



**HAL**  
open science

## Testing Theories of Reciprocity: Do Motivations Matter?

Luca Stanca, Luigino Bruni, Luca Corazzini

► **To cite this version:**

Luca Stanca, Luigino Bruni, Luca Corazzini. Testing Theories of Reciprocity: Do Motivations Matter?. *Journal of Economic Behavior and Organization*, 2009, 71 (2), pp.233. 10.1016/j.jebo.2009.04.009 . hal-00693819

**HAL Id: hal-00693819**

**<https://hal.science/hal-00693819>**

Submitted on 3 May 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Accepted Manuscript

Title: Testing Theories of Reciprocity: Do Motivations Matter?

Authors: Luca Stanca, Luigino Bruni, Luca Corazzini

PII: S0167-2681(09)00117-6  
DOI: doi:10.1016/j.jebo.2009.04.009  
Reference: JEBO 2376



To appear in: *Journal of Economic Behavior & Organization*

Received date: 22-5-2007  
Revised date: 20-4-2009  
Accepted date: 20-4-2009

Please cite this article as: Stanca, L., Bruni, L., Corazzini, L., Testing Theories of Reciprocity: Do Motivations Matter?, *Journal of Economic Behavior and Organization* (2008), doi:10.1016/j.jebo.2009.04.009

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# Testing Theories of Reciprocity: Do Motivations Matter?\*

Luca Stanca<sup>†</sup>, Luigino Bruni<sup>‡</sup>, Luca Corazzini<sup>§</sup>

Revised version - July 2008

## Abstract

One of the key issues for understanding reciprocity is how agents evaluate the kindness of an action. In this paper we investigate experimentally the hypothesis that the motivation driving an action is relevant for its perceived kindness and, as a consequence, for reciprocal behavior. In particular, we examine the hypothesis that, for a given distributional outcome, positive reciprocity is less strong in response to strategically motivated actions than to non-strategically motivated actions. Our results indicate that, both at the aggregate and the individual level, reciprocity is significantly stronger when strategic motivations can be ruled out. These findings suggest that intentions matter and, in particular, that models of intention-based reciprocity should take into account the nature of the motivations behind choices.

**Keywords:** Reciprocity, Intrinsic Motivation, Laboratory Experiments.

**JEL codes:** D63, C78, C91.

---

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

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

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

<sup>§</sup>Dipartimento di Scienze Economiche "Marco Fanno," University of Padua, Via Del Santo 33, 35123 Padova, Italy. Email: luca.corazzini@unipd.it

# Testing Theories of Reciprocity: Do Motivations Matter?\*

Luca Stanca,<sup>†</sup> Luigino Bruni,<sup>‡</sup> Luca Corazzini<sup>§</sup>

Final version, April 20, 2009

## Abstract

One of the key issues for understanding reciprocity is how agents evaluate the kindness of an action. In this paper we investigate experimentally the hypothesis that the motivation driving an action is relevant for its perceived kindness and, as a consequence, for reciprocal behavior. In particular, we examine the hypothesis that, for a given distributional outcome, positive reciprocity is less strong in response to strategically motivated actions than to non-strategically motivated actions. Our results indicate that, both at the aggregate and the individual level, reciprocity is significantly stronger when strategic motivations can be ruled out. These findings suggest that intentions matter and, in particular, that models of intention-based reciprocity should take into account the nature of the motivations behind choices.

**Keywords:** Reciprocity, Intrinsic Motivation, Laboratory Experiments.

**JEL codes:** D63, C78, C91.

---

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

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

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

<sup>§</sup>Department of Economic Science "Marco Fanno," University of Padua, Via del Santo, 33, 35123 Padova, Italy, and ISLA, Bocconi University, Via Roentgen, 1, 20136, Milan, Italy. Email: luca.corazzini@unipd.it

# 1 Introduction

A large number of experimental and field studies indicate that economic decisions are in many cases motivated not only by material self interest, but also by concerns for fairness. This evidence has led to the development of several recent theoretical models that incorporate fairness as a determinant of economic behavior (see e.g. Fehr and Gächter, 2000, Sobel, 2005, and Fehr and Schmidt, 2006, for recent surveys). Alternative theoretical approaches differ with respect to how fairness is defined. In particular, two main classes of models can be distinguished: models that focus on distributional concerns, and models that focus on intention-based reciprocity.<sup>1</sup>

In the distributional approach, fairness refers to the *distribution* of material payoffs. Economic agents are motivated not only by their own material gain, but also by how their payoff compares with that of other agents. Fehr and Schmidt (1999) assume that the utility of a subject depends on the difference between his own payoff and that of other subjects, so that agents have egalitarian preferences. Bolton and Ockenfels (2000) assume that the utility function of a subject depends on his own payoff relative to the average overall payoff, so that agents care about their own relative status. In these models, fairness-related preferences depend only on the final distribution of payoffs, so that agents are not concerned about how a given distribution has been obtained.

In the reciprocity approach, fairness refers to the *intentions* of other agents. Agents derive utility from rewarding kind actions and punishing unkind actions, even if this is costly in terms of material payoffs (e.g. Rabin, 1993; Dufwenberg and Kirchsteiger, 2004). Preferences depend on the perceived kindness of an action and, therefore, on the beliefs about other agents' intentions (why an agent has chosen a given action).<sup>2</sup> In these models, actions with identical outcomes may elicit different reciprocating responses depending on how they are interpreted. A key question for intention-based reci-

---

<sup>1</sup>In this paper we use the term reciprocity to refer to strong reciprocity, defined as the non-strategic conditional behavior to reward kind actions, and to punish unkind actions, even if this is costly for the reciprocating subject.

<sup>2</sup>Both distribution and intentions play a role in the models by Charness and Rabin (2002) and Falk and Fischbacher (2006). In particular, in the theory of reciprocity by Falk and Fischbacher (2006) the kindness of an action depends on both intentionality and the outcome of an action, where the latter is defined as the difference in the payoffs of the receiving and sending subjects.

procuity models is therefore *how* agents evaluate the kindness of a particular action.

One way of assessing the kindness of an action is to compare the action intentionally chosen with the alternative actions that could have been chosen, thus focusing on the strategy space of the first mover. Both intentionality, intended as free-will, and the set of alternative possibilities therefore may contribute to define the perceived kindness of an action. This implies two testable predictions. First, there should be no intention-based reciprocal behavior when the action of the first mover is not chosen intentionally, for example because it is the only available option or it is determined exogenously, by a disinterested third party or by chance. Second, the perceived kindness of an intentionally chosen action depends on the characteristics of the alternative actions that were available to the agent but were not chosen.

At the empirical level, a first group of experimental studies has investigated the role of intention-based fairness by focusing on the first prediction, testing the relevance of first mover's intentionality (the so-called attribution hypothesis). A control treatment where the sender can intentionally choose what action to take among a set of alternatives (thus signalling her intentions) is 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 randomly, as in Blount (1995) and Falk et al. (2008). The evidence, however, is mixed, and different results are obtained for positive and negative reciprocity.<sup>3</sup> It is important to observe that the notion of intentions investigated in this literature refers to the attribution of first mover's intentionality (free will).

A second group of experimental studies investigates the role of fairness intentions by focusing on the second testable prediction, testing the relevance of the alternative actions available to the first mover for the perceived kindness of a chosen action. In these studies, the strategy space of the first

---

<sup>3</sup>Bolton et al. (1998) study both positive and negative reciprocal behaviour, finding that distributional preferences are sufficient to explain observed reciprocal actions, whereas intentions play a marginal role. Blount (1995) finds significant evidence of attribution-based behavior only for negative reciprocity (see also Offerman, 2002). Charness (2004) compares a standard gift-exchange game to a treatment where the wage is determined randomly, finding that the slope of the relationship between wage and effort is significantly higher when wages are chosen by the employer. This lends some support to the role of intentions for positive reciprocity, although most of the reciprocal action can be attributed to distribution. Falk et al. (2008) find that the attribution of fairness intentions has a large and significant impact on both positive and negative reciprocal behavior.

mover is manipulated in ways that are strategically irrelevant, but potentially relevant for assessing the fairness of intentions (e.g. Andreoni et al., 2002, Brandts and Solà, 2001, Falk et al., 2003). These studies generally indicate that the perceived fairness of intentions is sensitive to alternative strategy spaces. Bolton and Ockenfels (2005) provide evidence that both distributional factors (relative shares) and strategy spaces (available actions) matter for fairness behavior. Bolton et al. (2005) study experimentally the influence of procedural fairness on the pattern of acceptance and resistance to different outcomes, finding that choice behavior is sensitive to procedural fairness. Overall, however, the evidence on the role of non-distributional factors for models of social preferences is not conclusive. In particular, what determines the perceived kindness of an action remains an open question.

In this paper we propose a new approach for assessing the relevance of intention-based theories of fairness. We formulate and test experimentally the hypothesis that the nature of the *motivations* driving an action plays an important role for its perceived kindness and, as a consequence, for the reciprocal response to that action. We therefore focus on the behavioral relevance for reciprocity of the type of motivation driving the action an agent is responding to. Following Falk and Fishbacher (2006), we propose a framework to model explicitly the effect of the nature of motivations on reciprocal behavior, and test the hypothesis that, for a given distributional outcome, an action is perceived to be less kind if it is strategically motivated (driven by the expectation of a higher future payoff), than if it is not strategically motivated.

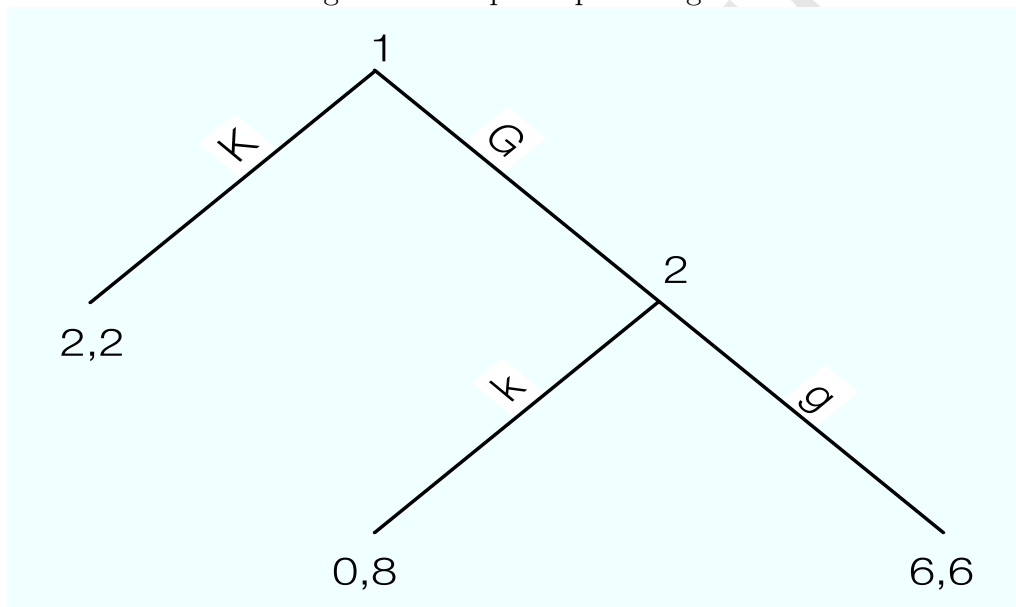
To clarify, consider as an example the sequential game in figure 1. Player 1 moves first, choosing between the actions  $K$  and  $G$ . If player 1 chooses  $K$ , the game ends. If player 1 chooses  $G$ , player 2 chooses between  $k$  and  $g$ , and the game ends. Payoffs in monetary units are indicated by the numbers at the end of each path.<sup>4</sup> Should  $G$  be perceived as a kind action? It clearly depends on the strategy space of player 1. If the action  $K$  was not available, so that the first mover could only choose  $G$ , his action would be perceived differently (presumably, as less kind). Intentionality should matter. Similarly, if the payoffs from choosing  $K$  were different, or there were other alternative actions with different payoffs, intentionally choosing  $G$  could be

---

<sup>4</sup>The game represented in figure 1 is a simplified version of the game used in the experiment, presented in detail in section 3. Payoffs, in particular, are intended to reflect those available to the subjects in the experiment.

perceived differently. It is important to observe, however, that the perceived kindness of action  $G$  also depends on what player 1 believes that player 2 will do. Assume that player 1 believes that player 2 can *only* choose  $k$ . Clearly, relative to the original setting, intentionally choosing action  $G$  would be perceived differently (presumably, as more kind) by player 2. This hypothesis, the different effect of strategic and non-strategic motivations on the perceived kindness of an action, has received relatively little attention in the empirical literature.

Figure 1: Sample sequential game



In order to test this hypothesis, we propose a new experimental design, based on a symmetric gift-exchange game, that allows to manipulate the beliefs of the first mover about the strategy space of the second mover. The game is in two phases: in phase 1 a sender gives to a receiver; in phase 2 the same game as in phase 1 is played with reversed roles. We compare two treatments varying the information sets of the players: in one treatment, in phase 1 subjects are not aware of phase 2; in the second treatment, in phase 1 subjects are aware of phase 2. As a consequence, in the first treatment giving in phase 1 cannot be strategically motivated, while in the second treatment giving in phase 1 can be strategically motivated, i.e. driven by the



expectation of a higher payoff in phase 2. The two treatments only differ with respect to the nature of the motivations of the sender in phase 1, whereas the distributional outcome of the sender's action and the sender's willingness to give are kept constant across treatments. We expect reciprocity to be stronger in the first treatment, where the sender's action cannot be perceived to be driven by strategic motivations, than in the second treatment, where it can be perceived to be strategically motivated.

The results indicate that the type of motivation behind the choice of an agent has a significant impact on the reciprocal behavior of other agents. When the experimental design rules out the attribution of strategic motivations, second movers' responses are characterized by significantly stronger positive reciprocity. This result holds both for strategy profiles and for actual decisions. In particular, at the individual level, a large number of second movers display unconditional behavior when first movers' can be strategically motivated.

The paper is structured as follows. Section 2 presents the theoretical framework 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 The Effect of Motivations on Reciprocity

In this section, following Falk and Fishbacher (2006), we propose a framework to model the effect of the motivations of an agent on the reciprocal behavior of another agent. Consider agent  $i$ , who is the second-mover in a one-shot sequential interaction with agent  $j$ . The utility function of agent  $i$  is assumed to depend not only on material payoffs ( $\pi_i$ ), but also on concerns for fairness, represented by a distribution component and a reciprocity component. The distribution component is expressed as the product of a distributional sensitivity parameter and a distribution measure. The reciprocity component is expressed as the product of a reciprocity parameter, a kindness term and a reciprocation term:

$$U_i(\pi_i, \pi_j) = \pi_i + \alpha_i \delta_i + \rho_i \phi_i \sigma_i \quad (1)$$

The parameter  $\alpha_i$  represents the agent's sensitivity to distributional factors. The distribution measure ( $\delta_i$ ) measures distributional fairness.<sup>5</sup> The reciprocity parameter ( $\rho_i$ ) represents the agent's sensitivity to reciprocity. The kindness term ( $\phi_i$ ) measures how kind the agent perceives the action undertaken by the other agent. The reciprocation term ( $\sigma_i$ ) measures the effect of the reciprocal action on the other agent's utility. Depending on the relative size of the parameters  $\alpha_i$  and  $\rho_i$ , and on the specification of  $\delta_i$ ,  $\phi_i$ , and  $\sigma_i$ , the distributional and intention-based reciprocity components may have a different relative weight in the agent's preferences.

Focusing on the reciprocity component, the question we are considering is what determines  $\phi$ , the perceived kindness of an action. In Falk and Fishbacher (2006), the kindness term depends on both the outcome of the action and the underlying intention:

$$\phi_i = \Delta_i \vartheta_i \quad (2)$$

where the outcome term  $\Delta_i$  is defined as the difference between the second mover's payoff and the first mover's payoff ( $\pi_i - \pi_j$ ), and the intention factor  $\vartheta_i$  is a coefficient between 0 and 1 that parametrizes the intentionality of the action, with  $\vartheta_i = 1$  describing a fully intentional action and  $\vartheta_i < 1$  an action not fully intentional.

In this paper we argue that the motivation driving an action is also relevant for its perceived kindness, so that the kindness term depends not only on the outcome of the action ( $\Delta$ ) and the intentionality the action ( $\vartheta$ ), but also on the nature of the motivation driving the action one is responding to (see e.g. Gouldner, 1960). Perceived kindness therefore depends not only on whether an action produced a favorable outcome and on whether the action was intentionally chosen, but also on the reason why the action was chosen.

The expression for the kindness term in (2) should be extended as follows:

$$\phi_i = \Delta_i \vartheta_i \mu_i \quad (3)$$

where  $\mu$  is the *motivation factor*, a parameter between 0 and 1 that characterizes the type of motivation driving an action. We assume that an action

---

<sup>5</sup>In Fehr and Schmidt (1999) the utility function depends negatively on the difference between the agent's payoff and the payoffs of the other agent. In Bolton and Ockenfels (2000) the utility function, strictly concave in the agent's share of total payoffs, depends negatively on the difference between the agent's payoff and the average payoff of other agents.

is perceived to be less kind if it is strategically motivated ( $\mu_i < 1$ ) than if it is not ( $\mu_i = 1$ ).<sup>6</sup>

This implies that, for a given outcome ( $\Delta$ ) and intention ( $\vartheta$ ), the kindness term  $\phi$  is smaller in response to strategically motivated actions (S), than to non-strategically motivated actions (NS):

$$\phi_S < \phi_{NS} \quad (4)$$

or, alternatively,

$$\phi_S = (1 - \beta) \phi_{NS} \quad (5)$$

with  $\beta > 0$ . As a consequence, strategic motivation of the first mover results in weaker reciprocity of the second mover as opposed to non-strategic motivation. The null hypothesis, that the nature of motivations is irrelevant for perceived kindness, can be formulated as  $\beta = 0$ , versus the alternative hypothesis that  $\beta > 0$ .<sup>7</sup>

In order to test this hypothesis, we have designed an experiment based on a gift-exchange game under two treatments. In one treatment, the motivation of the first mover can only be perceived as non-strategic, while in the second treatment it can also be perceived as strategic.

### 3 Experimental design and procedures

Our experiment is based on a symmetric version of the gift-exchange game (e.g. Fehr et al. 1993, Gächter 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 game, then present the two treatments, and finally illustrate the procedures of the experiment.

---

<sup>6</sup>An action is strategically motivated if it is driven by the expectation of a higher future payoff.

<sup>7</sup>Note that models based on guilt aversion, as in Battigalli and Dufwenberg (2007), or on trust responsiveness, as in Bacharach et al. (2007), would imply the opposite restriction  $\beta < 0$ . In these models, agents derive utility from meeting others' expectations about them.

### 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 should not 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.<sup>8</sup> In the *Information* treatment (I), *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

---

<sup>8</sup>Ben-Ner et al. (2004) use a similar design, based on a two-part dictator game in which dictators are unaware of the two-part structure, to study reciprocal behavior in a setting where strategic investment in reputation by first movers can be ruled out.

1 but with reversed roles. Therefore, this is simply a version of the standard gift-exchange game. In the *No-Information* treatment (NI), 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.<sup>9</sup>

The two treatments differ with respect to the motivation that may determine A's action in stage 1. In the I-treatment, A can give tokens to B for pure altruism, concern for efficiency, and/or because she expects that B will reciprocate, thus increasing her own overall payoff. Player A's motivation, can therefore be perceived by player B, at least partially, strategically motivated (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 cannot be perceived by player B as strategically motivated.

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 motivations 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* cannot be strategic: 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.

---

<sup>9</sup>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.

### 3.3 Procedures

In stage 2, when players B have to make their choice, 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, on the basis of responses to different actions of A, to distinguish between unconditional altruism and conditional altruism 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 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.<sup>10</sup>

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). 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.

---

<sup>10</sup>See Fischbacher et al. (2001) for a similar approach.

## 4 Predictions

If subjects are purely self-interested, in both treatments players B will choose to give zero tokens for any number of tokens received, since they are at the terminal node and gift-giving is costly. In the I-treatment, if it is common knowledge that all subjects are purely self-interested and rational, by backward induction the optimal choice of players A is to give zero in stage one. In 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 sub-game perfect equilibrium outcome is for all players to give zero: player A will choose  $a = 0$  and player B will choose  $b = 0$  for any  $a$ .

If subjects' preferences are characterized by concerns for fairness (determined by distributional factors or intention-based reciprocity), in both treatments the response of players B should depend positively on the amount sent by player A:

**Hypothesis 1** - If preferences are characterized by a concern for fairness, the amount sent by B in stage 2 is positively related to the amount sent by A in stage 1.

Our operational definition of fairness-dependent behavior 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.<sup>11</sup>

Note that if some players are motivated by concerns for fairness, 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 or efficiency, but also by the expectation that a reciprocating response could increase their own payoff. This is an additional reason why the focus of our analysis is on the responses of players B in the strategy method: this allows us to compare the two treatments, characterized by different motivations of first movers, 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.

---

<sup>11</sup>In order to enable a comparison of the two indicators, we also report Pearson correlation coefficients.

Let us turn to the main hypothesis of the experiment. In a relationship between a first mover and a second mover, if the action of the first mover can be strategically motivated, it is perceived by the second mover as less kind, so that positive reciprocity is elicited less strongly than if the action is not perceived to be strategically motivated. In our design, if the second mover is motivated only by distributional fairness, or the perceived kindness of an action does not depend on motivations, the positive relationship between the amounts sent by A in stage 1 and by B in stage 2 should be *the same* across treatments. On the contrary, if the motivation of an action matters for its perceived kindness, such positive relationship should be different across treatments.

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

Table 1 summarizes the predictions of alternative fairness models for the two treatments. The key hypothesis to be tested is that the reciprocating behavior of the second mover should be stronger in the NI-treatment, where strategic motivations of the first mover can be ruled out, than in the I-treatment, where strategic motivations of the first mover cannot be ruled out.

Table 1: Summary of predictions for player B

Model	NI-treatment	I-treatment
Standard prediction	$b = 0$	as in NI
Distribution-based fairness	$b$ rises with $a$	as in NI
Intention-based reciprocity:		
Only outcome	$b$ rises with $a$	as in NI
Outcome and attribution	$b$ rises with $a$	as in NI
Outcome, attribution <i>and</i> motivations	$b$ rises with $a$	$b$ rises with $a$ but less

*Note:* See discussion in section 4.



## 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 by e-mail using a list of voluntary potential candidates. None of the subjects had participated previously in trust or gift-exchange games. Sessions lasted approximately 45 minutes. No show-up fee was paid and the exchange rate was 2 tokens = 1 euro. The average payment was 14.9 euros, with payments ranging between 0 and 40 euros. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).

Table 2 reports summary statistics by treatment and player type. The average amount sent by A players is 4.46 tokens in the NI-treatment (median=3). This suggests that players A are not driven purely by self-interest, as they give almost 25 per cent of their endowment on average in the absence of extrinsic motivation. In the I-treatment the average amount sent by players A is 7.63 tokens (median=6.5). The difference between the average amounts sent by players A in the two treatments is statistically significant at the five per cent level, using a Mann-Whitney rank-sum test for a two-sided hypothesis based on 24 independent observations. This indicates that self-interest also plays a substantial role in determining the decisions of players A. Average responses by players B in strategy method are relatively similar in the two treatments (6.41 and 5.77, respectively), whereas the difference in actual decisions is much larger (5.29 and 2.79 in the NI- and I-treatment, respectively). However, both for strategies and decisions the difference between the amounts sent by B-players in the two treatments is not statistically significant, using a Mann-Whitney rank-sum test (p-values are 0.48 and 0.14, respectively).<sup>12</sup>

Let us turn to the test of hypothesis 1. Table 3 reports correlation coefficients between the responses of B players and the amounts sent by A players, within each treatment. If we consider the strategy profiles of B players (responses in strategy method), 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,

---

<sup>12</sup>For responses in strategy method, in order to have independent observations we computed the average amount sent by each B player and tested the null of no difference between treatments on 24 independent observations.

Table 2: Summary statistics of amounts sent, by player and treatment

	Means		Medians		N.Obs.
	Info	No Info	Info	No Info	
Amount sent by A	7.63	4.46	6.50	3.00	24
Amount sent by B (SM)	6.41	5.77	4.00	4.00	504
Amount sent by B (DM)	5.29	2.79	1.50	0.00	24

*Note:* Info = Information treatment. No Info = No Information treatment.

respectively).<sup>13</sup> These results clearly indicate that players B's strategies are characterized by a concern for fairness.

Table 3: Reciprocity within treatments

	Strategy method		Decision method	
	Info	No Info	Info	No Info
Spearman correlation	0.35**	0.52**	0.57**	0.76**
Pearson correlation	0.37**	0.46**	0.53*	0.70**

*Note:* Info = Information treatment. No Info = No Information treatment. Columns 1 and 2 report correlations for individual strategies, columns 3 and 4 report correlations for individual decisions. \* and \*\* indicate p-value <0.05 and <0.01, respectively, for a two-sided null hypothesis of zero correlation. All tests are based on 24 independent observations.

It could be observed that the positive relationship between the amounts sent and received might 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 correlations 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 since 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 3, correlation coefficients for the responses of players B to the actual amounts sent by players A (decision method), for the two treatments.

<sup>13</sup>In order to have independent observations 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.

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 and 0.76 for I- and NI-treatment, respectively). Pearson correlations are 0.53 and 0.70, respectively, in the two treatments.

**Result 1:** The null hypothesis of absence of fairness concerns for B subjects can be strongly rejected within each of the two treatments, both for strategies and actual decisions.

The results in table 3 provide a qualitative indication that fairness is elicited differently in the two treatments: correlation coefficients are larger in the NI-treatment, where strategic motivations can be ruled out, than in the I-treatment. We therefore turn to formal tests of hypothesis 2. We start by analyzing individual behavior, then examine aggregate behavior using regression analysis.

Since players B had to provide a response for each feasible action of players A, we can study the differences in reciprocating behavior between the two treatments at the individual level. Figure 2 displays the histogram of the Spearman correlations in each of the two treatments, and the corresponding cumulative distributions. Table 4 reports individual correlation coefficients between the responses of B players and each of the possible amounts sent by A players, for each of the two treatments. Both the table and the figure indicate that individual correlations are larger in the NI-treatment.

Table 5 reports the results of the test of the null hypothesis that correlation coefficients are the same in the two treatments. In particular, using a t-test, the null hypothesis of equal means across treatments is rejected for both Spearman and Pearson correlations (the p-value for a two-sided test is 0.05). Using a Mann-Whitney rank-sum test, as suggested by the nature of the distribution of the correlation coefficients and the small sample size, would lead to rejection of the equal median null hypothesis at the 20 per cent significance level (the p-values for a two-sided test are 0.19 for the Spearman correlation and 0.16 for the Pearson correlation coefficients, respectively). Kolmogorov-Smirnov test-statistics do not lead to reject the null of equality of distributions, but provide an indication that the distribution of correlations in the NI-treatment dominates the corresponding distribution for the I-treatment. These findings indicate that, at the individual level, reciprocity is stronger in the NI-treatment.

Individual strategy profiles, reported in figures 3 and 4, help to interpret this result: not only do subjects reciprocate more in the NI-treatment

Table 4: Individual Spearman correlations, by treatment

	Info-treatment	No Info-treatment
Pair 1	0.00	0.00
Pair 2	0.00	0.12
Pair 3	0.00	0.23
Pair 4	0.00	0.37
Pair 5	0.00	0.48*
Pair 6	0.00	0.74**
Pair 7	0.07	0.76**
Pair 8	0.17	0.78**
Pair 9	0.35	0.87**
Pair 10	0.61**	0.90**
Pair 11	0.61**	0.93**
Pair 12	0.68**	0.96**
Pair 13	0.84**	0.97**
Pair 14	0.92**	0.98**
Pair 15	0.94**	0.98**
Pair 16	0.95**	0.98**
Pair 17	0.99**	0.99**
Pair 18	0.99**	0.99**
Pair 19	1.00**	0.99**
Pair 20	1.00**	1.00**
Pair 21	1.00**	1.00**
Pair 22	1.00**	1.00**
Pair 23	1.00**	1.00**
Pair 24	1.00**	1.00**

*Note:* I = Information treatment. NI = No Information treatment. The figures reported are Spearman rank correlation coefficients, sorted by size, between the amount received and the amount sent by B players in strategy method. \* and \*\* indicate p-value <0.05 and <0.01, respectively.

Figure 2: Distribution of Spearman correlations, by treatment

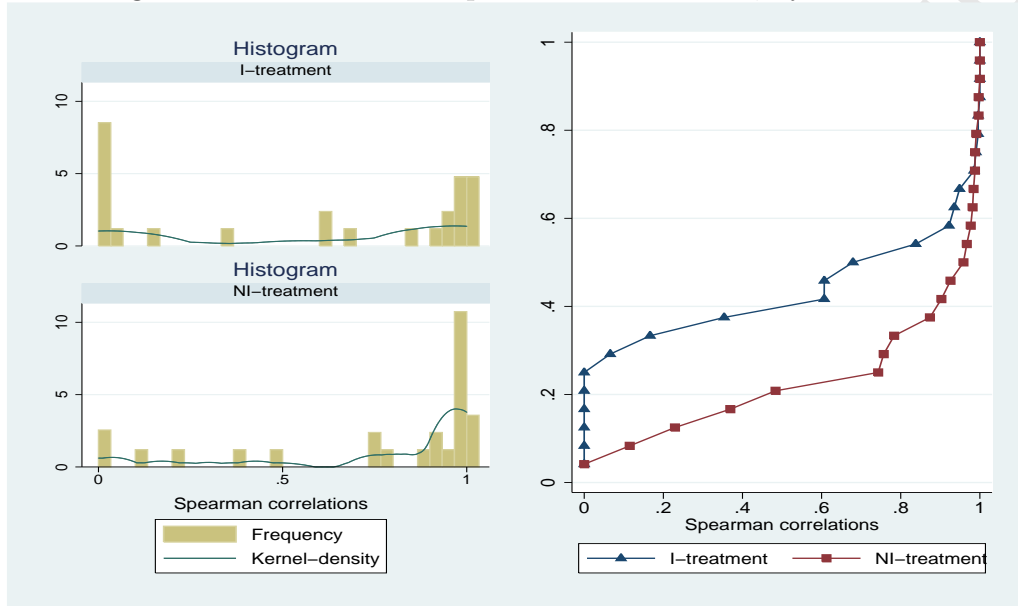


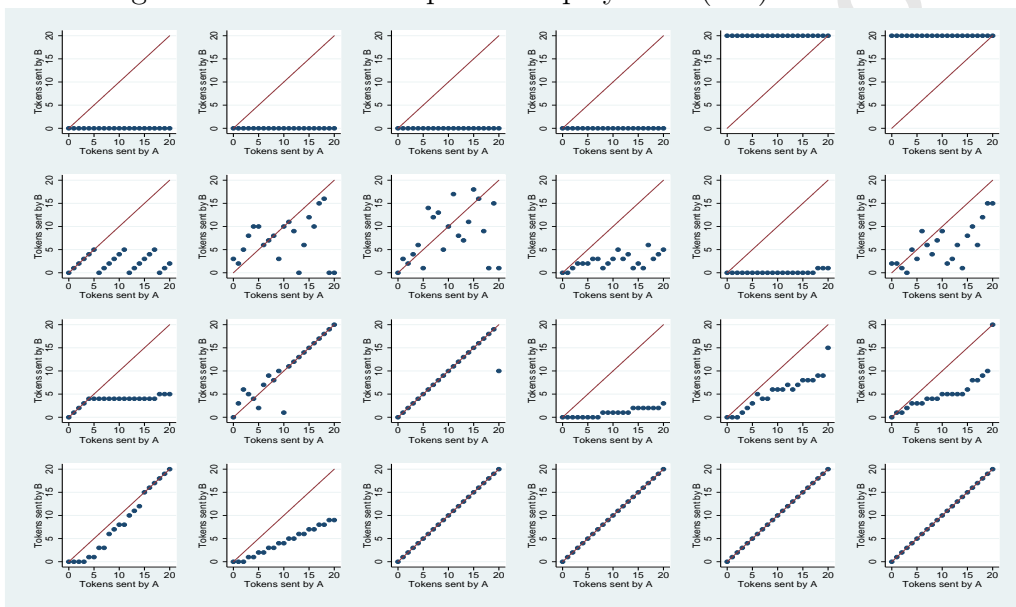
Table 5: Test for differences in distributions of correlations

	T-stat	P-val	U-stat	P-val	KS-stat	P-val
Spearman correlation	-1.88	0.05	-1.28	0.19	0.29	0.26
Pearson correlation	-1.86	0.05	-1.43	0.16	0.29	0.26

*Note:* the table reports results of alternative tests of the null hypothesis that average correlation coefficients are the same in the Information treatment and the No-Information treatment.

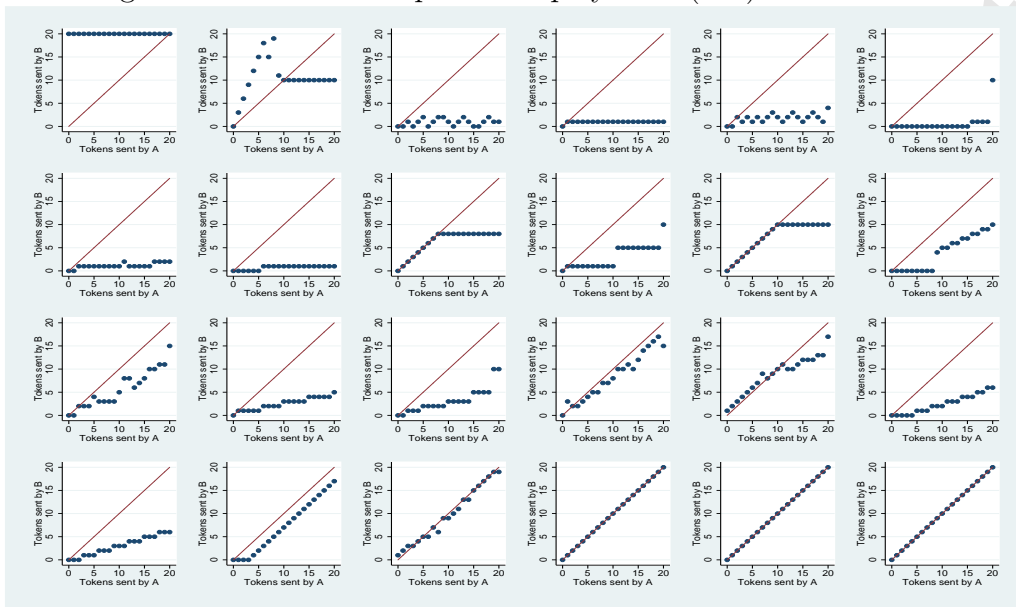
on average, but also a larger fraction of the subjects do not display any reciprocal behavior when strategic motivations cannot be ruled out. In the I-treatment, 6 out of 24 players B display an unconditional strategy profile (zero correlation with player A's decision), as opposed to only 1 out of 24 in the NI-treatment.

Figure 3: Individual responses of players B (SM): I-treatment



Let us now turn to aggregate behavior. We start 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 and 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. We also include a dummy variable for the NI-treatment and an interaction term to assess the differences between the two treatments. The results indicate that strategies explain about 60 per cent of the overall variability of decisions, and that choices in the decision method were more selfish than in the strategy method in both treatments (coefficient estimates are 0.73 and 0.94 in the Information and No-Information treatments, respectively). The difference in the coefficients between the two treatments indicates that decisions follow strategies more closely in the NI-

Figure 4: Individual responses of players B (SM): NI-treatment



treatment. However the estimated difference is not statistically significant (the corresponding t-statistic is 1.04). This indicates that there are no systematic differences in how subjects respond in strategy and decision method in the two treatments, thus providing support for the use of the strategy method.

Table 7 reports the results obtained by regressing the SM response of B players on A players' action, using observations from both treatments. We define a kind response for players B as sending a positive number of tokens, and an unkind response as sending zero tokens, so that the dependent variable is a dummy variable equal to 0 if  $b = 0$  and equal to 1 if  $b > 0$ . Parameter estimates are obtained using a probit estimator, and test statistics are based on errors clustered on pairs of subjects. We also include among the explanatory variables a dummy variable for the NI-treatment and an interaction term (action by A players 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 consider two specifications. In the first, the explanatory variable is the number of tokens sent by player A in stage 1. The results, reported in column

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-squared	0.61
Number of observations	48

*Note:* OLS estimates. t-statistics reported in brackets. \* and \*\* indicate p-value <0.05 and <0.01, respectively. I = Information treatment. NI = No Information treatment.

(1), indicate that the probability that B players respond kindly depends positively on the number of tokens received in stage 1. The coefficient for the interaction term is positive and strongly significant (t-statistic=2.69), consistently with the hypothesis that reciprocal behavior is stronger in the NI-treatment. In the second specification, the explanatory variable is defined as a dummy variable, equal to 1 if A sends a positive number of tokens and 0 otherwise. The results, reported in column (2), indicate that the probability that B players respond kindly is significantly higher if player A's action in stage 1 was kind. As above, the coefficient for the interaction term is positive, although only marginally significant (the t-statistic is 1.61). Overall, these results indicate that also at the aggregate level reciprocal behavior is stronger in the NI-treatment.

**Result 2:** The positive relationship between the amounts sent and received is stronger in the NI-treatment than in the I-treatment, both at aggregate and individual level.

## 6 Discussion

This paper presented an experimental investigation of the hypothesis that the motivations driving an action matter for its perceived kindness and, as



Table 7: Regressions of B's responses to A's actions (SM)

	(1)	(2)
Amount sent by A (dummy)	1.65** (4.75)	
Amount sent by A (dummy) * 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 (-0.40)	-0.15 (-0.53)
Constant	-0.97** (-3.14)	0.18 (0.90)
Pseudo- $R^2$	0.14	0.15
Number of observations	1008	1008

*Note:* Dependent variable: SM response by B subjects, defined as 0 if  $b = 0$  and 1 if  $b > 0$ . Probit estimates, t-statistics reported in brackets (standard errors clustered on pairs of subjects). \* and \*\* indicate p-value  $< 0.05$  and  $< 0.01$ , respectively. I = Information treatment, NI = No Information treatment.

a consequence, for the reciprocal response to that action. In particular, we tested the hypothesis that, for a given distributional outcome, an action is perceived to be less kind if it is strategically motivated than if it is not strategically motivated. In order to test this hypothesis, we proposed a new experimental design, based on a symmetric gift-exchange game, that allows us to manipulate the beliefs of the first mover about the strategy space of the second mover.

The results of the experiment indicate that when strategic motivations of the first mover can be ruled out, second movers' responses are characterized by stronger positive reciprocity. This result holds for both strategy profiles and actual decisions. The results at the individual level indicate, in particular, that a larger fraction of the subjects display unconditional behavior when strategic motivations cannot be ruled out.

These findings suggest several implications and directions for future research. First, at the theoretical level, concerns for fairness cannot be explained entirely by distributional factors. Intention-based reciprocity also matters. In particular, theoretical models of reciprocal behavior should take into account explicitly the type of motivation driving the action an agent is responding to. An action may elicit reciprocity differently depending on whether it is perceived as strategically motivated or not. At the empirical level, the perceived kindness of an action can be assessed not only on the basis of the strategy space of the first mover, but also on the beliefs of the first mover about the strategy space of the second mover.

Second, our results can be interpreted as providing an inter-personal extension of the concept of motivational crowding-out (e.g. Deci and Ryan, 2000). Motivational crowding-out is defined as the reduction of effort in activities carried out for intrinsic motivation when an instrumental reward, typically monetary, is introduced.<sup>14</sup> Several studies have documented the existence of motivational crowding-out (e.g. Cameron and Pierce, 1996, Eisenberger and Cameron, 1996, Deci et al., 2000, Deci and Ryan, 2000).<sup>15</sup> Self

---

<sup>14</sup>“Intrinsic motivation is defined as the doing of an activity for its inherent satisfactions rather than for some separable consequence. When intrinsically motivated a person is moved to act for the fun or challenge entailed rather than because of external prods, pressures, or rewards. [...] Extrinsic motivation is a construct that pertains whenever an activity is done in order to attain some separable outcome”, Deci and Ryan (2000, p. 56). The role of intrinsic motivation for optimal incentive contracts was examined by Krepes (1997) and Murdock (2002). See also the review in Lindenberg (2001).

<sup>15</sup>Frey (1997) examines the reduction of the effort of intrinsically motivated workers

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. Existing studies of motivational crowding-out have generally investigated differences in the behavior of one agent motivated by either intrinsic or extrinsic rewards. Our results indicate that the extrinsic motivation driving the action of an agent may have a crowding out effect on the intrinsically motivated economic behavior of *another* agent.

A third implication is related to the power of disinterested philanthropy in eliciting reciprocity in helped people (Margalit, 1996): 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, 1975), does gratuity induce more or less reciprocity than a contract (without enforcement)? Our experiment suggests that disinterested gifts elicit more reciprocity.

A final implication refers to interpersonal relationships, in particular within organizations. The literature on motivational crowding-out indicates that the introduction of extrinsic incentives in domains where intrinsic motivation are relevant can reduce effort and efficiency. If disinterested costly actions generate relatively more reciprocity, then it could be more efficient in workplaces, for instance, to manage interpersonal relations not only by means of contracts and incentives, but also 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.

---

when extrinsic rewards are introduced. Gneezy and Rustichini (2000) explore the effects of motivational crowding-out in an experiment on fund-raising. At the theoretical level, Harvey (2005) explains motivational crowding out within a principal-agent model where utility is interpreted as overall satisfaction.

## 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

- Andreoni, J., Brown, P., Vesterlund, L., 2002. What makes an allocation fair? Some experimental evidence. *Games and Economic Behavior* 40, 1-24.
- Bacharach, M., Guerra, G., Zizzo, D. 2007. Is trust self-fulfilling? An experimental study. *Theory and Decision* 63, 349-388.
- Battigalli, P., Dufwenberg, M., 2007. Guilt in games. *American Economic Review: Papers and Proceedings* 97, 171-176.
- Ben-Ner, A., Putterman, L., Kong, F., Magan, D., 2004. Reciprocity in a two-part dictator game. *Journal of Economic Behavior & Organization* 53, 333-352
- Blount, S., 1995. When social outcomes aren't fair: The effect of causal attributions on preferences. *Organizational Behavior and Human Decision Processes* 63, 131-144.
- 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., Ockenfels, A., 2000. ERC: A Theory of equity, reciprocity and competition. *American Economic Review* 90, 166-193.
- Bolton, G.E., Ockenfels, A., 2005. A stress test of fairness measures in models of social utility. *Economic Theory* 25, 957-982.
- Bolton, G.E., Brandts, J., Ockenfels, A., 2005. Fair procedures: Evidence from games involving lotteries. *Economic Journal* 115, 1054-1076.
- Brandts, J., Solà, C., 2001. Reference points and negative reciprocity in simple sequential games. *Games and Economic Behavior* 36, 138-157.
- Buchanan, J., 1975. The Samaritan's Dilemma. In: Phelps, E.S. (Ed.). *Altruism, Morality and Economic Theory*. New York: Russel Sage Foundation.
- Cameron, J., Pierce, W.D., 1996. Reinforcement, reward and intrinsic motivation: A meta-analysis. *Review of Educational Research* 64, 363-423.

- Charness, G., 2004. Attribution and reciprocity in a simulated labor market: An experimental investigation. *Journal of Labour Economics* 22, 665-688.
- Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. *Quarterly Journal of Economics* 117, 817-869.
- Deci, E.L., Koestner, R., Ryan, R.M., 2000. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin* 125, 627-668.
- Deci, E.L., Ryan, R.M., 1985. *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum.
- Deci, E.L., Ryan, R.M., 2000. Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemporary Educational Psychology* 25, 56-60.
- Dufwenberg M., Kirchsteiger, G., 2004. A Theory of sequential reciprocity. *Games and Economic Behavior* 47, 268-298.
- Eisenberger, R.W., Cameron, J., 1996. Detrimental effects of reward: Reality or myth?. *American Psychologist* 51, 1153-1166.
- Falk, A., Fehr, E., Fischbacher, U., 2008. Testing theories of fairness - Intentions matter. *Games and Economic Behavior* 62, 287-303.
- Falk, A., Fehr, E., Fischbacher, U., 2003. On the nature of fair behaviour. *Economic Inquiry* 41, 20-26.
- Falk, A., Fischbacher, U., 2006. A Theory of reciprocity. *Games and Economic Behavior* 54, 293-315.
- Fehr, E., Gächter, S., 2000. Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives* 14(3), 159-181.
- Fehr, E., Kirchsteiger, G., Riedl, A., 1993. Does fairness prevent market clearing? An experimental investigation. *Quarterly Journal of Economics* 108, 437-460.
- Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation, *Quarterly Journal of Economics* 114, 817-868.



- Fehr, E., Schmidt, K.M., 2006. The Economics of Fairness, Reciprocity and Altruism: Experimental Evidence. In: Kolm, S.C., Ythier, J.M. (Eds). Handbook of the Economics of Giving, Altruism and Reciprocity 1, 615-691. Amsterdam: Elsevier.
- Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10, 171-178.
- Fischbacher, U., Gächter, 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., Rustichini, A., 2000. Pay enough or don't pay at all. *Quarterly Journal of Economics* 115, 791-810.
- Gächter, S., Falk, A., 2002. Reputation and reciprocity: Consequences for the labour relation. *Scandinavian Journal of Economics* 104, 1-26.
- Gouldner, A., 1960. The norm of reciprocity. *American Sociological Review* 25, 161-178.
- 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* 26, 549-66.
- Kreps, D.M., 1997. Intrinsic motivation and extrinsic incentives. *American Economic Review* 87, 359-64.
- Lindenberg, S., 2001. Intrinsic motivation in a new light. *Kyklos* 54, 317-42.
- Margalit, A., 1996. *The Decent Society*. Harvard MA: Harvard University Press.
- McCabe, K.A., Rigdon, M.L., Smith, V.L., 2003. Positive reciprocity and intentions in trust games. *Journal of Economic Behavior and Organization* 52, 267-275.
- Murdock, K., 2002. Intrinsic motivation and optimal incentive contracts. *RAND Journal of Economics* 33, 650-71.

- Offerman, T., 2002. Hurting hurts more than helping helps. *European Economic Review* 46, 1423-1437.
- Rabin, M., 1993. Incorporating fairness into game theory and economics. *American Economic Review* 83, 1281-1302.
- Sobel J., 2005. Interdependent preferences and reciprocity. *Journal of Economic Literature* 43, 392-436.