DEPARTMENT OF ECONOMICS

UNIVERSITY OF MILAN - BICOCCA

WORKING PAPER SERIES

Testing Theories of Reciprocity: Does Motivation Matter?

Luca Stanca, Luigino Bruni, Luca Corazzini

No. 109 - March 2007

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

Testing Theories of Reciprocity: Does Motivation Matter?*

Luca Stanca[†], Luigino Bruni[‡], Luca Corazzini[§]

March 2007

Abstract

One of the key issues for understanding reciprocity is how people evaluate the kindness of an action. In this paper we argue that the motivation driving an action plays an important role for the reciprocating response to that action. We test experimentally the hypothesis that reciprocal behavior is stronger in response to actions driven by intrinsic motivation, as opposed to extrinsic motivation. Our results indicate that reciprocity is significantly stronger when extrinsic motivation can be ruled out, both at the aggregate and the individual level. These findings suggest that models of reciprocal behavior should take into account not only outcomes but also intentions and, in particular, motivations: the type of motivation of an action matters for its perceived kindness and, as a consequence, for reciprocity.

Keywords: Reciprocity, Intrinsic Motivation, Laboratory Experiments.
JEL codes: D63, C78, C91.

*We thank Tommaso Reggiani for excellent research assistance. We acknowledge financial support from Econometica.

[†]Corresponding author. Economics Department, University of Milan Bicocca. Piazza dell'Ateneo Nuovo 1, 20126 Milan, Italy. E-mail: luca.stanca@unimib.it

[‡]Economics Department, University of Milan Bicocca. Piazza dell'Ateneo Nuovo 1, 20126 Milan, Italy. E-mail: luigino.bruni@unimib.it

[§]School of Economics and Social Studies, University of East Anglia, Norwich, NR4 7TJ, United Kingdom. Istituto di Economia Politica "Ettore Bocconi", Università L. Bocconi, via Gobbi, 5, 20136, Milan, Italy. Email: luca.corazzini@unibocconi.it

1 Introduction

A large number of experimental and field studies have indicated that economic behavior is not always motivated by pure self-interest. In particular, economic interactions are in many cases characterized by reciprocity: the non-strategic conditional behavior to reward kind actions and to punish unkind actions, even if this is costly for the reciprocating subject (see e.g. Fehr and Gachter, 2000, Sobel, 2004, and Fehr and Schmidt, 2006, for recent surveys).

Among the main theoretical explanations of reciprocity, one class of models has focused on the properties of the *outcomes* of the actions one is responding to, as in the theories based on "inequity aversion" (e.g. Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000). A second class of models has emphasized the role of the *intentions* determining the actions (e.g. Rabin, 1993; Dufwenberg and Kirchsteiger, 1998; Charness and Rabin, 2000, Falk et al., 2003).[1]

At the empirical level, many authors have examined the roles played by outcomes and intentions in determining reciprocity. Experimental studies by Charness (1996), Bolton et al. (1998), Offerman (1999), and Cox (2000) have found that the fairness of intentions has a relatively small impact on positive reciprocity. Blount (1995) and Offerman (1999) have shown that intentions do play a significant role for negative reciprocity. Falk et al. (2000) have found that the attribution of fairness intentions has a large and significant impact on both positive and negative reciprocal behavior.

Most of this recent empirical literature has assessed the role of intentions by considering experimental designs where the presence or absence of intentions is used as the main treatment variable. In these studies, a control treatment where the sender can intentionally choose what action to take among a set of alternatives (thus signalling her intentions) is generally compared with a treatment where the sender cannot choose, either because she does not have alternative options, as in McCabe et al. (2003), or because her choice is determined exogenously, as in Blount (1995) and Falk et al., (2000).

In this paper we focus on the behavioral relevance of the *nature* of intentions for reciprocity. More precisely, we consider the effect that the *motivation* of an agent has on the reciprocating behavior of another agent. We provide experimental evidence on whether the intrinsic (non-instrumental) or extrinsic (instrumental) motivation of an action have different effects on

[1] Both intentions and outcomes matter for reciprocity in the model by Falk and Fishbacher (2006).

the reciprocating response to that action.[2] Our main hypothesis is that reciprocity is stronger in response to actions that are perceived as driven by intrinsic motivation, than in response to actions that are perceived as extrinsically motivated.[3]

In order to test this hypothesis, we examine experimentally a two-player sequential move symmetric gift-exchange game. Our design is such that, in one treatment, the first mover can only be driven by intrinsic motivation, while in the second treatment she can also be driven by extrinsic motivation. The two treatments only differ with respect to the nature of the intentions of the first mover, whereas the fairness of the outcome of the first mover's action is kept constant across treatments. We expect positive reciprocity to be stronger in the first treatment, where the sender's action can only be perceived to be driven by intrinsic motivation, than in the second treatment, where it can be perceived to be instrumental.

The results indicate that the type of motivation of an agent has a significant impact on other agents' reciprocal behavior. When the experimental design rules out the attribution of extrinsic motivation, second mover's responses are characterized by significantly stronger reciprocity. This result holds both for strategy profiles and actual decisions, and both at the aggregate and the individual level: a larger fraction of the subjects do not display any reciprocal behavior when extrinsic motivation cannot be ruled out.

The paper is structured as follows. Section 2 presents the theoretical motivation of the analysis. Section 3 describes the experimental design. Section 4 discusses the predictions to be tested. Section 5 presents the results. Section 6 concludes with a discussion of the main findings and the implications of the analysis.

[2]Following the psychological literature, in particular the Self Determination Theory of Deci and Ryan (2000), "to be intrinsically motivated means to engage in an activity because the activity itself is interesting and enjoyable" (Deci et al. 2007, p. 12). Behaviour motivated by extrinsic motivation, instead, "entails doing an activity because it leads to some outcome that is operationally separable from the activity itself. That is, extrinsic motivation concerns activities enacted because they are instrumental rather than because one finds the actions satisfying in their own right" (Deci et al. 2007, p. 13).

[3]In a different perspective, Falk et al. (1999) refer to reciprocity as an intrinsic motivation, that is, as an "inherent willingness to respond in kind which is in principle independent of extrinsic incentives" (p. 254).

2 Intrinsic Motivation and Reciprocity

In intention-based theories of reciprocity, agents reward actions that are perceived to be kind and punish actions that are perceived to be unkind. One consequence of the intention-based approach is that actions with identical outcomes may elicit different reciprocating responses depending on how they are interpreted. The key issue for understanding reciprocity is therefore to understand *how* people evaluate the kindness of a particular action. We argue that the type of motivation driving an action plays an important role for the reciprocal response to that action.

Much empirical evidence is available about the *motivational crowding-out effect*: the reduction of effort in activities carried out for intrinsic motivation when an instrumental reward, typically monetary, is introduced (e.g. Cameron and Pierce, 1996, Eisenberger and Cameron, 1996, Deci et al., 2000, Deci and Ryan, 2000). Self Determination Theory explains this phenomenon using the concept of *control* (Deci and Ryan, 1985): extrinsic reward is perceived as a form of control and, in intrinsically motivated people, it can reduce effort.[4]

Frey (1997) applied the concept of motivational crowding-out to economics, examining the reduction of the effort of intrinsically motivated workers when the relationship is transformed from intrinsic-based to extrinsic-based (typically, incentive-based). Gneezy and Rustichini (2000) explored the effects of motivational crowding-out in an experiment on fund-raising: the group of students not paid for collecting money performed better than the group that, instead, was paid. At the theoretical level, Harvey (2005) explained motivational crowding out within a principal-agent model where utility is interpreted as overall satisfaction.[5]

Existing studies on motivational crowding-out have generally investigated the differences in behavior motivated by intrinsic or extrinsic rewards *within the same person*. This research was not intended primarily to explore interactions or, in particular, reciprocity.[6] The intuition behind our study is to

[4] "Controlled regulation of behaviors is associated with poorer quality of performance and a lack of persistence once the controlling contingencies are removed" (Deci et al., 2007, p. 2).

[5] The role of intrinsic motivation for optimal incentive contracts was examined by Kreps (1997) and Murdock (2002). See also the review in Lindenberg (2001).

[6] Reeve and Jang (2006) identified specific autonomy-supportive teaching practices and found that these instructional behaviors, such as asking questions, being empathic, listening to students, being responsive to questions, encouraging students, providing positive feedback, and offering a meaningful rationale, were associated with students being more autonomously engaged in learning. In contrast, teachers who were directive, used con-

extend motivational crowding-out theory from "within person" to "between persons".

Following Falk and Fishbacher (2006), consider a utility function for agent i that depends not only on material payoffs (π_i), but also on a reciprocity component, expressed as the product of a reciprocity parameter, a "kindness term" and a "reciprocation term":

$$U_i = \pi_i + \rho_i \phi \sigma$$

The reciprocity parameter (ρ_i) represents the individual sensitivity to reciprocity. The kindness term (ϕ) measures how kind an agent perceives the action by another agent. The reciprocity term (σ) measures the effect of the reciprocal action on the agent's utility.

We argue that the motivation driving an action is relevant for its perceived kindness, so that the kindness term depends on the type of motivation driving the action one is responding to. In particular, for a given outcome, an action is perceived to be more kind if it is genuine (intrinsically motivated, IM) than if it is instrumental (extrinsically motivated, EM). This implies that $\phi_{EM} < \phi_{IM}$ or, alternatively, $\phi_{EM} = (1 - \delta) \phi_{IM}$, with $\delta > 0$. As a consequence, in an economic interaction, intrinsic motivation of the first mover results in stronger reciprocity of the second mover than extrinsic motivation.

We have designed an experiment in order to test this hypothesis ($\delta > 0$). Our constituent game is a version of the gift-exchange game. In one treatment, the motivation of the first mover can only be perceived as intrinsic, while in a second treatment it can also be perceived as extrinsic. If the perceived kindness of an action is related to the nature of its motivation, we would expect reciprocity to be stronger in the treatment where the sender's action can only be driven by intrinsic motivation, than in the treatment where it can also be driven by extrinsic motivation.

3 Experimental design and procedures

Our experiment is based on a symmetric version of the gift-exchange game (e.g. Fehr et al. 1993, Gachter and Falk, 2002). As illustrated below, this game has the advantage of making it easier for the reciprocating subject to interpret the nature of the other player's intentions, whose effect on reciprocity is the core of our analysis. We start by describing the details of the

trolling language, asked controlling questions, and gave solutions rather than hints and encouragement were experienced as autonomy diminishing. (Deci et al., 2007, 14).

game, then present the two treatments, and finally illustrate the procedures of the experiment.

3.1 The constituent game

We consider a two-player sequential move game that consists of two stages. At the beginning of the game both players (A and B) are given an endowment of 20 tokens. In the first stage, player A must choose the amount a (an integer between 0 and 20) she wants to send to player B; the amount sent is subtracted from the payoff of A, multiplied by 3 by the experimenter, and added to the payoff of B. In the second stage, player B must choose the amount (an integer between 0 and 20) she wants to send to player A; the amount sent is subtracted from the payoff of B, multiplied by 3 by the experimenter, and added to the payoff of A. Total payoffs are therefore $20 - a + 3b$ for player A and $20 - b + 3a$ for player B. For each player the minimum and maximum potential payoffs are 0 and 80 tokens, respectively.

Information feedback is as follows. At the end of stage 1, each subject is informed of her stage payoff in tokens. At the end of stage 2 each subject is informed of her stage payoff in tokens and of her total payoff in tokens and in euros. At the beginning of the game subjects are informed that there is no show-up fee, so that earnings are determined only by total payoffs, and that the exchange rate is 2 tokens = 1 euro.

A number of features of this game are intended to facilitate the reciprocating subject's interpretation of the other player's motivation. First, symmetry in the endowments eliminates the confounding effects of distributional aspects, that may arise for example in a trust game: since both players have the same endowment, inequality aversion cannot determine A's decision (it can motivate B's decision but, as explained below, in exactly the same way in the two treatments). Second, symmetry in the actions of the two players greatly simplifies the reciprocating subject's task of reading the other player's mind in order to interpret her motives. Third, the structure of the game is extremely simple, so as to enhance the saliency of the treatment.

3.2 Treatments

The treatment variable is the information set of the players. In the *information* treatment (I-treatment), *before* playing stage 1 all subjects (A and B) are informed that there will be a stage 2 that will be played with the same rules as in stage 1 but with reversed roles. Therefore, this is simply a version of the standard gift-exchange game. In the *no-information* treatment (NI-treatment), only *after* stage 1 has ended players (A and B) are informed that there will be a stage two where the same action will be taken with reversed roles. In this case, stage 1 is played as if it was the whole game, and stage 2 is played as a surprise sub-game.[7]

The two treatments differ with respect to the motivations that may determine A's action in stage 1. In the I-treatment, A can give tokens to B for pure altruism and/or because she trusts that B will reciprocate, thus increasing her own overall payoff. Player A's motivation, can therefore be perceived by player B as *intrinsic* but also *extrinsic* (that is, aimed at achieving a higher payoff through B's reciprocating response). In the NI-treatment, instead, A cannot give to B in order to obtain something else, given that stage 1 is played as if the game should then end. In this case player A's motivation can only be perceived by B as *intrinsic*.

Note that since all players receive the same information, in stage 2 players B are fully aware of the motivation driving players A's actions in stage 1. Given that in stage 2 players B have to take exactly the same action as players A in stage 1, it is particularly simple for them to interpret the nature of A's intentions in each of the two treatments. Since all other conditions are kept fixed, any differences in the reciprocating behavior of players B can be interpreted as the effect of the differences in player A's motivation.

It is important to observe that in both treatments the response of player B can only be driven by intrinsic motivation: since the game ends after player B's action, there is nothing "external" to be reached by her action. As a consequence, our analysis can be interpreted as a test of an inter-personal version of the motivational crowding out hypothesis: the extrinsic motivation of the *first* mover may crowd-out the intrinsically motivated reciprocating behavior of the *second* mover.

3.3 Procedures

We applied a variant of the strategy method in stage 2, when players B have to make their choice: player B had to provide a response for each feasible action of player A, *before* being informed of the actual choice of A (strategy method, henceforth SM). This allowed us to study the responses to each

[7]Note that subjects are simply given instructions about stage 1, without any explicit reference to the game ending thereafter, so that subjects are not in any way cheated by the announcement of stage 2. On the other hand, there is no reason why, when playing stage 1, subjects should expect stage 2 to follow. See the instructions in the appendix for details.

possible action of A and therefore, on the basis of responses to different actions of A, to distinguish between unconditional altruism and conditional altruism (reciprocity) in the strategies of B players. It is important to observe that the specific features of the NI-treatment, based on a surprise stage 2, imply that we cannot observe repeatedly subjects' reciprocating behavior over successive periods. This makes the application of the strategy method particularly appropriate in order to investigate reciprocal behavior within our experimental design.

After giving a response for each feasible action of player A, players B were informed of the actual action taken by A and had to provide a response (decision method, henceforth DM). Before players B made their choices with the two methods (SM and DM), all players were informed that the 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.[8]

This procedure based on responses by players B in both strategy and decision method allowed us to ensure that in the I-treatment players 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 (thus making salient the extrinsic motivation). It also allowed us to compare the consistency between the strategies of B players and their actual responses.

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 player A or B (each subject only played one role). To ensure public knowledge, instructions were distributed and read aloud (see Appendix 1). To ensure understanding of the experimental procedures, sample questions were distributed. 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.

4 Predictions

If it is common knowledge that all subjects are self-interested and rational, players B will choose to give zero tokens in both treatments, since they are second movers and gift-giving is costly. In the I-treatment, by backward induction, the optimal choice of players A is to give zero in stage one. In

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

the NI-treatment, the optimal choice of players A is again to give zero, since they play as if they were at the terminal node of the game. Therefore, in both treatments the subgame perfect equilibrium outcome is for all players to give zero: player A will choose $a = 0$ and player B will choose $b = 0$.

If subjects' behavior is characterized by positive reciprocity, so that they reward kind actions even if this is costly and cannot be strategic, in both treatments the response of players B should depend positively on the amount sent by player A:

> **Hypothesis 1. Reciprocity**: The amount sent by B in stage 2 is positively related to the amount sent by A in stage 1.

Our operational definition of reciprocity is based on the Spearman correlation coefficient between the amounts sent by A and B, rather than Pearson correlations, so as to avoid restricting the attention to linear dependence.[9] Nevertheless, in order to enable a comparison of the two indicators, in presenting the results we will also report Pearson correlation coefficients.

Note that if some players are motivated by reciprocity, and this is common knowledge, then the predictions for players A will differ in two treatments. In particular, players A should send more in the I-treatment, since they might be motivated not only by pure generosity, but also by the trust that a reciprocating response could increase their own payoff. This is a further reason why the focus of our analysis is mainly on the responses of players B in the strategy method: this allows us to compare the two treatments, characterized by different types of first mover's motivation, while controlling for differences in 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 robustness of our results.

Let us turn to the main hypothesis of the paper. In a relationship between a first mover and a second mover, if the action of the first mover is perceived by the second mover as extrinsically motivated, positive reciprocity is elicited less strongly than if the action is perceived as intrinsically motivated. In our design, if reciprocity is motivated only by the fairness of outcomes (e.g. inequity aversion), the positive relationship between the amounts sent by A in stage 1 and by B in stage 2 should be *the same* across treatments.

[9] Note that if, for example, reciprocity was motivated only by inequality aversion, this could imply a linear relationship between b and a, so that Pearson correlations could be used as an appropriate indicator of reciprocating behaviour. However, if reciprocity is motivated also (or only) by the fairness of intentions, the relationship between b and a is not necessarily linear, so that a non-parametric indicator such as the Spearman correlation is more appropriate.

On the contrary, if the motivation of an action matters for its perceived kindness, such relation could be different across treatments. In particular, if an action driven by extrinsic motivation is perceived as less kind than the same action driven by intrinsic motivation, then the positively reciprocating behavior of the responder should be stronger in the NI-treatment, where extrinsic motivation can be ruled out, than in the I-treatment, where extrinsic motivation cannot be ruled out.

> **Hypothesis 2. Effect of motivation on reciprocity:** The positive relation between the amounts sent by A in stage 1 and by B in stage 2 should be stronger in the NI-treatment than in the I-treatment.

5 Results

The experiment was conducted in the Experimental Economics Laboratory of the University of Milan Bicocca in January 2007. Participants were undergraduate students of Economics recruited using an email list of voluntary potential candidates. None of the subjects had participated previously in trust or gift-exchange games. Sessions lasted approximately forty five minutes. No show-up fee was paid and the exchange rate between the tokens and the euro was 2 tokens = 1 euro. The average payment was 14.9 euros, and payments ranged between 0 and 40 euros. The experiment was run using the experimental software z-Tree (Fischbacher, 1999).

Figures 1 to 3 display the histogram and the corresponding cumulative distributions, by treatment, for a and for b in both SM and DM. Table 1 reports median amounts sent, by treatment, for players A and for players B in both SM and DM. The median amount sent by A players is 3 tokens in the NI-treatment (mean=4.46). This suggests that players A are not driven purely by self interest, as they give almost 25 per cent of their endowment on average. The median amount sent by players A is 6.5 tokens in the I-treatment (mean=7.63). The difference between the two treatments indicates that self interest also plays a substantial role in determining the decisions of players A. The difference in the median of the distributions in the two treatments is statistically significant at the five per cent level, using a Mann-Whitney U-test for the relevant one-sided hypothesis. The median response by players B in strategy method is the same in the two treatments (4 tokens), whereas the median response in decision method is 0 in the NI-treatment and 1.5 in the I-treatment. In either case there are no significant differences between the two treatments.

Table 1: Median amount sent by treatment and significance of differences

	I	NI	U-stat	P-val	Nobs-I	Nobs-NI
Amount sent by A	6.50	3.00	1.90	0.06	24	24
Amount sent by B (SM)	4.00	4.00	-0.28	0.78	504	504
Amount sent by B (DM)	1.50	0.00	1.08	0.28	24	24

Note: I = Information treatment. NI = No Information treatment.

The figures reported in table 1 indicate that the differences in the distribution of responses by B players under the two decision methods are consistent with the hypothesis of reciprocal behavior: considering the strategy method, where differences in sending behavior by A's are controlled for, the median response by players B is the same in the two treatments; focusing on the decision method, where the B players respond to different amounts sent by A's, median responses are higher in the treatment where A's median input is higher (I-treatment).

Let us therefore turn to the test of hypothesis 1 (reciprocity). Table 2 reports correlation coefficients between the responses of B's and the amounts sent by A's, within each treatment. If we consider the strategy profiles of B players (responses in strategy method), the Spearman correlation coefficients are positive and strongly significant within each of the two treatments (0.35 and 0.52 for the I- and NI-treatment, respectively). Similar results are obtained for Pearson correlation coefficients (0.37 and 0.46 for the I- and NI-treatment, respectively). These results clearly indicate that players B's strategies are characterized by reciprocity.

Table 2: Reciprocity within treatments

	I (SM)	NI (SM)	I (DM)	NI (DM)
Spearman correlation	0.35	0.52	0.57	0.76
P-value	0.00	0.00	0.00	0.00
Pearson correlation	0.37	0.46	0.53	0.70
P-value	0.00	0.00	0.01	0.00

Note: I = Information treatment. NI = No Information treatment.

It is possible to argue that reciprocal behavior could be 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 could also be argued, more generally, that given that only 1 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 necessarily represent how they would respond to the actual choice of player A. We therefore also report, in table 2, correlation coefficients for the responses of B's to the actual amounts sent by A's (decision method), for the two treatments. The correlations 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 and 0.76 for I- and NI-treatment, respectively). Pearson correlations are 0.37 and 0.46, respectively, in the two treatments.

> **Result 1**: The null hypothesis of no reciprocity can be strongly rejected within each of the two treatments, both for strategies and actual decisions.

The results in table 2 provide a qualitative indication that, as expected, the strength of reciprocity is different in the two treatments: correlation coefficients are larger in the NI-treatment, where extrinsic motivation can be ruled out. We therefore turn to formal tests of hypothesis 2. We start by analyzing aggregate behavior using regression analysis.

Table 3 reports the results obtained by regressing the SM response of B on A's action, using observations from both treatments. We also include a dummy variable for the NI-treatment and an interaction term (A's action multiplied by the NI-treatment dummy). This allows us to assess the differences in reciprocating behavior between the two treatments at the aggregate level. We define a kind response as sending a positive number of tokens and an unkind response as sending zero tokens, so that players' actions are defined as binary variables. The parameters are estimated with a probit estimator, and test statistics are based on errors clustered on pairs of subjects.

The results for model (1), where both players' actions are defined as binary variables, indicate that the probability that B players respond kindly is significantly higher if player A's action in stage 1 was kind. More importantly, the coefficient for the interaction term is positive and significant at the 1 per cent level using the relevant one-sided test, indicating that reciprocal behavior is stronger in the NI-treatment. The results for model (2), where only players B's actions are defined as binary variables, indicate that the probability that B players respond kindly depends positively on the amount received by A in stage. As in the previous case, the coefficient for the interaction term is positive and strongly significant, consistently with the hypothesis that reciprocal behavior is stronger in the NI-treatment.

Table 3: Regressions of B's responses to A's actions (SM)

	(1)	(2)
Amount sent by A (binary)	1.65**	
	(4.75)	
Amount sent by A (binary) * NI-treatment dummy	0.80*	
	(1.61)	
Amount sent by A		0.04**
		(3.12)
Amount sent by A * NI dummy		0.10**
		(2.69)
NI-treatment dummy	-0.18	-0.15
	(-0.40)	(-0.53)
Constant	-0.97**	0.18
	(-3.14)	(0.90)
Pseudo-R^2	0.14	0.15
Number of observations	1008	1008

Note: Dependent variable: binary response by B (SM), probit estimates. t-statistics reported in brackets (standard errors clustered on subjects' pairs). * indicates p<0.05, ** indicates p<0.01. I = Information treatment. NI = No Information treatment.

Since players B had to provide a response for each feasible action of player A, we can also study the differences in reciprocating behavior between the two treatments at the individual level. Figure 4 displays the histogram of the Spearman correlations in the two treatments, and the corresponding cumulative distributions. Table 4 reports individual correlation coefficients between the responses of B's and each of the possible amounts sent by A's, for each of the two treatments. The coefficients are sorted in ascending order, and probability values are reported to assess their statistical significance. Both the table and the figures clearly indicate that individual correlations are larger in the NI-treatment.

Table 5 reports the results of the test of the null hypothesis that the distributions of correlation coefficients are the same in the two treatments. In particular, using a t-test, the null hypothesis of equal means across treatments is strongly rejected (p-value = 0.03 for a one- sided test, for both Spearman and Pearson correlation). Using a Mann-Whitney U-test, as suggested by the nature of the distribution of the correlation coefficients and the small sample size, also leads to rejection at the 10 per cent significance level (the p-values for a one-sided test are 0.10 for the Spearman correlation and 0.8 for the Pearson correlation coefficients, respectively). Kolmogorov-Smirnov

Table 4: Individual Spearman correlations, by treatment

	I-treatment	(P-value)	NI-treatment	(P-value)
Pair 1	0.00	1.00	0.00	1.00
Pair 2	0.00	1.00	0.12	0.62
Pair 3	0.00	1.00	0.23	0.32
Pair 4	0.00	1.00	0.37	0.10
Pair 5	0.00	1.00	0.48	0.03
Pair 6	0.00	1.00	0.74	0.00
Pair 7	0.07	0.78	0.76	0.00
Pair 8	0.17	0.47	0.78	0.00
Pair 9	0.35	0.12	0.87	0.00
Pair 10	0.61	0.00	0.90	0.00
Pair 11	0.61	0.00	0.93	0.00
Pair 12	0.68	0.00	0.96	0.00
Pair 13	0.84	0.00	0.97	0.00
Pair 14	0.92	0.00	0.98	0.00
Pair 15	0.94	0.00	0.98	0.00
Pair 16	0.95	0.00	0.98	0.00
Pair 17	0.99	0.00	0.99	0.00
Pair 18	0.99	0.00	0.99	0.00
Pair 19	1.00	0.00	0.99	0.00
Pair 20	1.00	0.00	1.00	0.00
Pair 21	1.00	0.00	1.00	0.00
Pair 22	1.00	0.00	1.00	0.00
Pair 23	1.00	0.00	1.00	0.00
Pair 24	1.00	0.00	1.00	0.00

Note: I = Information treatment. NI = No Information treatment. Correlation coefficients are sorted by size.

test-statistics do not lead to reject the null of equality of distributions, but provide an indication that the whole distribution of correlations in the NI-treatment dominates the corresponding distribution for the I-treatment.

Table 5: Test for differences in distributions of correlations

	T-stat	P-val	U-stat	P-val	KS-stat	P-val
Spearman correlation	-1.88	0.03	-1.28	0.10	0.29	0.13
Pearson correlation	-1.86	0.03	-1.43	0.08	0.29	0.13

Note: the table reports results of the test of the null hypothesis that the distributions of correlations coefficients are the same in the Information treatment and the No Information treatment.

These findings clearly indicate that, at the individual level, reciprocity is stronger in the NI-treatment. Individual strategy profiles, reported in figures 5 and 6, help to explain this result: not only do subjects reciprocate more on average in the NI-treatment, but also a larger fraction of the subjects do not display any reciprocal behavior when extrinsic motivation cannot be ruled out. In the I-treatment, 6 out of 24 responding subjects display an unconditional strategy profile (zero correlation with A's decision), as opposed to only 1 out of 24 in the NI-treatment.

> **Result 2**: Reciprocity is significantly stronger in the NI-treatment than in the I-treatment, both at aggregate and individual level.

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 in the two treatments. Table 6 reports regression results for the relationship between B players' actual decisions on their strategies: B's responses in DM are regressed by OLS on B's responses in SM corresponding to the same amount given by A.

The results indicate that strategies explain about 64 per cent of the overall variability of decisions. The coefficient for strategies is positive and significant in both treatments (0.73 and 0.94). The difference in the coefficients between the two treatments indicates that the decisions follow the strategies more closely in the NI-treatment, but is not statistically significant. This provides further evidence that the results reported above for the tests of the hypothesis that the nature of motivations matters for reciprocating behavior are not affected by the use of the strategy method.

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

	Response by B (DM)
B's strategy	0.73**
	(6.41)
B's strategy * NI-treatment dummy	0.21
	(1.04)
NI-treatment dummy	-1.63
	(-1.18)
Constant	1.40
	(1.40)
R^2	0.61
Number of observations	48

Note: OLS estimates. t-statistics reported in brackets. * indicates p<0.05, ** indicates p<0.01. I = Information treatment. NI = No Information treatment.

6 Conclusions

This paper presented an experimental investigation of the hypothesis that the motivation of an action matters for its perceived kindness, so that reciprocating responses may differ depending on whether an action is perceived as driven by intrinsic or extrinsic motivation. By extending the concept of motivational crowding-out to interpersonal relations and, in particular, economic interactions, we argued that extrinsic motivation by an agent may weaken the responding agent's reciprocal behavior.

In order to test this hypothesis, we designed a gift-exchange game where, in one treatment, the first mover can only be driven by intrinsic motivation, while in the second treatment she can also be driven by extrinsic motivation, while the fairness of the outcome of the first mover's action is kept constant across treatments. Our results indicate that when extrinsic motivation can be ruled out, reciprocal responses are indeed stronger. This finding holds both for strategy profiles and actual decisions, and both at the aggregate and the individual level. These results suggest several implications and directions for future research.

First, 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: an action may elicit reciprocity differently depending on whether it is perceived as driven by intrinsic or extrinsic motivation.

A second implication is related to the power of disinterested philanthropy in eliciting reciprocity in helped people (Margalit, 1982): is free riding less strong when people are engaged in mutual advantageous actions (like contracts), or when the helped person feels disinterested motivation in her partner? In other words, in a "Samaritan dilemma" (Buchanan, 1976), does gratuity induce more or less reciprocity than a contract (without enforcement)? Our experiment suggests that disinterested gifts elicit more reciprocity.

A third implication refers to interpersonal relationships, in particular within organizations. The literature on motivational crowding-out indicates that the introduction of instrumental incentives in domains where intrinsic motivation are at work can reduce effort and efficiency. If disinterested costly actions generate relatively more reciprocity, then it can be more efficient in workplaces, for instance, not to assign all interpersonal relations to contracts and incentives, but to leave some room for gratuity as a way to stimulate reciprocity and, as a consequence, cooperation. This would apply, in particular, to those domains where stakeholders ask for intrinsic motivation, such as value-based organizations, the non-profit sector, or caring. In such domains, to replace genuine relationality with contracts can reduce reciprocity. Our experiment can, therefore, offer new evidence and, possibly, theoretical hints for understanding such crowding-out phenomena.

7 Appendix: Instructions

This appendix reports the instructions distributed on paper to the subjects. Paragraph headings indicate in brackets if the given subsection is common to both treatments or is specific to the relevant treatment.

Instructions *[common to both 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 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 *[common to both treatments]*

- There are 24 subjects participating in this experiment.
- At the beginning of the experiment 12 couples of two participants will be formed randomly and anonymously. Within each couple, the two subjects will be randomly assigned two different roles: A and B.
- Therefore, each subject will interact exclusively with the other subject in her pair, without knowing her/her identity, and will have the role (A or B) assigned to him with equal probability at the beginning of the experiment.

How players interact *[NI-treatment]*

- Both A and B will receive an endowment of 20 tokens each.
- Player A will have to decide how many tokens (between 0 and 20) to send to player B.
- We will triple the amount sent, so that B will receive 3 tokens for each token sent by A.

- Therefore:
  - A will obtain 20 tokens minus the tokens sent to B;
  - B will obtain 20 tokens plus 3 times the tokens sent by A.

How players interact *[I-treatment]*

- Both A and B will receive an endowment of 20 tokens each.

The experiment will take place in 2 phases.

- PHASE 1
  - Player A will have to decide how many tokens (between 0 and 20) to send to player B.
  - We will triple the amount sent, so that B will receive 3 tokens for each token sent by A.
- PHASE 2
  - Subject B, having been informed of the amount sent to him by Player A in phase 1, will have to decide how many tokens (between 0 and 20) to send to player A.
  - We will triple the amount sent, so that A will receive 3 tokens for each token sent by B.
- Therefore, in total:
  - A will obtain 20 tokens minus the tokens sent to B in phase 1 plus 3 times the tokens sent by B in phase 2.
  - B will obtain 20 tokens plus 3 times the tokens sent by A in phase 1 minus the tokens sent to A in phase 2.

Instructions - phase 2 *[common to both treatments]*

- B has to decide how many tokens (between 0 and 20) to send to A, who sent a certain amount of tokens to B in phase 1.
- We will triple the amount sent, so that A will receive 3 tokens for each token sent by B.
- The choice of how many tokens B wants to send to A will be made with two different methods:
  - Method 1: *before being informed of how many tokens A sent to B in phase 1*, B has to decide how many tokens she wants to send to A for each of the possible amounts that A could have sent to him $(0, 1, ..., 20$ tokens). Since there are 21 possible cases, B has to make 21 choices.
  - Method 2: *after being informed of how many tokens A actually sent to B in phase 1*, B has to decide how many tokens she wants to send to A.
- After B players have made their 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 B had made, only the one corresponding to the actual decision of A will be used to determine the earnings.
  - If method 2 is selected, the single choice that B 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.

8 References

Blount, S. (1995). When social outcomes aren't fair: The effect of causal attributions on preferences. Organizational Behavior and Human Decision Processes LXIII, 131144.

Bolton, G. E., Brandts, J. and Ockenfels, A. (1998): Measuring Motivations for the Reciprocal Responses Observed in a Simple Dilemma Game, Experimental Economics 1, 207-220.

Bolton, G. E. and Ockenfels, A. (2000). ERC - A Theory of Equity, Reciprocity and Competition. American Economic Review, 90, 166–193.

Buchanan, J. (1975), The Samaritan's Dilemma, in Altruism, Morality and Economic Theory, ed. E.S. Phelps, 71-85, The Russel Sage Fundation, New York.

Cameron, J. and W.D. Pierce (1996) Reinforcement, Reward and Intrinsic Motivation: A Meta-Analysis, Review of Educational Research, 64, 363-423.

Charness, G. (1996): Attribution and Reciprocity in a Simulated Labor Market: An Experimental Investigation, Discussion paper, University of Berkeley.

Charness, G. and Rabin M. (2000): Social Preferences: Some Simple Tests and a New Model. Mimeo, University of California at Berkeley.

Cox, J. (2000): Trust and Reciprocity: Implications of Game Triads and Social Contexts, Working paper, University of Arizona.

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.

Deci E.L., and Ryan R.M. (2000) "The what and why of goal pursuits: human needs and the self determination of behavior ", Psychological Enquiry, 11, 227-268.

Deci E.L., R.M. Ryan, van Steenkiste M. (2007), "Self-determination theory and the explanatory role of pshycological needs in human well-being", in Bruni, L. et al. (eds.) "Happiness and Capability", Oxford University Press, forthcoming.

Dufwenberg M. and Kirchsteiger, G. (1999): A Theory of Sequential Reciprocity, mimeo, CentER for Economic Research, Tilburg.

Eisenberger, R.W. and J. Cameron (1996) Detrimental Effects of Reward: Reality or Myth?, American Psychologist, 51, p. 1153-1166.

Falk, Armin, Fehr, Ernst and Fischbacher, Urs, (2000) "Testing Theories of Fairness - Intentions Matter", Zurich IEER Working Paper No. 63.

Falk, A., Fehr, E., Fischbacher, U. (2003) On the nature of fair behaviour. Economic Inquiry. 41, 20-26.

Falk, F. and Fischbacher, U. (2006). A Theory of Reciprocity. Games and Economic Behavior 54(2), 293-315.

Falk, A., Gachter, S., Kovacs, J. (1999) Intrinsic Motivation and Extrinsic Incentives in a Repeated Game with Incomplete Contracts, Journal of Economic Psychology, vol. 20, n. 3, pp. 251-84.

Fehr, E. and Gächter, S. (2000): Fairness and Retaliation: The Economics of Reciprocity, Journal of Economic Perspectives 14, 159-181.

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. and Schmidt, hr, K. (1999): A Theory of Fairness, Competition, and Cooperation, Quarterly Journal of Economics 114, 817-868.

Fehr, E., and Schmidt, K.M. (2006) "The economics of Fairness, Reciprocity and altruism: Experimental evidence", in Handbook of the Economics of Giving, Altruism and Reciprocity, Eds: Varet, G., Ythier, Jean Mercier - Intriligator, Michael D. - Arrow, Kenneth J. - Kolm, S.-C.

Fischbacher, U. (1999), Zurich Toolbox for Readymade Economic Experiments, Working Paper No. 21, University of Zurich, Switzerland.

Fischbacher, U., Gachter, S., Fehr, E. (2001) Are people conditionally cooperative? Evidence from a public goods experiment, Economics Letters, 71, 397-404.

Frey, B. (1997) "Not just for Money: An Economic Theory of Personal Motivation", Brookfield: Edward Elgar.

Gneezy, U., and A. Rustichini "Pay Enough or Don't Pay At All." Quarterly Journal of Economics August 2000, 791-810.

Harvey, J.S. (2005) Why Did You Do That? An Economic Examination of the Effect of Extrinsic Compensation on Intrinsic Motivation and Performance, Journal of Economic Psychology, vol. 26, no. 4, pp. 549-66

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

Kreps, D.M. (1997) Intrinsic Motivation and Extrinsic Incentives, American Economic Review, vol. 87, no. 2, pp. 359-64.

Lindenberg, S. (2001) Intrinsic Motivation in a New Light, Kyklos, vol. 54, n. 2-3, 2001, pp. 317-42.

Margalit, A. (1996), The Decent Society, Harvard University Press, Harvard.

McCabe, K. A., Rigdon, M. L., Smith, V. L., (2003) Positive reciprocity and intentions in trust games. Journal of Economic Behavior and Organization 52 (2), 267-275.

Murdock, K. (2002) Intrinsic Motivation and Optimal Incentive Contracts, RAND Journal of Economics, vol. 33, n. 4, pp. 650-71.

Offerman, T. (1999): Hurting Hurts More that Helping Helps: The Role of the Self-Serving Bias, Working paper, University of Amsterdam.

Rabin, M. (1993). Incorporating Fairness into Game Theory and Economics. American Economic Review, 83, 1281 – 1302.

Reeve, J., & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. Journal of Educational Psychology, 98, 209-218.

Ryan, R. M., and Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55, 68-78.

Sobel J. (2004) "Interdependent Freferences and Reciprocity, Journal of Economic Literature, 2004: C7, D9.

Figure 1: Distribution of amounts sent by A, by treatment

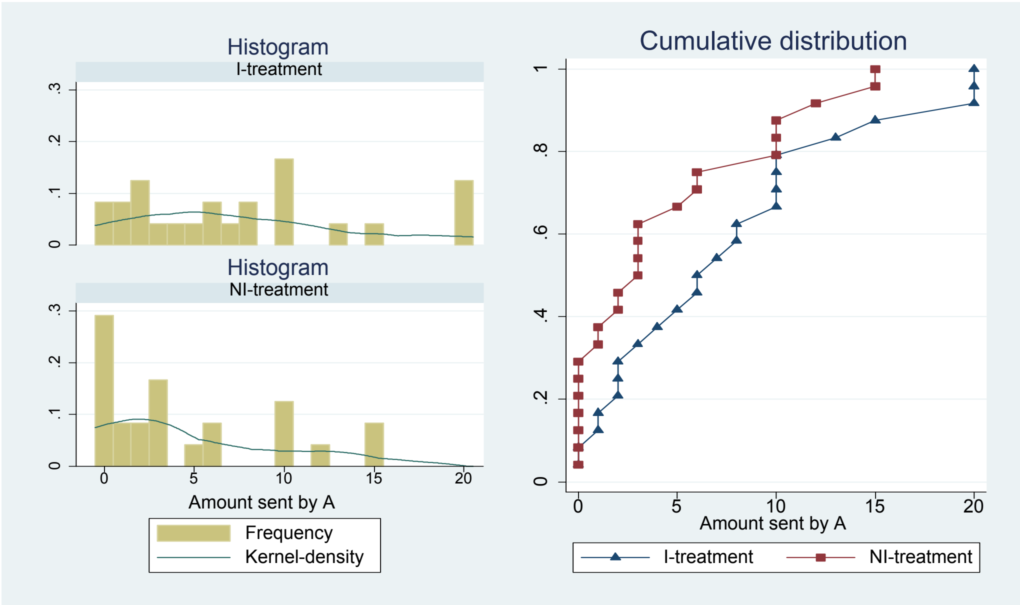

Figure 2: Distribution of amounts sent by B (SM), by treatment

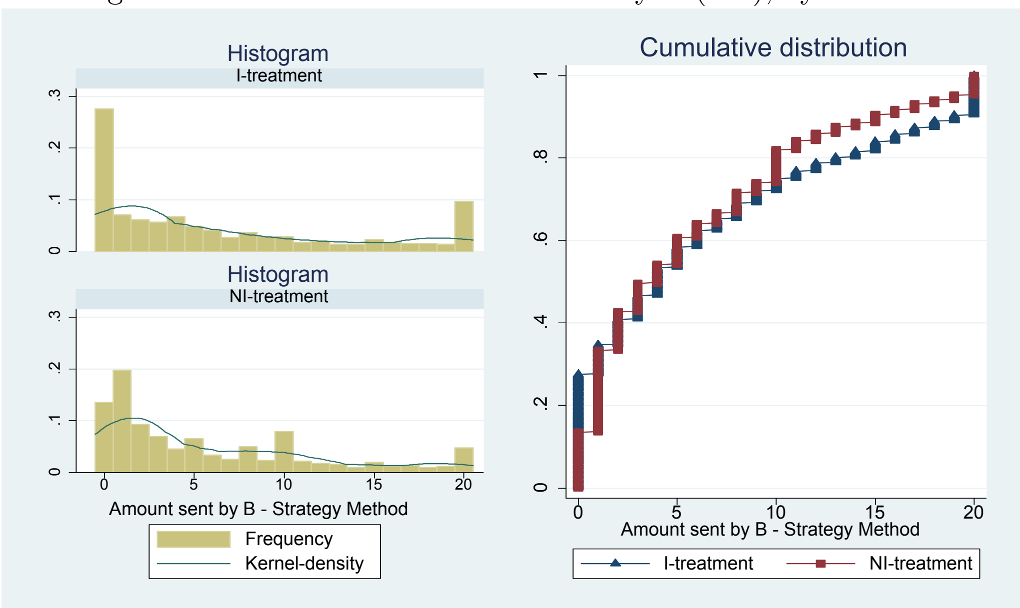

Figure 3: Distribution of amounts sent by B (DM), by treatment

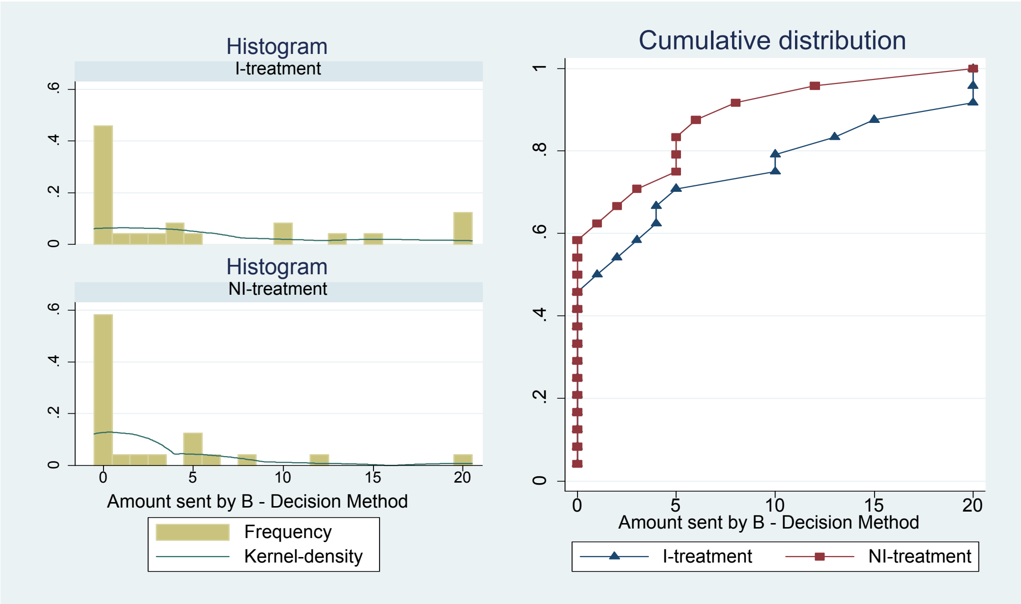

Figure 4: Distribution of Spearman correlations, by treatment

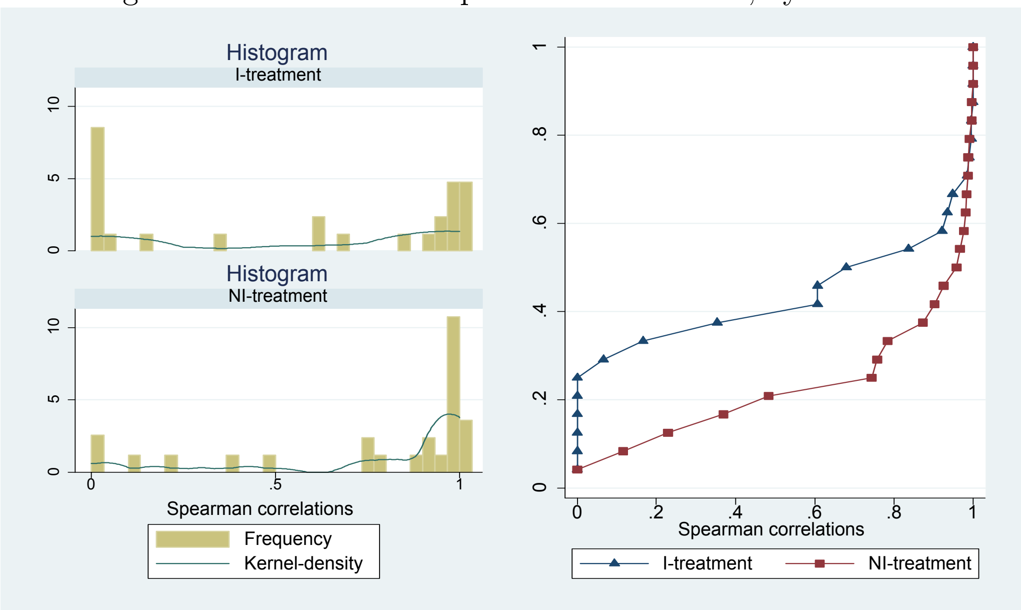

Figure 5: Individual responses of players B (SM): I-treatment

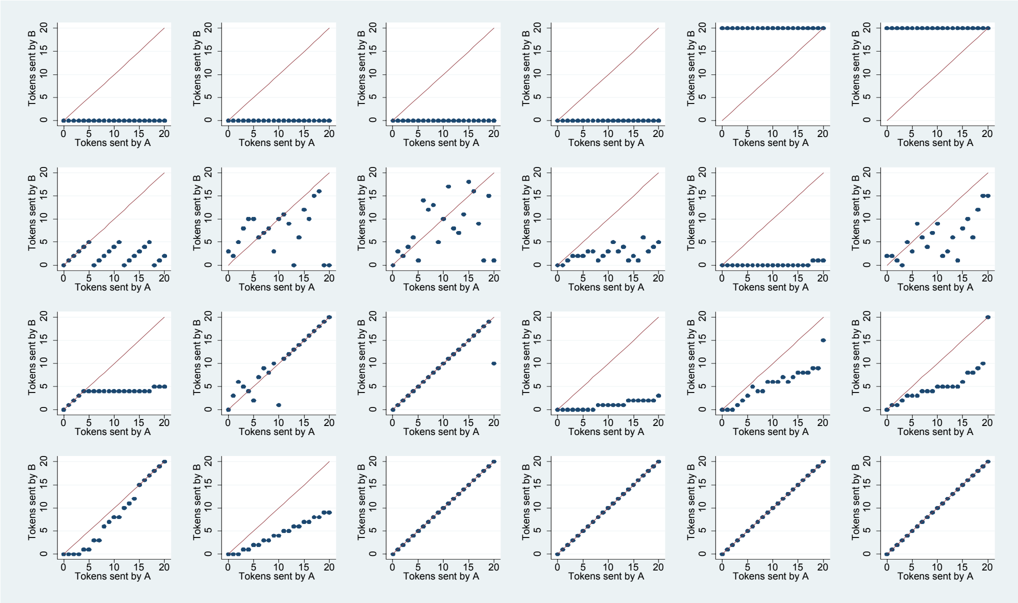

Figure 6: Individual responses of players B (SM): NI-treatment