An Experimental Test of Direct and Indirect Reciprocity in Case of Complete and Incomplete Information*

Martin Dufwenberg[1], Uri Gneezy[2], Werner Güth[3] & Eric van Damme[4]

May 22, 2000

Abstract: In this paper we report experimental results that relate to the reciprocity experiment of Berg, Dickhaut, and McCabe (1995). We consider direct reciprocity, which means to respond in kind to another person, and indirect reciprocity, understood as rewarding someone else. Another variation concerns the information about the multiplier of donations where we compare the benchmark case with a commonly known multiplier to a condition where the multiplier is known with certainty only to donors. Questions which we try to answer are: Will indirect reciprocity induce higher or lower donations?, will donors with the high multiplier "hide behind the small one?", how do receivers respond to different situations?

JEL codes: C72, C92

Key words: Trust, reciprocity, experiment, game

---

* This paper was conceived while we were all at or visiting the CentER for Economic Research at Tilburg University. We thank Harmut Kliemt for his encouraging and inspiring comments. Martin Dufwenberg and Eric van Damme thank the European Union's HCM-Network project "Games and Markets"; Uri Gneezy and Werner Güth gratefully acknowledge support by the European Union's TMR-project "Structural analysis of household savings and wealth positions over the life cycle".

[1] Department of Economics, Stockholm University

[2] Faculty of Industrial Engeneering and Management, Technion

[3] Department of Economics, Humboldt-University Berlin

[4] CentER for Economic Research, Tilburg University

1. Introduction

Consider the following case: X_2 receives a large amount from Y_1 with whom he had so far no contact at all. X_2 asks Y_1: "Why are you so generous?" Y_1 answers: "Somebody else – X_1! – has been nice to me and I wanted to be grateful." X_2 asks again: "Why didn't you reward X_1?" Y_1 answers: "Oh, I would have loved to. But this was impossible. You are the only one I could reward." X_2 concludes: "Now I understand. Thanks!"

What may appear like a fairy tale can be experimentally implemented so that we can learn whether the behavioral assumptions are realistic or not. Our starting point is the *trust game* introduced by Berg, Dickhaut, and McCabe (1995) (henceforth BDM). Here a donor can give as much as he wants of a monetary endowment to a receiver. The amount given by the donor is tripled, and then the receiver may return as much as he wants to the donor. Thus by trust in reciprocity (see Güth and Kliemt, 1994, for a theoretical discussion) players can achieve an enormous efficiency gain in the BDM trust game. They can triple the amount initially made available. BDM report evidence which they interpret as suggesting that "reciprocity exists as a basic element of human behavior" (p 122). Many other experimental studies claim to have established analogous results, and a few theoretical models for handling reciprocity in economic theory have been proposed. See Fehr & Gächter (2000) for a recent review of this literature.

We perform an experiment in which one treatment is a replication of the BDM set-up. We call this the *direct reciprocity treatment*. In another treatment we transform the BDM game such as to explore what we call indirect reciprocity. In this *indirect reciprocity treatment* a group consisting of four persons interact — two donors and two receivers. Instead of repaying their own donor, as in the direct reciprocity treatment, receivers can only reciprocate towards *the other* donor. Note that this is the

treatment which corresponds to the case described in the opening paragraph of this paper. Note also that the literature surveyed by Fehr & Gächter (2000) is not concerned with indirect reciprocity as we describe it here.

In a third treatment we allow for private information of donors about the factor by which donations are multiplied. In the BDM study, and in our first two treatments, the factor is 3. In our *incomplete information treatment* it is known to all participants that the multiplier is either 2 or 4, each with probability one half but *only donors* are informed about the value of this multiplier. Thus a donor, whose donation will be multiplied by the larger factor of 4, may attempt to "hide his greed" (see Güth, Huck, and Ockenfels, 1996, for experimental evidence) by choosing a donation which looks like a generous one for a multiplier of 2, if such a cunning stratagem is feasible. Private information about the multiplier of donations thus allows us to distinguish between intrinsically motivated donors and those who are only interested in an image of generosity.[5]

In the following section our experimental procedure will be explained in more detail. We then test whether indirect reciprocity induces at least as high donations as direct reciprocity, whether donors with the large multiplier attempt to hide behind the small multiplier, and how receivers react in the various treatments. We finally discuss in the light of these results why receivers reward at all.

2. Experimental design

To prevent any confounding effects a group of four individuals was mentioned in the instructions for each treatment (see Appendix A), even though effectively it was only

[5] Another interesting treatment might be to let only receivers know by which factor donations are multiplied. However, we shall not explore that possibility here.

in the indirect reciprocity case that the players were not engaging in two-player games. Let us denote by X_i for i = 1,2 the two donors and by Y_j for j = 1,2 the two receivers of each group. The distinction between direct and indirect reciprocity is graphically visualized by Figure II.1.

[Insert Fig II.1 here]

In Figure II.1 the variable x_i with $0 \leq x_i \leq e$ is what X_i gives away where e denotes X_i's endowment, i = 1,2. In all treatments e was equal to 10 "points", with each point worth 8 New Israeli Shekels (NIS).[6] What Y_i receives is $3x_i$. From this Y_i can return any amount y_i with $0 \leq y_i \leq 3x_i$. Whereas the receiver of y_i is X_i in case of direct reciprocity, it is X_j with $j \neq i$ when only indirect reciprocity is possible.

The multiplier of donations is commonly known in the complete information treatments. However, *only donors know* whether the value of this parameter is 2 or 4 in the incomplete information treatment. Receivers Y_i were told in their instructions for the incomplete information treatment that half of the donors have the large multiplier of 4 and the other half the smaller one of 2 and that this is commonly known. Thus the expected multiplier is 3 as in the treatments with complete information as well as in the original BDM study.

The monetary payoffs U_i of a donor i = 1,2 and V_i of a receiver i = 1,2 are in the cases of direct reciprocity and complete information

$$U_i = e - x_i + y_i$$
$$V_i = 3x_i - y_i$$

[6] At the time of the experiment, 3.5 NIS equaled approximately one US dollar.

and analogously in the cases of indirect reciprocity for i,j = 1,2, i ≠ j:

$$U_i = e - x_i + y_j$$
$$V_i = 3x_i - y_i$$

In the case of direct reciprocity and incomplete information the factor 3 in the definition of V_i has to be substituted by 2 or 4, respectively.

The "classical" solution would assume that players are only interested in their own monetary payoff: Since receivers Y_i act in subgames the structures of which are effectively dictator games, their optimal decision would be to choose $y_i = 0$ in each treatment. Anticipating that nothing will be given back, donors X_i will therefore avoid positive donations and choose $x_i = 0$. Formally, this solution behavior can be derived by once repeated elimination of (weakly) dominated strategies or as a subgame perfect equilibrium (assuming only equilibrium behavior would not rule out positive choices of y_i following positive donations x_i which are not chosen in some particular equilibrium).

The complete design is visualized in Figure II.2 showing that we do not want to test for interaction effects.

[Insert Fig II.2 here]

Further details of the experimental procedure should become evident from the Instructions (see Appendix A). After a pilot experiment at Uppsala University the main experiments were performed at the University of Haifa. In our (statistical) analysis we just use the results of the main experiments although those of the pilot experiment would have strengthened our conclusions.

In Figure II.3, using the format of Figure II.2, we list the average earnings in NIS, the time (in minutes needed), and the number of participants for each treatment.

[Insert Fig II.3 here]

3. Experimental results

A first impression of the experimental results can be gained from Table III.1 which, with treatments visualized as in Figure II.2, lists for all three treatments the averages, medians, standard deviations, and the numbers of observations both for donations x_i (left column) as well as for amounts returned y_i (right column).

[Insert Table III.1 here]

Whereas the interest rate, defined as

$$r = [(\text{mean } y) / (\text{mean } x) - 1] \cdot 100,$$

is negative (-6.62 %) for direct reciprocity, it is higher (+22.87 %) for indirect reciprocity. This difference of $1 + r = 1.2287$ and $1 + r = .9338$ is, however, not significant ($p = .4351$) with a Mann-Whitney test comparing the different samples of individual interest rates defined as $r_i = [(y_i / x_i) - 1] \cdot 100$ for direct reciprocity and $r_i = [(y_j / x_i) - 1] \cdot 100$ for indirect reciprocity.

Although direct reciprocity inspires more donations on average (the difference is insignificant — $z = .807$, $p = .4197$ with a Mann-Whitney test), these investments in trust in reciprocity are poorly rewarded. For indirect reciprocity such trusting investments seem to offer a reasonable return.

In case of incomplete information the interest rate r = -12.06% is again negative as for direct reciprocity with a commonly known multiplier. Here it is, of course, interesting to distinguish between pairs with the large multiplier of 4 and those with the small multiplier of 2 (see Table III.2).

[Insert Table III.2 here]

The positive difference of donations x_i for the small and the large multiplier is not significant (z = -1.321, p = 0.1864).

We can conclude:

- Indirect reciprocity induces only insignificantly smaller donations than direct reciprocity.

- Donors with the large multiplier do not donate significantly more, i.e. we observe "hiding of greed".[7]

- Receivers are more rewarding in the case of indirect reciprocity in the sense that they pay a positive interest rate r only in case of indirect reciprocity, whereas r is negative in both direct reciprocity treatments. However, this difference is not significant.

We find the last of these results rather surprising. We initially expected higher rates of return in the direct reciprocity treatment than in the indirect reciprocity treatment.

[7] Note that even if the multiplier μ is initially known only by the donor, many choices will unambiguously "signal" its value. For example, for a donation of 7 the receiver gets 14 or 28, and in either case the value of μ will be apparant. Nevertheless, the donor can always "hide" by donating for example 4.

Instead, there is no clear difference between the treatments. If anything, the difference goes the opposite way to what we expected!

4. On rewarding

If receivers reward positive donations, one may view donations as risky investments. If such investments do not pay as in case of the direct reciprocity treatments, one may ask, of course, whether this is due to overly optimistic expectations or to other motives, e.g. a desire for efficiency. To test this, one would have to elicit the beliefs of donors in order to find out how much they expect to get back. In a somewhat similar game, Dufwenberg and Gneezy (2000) elicit beliefs and report some evidence that first-movers may be motivated by efficiency rather than a hope for monetary gain. Conceivably, a similar effect could operate in the current design. Since we did not measure beliefs we do not discuss this further. Instead, we focus on the following issue: Why do some receivers reward?

Unlike in dictator games, where no reciprocity argument can explain positive donations,[8] all our treatments provide a justification for positive rewards, namely that receivers were rewarded before. The hypothesis "positive rewards ($x_i > 0$) trigger positive reactions ($y_j \geq x_i$) for i = j and i↑j" would apply to both, direct and indirect reciprocity. A distinction between direct and indirect reciprocity could be based on the degree of obligation which a positive donation implies. It seems reasonable to assume that one feels more obliged to reciprocate directly than indirectly. After all the direct reciprocity game suggests an implicitly agreed upon mutual exchange whereas any similar justification for the indirect reciprocity game appears rather farfetched.

[8] If one excludes indirect reciprocity in the sense that dictators reward receivers since they got their position for free (from the experimenters).

In the light of this, our finding that the receivers are most rewarding in the indirect reciprocity treatment is odd. An interesting train of thought, suggested to us by Hartmut Kliemt, can connect this result to the "Titmuss debate" on blood donation. In his 1971 book *The Gift Relationship*, Richard Titmuss expressed a strong preference for the British voluntary blood donation system over the market-based system in the U.S. One of his main arguments was that a market for blood, where those who receive blood pay their donors, would cause altruism to decay and the supply of blood to decrease. Titmuss' contention led to a heated debate in which many economists participated (see Stewart (1992) for an account and for a theoretical model which illustrates Titmuss' argument). Conceivably, Titmuss could have used our experimental finding to support his view. To see this, associate a receiver in the trust game with a person who donates blood, the amount he returns with his willingness to give blood, and the person who gets the returned money with the beneficiary of the blood donation. Accepting this analogy[9], our result suggests that the willingness to give blood is crowded out if a blood donor gets a monetary transfer from the person who gets his blood, a conclusion in line with Titmuss' contention.

We are not ready to offer a more detailed explanation for our result. A call for more research, specifically aimed at revealing what psychological forces may be at work, seems warranted. In this connection, let us note an aspect which might be important to consider when designing future related experiments. The obligation to reward if one is rewarded could be strengthened by allowing to observe others' behavior. The reason would be that one feels ashamed if one does not properly reward after a favor (see Hoffman et al., 1994, who claim less rewarding in double blind dictator experiments, and Bolton and Zwick, 1995, for other results). Our experiments were not double-blind, but interaction was anonymous. The direct reciprocity results seem to

[9] This is somewhat more obvious for direct (X sends money to a potential blood donor Y who can donate blood for X) than for indirect (all needy patients X give money hoping that this will induce potential blood donors to donate blood) reciprocity.

suggest that usually receivers feel not at all ashamed not to reward properly. Nevertheless one might perform an experiment where the essential treatment variation concerns how much the others (in one's 4 participant-group) learn about the own behavior. See Güth, Königstein, and Nehring (1999) for some relevant results.[10]

One may ask why receivers should feel at all obliged or ashamed in case of anonymous interaction (without a shadow of the future like it would exist in case of repeated trust games). An evolutionary justification requires that the receiver's moral type can be more or less reliably detected. This is ruled out by anonymity. It is possible, however, that we do not decide in each instance of life anew whether or not we trust, e.g. by donating, or reward trust, e.g. by rewarding donors. In other words: It could be that we cannot shift our behavioral gears between several separate games, but rather play one game of life. Consequently an evolutionary argument should not be based on a highly specific instance and its specific information conditions but on the general situation of human interaction where knowing others' moral types is rather likely. In other words: Experimental participants bring their general attitudes in the laboratory, for instance, a general obligation to reward trust (in reciprocity).

References

Berg, J., J. Dickhaut, and K. McCabe (1995): Trust, reciprocity, and social history, *Games and Economic Behavior* 10, 122-42.

Bolton, G. E. and R. Zwick (1995): Anonymity versus punishment in ultimatum bargaining, *Games and Economic Behavior* 10 (1), 95-121.

Dufwenberg, M. and U. Gneezy (2000): Measuring beliefs in an experimental lost wallet game, *Games and Economic Behavior* 30 (2), 163-182.

Fehr, E. and S. Gächter (2000), "Fairness and retaliation: The economics of reciprocity", forthcoming in the *Journal of Economic Perspectives*.

[10] Güth, Königstein, and Nehring consider an indirect reciprocity treatment where the essential new feature is that receivers are informed about the donations of *both* donors. Otherwise, these authors repeat the indirect as well as the direct reciprocity game to check whether the data are comparable.

Güth, W. and H. Kliemt (1994): Competition or co-operation - On the evolutionary economics of trust, exploitation and moral attitudes, *Metroeconomica* 45, 155-187.

Güth, W., S. Huck, and P. Ockenfels (1996): Two-level ultimatum bargaining with incomplete information – An experimental study, *Economic Journal* 106, 593-604.

Güth, W., M. Königstein, and K. Nehring (1999): On the effects of being more or less generous than others – An experimental study of direct and indirect reciprocity when social comparisons are possible, *Working Paper*, Humboldt-University of Berlin.

Hoffman, E., K. McCabe, K. Shachat, and V. Smith (1994): Preferences, property rights, and anonymity in bargaining games, *Games and Economic Behavior* 7 (3), 346-380.

Stewart, H. (1992): "Rationality and the market for human blood", *Journal of Economic Behavior and Organization* 19, 125-43.

Titmuss, R.M. (1971), *The Gift Relationship: From Human Blood to Social Policy*, London: George Allen & Unwin.

Appendix A:

Instructions for Player A in the indirect treatment

(Translated from Hebrew)

Welcome to this experiment in decision-making. In this experiment, you may earn some money that will be paid to you, privately and in cash, at the end.

The interaction in the experiment will be in groups of four participants, where the students in each group are called A, A', B, and B'. You are called Student A. At the beginning of the experiment both you and student A' will receive 10 "points". Students B and B' will not receive points.

You are asked to decide whether you want to send any amount of these 10 points to student B; and if so, how much. We will triple the amount you send and give it to student B; that is, for every 1 point that you send, student B will receive 3 points. Student A' will be asked to decide how much to send to student B' in a similar way.

Then we will ask student B' (who received the money from student A') to decide if (s)he wants to send to you any amount of the points (s)he received from A' and if so, how much. This amount will not be tripled. Student B will be asked if (s)he wants to send money to student A' in the same fashion.

This will end the experiment, and the money will be paid to you (for every point you will have at the end we will pay you NIS 8).

Your ID number: _____

Number of points you want to send to student B: _____ (Please remember that this amount should be between 0 and 10 points.)

Instructions for Player B in the indirect treatment

(Translated from Hebrew)

Welcome to this experiment in decision-making. In this experiment, you may earn some money that will be paid to you, privately and in cash, at the end.

The interaction in the experiment will be in groups of four participants, where the students in each group are called A, A', B, and B'. You are called Student B. At the beginning of the experiment both student A and student A' will receive 10 "points". You and student B' will not receive points.

Student A will be asked to decide whether to send any amount of these 10 points to you; and if so, how much. We will triple this amount and give it to you; that is, for every 1 point that student A will send you will receive 3 points. Student A' will be asked to decide how much to send to student B' in a similar way.

Then we will ask you to decide if you want to send any amount of the points to student A' (not to A) and if so, how much. This amount will not be tripled. Student B' will be asked if (s)he wants to send money to student A in the same fashion.

This will end the experiment, and the money will be paid to you (for every point you will have at the end we will pay you NIS 8).

Your ID number: _____

The amount of points sent to you by student A (after we tripled it): _____

Number of points you want to send to student B: _____ (Please remember that this amount should be between 0 and the amount you received.)

Appendix B: Individual decision data

	Full information				**Incomplete information**			
	Direct reciprocity		**Indirect reciprocity**		**Direct reciprocity**			
	x	***y***	***x***	***y***	***x***	**All *y***	***y* (*2)**	***y* (*4)**
1	0	0	0	0	0	0		0
2	0	0	0	0	0	0		0
3	0	0	0	0	1	0		0
4	1	0	0	0	1	0	0	
5	3	0	0	0	1	1	1	
6	3	3	1	0	2	0	0	
7	5	0	2	1	2	1	1	
8	5	5	2	2	4	4	4	
9	5	5	4	3	5	0	0	
10	5	7	5	0	5	2	2	
11	5	7	5	3	5	3		3
12	5	8	5	5	5	5		5
13	5	6	5	5	5	5		5
14	6	9	5	5	5	5	5	
15	6	10	5	10	5	7		7
16	6	3	5	15	5	7	7	
17	7	7	7	5	5	10		10
18	7	7	7	7	5	10		10
19	7	9	7	10	5	10	10	
20	9	7	8	7	6	0	0	
21	9	10	8	8	6	3		3
22	10	0	8	12	7	7	7	
23	10	0	9	14	7	7	7	

24	10	5	10	5	7	14		14
25	10	10	10	10	8	4	4	
26	10	10	10	10	8	5		5
27	10	15	10	15	10	0		0
28	10	15	10	30	10	5		5
29					10	9	9	
30					10	10		10
31					10	10		10
32					10	10	10	
Average	6.04	5.64	5.29	6.5	5.47	4.81	5.44	5.44
Std.dev.	3.28	4.55	3.54	6.71	3.07	4.07	3.75	4.40

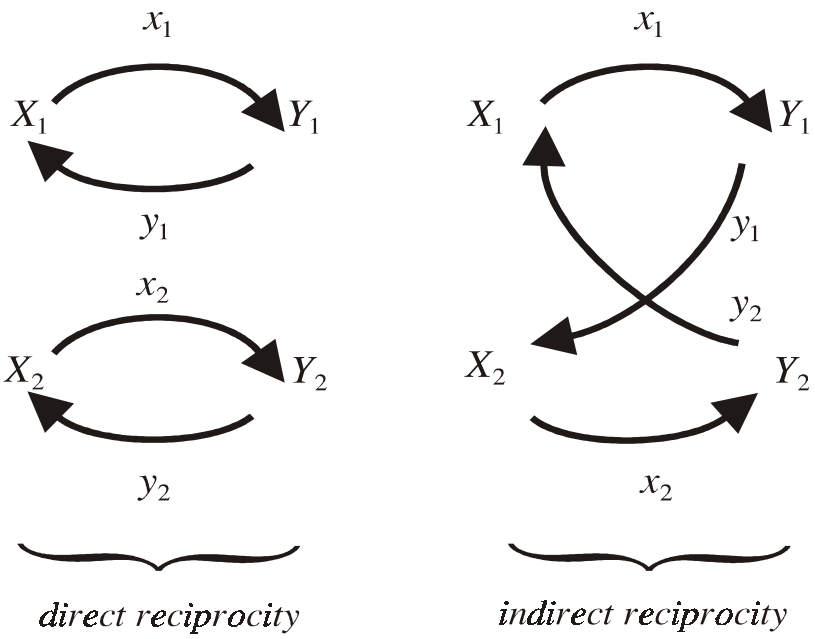

Figure II.1. A graphical illustration of the exchanges with direct and indirect reciprocity

Multiplier information	Reciprocity	
	Direct	**Indirect**
Complete (multiplier commonly known)		
Incomplete (multiplier known only to donors)		

Figure II.2. The three treatments

Multiplier information	Reciprocity	
	Direct	**Indirect**
Complete	176.48	164.64
	≈30	≈30
	28	28
Incomplete	307.24	Average earnings
	≈30	Time needed
	32	Number of observations

Figure II.3. Average earnings (NIS), time needed (in minutes) and number of participants in all three treatments.

Multiplier information	Reciprocity							
	Direct				Indirect			
	x		*y*		*x*		*y*	
Complete	6.04	6	5.64	7	5.29	5	6.5	5
	3.28	28	4.55	28	3.54	28	6.7	28
Incomplete	5.47	5	4.81	5				
	3.07	32	4.07	32				

Table III.1. Averages, medians (upper line) and standard deviations, numbers of observations (lower line) for all three treatments, depicted as in Figure II.2

		Multiplier							
		μ = **2**				μ = **4**			
		x		*y*		*x*		*y*	
Mean	**Median**	5.2	5	4.2	4	5.9	5	5.4	5
Std.dev.	**# of obs**	2.8	16	3.3	16	3.6	16	4.3	16

Table III.2. The results of the incomplete information treatment separately for pairs with small (2) and large (4) multiplier