DEPARTMENT OF ECONOMICS

UNIVERSITY OF MILAN - BICOCCA

WORKING PAPER SERIES

Measuring Indirect Reciprocity: Whose Back Do We Scratch?

Luca Stanca

No. 131 – December 2007

Dipartimento di Economia Politica
Università degli Studi di Milano - Bicocca
http://dipeco.economia.unimib.it

Measuring Indirect Reciprocity: Whose Back Do We Scratch?*

Luca Stanca†
Department of Economics
University of Milan Bicocca

December 2007

Abstract

This paper presents an experimental investigation of strong indirect reciprocity. We examine both generalized indirect reciprocity (if A helps B then B helps C) and social indirect reciprocity (if A helps B then C helps A), in a setting where reciprocal behavior cannot be explained by strategic motivations. We also consider a treatment for direct reciprocity, as a benchmark, and use a variant of the strategy method to control for differences in first movers' actions across treatments. We find evidence of strong reciprocity within each treatment, both for strategies and decisions. Generalized indirect reciprocity is found to be significantly stronger than social indirect reciprocity and, interestingly, direct reciprocity. This finding is interpreted as reflecting the relevance of first movers' motivation for second movers' reciprocal behavior.

Keywords: Reciprocity, Experimental Economics.
JEL codes: C78, C91, C92.

*I thank Tommaso Reggiani for excellent research assistance and Luca Corazzini for help in writing the z-tree programs. I also thank Beatrice Petrovich and Marco Mantovani for assistance in the experimental sessions.

†Piazza dell'Ateneo Nuovo 1, 20126 Milan, Italy. E-mail: luca.stanca@unimib.it

1 Introduction

Reciprocity plays an important role in several economic interactions, as these are often based on the expectation that future benefits will be delivered by another agent. Given that information asymmetries and monitoring costs can make the use of enforceable contracts impossible, economic relations often have to based on implicit contracts whose enforcement is based on reciprocity (see e.g. Fehr et al., 2002, for a discussion). More generally, reciprocity has long been identified in natural and social sciences as an important mechanism in the evolution of cooperation. If individuals interact repeatedly, reciprocating behaviour can induce cooperative behaviour even in a population of purely self-interested individuals (e.g. Hamilton, 1964, Trivers, 1971, Axelrod and Hamilton, 1981, Ridley, 1996, Friedman, 1971, Fudenberg and Maskin, 1986, Binmore, 1998).

Most research on reciprocal behavior has focused on direct reciprocity in bilateral interactions (if I scratch your back you'll scratch mine). More recently, indirect reciprocity has also been identified as an important explanation of cooperative behavior (e.g. Alexander, 1987, Nowak and Sigmund, 1998, Leimar and Hammerstein, 2001).[1] Two mechanisms of indirect reciprocity can be distinguished when the interaction also includes a third actor, not involved in the original exchange. First, social indirect reciprocity, where a kind (or unkind) action is reciprocated *by* a third agent (if I scratch your back someone else will scratch mine). Second, generalized indirect reciprocity, where a kind (or unkind) action is reciprocated *towards* a third agent (if I scratch your back you'll scratch someone else's).

Most of the existing experimental evidence on indirect reciprocity focuses on iterated interactions, where there can be a strategic incentive to reciprocate indirectly (e.g. Wedekind and Milinski, 2000, Engelmann and Fischbacher, 2003, Seinen and Schram, 2004, Bolton, Katok and Ockenfels, 2004, Greiner and Levati, 2005). Relatively less evidence is available on strong indirect reciprocity (e.g. Dufwenberg et. al, 2000, Guth et al., 2001), and the results are generally not conclusive.[2] Whether economic behavior is characterized by strong indirect reciprocity remains an open question. However, indirect reciprocity is indeed an increasingly widespread and relevant phenomenon in economic and social interactions, since iterated exchanges between partners tend to be replaced by one-shot interactions between strangers, as in web-based auctions and other forms of e-commerce.

The objective of this paper is to provide an experimental investigation of both generalized and social indirect reciprocity in a setting where reciprocal behavior cannot be explained by strategic motivations. We compare three treatments in a

[1] See Nowak and Sigmund, 2005, for a comprehensive review of the theoretical and empirical literature on indirect reciprocity).

[2] See also e.g. Fehr and Gachter (2002), Boyd et al. (2003), Carpenter and Matthews (2004), Brandt et al. (2006) for recent research on altruistic punishment.

between-subject design: generalized indirect reciprocity, social indirect reciprocity and, as a benchmark, direct reciprocity. We consider a one-shot gift-exchange game, using a variant of the strategy method to compare second movers' actions while controlling for differences in senders' actions. We find evidence of strong reciprocity within each of the three treatments, both for strategies and actual decisions. In the comparison between treatments, generalized indirect reciprocity is found to be significantly stronger than both social indirect and, quite surprisingly, direct reciprocity. We interpret this result as reflecting the relevance of first movers' extrinsic or intrinsic motivation for second movers' reciprocal behavior.

The paper is structured as follows. Section 2 provides a brief critical review of the existing experimental literature on indirect reciprocity. Section 3 describes the hypotheses to be tested, and the experimental design and procedures. Section 4 presents the results. Section 5 concludes with a discussion of the main findings and the implications of the analysis.

2 Measuring Indirect Reciprocity

Reciprocity can be defined as the conditional behaviour to return helpful and harmful acts in kind, even when this is costly for the reciprocator. Direct reciprocity refers to a sequential interaction between two agents, and it indicates that if A helps (harms) B, then B helps (harms) A at his own cost. Note that in direct reciprocity there is a double coincidence between the first mover and the recipient of the reciprocating action, and between the recipient of the first action and the reciprocating subject. Indirect reciprocity refers to the conditional behaviour to respond to helpful and harmful acts in kind in a sequential interaction that involves a third agent, so that there is no such double coincidence. Indirect reciprocity can take two different forms.[3]

First, cooperative (punishing) behaviour may be based on own prior experiences, in a situation where the reciprocating subject is the recipient of the first move, but the first mover is not the recipient of the reciprocating action. A kind (or unkind) action can therefore be reciprocated *towards* a third agent, not involved in the original interaction: if A helps (harms) B, then B helps (harms) C. We refer to this form of reciprocity as *generalized indirect reciprocity*:[4]

Definition 1 *Positive (negative)* ***generalized*** *indirect reciprocity is a behavior to adopt a helpful (harmful) action towards someone else, at one's own material cost, because some other person's intentional behaviour was perceived to be helpful (harmful) to oneself.*

[3]Nowak and Sigmund (2005) define "upstream reciprocity" reciprocal behavior based on prior experiences and "downstream reciprocity" reciprocal behavior based on reputation.

[4]Rutte and Taborsky (200x) and Pfeiffer et al. (2004) use the term generalize reciprocity to describe individuals who base cooperative behaviour on prior experiences, irrespective of the identity of their partners.

Second, cooperative (punishing) behaviour may be based on reputation, in a situation where the first mover is the recipient of the reciprocating action, but the reciprocating subject is not the recipient of the first move. A kind (or unkind) action can therefore be reciprocated *by* a third agent, not involved in the original interaction: if A helps (harms) B, then C helps (harms) A. We refer to this form of reciprocity as *social indirect reciprocity*:[5]

Definition 2 *Positive (negative)* ***social*** *indirect reciprocity is a behavior to adopt a helpful (harmful) action towards someone else, at one's own material cost, because that person's intentional behaviour was perceived to be helpful (harmful) to some other person.*

In repeated interactions, reciprocal behaviour can be strategic, if it is motivated by the expectation of positive net benefits in the long run. This applies not only to direct reciprocity but also, for different reasons, to both types of indirect reciprocity. Similarly to direct reciprocity, social indirect reciprocity opens the way to strategic reputation building. Self-interested individuals who would otherwise not act cooperatively, might do so in order to increase the probability of being reciprocated by a third party. Reputation can provide a strong mechanism to sustain cooperation, even when the group size is large. Generalized indirect reciprocity may lead to cooperation because previous interactions with anonymous partners provide information about the overall level of cooperation within the group. The expected future benefits from cooperation can outweigh the immediate costs for the reciprocator in interactions within small groups, although this strategic incentive tends to disappear as the group size increases (Boyd and Richerson, 1989, Pfeiffer et al., 2005).

In one-shot interactions, in the absence of expected future benefits of costly reciprocating actions, reciprocity cannot be strategic. Strong reciprocity can be defined as the conditional behaviour to return helpful and harmful acts in kind, even when this is costly and provides neither present nor future net benefits for the reciprocator (see Gintis, 2000, Bowles and Gintis, 2001, Fehr et al., 2002). Strong reciprocity applies not only to direct reciprocity, but also to both generalized and social indirect reciprocity. In the remainder of this section we briefly review the economic experimental literature on indirect reciprocity. While most of this literature considers settings where there can be a strategic incentive to reciprocate indirectly, relatively less evidence is available on strong indirect reciprocity.

Social indirect reciprocity has been analyzed extensively in experiments based on the repeated helping game, developed by Nowak and Sigmund (1998).[6] Helping decisions are generally shown to be significantly affected by the information about

[5]Carpenter and Matthews (2004) define (negative) social reciprocity as "the act of demonstrating one's disapproval, at some personal cost, for the violation of widely-held norms".

[6]The game involves a series of asymmetric single-stage helping games based on random pairing in large groups. Within pairs, donors are given the opportunity of giving a certain amount of money to the recipient, having some information about the past actions of their

number of times the recipient has given and kept in the past (image score). Evidence for strategic reputation building is also found in many experiments: helping rates rise if agents know that their action will affect their own reputation.[7] Social indirect reciprocity thus provides an important mechanism for sustaining cooperative behavior in iterated interactions.

Wedekind and Milinski (2000) report evidence of social indirect reciprocity in a helping game repeated over six rounds. They find that donors are more likely to give to those who have been generous in the past, so that image scoring can significantly promote cooperation. Seinen and Schram (2004) find strong evidence of social indirect reciprocity in a repeated helping experiment involving random pairing in large groups with a large number of rounds. Helping decisions are significantly affected by the information about previous decisions of the recipients. It should be noted, however, that these results can at least partially be attributed to direct reciprocity, given that subjects interact repeatedly with the opponents chosen randomly from a group of 14 over 90 rounds.[8]

Bolton, Katok and Ockenfels (2005) investigate cooperation among strangers with limited information about reputation, focusing on the effects of differing types of recursive information: no information on image score, first order information and second order information.[9] In the no-information treatment the donor chooses without knowing the receiver's history. In the first order information treatment the donor knows whether the receiver helped when last in the role of donor. In the second order information treatment the donor also knows whether the receiver last helped someone who last helped or did not help. Subjects are matched in pairs for each of fourteen rounds, knowing that no two subjects are matched together more than once, so that any reciprocal behavior is necessarily indirect. The main results of the analysis are that not only providing information about partners' immediate past action increases cooperation, but also providing recursive information about the action of partners' previous partners further promotes cooperation.

Engelmann and Fischbacher (2002) investigate indirect reciprocity and strategic reputation building in a similar repeated helping experiment. In their experimental design, in any period only half of the players have a public image score, and therefore have a strategic incentive to help. This allows to distinguish between strategic indirect reciprocity and strong indirect reciprocity. They find evidence of strong indirect reciprocity, but also clear evidence of strategic reputation building: the average helping rate of donors with a public score is more than twice the

receiver. To eliminate the confounding effects of direct reciprocity, players should know that their recipients will never be their donors in future interactions.

[7]The propensity to help is also generally higher in earlier rounds, when reputation has a higher impact on future earnings.

[8]Subjects were actually told that the group size was twice the actual size. As a consequence, as observed by Bolton et al. (2005), the extent to which the results truly reflect indirect reciprocity depends on what subjects believed about the rematch probability.

[9]Subjects are randomly matched in pairs for each of 14 rounds, so that no two subjects are ever matched together more than once. Each player is mover or receiver half of the times.

average helping rate of donors without image score.

Carpenter and Matthews (2004) test for the existence of social reciprocity in a public goods experiment where players are allowed to monitor the decisions made by other players and punish them at a cost (Fehr and Gachter, 2002, Boyd et al., 2003, Brandt et al., 2006). To distinguish between social reciprocity, intended as a normative response, and other punishment explanations they consider a treatment in which players could monitor and sanction other members of their group, and a treatment in which players could monitor and punish members of other groups. Although the norm violation motivation for punishing outside the own group is weaker, the experimental evidence confirms the existence of social reciprocity.

A relatively smaller number of experiments have focused on generalized indirect reciprocity, showing that players who have just received a donation are more likely to give a donation in turn.[10] Greiner and Levati (2005) analyze generalized indirect reciprocity in trust games played in a chain of n players. In their experiment, subjects have information only about what has been given to them, without information on the person they must send to or any other individual in their group. The authors repeat the game a finite number of times and vary the rematching procedure, distinguishing a partners condition and a strangers condition, where players should not have an incentive to play strategically.[11] They also consider a condition where all players decide simultaneously, knowing the amount the received in the previous round, in addition to the one where they play consecutively, knowing the amount they received in the same round. The group size is varied, to check the hypothesis that indirect reciprocity is stronger in smaller groups. The main result is that, although strategic considerations and group size do matter, generalized indirect reciprocity can enhance cooperation in cyclical networks.

Dufwenberg et al. (2001) examine strong indirect reciprocity using a variant of the trust game, as in Berg et al. (1995). One treatment is based on a standard trust game aimed at capturing direct reciprocity. In the indirect reciprocity treatment, subjects interact in groups of four (two donors and two receivers), with crossing sender-responder pairs. Instead of repaying his own sender, as in the direct reciprocity treatment, each responder can only return money to the other sender (without knowing the amount transferred by that sender). Therefore, responders can only reciprocate towards the other donor (generalized indirect reciprocity).[12] The results indicate that direct reciprocity produces higher dona-

[10] See Boyd and Richardson (1989) and Pfeiffer et al. (2004) for theoretical analyses.

[11] It should be observed that in the strangers sessions, although new groups of 3 players were formed in each of the 40 repetitions, this does not rule out the possibility to be matched twice with the same player. Indeed, matching groups were further restricted in order to increase the number of independent observations, although participants were not informed about such restrictions.

[12] In a third treatment, the authors allow for private information of donors about the factor by which donations are multiplied. This allows to distinguish between intrinsically motivated donors and those who are only interested in an image of generosity.

tions by senders on average, but these are poorly rewarded by responders. Rates of return by respondents are higher in the generalized indirect reciprocity treatment. More specifically, the interest rate is positive only in the case of indirect reciprocity, whereas it is negative in both direct reciprocity treatments, although, the difference is not statistically significant. In this paper reciprocity is measured in terms of return rates, defined as $r_i = \frac{y_i}{x_i} - 1$ or $r_i = \frac{y_i}{x_j} - 1$ for direct and indirect reciprocity, respectively. However, average return rates do not provide an appropriate measure of reciprocity, given that they could be reflecting unconditional behaviour.[13]

Guth et al. (2001) examine strong indirect reciprocity in a similar setting. A standard investment game with direct reciprocity is used as a control treatment. In the indirect reciprocity treatment (Indirect reward Full Information), each responder can only return money to the other sender, instead of repaying his own sender as in the direct reciprocity treatment, knowing both the amount sent by his own sender and the amount sent by the other sender.[14] The main difference with respect to Duwfenberg et al. (2001) is that here reciprocating subjects also know what their recipients did in the first stage. The results indicate that investments are strongly reduced when indirect reward is compared with direct reward. It is important to observe that, given the features of the experimental design, the indirect reciprocity treatment does not allow to distinguish between generalized and social indirect reciprocity. Both motives may be driving responders, given that they may respond to the amount received and, at the same time, to the amount sent in the first stage by they person they are responding to.

3 The experiment

3.1 Design

The experiment is based on three treatments in a between-subject design: direct reciprocity, indirect generalized reciprocity and indirect social reciprocity. In all treatments subjects interact in groups of four: A1, A2, B1 and B2. They play a two-stage gift-exchange game (e.g. Fehr et al. 1993, Gachter and Falk, 2002), where all subjects have the same endowment of 20 tokens. In the first stage, A1 (A2) must choose the amount a_1 (a_2), an integer between 0 and 20, she wants to send to player B1 (B2); the amount sent is subtracted from the payoff of A1 (A2),

[13]The finding that the average interest rates is higher in the generalized indirect reciprocity treatment could indeed be explained by the fact that all receivers are *unconditionally* less altruistic in the direct reciprocity treatment (possibly because they take into account the extrinsic motivation of senders).

[14]A third treatment is a slightly modified version of the direct reciprocity treatment, where the interaction between pairs is purely informational. Each receiver is informed of the amounts sent by both senders. This allows to assess whether direct reciprocity is influenced by information about investments received by other second movers (social comparison).

multiplied by 3 by the experimenter, and added to the payoff of B1 (B2). In the second stage, players B1 and B2 must choose the amount (an integer between 0 and 20) they want to send to players A1 and A2. The details of stage 2 differ in the three treatments:

1. In the *direct reciprocity* treatment (DIR), B_i is informed of the amount she received from A_i in stage 1 and must choose the amount b_i, an integer between 0 and 20, she wants to send to A_i; the amount sent is subtracted from the payoff of B_i, multiplied by 3 by the experimenter, and added to the payoff of A_i.

2. In the *indirect generalized reciprocity* treatment (IRG), B_i is informed of the amount she received from A_i in stage 1 and must choose the amount b_i, an integer between 0 and 20, she wants to send to A_j; the amount sent is subtracted from the payoff of B_i, multiplied by 3 by the experimenter, and added to the payoff of A_j.

3. In the *indirect social reciprocity treatment* (IRS), B_i is informed of the amount sent by A_j in stage 1 and must choose the amount b_i, an integer between 0 and 20, she wants to send to player A_j; the amount sent is subtracted from the payoff of B_i, multiplied by 3 by the experimenter, and added to the payoff of A_j.

Total payoffs are $20 - a_i + 3b_i$ for player A_i and $20 - b_i + 3a_i$ for player B_i in DIR, and $20 - a_i + 3b_j$ for player A_i and $20 - b_i + 3a_i$ for player B_i in IRG and IRS. For each player the minimum and maximum potential payoffs are 0 and 80 tokens, respectively.

In stage 2 we applied a variant of the strategy method (henceforth SM): player B had to provide a response for each feasible action of player A, *before* being informed of the actual choice of A. This allowed us to study the responses to each possible action of A and, therefore, to distinguish between unconditional altruism and conditional altruism (positive reciprocity) in the strategies of B players.

After providing a response for each feasible action of player A, players B were informed of the actual action taken by A and had to choose a response (decision method, henceforth DM). Before players B made their choices with the two methods (SM and DM), all players were informed that payoffs would be determined on the basis of one of the two methods, to be selected randomly by publicly tossing a coin. After players B had made their decisions in both SM and DM, the method to determine the payoffs was selected on the basis of the outcome of the coin toss.[15] This procedure based on responses by players B in both strategy and decision method allowed us to compare the consistency between the strategies of B players and their actual responses. It also allowed us to ensure that players

[15]See Fischbacher et al. (2001) for a similar approach.

A could choose their action in stage 1 knowing that in stage 2 players B would choose their action having been informed of the actual action taken by A in stage 1.

3.2 Procedures

We run two sessions for each treatment, with 24 subjects participating in each session, for a total of 96 subjects. In each of the four sessions, subjects were randomly assigned to a computer terminal at their arrival and, before the game started, to their role as A or B (each subject only played one role). In order to ensure public knowledge, instructions were distributed and read aloud (see Appendix 1). Sample questions were distributed to ensure understanding of the experimental procedures. Answers were privately checked and, if necessary, explained to the subjects, and the experiment did not start until all subjects had answered all questions correctly.

The experiment was conducted in the Experimental Economics Laboratory of the University of Milan Bicocca in November 2007. Participants were undergraduate students of Economics recruited by e-mail using a list of voluntary potential candidates. Sessions lasted approximately 45 minutes. No show-up fee was paid and the exchange rate between was 2 tokens per euro. The average payment was about 14 euros, and payments ranged between 0 and 40 euros. The experiment was run using the experimental software z-Tree (Fischbacher, 1999).

3.3 Hypotheses

If subjects are purely self-interested and rational, players B will choose to give zero tokens in all treatments, since they are second movers and gift-giving is costly. If subjects A expect all subjects B to be purely self-interested, by backward induction the optimal choice of players A is also to give zero in stage one. Therefore, under the assumption that rationality is common knowledge, in all treatments the subgame perfect equilibrium is for all players to give zero.

If the behavior of subjects B is characterized by positive strong reciprocity, so that they reward kind actions even if this is costly and cannot be strategic, the response of subjects B depends positively on the amounts sent by subjects A. This provides the first null hypothesis, the absence of strong reciprocity, to be tested within each of the three treatments:

H1a

$$\begin{aligned} H_0 &: \rho_{DIR} = 0 \\ H_1 &: \rho_{DIR} > 0 \end{aligned}$$

H1b

$$H_0 \ : \ \rho_{IRG} = 0$$
$$H_1 \ : \ \rho_{IRG} > 0$$

H1c

$$H_0 \ : \ \rho_{IRS} = 0$$
$$H_1 \ : \ \rho_{IRS} > 0$$

where ρ is the correlation between the amounts sent by subjects A and the corresponding amounts sent by subjects B within each treatment. Note that our operational definition of reciprocity is based on Spearman correlation coefficients, rather than Pearson correlations, so as to avoid restricting the attention to linear dependence. Nevertheless, in order to enable a comparison of the two indicators, in presenting the results we will also report Pearson correlation coefficients.

Turning to the main objective of the analysis, we compare the strength of reciprocating behaviour across the three treatments. In particular, we test the hypothesis that direct reciprocity is stronger than both generalized and social indirect reciprocity, given that only in the former there is a "double coincidence" of roles. We also compare the two indirect reciprocity treatments and test the hypothesis that outcomes are more important than intentions for eliciting reciprocal behaviour, so that social indirect reciprocity (where the payoff of the reciprocator is not affected by the first mover's action) should be less strong than generalized indirect reciprocity (where the payoff of the first mover is not affected by the reciprocator's action):

H2a

$$H_0 \ : \ \rho_{DIR} = \rho_{IRG}$$
$$H_1 \ : \ \rho_{DIR} > \rho_{IRG}$$

H2b

$$H_0 \ : \ \rho_{DIR} = \rho_{IRS}$$
$$H_1 \ : \ \rho_{DIR} > \rho_{IRS}$$

H2c

$$H_0 \ : \ \rho_{IRG} = \rho_{IRS}$$
$$H_1 \ : \ \rho_{IRG} > \rho_{IRS}$$

Note that if subjects A can expect with positive probability subjects B to display strong reciprocity, then the predictions for subjects A would differ across treatments. In particular, subjects A should send more in the DIR treatment than in the IRG treatment and, although to a lesser extent, more in the IRS treatment than in the IRG treatment, since they might be motivated by the trust that a reciprocating response (direct or indirect, respectively, in the two cases) could increase their own payoff. This strategic motivation for subjects A is instead absent in the IRG treatment. The expected differences in the behavior of first movers lead us to focus our analysis of reciprocal behavior on the responses of players B in strategy method. This allows us to compare the three treatments while keeping constant the sending behavior of players A. Nevertheless, we also analyse reciprocity in the actual responses of players B (decision method) in order to provide a check of the consistency between strategies and decisions.

4 Results

Figure 1 displays average amounts sent in tokens, by treatment, for subjects A and B in both SM and DM. Table 1 reports the corresponding figures, while also providing mean and median contributions for each input level in strategy method. The average amounts sent by A players are 7.6, 6.3 and 7.5 tokens in the DIR, IRG and IRS treatments, respectively. This suggests that strategic motivations do play a role for subjects A, as the amount they send is lowest when strategic motivation is absent (IRG), and highest when reciprocity can be expected to be stronger (DIR), although the differences are relatively small.

The average amounts sent by B players in decision and strategy method are 5.3 and 6.4 in the DIR treatment, 3.9 and 8.4 in the IRG treatment and 3.6 and 4.6 in the IRS treatment. These figures are consistent with the hypothesis of strongly reciprocal behavior: focusing on the decision method, where B players respond to different amounts sent by A's, the average response by subjects B is higher in the treatment where A's input is higher (DIR). Once differences in sending behavior by A's are controlled for, responses by players B are highest in IRG and lowest in IRS: the average amounts sent by B players in strategy method are 6.4 in DIR, 8.4 in IRG and 4.6 in IRS. A similar pattern applies to median values.

Table 1

Table 2 reports Mann-Whitney tests of the hypothesis that contributions are the same across pairs of treatments, versus the two-sided alternative that they are different, based on 24 independent observations. The results indicate that the differences between treatments are not statistically significant for subjects A and for subjects B in decision method. However, there are significant differences in the amounts sent by subjects B in strategy method. In particular, the amounts sent in IRG are significantly higher than those in both DIR and IRS, whereas differences

between DIR and IRS are not statistically significant. This result holds both on aggregate (averaging over all 21 possible actions by subject A) and for responses to individual actions by subjects A.[16]

Table 2

Result 1: The amounts sent by subjects B in IRG are significantly higher than those in DIR and IRS, both on average and by individual amount received.

Figures 2 and 3 further illustrate these results by comparing across treatments mean and median responses by B subjects in strategy method for each possible action of A subjects. Average responses in DIR and IRG are very similar up to about 5 tokens sent by subject A. Thereafter, average contributions in the two treatments diverge, as they remain on a steeply rising trend in IRG, whereas they rise at a significantly lower rate in DIR. Contributions in IRG are instead systematically lower for all possible amounts sent by A's. A similar pattern applies to median contributions, displayed in figure 3, where the difference between IRG and DIR is even more pronounced and the one between DIR and IRS relatively small.

Figures 2-3

We now turn to the test of hypothesis 1 (strong reciprocity within treatments). Table 3 reports correlations between the responses of subjects B and the amounts sent by subjects A, within each treatment, and the corresponding p-values for the null hypothesis of zero correlation. If we consider responses in strategy method, Spearman correlation coefficients are positive and strongly significant within each of the three treatments (0.35, 0.64 and 0.49 for the DIR, IRG and IRS treatments, respectively).[17] Similar results are obtained for Pearson correlation coefficients (0.37, 0.65 and 0.50). These results clearly indicate that players B's strategies are characterized by reciprocity.

Table 3

It could be observed that reciprocal behavior might have been to some extent induced, or enhanced, by the use of the strategy method: given that players B are faced with a choice for *each of the feasible actions* of players A, this might artificially lead to stronger reciprocity than if players B were to make only one choice, in response to the single actual decision made by A. It is also possible,

[16]For aggregate responses in strategy method, we computed the average amount sent by each B player and tested the null hypothesis of no difference between treatments on 24 independent observations.

[17]We computed individual correlation coefficients by pairs of subjects, and tested the null hypothesis of zero correlation using a sign test based on 24 independent observations.

more generally, that given that only one of the 21 feasible actions by players A has actually been chosen, the strategy profile of players B as expressed in their SM choices does not represent how they would respond to the actual choice of player A. We therefore also report, in table 3, correlation coefficients for the responses of players B to the actual amounts sent by players A (decision method), for each of the three treatments. The correlation coefficients are, as in the previous case, positive and strongly significant within each of the two treatments, and are indeed larger than in the case of SM responses (0.57, 0.47 and 0.56 for the DIR, IRG and IRS treatments, respectively). Similar results apply to Pearson correlation cofficients.

> **Result 2**: The null hypothesis of no reciprocity can be strongly rejected within each of the three treatments, both for strategies and actual decisions.

The results in table 3 provide a qualitative indication that reciprocity is elicited differently in the three treatments. We therefore turn to formal tests of hypothesis 2. We start by analyzing aggregate behavior using regression analysis, to then examine reciprocating behavior at individual level.

Table 4 reports the results obtained by regressing the response of B players on A players' action, in either DM or SM, using observations from all three treatments. Parameter estimates are obtained by OLS, test statistics are based on heteroskedasticity-robust standard errors, with errors clustered on pairs of subjects for equations estimated on SM observations.

Table 4

In the first specification, observations from all three treatments are pooled together and common parameters are estimated separately for DM and SM. The results, reported in columns (1) and (2), indicate that, averaging over all treatments, the coefficient measuring reciprocity is similar across decisions and strategies (0.48). Next, we also include dummy variables and interaction terms (amount sent by A players multiplied by the treatment-specific dummy) for each of the two indirect reciprocity treatments. This allows us to assess the differences in reciprocating behavior between the two treatments at the aggregate level. While there are no significant differences across treatments in the estimates for decisions (column 3), the estimates for strategies, reported in columns (4), show that the coefficient for the IRG interaction term is positive and significant at the 5 per cent level. The coefficient for the IRS interaction term is instead negative and not significant. This indicates that reciprocal behavior is significantly stronger in IRG than in DIR and, a fortiori, than in IRS.

> **Result 3**: On aggregate, reciprocity is significantly stronger in the IRG-treatment than in the IRS-treatment and the DIR-treatment.

Since subjects B had to provide a response for each feasible action of players A, we can also study differences in reciprocating behavior between the two treatments at individual level. Table 5 reports Mann-Whitney rank-sum tests of the null hypothesis that correlation coefficients are the same across pairs of treatments, against the relevant one-sided alternative. The results lead to rejection of the null hypothesis of equal median at the 5 per cent significance level in the comparison between DIR and IRG and at the 10 per cent significance level in the comparison between DIR and IRS.

Table 5

> **Result 4**: At individual level, reciprocity is significantly stronger in the IRG-treatment than in both the IRS-treatment and the DIR-treatment.

In order to help interpret this result, figures 4 to 6 display the histogram and the corresponding cumulative distributions of the individual amounts sent, by treatment, for players A and for players B in both SM and DM.

Figures 4-6

Finally, we check the robustness of the results by examining whether there are any systematic differences in how subjects respond in strategy and decision method. Table 6 reports regression results for the relationship between B players' actual decisions and their strategies: responses in DM are regressed on responses in SM corresponding to the same amount given by A. As above, we also include dummy variables and interaction terms (amount sent by A players multiplied by the treatment-specific dummy) for each of the two indirect reciprocity treatments, either individually (columns 1 and 2) or jointly (column 3). Strategies explain about 55 per cent of the overall variability of decisions of subjects B. The coefficient for strategies is about 0.7 and statistically significant in all treatments. The coefficients of the interaction terms indicate that, relative to the DIR treatment, decisions follow strategies more closely in the IRS-treatment (0.13) and less closely in the IRS treatment (-0.04). However, these differences between treatments are not statistically significant.

Table 6

5 Discussion

This paper presented an experimental investigation of strong indirect reciprocity, in a setting where reciprocal behavior cannot be explained by strategic motivations. We compared generalized indirect reciprocity, social indirect reciprocity and, as a benchmark, direct reciprocity. We used a variant of the strategy method

to compare second movers' actions while controlling for differences in senders' actions. The results provide several challenges to the paradigm of rationality that dominates economics.

Within treatments, we found evidence of non-strategic reciprocal behavior not only for direct reciprocity, but also for social and generalized indirect reciprocity. The finding of social indirect reciprocity in the absence of iterated interactions supports the existing evidence from repeated helping games, indicating that reputation provides an effective incentive for cooperative behavior of self-interested individuals for whom it is not in their short-term interest to cooperate. The finding of strong generalized indirect reciprocity indicates that reciprocal cooperation is not necessarily based on specific knowledge about the partner's kindness, but any prior experience of cooperation may be relevant.

Between treatments, generalized indirect reciprocity is found to be significantly stronger not only, as expected, than social indirect reciprocity but, quite surprisingly, than direct reciprocity. One possible explanation for this finding is based on how agents evaluate the kindness of an action and, more specifically, the idea that the motivation of an action may be relevant for its perceived kindness. If an action is perceived to be more kind when intrinsically rather than extrinsically motivated, reciprocity should be stronger in response to actions that can only be driven by intrinsic motivation, than to actions that can also be driven by extrinsic motivation (Deci and Ryan, 1985, Deci et al., 2000).[18] In our experiment, the responder's willingness to cooperate in the direct reciprocity treatment may be crowded out by the extrinsic motivation of the sender, which is instead absent in the generalized indirect reciprocity.

Overall, our results suggest several implications and directions for future research on reciprocity and cooperation. First, at the empirical level, experimental and field evidence of direct reciprocity could be interpreted in terms of generalized reciprocity. More generally, empirical studies should attempt to identify what determines the perceived kindnesss of an action in determining of reciprocal behavior. Second, at the theoretical level, reciprocity cannot be explained by models that focus only on the outcomes of the actions one is responding to. Theoretical models of reciprocal behavior should also take into account intentions and, in particular, consider explicitly the type of motivation driving an action. Finally, theoretical approaches attempting to explain cooperation in an evolutionary context should pay more attention to the potential role of generalized reciprocity. Generalized reciprocity may represent a more general mechanism leading to cooperation than direct and indirect reciprocity, which require individual recognition and specific social memory.

[18]Focusing on direct reciprocity, Stanca et al. (2007) find that reciprocal behavior is significantly stronger when extrinsic motivation can be ruled out, a result that extends the concept of motivational crowding-out from *within* subject to *between* subjects.

6 Appendix: Instructions

This appendix reports the instructions distributed on paper to the subjects for the indirect reciprocity treatments (IRG and IRS). We indicate in brackets parts that are specific to a given treatment.

Instructions *[common to all treatments]*

- Welcome and thanks for participating in this experiment.
- During the experiment you are not allowed to talk or communicate in any way with other participants. If at any time you have any questions raise your hand and one of the assistants will come to you to answer it.
- By following the instructions carefully you can earn an amount of money that will depend on your choices and the choices of other participants.
- At the end of the experiment the number of tokens that you have earned will be converted in euros at the exchange rate 2 tokens = 1 euro. The resulting amount will be paid to you in cash.

General rules

- There are 24 subjects participating in this experiment.
- In the experiment you will interact in a group of 4 subjects.
- You will not be informed about the identity of the other subjects in your group, nor will they be informed about yours.

How players interact

- The four subjects in your group will be randomly assigned to one of four roles: A1, A2, B1, and B2.
- Each subject will receive an endowment of 20 tokens.
- The experiment will take place in 2 phases.

PHASE 1

- A1 chooses how many tokens (between 0 and 20) to send to B1. At the same time, A2 chooses how many tokens (between 0 and 20) to send to B2.
- We will triple the amount sent, so that B1 and B2 will receive 3 times the tokens sent by A1 and A2, respectively.

PHASE 2

- (IRG) In phase 2 B1, informed of the amount that she received from A1 in stage 1 will choose how many tokens (between 0 and 20) to send to A2. At the same time B2, informed of the amount that she received from A2 in stage 1 and will choose how many tokens (between 0 and 20) to send to A1.
- (IRS) In phase 2 B1, informed of the amount that A2 sent to B2 in stage 1 and will choose how many tokens (between 0 and 20) to send to A2. At the same time B2, informed of the amount that A1 sent to B1 in stage 1, will choose how many tokens (between 0 and 20) to send to A1.
- We will triple the amounts sent, so that A2 will receive 3 times the tokens sent by B1 and A1 will receive 3 times the tokens sent by B2.
- The experiment will end. Overall earnings for each subject will be determined as the sum of the earnings obtained in the two phases:
  - A1: 20 tokens minus the tokens she sent to a B1 in phase 1 plus 3 times the tokens sent to her by B2 in phase 2.
  - A2: 20 tokens minus the tokens she sent to a B2 in phase 1 plus 3 times the tokens sent to her by B1 in phase 2.
  - B1: 20 tokens plus 3 times the tokens sent to her by A1 in phase 1 minus the tokens she sent to A2 in phase 2.
  - B2: 20 tokens plus 3 times the tokens sent to her by A2 in phase 1 minus the tokens she sent to A1 in phase 2.

Instructions ON SCREEN - phase 1

- You are A1. You have to decide how many tokens (between 0 and 20) to send to B1. We will triple the amount sent, so that A2 will receive 3 tokens for each token you send.
- (IRG) In phase 2 B1 will be informed of the amount that she received from you in stage 1 and will choose how many tokens (between 0 and 20) to send to A2.
- (IRS) In phase 2 B2 will be informed of the amount that you sent to B1 in stage 1 and will choose how many tokens (between 0 and 20) to send to you.

Instructions ON SCREEN - phase 2

- You are B1. You have to decide how many tokens (between 0 and 20) to send to A2. We will triple the amount sent, so that A2 will receive 3 tokens for each token you send.

- (IRG) Your choice will be made with two different methods:
  - Method 1: *before being informed of how many tokens A1 actually sent you in phase 1,* you have to decide how many tokens to send to A2 for each of the possible amounts that A1 could have sent to you in phase 1 $(0, 1, ..., 20$ tokens). Since there are 21 possible cases, you have to make 21 choices.
  - Method 2: *after being informed of how many tokens A1 actually sent you in phase 1*, you have to decide how many tokens you want to send to A2.
- After you have made your choice with both methods, earnings will be determined on the basis of one of the two methods, selected randomly.
  - If method 1 is selected, of the 21 choices that you had made, only the one corresponding to the actual decision of A1 will be used to determine the earnings.
  - If method 2 is selected, the single choice that you had made will be used to determine the earnings.
- (IRS) Your choice will be made with two different methods:
  - Method 1: *before being informed of how many tokens A2 actually sent to B2 in phase 1,* you have to decide how many tokens to send to A2 for each of the possible amounts that she could have sent to B2 in phase 1 $(0, 1, ..., 20$ tokens). Since there are 21 possible cases, you have to make 21 choices.
  - Method 2: *after being informed of how many tokens A2 actually sent to B2 in phase 1*, you have to decide how many tokens you want to send to A2.
- After you have made your choice with both methods, earnings will be determined on the basis of one of the two methods, selected randomly.
  - If method 1 is selected, of the 21 choices that you had made, only the one corresponding to the actual decision of A1 will be used to determine the earnings.
  - If method 2 is selected, the single choice that you had made will be used to determine the earnings.

The experiment will end and overall earnings for each subject will be determined as the sum of the earnings obtained in phase 1 and in phase 2.

7 References

Alexander, R.D. (1987) "The Biology of Moral Systems", Aldine de Gruyter, New York.

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

Axelrod, R. and W.D. Hamilton (1981) The evolution of cooperation. Science, CCXI, p. 1390-1396.

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

Binmore, K.G. (1998) Game Theory and the Social Contract, II: Just Playing. MIT Press, Cambridge, MA.

Bolton, G.E., Katok, E., and Ockenfels, A. (2005) "Cooperation among Strangers with Limited Information about Reputation", Journal of Public Economics, 89, 2005, 1457-1468.

Bonein, A., and D. Serra (2007) "Another experimental look at reciprocal behavior: indirect reciprocity," Working Papers 07-04, LAMETA, Universtiy of Montpellier.

Boyd, R., and Richerson, P.J. (1989) The evolution of indirect reciprocity. Social Networks, 11, 213-236 (1989).

Boyd R, Gintis H, Bowles S, Richerson PJ (2003) The evolution of altruistic punishment. Proceedings National Academy of Science, U S A 100: 3531-3535.

Bowles, S., and Gintis, H. (2001) The Evolution of Strong Reciprocity, Discussion Paper, University of Massachusetts at Amherst.

Brandt H, Hauert C, Sigmund K (2006) Punishing and abstaining for public goods. Proc Natl Acad Sci U S A 103: 495-497.

Carpenter, J.P. and Matthews, P.H. (2004) "Social Reciprocity", IZA Discussion Paper No. 1347.

Deci E.L., Koestner, R., and R.M. Ryan (2000) "A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation, Psychological Bulletin, 125, p.627-668.

Deci, E. L., and Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. New York: Plenum.

Dufwenberg, M., U. Gneezy, W. Güth, and E. van Damme (2001) "Direct versus Indirect Reciprocity: An Experiment", Homo Oeconomicus, 2001, vol. 18, pages 19-30.

Engelmann, D. and Fischbacher, U. (2002) Indirect Reciprocity and Strategic Reputation Building in an Experimental Helping Game. Working Paper n. 132, Institute for Empirical Research in Economics, University of Zurich.

Fehr, E., Kirchsteiger, G. and Riedl, A. (1993), Does Fairness Prevent Market Clearing? An Experimental Investigation, Quarterly Journal of Economics 108, 437-460.

Fehr, E., Fischbacher, U., and Gächter, S. (2002) Strong Reciprocity, Human Cooperation and the Enforcement of Social Norms", Human Nature, 13, p. 1-25.

Fehr, E. and Gächter, S. (2000), Fairness and Retaliation: The Economics of Reciprocity. Journal of Economic Perspectives, Summer, 14, pp. 159-81.

Fehr E, Gächter S (2002) Altruistic punishment in humans. Nature 415: 137-140.

Friedman, J. (1971) A Noncooperative Equilibrium for Supergames, Review of Economic Studies 38: 1-12.

Fudenberg D. and E. Maskin (1986) The Folk Theorem in Repeated Games with Discounting or with Incomplete Information, Econometrica 54: 533-556.

Gachter, S. and Falk, A. (2002), "Reputation and Reciprocity: Consequences for the Labour Relation", Scandinavian Journal of Economics 104(1), 1-26.

Gintis, H. (2000) Strong Reciprocity and Human Sociality, Journal of Theoretical Biology 206: 169-179.

Greiner, B., and Levati, M.V., (2005) "Indirect reciprocity in cyclical networks: An experimental study," Journal of Economic Psychology, Elsevier, vol. 26(5), pages 711-731, October.

Guth, W., Konigstein, M., Marchand, N., K. Nehring, (2001) Trust and Reciprocity in the Investment Game with Indirect Reward, Homo Oeconomicus, vol. 18, pages 241-262.

Hamilton, W.D. (1964) Genetical Evolution of Social Behavior I, II, Journal of Theoretical Biology 7(1): 1-52.

Leimar O, Hammerstein P (2001) Evolution of cooperation through indirect reciprocity. Proceedings of the Royal Society Biological Sciences, 268: 745-753.

Nowak, M. and Sigmund, K. (1992) Tit for tat in heterogeneous populations, Nature 355, 250253.

Nowak, M. and Sigmund, K. 1993 A strategy of winstay, lose shift that outperforms tit-for-tat in the Prisoners dilemma game. Nature 364, 5658.MA, 1992.

Nowak, M.A., Sigmund, K., 1998. The dynamics of indirect reciprocity. Journal of Theoretical Biology, 194, 561-574.

Nowak, M.A., and Sigmund, K. (2005) Evolution of indirect reciprocity, Nature, Vol 437, n. 27, p. 1291-1298.

Pfeiffer, T., Rutte, C., Killingback, T., Taborsky, M., and Bonhoeffer, S. (2005) Evolution of cooperation by generalized reciprocity. Proceedings of the Royal Society Biological Sciences, 272, 1115-1120.

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

Rutte C., Taborsky M. (2007) Generalized Reciprocity in Rats. Public Library of Science, Biology 5(7): e196 doi:10.1371/journal.pbio.0050196

Seinen, Ingrid; Schram, Arthur (2006), Social Status and Group Norms: Indirect Reciprocity in a Repeated Helping Experiment, European Economic Review, vol. 50, no. 3, pp. 581-602.

Stanca, L., Bruni, L., and L. Corazzini (2007) Testing Theories of Reciprocity: Does Motivation Matter?, Working Paper n. 109, Economics Department, University of Milan Bicocca.

Trivers, R. The evolution of reciprocal alruism. Quarterly Review of Biology, XLVI (1971), 35-57.

Wedekind, C. and Milinski, M. (2000) Cooperation through image scoring in humans, Science 288, 850-852.

Table 1: Mean and median amounts sent in tokens, by treatment

	Means			Medians		
	DIR	IRG	IRS	DIR	IRG	IRS
Subject A	7.63	6.29	7.54	6.50	4.50	10.00
Subject B - DM	5.29	3.92	3.63	1.50	2.50	4.00
Subject B - SM all subjects	6.41	8.43	4.57	4.00	8.00	3.00
Subject B - SM 0	1.88	1.38	0.29	0.00	0.00	0.00
Subject B - SM 1	2.42	2.08	0.67	1.00	1.00	0.00
Subject B - SM 2	2.92	2.79	1.25	1.50	2.00	1.00
Subject B - SM 3	3.50	3.75	1.46	2.50	3.00	1.00
Subject B - SM 4	4.25	4.67	1.88	4.00	4.00	2.00
Subject B - SM 5	4.21	4.83	2.88	3.00	5.00	2.50
Subject B - SM 6	5.25	6.25	2.96	4.50	6.00	3.00
Subject B - SM 7	5.46	6.92	3.58	4.00	7.00	3.50
Subject B - SM 8	5.75	8.13	3.88	4.00	8.00	3.50
Subject B - SM 9	5.88	7.92	4.50	4.50	9.00	4.50
Subject B - SM 10	6.46	9.04	5.04	5.50	10.00	5.00
Subject B - SM 11	7.29	9.96	5.08	5.50	11.00	5.00
Subject B - SM 12	6.96	10.17	5.54	6.00	10.00	5.00
Subject B - SM 13	7.04	10.75	5.92	6.00	12.50	5.00
Subject B - SM 14	7.54	11.38	6.42	6.00	13.50	5.50
Subject B - SM 15	8.96	11.92	6.88	8.00	15.00	5.50
Subject B - SM 16	9.25	12.17	7.17	9.00	14.00	5.50
Subject B - SM 17	9.58	11.96	7.33	8.00	14.50	5.50
Subject B - SM 18	9.67	13.25	7.63	9.00	16.50	6.00
Subject B - SM 19	10.17	13.54	7.92	9.50	17.50	6.50
Subject B - SM 20	10.25	14.25	7.63	9.50	18.50	6.00

Note: Means and medians are calculated over 24 individual observations within each treatment (504 observations for SM all subjects).

Table 2: Test for differences in contributions between treatments

	IRG-IRS		DIR-IRS		DIR-IRG	
	U-test	P-value	U-test	P-value	U-test	P-value
Subject A	-1.20	0.23	-0.31	0.75	0.97	0.33
Subject B - DM	0.06	0.95	0.08	0.94	-0.19	0.85
Subject B - SM	3.40	0.00	0.94	0.35	-2.04	0.04
Subject B - SM 0	2.33	0.02	0.92	0.36	-1.37	0.17
Subject B - SM 1	3.14	0.00	1.26	0.21	-1.68	0.09
Subject B - SM 2	2.97	0.00	0.82	0.41	-1.96	0.05
Subject B - SM 3	3.66	0.00	1.24	0.22	-1.97	0.05
Subject B - SM 4	3.94	0.00	1.72	0.09	-1.91	0.06
Subject B - SM 5	2.66	0.01	0.44	0.66	-2.21	0.03
Subject B - SM 6	3.70	0.00	1.34	0.18	-1.80	0.07
Subject B - SM 7	3.71	0.00	0.98	0.32	-2.08	0.04
Subject B - SM 8	3.89	0.00	1.02	0.31	-2.28	0.02
Subject B - SM 9	3.16	0.00	0.65	0.52	-2.19	0.03
Subject B - SM 10	2.86	0.00	0.61	0.54	-2.03	0.04
Subject B - SM 11	3.30	0.00	1.21	0.23	-1.61	0.11
Subject B - SM 12	2.88	0.00	0.65	0.52	-1.97	0.05
Subject B - SM 13	3.08	0.00	0.51	0.61	-2.17	0.03
Subject B - SM 14	3.00	0.00	0.42	0.68	-2.11	0.03
Subject B - SM 15	2.98	0.00	1.00	0.32	-1.51	0.13
Subject B - SM 16	2.77	0.01	0.85	0.39	-1.53	0.13
Subject B - SM 17	2.34	0.02	0.96	0.34	-1.33	0.18
Subject B - SM 18	2.89	0.00	0.76	0.45	-1.58	0.11
Subject B - SM 19	2.78	0.01	0.75	0.45	-1.41	0.16
Subject B - SM 20	2.98	0.00	0.69	0.49	-1.46	0.14

Note: Mann-Whitney test statistics are based on 24 independent observations for each treatment. P-values refer to the relevant two-sided hypothesis. For strategies on aggregate (third row), we computed the average amount sent by each B player and tested the null hypothesis of no difference between treatments over 24 independent observations.

Table 3: Strong reciprocity within treatments

	Strategies			Decisions		
	DIR	IRG	IRS	DIR	IRG	IRS
Spearman correlation	0.35	0.64	0.49	0.57	0.47	0.56
P-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
Pearson correlation	0.37	0.65	0.50	0.53	0.59	0.53
P-value	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)

Note: For responses in strategy method we computed individual correlation coefficients by pairs of subjects, and tested the null hypothesis of zero correlation, against the relevant one-sided alternative, using a sign-test based on 24 independent observations.

Table 4: Regressions of B's responses to A's actions (SM, DM)

	(1)	(2)	(3)	(4)
	DM	SM	DM	SM
Tokens sent by A	0.48**	0.48**	0.61**	0.42**
	(4.19)	(10.67)	(2.71)	(4.95)
Tokens sent by A * IRG dummy			-0.20	0.22*
			(-0.67)	(1.98)
Tokens sent by A * IRS dummy			-0.23	-0.02
			(-0.98)	(-0.17)
IRG treatment dummy			0.69	-0.19
			(0.29)	(-0.15)
IRS treatment dummy			0.14	-1.66
			(0.07)	(-1.38)
Constant	0.88	1.62**	0.64	2.24*
	(1.05)	(3.73)	(0.31)	(1.93)
R^2	0.28	0.23	0.31	0.31
Number of observations	72	1512	72	1512

Note: The dependent variable is the response by subjects B. Coefficients are estimated by OLS, t-statistics reported in brackets. Test statistics are based on heteroskedasticity-robust standard errors, with errors clustered on pairs of subjects for equations estimated on SM observations. * indicates $p<0.05$, ** indicates $p<0.01$.

Table 5: Test for differences in reciprocity between treatments

	IRG-IRS	DIR-IRS	DIR-IRG
Spearman correlations - U-test	1.42	-0.76	-1.74
Spearman correlations - P-value	(0.08)	(0.22)	(0.04)
Pearson correlations - U-test	1.26	-1.10	-1.88
Pearson correlations - P-value	(0.11)	(0.14)	(0.03)

Note: the figures reported are Mann-Whitney rank-sum tests of the null hypothesis that correlation coefficients are the same across pairs of treatments. P-values refer to the relevant one-sided hypothesis (see section 3).

Table 6: Regressions of B's responses on B's strategies

	(1)	(2)	(3)
B's strategy	0.76**	0.71**	0.73**
	(8.28)	(8.32)	(6.89)
B's strategy * IRG-treatment dummy	-0.07		-0.04
	(-0.42)		(-0.23)
IRG-treatment dummy	-0.98		-1.27
	(-0.78)		(-0.88)
B's strategy * IRS-treatment dummy		0.14	0.13
		(0.64)	(0.55)
IRS-treatment dummy		-0.08	-0.74
		(-0.06)	(-0.53)
Constant	1.12	0.74	1.40
	(1.71)	(1.06)	(1.51)
R^2	0.56	0.54	0.54
Number of observations	72	72	72

Note: OLS estimates. t-statistics reported in brackets (heteroskedasticity-robust standard errors). * indicates p<0.05, ** indicates p<0.01.

Figure 1: Mean amounts sent, by treatment

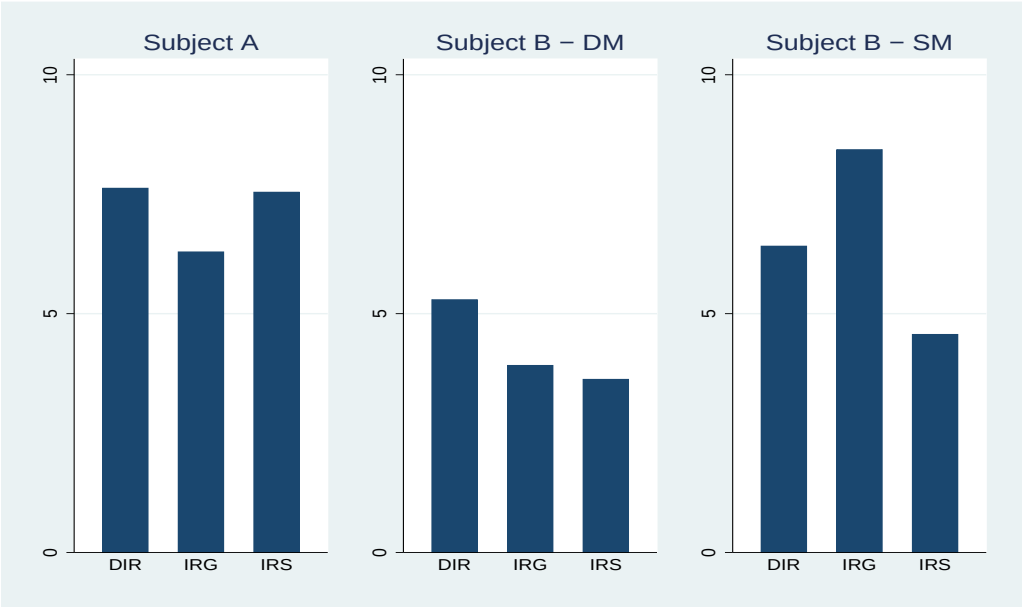

Figure 2: Mean responses of players B (SM)

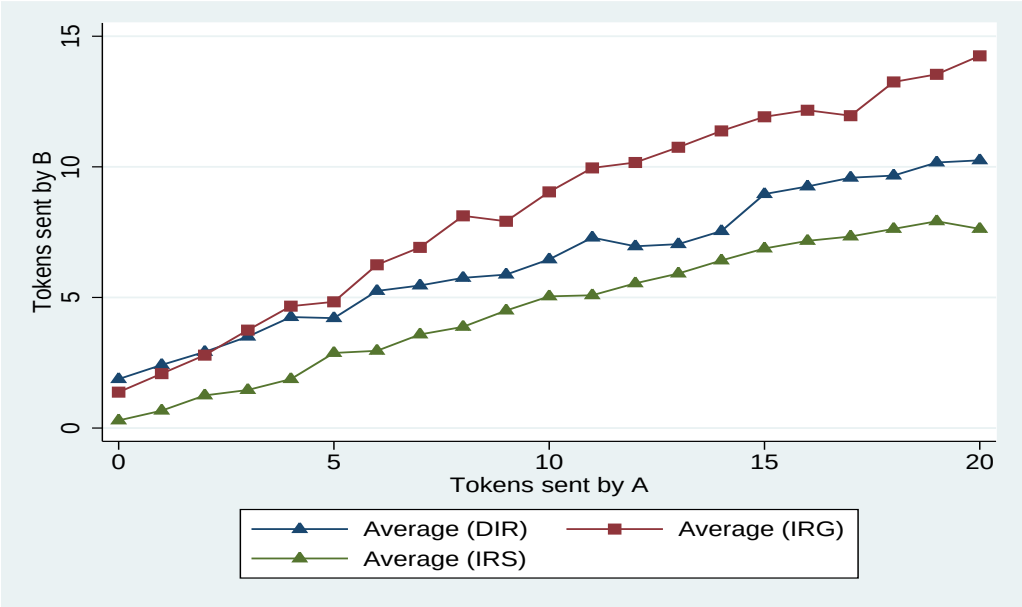

Figure 3: Median responses of players B (SM)

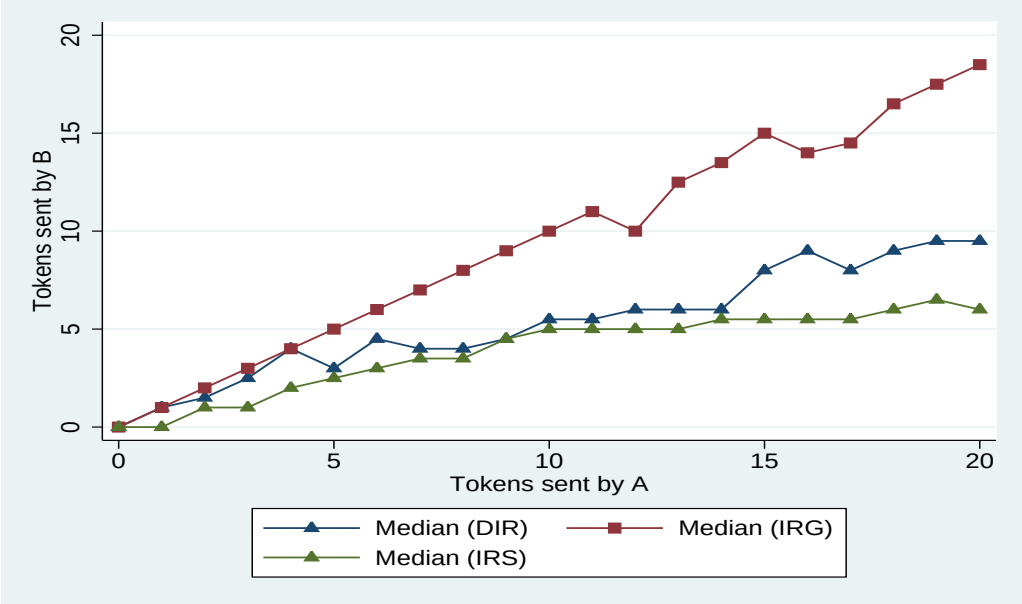

Figure 4: Distribution of amounts sent by A, by treatment

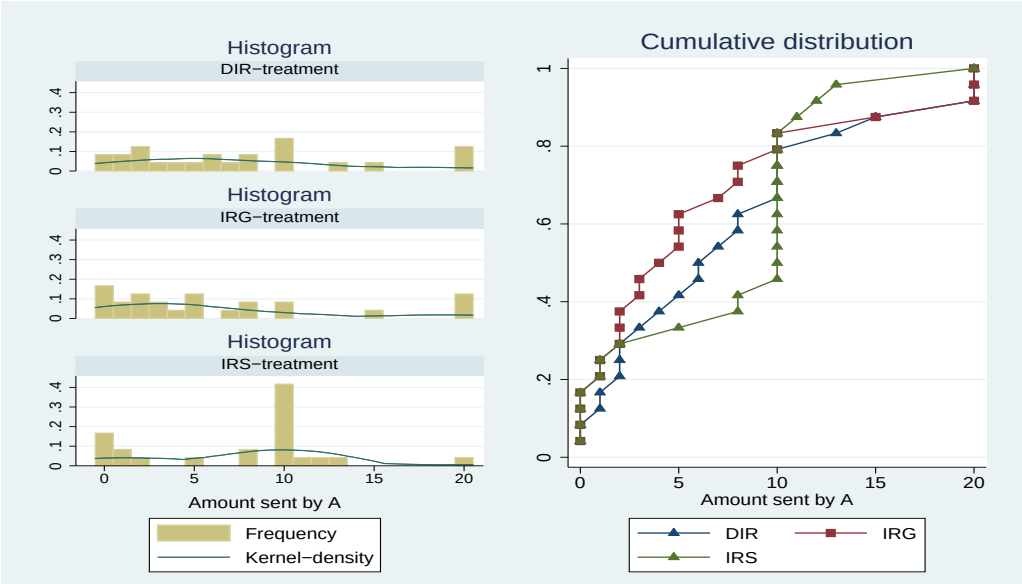

Figure 5: Distribution of amounts sent by B (SM), by treatment

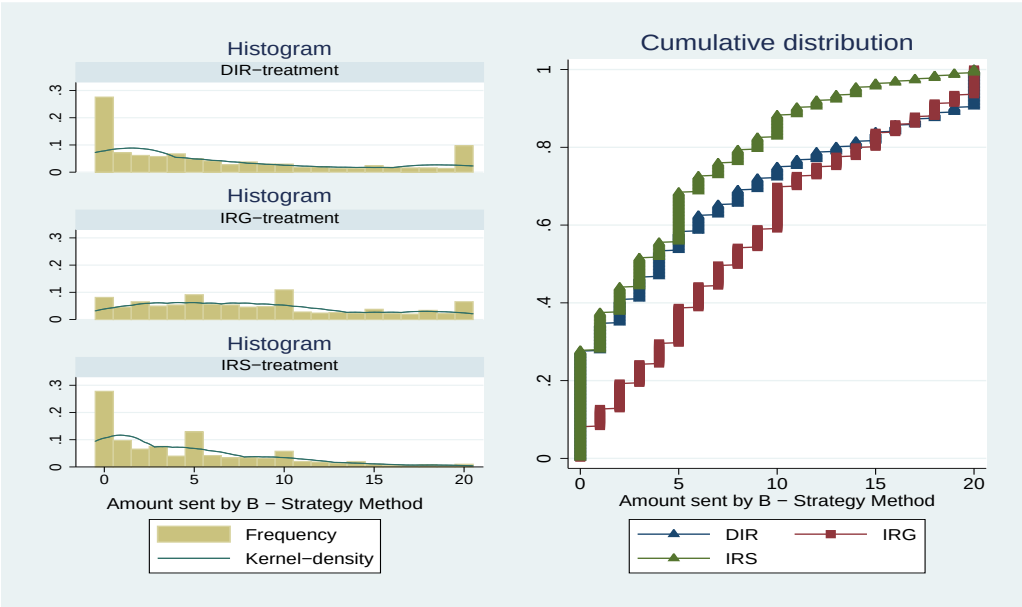

Figure 6: Distribution of amounts sent by B (DM), by treatment