University of Innsbruck

Working Papers
in
Economics and Statistics

Choosing the carrot or the stick? – Endogenous institutional choice in social dilemma situations

Matthias Sutter, Stefan Haigner and Martin Kocher

2008-07

Choosing the carrot or the stick? – Endogenous institutional choice in social dilemma situations[#]

Matthias Sutter[+], Stefan Haigner[§] and Martin G. Kocher[*]

Abstract

We analyze an experimental public goods game in which group members can *endogenously* determine whether they want to supplement a standard voluntary contribution mechanism with the possibility of rewarding or punishing other group members. We find a large and positive effect of endogenous institutional choice on the level of cooperation in comparison to exogenously implemented institutions. This suggests that participation rights enhance cooperation in groups. With endogenous choice, groups typically vote for the reward option, even though punishment is actually more effective in sustaining high levels of cooperation. Our results are evaluated against the predictions of social preference models.

JEL classification: C72, C91, C92

Keywords: public goods, punishment, reward, endogenous institutional choice, voting, experiment

This version: 3rd May 2008

[#] This paper is a substantially revised and extended version of our CEPR Working Paper No. 5497 from February 2006. We would like to thank Simon Gächter, Rudolf Kerschbamer, Michael Kosfeld, Stephan Kroll, Wolfgang Luhan, Nikos Nikiforakis, Axel Ockenfels, Louis Putterman, Christina Strassmair and Florian Wakolbinger for very helpful comments and Olga Mellizo for language editing. We are grateful to the Center for Experimental Economics at the University of Innsbruck (sponsored by *Raiffeisen-Landesbank Tirol*) and the *Austrian National Bank* (Jubiläumsfondsprojekt No. 9879) for financial support.

[+] University of Innsbruck and University of Gothenburg. Address for correspondence: Department of Public Finance, University of Innsbruck, Universitaetsstrasse 15, A-6020 Innsbruck, Austria. e-mail: matthias.sutter@uibk.ac.at.

[§] University of Innsbruck.

[*] University of Innsbruck and University of Munich.

1 Introduction

In this paper we study cooperation in social dilemma situations using an experimental public goods game. In contrast to most of the previous work, we let subjects decide endogenously on the institutions governing their interactions. In particular, subjects can vote on whether they want to supplement a standard voluntary contribution mechanism (VCM) with the possibility of rewarding or punishing other subjects or whether they prefer to use the standard VCM. We find that the endogenous institutional choice has a large and positive effect on cooperation levels in the public goods game compared to a setting in which these institutions have been determined exogenously (by the experimenter). This suggests that participation rights foster cooperation, even beyond the level that can be achieved through reward and punishment when these institutions have been implemented exogenously.

By varying the effectiveness of the available punishment or reward technology (defined as the ratio of monetary consequences of being punished or rewarded to the costs of punishing or rewarding), we are able to show that the reward option is chosen almost exclusively when reward or punishment are highly effective, i.e. have a high ratio. However, the standard VCM is the preferred institution when the technology is less effective. The punishment option – though very effective in raising contributions – is only very rarely chosen endogenously, indicating that most subjects prefer an institutional environment without punishment.

We compare our results to predictions based on standard game theoretic analysis (under the assumptions of rationality and selfishness) and predictions derived from two prominent social preference models (Fehr and Schmidt, 1999, and Charness and Rabin, 2002). While the standard game theoretic model essentially fails to explain contributions and institutional choices, the models with social preferences are able to predict many, but not all behavioral regularities in our experiment. In particular, the inequity aversion model of Fehr and Schmidt (1999) largely fails to account for the observed behavior under the reward institution, whereas the model of Charness and Rabin (2002) can only explain the power of the punishment institution for increasing cooperation when punishment is exclusively driven by reciprocity.

Our paper is related to a large number of studies dealing with the conditions and institutions that foster cooperation in social dilemma situations. In recent years, the effects of reward and especially of punishment on cooperation in social dilemma games have received considerable attention, for instance in Fehr and Gächter (2000, 2002), Andreoni et al. (2003), Masclet et al. (2003), or Sefton et al. (2007). These and most subsequent papers discussed in the next section share the feature that the institution governing the interaction of subjects has been determined exogenously by the experimenters. Although these studies have provided illuminating and very important insights into the determinants of cooperation in groups, they are unable to answer two straightforward questions that are central to our paper: (1) Which institutions do subjects actually prefer and choose themselves in a social dilemma? (2) How does the endogenous choice of rules affect group members' interaction and their level of cooperation? The answers to both questions are important in order to judge whether the inferences about the treatment effects of different – but exogenously imposed – institutions are reliable and can be generalized to situations where subjects can self-select or avoid certain institutions through an endogenous choice.

In real life, groups often establish themselves the (explicit or implicit) rules that govern their interaction in social dilemmas. On a large-scale international level, for instance, one might think of the United Nations Framework Convention on Climate Change, the so-called Kyoto-protocol, or the United Nations Security Council. Representatives in these organizations establish rules – and possible sanctions in case of their violation – on how to provide the public good of carbon dioxide emission reduction or global security. Regarding small group interactions, work teams in companies or sports teams typically develop social norms on how to deal with 'rate busters' or on how to reward team members for contributions to the team success. In the private sphere, people sharing a flat decide on whether they want to impose rewards or sanctions for the private contributions to public goods like cleaning, shopping, cooking, disposing garbage, and so on. Obviously, reward and punishment do not always exist as a monetary fine or a monetary gift, but also in the form of social approval or disapproval or in many other forms that might be at least as effective.

Seen from a broader perspective, our approach of endogenous institutional choice is linked to the question whether democratic participation rights have an influence on individual behavior. Frey (1994), Frey et al. (2004) or Pommerehne and Weck-Hannemann (1996), for instance, have studied the effects of the degree of direct-democratic participation rights (as a form of endogenous institutional choice) on the level of cooperation among citizens and between citizens and the public sector. Their findings indicate that stronger direct-democratic participation rights increase cooperation and reduce, for example, tax evasion (see also Feld and Tyran, 2002). In our experiment we can examine how the decision to participate in the vote on the institution is linked to a subject's contribution level in the public goods game and to her punishing and rewarding behavior. Hence, we can establish a link between voting and economic decision making. We are also able to measure the economic benefit of an endogenous institutional choice by comparing contributions under the endogenous choice with contributions under identical, but exogenously determined, institutions (a comparison that is hardly possible with field data). Our results provide clear evidence that endogenous institutional choice yields substantial benefits. In other words, we observe a clear democratic participation rights-premium to cooperation.

The remainder of this paper is structured as follows. In Section 2 we give a brief overview of the related literature. Section 3 presents the public goods game and the different institutions that can be implemented in our setting. Section 4 describes our experimental design in detail, and in Section 5 we derive several predictions on contribution levels and on voting behavior. Section 6 reports the experimental results, and finally Section 7 discusses our findings and concludes the paper.

2 Related literature

Our paper builds upon a considerable number of studies on the effects of *exogenously* imposed reward and punishment options in social dilemmas. Fehr and Gächter (2000, 2002) have analyzed the contributions to a public good when subjects can punish each other after having observed the other group members' contributions. Even though using the punishment option is at odds with selfish money-maximization in their experimental design, it is used rather frequently, thereby inducing very high levels of cooperation.[1]

[1] In political science or social psychology, related results have been documented earlier by Yamagishi (1986) or Ostrom et al. (1992), for example. Recently, the emerging field of neuro-economics has examined the physiological foundations of punishment. De Quervain et al. (2004) have shown that using the punishment option actually triggers some kind of reward for the punishing subject, as it stimulates the very area in the brain that is also activated when one is being rewarded. This finding is one rationalization of the fact that punishment is even used in one-shot interactions (respectively in the last round of repeated

A very important feature when analyzing the impact of punishment is the demand for punishment under varying prices and/or effectiveness. While some papers use only one punishment technology, several other papers, including ours, compare the behavioral impact of differently effective punishment (and reward) mechanisms. Only a few papers concentrate on eliciting a complete demand function for punishment by a systematic variation of the prices of punishment. Anderson and Putterman (2006) as well as Carpenter (2007) show that the law of demand holds: the quantity of punishment demanded is lower with a higher price.

Kroll et al. (2007) study the role of punishment in yet another context where group members can vote on non-binding minimum contributions to a public good. When failure to comply with the non-binding minimum can be punished, contributions increase significantly. Thus, it is not the outcome of the non-binding vote that is important for higher contributions in such a setting, but the possibility of punishing defectors.[2]

One of the few papers that take reward into account is Sefton et al. (2007). They have included a reward option in their public goods experiments. Like punishment, reward has been found to increase cooperation in comparison to the standard VCM-setting. Reward, however, seems to be less suitable for sustaining high levels of cooperation, as the decay of contributions over time is faster with the reward option than with the punishment option. Hence, using the carrot (i.e., reward) seems to be less efficient for enhancing cooperation in a group than the stick (i.e., punishment). A similar result is obtained by Fehr and Gächter (2003). They find punishment to be more effective than reward in mitigating the free-rider problem and also observe a strong impact of the effectiveness of reward and punishment. Andreoni et al. (2003) provide complementary results in a bargaining game where a proposer can offer a portion of a pie to a responder, and the responder can reward and/or sanction the proposer for her offer. It turns out that the distribution of offers is much

interactions) or when only third parties that are materially unaffected by the actions of the players can punish (Fehr and Fischbacher, 2004).

[2] Masclet et al. (2003) have addressed the question whether the monetary consequences of punishment strengthen cooperation or whether the expression of disapproval of another subject's behavior through punishment is possibly equally or even more important. They have compared a public goods game with monetary sanctions to one with only symbolic, non-monetary sanctions. Cooperation remains higher with the monetary sanctions, but even the non-monetary sanctions raise cooperation levels above those in the standard VCM. Rege and Telle (2004) also report a positive effect of (non-monetary) social (dis)approval on cooperation in a public goods game.

more affected by the punishment option than by the reward option, even though the latter also has an effect.[3]

Decker et al. (2003) have been the first to let subjects choose endogenously between different voting rules for executing the punishment of group members. They let group members propose their desired punishment intensity for each group member, and then they let subjects submit bids in a first-price auction whether the execution of punishment shall require a simple majority or unanimity. It turns out that a unanimity rule is preferred (and receives higher bids). A related paper by Ertan et al. (2005) shows that most groups rule out the possibility of punishing high contributors, but they implement an institution where the punishment of low contributors is possible. Contrary to our paper – and two others discussed below – neither Decker et al. (2003) nor Ertan et al. (2005) allow groups to choose whether they want to have a punishment option at all. In addition to that, both papers do not compare their endogenous treatments with behavior under identical, but exogenously imposed institutions. Consequently, they did not assess the effects of endogenous institutional choice on cooperation, nor have they considered a reward option.

There are two recent papers that are more directly related to the research questions in this paper. The first one by Botelho et al. (2005) extends the design of Fehr and Gächter (2000, 2002) by allowing subjects to choose directly whether they want to operate under a punishment institution or under a standard VCM. After subjects have experienced both the VCM and the VCM with the punishment option in a (partly perfect) stranger design, they can choose the governing institution for one final round (with substantially increased salience). Botelho et al. (2005) find that almost none of their cohorts opt for the punishment institution. From an ex-post point of view, this choice is justified because the overall efficiency of interactions (including punishment costs) is significantly lower in a world with punishment than without. Compared to Botelho et al. (2005), our experiment considers not only a punishment, but also a reward institution, which will be shown to be rather attractive among groups when it is compared to the punishment institution. Furthermore, we use a partner design because we are mainly interested in the institutional choice of groups whose members will interact repeatedly after the initial decision on the

[3] Reward in another interactive context – in the form of (voluntary) bonus contracts in a principal-agent relationship – is studied, e.g., by Fehr et al. (2007) or Fehr and Schmidt (2004). Abbink et al. (2000) use the moonlighting game to analyze punishment and reward, and Croson and Konow (2008) a variant of the dictator game with reward and punishment.

governing institution. Finally, in contrast to Botelho et al. (2005) our subjects may abstain from participating in the vote. We are, therefore, able to draw inferences on the relationship between vote participation and behavior in the public goods game.

Gürerk et al. (2006) study the *self-selection of individuals* into groups with an *exogenously given* institution. They find that subjects self-select into groups where they can both punish and reward other group members much more often than into groups with a standard VCM procedure only. Our approach is different in several aspects: First, we let *groups self-select the institution* that shall apply to their interaction. Therefore, our approach resembles a choice of rules within a fixed group in contrast to the voting by feet-approach of Gürerk et al. (2006). Second, we compare the attractiveness of reward and punishment when they are equally effective and when subjects can choose directly between both institutions. In Gürerk et al. (2006) punishment had much stronger consequences than reward. Third, in Gürerk et al. (2006) subjects could switch between different institutions as often as they wished to without any costs. In our design, we implement small costs of voting in order to be able to assess potential effects of vote participation on cooperation, and sanctioning and rewarding behavior. Group members are not required to participate in this vote, but those abstaining are bound by the collective decision of voters in their group.

3 Our public goods setup

Let $I = \{1, 2, \ldots, n\}$ denote a group of n subjects who interact in T periods. In each period $t \in \{1, 2, ..., T\}$ individual $i \in I$ receives an endowment E which can be allocated either to a private good or a public good. The voluntary contribution of individual i to the public good in period t must satisfy $0 \leq c_{i,t} \leq E$. Let C_t denote the sum of all group members' contributions (i.e. $C_t = \sum_{j=1}^{n} c_{j,t}$). Individual member i's payoff from her contribution in period t is given by

$$\pi_{i,t} = E - c_{i,t} + \gamma C_t \qquad (1)$$

The marginal per capita return (MPCR) from investing into the public good is denoted as γ, which satisfies $0 < \gamma < 1 < n\gamma$, meaning that the self-interested choice and the socially optimal one are in conflict.

We consider three variants of this basic game: (*i*) the standard voluntary contribution mechanism (VCM) as it has just been described, (*ii*) the VCM with a punishment option, and (*iii*) the VCM with a reward option. In the standard VCM, all decisions are made simultaneously and each period has only one stage (the contribution stage). In the VCM with punishment (reward) each period has two stages, first the contribution stage and, thereafter, a punishment (reward) stage. In the second stage, group members receive information about the other members' contributions and can punish (reward) each of the other group members. Punishment is costly for the punisher and the punished subject. Reward is costly for the rewarding subject and beneficial for the rewarded subject.

We implement reward or punishment as a *binary* decision. Hence, a group member cannot choose between different intensities of reward or punishment[4]. The effectiveness of reward or punishment, however, is exogenously varied. We use the term "leverage" and the symbol L to refer to different technologies. We consider $|L| = l$ or $|L| = h$, with $h > l$. A higher leverage means that the same choice of punishing or rewarding another subject has larger monetary consequences for the punished or rewarded subject (keeping constant the costs for the subject who punishes or rewards).

Taking into account the monetary consequences of the second stage in each period yields the following payoff function for member i in the VCM with punishment, or reward.

$$\pi_{i,t} = E - c_{i,t} + \gamma C_t + L\sum_{k \neq i} p_{ik,t} - \sum_{h \neq i} p_{hi,t} \qquad (2)$$

where $p_{hk,t} = 1$ if member k has punished (rewarded) member h in period t and zero otherwise.[5] The sign of the parameter L indicates whether punishment ($L < 0$) or reward ($L > 0$) applies. The absolute value of L captures the effectiveness of the punishment (reward)

[4] The reason to choose such a simple form of reward and punishment is that we wanted to make sure that subjects fully understand the available institutions in the endogenous treatments. Any more complicated reward or punishment technology would have made the choice task of the participants more difficult and would not have added important further insights. Our simple, binary punishment or reward technology has the additional advantage that it is very weak – contrary to all other studies where the intensity of reward or punishment could be varied substantially. If we find effects of punishment and reward with our weak, binary, technology, then it seems reasonable that a stronger technology with a broader range of intensities would also have positive effects on contribution levels.

[5] Normalizing the cost of punishment (reward) is always possible by an appropriate choice of L. Note that if the last two terms in equation (2) are dropped (or the punishment and reward option are not used), we are back in the standard VCM given by equation (1).

technology, with a larger $|L|$ indicating a stronger effect (a higher leverage) of one's own actions on the profits of others.[6]

4 Experimental design and procedure

Our experimental treatments were based on the public goods games introduced in the previous section. The parameters were set up as follows: group size $n = 4$, endowment per period $E = 20$ ECU (experimental currency unit), MPCR $\gamma = 0.4$, and number of periods $T = 10$. We used a design that systematically varies the following three factors:

- F1. The institution governing the provision of the public good: either (*i*) the *standard VCM*, (*ii*) the VCM with *punishment*, or (*iii*) the VCM with *reward*.
- F2. The leverage of punishment or reward: either (*i*) $|L| = 1$ or (*ii*) $|L| = 3$.[7]
- F3. The way in which the institution is determined: either (*i*) *exogenously* by the experimenter or (*ii*) *endogenously* by the group itself through a voting procedure.

Table 1 summarizes our experimental treatments and the number of independent observations (i.e. groups) in each treatment. We used a partner matching, which yields one statistically independent observation per group of four subjects.

Table 1 about here

[6] Needless to say, it is impossible to design a straightforward mechanism where punishment and reward are equivalent in terms of efficiency, unless the Nash equilibrium (assuming selfishness and common knowledge of rationality) is played. By nature, costly punishment makes everybody worse off if used, whereas costly reward has the potential of increasing overall efficiency (if $|L| > 1$). Since all real-world sanction and reward mechanisms have this feature, it seems natural to preserve it when studying the choice between reward and punishment in the laboratory.

[7] Though one can find a wide range of technologies for the *punishment* option (ranging from a 1:1 ratio in Sefton et al., 2007, to a 1:5 ratio in Andreoni et al., 2003), the comparative statics of different technologies on contribution levels have only recently been examined systematically. See Anderson and Putterman (2006), Casari and Luini (2005), Egas and Riedl (2008), Masclet and Villeval (2008), Nikiforakis and Normann (2008) or the survey in Casari (2005). The only paper that we are aware of which is attempting to analyze different *reward* technologies in a very similar setting to ours is Fehr and Gächter (2003).

The first two factors of our design have been explained in the previous section. The way in which the institution is determined (factor F3) requires some additional explanation. In the exogenous treatments, the institutions were simply pre-determined by the experimenter. Participants in these sessions were informed in the experimental instructions about the institution under which they were playing the public goods game.

In the endogenous treatments, participants were introduced to all three possible institutions and could determine the institution prevailing in their group by the following procedure: First, subjects had to decide whether they wanted to participate in a costly vote on the institution. Voters incurred a small *one-time* fee of 10 ECU (equal to 5% of the sum of the individual endowment over the ten periods). Non-voters did not incur any costs, but were aware that the decision of voters would be binding for them. The voting cost allows us to get rid of theoretical cases of indifference regarding voter turnout and to analyze connections between voter turnout and behavior in the public good game. Then, the following approval voting protocol with simultaneous choices of all voters was applied: Each voter had to indicate for each available institution (*standard-VCM*, *punishment*, or *reward*) whether she was willing to accept this institution or not. It is important to note here that the endogenous choice was always restricted to a given leverage $|L|$. Hence, in one endogenous experimental treatment subjects could choose between the standard VCM, punishment with $|L| = 1$, or reward with $|L| = 1$. In the other endogenous experimental treatment, the leverage was always $|L| = 3$.

The voting outcome was determined as follows: If unanimous consent for any of the three institutions was not reached in the first ballot, the vote was repeated until unanimity (for at least one institution) was achieved.[8] In the case of a unanimous consent of all voters for exactly one institution, this institution was applied in the public goods game for all ten periods. If more than one institution had received unanimous support, then the institution was determined by a random draw from the tied institutions.[9] Finally, if none of the four

[8] There was no upper limit on the number of voting rounds. In the results section we will show that most groups reached a consensus rather soon. Even the groups with the highest number of voting rounds before reaching an agreement finished the voting procedure in less than ten minutes (because the software yielded instantaneous feedback after the last voter in a group had confirmed her voting decisions). Remember that the costs of participating in the vote were one-time and, thus, independent of the number of voting rounds.

[9] We do not argue that our voting mechanism is actually applied to real-world institutions for social dilemmas. We are confident, however, that its structure comes close to more or less tacit agreements on

group members had participated in the vote, then the institution was determined randomly in order to avoid creating a bias by imposing a default institution.

Voters were always informed about the total number of voters in their group, and after each voting round they were also informed about the number of votes each institution had received. Voting decisions were secret in the sense that they could not be linked to individuals or individual behavior in the consecutive public goods game.

One noteworthy feature of our design is that there was not a history of the VCM before participants decided on whether to vote and what to vote for. While one might claim that this is a disadvantage of our design – because subjects had to make an institutional choice before having experienced a particular institution – we believe it is rather a strength, given that we are mainly interested in the kind of rules that subjects prefer when facing a social dilemma situation. As such, our focus is on the initial choice of rules, and not on how the choice of rules depends on previous experience with different rules.[10]

In the punishment or reward stage that followed the contribution stage, subjects were informed about the contributions of all group members, and could decide on which members to punish or reward (depending on the prevailing institution). It was possible to punish or reward any other member and also more than one group member. At the end of each period subjects were again informed about the other members' contributions, and about the decisions on punishment and reward (if applicable) and the resulting payoffs. Each member of a group had a unique ID. Contributions to the public good could be assigned to this ID throughout the experiment. Decisions on reward and punishment, however, could not be linked to the ID. Hence, subjects could not identify which other group member had punished or rewarded them. This was done in order to avoid retaliation

---

the implementation of sanctions and rewards in many small groups in real life and also to the decision making in institutional arrangements in larger international organizations, like in the context of the Kyoto protocol or World Trade Organization negotiation rounds (where unanimity is often required to implement an institution).

[10] In order to check whether more experience with the game could lead to different results regarding the endogenous choice, we asked subjects in a post-experimental questionnaire whether they would stick to their institutional decision if the experiment had continued or whether they would now support another institution. The results from this questionnaire reinforce our conclusions based on initial choices and are presented in detail in Section 6.2.1.

(of one member punishing another for having been punished by the other member in the past).[11]

In the exogenous treatments we had in total 200 participants, yielding ten independent observations for each of the five different exogenous treatments (i.e. VCM, punishment with $L = -1$ or $L = -3$, and reward with $L = 1$ or $L = 3$). In the endogenous treatment we let 160 participants choose their desired institution when the leverage of reward or punishment was fixed at $|L| = 1$. Hence, we have 40 independent observations for the institutional choice with the low leverage. In addition, we had another 80 participants (i.e. 20 independent groups) choosing the institution under the high leverage $|L| = 3$.

In sum, 440 students drawn from the general undergraduate population at the University of Innsbruck participated in the computerized experiment (using z-Tree by Fischbacher, 2007). No subject was allowed to participate in more than one session and none of them had participated in a public goods experiment before.[12] Decisions were taken anonymously in cubicles and communication among participants was not permitted. The average session length was less than one hour, and subjects earned about 10 € on average.

5 Theoretical predictions

This section presents theoretical predictions for three different models of subjects' preferences. The first model is the standard *homo oeconomicus*-model, which assumes that decision makers are rational and purely selfish. The two other models belong to the group of outcome-based social preference models (at least in the way in which they are implemented below). Although there are many other models of social preferences (among them also intention-based models), the two models considered here are among the most prominent ones, they share some desirable properties for games where punishment or reward are involved, and they are tractable for deriving clear predictions, in particular for the case of our vote on institutions. It is important to note that our experiment is by no

[11] Such "counter-punishment" is documented and analyzed in Denant-Boemont et al. (2005) and Nikiforakis (2008), for instance.

[12] The experimental instructions for endogenous choice under $|L| = 3$ can be found in Appendix A. The instructions for the other treatments are analogous and available upon request. All instructions were context-free and neutrally framed.

means designed to deliver a rigorous test of the empirical validity of the models. Their application to our setting simply allows us to organize our theoretical predictions in a systematic way.

First, the model of inequity aversion by Fehr and Schmidt (1999) assumes that subjects compare their own material payoff with the material payoff of each other member in their reference group.[13] In particular, the utility of subject i in a set of n subjects depends on the vector of monetary payoffs $\pi = (\pi_1, \ldots, \pi_n)$ in the following form:

$$U_i(\pi) = \pi_i - \alpha_i \frac{1}{n-1} \sum_{j \neq i} \max\{\pi_j - \pi_i, 0\} - \beta_i \frac{1}{n-1} \sum_{j \neq i} \max\{\pi_i - \pi_j, 0\} \qquad (3)$$

with $\beta_i \leq \alpha_i$ and $0 \leq \beta_i < 1$. The second term on the right-hand side of equation (3) captures the utility losses from disadvantageous inequality towards other group members (with weight α_i), and the third term the losses from advantageous inequality (with weight β_i). A desire to reduce inequality among group members may be one of the motivations for using instruments like punishment or reward. Therefore, we consider the model of Fehr and Schmidt (1999) as a straightforward choice to derive predictions for our setting.[14] Note that the model of Fehr and Schmidt (1999) nests the *homo oeconomicus*-model of pure self-interest when it is assumed that $\alpha_i = \beta_i = 0$.

The model by Charness and Rabin (2002) assumes that subjects care about social welfare. Their interpretation of social welfare includes the minimum payoff in the group and the sum of all group members' payoffs. More specifically, in their general model Charness and Rabin (2002, see their Appendix 1) define a subject's preferences in a set of n subjects with $\pi = (\pi_1, \ldots, \pi_n)$ as the vector of monetary payoffs in the following form:

$$U_i(\pi) = (1-\lambda)\pi_i + \lambda\left[\delta \min(\pi_1, \ldots, \pi_n) + (1-\delta)(\pi_1 + \pi_2 + \ldots + \pi_n)\right] \qquad (4)$$

The weight a subject assigns to social welfare relative to the own payoff is measured by $\lambda \in [0,1]$. Note that setting $\lambda = 0$ captures the case of pure self-interest as in the *homo oeconomicus*-model. The weight $\delta \in [0,1]$ is attached to the "maximin"-aspect of social

[13] A model of social preferences where subjects compare their own payoff to the *average* payoff in their group (as in another influential model by Bolton and Ockenfels, 2000, for instance) seems less straightforward to be used in our context. This has been shown by Falk et al. (2005) in a similar setting.

[14] In fact, Fehr and Schmidt (1999) have applied their model to a public goods game with punishment. They have not used it for rewards, though, nor for a setting with endogenous choice between reward and punishment.

welfare, and the weight $1-\delta$ applies to total-surplus maximization. Since social welfare aspects seem particularly relevant for the reward institution with $L = 3$ we regard the model by Charness and Rabin (2002) as a valuable complement to the Fehr and Schmidt (1999)-model for deriving theoretical predictions. Note that we do not consider the version of the Charness and Rabin (2002)-model that takes intentions into account, because it does not seem suitable to derive clear theoretical predictions from it for the institutional vote in our setup.

The main predictions from the three models are summarized in Table 2. They rely on the assumptions of common knowledge of rationality and risk neutrality. The finitely repeated game is solved through backward-induction. For the sake of succinctness, we discuss in this section only the main intuition for the theoretical results of Table 2. The interested reader is referred to Appendix B for formal derivations and proofs.

Table 2 about here

5.1 Standard preferences

Applying standard preferences yields straightforward predictions (see the first column in Table 2). A rational money-maximizing agent will abstain from costly punishment or reward. Hence, contributions do not depend on the availability of punishment or reward, irrespective of their leverage. Since free-riding is a dominant strategy due to $\gamma < 1$, all treatments will lead to zero contributions. As a consequence, subjects should be indifferent between the available institutions in the endogenous treatments and, thus, abstain from voting in order to save the voting costs.

5.2 Inequity aversion – Fehr and Schmidt (1999)-preferences

Incorporating inequity aversion in the utility function leads to equilibria with positive contributions in the standard VCM, as Fehr and Schmidt (1999, p. 839 and pp. 860-863) have shown. The main intuition behind their result is that if a subject's aversion against *advantageous* inequity is strong enough, she may refrain from free-riding. In formal terms, $\gamma + \beta_i \geq 1$ is a precondition for positive contributions in equilibrium. However, as long as a group of four members (as is the case in our experiment) has at least one single free-rider with $\gamma + \beta_i < 1$, positive contributions cannot be sustained in equilibrium (see [A] in the second column of Table 2). Only if all group members satisfy $\gamma + \beta_i \geq 1$, any positive contribution from the interval [0, 20] constitutes an equilibrium (see [B] in the second column of Table 2).[15]

The same equilibria [A] and [B] also prevail with punishment or reward under the low leverage $|L| = 1$, and even with reward under $L = 3$. The low-leverage punishment does not constitute a credible threat to induce free-riders (who satisfy $\gamma + \beta_i < 1$) to contribute positive amounts. Similarly, the incentive to receive rewards is not sufficient to make free-riders contribute under both possible levels of leverage in our experiment. Hence, if there is at least one free-rider in the group, cooperative equilibria do not exist in our setup for these three conditions.

Only with a high-leverage punishment option it is possible for cooperators (who satisfy $\gamma + \beta_i \geq 1$) to enforce positive contributions also from free-riders. Let n' denote the number of cooperators in a group, then $\bar{c} = n' L / (1 - \gamma) > 0$. This result is driven by the fact that free-riders are, then, better off contributing and thus avoiding punishment instead of free-riding and getting punished. Punishment is a credible threat, because cooperators are able to reduce the disadvantageous inequality caused by actual free-riding through the efficient punishment option. Since cooperators suffer from inequality they would be willing to bear the cost of punishment to reduce it. In equilibrium, group members contribute and punishment is never used.

Concerning the incentives to participate in the vote on the institution in the endogenous treatments, it is important to note that the predictions on contributions are the same for the standard VCM and punishment or reward with the low leverage $|L| = 1$. Hence, subjects have no incentive to vote under $|L| = 1$, as the institution does not make a difference (see panel (B) in Table 2).

With the high-leverage and a punishment option, it is possible to induce the free-riders to contribute a positive amount up to $\bar{c}$ (the latter threshold depending on the

[15] This latter case is empirically rather unlikely, though. If we take, for the purposes of illustration, the assumptions about the distribution of social preference parameters as reported in Fehr and Schmidt (1999, p. 844), there will be 60% of subjects with $\beta_i < 0.6$. The latter implies that only in about 2.6% of all cases *all* group members will satisfy $\beta_i \geq 0.6$ (if $n = 4$) and that in the remaining 97.4% of cases the only equilibrium will be zero contributions (because at least one member satisfies $\beta_i < 0.6$).

number of cooperators in the group). Hence, cooperators should vote for punishment to be able to enforce positive contributions from free-riders.

5.3 Social welfare – Charness and Rabin (2002)-preferences

If subjects care for social welfare, they may contribute positive amounts in the standard VCM. The Charness and Rabin (2002)-model would predict that subjects who weigh social welfare at least as strongly as their own payoffs (implying $\lambda_i \geq 0.5$) will contribute their full endowment (see the right-most column in Table 2). For all subjects with $\lambda_i < 0.5$, complete free-riding is a dominant strategy. Hence, there are asymmetric equilibria in the standard VCM when subjects have Charness and Rabin (2002)-preferences.[16]

Adding a punishment option does not change any of these predictions. The intuition for the irrelevance of the punishment option is straightforward: Relatively selfish group members (those with $\lambda_i < 0.5$) would contribute positive amounts if the threat of getting punished for zero contributions was credible. However, a punishing group member will reduce her own payoff and the group's total payoff. Furthermore, punishment will never increase the minimum payoff in the group but rather decrease it in many cases. Taken together, a subject suffers a utility loss from punishment in any case (disregarding possible intentional reasons to punish), and will thus not execute it. Consequently, punishment cannot induce any contributions in excess to those under the VCM.

However, the reward option is a useful tool to induce positive contributions with Charness and Rabin (2002)-preferences. This is straightforward for the case of L = 3, where subjects who care much for social welfare (who satisfy $\lambda_i \geq 0.5$ and thus contribute c_i = 20; their number is denoted by n'') can enforce a positive contribution of $\bar{c} = n''L/(1-\gamma)$ even from subjects who care less for social welfare (with $\lambda_i < 0.5$). This enforcement rests on rewarding the latter subjects for contributing $\bar{c}$. For the case of L = 1, the same argument applies, but reward can only be part of an equilibrium strategy (and thus $\bar{c}$ can only be enforced) if there are n'' subjects who care quite a lot for social welfare and largely disregard their own payoff (by satisfying $\lambda_i \geq 0.75$).

[16] Note that if all group members satisfy $\lambda_i < 0.5$, all group members (symmetrically) free-ride.

These predictions for contributions imply that subjects who care a lot for social welfare should participate in the vote in our endogenous treatments and support the reward option in order to be able to enforce positive contributions from subjects who care little for social welfare (and who do not participate in the vote). Note that the latter group of subjects does only contribute in the VCM with reward, but not in the standard VCM or the VCM with punishment.[17]

6 Experimental results

6.1 Behavior with exogenously determined institutions

Figure 1 displays the time trend of contributions over the ten periods in the different treatments with exogenous institutions, and panel (A) of Table 3 reports the overall averages. Contributions are lowest in the standard VCM-treatment, where subjects contribute on average 32.5% of their endowment (= 6.50 out of 20 ECU). This is clearly above the zero contribution level expected under the standard approach of common knowledge of rationality and selfishness, but it is not ruled out when taking into account social preferences. With the low leverage $|L| = 1$, the average contributions are only slightly higher with punishment (8.79) or with reward (6.62) than in the standard VCM. None of the non-parametric pairwise comparisons of contributions yields a significant difference, though. This is in accordance with the predictions based on Fehr and Schmidt (1999)-preferences, and it is not ruled out using the Charness and Rabin (2002)-model either, as long as no group member puts a very strong weight on social welfare by satisfying $\lambda_i \geq 0.75$. From panel (B) of Table 3 it is also easy to see that profits are not significantly different between the standard VCM and the VCM with punishment or reward under $|L| = 1$.

Result 1. *When the institution is fixed exogenously and when* $|L| = 1$*, the institution does not influence the level of contributions. Hence, reward and punishment options do not lead to higher contributions than the standard VCM when their effectiveness is low.*

Figure 1 and Table 3 about here

[17] This conclusion relies on the assumption that the voting costs are small. See Appendix B for more details.

When we look at contribution levels under the high leverage $|L| = 3$, we find significantly higher contributions with punishment (12.11) or reward (11.20) than in the standard VCM (6.50) ($p < 0.05$ for both pairwise comparisons; two-sided Mann-Whitney U-test[18]; $N = 20$ in each case). We find no significant difference in contribution levels between the punishment and reward treatments, though.

When we examine the impact of the leverage for a given institution, i.e., the effectiveness of the mechanisms, we find significantly higher contributions under the high leverage ($|L| = 3$) than under the low leverage ($|L| = 1$) both for punishment (12.11 vs. 8.79; $p < 0.05$; Mann-Whitney U-test, $N = 20$) and for reward (11.20 vs. 6.62; $p < 0.05$; Mann-Whitney U-test, $N = 20$). This has been expected for punishment with Fehr and Schmidt (1999)-preferences and for reward with Charness and Rabin (2002)-preferences.

Result 2. *When the institution is fixed exogenously and when punishment or reward are highly effective ($|L| = 3$), both institutions trigger significantly higher contribution levels than the standard linear VCM or the punishment and reward institutions under $|L| = 1$.*

Checking profits and thus the important issue of the efficiency of different institutions under $|L| = 3$, panel (B) of Table 3 reveals that average profits are significantly higher under reward (29.37) than under both punishment (24.83) and the standard VCM (23.90). The underlying reason for this is that reward with high effectiveness is efficiency enhancing, whereas actual punishment has rather high social costs (of 4 ECU per instance of punishment). The latter fact explains why profits with punishment under $|L| = 3$ are not higher than in the standard VCM even though average contributions are almost twice as high.

Turning to the frequency of using the punishment and reward options in the exogenous treatments, we can provide evidence that both are used more often in the high leverage condition $|L| = 3$ than in the low leverage condition, but only significantly so with reward. Rewards are assigned in 44% of all possible cases under $|L| = 3$ and only in 10% of cases under $|L| = 1$ ($p < 0.05$; Mann-Whitney U-test, $N = 20$). This is a clear indication that subjects exploit the potential efficiency gains under $|L| = 3$ to a large extent. The

[18] All non-parametric tests reported in this paper are two-sided, and the overall average contribution in each group is always treated as a single, statistically independent observation.

leverage has, however, no significant effect on the frequency of punishment, which is used in 20% of all cases under $|L| = 3$ and in 15% under $|L| = 1$.

Result 3. *Regarding the actual frequency of using punishment or reward, subjects are much more responsive to a higher leverage in the reward institution than in the punishment institution. Under* $|L| = 3$ *rewarding is significantly more frequent than under* $|L| = 1$.

Figures 2 and 3 about here

Figures 2 and 3 provide a more detailed picture of the patterns of reward and punishment, depending upon the leverage. They show the relative frequency with which a subject *j* is rewarded or punished by subject *i*, contingent on the absolute difference between subject *i*'s and subject *j*'s contribution. Concerning the frequency of punishment, we can discern from the left-most bar in Figure 2, for instance, that in 46% of the cases in which member *j* contributes 14 to 20 tokens less than member *i* punishment of member *j* by member *i* is observed. In general, there is a decrease in the relative frequency of punishment when moving from left to right in Figure 2. This means that members *j* are more frequently punished the less they contribute in comparison to member *i.* Like Fehr and Gächter (2000), we also observe a positive, but relatively small number of low contributors that punish high contributors.[19]

The picture is different for rewards (in Figure 3), where one is generally more likely to be rewarded the higher one's contribution in comparison to the contribution of the rewarding member. Under the high leverage, though, it seems that the subjects with high contributions reward each other, which leads to a peak in the central deviation range of [-2, 2] between member *j*'s and member *i*'s contributions. Finally, note that under the high leverage the probability of getting rewarded is generally very high if one contributes at least as much as the potentially rewarding subjects.

[19] Herrmann et al. (2008) use the term antisocial punishment for such cases, and they demonstrate that antisocial punishment is widespread around the world, though it is contained by strong norms of civic cooperation and positive attitudes towards the rule of law.

Result 4. *The likelihood of getting punished is inversely related to a subject's contribution. Reward is most often granted for subjects that contribute at least as much as or more than the rewarding subject.*

6.2 Behavior with endogenously determined institutions

6.2.1 Participation in the vote and institutional choice

Seventy one out of 160 subjects (44%) participate in the voting procedure in the low-leverage condition. In the high-leverage condition, 48 out of 80 subjects (60%) take part in the costly vote. In fact, voter turnout is significantly higher in the high-leverage condition ($\chi^2 = 5.21$, $p < 0.05$), as qualitatively expected if subjects have social preferences of either the Fehr and Schmidt (1999)- or the Charness and Rabin (2002)-type. Obviously, the standard prediction of abstention from the vote fails.

Result 5. *Voter turnout is significantly higher when reward and punishment are more effective (|L| = 3) than when they are less effective (|L| = 1).*

In panels (A) and (B) of Table 4 we report the number of voting rounds until a unanimous agreement was reached, contingent upon the number of group members that participated in the voting procedure. Of course, all groups with only one voter finish the procedure after the first round because the single voter can implement his or her most desired institution. Across both leverage levels, 46 out of 57 groups reach an agreement within the first four voting rounds. In the low-leverage condition, the maximum number of voting rounds is 26, whereas in the high-leverage condition three groups need more than 50 voting rounds to reach an agreement, with a maximum of 103 voting rounds in one group (which finally settled – in less than 10 minutes – for the standard VCM).

Table 4 about here

Panel (C) of Table 4 reports the preferences of voters. More precisely, it shows the number of times each possible institution (VCM, punishment or reward) is supported under both leverages in the very first voting round. Remember that voters could support more than one institution. The standard VCM is most often supported when $|L| = 1$, whereas the VCM with reward is most preferred when $|L| = 3$, and the distribution of preferences is significantly different between both leverages ($\chi^2 = 13.09$, $p < 0.01$). The rightmost column shows the frequency with which a voter's preference coincided with the actually chosen institution. We see that in total 97 out of 119 voters (i.e. 82%) experienced the very institution that they had supported in the first round of voting.

Table 1 reveals the chosen institutions in the two different leverage conditions. With the low-leverage ($|L| = 1$), 25 out of 40 groups choose the standard VCM, five groups opt for the VCM with punishment and ten groups agree on the VCM with reward. The distribution of chosen institutions is significantly different from a random one and also markedly different from the distribution in the high-leverage condition ($\chi^2 = 19.61$, $p < 0.01$). There were three groups in which none of the four members voted. Coincidentally, the random determination of the institution yielded one group each for the standard VCM, the punishment institution and the reward institution.

With the high leverage ($|L| = 3$), 17 out of 20 groups choose the VCM with reward and the remaining three groups the standard VCM. No group opts for the VCM with punishment in the high-leverage condition. The distribution of chosen institutions with the high leverage is completely at odds with the predictions from the Fehr and Schmidt (1999)-model, but fits the predictions from the Charness and Rabin (2002)-model very well. Subjects obviously aim for the option to reward each other mutually when the reward option increases overall efficiency. Note that the equilibrium in which each group member rewards every other group member is actually quite frequent with 21% of all cases (36 out of 170). Not surprisingly, there is not a single such case of mutual rewarding under the endogenously chosen reward with $|L| = 1$.

***Result 6**. The distribution of institutional choices depends significantly on the leverage of reward and punishment. Given a high leverage, 85% of groups endogenously choose the VCM with reward, whereas with a low leverage the majority of groups (63%) agree on the standard VCM.*

In order to check the robustness of our results regarding the endogenous institutional choice and subjects' satisfaction with their chosen roles, we asked them at the end of the experiment whether they would vote again and, if so, for which institution they would vote if the experiment were to be continued. The data from this questionnaire reveal that 72% of voters (86 out of 119) would vote again for exactly the same institution. Only 13% of voters would abstain from voting, and 15% would switch to a different institution.[20]

The answers of non-voters are markedly different, because 74 out of 121 non-voters (61%) indicated that they would participate in the vote if the experiment were repeated. Most of them would have voted for punishment under $|L| = 1$ (27 out of 47) and for reward under $|L| = 3$ (22 out of 27). This clearly indicates that our results on the endogenous institutional choice would have been very similar and our conclusions corroborated if it had taken place in a later stage after several periods of experience with the VCM.

6.2.2 Contributions in the endogenous treatments

Figure 4 shows the time trend of contributions over the ten periods in the different treatments with endogenous institutional choice, and panel (A) of Table 5 reports the overall averages. In the low-leverage condition ($|L| = 1$), contributions are lowest in the standard VCM (5.48), intermediate in the VCM with reward (8.53) and clearly highest in the VCM with punishment (16.13). Each pairwise comparison yields significant differences ($p < 0.05$ in each case; Mann-Whitney U-tests). This means that with an endogenous institutional choice, the possibility of rewarding and punishing increases the level of cooperation significantly in comparison to the standard VCM, even though the leverage of reward and punishment is low. Punishment works best here, as it is associated with average contributions that are almost twice as high as with reward.

In the high-leverage condition, no group agreed on the punishment option. Groups with the reward institution have significantly higher contribution levels (14.59) than groups in the standard VCM (9.17) ($p < 0.05$; Mann-Whitney U-test, $N = 20$) or groups with reward under $|L| = 1$ (8.53; $p < 0.05$, $N = 27$).

[20] Half of those subjects who would abstain from voting had voted for the standard VCM in the beginning. The rest of those indicating to abstain were almost equally split between voters who had in the beginning preferred either punishment or reward. Regarding voters who would switch to another institution, there is no clear pattern. Seven out of ten subjects who had voted for the standard VCM in the beginning would switch to reward, the other three to punishment. Five out of seven subjects who had preferred reward initially would vote for punishment, the two other for the standard VCM. Only one subject who had voted for punishment would vote for a different institution, namely reward.

Result 7. *When the institution is determined endogenously, we find for both leverages that reward leads to higher contributions than the standard VCM. The same holds true for punishment with the low leverage. Since there is not a single group that chooses the punishment option under the high leverage, we cannot assess its effect on contributions there.*

Figure 4 and Table 5 about here

Average profits per group member and period are shown in panel (B) of Table 5. They are always significantly lower in the standard VCM than with punishment or reward. Under the low leverage, average profits are about 15% higher with punishment than with reward ($p < 0.05$; Mann-Whitney U-test; $N = 15$).

Figures 5 and 6 show the relative frequency of punishment and reward, depending on the difference in contributions of group members i and j. Since the patterns are very similar to the ones observed in Figures 2 and 3, we refer to the discussion of these figures in Section 6.1. Overall, the reward option is again more often used under $|L| = 3$ than under $|L| = 1$ (55% vs. 17%; $p < 0.05$; Mann-Whitney U-test, $N = 27$). In the punishment treatment under the low leverage, subjects use the punishment option only in 10% of all cases. Nevertheless, they can sustain the highest contribution levels.

Figures 5 and 6 about here

6.2.3 How voting relates to contributions and the use of sanctions

We have already shown that 44% (under $|L| =1$), respectively 60% (under $|L| =3$), of subjects participated in the vote on the institution. In this subsection, we analyze whether the decision to participate in the vote is systematically related to a subject's contributions in the public goods game. One intuitive conjecture would be that subjects who participate in the vote are those who care relatively more for the common interest of the group because they try to implement an institution that they consider to be the best. In fact, using the Charness and Rabin (2002)-model we have been able to show that subjects who care relatively much for social welfare are expected to participate in the vote (see Table 2 and Appendix B). The same subjects are also expected to contribute more to the public good.

Hence, one should expect higher contributions from voters than from non-voters. Table 6 reports the results of a censored tobit regression where we regressed a member's contribution on her decision to participate in the vote (1 = voter), on the selected institution (separate dummies for punishment and reward), on the leverage (1 if $|L| = 3$), and on the number of voters in a group.

Table 6 about here

Column [1] of Table 6 presents the marginal effects of the independent variables and considers only the contributions in the very first period as the dependent variable. It shows that the dummy for voters is significantly positive, despite controlling for other institutional variables within a group. Voters contribute about 4 ECU (or 20% of their endowment) more than non-voters. Hence, participation in the vote is a significant indicator of a subject's initial level of cooperation.[21]

Interestingly enough, this voter-effect disappears very quickly, as it is obviously overpowered by the evolution of decisions within a group. The dummy for voters already ceases to be significant in period 2 (see column [2] of Table 6). The dummy also remains insignificant if we take a subject's overall average contributions over all ten periods as the dependent variable (see column [3]). It is comforting to note, though, that the other variables of interest (punishment, reward and leverage) always remain significantly positive in columns [2] and [3].[22]

Table 7 about here

In Table 7 we report a panel probit regression of a member's decision to punish or reward on the following variables: member *i* being a voter, the number of voters in the

[21] It is important to note that the documented "voter"-effect does not depend on whether a voter faced her preferred institution (i.e. the one that a voter supported in the first voting round) or not. Including a dummy for being "overruled" (meaning that a voter finally accepted an institution that she did not support in the beginning) does not change the magnitude and significance of the "voter"-dummy. The "overruled"-dummy itself is not significant.

[22] A Poission regression model that accounts for the integer values of the dependent variable yields basically the same results as those presented in Table 6.

group, member *i*'s contribution to the public good, and the difference between member *i*'s and member *j*'s contributions in a given period. We find that voters are much more likely to reward other group members than non-voters, both under $|L| = 1$ and $|L| = 3$. Voters do not have a higher probability than non-voters to punish other group members, though. The average contribution within a group has a significantly negative impact on the likelihood of punishment, and a significantly positive influence on the likelihood of reward. Furthermore, the likelihood of member *i* to punish member *j* is significantly decreasing with the difference between member *j*'s and member *i*'s contribution, and the likelihood of rewarding is increasing with this difference.

Result 8. *Subjects who participate in the vote on the institution contribute significantly more to the public good than non-voters in the beginning of the experiment. From period 2 on, this effect disappears due to interaction within a group. Voters reward others more often than non-voters, but they do not punish more often than non-voters.*

6.3 The effect of an endogenous institutional choice on contributions

A comparison of the corresponding cells in Table 3 and Table 5 shows immediately that average contributions are always higher in the endogenous treatments than in the corresponding exogenous treatments in case a reward or punishment option exists. The difference is most remarkable and significant for the VCM with punishment under $|L| = 1$, where contributions are on average 16.13 in the endogenous case and 8.79 in the exogenous case ($p < 0.01$; Mann-Whitney U-test; $N = 15$). In the VCM with reward, the differences are weakly significant (6.62 vs. 8.53 under $|L| = 1$; $p = 0.10$, $N = 20$; and 11.20 vs. 14.59 under $|L| = 3$; $p = 0.09$, $N = 27$), lending support to the conclusion that an endogenous institutional choice is a device that increases the level of cooperation in comparison to exogenously imposed institutions. This is summarized in our final result.

Result 9. *Contributions are higher when reward and punishment are endogenously chosen than when these institutions are implemented by an external authority.*

Result 9 seems all the more remarkable as it is independent of the number of voters in a group. In fact, there is no significant correlation between the average contribution and the number of voters in a group, as the partial correlation coefficient – controlling for leverage and institution – is only 0.06 ($p > 0.3$). There is also no significant relationship between the number of voting rounds before implementing an institution and the average contribution in a group. Furthermore, the relative frequencies of reward, or punishment, are not significantly different between the exogenous and the endogenous treatments, given the low leverage $|L| = 1$. Only for the high leverage we find a higher relative frequency of reward in the endogenous treatment than in the exogenous one (55% vs. 44%; $p < 0.01$; Mann-Whitney U-test; $N = 27$). In sum, we regard the evidence as an indication that it is the endogenous choice of the reward or punishment option *per se* that positively affects the contribution levels.

7 Discussion and conclusion

Previous studies have been able to demonstrate that an option to punish defectors and reward cooperators in public goods games has a positive influence on the level of cooperation. Since the reward or punishment options have typically been implemented exogenously by the experimenters, it has remained an open question as to whether subjects would actually choose such institutions if given the choice and whether the opportunity to choose the desired institution endogenously would have any additional impact on the level of cooperation.

In this paper, we have shown that there is a positive effect of an endogenous choice of a reward or punishment mechanism on the level of cooperation within groups. One may call this effect a democratic participation rights-premium to cooperation. This premium adds to the cooperation-increasing effect of reward or punishment when they are imposed exogenously. We consider the demonstration of the existence of this premium the first important finding of our paper. Dal Bó et al. (2007) provide evidence that reinforces our conclusion regarding the existence of the democratic participation rights-premium. They use a prisoner's dilemma game and show that an endogenous choice of a policy governing interaction increases cooperation.

One possible explanation for the higher levels of cooperation with endogenous institutional choice in our experiment might be due to voters *intentionally* choosing an institution. This might be perceived – by non-voters as well as other voters – as a signal that these voters actually plan to use the instruments at hand. Consequently, in the case of the VCM with punishment, the collective choice might pose a more credible threat to free-riders than the exogenous implementation. Furthermore, the threat is already salient in the first period to a greater extent, whereas with exogenous institutions it might take a few periods to deter free-riding by building up a punishment reputation. Hence, selfish subjects might be more reluctant to free-ride in the endogenous treatment. In the case of the VCM with reward, the signal from the vote might induce subjects to expect rewards in case they contribute relatively high amounts. In fact, the reward option is used significantly more often in the endogenous treatment than in the exogenous treatment when the leverage is high. In a similar vein, the probability of getting punished when contributing significantly less than the potential punisher is much higher in the endogenous treatment than in the exogenous treatment (which becomes immediately apparent when comparing Figures 2 and 5).

Therefore, from a behavioral perspective the endogenous choice of an institution seems to contain more than just the determination of a given set of rules. It can be considered as establishing a social norm on how to behave and treat each other. It implies the signal of actively supporting – and probably applying – an available instrument. Given that the endogenous choice of the standard VCM does not provide group members with any instrument to express their (dis)approval with the other members' contributions, it is also not surprising that contributions in the standard VCM do not differ between the exogenous and the endogenous case.

One straightforward implication of our findings concerning the positive effects of an endogenous choice is that implementing an endogenous institutional choice may provide a welfare-enhancing substitute for increasing the leverage of punishment through an exogenous authority. Our results in the exogenous treatments as well as recent studies by Anderson and Putterman (2006), Carpenter (2007), Casari and Luini (2005), Egas and Riedl (2008), Fehr and Gächter (2003) or Nikiforakis and Normann (2008), for instance, have shown that contributions to a public good typically increase when the effectiveness of punishment is increased *exogenously*. However, a high leverage of punishment usually implies high social costs because it potentially decreases overall efficiency. In our experiment we are able to provide evidence that instead of an increase in the effectiveness of punishment one might resort to an endogenous choice of the rules. Note that we find contributions in the *endogenous* treatment with punishment under $|L| = 1$ to be even significantly higher than contributions in the *exogenous* punishment treatment with $|L| = 3$ ($p < 0.05$; Mann-Whitney U-test; $N = 15$).

Interestingly enough, the higher *aggregate* cooperation levels in the endogenous treatments do *not* depend on the number of voters in a group (i.e. the voter turnout), but rather on the mere opportunity of an endogenous choice. On the individual level, however, we have found that the voluntary (but costly) participation in the vote is a good indicator of a subject's level of cooperation at the beginning of the experiment. Voters have been estimated to contribute in the first period about 4 ECU (i.e. 20% of their endowment) more to the public good than non-voters. This is a noteworthy effect because it implies that the decision to participate in a vote can be linked to economic behavior. Voters, however, do not contribute significantly more than non-voters from the second period on, which is a possible indication that voters are conditional cooperators (Fischbacher et al., 2001; Fischbacher and Gächter, 2006; Gächter, 2007) who react to the contributions of non-voters. Consequently, the behavior of voters and non-voters converges rather quickly after both types of subjects have interacted with each other, even though subjects could not link other group members' contributions to their voting decision. The convergence of voters' and non-voters' contributions implies that the prevailing institutions overpower the vote-participation effect on the aggregate level.

Another important finding of our paper is that the choice of institutions seems to depend heavily on their effectiveness. When reward and punishment have a high leverage, groups choose almost exclusively the reward option. A straightforward explanation for this choice behavior might be the intention to exploit the potentially large efficiency gains through mutual rewards. Another explanation for groups shying away from the punishment institution might be the fear of high efficiency costs if punishment is actually applied under $|L| = 3$. Furthermore, subjects could simply be myopic in the sense of overestimating the costs of being punished that usually accrue in the first periods of a public goods game and underestimating the potentially positive effects of the threat of punishment in raising contribution levels in the long run.

When the leverage is low, however, a clear majority of groups prefer the standard VCM over the reward or the punishment mechanism. The main reason for preferring the standard VCM might be the expectation of subjects that reward or punishment with $|L| = 1$ may have little or no effect at all, as is actually true when reward and punishment with the low leverage are implemented *exogenously*. If the institution is chosen *endogenously*, though, reward or punishment increase the contribution levels even with a low leverage. In fact, total profits have been significantly higher both with punishment and with reward than with the standard VCM (see Panel (B) of Table 5). Hence, from an ex-post perspective it would have paid off, on average, to vote for punishment or reward even with the low leverage.

Our results on the relative attractiveness of reward in comparison to punishment complement the recent literature on endogenous institutional choice in public goods games. Neither Botelho et al. (2005) and Ertan et al. (2005) nor Gürerk et al. (2006) have been able to address this issue in full depth because Botelho et al. (2005) and Ertan et al. (2005) only considered punishment and Gürerk et al. (2006) did not allow for a choice between reward and punishment. Our results on the relative *un*attractiveness of punishment are basically in line with the findings of Botelho et al. (2005), and they seem to be compatible with Ertan et al.'s (2005) finding that groups never allow punishment of high contributors. Botelho et al. (2005) argue that subjects opted for the standard VCM because subjects had experienced it as more efficient than the punishment institution. In our experiment, groups favored the reward institution over punishment possibly because of the higher efficiency of reward. Both findings should be considered complementary, indicating that subjects do not find the implementation of a punishment option attractive *if* they have an alternative. If they do not have any alternative like in the exogenous case, the punishment option is frequently used and has a positive effect on contributions (Fehr and Gächter, 2000, 2002), as long as the effectiveness of punishment is sufficiently high (Nikiforakis and Normann, 2008). Nevertheless, it is usually associated with ambiguous effects on overall efficiency. If subjects consider efficiency as important – as assumed in the model of Charness and Rabin (2002) and as documented in simple distribution experiments by Engelmann and Strobel (2004) – the relative attractiveness of the reward mechanism seems fairly intuitive.

It is not a contradiction to our results that punishment mechanisms seem to be much more widespread in reality than reward mechanisms. While there is a natural limit to the leverage of reward (though, we would argue that this limit is clearly above $L = 1$, at least in terms of potential utility created), the leverage of punishment can be almost infinite. Nevertheless, we believe that reward in small groups (e.g., in the form of social approval) is much more widespread in reality than economists tend to believe and it should probably also be applied more often in formal contexts (see, e.g., the literature on the possibility of rewarding honest tax payers; examples are Alm et al., 1992, or Feld et al., 2006).

We acknowledge, of course, that the implementation of the institutional choice in our experiment is slightly artificial, as compared to how informal sanctions are implemented in small groups in reality. Often, an implicit agreement rather than a formal vote is sufficient to determine whether members of a group tend to resort to individual punishment or individual reward (like in sports teams, for instance). Also, in many cases in reality both punishment and reward are feasible at the same time. However, as often in experiments, by abstracting a little bit when choosing our design, we believe to have gained considerably in terms of possible analytical depth, especially when it comes to be able to detect links between voting and contribution behavior. Obviously, implementing other experimental environments or even conducting field experiments in order to study the effects of endogenous institutions in social dilemmas might be interesting avenues of research for the future.

Regarding the theoretical predictions, it is not surprising that a model with purely selfish and rational agents is not consistent with our experimental results. The models with social preferences fit our data much better. Both the inequity aversion-model of Fehr and Schmidt (1999) and the social welfare-model of Charness and Rabin (2002) have their strengths and weaknesses, though, when it comes to predicting behavior in public goods games. The Fehr and Schmidt (1999)-model clearly predicts that punishment and reward should not matter for contributions when the leverage is low, and this is what we have found in the exogenous treatments. The model is also particularly suited to explain the effects of the punishment option under the high leverage, because highly efficient punishment can easily reduce inequity. The latter effect makes the threat of punishment credible and keeps contributions relatively high. However, the model largely fails to predict the observed behavior under the reward institution. This is much better captured by the model of Charness and Rabin (2002), who allow for efficiency-minded subjects. A striking feature of applying Charness and Rabin (2002) to our setup is the fact that it predicts that subjects who care more for social preferences will participate in the vote. First round data from our experiment have confirmed this prediction, thereby linking voting decisions to contributions. The Charness and Rabin (2002)-model in its outcome-based version is less suitable for predicting the effects of punishment, because subjects who care for social welfare (but not inequity) will never execute the punishment option. Both social preference models in the variants used in our paper are not too successful in explaining the effects of endogenous institutional choice. One straightforward explanation for the latter limitation is that they ignore intentions and the emotional as well as physiological consequences of the act of punishing or rewarding (de Quervain et al., 2004; Casari and Luini, 2006; Brandts and Rivas, 2007). In fact, a large fraction of subjects seem to use the punishment and reward options either to reciprocate kind or unkind decisions of others or to express their (dis)approval with the contributions of others, independently from the effects on the final distribution of profits. Such behavior can be called expressive punishment or expressive reward, and it cannot be captured directly by outcome-based models of social preferences, like those of Fehr and Schmidt (1999) and Charness and Rabin (2002). Applying intention-based models (for example in the spirit of Rabin, 1993; Charness and Rabin, 2002, in its most comprehensive version; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006) can possibly fill this gap in the future, even though these models can often not provide clear-cut predictions in more complex decision environments.

8 References

Abbink, K., Irlenbusch, B., Renner, E. (2000), The moonlighting game. An experimental study on reciprocity and retribution. *Journal of Economic Behavior and Organization* 42: 265-277.

Alm, J., McClelland, G. H., McKee, M. (1992), Deterrence and beyond: Toward a kinder, gentler IRS. In: Slemrod, J. (ed.), *Why People Pay Taxes*. Ann-Arbor, University of Michigan Press: 311-329.

Anderson, C. M., Putterman, L. (2006), Do non-strategic sanctions obey the law of demand? The demand for punishment in the voluntary contribution mechanism. *Games and Economic Behavior* 54: 1-24.

Andreoni, J., Harbaugh, W., Vesterlund, L. (2003), The carrot or the stick: Rewards, punishments and cooperation. *American Economic Review* 93: 893-902.

Bolton, G., Ockenfels, A. (2000), ERC: A theory of equity, reciprocity and competition. *American Economic Review* 90: 166-193.

Botelho, A., Harrison, G. W., Costa Pinto, L. M., Rutström, E. E. (2005), Social norms and social choice. University of Central Florida, Working Paper.

Brandts, J., Rivas, M. F. (2007), On punishment and well-being. UFAE and IAE Working Paper 705.07. University Autonoma Barcelona.

Carpenter, J. (2007), The demand for punishment. *Journal of Economic Behavior and Organization* 62: 522-542.

Casari, M. (2005), On the design of peer punishment experiments. *Experimental Economics* 8: 107-115.

Casari, M., Luini, L. (2005), Group cooperation under alternative peer punishment technologies: An experiment. University of Siena, Working Paper 2/2005.

Casari, M., Luini, L. (2006), Peer punishment in teams: Emotional or strategic choice? Purdue University, Working Paper.

Charness, G., Rabin, M. (2002), Understanding social preferences with simple tests. *Quarterly Journal of Economics* 117: 817-869.

Croson, R., Konow, J. (2008), Social preferences and moral biases. *Journal of Economic Behavior and Organization*, forthcoming.

Dal Bó, P., Foster, A., Putterman, L. (2007), Institutions and behavior: experimental evidence on the effects of democracy. Brown University, Working Paper.

Decker, T., Stiehler, A., Strobel, M. (2003), A comparison of punishment rules in repeated public good games. *Journal of Conflict Resolution* 47: 751-772.

Denant-Boemont, L., Masclet, D., Noussair, C. (2005), Anonymity in punishment, revenge and cooperation: A public good experiment. Emory University, Working Paper.

de Quervain, D., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., Fehr, E. (2004), The neural basis of altruistic punishment. *Science* 305 (27 August 2004): 1254-1258.

Dufwenberg, M., Kirchsteiger, G. (2004), A theory of sequential reciprocity. *Games and Economic Behavior* 47: 268-298.

Egas, M., Riedl, A. (2008), The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of the Royal Society B – Biological Sciences*, forthcoming.

Engelmann, D., Strobel, M. (2004), Inequity aversion, efficiency, and maximin preferences in simple distribution experiments. *American Economic Review* 94: 857-869.

Ertan, A., Page, T., Putterman, L. (2005), Can endogenously chosen institutions mitigate the free-rider problem and reduce perverse punishment? Brown University, Working Paper.

Falk, A., Fehr, E., Fischbacher, U. (2005), Driving forces behind informal sanctions. *Econometrica* 73: 2017-2030.

Falk, A., Fischbacher, U. (2006), A theory of reciprocity. *Games and Economic Behavior* 54: 293-315.

Fehr, E., Fischbacher, U. (2004), Third party punishment and social norms. *Evolution and Human Behavior* 25: 63-87.

Fehr, E., Gächter, S. (2000), Cooperation and punishment in public goods experiments. *American Economic Review* 90: 980-994.

Fehr, E., Gächter, S. (2002), Altruistic punishment in humans. *Nature* 415 (10 January 2002): 137-140.

Fehr, E., Gächter, S. (2003), Self-governing solutions to collective action problems: Rewards versus punishments. Preliminary draft.

Fehr, E., Klein, A., Schmidt, K. (2007), Fairness and contract design, *Econometrica* 75: 121-154.

Fehr, E., Schmidt, K. (1999), A theory of fairness, competition and cooperation. *Quarterly Journal of Economics* 114: 817-868.

Fehr, E., Schmidt, K. (2004), Fairness and incentives in a multi-task principal-agent model. *Scandinavian Journal of Economics* 106: 453-474.

Feld, L, Frey, B., Torgler, B. (2006), Rewarding honest taxpayers? Evidence on the impact of rewards from field experiments. CREMA, Working Paper No. 2006-16.

Feld, L., Tyran, J.-R. (2002), Tax evasion and voting: An experimental analysis. *Kyklos* 55: 197-221.

Fischbacher, U. (2007), z-Tree - Zurich toolbox for readymade economic experiments – experimenter's manual. *Experimental Economics* 10: 171-178.

Fischbacher, U., Gächter, S. (2006), Heterogeneous social preferences and the dynamics of free riding. CeDEx Discussion Paper No. 2006-01.

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

Frey, B. S. (1994), Direct democracy: Politico-economic lessons from Swiss experience. *American Economic Review, Papers and Proceedings* 84: 338-342.

Frey, B. S., Benz, M., Stutzer, A. (2004), Introducing procedural utility: Not only what, but also how matters. *Journal of Institutional and Theoretical Economics* 160: 377-401.

Gächter, S. (2007), Conditional cooperation: Behavioral regularities from the lab and the field and their policy implications. In Frey, B. S., Stutzer, A. (eds), *Economics and Psychology. A Promising New Cross-Disciplinary Field*. CESifo Seminar Series, Cambridge/Mass., MIT Press: 19-50.

Gürerk, Ö., Irlenbusch, B., Rockenbach, B. (2006), The competitive advantage of sanctioning institutions. *Science* 312 (7 April 2006): 108-111.

Herrmann, B., Thöni, C., Gächter, S. (2008), Antisocial punishment across societies. *Science* 319 (7 March 2008): 1362-1366.

Kroll, S., Cherry, T. L., Shogren, J. F. (2007), Voting, punishment, and public goods. *Economic Inquiry* 45: 557-570.

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

Masclet, D., Villeval, M.-C. (2008), Punishment and inequality: a public good experiment. *Social Choice and Welfare*, forthcoming.

Nikiforakis, N. (2008), Punishment and counter-punishment in public-good games: Can we really govern ourselves? *Journal of Public Economics* 92: 91–112.

Nikiforakis, N., Normann, H.-T. (2008), A comparative statics analysis of punishment in public-good experiments. *Experimental Economics*, forthcoming.

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

Pommerehne, W. W., Weck-Hannemann, H. (1996), Tax rates, tax administration and income tax evasion in Switzerland. *Public Choice* 88: 161-170.

Rabin, M. (1993), Incorporating fairness into game theory and economics. *American Economic Review* 83: 1281-1302.

Rege, M., Telle, K. (2004), The impact of social approval and framing on cooperation in public good situations. *Journal of Public Economics* 88: 1625-1644.

Sefton, M., Shupp, R., Walker, J. (2007), The effect of rewards and sanctions in the provision of public goods. *Economic Inquiry* 45: 671-690.

Yamagishi, T. (1986), The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology* 51: 110-116.

9 Tables and figures

Table 1: Treatment characteristics and number of independent observations

Determination of the institution	Leverage	Institution: Standard VCM	Institution: Punishment	Institution: Reward
Exogenous	$\|L\|=1$	$N = 10$*	$N = 10$	$N = 10$
	$\|L\|=3$		$N = 10$	$N = 10$
Endogenous#	$\|L\|=1$	$N = 25$	$N = 5$	$N = 10$
	$\|L\|=3$	$N = 3$	–	$N = 17$

* There is no leverage in the standard VCM because there are no sanctions available.

The table indicates the actually chosen institutions in the endogenous treatments. In sum, we had 40 groups in the endogenous treatments with $|L|=1$, and 20 groups with $|L|=3$.

Table 2: Predictions for contributions and voting with different preference models

Panel (A) – *Contributions*	Standard preferences: Selfish payoff-maximization	Fehr and Schmidt (1999)-preferences: Inequity aversion*	Charness and Rabin (2002)-preferences: Social welfare maximization[#]
Standard VCM	$c_i = 0 \;\forall\; i$	[A] $c_i = 0 \;\forall\; i$ if $\exists\, j$ with $\beta_j < 0.6$, or [B] $c_i = c \in [0, 20]$ if $\beta_i \geq 0.6 \;\forall\; i$ [†]	$c_i = 0 \;\forall\; i$ with $\lambda_i < 0.5$ and $c_i = 20 \;\forall\; j$ with $\lambda_j \geq 0.5$ (and $\delta_j \leq 1 - 1/2\lambda_j$)
Punishment with $\|L\| = 1$	$c_i = 0 \;\forall\; i$	Same as Standard VCM	Same as Standard VCM
Reward with $\|L\| = 1$	$c_i = 0 \;\forall\; i$	Same as Standard VCM	$0 \leq c_i \leq \bar{c} = n''L/(1-\gamma)$, with $x = 0.75$, otherwise (if no member satisfies $\lambda_i \geq 0.75$) same as Standard VCM
Punishment with $\|L\| = 3$	$c_i = 0 \;\forall\; i$	$0 \leq c_i \leq \bar{c} = n'L/(1-\gamma)$	Same as Standard VCM
Reward with $\|L\| = 3$	$c_i = 0 \;\forall\; i$	Same as Standard VCM	$0 \leq c_i \leq \bar{c} = n''L/(1-\gamma)$, with $x = 0.5$, otherwise (if no member satisfies $\lambda_i \geq 0.5$) it holds $c_i = 0 \;\forall\; i$.
Panel (B) – *Incentives for voting*			
Endogenous choice with $\|L\| = 1$	No vote participation	No vote participation	Group members with $\lambda_i \geq 0.75$ have an incentive to participate in the vote and support *reward* [‡]
Endogenous choice with $\|L\| = 3$	No vote participation	Group members with $\beta_i \geq 0.6$ have an incentive to participate in the vote and support *punishment* [‡]	Group members with $\lambda_i \geq 0.5$ have an incentive to participate in the vote and support *reward* [‡]

i Group member *i*, with four members in the group.

* The variable n' indicates the number of group members with $\beta_i \geq 0.6$. In the Fehr and Schmidt (1999)-model, the parameter β measures a subject's sensitivity to advantageous inequality.

[#] The variable n'' indicates the number of group members with $\lambda_i \geq x$ (where x constitutes a certain threshold). The parameter λ measures in the Charness and Rabin (2002)-model the weight a subject assigns to social welfare relative to own payoff (which receives weight $1 - \lambda$). "Social welfare" captures either a concern for the minimum payoff in the group (with weight δ) or a concern for the sum of all group members' payoffs (with weight $1 - \delta$).

[†] Taking for the purpose of illustration the distribution of parameter β reported in Fehr and Schmidt (1999, p. 844), the case of $\beta_I \geq 0.6 \;\forall\; i$ occurs in 2.6% of cases, whereas the case of at least one group member j satisfying $\beta_j < 0.6$ occurs in 97.4% of four-person groups.

[‡] Note that a coordination problem arises in the voting stage, since in equilibrium only one of the group members with an incentive to vote should actually vote because voting is costly.

Table 3: Average contributions and profits in the *exogenous* treatments

	Leverage	Standard VCM	Punishment	Reward
Panel (A) – Contributions	|*L*|=1	6.50$^{\#,*}$ (*N* = 10)	8.79 (*N* = 10)	6.62 (*N* = 10)
	|*L*|=3		12.11$^{\#}$ (*N* = 10)	11.20* (*N* = 10)
Panel (B) – Profits	|*L*|=1	23.90*	24.37	23.97
	|*L*|=3		24.83§	29.37*,§

[#] Significant difference ($p < 0.05$) between standard VCM and punishment.

* Significant difference ($p < 0.05$) between standard VCM and reward.

[§] Significant difference ($p < 0.05$) between punishment and reward.

Table 4: Number of voting rounds and voting behavior in first round

Panel (A) – Voting rounds with |L|=1

Number of voters in a group (*N* = total number of groups)	Reaching an unanimous vote in voting round …								
	1	2	4	7	9	10	11	13	26
1 (*N* = 15)	15								
2 (*N* = 13)	9	1		1		1		1	
3 (*N* = 6)	3	1					1		1
4 (*N* = 3)		1	1		1				

Panel (B) – Voting rounds with |L|=3

Number of voters in a group (*N* = total number of groups)	Reaching an unanimous vote in voting round …							
	1	2	3	11	13	51	61	103
1 (*N* = 5)	5							
2 (*N* = 6)	5		1					
3 (*N* = 5)	2	1			1		1	
4 (*N* = 4)			1	1		1		1

Panel (C) – Voting behavior in first voting round

Leverage	Number of voters	Support for … (multiple support possible)	VCM	Punishment	Reward	1st preference = actual institution
|L|=1	71		47	17	28	59
|L|=3	48		18	8	38	38

Table 5: Average contributions and profits in the *endogenous* treatments

	Leverage	Standard VCM	Punishment	Reward
Panel (A) – Contributions	$\|L\|$=1	5.48$^{\#,*}$ (N = 25)	16.13$^{\#,§}$ (N = 5)	8.53*,§ (N = 10)
	$\|L\|$=3	9.17* (N = 3)	-	14.59* (N = 17)
Panel (B) – Profits	$\|L\|$=1	23.29$^{\#,*}$	29.25$^{\#,§}$	25.12*,§
	$\|L\|$=3	25.51*	-	32.05*

$^{\#}$ Significant difference ($p < 0.05$) between standard VCM and punishment.

* Significant difference ($p < 0.05$) between standard VCM and reward.

§ Significant difference ($p < 0.05$) between Punishment and reward.

Table 6: Vote participation and contributions in the public goods game (censored tobit)

Independent variables	Dependent variable: Contributions in … first period	second period	overall (Periods 1-10)
Voter (= 1)	4.01*	2.67	0.95
Punishment institution (= 1)	5.98**	15.44**	11.91**
Reward institution (= 1)	5.35*	10.98**	4.15**
Leverage (= 1 if $\|L\| = 3$)	-2.09	0.35	5.79**
Number of voters in group	0.94	0.15	-0.12
Intercept	3.84*	1.30	4.73*

** Significant at 1% level; * significant at 5% level.

Table 7: Punishing or rewarding behavior (panel probit-regression)

	Dependent variable: Decision to…		
Independent variables	punish with |L| = 1	reward with |L| = 1	reward with |L| = 3
Voter (= 1)	-0.173	0.217*	0.340**
Number of voters in group	0.238*	0.265**	-0.054
Group contribution	-0.006**	0.002**	0.013**
Difference between other member's and own contribution	-0.098**	0.035**	0.022**
Observations	600	1200	2040

** Significant at 1% level; * significant at 5% level.

Figure 1. Average contributions in the exogenous treatments

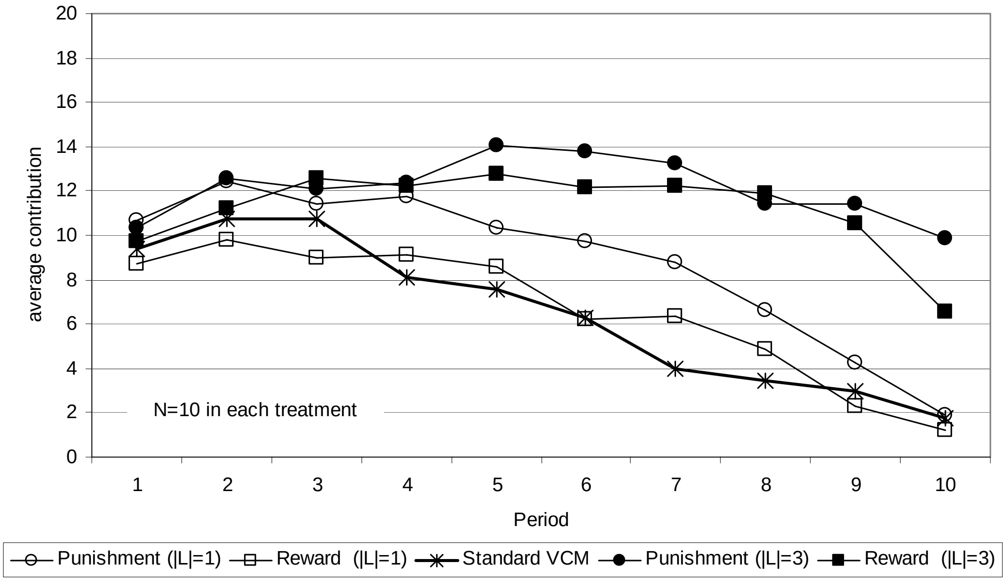
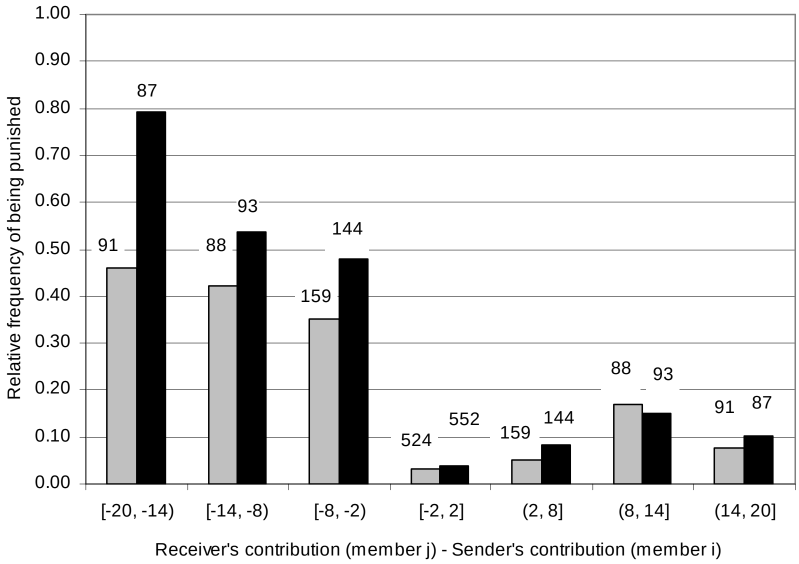


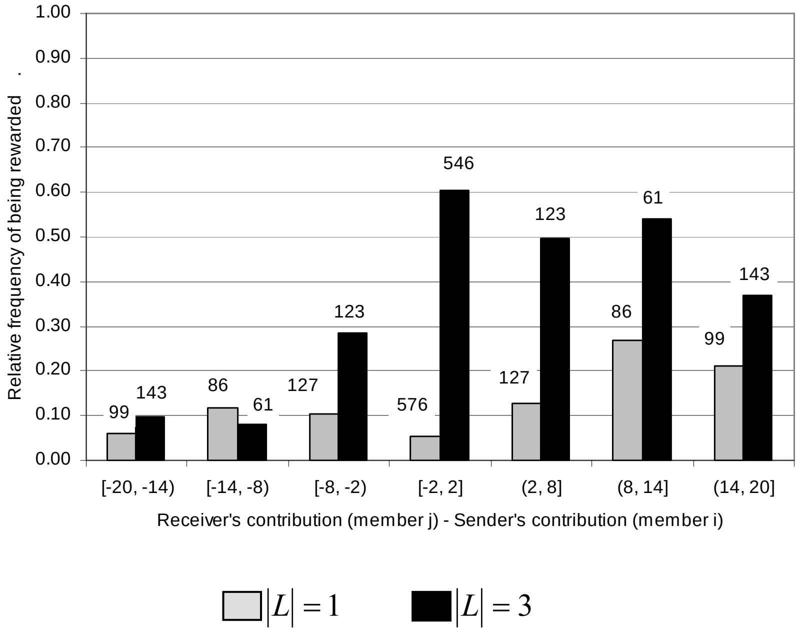
Figure 2. Punishment in the exogenous treatments



$|L| = 1$ $|L| = 3$

Numbers above bars indicate the number of observations.

Figure 3. Reward in the exogenous treatments



Numbers above bars indicate the number of observations.

Figure 4. Average contributions in the endogenous treatments

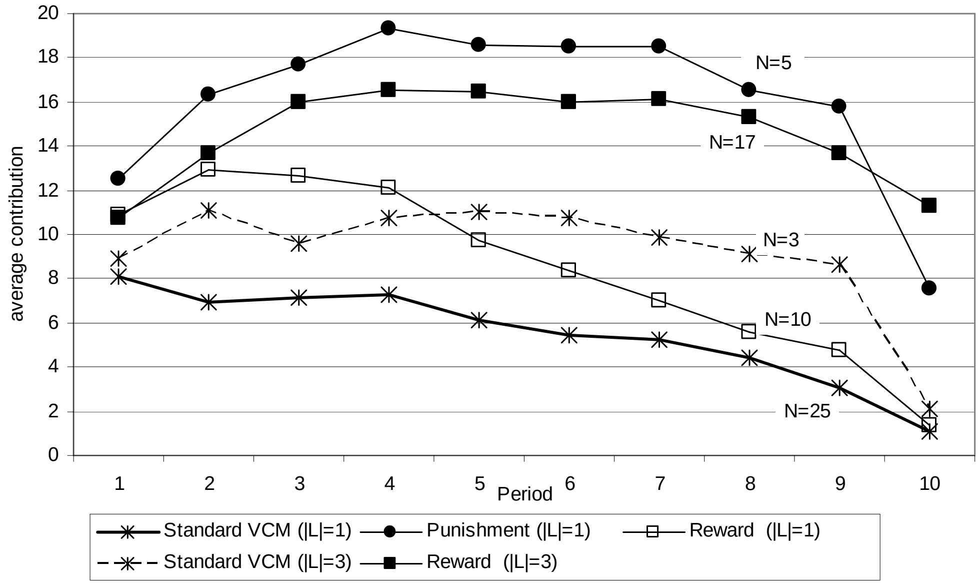
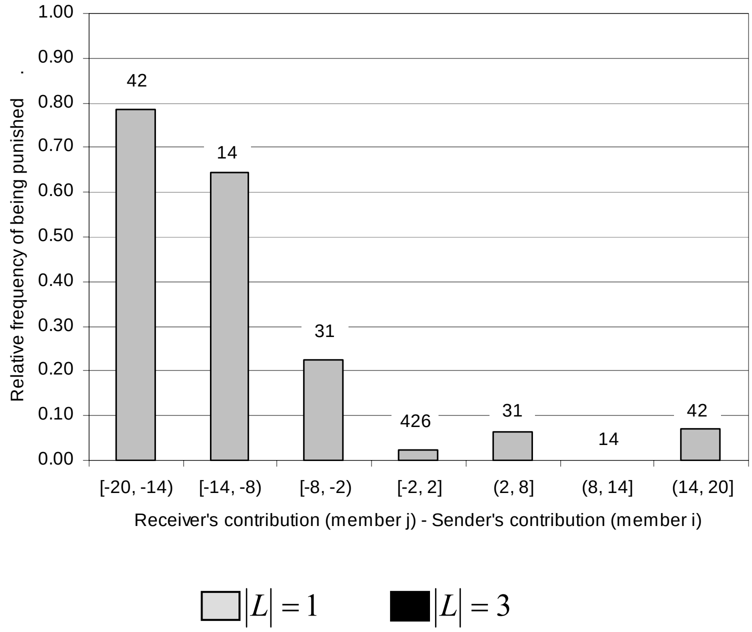


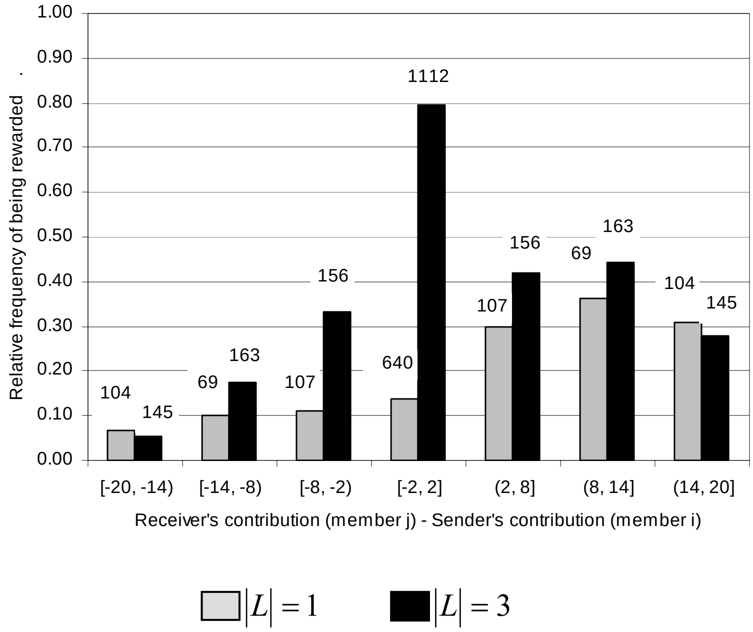
Figure 5. Punishment in the endogenous treatments



Numbers above bars indicate the number of observations.

Note: There are only data for $|L|$=1 because punishment was never chosen with $|L|$=3.

Figure 6. Reward in the endogenous treatments



Numbers above bars indicate the number of observations.

Appendix A: Instructions for the endogenous choice with $|L|$ = 3

(Originally in German –instructions for other treatments are available upon request) [not intended for publication; can be made available on our website]

Welcome to the experiment. Please refrain from talking to other participants from now on.

Groups of 4 persons and 10 periods

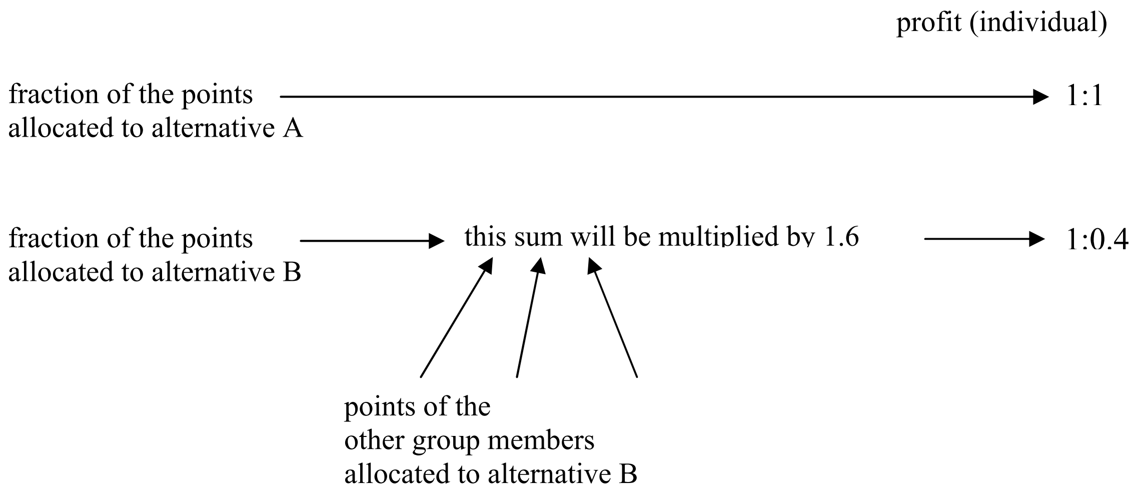
At the beginning of the experiment you are randomly assigned to a group of 4 subjects that will remain the same throughout the whole experiment. The whole experiment lasts 10 periods.

Basic decision

At the beginning of each period you will receive an endowment of 20 tokens. These tokens can be allocated into two alternatives, denoted A and B. The sum of the tokens allocated to A and B must equal 20. You cannot save any tokens or transfer them to a later period.

Profits from both alternatives

You will receive the tokens allocated to alternative A back one-to-one as one part of your profit.
The tokens allocated to alternative B are summed up with the contributions of the other group members to alternative B. The sum of all tokens allocated to alternative B in your group is multiplied by 1.6 and will be equally distributed– independent of one's particular contribution – among all group members. This means that you – and every other group member – will receive 0.4 tokens in return (=1.6/4) for every token allocated to alternative B.
The following figure summarizes the profits from both alternatives. The sum of profits from both alternatives gives your total earnings in this period.



Three possible institutions

Institution number one

Under this institution you have to make your basic decision only. Your profit per period is as described above, depending on how much you have allocated to alternative A and how much the group altogether has allocated to alternative B. After the basic decision has been entered by all subjects, you will be informed on the next screen about the decisions of all other group members.

Institution number two

This is the same as institution number one, with the following additional feature: After having observed the other members' contributions you have the option to assign points to other members. You can assign points to as many other members as you wish, but you can only assign one point per member. Assigning a point has the following consequences: The person assigning a point bears the cost of one token (this cost will be subtracted from the period payoff). The person receiving a point bears the cost of three tokens (also to be subtracted from her period payoff).

Institution number three

This is also the same as institution number one, with the following additional feature: After having observed the other members' contributions you have the option to assign points to other members. You can assign points to as many other members as you wish, but you can only assign one point per member. Assigning a point has the following consequences: The person assigning a point bears the cost of one token (this cost will be subtracted from the period payoff). The person receiving a point receives three tokens (to be added to her period payoff).

Which institution is valid for your group?

Before the experiment starts, you have the option to participate in a vote on the institution that will be valid in your group for all 10 periods. You are not required to participate in the vote, but even in that case the voting outcome in your group will be binding for you as well. Participation in the vote has a one-time cost of 10 tokens (which will be subtracted from your first period's earnings).

If you participate in the vote you have to indicate for each of the three institutions introduced above whether you support this institution (click on "Yes") or not (click on "No"). You can support as many institutions as you want. Before the vote starts, you will see the number of voters in your group. When all group members that participate in the vote have made their decisions, you will see on a succeeding screen how many group members supported any of the three institutions.

1. If one and only one institution is unanimously supported by all voters in your group, then this institution will be applied for the 10 periods.

2. If two or three institutions are unanimously supported, then the institution will be chosen randomly from those institutions unanimously supported by all voters in your group.
3. If no institution has unanimous support, then the voting procedure will be repeated.

The voting procedure is repeated until the unanimous support of at least one institution is reached.

If no group member participates in the vote, the institution is chosen randomly with equal probabilities for each institution.

[In order to speed up the voting procedure, we ask you to pay close attention to the status bar where you will see either the message “waiting for others” or “please vote”.]

Screen sequences

In the upper part of the screen you will see your group and subject number. On all screens you will also be able to see the prevalent institution in your group.

Independent of the institution used, you always have to make the basic allocation decision. Then, you will be shown the basic decisions of all group members on the next screen. Under institution one you will also see the profits. Under institutions two and three you can then enter the points if you wish to assign some. As the final screen in institutions two and three of each period, you will see again the basic decisions of all group members (linked to their ID) and the sum of points you have assigned and received. You are not able to see how many points other group members have assigned or received or who has assigned points to you. The profits are calculated as described above.

Exchange rate. At the end of the experiment all tokens you have earned are summed up and converted in real money at a rate of **10 tokens = 0.40 EURO**.

Appendix B: Predictions and propositions [not necessarily for publication; this material can be made available on our website]

B1. Theoretic predictions when agents are purely selfish

B1.1 Standard voluntary contribution mechanism

The payoff for member *i* in period *t* is given by $\pi_{i,t} = E - c_{i,t} + \gamma C_t$, where $C_t = \sum_{j=1}^{n} c_{j,t}$ denotes the sum of contributions within a group. Given that $0 < \gamma < 1 < n\gamma$, it follows that the marginal return for member *i* from investing into the public good is negative (since $\partial \pi_{i,t} / \partial c_{i,t} = -1 + \gamma < 0$). Under the assumptions of selfishness and common knowledge of rationality, the only subgame-perfect equilibrium is to contribute zero in each single period (i.e. $c_{i,t} = 0\ \forall t$).

B1.2 VCM with punishment or reward

Since punishment (reward) is costly (both with $|L| = 1$ and $|L| = 3$), no member will ever punish (reward) another member in the second stage of the final period. Hence, contributions in the first stage of the final period cannot be affected by the availability of punishment (reward). Unraveling the same logic back to the first period yields zero contributions as the only subgame-perfect equilibrium. Given that equilibrium contributions are identical under all available institutions, it follows that in the endogenous treatments all members should abstain from voting in order to save the voting costs.

B2. Theoretic predictions based on the Fehr and Schmidt (1999)-model

The subsequent analysis assumes social preferences of the Fehr-Schmidt type (see equation (3) in section 5), and common knowledge of rationality.

B2.1 Standard voluntary contribution mechanism

Proposition 1*: Equilibrium contributions in the standard linear VCM depend on the distribution of social preferences in a group. If at least one member in a four-person group cares relatively little about advantageous inequality (i.e., $\beta_i < 0.6$), the only equilibrium is* *complete free-riding ($c_{i,t}$ = 0) of all members. Otherwise, there is a multiplicity of equilibria with all members contributing $c_{i,t} = c \in [0, E]$.*

Fehr and Schmidt (1999) show (in their Proposition 4 on p. 839) for the standard VCM that

(a) If $\gamma + \beta_i < 1$ for member *i*, then it is a dominant strategy for that member to free-ride and choose $c_{i,t} = 0$ in all periods.

(b) Let *f* denote the number of members with $\gamma + \beta_i < 1$, where $0 \leq f \leq n$. If $f/(n-1) > \gamma/2$, then there is a unique equilibrium with $c_{i,t} = 0$ for all members *i*.

(c) Only if $f/(n-1) < (\gamma + \beta_j - 1)/(\alpha_j + \beta_j)$ for all members *j* with $\gamma + \beta_j \geq 1$ (the latter being called '*cooperators*' in the following), there is a multiplicity of equilibria where members with $\gamma + \beta_i < 1$ contribute zero and the others contribute $c_{i,t} = c \in [0, E]$. Under this condition it must hold that $(\gamma + \beta_j - 1)/(\alpha_j + \beta_j) < \gamma/2$.

For the intuition and the formal proofs we refer the reader to Proposition 4 of Fehr and Schmidt (1999, p. 839 and pp. 860-863). Note that since in c) we face multiple equilibria already in the one-shot game, there might be more equilibria than laid out in c) in the finitely repeated game.[23] From part (b) it follows for our case with $\gamma = 0.4$ and $n = 4$ that if $f/(n-1) > 0.2$ (i.e. the fraction of members with a $\beta_i < 0.6$ is larger than 20%), there is no equilibrium with positive contributions in the repeated standard VCM. This means that one member with $\gamma + \beta_i < 1$ suffices to wipe out cooperation in the whole group.

B2.2 Low leverage ($|L| = 1$)

B2.2.1 Punishment

Proposition 2*: Consider the VCM with punishment and $|L| = 1$. Complete free-riding is the only equilibrium when at least one group member satisfies $\beta_i < 0.6$ (which yields the same*

[23] This reservation is also valid for many of the cases discussed below.

prediction as in the VCM). Only if all members satisfy $\beta_i \geq 0.6$, any positive contribution can be sustained as an equilibrium.[24]

Let us denote the number of cooperators (who satisfy $\gamma + \beta_i \geq 1$) in a group by *n'*. For reasons of generality, let us define *k* as the costs of punishing or rewarding another subject. Note that *k* is normalized to unity ($k = 1$) in our experiment. Consider a strategy where all members contribute $c > 0$ and where all cooperators punish any member who contributes $c_{i,t} < c$. For this strategy to be an equilibrium we must show that (*i*) it does not pay for any member to free-ride and contribute less than *c* and that (*ii*) cooperators have an incentive to punish those who contribute $c_{i,t} < c$, i.e. that the punishment threat is credible.

(*i*) Free-riding on the other members' contributions *c* by choosing $c_{i,t} < c$ generates a monetary gain of $(c - c_{i,t})(1-\gamma)$ for the free-rider (relative to those contributing *c*). If the *n'* cooperators punish the free-rider, the latter suffers a monetary loss of *n'L*. The maximum gain from a deviation from *c* is given for $c_{i,t} = 0$. The resulting gain is smaller than the loss from being punished as long as the following condition is satisfied.

$$c \leq \bar{c} = \frac{n'L}{(1-\gamma)}. \tag{B1}$$

Hence, no member has an incentive to deviate from contributing *c* as long as the latter condition is fulfilled. What remains to be shown is whether the threat of punishing those members who would consider contributing $c_{i,t} < c$ is credible.

(*ii*) For the threat to be credible, a cooperator's utility from punishing must be larger than her utility from not punishing, under the assumption that the other $(n'-1)$ cooperators stick to their punishment strategy. Hence, we have to check the following inequality for member *i*.

$$-k - \frac{\alpha_i}{n-1}(n-n'-1)k - \frac{\alpha_i}{n-1}(\bar{c} - c + k - n'L) \geq -\frac{\alpha_i}{n-1}(\bar{c} - c - (n'-1)L) - \frac{\beta_i}{n-1}(n'-1)k \tag{B2}$$

The first term on the left hand side denotes the costs of punishing a deviating member. The second term indicates the disadvantageous inequality towards those members who contribute *c*, but who do not punish deviating members. The third term captures the remaining disadvantageous inequality towards the deviating member who gets punished by

[24] This latter case is rather unlikely. See Table 2 for details.

all *n'* cooperators. The first term on the right-hand side denotes the disadvantageous inequality towards the deviating member if member *i* does not punish, but the other *n'* – 1 cooperators punish. The second term is due to the advantageous inequality of member *i* towards the *n'* – 1 cooperators who punish. Rearranging and simplifying yields the following condition to be satisfied in order to make punishment a credible threat.

$$\frac{L}{k} \geq (n - n') + \frac{1}{\alpha_i}\left[(n-1) - \beta_i (n'-1)\right] \tag{B3}$$

Note that the left-hand side of equation (B3) yields $L/k = 1$ in our experiment in the low-leverage treatment ($|L| = 1$). From that it follows that condition (B3) cannot be satisfied as long as $n' < n$. Hence, the threat of punishment is not credible under $|L| = 1$ if $n' < n$. As a consequence, the same predictions as in the standard VCM apply. Only if $n' = n$, i.e. if all group members are cooperators (with $\gamma + \beta_i \geq 1$), the threat of punishment is credible and any contribution level $c \leq \bar{c}$ can be enforced.

B2.2.2 Reward

***Proposition 3**: Consider the VCM with reward and $|L| = 1$. If at least one group member satisfies $\beta_i < 0.6$, zero contributions are the only equilibrium. Only when all members are rather averse to advantageous inequality ($\beta_i \geq 0.6$ for all i) then positive contributions $c_{i,t} = c \in [0, E]$ are feasible in equilibrium. Using the reward option can only be part of an equilibrium strategy if all members satisfy $\beta_i \geq 0.75$.*[25]

Suppose all members contribute $c \in [0, E]$. Note that the *n* – *n'* members with $\gamma + \beta_i < 1$ will never reward other members. A cooperator, then, has to consider the following inequality in order to decide whether to reward (left-hand side) or not (right-hand side).

$$(n'-1)L - (n-1)k - \frac{\alpha_i}{n-1}\left[L + (n-1)k\right](n-n') \geq (n'-1)L - (n'-1)\frac{\beta_i}{n-1}\left[L - (n'-1)k\right] \tag{B4}$$

[25] Obviously, using the reward strategy with $|L| \leq 1$ does not increase overall efficiency in terms of monetary payoffs.

The first term on the left-hand side captures the gains from being rewarded by the other $n' - 1$ cooperators. The second term shows the costs from rewarding all other members. The third term denotes the disadvantageous inequality towards the $n - n'$ members with $\gamma + \beta_i < 1$ who do not reward others. The first term on the right-hand side shows again the gains from being rewarded by the other cooperators. The second term indicates the advantageous inequality towards those $n' - 1$ cooperators. Rearranging terms yields the following condition that has to be satisfied in order to make cooperators reward the other members.

$$(n-1)k + \frac{1}{n-1}\left[L + (n-1)k\right]\left[\alpha_i (n-n') - \beta_i (n'-1)\right] \leq 0 \qquad \text{(B5)}$$

This condition is never satisfied for $n' < n$. Therefore, there are no equilibria in which reward is part of the equilibrium strategy. Thus, for $n' < n$ we have the same set of equilibria as in the standard VCM.

For $n' = n$ we find that mutual rewarding can be part of an equilibrium strategy. To show this, we examine a situation where all members contribute c and reward each other. If member i deviates and does not reward the other $(n-1)$ members, she saves rewarding costs of $(n-1)k$. But at the same time member i suffers from advantageous inequality, expressed by $\frac{\beta_i}{n-1}\left[L + (n-1)k\right](n-1)$. Sticking to reward is thus optimal if

$$(n-1)k \leq \frac{\beta_i}{n-1}\left[L + (n-1)k\right](n-1) \qquad \text{(B6)}$$

For $|L| = 1$ and $k = 1$ we must have $\beta_i \geq 0.75$ for condition (B6) to hold. If all group members satisfy $0.6 < \beta_i < 0.75$, any positive contribution $c_{i,t} = c \in [0, E]$ can still be sustained as an equilibrium, but mutual reward is not part of an equilibrium.

B2.2.3 Voting under the low leverage (|*L*| = 1)

Proposition 4*: Suppose group members can vote between the standard VCM, the VCM with punishment, and the VCM with reward under |L| = 1. If voting is costly, subjects do not participate in the vote because contributions are expected to be the same for the standard VCM and reward or punishment with |L| = 1.*

Given the positive costs of voting, Propositions 2 and 3 imply Proposition 4.

B2.3 High leverage (|L| = 3)

B2.3.1 Punishment

***Proposition 5**: Consider the following strategy in the VCM with punishment and |L| = 3. In the first stage of each period each member contributes $c \in [0,E]$. If all members do so, there is no punishment in the second stage. This constitutes a subgame-perfect equilibrium if the following two conditions are satisfied:*

(i) $L \geq \frac{(1-\gamma)(c-c_{i,t})}{n'}$, *with* $c_{i,t} < c$ *denoting a free-rider's contribution, and*

(ii) $\frac{L}{k} \geq (n-n') + \frac{1}{\alpha_i}\left[(n-1) - \beta_i(n'-1)\right]$ *for all n' cooperators.*

The strategy constitutes a continuum of equilibria in the game where $c \leq \bar{c} = Ln'/(1-\gamma)$ *can be enforced, but where no punishment actually occurs. Consequently, contributions are expected to be higher with punishment under |L| = 3 than under |L| = 1 and higher than in the standard VCM.*

Under the high-leverage |L| = 3 condition (B3) from section B2.2.1 can be satisfied even if $n' < n$ (depending on a member's α_i and β_i). Hence, punishment can be a credible threat even if not all group members are cooperators. Therefore, it is possible to enforce $\bar{c}$. All other equilibria discussed for the case of punishment with |L| = 1 remain also valid for |L| = 3.

B2.3.2 Reward

***Proposition 6**: Consider the VCM with reward and |L| = 3. Equilibrium predictions in this institution are identical to those with reward under |L| = 1. Yet, equilibria where all members reward each other are easier to achieve than under |L| = 1 because they require* $\beta_i \geq 0.6$ *instead of* $\beta_i \geq 0.75$ *for all members.*

From condition (B5) in section B2.2.2 it can be seen that even under the high leverage treatment it is not possible to enforce any $c \in [0, E]$ through the use of rewards, since it is better even for cooperators to abstain from rewarding as long as $n' < n$. Only if $n' = n$ mutual rewarding can be part of an equilibrium strategy. From condition (B6) and the assumption that $n' = n$ it follows that mutual rewarding is part of an equilibrium strategy already if $\beta_i \geq 0.6$ for all members (which is a less restrictive condition than under $|L| = 1$, where it was necessary to satisfy $\beta_i \geq 0.75$ for all members).

B2.3.3 Voting under the high leverage ($|L| = 3$)

Proposition 7*: Suppose group members can vote between the standard VCM, the VCM with punishment, and the VCM with reward under $|L| = 3$. Assume that the distribution of social preferences follows the note in Table 2 as described above. Then, subjects are expected to support only the punishment institution and voter turnout will be higher under $|L| = 3$ than under $|L| = 1$.*

Concerning voting behavior in our endogenous treatment under $|L| = 3$, it is important to note first that the punishment institution is the only one in which a contribution level $\bar{c} > 0$ can actually be enforced (with the magnitude of $\bar{c}$ depending on the number of cooperators in the group). In the standard VCM or with reward, positive contributions may be sustained in equilibrium, but only if none of the group members satisfies $\beta_i < 0.6$. Taking the distribution of social preferences from Fehr and Schmidt (1999), 97.4% of four person-groups would have at least one member with $\beta_i < 0.6$, and therefore the only equilibrium would be zero contributions. In the punishment institution, though, it suffices to have one subject with $\beta_i \geq 0.6$ to be able to enforce a contribution of $\bar{c} = 5$. Given the distributional assumptions from Fehr and Schmidt (1999), about 87% of groups should have at least one subject with $\beta_i \geq 0.6$. It can be easily verified that for almost any meaningful assumption about the distribution of social preferences among subjects, it is optimal (in expected terms) to participate in the vote (and incur the one-time cost of 10 ECU) and support only the punishment institution.

Finally, we note that the voting stage entails a coordination problem. Given our approval voting procedure, in equilibrium only one subject per group should go to the ballot and vote for the punishment mechanism in $|L| = 3$ in order to save on voting costs.

B3. Theoretic predictions based on the Charness and Rabin (2002)-model

The subsequent analysis assumes (i) that n'' subjects (henceforth called cooperators) out of *n* group members have social preferences of the Charness-Rabin type (see equation (4) in section 5), (ii) $(n - n'')$ subjects care only for their own monetary payoff, and (iii) common knowledge of rationality applies.

B3.1 Standard voluntary contribution mechanism

Proposition 8*: Purely selfish players will never contribute positive amounts in the standard VCM. Cooperators will contribute their entire endowment to the public good if* $\lambda_i \geq 0.5$ and $\delta_i \leq 1 - \frac{1}{2\lambda}$.

For a purely selfish players it is still a dominant strategy to contribute $c = 0$, since the presence of players with social preferences does not change the incentives for selfish players. A cooperator has to determine whether contributing a positive amount increases his utility (see equation (4) in section 5) or not. Positive contributions decrease the cooperator's monetary payoff (by $1-\gamma$), increase the total payoff in the group (by $n\gamma - 1$), and decrease the minimum payoff in the group (which is the cooperator's payoff) by $1-\gamma$. Taking the derivative of equation (4) with respect to *c* yields the following inequality $(1-\lambda)(-1+\gamma)+\lambda[\delta(-1+\gamma)+(1-\delta)(-1+n\gamma)] \geq 0$. After some rearrangement this simplifies to

$$\delta \leq \frac{1.2\lambda - 0.6}{1.2\lambda} = 1 - \frac{1}{2\lambda} \quad \text{(B7)}$$

From (B7) it follows that the cooperator's concern for social welfare must satisfy $\lambda_i \geq 0.5$, and that δ_i must satisfy particular conditions (in relation to λ_i) such that a cooperator has a positive incentive to contribute. If so, then cooperators contribute their full endowment E. if (B7) is not satisfied, all group members contribute zero.

B3.2 Punishment

***Proposition 9**: Consider the VCM with punishment and $|L| \geq 1$. The existence of the punishment option cannot induce any contributions in excess of those which would also be contributed without the availability of punishment.*

A selfish player will only cooperate if she is threatened by punishment in case of defection. In order to enforce contributions $c > 0$ from the selfish players, the gain from cooperation must exceed the gain from defection (i.e. contributing zero). That is $E - c + \gamma C \geq E + \gamma C - \gamma c - n''L$, which yields $c \leq \bar{c} = n''L/(1-\gamma)$ after rearrangement. Thus, in general a selfish player has an incentive to contribute $0 < c \leq \bar{c}$ in case punishment by the cooperators is a credible threat. However, if a cooperator decided in the second stage to punish the defector she would be worse off than under no punishment, because punishment reduces (i) her own payoff, (ii) the minimum payoff in the group, and (iii) the group's total payoff. Thus, punishment cannot constitute a subgame-perfect equilibrium in the second stage and cannot induce any contributions in excess to those under the VCM.

B3.3 Reward

***Proposition 10**: Consider the VCM with reward and $|L| \geq 1$. In the case of $L = 3$, subjects who satisfy $\lambda_i \geq 0.5$ (and thus contribute $c_i = 20$; their number is denoted by n') can enforce a positive contribution of $\bar{c} = n''L/(1-\gamma)$ from subjects who care less for social welfare (with $\lambda_i < 0.5$). This enforcement rests on rewarding the selfish subjects for contributing $\bar{c}$. For the case of $L = 1$, the same argument applies, but reward can only be* *part of an equilibrium strategy (and thus $\bar{c}$ can only be enforced) if there are n'' subjects who satisfy the more demanding condition $\lambda_i \geq 0.75$.*

Similar to the case of punishment, a selfish player compares the payoff from contributing and getting rewarded in the second stage with the payoff from defection in the first stage. Therefore, we have to assess the following inequality: $E - c + \gamma C + n''L \geq E + \gamma C - \gamma c$. Again, after rearranging, we obtain the condition $c \leq \bar{c} = n''L/(1-\gamma)$. Thus, a selfish player cooperates if this condition is fulfilled, i.e. in case reward is credible, since then her payoff from cooperation exceeds her payoff from defection.

It remains to be shown whether it constitutes an equilibrium for a cooperator to reward $(n-1)$ group members in the second stage, even though she will, in turn, only be rewarded by the $(n''-1)$ other cooperators in the group. This is equivalent to check whether the inequality (B8) holds:[26]

$$\begin{aligned} &(1-\lambda)[(n'-1)L-(n-1)k]+ \\ &\lambda[\delta((n'-1)L-(n-1)k)+(1-\delta)\{n'[(n'-1)L-(n-1)k]+(n-n')(n'L)\}] \geq \\ &(1-\lambda)(n'-1)L+ \\ &\lambda[\delta((n'-1)L-L-(n-1)k)+(1-\delta)\{(n'-1)[(n'-1)L-L-(n-1)k]+(n-n'+1)(n'-1)L\}] \end{aligned} \qquad \text{(B8)}$$

After rearrangement and simplification we arrive at condition (B9):

$$\frac{L}{k} \geq \frac{(1-\lambda)(n-1)+\lambda(1-\delta)}{\lambda} \qquad \text{(B9)}$$

In the low-leverage treatment this is satisfied for $\delta = [0,1]$ and $\lambda = [3/4,1]$. Similarly, in the high-leverage treatment it is satisfied for $\delta = [0,1]$ and $\lambda = [0.5,1]$, provided that $5 + \delta \geq 3/\lambda$. Hence, as long as λ and δ satisfy these conditions, reward constitutes a credible incentive to elicit contributions up to $\bar{c}$ from selfish players.

---

[26] As the payoff differences occur in the second stage, we omit for simplicity all the variables from the first stage that cancel out each other, i.e. $E, \gamma C, \gamma c$.

B3.4 Voting under the low leverage (|*L*| = 1)

Proposition 11*: Suppose group members can vote between the standard VCM, the VCM with punishment, and the VCM with reward under |L| = 1. If voting costs are sufficiently small*[27]*, group members with* $\lambda_i \geq 0.75$ *participate in the vote and support reward*

Propositions 9 and 10 imply Proposition 11. Note again that subjects who care a lot for social welfare should participate in the vote in our endogenous treatments and support the reward option in order to be able to enforce positive contributions from subjects who care little for social welfare (and who do not participate in the vote). This latter group of subjects does only contribute in the VCM with reward but not in the standard VCM or the VCM with punishment.

B3.5 Voting under the high leverage (|*L*| = 3)

Proposition 12*: Suppose group members can vote between the standard VCM, the VCM with punishment, and the VCM with reward under |L| = 3. If voting costs are negligible, group members with* $\lambda_i \geq 0.5$ *participate in the vote and support reward*

Again, Propositions 9 and 10 imply Proposition 12. As a final remark, remember that given our approval voting procedure, in equilibrium only one subject per group should go to the ballot and vote for the reward mechanism in |*L*| = 3 in order to save on voting costs.

[27] Note that in our design voting costs were, indeed, small. We are not aware of any calibration results for the parameters in the Charness and Rabin (2002)-model. With them it would be possible to weigh voting costs against the expected advantage of implementing the reward option.

University of Innsbruck – Working Papers in Economics and Statistics
Recent papers

2008-07 **Matthias Sutter, Stefan Haigner and Martin Kocher:** Choosing the carrot or the stick? – Endogenous institutional choice in social dilemma situations.

2008-06 **Paul A. Raschky and Manijeh Schwindt:** Aid, Catastrophes and the Samaritan's Dilemma.

2008-05 **Marcela Ibanez, Simon Czermak and Matthias Sutter:** Searching for a better deal – On the influence of group decision making, time pressure and gender in a search experiment.

2008-04 **Martin G. Kocher, Ganna Pogrebna and Matthias Sutter:** The Determinants of Managerial Decisions Under Risk.

2008-03 **Jesus Crespo Cuaresma and Tomas Slacik:** On the determinants of currency crises: The role of model uncertainty.

2008-02 **Francesco Feri:** Information, Social Mobility and the Demand for Redistribution.

2008-01 **Gerlinde Fellner and Matthias Sutter:** Causes, consequences, and cures of myopic loss aversion - An experimental investigation.

2007-31 **Andreas Exenberger and Simon Hartmann:** The Dark Side of Globalization. The Vicious Cycle of Exploitation from World Market Integration: Lesson from the Congo.

2007-30 **Andrea M. Leiter and Gerald J. Pruckner:** Proportionality of willingness to pay to small changes in risk - The impact of attitudinal factors in scope tests.

2007-29 **Paul Raschky and Hannelore Weck-Hannemann:** Who is going to save us now? Bureaucrats, Politicians and Risky Tasks.

2007-28 **Harald Oberhofer and Michael Pfaffermayr:** FDI versus Exports. Substitutes or Complements? A Three Nation Model and Empirical Evidence.

2007-27 **Peter Wechselberger, Stefan Lang and Winfried J. Steiner:** Additive models with random scaling factors: applications to modeling price response functions.

2007-26 **Matthias Sutter:** Deception through telling the truth?! Experimental evidence from individuals and teams. *Revised version accepted for publication in The Economic Journal.*

2007-25 **Andrea M. Leiter, Harald Oberhofer and Paul A. Raschky:** Productive disasters? Evidence from European firm level data.

2007-24 **Jesus Crespo Cuaresma:** Forecasting euro exchange rates: How much does model averaging help?

2007-23 **Matthias Sutter, Martin Kocher and Sabine Strauß:** Individuals and teams in UMTS-license auctions.

2007-22 **Jesus Crespo Cuaresma, Adusei Jumah and Sohbet Karbuz:** Modelling and Forecasting Oil Prices: The Role of Asymmetric Cycles.

2007-21 **Uwe Dulleck and Rudolf Kerschbamer:** Experts vs. discounters: Consumer free riding and experts withholding advice in markets for credence goods.

2007-20 **Christiane Schwieren and Matthias Sutter:** Trust in cooperation or ability? An experimental study on gender differences. *Conditionally accepted for publication in: Economics Letters.*

2007-19 **Matthias Sutter and Christina Strassmair:** Communication, cooperation and collusion in team tournaments – An experimental study. *Revised version forthcoming in: Games and Economic Behavior.*

2007-18 **Michael Hanke, Jürgen Huber, Michael Kirchler and Matthias Sutter:** The economic consequences of a Tobin-tax – An experimental analysis.

2007-17 **Michael Pfaffermayr:** Conditional beta- and sigma-convergence in space: A maximum likelihood approach.

2007-16 **Anita Gantner:** Bargaining, search, and outside options. *Published in: Games and Economic Behavior, Vol. 62 (2008), pp. 417-435.*

2007-15 **Sergio Currarini and Francesco Feri:** Bilateral information sharing in oligopoly.

2007-14 **Francesco Feri:** Network formation with endogenous decay.

2007-13 **James B. Davies, Martin Kocher and Matthias Sutter:** Economics research in Canada: A long-run assessment of journal publications. *Revised version published in: Canadian Journal of Economics, Vol. 41 (2008), 22-45.*

2007-12 **Wolfgang Luhan, Martin Kocher and Matthias Sutter:** Group polarization in the team dictator game reconsidered. *Revised version forthcoming in: Experimental Economics.*

2007-11 **Onno Hoffmeister and Reimund Schwarze:** The winding road to industrial safety. Evidence on the effects of environmental liability on accident prevention in Germany.

2007-10 **Jesus Crespo Cuaresma and Tomas Slacik:** An "almost-too-late" warning mechanism for currency crises.

2007-09 **Jesus Crespo Cuaresma, Neil Foster and Johann Scharler:** Barriers to technology adoption, international R&D spillovers and growth.

2007-08 **Andreas Brezger and Stefan Lang:** Simultaneous probability statements for Bayesian P-splines.

2007-07 **Georg Meran and Reimund Schwarze:** Can minimum prices assure the quality of professional services?.

2007-06 **Michal Brzoza-Brzezina and Jesus Crespo Cuaresma:** Mr. Wicksell and the global economy: What drives real interest rates?.

2007-05 **Paul Raschky:** Estimating the effects of risk transfer mechanisms against floods in Europe and U.S.A.: A dynamic panel approach.

2007-04 **Paul Raschky and Hannelore Weck-Hannemann:** Charity hazard - A real hazard to natural disaster insurance. *Revised version forthcoming in: Environmental Hazards.*

2007-03 **Paul Raschky:** The overprotective parent - Bureaucratic agencies and natural hazard management.

2007-02 **Martin Kocher, Todd Cherry, Stephan Kroll, Robert J. Netzer and Matthias Sutter:** Conditional cooperation on three continents.

2007-01 **Martin Kocher, Matthias Sutter and Florian Wakolbinger:** The impact of naïve advice and observational learning in beauty-contest games.

University of Innsbruck

Working Papers in Economics and Statistics

2008-07

Matthias Sutter, Stefan Haigner and Martin Kocher

Choosing the carrot or the stick? – Endogenous institutional choice in social dilemma situations

Abstract

We analyze an experimental public goods game in which group members can endogenously determine whether they want to supplement a standard voluntary contribution mechanism with the possibility of rewarding or punishing other group members. We find a large and positive effect of endogenous institutional choice on the level of cooperation in comparison to exogenously implemented institutions. This suggests that participation rights enhance cooperation in groups. With endogenous choice, groups typically vote for the reward option, even though punishment is actually more effective in sustaining high levels of cooperation. Our results are evaluated against the predictions of social preference models.

ISSN 1993-4378 (Print)
ISSN 1993-6885 (Online)