Carrots without Bite: On the Ineffectiveness of 'Rewards' in sustaining Cooperation in Social Dilemmas

Stoop, Jan and van Soest, Daan and Vyrastekova, Jana

Erasmus University Rotterdam, Department of Applied Economics, the Netherlands., VU University Amsterdam and Tilburg University, Radboud University Nijmegen, Department of Economics, Nijmegen, the Netherlands

11 March 2011

Carrots without Bite: On the Ineffectiveness of 'Rewards' in sustaining Cooperation in Social Dilemmas*

Jan Stoop†, Daan van Soest‡, and Jana Vyrastekova§

March 11, 2011

Abstract

Peer-to-peer sanctions increase cooperation in multi-person social dilemmas (Fehr & Gächter (2000)), but not when subjects have the option to retaliate (Nikiforakis (2008)). One-shot peer-to-peer rewards have been found to enhance efficiency too (Vyrastekova & van Soest (2008), Rand et al. (2009a)), but it is an open question whether the positive impact on cooperation is weakened or strengthened when we allow for counterrewarding. We examine the impact of possible reciprocity in rewarding on cooperation in a non-linear public bad game, and find that efficiency in the social dilemma is equally low as absent any reward options. We hypothesize that subjects are unwilling to sever mutually profitable bilateral exchanges of reward tokens to induce cooperation in the social dilemma, and identify the underlying mechanism by comparing behavior across three matching protocols.

Key words: Social dilemmas, economic experiments, rewards.
JEL Classification: C72, C92, D74.

*We thank the Netherlands Organisation for Scientific Research, NWO, for financial support as part of the program on Evolution and Behavior. We are also grateful to Chris Müris and David Voňka for their comments on an earlier version of this paper.

†Erasmus University Rotterdam, Department of Applied Economics, the Netherlands.

‡VU University Amsterdam and Tilburg University; corresponding author. Please send all comments to D.P. van Soest, Tilburg University, Department of Economics, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: d.p.vansoest@uvt.nl.

§Radboud University Nijmegen, Department of Economics, Nijmegen, the Netherlands.

1 Introduction

Over the past two decades, many economic experiments have been conducted to assess the relative effectiveness of self-regulatory instruments in sustaining cooperation in multi-person social dilemma situations, such as linear public good games and non-linear public bad games. Instruments tested in laboratory experiments include ostracism (Masclet (2003), Maier-Rigaud et al. (2010)), peer-to-peer rewards (Sefton et al. (2007), Vyrastekova & van Soest (2008), Rand et al. (2009a)), and verbal expressions of approval or disapproval (Masclet et al. (2003)). Most attention, however, has been paid to the effectiveness of peer-to-peer punishments; see for example Yamagishi (1988), Ostrom et al. (1992), and Fehr and Gächter (2000, 2002). Offering subjects the opportunity to impose (monetary) sanctions on their peers significantly increases the efficiency of public good provision, and this is even the case if punishments are not only costly to the punished, but also to the subject imposing them (Gächter et al. (2008)). To economists, these results are surprising because the experimental games are set up such that subjects should not be willing to provide the second-order public good of punishing free-riders in any of the periods, and hence efficiency in the social dilemma should be equally low with and without the opportunity to impose punishments.

The external validity (or real world relevance) of the experimental punishment mechanism results has been challenged on two grounds. The first is that 'sticks' may not be used so eagerly if there is an opportunity for revenge. Nikiforakis (2008) conducted a public good game experiment with two punishment stages rather than just one, so that subjects can use the second punishment stage to directly reciprocate to any sanctions received in the first. The consequences are quite dramatic. Faced with the threat of potential retaliation hardly any sanctions are imposed in the first punishment stage, and hence the efficiency in the multi-person social dilemma stage does not differ from the efficiency level that materializes absent any punishment stages.[1] Hence, peer-to-peer punishments may be able to sustain coopera-

[1]See Denant-Boemont et al. (2007) and Nikiforakis & Engelmann (2011) for additional

tion in the real world, but only if punishers can hide their identity to those being punished (see also Rand et al. (2009a)).

The second criticism regarding the real-world relevance of peer-to-peer punishments is that in most societies, the use of force is the exclusive right of the government: typically, individual citizens are allowed to neither impose physical nor monetary punishments on their peers (Vyrastekova & van Soest (2008)). That means that peer-to-peer rewards may be empirically more relevant than peer-to-peer punishments, and a relatively small literature has emerged analyzing the effectiveness of rewards in sustaining cooperation (Sefton et al. (2007), Vyrastekova & van Soest (2008), Rand et al. (2009a), and Sutter et al. (2010)). When using the same design features as the standard punishment experiment, rewards are observed to increase cooperation in the social dilemma stage if and only if the benefits of receiving a reward are larger than the costs of giving it — but less so than in case of punishments.

This paper contributes to the literature on the effectiveness of the reward mechanism in sustaining cooperation in multi-person social dilemmas by exploring to what extent offering subjects the opportunity to counter-reward increases or decreases the mechanism's effectiveness. While individual agents generally have strong incentives to hide their identity in case they punish another agent in a social dilemma, the opposite holds in case of rewards; the benefactor usually has good reasons to reveal her identity to the recipient. Also, in most real-world social dilemmas agents are likely to be well aware of the history of (at least a subset of) their fellow agents' behavior in the social dilemma as well as of the history of whom they received 'rewards' from (in the form of gifts, but possibly also in the form of help minding one's children, help with crop harvesting, etc.); see also Rand et al. (2009a). Do rewards improve efficiency in the social dilemma in such a setting, even when subjects can reciprocate to rewards received before?

To answer this question, we analyze the behavior of subjects in a finitely repeated game. In every period, subjects first decide on their investments

analyses of the underlying mechanism. For a cross-cultural analysis of the factors inducing subjects to engage in retaliation (or anti-social punishment), see Hermann et al. (2008).

in a standard non-linear public bad game, after which they can decide how many reward tokens (of a limited budget) they send to each of the four other members of their group. The costs of sending a reward token are smaller than the benefits of receiving one, so bilaterally exchanging reward tokens is a profitable enterprize by itself. As is the case in many real world instances, the design allows each subject to base her reward decisions not only on her fellow group members' behavior in the multi-person social dilemma in the current period, but also on the number of reward tokens she received from them in previous reward stages. That means that we use the so-called Partner Fixed (PF) matching protocol, where Partner refers to the fact that group composition remains unchanged throughout the experiment, while Fixed refers to the fact that each subject receives a unique identity label that is constant throughout the experiment too.

We hypothesize that rewards may not be able to sustain cooperation in the social dilemma. If subjects can condition their decision to send reward tokens not just on (the history of) their peers' behavior in the social dilemma but also on (the history of) rewards received, which of the two — if any — will they reciprocate to? Or, stated otherwise, are subjects willing to potentially jeopardize a mutually profitable (bilateral) exchange of reward tokens by withholding rewards if another agent decides to act less cooperatively in the social dilemma? If subjects do not view the decrease in the number of reward tokens received as a just punishment for their acting less cooperatively in the social dilemma stage, they may retaliate by withholding rewards too. In that sense, withholding rewards can be viewed as a second-order public good, and the question is whether or not subjects are willing to provide it.

We test this hypothesis using two different treatments. One is a treatment in which every social dilemma stage (the non-linear public bad game) is followed by a single reward stage, the 1SR-PF treatment (where 1SR refers to the fact that there is just one reward stage in every period, and where PF indicates that we use the Partner Fixed matching protocol). The second treatment is a game in which a period consists of a social dilemma stage followed by two reward stages (rather than just one). This game captures the idea that in the real world the frequency with which rewards can be exchanged may well be higher than the frequency in which agents make decisions regarding their behavior in social dilemmas (or the frequency with which they receive information on the behavior of their peers in these social dilemma situations). We refer to this experimental game as the 2SR-PF treatment, as the matching protocol remains Partner Fixed.

To date many social dilemma experiments have been run, and our setup is closest to the designs implemented by Vyrastekova & van Soest (2008) and Rand et al. (2009a). As is the case in our experiment, rewards are 'efficiency enhancing' in these two studies in that the payoffs of the recipient of a reward increase by three points while the costs of giving it are just one point. Our study differs from these two because we offer subjects two opportunities for rewarding in every period rather than just one so that they can 'counterreward'. In addition, our study improves on that by Vyrastekova & van Soest (2008) by allowing subjects to condition their rewarding decisions on the complete history of play (by implementing the Partner Fixed matching protocol, as is also done by Rand et al. (2009a)) rather than just on social dilemma play in the current period. And while Rand et al. (2009a) constrain their subjects' choice space to the decision, vis-a-vis each fellow subject, to give him a reward, yes or no, subjects have complete freedom in rewarding in our design. They can choose to give no rewards, to spread them equally, or to give them all to one fellow group member — or anything in between. That means that subjects do not need to solicit the cooperation of *all* other group members to obtain the maximum benefits from bilateral cooperation — selecting only a few partners (or maybe just one) to exchange reward tokens with may suffice.

When designing the experiment, we expected the results of the analysis to be sensitive to whether we would allow for two or just one reward stage per period. We hypothesized that rewards may be able to sustain cooperation in the 1SR-PF treatment but not in the 2SR-PF treatment because the second reward stage might shift our subjects' attention away from their peers' behavior in the social dilemma and towards their behavior in the rewarding stages. These predictions did not play out in practice, however, as behavior in the social dilemma stage was very similar in the two treatments — rewards are found to be unable to increase efficiency in the social dilemma above that achieved absent any reward options, independent of whether there is one reward stage, or two.

While both Vyrastekova & van Soest (2008) and Rand et al. (2009a) find that rewards can sustain cooperation in social dilemma situations, we thus come to the exact opposite conclusion. Even though in our experiment the average number of reward tokens sent by each subject is high and even increasing as the game progresses (as is the case in Rand et al. (2009a)), efficiency in the non-linear public bad game is low — even lower than predicted by standard game theory. Indeed, we find that subjects establish relationships with one another in which each partner systematically sends reward tokens to the other. These mutually profitable partnerships are formed early on in the experiment and are long-lasting. We also find that the establishment of these connections is largely independent of the partners' behavior in the social dilemma in the early periods of the experiment. Hence, subjects reciprocate to rewards received — not to their peers' behavior in the social dilemma.

This paper is organized as follows. In section 2 we present the two experimental games that make up the main treatments of this experiment, as well as the three matching protocols implemented. In section 3 we present the data for the 2SR-PF sessions as this matching protocol is empirically the most relevant one. In section 4 we present the results of the other two matching protocols, as they provide additional support for our claim that rewards are not likely to be effective in sustaining cooperation in real-world social dilemmas when agents can reap the full benefits of the exchange of rewards by forming long-lasting partnerships with just a limited number of fellow community members. We explain why our conclusions are opposite to those obtained by Vyrastekova & van Soest (2008) and Rand et al. (2009a) in the concluding section 5.

2 The game and experimental procedure

In this section we present the experimental design. Section 2.1 presents the model, and section 2.2 describes the experimental procedure.

2.1 The experimental game

In line with the game developed by Ostrom et al. (1992), we implement a non-linear public bad game with $N > 1$ identical players. The game is repeated $T \geq 1$ times, and in every period $t = 1, \ldots, T$ each player $i \in \{1, 2, \ldots, N\}$ can allocate a fixed amount of 'effort', e, between a social dilemma activity and an alternative economic activity, the outside option. We use $x_{i,t}$ to denote the amount of effort player i puts into the social dilemma activity in period t, where $x_{i,t}$ is an integer number between 0 and e. The marginal return on the amount of effort allocated to the outside option, $e - x_{i,t} \geq 0$, is constant and equal to w. The private marginal benefits of effort allocated to the social dilemma activity are equal to $A - BX_t$, where $X_t \equiv \sum_{i=1}^{N} x_{i,t}$. The baseline game consists of one stage only, the social dilemma stage, which we will refer to with superscript $s1$. Player i's total payoffs in stage $s1$ of period t are thus equal to:

$$\pi_{i,t}^{s1} = w(e - x_{i,t}) + [A - BX_t]\, x_{i,t}. \tag{1}$$

Because $\partial \pi_{i,t}^{s1} / \partial x_{j,t} < 0$ for all $j \neq i$, this game is a (non-linear) public bad game. If $T = 1$, the symmetric individual Nash effort level is $x^{NE} = (A - w)/B(N+1)$, while the socially optimal individual effort level is equal to $x^{SO} = (A - w)/2BN$. Since $x^{NE} > x^{SO}$ if $N > 1$, there is a social dilemma. If the game is repeated a finite number of times ($T \geq 2$), the standard game-theoretic prediction is that all players choose the Nash equilibrium effort x^{NE} in all periods $1, \ldots, T$. Using backward induction, if it does not pay to cooperate in the last period of a finitely repeated game, it does not pay to cooperate in any previous period either.

The game described above captures a social dilemma in which there are no instruments to affect the behavior of one's peers other than one's own social dilemma effort level. Hence it serves as a baseline against which we can test the impact of players having the opportunity to reward their peers. We refer to this baseline game as 0SR, reflecting that there are zero reward stages in this game.

The game that allows for rewarding is modeled as follows. The first stage ($s1$) in this game is identical to the (first) stage of the baseline game (0SR), and hence a player's payoffs in this stage are given by equation (1). The social dilemma stage is then followed by either one reward stage, $s2$ or by two (identical) reward stages, $s2$ and $s3$. We will refer to these games as 1SR and 2SR, respectively, reflecting that these games have either one reward stage, or two. A reward stage is set up as follows. Each of the N players receive z reward tokens which she can keep herself, or give to one or more of her fellow group members. Every token that the player keeps, increases her payoffs by 1 point. Every token that is sent to a fellow group member, increases that group member's payoffs by r points, where $r > 1$. Note that while this assumption seems restrictive, it likely to be met in many different situations (see Vyrastekova & van Soest (2008), and Rand et al. (2009a, 2009b)). 'Rewards' can be thought of as gifts (financial, in kind, or time) that increase the recipient's welfare. People's marginal valuation of objects may well differ, and their marginal valuation of money can differ too. And time constraints may also result in people valuing time differently; if community members undertake, say, agriculture in addition to being active in fishing at a lake (the social dilemma activity), rewards can take the form of assisting a fellow community member getting his harvest of the land in time. If not all crops are ready for harvest at the same time, time constraints differ between community members, and so do their marginal values of time. Hence, the recipient's valuation of the 'reward' may well be higher than the provisioning cost incurred by the benefactor.

So, we assume that $r > 1$, and player i's payoffs in stage s ($s = \{s2\}$ in 1SR, $s = \{s2, s3\}$ in 2SR) in period t are given by:

$$\pi_{i,t}^{s} = z - \sum_{j \neq i} p_{ij,t}^{s} + r \sum_{j \neq i} p_{ji,t}^{s}, \tag{2}$$

where $p^s_{ij,t}$ is the number of reward tokens that player i sends to player j ($j \neq i$) in stage s in period t. Hence, the total individual payoffs in period t of the 1SR and 2SR game are $\pi^{1SR}_{i,t} = \pi^{s1}_{i,t} + \pi^{s2}_{i,t}$ and $\pi^{2SR}_{i,t} = \pi^{s1}_{i,t} + \pi^{s2}_{i,t} + \pi^{s3}_{i,t}$, respectively.

Aggregate payoffs are maximized if all players (i) choose effort level $x^{SO} = (A - w)/2BN$ in every period, and (ii) always send all their z reward tokens in both reward stages to their fellow group members, because $r > 1$. The standard game-theoretic predictions are, however, that no reward tokens are sent in either $s2$ or $s3$ in any period of 2SR (i.e., $p^{s2}_{ij,t} = p^{s3}_{ij,t} = 0$ for all $j \neq i$, and for all $t = \{1, \ldots, T\}$). Applying backward induction there is no reason for a selfish player to send reward tokens in $s3$ of period T, and hence there is no reason to send reward tokens in stage $s2$ of that period either. If all players are selfish, there is also no reason to choose any effort level other than the Nash equilibrium one, x^{NE}, in period T, and hence there are no reasons to send reward tokens in either of the two reward stages in period $T - 1$ either. That means that the game unravels, and efficiency in the social dilemma activity ($s1$) is equal to the non-cooperative level independent of whether or not players have the opportunity to send reward tokens. And the same reasoning holds for 1SR, giving rise to the same game-theoretic prediction that the Nash equilibrium obtains in every stage and in all periods.

According to social orientation tests, only about 30 percent of humans behave consistently with the assumption of 'homo economicus' in laboratory experiments; see for example Fischbacher et al. (2001). Altruists may be willing to always give rewards, because it gives rise to warm glow and/or because it increases group welfare; conditional cooperators may use the reward stages 'properly' by giving rewards to those players who act cooperatively in the social dilemma stage. Thus, if players are endowed with a richer set of preferences than homo economicus, the above standard game-theoretic predictions may be refuted. It may also be the case that players are predominantly interested in their own material welfare, but that the above predictions do not play out because players realize that others may be willing to reciprocate to rewards received (see **?**). Hence, they may decide to establish bilateral ties of cooperation by exchanging reward tokens rather than to use the reward tokens to sustain cooperation in the social dilemma.

In real world social dilemmas agents typically have good knowledge of the (past and present) behavior of (at least a subset of) their fellow community members in the social dilemma activity, and also whether and from whom they received 'rewards' (in the form of gifts, or help) in the present and past. That means that from the range of matching protocols typically used in economic experiments, the Partner Fixed protocol is the most plausible one. In this matching protocol, group membership does not change throughout the experimental session, and also identity labels remain fixed not only within but also between periods. In this setup, all of the above reasons to send reward tokens may materialize, and we can assess the net result of their interaction by comparing the efficiency in the social dilemma stage in the 1SR and 2SR treatments to that in the 0SR treatment. To have an adequate benchmark, participants play either the 0SR and 1SR treatments or the 0SR and 2SR treatments sequentially in every Partner Fixed (PF) session, with 0SR being played first.

However, we can gain additional insight into the relevance of the various uses of reward tokens by having players play the game using two alternative matching protocols. In one, group composition remains constant throughout the experiment but identity labels are randomly changed between periods (Partner Random, PR). In sessions with this PR matching protocol, players cannot base their reward decisions on whether or not they received rewards from a specific fellow group member in the past, but they can use their reward tokens to encourage fellow group members to continue acting cooperatively in the social dilemma. Hence, if the latter type of use of rewards is the dominant one, play in the PR and PF sessions should be identical. In the second alternative matching protocol, new groups are formed randomly in every period (the so-called Stranger (S) matching protocol). Here, tokens may be given as a reward for acting cooperatively in the social dilemma stage, but such rewarding behavior can not be motivated by subjects expecting to benefit themselves from their peers' cooperative behavior in future periods. So, whereas the 0SR, 1SR and 2SR treatments with PF matching are the most important ones, we also implement the two treatments using PR and S matching protocols as they allow us to identify the dominant motivation behind the use of reward tokens — if they are used at all.

2.2 Experimental design

The experiments were conducted at Tilburg University's CentER laboratory in the Fall of 2008 and in the Spring of 2009. Subjects were students with different nationalities and with backgrounds in business, economics, law, or social sciences. Each subject participated in only one session. The experimental parametrization of the game is given in Table 1, and Table 2 presents the associated socially optimal and Nash equilibrium levels. Sessions lasted roughly two hours, and average earnings were €15.96 including a €5 show-up fee. All decisions were mediated via z-Tree (Fischbacher (2007)).

Variable	Description	Value
N	number of individuals per group	5
T	number of periods of the stage game	15
w	return on investments in the private activity	0.5
A	parameter of the social dilemma's revenue function	11.5
B	parameter of the social dilemma's revenue function	0.15
e	individual endowment of effort	13
z	individual endowment of 'reward' tokens	12
r	value of reward tokens received	3

Table 1 Experiment parameterization.

In each session, subjects played the 0SR treatment as well as either the 1SR or the 2SR treatment, and within a session all games were implemented using the same matching protocol (Partner Fixed, Partner Random, or Stranger). In the instructions participants were informed about the matching process in their session, and games 0SR and 1SR/2SR were referred to as Task 1 and Task 2, respectively. Participants were informed that they would participate in two tasks, but they received the instructions

Variable	Description	Value
x^*	symmetric individual socially optimal effort level	6
X^*	aggregate socially optimal effort level	30
x^{NE}	individual Nash equilibrium effort level	10
X^{NE}	aggregate Nash equilibrium effort level	50
$p_{ij}^{SO,s}$	indiv. socially optimal no. of reward tokens sent in every stage	12
$p_{ij}^{NE,s}$	indiv. Nash equilibrium no. of reward tokens sent in every stage	0

Table 2 Social optimum and Nash equilibrium values of all decision variables for the given experiment parametrization.

for Task 2 only after Task 1 was finished.[2] The tasks were framed neutrally. The effort decision was described as 'investing tokens in option 1 or 2', where the first represented the social dilemma activity and the second the outside option (with constant marginal benefits w). In Task 1, subjects played 15 periods of 0SR. Subjects were shown equation (1), but they were also given a payoff table in which they could look up, for every aggregate amount of effort put in by the other group members, what payoffs they would earn for a specific amount of effort invested. We did not inform the subjects about the socially optimal or the Nash equilibrium effort levels. Before the start of the experiment subjects were presented with a short test; the participants answered all questions correctly without much difficulty.

After Task 1, the same sequence of events took place for Task 2, consisting of 15 periods of either 1SR or 2SR. Participants were informed that after having made the same decision as in Task 1 (allocating tokens to options 1 and 2), either one or two more decisions were to be made in every period — depending on whether they were to play 1SR or 2SR. The decision(s) consisted of allocating a budget of tokens between other group members and themselves, where any token kept increased one's payoffs by 1 point, and any token given increased the recipient's payoffs by 3 points. Hence, the 'rewarding' decision problems were framed neutrally too.

The information structure in every period of Task 1 (0SR) and Task 2 (1SR or 2SR) was as follows. At the end of stage 1 of Task 1, subjects were informed about the individual effort decisions of all other group members,

[2]The instructions are available upon request from the authors.

and about their associated profits. In Task 2, subjects received the same information as in Task 1, but they were also informed, at the end of every reward stage, about the number of reward tokens they had received from other subjects as well as about the associated payoff consequences.

As explained above, 0SR followed by 2SR was implemented using three different matching protocols (PF, PR and S) while 0SR followed by 1SR was run just using the PF matching protocol. The four session types are summarized in Table 3. Comparing 2SR across the three different matching protocols allows us to better understand the mechanism giving rise to rewarding behavior, and hence we will focus our attention on the 2SR results.

Session	Subjects	Groups	Average Earnings
Partner, 1SR, fixed identity labels	50	11	€12.60
Partner, 2SR, fixed identity labels	50	10	€19.60
Partner, 2SR, random identity labels	55	11	€18.11
Stranger, 2SR	80	4 sessions	€14.30

Table 3 Summary information on the number of participants and amounts of money earned in the three session types.

3 Analysis of play in the PF sessions

The key question to be addressed in this question is whether the option to send rewards increases efficiency in the multi-person social dilemma (stage $s1$) – if subjects have the option to reciprocate not just to their peers' behavior in $s1$, but also to rewards received in previous periods. If the option to reward is introduced when using the Partner Fixed matching protocol, is the resulting level of efficiency higher than absent any reward possibilities? In this section we compare efficiency in 1SR and 2SR to that materializing in 0SR.

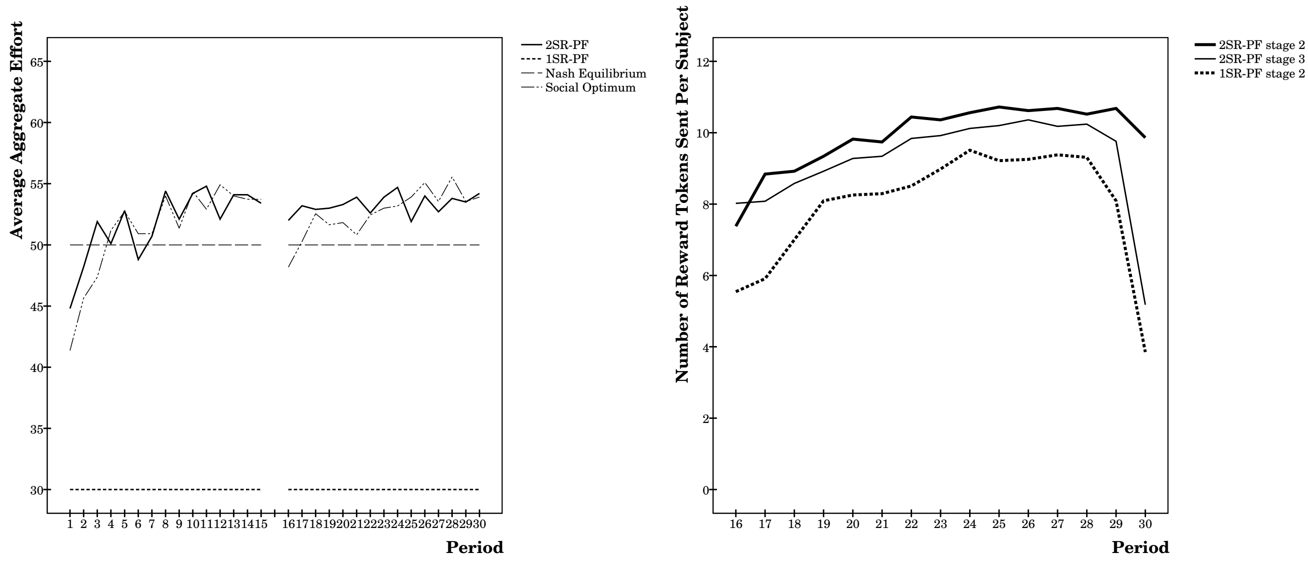
In Figure 1(a) we present the aggregate effort (averaged over all groups) in the social dilemma stage in periods 1-15 of the PF sessions, as well as that in periods 16-30. Comparing the aggregate effort in 1SR-PF (averaged over all 15 periods) to that in 0SR-PF, the relevant Wilcoxon test (with

$N_1 = N_2 = 11$) yields a p-value of 0.28. Similarly, even when omitting the first three periods of 0SR to account for learning, aggregate effort in 2SR (averaged over all 15 periods) is not below that in 0SR ($p = 0.58$ according to a Wilcoxon test with $N_1 = N_2 = 10$).[3] Hence, these results indicate that the opportunity to send reward tokens (whether it is just one opportunity per period or two) does not affect efficiency in the social dilemma activity. The only important difference between 1SR and 2SR is that in the latter treatment average aggregate effort is essentially constant over periods 16-30 while in the 1SR treatment it takes longer for average effort to stabilize. Indeed, in case of 2SR the difference in average aggregate effort levels between periods 15 and 16 is not significant ($N_1 = N_2 = 10$, $p = 0.72$) while there is a significant increase in cooperation when comparing effort in those two periods in case of 1SR ($N_1 = N_2 = 11, p = 0.056$). However, even in 1SR convergence is pretty quick because effort in periods 15 and 18 are already statistically indistinguishable, and even though the fall in effort between periods 15 and 16 is statistically significant, average effort in that period is very close to the Nash equilibrium prediction (as it is equal to 9.6). Hence, behavior in periods 16-30 is quite similar in 1SR and in 2SR, and efficiency in either treatment is nog significantly different from that in 0SR.[4]

Thus, we find no evidence that adding one or two reward stages to a standard public bad game increases efficiency in the social dilemma. This is not due to subjects' refusing to use the reward options in either 1SR or 2SR, as standard game theory would predict. On the contrary, Figure 1(b) shows that in period 16, on average, subjects give away between half (in $s2$ of 1SR) and two-thirds (in $s2$ of 2SR) of their endowment of reward tokens in the relevant stage, and also that the number of reward tokens sent increases over time – in case of $s2$ decisions in 2SR the number of reward tokens sent

[3]When including the first three periods of 0SR, average effort in the 2SR treatment only just fails to be significantly *higher* than that in 0SR ($p = 0.11$).

[4]Indeed, the results of the 1SR-PF treatment are statistically indistinguishable from those in 2SR-PF. For the null hypothesis of play in 1SR-PF and 2SR-PF being identical the associated p-value for the average aggregate effort levels in periods 1-15 is equal to 0.756 (according to the relevant Mann-Whitney U test, with $N_1 = 11, N_2 = 10$), and for the average aggregate effort level in periods 16-30 this test yields a p-value of 0.863.



(a) **(b)**

Figure 1 **(a)** Average aggregate effort in the social dilemma stage in the PF sessions. **(b)** Average number of reward tokens sent per subject in stage 2 and stage 3 in the PF sessions.

even approaches the maximum of 12 tokens as the game proceeds. And the average number of reward tokens sent in $s3$ of every period of 2SR is only just below that sent in $s2$ — except for the very last period. This gives rise to the following three results.

Result 1 Behavior in the social dilemma stage of 1SR-PF and 2SR-PF is, on average, even less cooperative than predicted by standard game theory, but the average subject gives away more than half (two-thirds) of her endowment of reward tokens in the first reward stage of 1SR (2SR) in all periods.

Result 2 While efficiency in the social dilemma stage remains low in both 1SR-PF and 2SR-PF, the average number of tokens sent increases over time in all reward stages.

Result 3 In 2SR-PF, the number of reward tokens sent in $s3$ is smaller than that sent in $s2$, but not substantially so (except for the very last period).

Results 1 and 2 suggest that it is unlikely that the decision to send reward tokens is motivated by a desire to compensate one's peers for their cooperative behavior in the social dilemma. Instead, the temporal increase in rewards exchanged (result 2) and the fact that – in 2SR-PF – almost an equal number of reward tokens are sent in the second reward stage as in the first (result 3) suggest that subjects (i) recognize that exchanging reward tokens is profitable, and (ii) base their decision to send reward tokens more on the history of reward tokens received than on the development of cooperation in the social dilemma stage.[5]

However, the above results are obtained on the basis of aggregate data, and these may hide important differences at the individual level. For example, it may be the case that individuals frequently change their decisions to send reward tokens in response to changes in effort levels chosen by their peers, with those decreasing (increasing) their effort levels facing an increase (decrease) in the number of reward tokens received. Below, we present the analyses for 2SR-PF.[6]

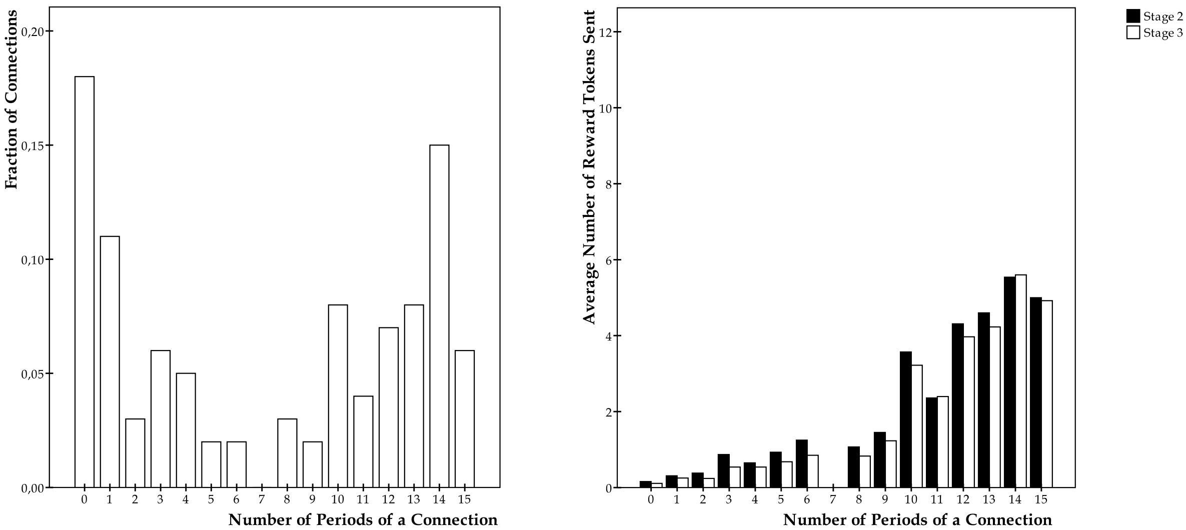
To test the hypothesis that subjects are unwilling to provide the second-order public good of severing mutually profitable bilateral exchange relationships with free riders, we first analyze the persistence in the number of rewards exchanged between subjects in 2SR-PF. We introduce the following definition:

[5]The negligible difference in the number of reward tokens sent in $s2$ and $s3$ provides additional evidence that play in 2SR-PF and 1SR-PF are very similar. When comparing the number of reward tokens sent in the last reward stage of either treatment (that is, $s2$ in 1SR-PF and $s3$ in 2SR-PF), the p-value of the relevant Mann-Whitney U test equals 0.152 (with $N_1 = 11, N_2 = 10$).

[6]As suggested by Figure 1, behavior in 1SR-PF and 2SR-PF treatments are very similar, and probing the data all the evidence suggests that the underlying mechanisms are the same too. Because adjustment occurs faster in 2SR-PF than in 1SR-PF while the second reward stage also allows us to better identify the underlying mechanism when running alternative matching protocols (2SR-PR and 2SR-S; see section 4), we prefer clarity to completeness, and focus our discussion of the results on just 2SR-PF. The results for 1SR-PF are, however, available upon request.

Definition Subjects i and j ($j \neq i$) are said to have a connection of length τ in period t, measured by $Connection_{ij,t} = \tau$, if τ is the number of periods between periods 16 and t in which i sent a strictly positive number of reward tokens to j in both $s2$ and $s3$, and vice versa.

Figure 2(a) shows the frequency with which connections with a certain duration occur in the data, evaluated in period 30. Although there are quite a few short-run connections, the persistence in rewarding and counterrewarding is remarkable. Almost fifty percent of all connections have a length of 10-15 periods — all but one subject were involved in at least one such a relationship. Consistent with intuition, Figure 2(b) indicates that the number of tokens sent is larger the longer the connection is in place.[7]



(a) **(b)**

Figure 2 **(a)** Fraction of connections that last τ periods. **(b)** Average number of reward tokens sent between two subjects in a connection which lasts τ periods.

We thus find that connections are long-lasting — even though efficiency

[7]The Spearman correlation coefficient between the length of the connection and the average number of tokens sent is 0.90 in the first reward stage ($N = 100, p < 0.01$), and 0.91 for the second reward stage ($N = 100, p < 0.01$).

in the social dilemma stage ($s1$) is poor. The persistence in 'rewarding' raises the question how connections are formed. What is the role of the behavior in the social dilemma stage in every period? Is it really true that the number of reward tokens received is independent of a subject's effort decisions? To analyze this, we use regression analysis to explain the number of rewards sent in the two reward stages.

Let us first analyze the decisions of subject i to send reward tokens to subject j ($j \neq i$) in the two reward stages of the first period ($t = 16$). The key explanatory variables here are whether or not subject j acted cooperatively in the social dilemma stage of the first period. A natural benchmark is the average effort level of the $N - 1$ other subjects in a group; $\overline{x}_{-j,t} \equiv \sum_{i \neq j}(x_{i,t}/(N-1))$. Let us define cooperation (non-cooperation) as subjects choosing an effort level that is below (above) their group's average as measured by $\text{Max}\{0, \overline{x}_{-j,t} - x_{j,t}\}$ ($\text{Max}\{0, x_{j,t} - \overline{x}_{-j,t}\}$). These variables are included in the analysis of both $s2$ and $s3$. In addition, we also include $p^{s2}_{ji,t=16}$ as an explanatory variable in $s3$ of period 16.

The results are reported in the first two columns of Table 4. Column (i) shows that subjects are quite prone to sending reward tokens in $s2$ — as evidenced by the magnitude of the intercept — but slightly less so to subjects who put in more effort in the social dilemma stage than the average other group member.[8] Column (ii) shows that $p^{s3}_{ij,t}$ is not directly affected by subject j's (relative) effort decision, but that it is larger the more reward tokens subject j sent to subject i in $s2$ ($p^{s2}_{ji,t}$).[9]

Next, we analyze behavior in the second period of 2SR ($t = 17$); see columns (iii) and (iv) in Table 4. We use the same controls to explain $p^{s2}_{ij,t}$ as in columns (i) and (ii), but we also add the lagged number of rewards received as explanatory variables (that is, $p^{s3}_{ji,t-1}$ in column (iii), and $p^{s2}_{ji,t-1}$ in column (iv)). The results are striking. The decision to send reward tokens

[8]The intercept is about 2. Because $N - 1 = 4$, subjects send, on average, 8 of their 12 reward tokens in $s2$, and only take off 0.24 reward tokens for every unit of effort other subjects put in above the group's average.

[9]Note that $x_{i,t}$ may affect $p^{s3}_{ij,t}$ via $p^{s2}_{ji,t}$. However, this indirect effect is likely to be small because of the relatively large intercept and the very low R^2 of the $s2$ regression presented in column (i).

Dependent variable: Reward tokens sent by subject i to subject j in stage 2 and stage 3 in the PF sessions								
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
	Period 16		Period 17		Period 18		Period 25	
	$s2$	$s3$	$s2$	$s3$	$s2$	$s3$	$s2$	$s3$
$\text{Max}\{0, \overline{x}_{-j,t} - x_{j,t}\}$	0.049	0.023	0.258	0.196	0.249	0.196	0.068	0.006
	(0.070)	(0.052)	(0.144)	(0.191)	(0.145)	(0.148)	(0.075)	(0.026)
$\text{Max}\{0, x_{j,t} - \overline{x}_{-j,t}\}$	-0.242^{**}	-0.043	-0.241	-0.062	-0.040	0.042	-0.032	-0.025
	(0.103)	(0.107)	(0.168)	(0.186)	(0.103)	(0.139)	(0.054)	(0.041)
$p^{s2}_{ji,t}$		0.461^{***}		0.572^{***}		0.671^{***}		0.602^{***}
		(0.107)		(0.096)		(0.138)		(0.173)
$p^{s3}_{ji,t-1}$			0.541^{***}	0.107	0.785^{***}	0.209	0.944^{***}	0.365^{*}
			(0.059)	(0.091)	(0.081)	(0.122)	(0.036)	(0.179)
Constant	2.096^{***}	1.180^{***}	1.107^{***}	0.400	0.444^{**}	-0.001	0.248	0.035
	(0.222)	(0.204)	(0.305)	(0.294)	(0.165)	(0.169)	(0.151)	(0.118)
N	200	200	200	200	200	200	200	200
R^2	0.0241	0.1950	0.3342	0.4837	0.5999	0.7100	0.8371	0.8980

Table 4 OLS regression estimates of the number of reward tokens sent in the first reward stage ($s2$) and in the second reward stage ($s3$) in the first three periods of 2SR, and in its tenth period (i.e., $t = 16 - 18$, and 25). Standard errors, clustered at the group level, are reported between parentheses. ***: significant at the 1%-level, **: significant at the 5%-level, *: significant at the 10%-level.

in both $s2$ and $s3$ is independent of the recipient's behavior in $s1$, while the coefficients on the number of rewards received in the previous reward stage ($p^{s3}_{ji,t-1}$ in $s2$, and $p^{s2}_{ji,t-1}$ in $s3$) are positive and significant.

We replicate this analysis for the third period (t=18) and, arbitrarily, for the tenth (t=25), and the same pattern emerges; see columns (v)-(viii) in Table 4. Subject j's behavior in $s1$ does not affect $p_{ij,t}$ in $s2$ or $s3$; what matters is the number of reward tokens received from j in the previous reward stage. Additional support for this conclusion comes from the temporal pattern of the magnitudes of the intercept, and of the specifications' coefficients of determination (R^2). The 'exogenous' propensity to send reward tokens decreases as the game proceeds, while the specification's explanatory power increases substantially.

We summarize these results as follows.

Result 4 In the regressions explaining $p^{s2}_{ij,t}$ and $p^{s3}_{ij,t}$, we find that (i) effort only affects rewarding decisions in $t = 16$, (ii) the coefficient on $p_{ji,t}$ in the previous reward stage increases over time while the 'exogenous' propensity to send reward tokens decreases, and (iii) the explanatory power of past rewards received increases over time.

Hence, we conclude that the participants in the PF sessions do not use the rewards to enforce cooperation in the social dilemma in 2SR-PF. Instead, they use them to increase their own private earnings by establishing bilateral exchange relationships. Similar patterns are observed in 1SR-PF — albeit less clear cut. The key difference between 1SR-PF and 2SR-PF is that the higher frequency of the option to send reward tokens speeds up the formation of connections, but otherwise the qualitative results are the same.

4 Additional evidence on the motivation to use reward tokens

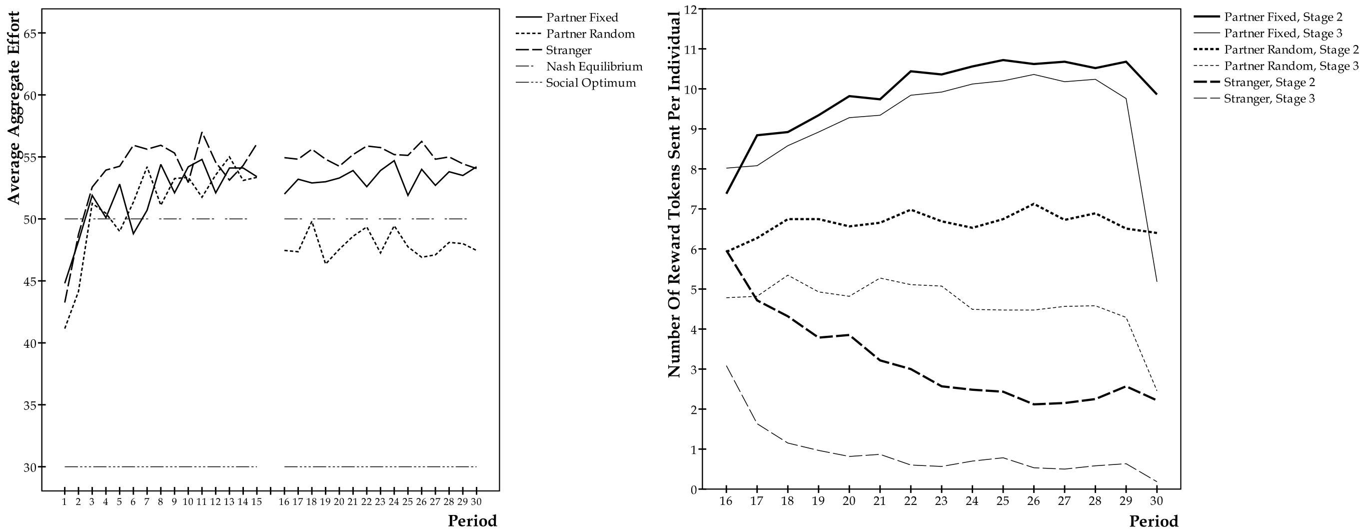
We run 2SR using two alternative matching protocols, Partner Random (PR) and Stranger (S). We do so to verify our conclusion that rewards are not being used to enforce cooperation in the social dilemma, but also to gain insight into the underlying mechanism. Subjects receive information on their peers' past behavior in neither the PR or S sessions, but they can still base their decision to reward on their peers' behavior in $s1$ — to sustain (or enforce) future cooperation, or just to (non-strategically) reward fellow group members' kind actions in $s1$. Of course, because of the random rematching between periods, subjects in the S sessions have less incentives to send reward tokens than those in the PR and PF sessions — but they may still decide to do so in 2SR. For that reason, we focus our attention on the 2SR game rather than on that of 1SR. But in the 2SR-PF and 2SR-PR sessions the incentives to reward are equally strong. In fact, if enforcement and non-strategic rewarding are the main motivations behind the use of reward tokens, play in all three stages of 2SR-PR should be identical to that in 2SR-PF.

In Figure 3(a) we present the results of the average aggregate amount of effort invested in $s1$ in the 2SR-PR and 2SR-S sessions, and the average numbers of reward tokens sent in $s2$ and $s3$ are shown in Figure 3(b). For ease of comparison we also include the results for the 2SR-PF sessions.

As stated above, play in the Fixed and Random Partner protocols should be identical if reward tokens are used exclusively to enforce cooperation. The data reject this hypothesis because the average aggregate effort in 2SR-PR is below that in 2SR-PF (albeit marginally so because $p = 0.105$ according to the relevant Mann-Whitney U test with $N_1 = 10, N_2 = 11$) while $p^{s2}_{ij,t}$ and $p^{s3}_{ij,t}$ are significantly higher in 2SR-PF than in 2SR-PR ($p < 0.001$ in both cases, as indicated by Mann-Whitney U tests with $N_1 = 10, N_2 = 11$).

Result 5 Even though the numbers of reward tokens sent in $s2$ and $s3$ in 2SR-PR are about half of those in 2SR-PF, efficiency in $s1$ of 2SR-PR is higher than that in 2SR-PF.

So, play in the 2SR-PR and 2SR-PF sessions seems to differ, and hence we turn to the question whether subjects in the 2SR-PR and 2SR-S sessions send their reward tokens to the ones investing least in the social dilemma stage — at least in period 16. We calculate the percentages of subjects



(a) **(b)**

Figure 3 **(a)** Average aggregate effort in the social dilemma stage in the 2SR-PF, 2SR-PR and 2SR-S sessions. **(b)** Average number of reward tokens sent per subject in stage 2 and stage 3 in the three matching protocols.

choosing to indiscriminately spread their rewards over all four other members of their group (i.e., the opposite of selective use of reward tokens). In the 2SR-PF and 2SR-S sessions these are 34 percent and 42 percent, respectively, while only 16 percent of the subjects does so in the 2SR-PR sessions.

Result 6 In $s2$ of period 16 of the 2SR-S sessions, no less than 40 percent of the subjects sends an equal number of reward tokens to all four other group members. In the 2SR-PR and 2SR-PF sessions these percentages are respectively 16 and 34 percent, which means that the use of reward tokens in the 2SR-PR sessions is more selective than that in the 2SR-PF sessions.

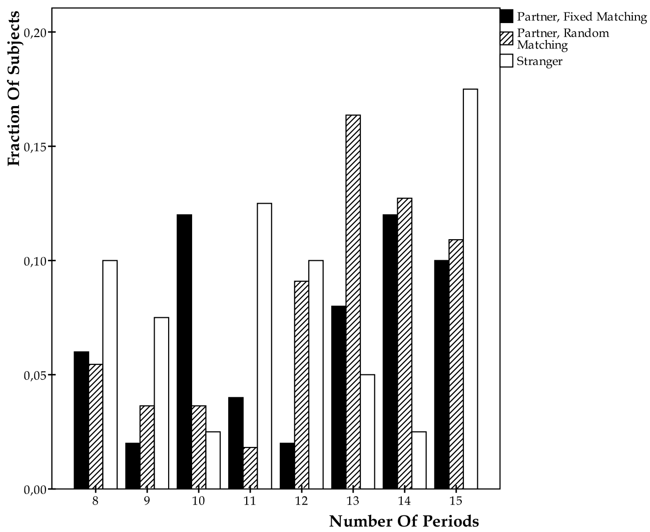
Hence, we find that the option to send rewards is not used very selectively in the S sessions (Result 6). Also, the play in the 2SR-PF and 2SR-PR sessions are not identical because (i) efficiency in $s1$ of 2SR-PR only just fails to be significantly higher than that in 2SR-PF while the number of reward tokens sent in $s2$ and $s3$ are significantly lower (Result 5), and (ii) the decision to send rewards in the 2SR-PR sessions is substantially more selective (Result 6). These observations do not support the hypothesis that reward tokens are used to enforce cooperation in the social dilemma stage. However, they also send a conflicting message. Result 6 suggests that subjects in 2SR-S try to find partners willing to reciprocate to rewards received within the same period, while Result 5 suggests that there is a real efficiency improvement associated with the more selective use of reward tokens in 2SR-PR. In the remainder of this section we try to reconcile these two results.

Let us first have a closer look at the subjects' individual behavior in the social dilemma stage in each of the three session types. The average variances in effort within groups over periods 16-30 are 2.1, 2.5 and 1.9 in the PF, PR and S sessions of 2SR, respectively. The within-group variance is highest in 2SR-PR[10], and closer inspection of the temporal pattern (available upon request) reveals that it does not really decline over time either. Thus, we find an important difference in play between 2SR-PF and 2SR-PR. While effort decisions and the number of rewards sent should be identical if the predominant use of rewards is to sustain cooperation in the social dilemma, we find that convergence to symmetric effort levels is least strong in 2SR-PR.

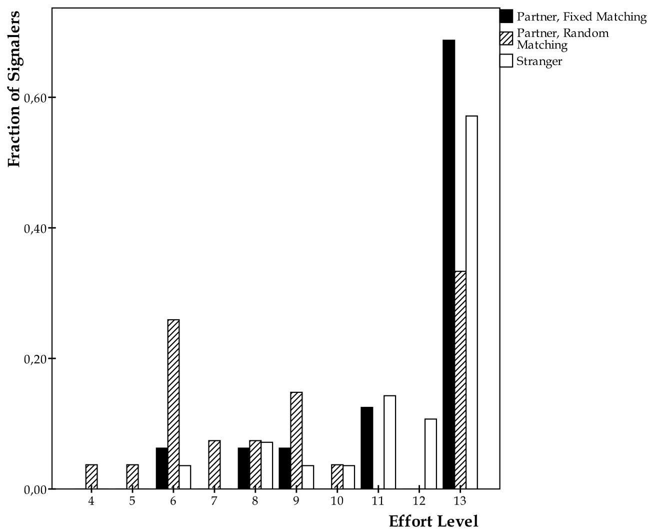
Result 7 Compared to the other two session types in 2SR, we find that in 2SR-PR the within-group variance in effort remains highest.

To further explore the differences in within-group convergence of effort levels between the three matching protocols, we calculate (i) the number of periods in which a subject chooses a particular effort level in each of the three session types, and (ii) conditional on choosing the same effort level for a number of periods, what effort level was chosen. The results are shown in Figure 4(a) and Figure 4(b), respectively.

[10]Using a Mann-Whitney test ($N_1 = N_2 = 15$), the variance in 2SR-PR is significantly higher than in 2SR-PF and 2SR-S at $p < 0.01$ in both cases.



(a)



(b)

Figure 4 **(a)** Fraction of subjects who choose the same effort level in $s1$ for 8 periods or more in 2SR. **(b)** Distribution of effort levels chosen by subjects who choose the same effort level for 12 periods or more.

Figure 4(a) presents the frequency of subjects choosing the same effort level for eight periods or more in the 2SR treatment. Again, we find important differences in play between 2SR-PR on the one hand, and 2SR-PF and 2SR-S on the other. Almost 50 percent of the subjects in 2SR-PR choose the same effort level for 12 periods (out of a maximum of 15) or more, while the numbers in 2SR-PF and 2SR-S are 32 and 35 percent, respectively. Conditional on choosing the same effort level for 12 periods or more, Figure 4(b) presents the distribution of effort levels chosen. In total, 64 percent of the subjects in 2SR-PR pick effort levels strictly below the Nash equilibrium level ($x \leq 9$), while the subjects of 2SR-S and 2SR-PF are clearly overrepresented at effort levels above the Nash equilibrium ($x \geq 10$), with frequencies of 82 percent and 81 percent, respectively.

Result 8 Compared to the other two session types, subjects in 2SR-PR revise their effort decisions less frequently, and they also tend to choose lower effort levels.

So we find that in 2SR-PR (i) the within-group variance in effort remains highest (Result 7), (ii) the within-subject variance of effort is lowest (Result 8), (iii) the use of reward tokens in $s2$ is most selective (Result 6), and (iv) efficiency in $s1$ is highest (Result 5). Combined with the fact that $p^{s2}_{ij,t}$ and $p^{s3}_{ij,t}$ in 2SR-PR remain fairly constant (see Figure 3(b)), one explanation might be that subjects use their effort decisions in $s1$ to signal their identity to overcome the problem of subject identifiers being reshuffled between periods — in order to still be able to establish mutually profitable bilateral exchange relationships. We offer two pieces of evidence for this: (i) non-parametric tests regarding reciprocity in the number of reward tokens sent between 'signalers', and (ii) regression analyses aimed at explaining the use of reward tokens in $s2$ and $s3$. Let us first define a 'signaler':

Definition A 'signaler' in 2SR-PR is a subject who chooses the same effort level in $s1$ for twelve periods or more.[11]

Our first piece of evidence supporting our signalling hypothesis is the way in which two signalers exchange reward tokens. If tokens are used as a way to sustain cooperation in the social dilemma or to non-strategically reward 'good behavior' in $s1$, one expects the 'partner' with a higher (lower) effort level in the social dilemma stage to give more (fewer) reward tokens than the other 'partner'. If subjects simply view their partner's effort level as a signal of their identity, there would be no systematic difference in the number of reward tokens sent by the two partners.

We test this by analyzing the number of reward tokens exchanged between all pairs of signalers with unequal effort levels. In neither $s2$ nor $s3$ can we reject the null hypothesis of no difference in the average number of rewards sent by either partner, with p-values of 0.33 and 0.61 respectively, according to the relevant Wilcoxon tests (with $N_1 = N_2 = 20$).

Result 9 Differences in effort levels chosen by two 'signalers' in 2SR-PR do not affect the net flow of reward tokens exchanged.

[11] All conclusions are robust against using other cutoffs — results available upon request.

Second, we try to explain $p^{s2}_{ij,t}$ and $p^{s3}_{ij,t}$ in 2SR-PR using a similar setup as reported in Table 4. The controls used for stage 1 behavior are $\text{Max}\{0, \overline{x}_{-j,t} - x_{j,t}\}$ and $\text{Max}\{0, x_{j,t} - \overline{x}_{-j,t}\}$. Also, the variable $p^2_{ji,t}$ is included in the regressions for stage 3. This variable measures the direct reciprocity among subjects within periods. However, in contrast to the PF sessions (both 1SR and 2SR), subjects cannot directly reciprocate the number of reward tokens received between periods because identity labels are changed between periods. Therefore, we have included variables that capture the signaling mechanism that subjects may use. The first variable is $I(\text{Signal}_{j,t})$, which has a value of 1 if the effort level of subject j in period t, $x_{j,t}$, has been among the effort levels that subject i observed in period $t-1$.[12] Clearly, if this variable shows up significantly, the data provide support for the hypothesis that effort levels are used to signal one's identity. We also calculate $I(\text{Signal}_{j,t}) \times x_{j,t}$ to check whether the strength of the signal is inversely related to the level chosen: the lower the signal, the more costly it is, and hence the more trustworthy the signaller may be. Finally, we have included the interaction term $I(\text{Signal}_{j,t}) \times p^3_{j,t-1}$. This variable links current rewarding with the number of reward tokens subject i has received in stage 3 of the previous period from a group member who potentially is a signaler. We report the results of periods 17, 18, 19, and 25 in Table 5.[13]

The results are as follows. First, subjects seem to condition their stage 2 and stage 3 rewards on stage 1 behavior in the 2SR-PR sessions, as opposed to the 2SR-PF sessions where stage 1 behavior had no effect at all. However, the economic importance of this mechanism is quite small. In many periods the coefficients on the $s1$ variables fail to be statistically significant, and if

[12] That means that $I(\text{Signal}_{j,t}) = 1$ if $x_{j,t} = \{x_{1,t-1}, x_{2,t-1}, ..., x_{5,t-1}\}$, and zero otherwise. Hence, $I(\text{Signal}_{j,t}) = 1$ if $x_{j,t} = x_{1,t-1}$, but because subjects have no means of inferring other subjects' identities other than by their effort decisions, the signalling variable is also equal to 1 if a subject chooses the same effort level – accidentally, or on purpose – as one (or more) of their fellow group members in the previous period.

[13] We do not include period 16 (as done in Table 5) in Table 5 because of the lagged variables in the regression analysis. We did run the same regression as in columns (i) and (ii) of that table for period 16, and we find that stage 1 behavior has a significant impact on stage 2 rewarding, but not on stage 3. The magnitude of these variables are comparable to those reported in Table 5. The variable p^2_{ji} is highly significant and has a value of 0.70.

Dependent variable: Reward tokens sent by subject i to subject j in stage 2 and stage 3 in the PR sessions								
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
	Period 17		Period 18		Period 19		Period 25	
	$s2$	$s3$	$s2$	$s3$	$s2$	$s3$	$s2$	$s3$
$\text{Max}\{0, \overline{x}_{-j,t} - x_{j,t}\}$	0.208	−0.020	0.085	0.042	0.215**	0.088	0.168*	0.053
	(0.132)	(0.088)	(0.078)	(0.072)	(0.084)	(0.053)	(0.090)	(0.032)
$\text{Max}\{0, x_{j,t} - \overline{x}_{-j,t}\}$	−0.220*	−0.110**	−0.361**	−0.036	−0.206**	−0.019	−0.091	0.016
	(0.106)	(0.041)	(0.116)	(0.026)	(0.087)	(0.059)	(0.142)	(0.048)
$p^2_{ji,t}$		0.729***		0.606***		0.598***		0.698***
		(0.075)		(0.110)		(0.095)		(0.081)
$I(\text{Signal}_j)$	1.438	0.693	1.212	1.162*	1.679***	1.211*	2.192**	0.625*
	(1.398)	(0.856)	(0.977)	(0.597)	(0.448)	(0.561)	(0.848)	(0.341)
$I(\text{Signal}_j) \times x_j$	−0.167	−0.074	−0.155*	−0.107*	−0.170***	−0.123*	−0.227***	−0.075*
	(0.127)	(0.083)	(0.077)	(0.054)	(0.050)	(0.056)	(0.063)	(0.038)
$I(\text{Signal}_j) \times p^3_{ji,t-1}$	0.265**	0.075	0.335*	0.114	0.291***	0.200*	0.200**	0.073*
	(0.086)	(0.053)	(0.178)	(0.068)	(0.084)	(0.094)	(0.070)	(0.033)
Constant	1.466***	0.159	1.924***	0.124	1.367***	−0.065	1.431**	−0.129
	(0.349)	(0.164)	(0.333)	(0.266)	(0.167)	(0.153)	(0.496)	(0.179)
N	220	220	220	220	220	220	220	220
R^2	0.1999	0.6755	0.3110	0.5296	0.2883	0.5912	0.1819	0.7053

Table 5 OLS regression estimates of the number of reward tokens sent in the first reward stage ($s2$) and in the second reward stage ($s3$) in period 2-4 of 2SR-PR, and in its tenth period (i.e., $t = 17 - 19$, and 25). Standard errors, clustered at the group level, are reported between parentheses. ***: significant at the 1%-level, **: significant at the 5%-level, *: significant at he 10%-level.

they are, $s1$ behavior tends to increase or decrease the number of rewards received in either stage 2 or stage 3 with less than one token (because $|\overline{x}_{-j,t} - x_{j,t}| \leq 1$ in about 78 percent of the cases).

Second, the dummy variable $I(\text{Signal}_{j,t})$ increases in magnitude as the game progresses when it comes to stage 2 rewards. The same holds for the interaction term $I(\text{Signal}_{j,t}) \times x_{j,t}$: subjects choosing higher effort levels as a signal tend to attract fewer reward tokens than signalers with lower effort levels. All in all, this means that signalers attract more reward tokens than non-signalers, but that it takes some periods before the subjects learn that signaling is a profitable strategy. Note that in stage 3, the magnitude and significance of these signaling variables tend to be small. This makes intuitive sense, because in $s3$ subjects have a direct test of a group member's trustworthiness: the number of reward tokens received in stage 2 ($p^2_{ji,t}$). This variable is large and highly significant in $s3$ in all periods, comparable to the results of the 2SR-PF sessions. The importance of the variable ($p^2_{ji,t}$) is reflected in two other ways. The coefficients of determination (R^2) in the regressions for stage 3 rewarding are much greater than those of stage 2, and the constant term becomes insignificant in stage 3, implying that the exogenous propensity to reward as observed in stage 2 is no longer present.

Third, further evidence that signaling is profitable comes from the positive sign of the coefficient on $I(\text{Signal}_{j,t}) \times p^3_{ji,t-1}$ — subjects try to reciprocate reward behavior between periods. However, as expected and in line with the signaling hypothesis, this variable tends to be insignificant in the $s3$ regressions.

We summarize our findings as follows:

Result 10 In the regression explaining $p^2_{ij,t}$ and $p^3_{ij,t}$ in 2SR-PR, we find that (i) stage 1 effort has only a very weak impact on the number of rewards received, (ii) subjects who signal their identity by choosing the same effort level as in the previous period receive more reward tokens, (iii) a large part of the variation in stage 3 rewarding is explained by $p^2_{ji,t}$ which shows that subjects use reward tokens mainly as a way to bilaterally profit from an exchange of rewards, and (iv) for signalers,

$p^3_{ji,t-1}$ is positive and significant in $s2$ rewarding behavior.

5 Discussion and conclusions

In society, behavior of agents is embedded in a system of interpersonal relations where individual welfare depends on activities that require multilateral cooperation, and also on alternative economic activities that only require bilateral cooperation (Granovetter (1985), Bowles & Gintis (2002)). Examples of the former are common property resources like fish or water, and examples of the latter include helping others harvesting their crops, and child minding. Ideally all agents cooperate in both types of activities; in a less ideal world, refusing to cooperate bilaterally (that is, withholding 'rewards') can be used as an instrument to enforce cooperation in the multi-person social dilemma.

In this paper we experimentally test whether indeed 'rewards' can sustain cooperation in a multi-person social dilemma. We implement a non-linear public bad game with two stages in which subjects have the option to send 'reward tokens' to their fellow group members. We implement three different matching protocols (Partner Fixed, Partner Random and Stranger). All three allow subjects to condition their rewarding decisions in both reward stages on their peers' behavior in the social dilemma stage of the period, but also, in the second reward stage, on the number of rewards received from their peers in the first rewarding stage. Compared to the Partner Random sessions, there is less reason in the Stranger sessions to reciprocate to either decision of one's peers, while in the Partner Fixed sessions (in both 1SR and 2SR) subjects can take into account the entire history of their peers' decisions — not just those taken in the current period.

The results of our laboratory experiments suggest that having the option to selectively increase (or not to increase) one's fellow group members' welfare does not increase efficiency in the multi-agent social dilemma. While efficiency in the social dilemma stage in the PF sessions remains low throughout the experiment in both 1SR and 2SR, the propensity to send reward tokens increases over time. We find that this increased propensity to send rewards is due to subjects positively reciprocating to the number of reward tokens received from fellow group members in the previous reward stage (either in the current or in the previous period), while decisions in the social dilemma stage are found to play a role in the first period only. Subjects' propensity to positively reciprocate to their peers' bilateral activities is larger than their propensity to negatively reciprocate to their peers' behavior in the social dilemma.

Additional support from this claim comes from the analysis of play in 2SR when using the Stranger and Partner Random matching protocols. Even when new groups are created at the beginning of every new period (in the 2SR-S sessions), a substantial share of the participants equally divide whatever number of reward tokens they give over all four other members of their group — independent of their peers' behavior in the social dilemma. And if enforcement were the main motivation to use rewards, play in the two Partner matching protocols (2SR-PF and 2SR-PR) should be identical. To be able and willing to use rewards to enforce cooperation, subjects need to be (i) ensured of future interactions with the same subjects, and (ii) able to observe the behavior of one's peers in the social dilemma stage of the current period. These conditions are met in both the Partner Fixed and Partner Random sessions, but still we find marked differences in play between the two — the drive to engage in the exchange of reward tokens is so strong that subjects actually use their decisions in the social dilemma to signal their identity in the Partner Random sessions.

So, from the analysis of play in all three session types we find that subjects attach more weight to establishing bilateral cooperation in the reward stages than to affecting their peers' behavior in the social dilemma stage. These conclusions are diametrically opposite of those drawn by Vyrastekova & van Soest (2008) and Rand et al. (2009a), even though our design is close to the ones developed in these studies. However, Vyrastekova & van Soest (2008) note that cooperation decays over time in their experiment — having two reward options per period (in our setup) rather than one (as in theirs) just speeds up this process. And in Rand et al. (2009a), subjects can decide to incur a cost of 4 points to increase the payoffs of a fellow group member by 12 points, yes or no, and this decision can be made separately for each of the other members of one's group. That means that subjects need to elicit bilateral cooperation of *all* their peers to maximize their returns from the 'reward activity' while in our setup establishing a bilateral exchange with one (trustworthy) subject suffices. Indeed, in our experiment subjects can always find a subject whose behavior in the social dilemma is close to theirs, and still achieve maximum efficiency in the 'reward activity' independent of how they themselves behave in the social dilemma. That means that the 'carrot' loses its bite — it can be reaped by all, not by just those who act cooperatively in the social dilemma.

We conclude that rewards are unlikely to sustain cooperation in real-world multi-person social dilemmas. Unlike the predictions of standard game theory this is not due to the fact that agents are unwilling to cooperate. On the contrary: agents are too eager to cooperate in small-scale settings — so eager that they do so at the expense of the larger-scale social dilemma.

References

Bowles, S. & Gintis, H. (2002). Social capital and community governance. *Economic Journal*, *112*, 416–436.

Denant-Boemont, L., Masclet, D., & Noussair, C. (2007). Punishment, counterpunishment and sanctions in a social dilemma experiment. *Economic Theory*, *33*, 145–167.

Fehr, E. & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *The American Economic Review*, *90*(4), 980–994.

Fehr, E. & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*, 137–140.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–178.

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? evidence from a public goods experiment. *Economics Letters*, *71*(3), 397–404.

Gächter, S., Renner, E., & Sefton, M. (2008). The long-run benefits of punishment. *Science*, *322*, 1510.

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, *91*, 481–510.

Hermann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science*, *319*, 1362–1367.

Maier-Rigaud, F., Martinsson, P., & Staffiero, G. (2010). Ostracism and the provision of a public good: Experimental evidence. *Journal of Economic Behavior and Organization*, *73*(3), 387–395.

Masclet, D. (2003). Ostracism in work teams: A public good experiment. *International Journal of Manpower*, *24*(7), 867–887.

Masclet, D., Noussair, C., Tucker, S., & Villeval, M. (2003). Monetary and nonmonetary punishment in the voluntary contributions mechanism. *The American Economic Review*, *93*(1), 366–380.

Nikiforakis, N. (2008). Punishment and counterpunishment in public goods games: Can we govern ourselves? *Journal of Public Economics*, *92*, 91–112.

Nikiforakis, N. & Engelmann, D. (2011). Altruistic punishment and the threat of feuds. *Journal of Economic Behavior and Organization*, *Forthcoming*.

Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a sword: Self governance is possible. *American Political Science Review*, *86*(2), 404–417.

Rand, D., Dreber, A., Ellingsen, T., Fudenberg, D., & Nowak, M. (2009a). Positive interactions promote public cooperation. *Science*, *325*, 1272–1275.

Rand, D., Dreber, A., Ellingsen, T., Fudenberg, D., & Nowak, M. (2009b). Weighing reward and punishment: Reply. *Science*, *326*(18 December), 1632–1633.

Sefton, M., Shupp, R., & Walker, J. (2007). The effects of rewards and sanctions in provision of public goods. *Economic Inquiry*, *45*(4), 671–690.

Sutter, M., Haigner, S., & Kocher, M. (2010). Choosing the stick or the carrot? Endogenous institutional choice in social dilemma situations. *Review of Economic Studies*, *forthcoming*.

Vyrastekova, J. & van Soest, D. (2008). On the (in)effectiveness of rewards in sustaining cooperation. *Experimental Economics*, *11*, 53–65.

Yamagishi, T. (1988). The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology*, *51*, 110–116.