Testing theories of reciprocity: Do motivations matter?

Luca Stanca [a,*], Luigino Bruni [a], Luca Corazzini [b,c]

[a] *Economics Department, University of Milan Bicocca, Piazza dell'Ateneo Nuovo 1, 20126 Milan, Italy*
[b] *Department of Economic Science "Marco Fanno," University of Padua, Via del Santo, 33, 35123 Padova, Italy*
[c] *ISLA, Bocconi University, Via Roentgen, 1, 20136 Milan, Italy*

ARTICLE INFO

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.

© 2009 Elsevier B.V. All rights reserved.

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; 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.[1]

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.

---

* Corresponding author. Tel.: +39 0264483155; fax: +39 0264483085.
E-mail addresses: luca.stanca@unimib.it (L. Stanca), luigino.bruni@unimib.it (L. Bruni), luca.corazzini@unipd.it (L. Corazzini).

[1] 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.

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).[2] In these models, actions with identical outcomes may elicit different reciprocating responses depending on how they are interpreted. A key question for intention-based reciprocity 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.[3] 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 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 Fischbacher (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 Fig. 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.[4] 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 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.

---

[2] 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.

[3] 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.

[4] The game represented in Fig. 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.