k}_3 = 4, \hat{k}_4 = 4$. Finally, we once again see that there are large differences across groups. A Kruskal–Wallis test shows that groups are significantly different at the 0.01 level. (From individual comparisons ($\alpha = 0.05$) we can conclude that $\bar{k}_4 > \bar{k}_1$; $\bar{k}_4 > \bar{k}_2$ and $\bar{k}_3 > \bar{k}_1$).

Reported norms are consistent with the donor's own status. A lower own status corresponds to a lower norm. Up to round 90 subjects can keep their own status above the norm they report using for recipients 88% of the time. In case they had used a higher norm than the reported one, this would have led to an average own status below that norm in all sessions. In that case, on average, conditional cooperators would punish each other for punishing defectors, which would lead to a cascade of passes until nobody helps anymore.

Next, we study the development of behavior over time. To do so, we analyzed the data for different time intervals separately. In all 4 groups we find the same pattern: 79–84% of all choices can be explained by the reported cut points between rounds 20 and 90 (68% in earlier rounds). Distinguishing recipient scores above and below the cut point, 88% of the time subjects help recipients with a social status s_r equal to or higher than their cut point k_r (79% in earlier rounds). When matched with recipients with $s_r < k_r$, donors help 31% of the cases (69% in earlier rounds).

For the subjects with a reported cut point k_r, table V provides data about the relation between k_r and the scores s_r of the recipients they were matched with. In table Vb we show the fraction of recipients with s_r lower than, equal to or higher than k_r. It shows that a large majority (80%) of the subjects is matched with recipients having a score greater than or equal to the cut point. Table Va gives the fraction of helpful choices made by donors matched with recipients with various s_r's.

Table V shows a number of interesting patterns. First of all, in the first 20 rounds, the reported cut point does not predict choices very well. The highest frequency of *help* is found for the situation where the recipients have a score that is just below the cut point, *i.e.*, when the cut-point strategy predicts that the donor will not help. This is in line with our observation that cut points are adjusted early on in order to allow the own status to fit the own norm. Second, after round 20, levels of helping are increasing in the social score of the

recipient (with only two exceptions). Hence, the norm qualitatively describes the donors' choices. Third, after round 70 the fraction of *help* declines for any social status of the recipient. At this stage of the experiment, subjects are less inclined to help in general. Reciprocation is losing some of its impact.

Even in rounds 21–70, the cut points are far from perfect. Especially the high fractions of helping when the recipient has a score that is 1 below the reported norm might seem surprising (note, however, that this score is only observed 7% of the time). This can be explained by considering the strategies distinguished in the previous section. Many of the subjects reporting a cut point for the recipient's status were found to (also) use a cut point for their own score. This might be especially relevant in cases where the recipient's status was 'close' to the norm. In that case, the own score might become the dominant concern, as in strategy 6. In future experiments, we hope to use a design that will allow us to discover more subtleties in individual strategies.

From these observations, we conclude that the reported cut point most closely refers to the norm used in rounds 21–70. In earlier rounds, cut points have not been established and in later rounds they play a less important role.

VI Summary and conclusions

This paper has presented clear evidence that indirect reciprocity plays an important role in our experimental setting. This setting was based on recent theoretical literature stressing the importance that indirect reciprocity can have in large groups. In the (experimental) economic literature, the main focus has been on direct reciprocity. There is an abundance of evidence showing the existence of behavior that might be interpreted as direct reciprocity. This paper has shown that indirect reciprocity may be an equally important phenomenon. This behavior may lead to stable cooperative regimes, in which the majority of the population helps.

In our experiments, indirect reciprocity shows up in the donor's sensitivity to the recipient's social status. It is therefore rational for individuals to build up a good reputation. This is indeed observed in our data: the own social status is important for the decision whether or not to help, when this status is passed on to future donors. Both when estimating strategies and when analyzing responses to the questionnaire, we observed that many subjects used cut points with respect to their own and the recipient's status. In the strategy analysis we distinguished 6 strategies: always help, never help, a cut-point rule based on the social status of the recipient, a cut-point rule based on the own social status and two combinations of these cut-point rules (see table III). We found that 54% of the subjects used a cut-point strategy. Of these subjects, 48 %-point used the social status of the recipient when making their decision and 36 %-point used their own social status. The fraction of subjects basing their decision on the social status of the recipient was high enough to make investment in the own status worthwhile. That is, given the cut points used, the expected payoffs from keeping the own score high enough was larger than the expected payoff from never helping.

However, the cut points differ across groups. As a consequence, a good reputation in one group does not have to be good enough to be helped in another group. The cut points or norms develop in an early stage of our experiment

and are related to the distinct composition of strategy types across groups. We argued that conditional cooperators need to pass more often when they encounter unconditional defectors and therefore need to lower their own social status. The norms that subjects use for the recipient are related to their own social status. Norms are lower or equal to the own social status 88 % of the time. Subjects lower the norm they use for the recipient when they cannot obtain that status themselves. This explains why social statuses within groups differ less then across groups.

Our experiments are not intended to provide insights into the motivations underlying reciprocity. These motivations may be related to other regarding preferences. If this is the case, it is not quite clear why subjects would use cut points with respect to the recipient's social status, however. Take fairness, for example. If this is the motivation driving behavior, a subject is interested in the own earnings and the relationship between the own earnings and those of the other. To some extent this might explain why a subject would help a recipient with a high status. This status indicates that the recipient has helped in the previous six rounds and, therefore, has earned less than if (s)he had passed. The information is limited, however. First of all, information is only provided about the recipient's earnings in six rounds. Second, no information is provided about the previous rounds were the recipient had been a recipient. One result that is difficult to explain with fairness is that subjects help more when the status is passed on than when it is not. Models of fairness that include the possibility of reciprocity, as developed by Levine [1998] for example, may be better suited to explain the behavior we observe. Though once again, additional assumptions are needed to explain the differences between sessions with and without information. Our conclusions about motivations are restricted to this type of general findings. Other experimental designs are needed for further insights.

This study is meant to be a starting point in the understanding of how cooperative regimes can evolve in large societies. The economic consequences of indirect reciprocity and endogenous norms are obvious and warrant further research in this area. We intend to continue this research by studying the development of norms related to the donor's and the recipient's status in more detail. By applying a strategy method and restricting information flows, we hope to gather more information about the way in which norms are adjusted to observed behavior within the group.

Acknowledgements
We thank Arno Riedl and Joep Sonnemans for inspiring discussions about the experimental design as well as for useful suggestions for the paper and Sido Mylius and Peter Wakker for critically reading the manuscript.

REFERENCES

Alexander, R.D. *The Biology of Moral Systems*. Aldine de Gruyter, New York, 1987. 3, 4

Andreoni, J. and J.H. Miller. Rational cooperation in the finitely repeated pris-

oner's dilemma: experimental evidence. *The Economic Journal*, CIII (1993): 570–585. 2

Arrow, K. *The Limits of Organization*. Norton, New York, 1974. 1

Axelrod, R. *The Evolution of Cooperation*. Basic Books, New York, 1984. 3

———, and W.D. Hamilton. The evolution of cooperation. *Science*, CCXI (1981): 1390–1396. 3

Berg, J., J. Dickhaut, and K. McCabe. Trust, reciprocity and social history. *Games and Economic Behavior*, X (1995):122–142. 2

Binmore, K.G. *Fun and Games: a Text on Game Theory*. Heath and Co., Lexington, MA, 1992. 4

———, *Game Theory and the Social Contract, II: Just Playing*. MIT Press, Cambridge, MA, 1998. 3

Bolton, G.E. and A. Ockenfels. A theory of equity, reciprocity, and competition. *American Economic Review*, XC (1992):166–193, 2000. 2

Boyd, R. and J.P. Lorberbaum. No pure strategy is evolutionarily stable in the repeated prisoner's dilemma. *Nature*, CCCXXVII (1987):58–59. 3

Boyd, R. and P.J. Richerson. The evolution of indirect reciprocity. *Social Networks*, XI (1989):213–236. 4

Brandts, J. and A. Schram. Cooperative gains and noise in public goods experiments; applying the contribution function approach. *Journal of Public Economics*, VII–IX (2000):399–427. 2

Cooper, R., D.V. DeJong, R. Forsythe, and T.W. Ross. Cooperation without reputation: Experimental evidence from prisoner's dilemma games. *Games and Economic Behavior*, XII (1996):187–218. 2

Dufwenberg, M., U. Gneezy, W. Güth, and E. van Damme. An experimental test of direct and indirect reciprocity in case of complete and incomplete information. forthcoming in *Homo Oeconomicus*, 2000. 5

Ellison, G. Cooperation in the prisoner's dilemma with anonymous random matching. *Review of Economic Studies*, LXI (1994):567–588. 5

Fehr, E. and S. Gächter. Cooperation and punishment — an experimental analysis of norm formation and norm enforcement. Forthcoming in the *American Economic Review*, 1996. 2

———, and ———,. Reciprocity and economics: The economic implications of *Homo Reciprocans*. *European Economic Review*, XLII (1998):845–859. 2

Fehr, E., S. Gächter and G. Kirchsteiger. Reciprocity as a contract enforcement device: experimental evidence. *Econometrica*, LXV (1997):833–860. 2

Fehr, E., E. Kirchler, A. Weichbold, and S. Gächter. When social norms overpower competition — gift exchange in experimental labour markets. *Journal of Labor Economics*, XVI (1998):324–351. 2

Fehr, E., G. Kirchsteiger, and A. Riedl. Does fairness prevent market clearing? An experimental investigation. *Quarterly Journal of Economics*, CVIII (1993): 437–460. 2

Fehr, E., and K.M. Schmidt. Fairness, incentives, and contractual choices. *European Economic Review*, XLIV (2000):1057–1068. 2

Foster, D. and H.P. Young. Stochastic evolutionary game dynamics. *Theoretical Population Biology*, XXXVIII (1990):219–232. 3

Gächter, S. and A. Falk. Reputation or reciprocity? Working paper, University of Zürich, 1997. 2

Güth, W., R. Schmittberger, and B. Schwarze. An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, III (1982): 367–388. 2

Heijden, E.C.M., van der. *Altruism, Fairness and Public Pensions*. PhD thesis, Tilburg University, The Netherlands, 1996. Chapter 6. 4

Kahneman, D., J.L. Knetsch, and R.H. Thaler. Fairness and the assumptions of economics. *Journal of Business*, LIX (1986):285–300. 4

Kandori, M. Social norms and community enforcement. *Review of Economic Studies*, LIX (1992):53–80. 5

Leimar, O. Repeated games: a state space approach. *Journal of Theoretical Biology*, CLXXXIV (1997):471–498. 4

Levine, D. Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics*, I (1998):593–622. 2, 18

McKelvey, R.D. and T.R. Palfrey. An experimental study of the centipede game. *Econometrica*, LX (1992):803–836. 2

Nowak, M.A. and K. Sigmund. The alternating prisoner's dilemma. *Journal of Theoretical Biology*, CLXIV (1993):471–489. 4

———, and ———, A strategy of win-stay, loose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, CCCLXIV (1994):56–57. 3

———, and ———, The dynamics of indirect reciprocity. *Journal of Theoretical Biology*, CXCIV (1998a):561–574. 4

———, and ———, Evolution of indirect reciprocity by image scoring. *Nature*, CCCXCIII (1998b):573–577. 4

Pollock, G.B. and L.A. Dugatkin. Reciprocity and the evolution of reputation. *Journal of Theoretical Biology*, CLIX (1992):25–37. 4

Rabin, M. Incorporating fairness into game theory and economics. *American Economic Review*, LXXXIII (1993):1281–1302. 2

Ridley, M. *The Origins of Virtue: Human Instincts and the Evolution of Cooperation*. Penguin Books Ltd, London, 1996. 3

Roth, A., V. Prasnikar, M. Okuno-Fujiwara, and S. Zamir. Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh and Tokyo. *American Economic Review*, LXXXI (1991):1068–1095. 2

Schram, A. Sorting out the seeking: The economics of individual motivations. *Public Choice*, CIII (2000):231–258. 2

Selten, R. and P. Hammerstein. Gaps in Harley's argument on evolutionarily stable learning rules and in the logic of "tit for tat". *Behavioral Brain Sciences*, VII (1984):115–116. 3

Siegel, S. and N.J. Castellan, Jr. *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill Book Company, Singapore, 2nd edition, 1988. 10

Sugden, R. *The Economics of Rights, Cooperation and Welfare*. Blackwell, Oxford, 1986. 4

Trivers, R. The evolution of reciprocal alruism. *Quarterly Review of Biology*, XLVI (1971):35–57. 2, 3, 4

Wedekind, C. Give and ye shall be recognized. *Science*, CCLXXX (1998):2070–2071. 4

Wedekind, C. and M. Milinski. Cooperation through image scoring in humans. *Science*, CCLXXXVIII (2000):850–852. 5

Young, P.H. *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton University Press, Princeton, New Jersey, 1998. 7

APPENDIX

INSTRUCTIONS EXPERIMENT

[*What follows is a translation of a sample of the instructions of the experiment.*[12]]

Introduction

Welcome to this experiment in decision making. In this experiment you will earn money. At the start of the experiment you will receive 30 guilders. During the experiment you may win or lose money dependent on your own decisions and the decisions of other participants. At the end, your final earnings will be paid to you privately. Your decisions are anonymous. They will not be attached to your name. [...] You are not allowed to speak with other participants or to communicate in any other way. If you want to ask a question, please raise your hand.

[*A short introduction on how to use the computer followed, the instructions continued with the description of the experiment.*]

Design experiment

This experiment consists of at least 90 rounds. From round 90 on a next round will start with a probability of 90%. After the experiment has finished, the computer will choose 20 rounds randomly. Only the earnings from these 20 rounds will be paid to you.

At the beginning of every round the participants will be randomly divided into pairs. The probability to form a pair with a specific other participant is the same for all participants in every round. However, the probability to form a pair twice in a row with the same participant is very small.

One of the two participants will play role A, the other role B. Which role you play is also determined randomly in every round. Only when appointed role A you will have to make a decision, if you have role B, you will not have to do anything that round.

Choices

Being participant A you will have 2 alternatives, which are shown schematically in the table below.

When you choose 'yellow', the participant with role B that you are matched with will get 250 cents and you will lose 150 cents. If you choose 'blue' neither of you will gain or lose money in that round.

	Alternatives	
	yellow	blue
Earnings for A (in cents)	−150	0
Earnings for B (in cents)	250	0

[12]The data and the complete translation of the instructions will be sent by the authors on request.

Information

Before you are asked to make a choice, you will receive information about what the participant you are matched with in this round has chosen in earlier rounds. You will see a summary of the at most 6 most recent decisions that participant B made, when he or she was appointed role A in earlier rounds. This information looks as follows:

Participant B chose as follows in earlier rounds when in role A:
… times yellow … times blue.

The total number of choices equals at most 6. In case participant B has never been in role A, zeros will appear at the dots in the scheme shown above. Only participant A will get this information, participant B will only see how much (s)he earned in this round.

So, if you are appointed role B, the participant A that you are matched with will also get to see your last 6 decisions when you were in role A. The information about you will not change in the periods in which you have role B, as you will not make any choices then.

[*At this point subjects had to answer 5 questions testing understanding. The instructions continued with a description of the windows that would appear during the experiment:*

*On the left part of each computer-screen a permanent box was shown, with the round number, the own score (a summary of the last 6 choices, consisting of a blue and a yellow number) and an overview of all outcomes so far. The colors of the numbers shown corresponded with the choice (*yellow *or* blue*), the choice-maker was coded by a letter 'Y' (you) and 'O' (other).*

At the beginning of each round a window popped up with a message about which role was appointed to the subject. The A-participant first got a window at the top right with information about former choices of participant B and after a few seconds another window at the bottom right in which (s)he had to make her choice. Every decision had to be confirmed. The B-participant only received a message concerning the outcome of that round.]