

# Public goods provision and sanctioning in privileged groups

## Citation for published version (APA):

Riedl, A. M., & Reuben, E. (2009). Public goods provision and sanctioning in privileged groups. *Journal of Conflict Resolution*, 53(1), 72-93. <https://doi.org/10.1177/0022002708322361>

## Document status and date:

Published: 01/01/2009

## DOI:

[10.1177/0022002708322361](https://doi.org/10.1177/0022002708322361)

## Document Version:

Publisher's PDF, also known as Version of record

## Document license:

Taverne

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

---

# Public Goods Provision and Sanctioning in Privileged Groups

Ernesto Reuben

*Kellogg School of Management*

*Northwestern University*

Arno Riedl

*CESifo, IZA, and Department of Economics*

*Maastricht University*

In public-good provision, privileged groups enjoy the advantage that some of their members find it optimal to supply a positive amount of the public good. However, the inherent asymmetric nature of these groups may make the enforcement of cooperative behavior through informal sanctioning harder to accomplish. In this article, the authors experimentally investigate public-good provision in normal and privileged groups with and without decentralized punishment. The authors find that compared to normal groups, privileged groups are relatively ineffective in using costly sanctions to increase everyone's contributions. Punishment is less targeted toward strong free riders, and they exhibit a weaker increase in contributions after being punished. Thus, the authors show that privileged groups are not as privileged as they initially seem.

**Keywords:** *privileged groups; public goods; punishment; cooperation; collective action*

In public-good provision, the incentives to free ride on the contributions of others generally lead to underprovision of the public good. In his seminal work on collective action, Olson (1965) provides a typology of groups that helps to explain why some groups nevertheless perform better than others in collective-good provision. In particular, he identifies a type of group in which at least one member "has an incentive to see that the collective good is provided, even if he has to bear the full burden of providing it himself" (Olson 1965, 50). In such groups, one or more individuals receive a marginal benefit from the public good that exceeds their marginal cost of providing it. Hence, in such groups, at least some level of the public good is voluntarily produced. Correspondingly, Olson also calls them privileged groups.

---

**Authors' Note:** We thank Enrique Fatas, Rachel Croson, Dan Levin, Jan Potters, Louis Putterman, Frans van Winden, and particularly, Simon Gächter for useful comments and suggestions. Financial support from the Dutch Science Foundation (NWO) through the Evolution & Behavior grant 051-12-012 and from the EU-Marie Curie RTN ENABLE (MRTM-CT-2003-505223) is gratefully acknowledged. All materials for replication are available at <http://jcr.sagepub.com/> and at the authors' personal Web pages.

The theoretical literature on collective action generated important insights about when groups can be regarded as privileged. It has identified important determinants that turn nonprivileged groups into (fully) privileged ones, and hence, circumstances that mitigate underprovision of public goods (see, e.g., Sandler 1992). Considerably less attention has been devoted to the issue of underprovision of public goods in privileged groups, perhaps because it seems less problematic in such groups. A main purpose of this article is to study precisely this problem and show that privileged groups may not always be as privileged as they seem.

In our view, there are at least three reasons that call for further exploration of voluntary public-goods provision in privileged groups. First, it is clear that many groups confronted with the problem of providing public goods are privileged in the sense of Olson (1965). Contributing to the fight against international terrorism is an obvious example. The United States and the United Kingdom, for instance, can be assumed to contribute because they are likely targets, and hence, their governments perceive the benefits of contributing as being larger than the costs. In contrast, some continental European countries may see themselves as unlikely victims, and hence, perceive the costs as being larger than the benefits. The same could be said of most cases of international security, such as containment of “rouge” regimes. Another example is the so-called commons-based peer-production model, in which a large number of anonymous people with varying degrees of interest voluntarily contribute to a common project.<sup>1</sup>

Second, even though there are some voluntary contributions in privileged groups, underprovision will occur as long as not all individual incentives are aligned with maximization of social welfare.<sup>2</sup> That such underprovision is not only a theoretical issue is vividly revealed in national and international political discussions regarding contributions to fighting terrorism and global warming, for instance. As for ordinary, nonprivileged, groups, this asks for studying appropriate mechanisms overcoming the problem of suboptimal collective-good provision.

Third, experimental research has recently identified important informal determinants of voluntary contributions, such as fairness-oriented and reciprocal inclinations. These determinants, however, may work out differently in privileged groups. The asymmetric nature of such groups may cause conflicting perceptions of fairness and reciprocity norms, which in turn may undermine the otherwise very successful enforcement mechanism of decentralized individual punishment (Ostrom, Walker, and Gardner 1992; Fehr and Gächter 2000).

This article is the first to experimentally investigate voluntary contributions to a public good in privileged groups vis-à-vis normal, nonprivileged, groups. In particular, we compare behavior of group members with different incentives to contribute in situations with and without the option to costly punish free riders. Our results indicate that without punishment, there is substantial underprovision in privileged groups, albeit contributions are higher than in normal groups. With punishment, privileged groups lose their privileged status completely. Contributions to

the public good are not higher than in nonprivileged groups, and there is still substantial underprovision. In addition, inequality in privileged groups increases significantly when punishment is possible. Hence, the otherwise powerful institution of decentralized punishment loses some of its bite in privileged groups. The identification of the underlying reasons for this observation is another aim of our article.

We conducted a linear public-goods experiment (Isaac, Walker, and Thomas 1984) in which, in some treatments, participants are allowed to punish each other after having observed individual contributions to the public good (similar to Fehr and Gächter 2000). In normal groups, material incentives are such that all participants have a dominant strategy to free ride. In privileged groups, there is one participant whose material incentives are such that full contribution is a dominant strategy, whereas other group members' dominant strategy is to free ride completely. For convenience, we refer to participants with the different material incentives to contribute as high- and low-benefit participants.<sup>3</sup>

Note that under the standard assumption of (common knowledge of) narrow selfish preferences, only high-benefit participants are predicted to contribute (fully) to the public good. In addition, neither high- nor low-benefit participants are predicted to punish. However, there is plenty of evidence that in normal groups, first, members have some (initial) inclination to contribute positive amounts, and second, individual costly punishment is undertaken and effective in decreasing free riding.<sup>4</sup> These behavioral regularities are often attributed to members' having a preference for fair outcomes or reciprocal actions. From the outset, however, it is unclear if these types of social preferences can also increase contributions in privileged groups. First, since high-benefit members benefit disproportionately from contributing to the public good, it is unclear if such behavior should be interpreted as kind toward others or as merely selfish. Since intentions are an important part of reciprocity (Falk, Fehr, and Fischbacher 2008), such ambiguity may make low-benefit members unwilling to reciprocate high contributions of high-benefit members and high-benefit members unwilling to punish low contributions of low-benefit members. Second, since contributions of low-benefit members can increase income differences with high-benefit members, preferences for equality support outcomes in which low-benefit members do not contribute to the public good.<sup>5</sup>

Our results indeed show that differences in punishment as well as in the response to being punished can explain why privileged groups fail to outperform normal groups. The main reason is that the punishment institution produces a much smaller increase in contributions by low-benefit members in privileged groups. This in turn can be attributed to relatively unstructured punishment behavior of high-benefit members in combination with a weak response in contributions to being punished by low-benefit members.

Experimental studies have investigated the role of incentives to cooperate (i.e., the marginal per capita return to contribute) in public-goods settings quite extensively. From these studies, it is well known that changes in the marginal per capita

return have significant effects on contribution behavior (see Ledyard 1995). A few studies have even made cooperation the dominant strategy (Saijo and Nakamura 1995; Palfrey and Prisbrey 1996, 1997). Our article is also related to experiments investigating the effects of heterogeneous endowments in social dilemmas. The early findings of this literature are reviewed in Ostrom, Gardner, and Walker (1994) and Ledyard (1995), in which it is reported that inequality leads to lower cooperation levels. However, more recent studies have found mixed results: some report a negative effect of inequality (van Dijk, Sonnemans, and van Winden 2002; Cherry, Kroll, and Shogren 2005), others a positive effect (Chan et al. 1996; Buckley and Croson 2006) or no effect (Chan et al. 1999; Sadrieh and Verbon 2006). There are also a few studies using field experiments. Cardenas (2003) finds that real-life income inequalities between Colombian villagers make the use of communication less effective in increasing cooperation inside the lab. Visser and Burns (2006) report that among South African fishermen, lab-induced income inequalities do not mitigate the effectiveness of punishment for promoting cooperation. The work most closely related to this article is that of Fisher et al. (1995) and Brandts and Schram (2001). In both studies, the effect of unequal incentives to contribute within a group is explored. However, neither of the two investigates privileged groups as we do.

The rest of the article is organized as follows. In the next section, we describe the design of the experiment and derive qualitative predictions for theoretical models, assuming fairness preferences and reciprocity. In the section titled Voluntary Contributions, we present the results concerning participants' contribution and punishment behavior. The final section closes with a summary of our main results and concludes.

## Experimental Design

The basic game implemented in the experiment is a linear public-goods game. Each participant takes part in only one of four experimental treatments. The treatments differ in two dimensions: first, the public-goods game is either followed by a decentralized punishment stage a la Fehr and Gächter (2000, punishment condition), or there is no such punishment stage (baseline condition); second, members of a group differ with respect to the marginal benefit they receive from the public good (privileged group condition) or do not differ in this respect (normal group condition). The game is played by the same group of three participants for ten consecutive periods.

The public-goods game consists of a contribution stage in which each participant,  $i \in \{1, 2, 3\}$ , receives an endowment of twenty tokens. Participants simultaneously decide how many tokens,  $c_i \in [0, 20]$ , they contribute for the provision of the public good. The marginal monetary benefit for participant  $i$  for every token contributed by

any group member is given by  $\alpha_i$ , whereas the marginal monetary benefit from not contributing a token equals 1. Hence, from a purely monetary perspective, participant  $i$ 's dominant strategy is to contribute nothing if  $\alpha_i < 1$ . In contrast, if  $\alpha_i > 1$ , then participant  $i$ 's dominant strategy is to contribute the whole endowment. Additionally, if  $\sum_i \alpha_i > 1$ , then each token contributed increases the sum of earnings in the group, and hence, monetary surplus maximization is attained if everyone contributes all twenty tokens.<sup>6</sup>

In the punishment condition, the contribution stage of the public-goods game is followed by a punishment stage. In the punishment stage, participants are first informed of the contributions of all other group members. Thereafter, each participant  $i$  simultaneously decides how many punishment points,  $p_{ij} \in [0, 10]$ , to assign to each other participant  $j$ . Each punishment point costs the punisher one token and reduces the earnings of the punished participant by three tokens.<sup>7</sup> After the punishment decision, participants are informed of the total number of punishment points assigned to them by other members of their group. As in Fehr and Gächter (2000), participants receive no information about individual punishment behavior of other group members. At the end of a period, earnings of a participant  $i$  are given by<sup>8</sup>

$$\pi_i = 20 - c_i + \alpha_i \sum_j c_j - 3 \sum_{j \neq i} p_{ji} - \sum_{j \neq i} p_{ij}.$$

Importantly, in some treatments, groups differ in regard to the values of marginal benefits,  $\alpha_i$ , assigned to members a group. In normal groups, all members have the same  $\alpha_i = 0.5$ . In privileged groups, at the beginning of the experiment, a member is randomly chosen to receive an  $\alpha_j = 1.5$ . The two other members in a privileged group have the same  $\alpha_i = 0.5$ . Each member keeps his or her assigned marginal benefit for all ten periods. For convenience, we refer to members with an  $\alpha_i = 0.5$  as low-benefit members and to members with an  $\alpha_j = 1.5$  as high-benefit members. Clearly, from a monetary perspective, high-benefit members' dominant strategy is to contribute all their endowment. Note that the presence of a high-benefit member precisely mirrors Olson's (1965) definition of privileged groups as being groups in which at least one member is willing to bear the cost of providing some level of the public good on his own. Note that because of the presence of low-benefit members, privileged groups are also prone to underprovision of the public good. The four different treatments are summarized in table 1.

Normal and privileged groups differ only in the absence or presence of a high-benefit member instead of a low-benefit member. Thus, by comparing these group types, we can investigate the effect of the high-benefit member on overall contribution levels when punishment is or is not available. Furthermore, since low-benefit members face the same monetary incentives in both types of groups, any differences in behavior of these members across group types must be caused by the presence of the high-benefit member.

**Table 1**  
**Experimental Treatments**

	Baseline	Punishment
Normal groups	Contribution-only stage $\alpha_i = 0.5$ for all participants	Contribution and punishment stages $\alpha_i = 0.5$ for all participants
Privileged groups	Contribution-only stage $\alpha_i = 1.5$ for one participant $\alpha_i = 0.5$ for others	Contribution and punishment stages $\alpha_i = 1.5$ for one participant $\alpha_i = 0.5$ for others

Under the assumption that individuals are selfish money maximizers, nobody is predicted to contribute a positive amount to the public good in normal groups. However, the evidence from previous public-goods experiments suggests some voluntary contribution. Typically, in early periods, members contribute substantial positive amounts that decline to low levels with repetition (see, e.g., Ledyard 1995). In privileged groups, in the baseline treatment, low-benefit individuals face the same monetary incentives as members in the normal groups. In contrast, high-benefit individuals earn more money the more they contribute to the public good. Therefore, when taking only material incentives into account, it is predicted that public-good provision is higher in privileged groups than in normal groups and that high-benefit members contribute all their endowment, whereas low-benefit members contribute nothing. Note, however, that the simultaneous presence of high- and low-benefit members may interact with nonmonetary motives of voluntary contribution in a nontrivial way. This is shortly discussed below.

Since punishment is costly, and hence, not credible for rational individuals maximizing their own payoff, predictions based on this assumption are the same as without a punishment option. However, there is considerable evidence that punishment is used to enforce high and stable contribution levels (Fehr and Gächter 2000). Various models of generalized preferences have been proposed to explain contribution and punishment behavior in a range of experiments. In the following, we briefly discuss some general predictions regarding punishment and voluntary contributions of such models and relate it to our experimental setup with privileged groups.

For normal groups, the predictions of outcome-based models of inequity aversion are that contribution levels will be relatively low in the baseline treatment and can be high in the punishment treatment (e.g., Fehr and Schmidt 1999; Bolton and Ockenfels 2000). For privileged groups, the predictions are different. In the case of no punishment possibilities, any increase of contributions by the high-benefit individual increases his income and does not alter income differences. Hence, high-benefit individuals are predicted to contribute all their endowment, irrespective of their other-regarding preferences. In contrast, any increase in contribution by low-benefit individuals decreases their income and increases income differences. Consequently, low-benefit individuals are predicted not to contribute. Importantly, introducing the

punishment opportunity does not change this prediction. The reason is that in the mentioned models of outcome-based inequity aversion, the only credible motivation for costly punishment is the reduction of income differences. However, in privileged groups, increasing contributions by a low-benefit individual increases inequality. Therefore, high-benefit individuals who are inequity averse have no incentive to force low-benefit group members to contribute.<sup>9</sup> Hence, in contrast to normal groups, the opportunity to punish is not predicted to increase contributions in privileged groups.

Predictions of intention-based models of social preferences are more ambiguous. In normal groups, high and low contributions have a straightforward interpretation as being a kind or an unkind action. This provides an incentive to punish low contributors and not to punish high contributors. In privileged groups, however, it is difficult to judge the kindness of actions of high-benefit individuals. In particular, high contributions directly benefit such individuals, and hence, may be interpreted as a selfish act. Consequently, in intention-based models, the prediction for privileged groups depends critically on how intentions are modeled. For example, in Falk and Fischbacher (2006), kindness depends on relative income positions. Hence, low-benefit individuals will not regard high contributions by high-benefit individuals as kind, and high-benefit individuals will not consider low contributions by low-benefit individuals as unkind and thus deserving punishment.

In summary, purely outcome-based as well as intention-based models of social preferences suggest that introducing punishment opportunities has a weaker effect on contributions in privileged groups than in normal groups.

## Results

The computerized experiment was conducted at the CREED laboratory of the University of Amsterdam. The experiment was programmed with z-Tree (Fischbacher 2007). In total, eighty-one people participated in the experiment. Twenty-one (eighteen) participated in the baseline treatment without (with) punishment, and twenty-four (eighteen) participated in the privileged-group treatment without (with) punishment. This generated between six and eight independent observations per treatment. Each treatment was run in a separate session. A typical experimental session took one hour. On average, participants were paid out €13.76 (approx. \$17.50). More detailed experimental procedures and the instructions can be found online at this journal's Web site as well as in the authors' personal Web pages. The online materials also include the software programs used, detailed descriptive statistics, and all the data required to replicate the results.

### Voluntary Contributions

In this section, we first discuss contribution behavior at the group level. Thereafter, we zoom into the behavior of low- and high-benefit participants.

**Figure 1**  
**Contributions in the Different Groups and Treatments**

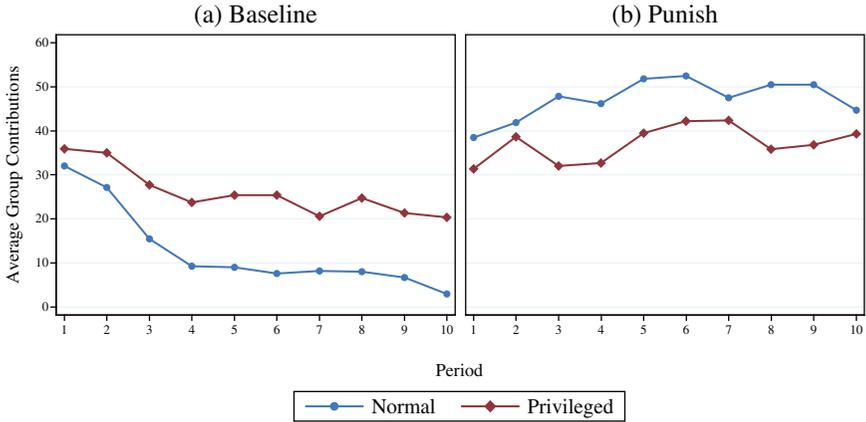


Figure 1 depicts the groups’ mean total contributions during periods in (a) the baseline treatments and (b) the punishment treatments for both group types. In keeping with evidence from comparable public-goods experiments, contributions to the public good are relatively low and decrease over time in the baseline treatment with normal groups. At a higher level, similar dynamics of contributions are observed in privileged groups. On average, normal groups contribute 12.63 tokens, whereas privileged groups contribute 26.04 tokens. A nonparametric Mann-Whitney test clearly rejects equality of distributions ( $p = .001$ ).<sup>10</sup>

In the baseline treatment, rank order correlations (Spearman’s  $\rho$ ) of contributions on periods indicate that voluntary contributions significantly decline over time in normal as well as privileged groups ( $\rho < -.462, p < .001$ ). However, compared to privileged groups, contributions decline to a lower level in normal groups. This difference may be attributed to convergence toward different equilibrium contribution levels. Indeed, in normal groups, contributions decline to (almost) zero. In privileged groups, last-period contributions are, on average, 20.37 tokens. This is indistinguishable from what is obtained when the high-benefit participant contributes everything and the two low-benefit participants contribute nothing.

In the punishment treatment, contribution levels are higher than in the baseline treatment, and they do not decline over time for both group types. Note that in contrast to the baseline without punishment, it is no longer the case that contributions are highest in privileged groups. In fact, the average total contribution of normal groups is 47.19, whereas it is only 37.08 in privileged groups. Although a Mann-Whitney test does not reject the hypothesis that average contributions come from

the same distribution ( $p = .120$ ), it is intriguing that privileged groups lose their privileged status once punishment is possible. Additionally, according to Spearman's rank-order correlations, contributions stay constant over time in privileged groups ( $\rho = .149$ ,  $p = .129$ ) but exhibit a significant upward trend in normal groups ( $\rho = .248$ ,  $p = .029$ ).

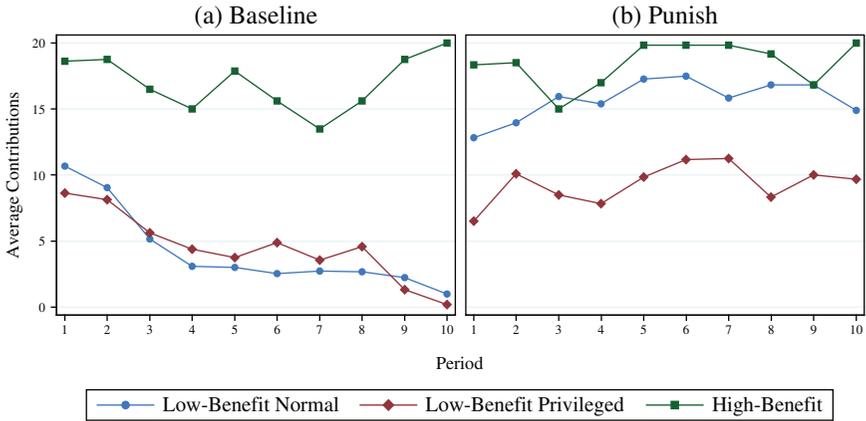
The quantitative effect of punishment on contributions also strongly differs between group types. In normal groups, the increase in contributions because of punishment is large (34.56 tokens) and statistically significant ( $p < .001$ ). For privileged groups, the increase is also statistically significant ( $p = .041$ ) but quantitatively rather small (11.04 tokens). The difference in the effectiveness of punishment is also illustrated by the fact that in normal groups, full contribution by all group members is achieved in 35.5 percent of all periods, whereas in privileged groups, it is observed only in 6.7 percent. Hence, although the opportunity to punish increases contributions in both types of groups, the magnitude of this effect is strikingly different between normal and privileged groups. Particularly, it is much stronger in groups in which all members benefit equally from the public good. We summarize the findings discussed so far in our first result.

*Result 1: Voluntary contributions in normal and privileged groups.* In the absence of punishment, there is underprovision of the public good in normal as well as privileged groups, although voluntary contributions are higher in the latter. If punishment is possible, there is still underprovision but at a lower level, and the difference between group types disappears. Punishment possibilities increase contributions in both group types, but the effect of punishment is much smaller in privileged groups.

Given the diametrically opposed material incentives for high- and low-benefit members, it is to be expected that their contribution behavior differs. Figure 2 shows that this is indeed the case. In privileged groups, high-benefit members contribute significantly more than low-benefit members in treatments with and without punishment. For high- and low-benefit members, the average contribution levels are 17.03 and 4.50 tokens, respectively, in the baseline treatment, and 18.43 and 9.32 tokens, respectively, in the punishment treatment. In both cases, equality of distribution is rejected at any conventional level of statistical significance ( $p < .005$ ).

Remarkably, high-benefit participants do not follow their dominant strategy to contribute the full endowment of twenty tokens. This behavior is particularly pronounced and statistically significant in the baseline treatment without punishment ( $p = .010$ ). Recall that for high-benefit participants, contributing the whole endowment is a dominant strategy in absolute as well as relative material terms, and it is also the best for their group's total payoff. Therefore, it is impossible to explain the observed deviations from full contribution on the basis of outcomes only. High-benefit participants act (at least to some extent) conditionally cooperatively in terms of contributions. It appears that they "punish" low-benefit participants for

**Figure 2**  
**Contributions by Low- and High-Benefit Participants in Each Treatment**



low contributions by not contributing all their endowment. To test this, we regress the contribution of the high-benefit participant on the sum of contributions of the low-benefit participants in the previous period. In addition, we control for a possible end-game effect by including a dummy variable for period ten.<sup>11</sup> The regression results show a positive and statistically significant coefficient for the lagged contributions of low-benefit participants ( $p = .033$ ) and for the last-period dummy ( $p < .001$ ). This clearly indicates conditional cooperative behavior of high-benefit participants up to the last period or periods. It also explains why their contributions decrease over time until the end-game effect sets in (see Figure 2a). In the punishment treatment, the coefficient for the lagged contributions of low-benefit participants is statistically not significantly different from zero ( $p = .789$ ). This difference between treatments is consistent with the interpretation that in the presence of the more effective explicit punishment possibility, there is no need for high-benefit participants to resort to lowering contributions as a means of disciplining low-benefit participants.<sup>12</sup> In the next section, we provide evidence showing that high-benefit participants indeed do punish low-benefit participants.

In the no-punishment treatment, low-benefit participants behave very similarly in normal and in privileged groups (see Figure 2a). Average contributions in both types of groups are almost equal (4.21 and 4.50 in normal and privileged groups, respectively) and statistically indistinguishable ( $p = .347$ ). Contributions of low-benefit participants also decline significantly over time in both types of groups (Spearman's  $\rho = -.584$  and  $-.590$ ,  $p < .001$ ). In contrast, in the punishment treatment, contributions of low-benefit participants differ starkly in normal and in

privileged groups (see Figure 2b). Average contributions amount to 15.73 tokens in normal groups but are only 9.32 tokens in privileged groups. A Mann-Whitney test clearly rejects equality of distributions ( $p = .033$ ). Interestingly, this difference in contributions is present already in the first period. In period 1, the average contribution in normal groups is 12.83, which is almost twice as high as in privileged groups, in which it is only 6.50 tokens ( $p = .004$ ). Moreover, this difference does not vanish but rather increases over time. Spearman rank-order correlations show that in normal groups, contributions exhibit a statistically significant upward trend ( $\rho = .247$ ,  $p = .028$ ), whereas no such trend can be detected in privileged groups ( $\rho = .104$ ,  $p = .215$ ).

Clearly, in privileged groups, contributions of low-benefit members are significantly higher in the punishment treatment than in the baseline treatment ( $p = .041$ ). However, in comparison to the treatment effect for normal groups, the observed increase in contributions is rather meager. For normal groups, the average increase in contributions from the baseline treatment to the punishment treatment is 11.52 tokens, whereas it is only 4.82 tokens for low-benefit members of privileged groups ( $p = .021$ ). Together, these observations indicate that the effectiveness of punishment concerning low-benefit participants strongly differs between group types to the disadvantage of privileged groups. We summarize our findings in the following result.

*Result 2: Contributions of high- and low-benefit participants.* Without punishment opportunities, although high-benefit participants contribute significantly more than low-benefit participants, there is significant underprovision of the public good by high-benefit participants. With punishment opportunities, contributions of both types of participants increase. However, contributions of low-benefit participants in privileged groups reach only 60 percent of the contribution level in normal groups.

## **Punishment**

The amount of tokens lost because of punishment, per participant and period, is similar in privileged and normal groups (3.72 and 3.23,  $p = .409$ ). In privileged groups, low-benefit members are punished slightly more than high-benefit members (4.03 vs. 3.10). However, this difference is not statistically significant ( $p = .310$ ). Over time, punishment decreases significantly in both types of groups (Spearman's  $\rho < -.365$ ,  $p < .005$ ). Allocated punishment points are also similar across types. On average, low-benefit participants in normal and in privileged groups deal out 0.54 and 0.51 punishment points, respectively. High-benefit participants in privileged groups assign slightly more punishment points (0.84). The differences are statistically not significant ( $p > .155$ ).

In the remainder of this subsection, we discuss possible determinants of punishment and how they differ, first between group types and, second, between high- and low-benefit members of privileged groups. Thereafter, we investigate if and

how high- and low-benefit members react with their public-good contributions to received punishment. This will shed light on the observed differences in contributions of low-benefit participants in normal and privileged groups.

For the analysis of possible determinants of punishment behavior, we apply Tobit regressions. The dependent variable is the amount of punishment points,  $p_{ij}$ , participant  $i$  allocates to participant  $j$ .<sup>13</sup> As explanatory variables, we use the deviation of  $j$ 's contribution from  $i$ 's contribution and the deviation of  $j$ 's contribution from the contribution of the third participant in the group  $k$ , where we allow for different coefficients for positive and negative deviations. In addition, we control for  $j$ 's contribution for cases in which all participants contributed their whole endowment (this accounts for 21 percent of all periods but only 9 percent of all punishment points) and for a time trend. For privileged groups, we run separate regressions for high- and low-benefit participants. In the regressions for low-benefit participants, we included a dummy variable if the target person  $j$  is a high-benefit participant and also interact this variable with the period of play.<sup>14</sup> The results are reported in Table 2.

We first discuss behavior of low-benefit participants in normal and privileged groups (columns 2 and 3 in the table). In line with other studies, in both groups, negative deviations from own contributions are strongly and statistically significantly punished. There is no statistically significant difference between normal and privileged groups (Wald test,  $p = .982$ ). Positive deviations are punished only in normal groups. Such "perverse" punishment of high contributors occurs at a relatively low level and is also found by other authors (see, e.g., Cinyabuguma, Page, and Putterman 2006; Gächter and Herrmann 2007; Egas and Riedl 2008). Probably the most interesting difference between groups is the punishment response toward differences in contributions between the target person  $j$  and the third person  $k$ . In normal groups, a positive deviation of  $j$  compared to  $k$  leads to a strong and significant decrease in punishment, whereas in privileged groups, no such effect is detected (Wald test for the difference between treatments,  $p = .004$ ). In privileged groups, the highest contributor is almost always the high-benefit member. Therefore, the above result indicates that relatively high contributions of such a participant are not rewarded by less punishment. Additionally, low-benefit participants appear to punish high-benefit participants simply for being high-benefit participants, especially at the beginning of the game (see table 2, column 3).

In contrast to low-benefit participants, punishment behavior of high-benefit participants is largely independent of relative contributions except that they reward full contributions of low-benefit participants with a strong and significant decrease in punishment. High-benefit participants also punish less in later periods, as indicated by the statistically significantly negative linear time trend.

*Result 3: Punishment behavior in normal and privileged groups.* In both types of groups, low-benefit members' punishment behavior is sensitive to relative

**Table 2**  
**Punishment Assigned to  $j$  by  $i$  Depending on Differences in Contributions**

Dependent Variable: Punishment Given by $i$ to $j$ ( $p_{ij}$ )	Normal Groups		Privileged Groups	
	All		Low-Benefit	High-Benefit
Contribution of $j$	0.121	(1.59)	0.120	(1.27)
Positive deviation from $c_i$	0.148***	(2.58)	-0.105	(0.93)
Negative deviation from $c_i$	0.312***	(4.38)	0.420***	(3.10)
Positive deviation from $c_k$	-0.537***	(3.76)	-0.137	(1.21)
Negative deviation from $c_k$	0.118	(1.30)	0.058	(0.55)
Fully contributing group	-4.360***	(4.74)	2.977	(1.17)
Period	-0.081	(1.15)	-0.019	(0.10)
$j$ is a high-benefit participant			5.000***	(3.39)
$j$ is a high-benefit participant $\times$ period			-0.696***	(3.68)
Constant	-3.253**	(2.17)	-6.292***	(2.60)
# Observations	360		240	
Log-likelihood	-271.964		-191.053	
$\chi^2$	124.880		38.652	
			2.721	(1.22)
			120	
			-140.377	
			56.709	
			-0.457**	(2.43)
			-0.086	(0.74)
			-0.311	(0.93)
			-0.074	(0.73)
			-0.166	(1.53)
			0.132	(0.97)
			-14.606***	(7.11)
			-0.457**	(2.43)

Note: Numbers in parentheses are  $t$ -statistics estimated with White's heteroscedasticity consistent covariance matrix estimator; asterisks indicate significance at the 1 percent level (\*\*\*), 5 percent level (\*\*), and 10 percent level (\*).

contributions. They severely punish deviations from their own contribution. Additionally, in normal groups, low-benefit members punish less if the target person contributes more than the third group member. In contrast, high-benefit members' punishment behavior is insensitive to the relative contributions of low-benefit members.

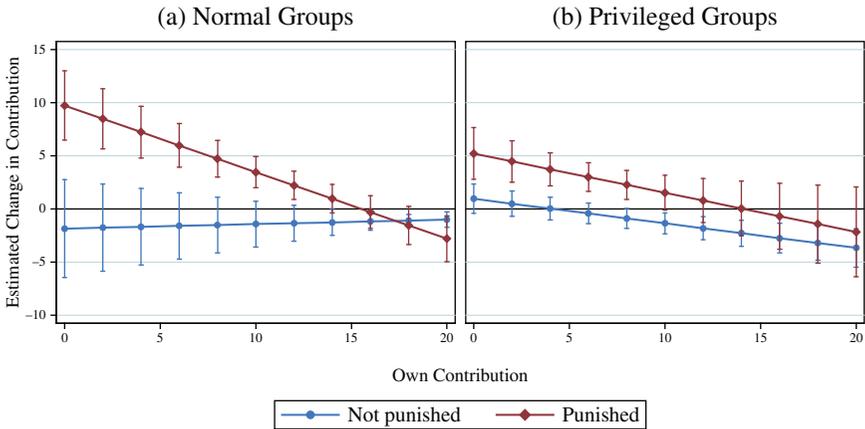
On average, high-benefit participants do not punish less than low-benefit participants. Therefore, it is likely that the unstructured punishment pattern of high-benefit participants regarding contributions of low-benefit participants plays a key role in explaining the comparatively low contributions of low-benefit participants in privileged groups.

Another important aspect for the effectiveness of punishment in increasing contributions is the response of those being punished. Recall that even when punished, the motivation of low-benefit participants to contribute to the public good might be different between normal and privileged groups. For example, recall that in models of inequity aversion, low-benefit types have little incentive to contribute in privileged groups because any increase in contributions decreases their earnings and increases inequality, especially vis-à-vis the high-benefit member. To investigate the response to being (not) punished and its potentially differential characteristics across group types, we have estimated the mean change in contributions from period  $t$  to period  $t + 1$ , depending on whether the low-benefit participant was punished in period  $t$  or not (controlling for the participant's contribution and the contribution of others in period  $t$ ). For each group type, we ran two separate ordinary least squares (OLS) regressions, one for periods in which the low-benefit participant was punished and another for periods in which this participant was not punished.

Figure 3a visualizes the estimation results for normal groups and Figure 3b for privileged groups.<sup>15</sup> The lines show the estimated change in contributions as a function of the (low-benefit) participant's contribution in period  $t$  for cases in which the participant was punished (black  $\diamond$ ) or not punished (gray  $\circ$ ) in period  $t$ . When not punished, low-benefit participants have, on average, a tendency to decrease their contributions in both types of groups, particularly if they contributed a large amount. In both normal and privileged groups, members who are not punished and contributed more than fourteen tokens significantly decrease their contribution (Wald tests  $p < .046$ ), although one can see that the decrease is not large in an absolute sense. If we compare the estimated change in contribution between group types, we find it to be statistically different only at very high contribution levels (contributions above seventeen tokens, Wald tests  $p < .033$ ). In this case, participants in privileged groups decrease their contribution more if not punished.

When punished, low-benefit participants respond similarly in both types of groups only if they contributed a large amount of their endowment. For contribution levels of fourteen tokens or more, participants do not respond to punishment by significantly increasing contributions, in normal or in privileged groups (Wald tests,  $p > .123$ ).

**Figure 3**  
**Mean Change in Contribution by Low-Benefit Participants,**  
**Depending on Previous Contributions and Punishment**



In stark contrast, low contributors respond to punishment by significantly increasing their contribution, and they do so differently in normal and privileged groups. For instance, after being punished, participants who free ride completely (zero contribution) increase contributions about twice as much in normal groups as in privileged groups. Wald tests confirm that for contribution levels of nine tokens or less, punished low-benefit participants increase their contributions significantly less in privileged groups than in normal groups ( $p < .047$ ). Thus, over a large range of contribution levels, punishment produces a much weaker positive response in privileged groups than in normal groups. We summarize these findings in the following result.

*Result 4: Response to punishment in normal and privileged groups.* In both types of groups, low-benefit participants who contribute low amounts respond with a significant increase in contributions if punished. Over a large range of low contribution levels, this effect is significantly and substantially weaker in privileged groups.

The fact that high-benefit participants in privileged groups exhibit a relatively unstructured punishment behavior (Result 3), together with the above finding that strongly free-riding low-benefit participants exhibit a much weaker response to punishment in privileged groups, can explain why the opportunity to punish has a much smaller effect on the contributions of low-benefit members in privileged groups compared to normal groups (Result 2).

## Efficiency and Inequality

The socially beneficial effect of the opportunity to punish free riders is that it increases contributions to the public good, and thereby, also total monetary earnings. The downside is that if costly punishment is actually administered, it destroys resources, and hence, decreases total monetary surplus. In this section, we examine how punishment affects earnings of low- and high-benefit members in normal and privileged groups.

Taken across all periods, in the baseline treatment with normal groups, average earnings per member and per period are equal to 22.10 tokens, which is statistically not significantly different from the 23.67 when punishment is possible ( $p = .366$ ). However, the dynamics of earnings reveal a clear divergence between the baseline and the punishment treatment. In the baseline treatment, earnings significantly decrease over time, whereas they show a clear upward trend in the punishment treatment (Spearman's  $\rho = -.584$ ,  $p < .001$ , and  $\rho = .322$ ,  $p = .006$ ). Furthermore, earnings in the second half (periods six through ten) of the public-goods game are significantly larger in the punishment treatment compared to the baseline treatment (21.11 tokens and 24.86 tokens,  $p = .017$ ).

In privileged groups, the effect of punishment on earnings is more ambiguous. When taken across all periods, average earnings per participant and period are similar in treatments without and with punishment (33.01 and 33.58 tokens,  $p = .426$ ). Earnings also show a decreasing trend in the baseline treatment and an increasing trend in the punishment treatment (Spearman's  $\rho = -.463$ ,  $p < .001$ , and  $\rho = .254$ ,  $p = .025$ ). However, unlike in normal groups, in later periods (six through ten), the earnings in the punishment treatment are only marginally significantly larger than in the baseline treatment (31.25 tokens without punishment and 36.23 tokens with punishment,  $p = .091$ ). The reason for this difference is that compared to normal groups, low-benefit participants in privileged groups increase their contributions less in the punishment treatment.

Punishment not only affects total earnings but also how these earnings are distributed. A natural and simple measure of inequality is the standard deviation of earnings per period within each group. In normal groups, the average standard deviation across periods is 4.76 in the baseline treatment and 3.43 when punishment is possible. The difference is statistically not significant ( $p = .183$ ). In privileged groups, however, punishment significantly increases inequality. Whereas in the baseline treatment, the average standard deviation is 11.56, it increases to 23.69 in the punishment treatment ( $p = .041$ ).

The underlying reason for increased inequality in privileged groups is that punishment has a stark differential effect on the earnings of low- and high-benefit participants. High-benefit participants gain from the punishment institution: their earnings increase from 42.01 tokens in the baseline treatment to 52.38 tokens in the punishment treatment ( $p = .071$ ). In contrast, low-benefit participants in privileged

groups do not gain at all from the punishment institution. In effect, compared to the baseline treatment, their earnings are significantly lower in the punishment treatment (28.51 tokens without punishment and 24.18 tokens with punishment,  $p = .004$ ). If we compare earnings of low-benefit participants in normal and privileged groups, we find that in the treatment without punishment, low-benefit participants earn significantly more in privileged groups ( $p < .001$ ), but this is no longer the case in the treatment with punishment ( $p = .409$ ). In other words, low-benefit participants “exploit” the high-benefit participant when punishment is not possible but are no longer able to do so when punishment is introduced. This is important as it illustrates that privileged groups are not necessarily characterized by the exploitation of the “rich by the poor,” as is usually claimed (Olson 1965; Sandler 1992). We summarize the findings of this section in our final result.

*Result 5: Efficiency and inequality.* In normal groups, the opportunity to punish leads to higher overall earnings with no effect on their distribution. In privileged groups, punishment does not significantly increase total earnings but strongly increases inequality at the cost of low-benefit members. Consequently, with punishment, low-benefit members do not benefit from being part of a privileged group.

## Discussion and Conclusion

In this article, we provide evidence that Olson’s (1965) privileged groups are not as privileged as conjectured, especially if punishment is possible. In line with standard theoretical predictions, contributions are higher in privileged groups than in normal, nonprivileged groups, and underprovision of the public good is observed in both group types. Surprisingly, in privileged groups, this is not only because of free riding by low-benefit members but also because of nonoptimal contributions of members whose dominant strategy is to contribute fully. We observe that high-benefit members reciprocate to low contributions of low-benefit members by lowering their contributions, despite the fact that contributing fully maximizes absolute and relative individual earnings as well as group earnings. When a punishment option is introduced, privileged groups lose their privileged status completely. Total contributions in normal and privileged groups are virtually identical. In addition, the introduction of punishment significantly increases earnings inequality in privileged groups as high-benefit members avoid being exploited by low-benefit members. We identify the following behavioral features that account for this relative disadvantage of privileged groups. With a punishment option, low-benefit members in privileged groups increase their contributions to a much smaller extent than their counterparts in nonprivileged groups, in which all have the same incentive to free ride. The reason for this is twofold. First, high-benefit members exhibit an unstructured punishment pattern, and second, free-riding low-benefit members

increase their contributions in response to punishment to a much smaller extent in privileged groups than in normal groups.

Formal models of other-regarding preferences capture, to some extent, the possibility that punishment is less effective in privileged groups than in normal groups. However, the theoretical argument supporting this result is not in agreement with our empirical observations. These models predict that low-benefit members do not contribute because high-benefit members who have a preference for equality are not willing to punish them. In contrast, our results show that high-benefit members do punish free-riding low-benefit members (albeit in a relatively unstructured way) but that free-riding low-benefit members are very reluctant to increase contributions despite the fact that they are punished. This is an interesting finding because it demonstrates that members in one role (high-benefit) may be willing to undertake a costly action (punishing), although it increases earnings inequalities, but that members in another role (low-benefit) are unwilling to take a costly action (increasing contributions) because it increases inequality. This strongly suggests that an individual's disutility from punishment depends not only on the received punishment points but also on who is punishing and whether punishment is perceived as fairly administered. To account for such behavior, a more general theoretical model of behavior is called for. Our results clearly indicate that such a model should take into account important heterogeneity parameters and the roles in which members are immersed.

Our results also highlight the importance of a common notion of "fair" or "kind" behavior for mechanisms such as decentralized punishment to be effective. In homogeneous groups, contributions to a public good are unambiguously kind because they increase everyone's payoff at a cost for the contributing individual. In privileged groups, however, high contributions by high-benefit members may be perceived as selfish (because they increase the contributor's earnings), unfair (because they increase earning inequalities), or kind (because they increase the earnings of others). Low contributions of low-benefit members can similarly be interpreted in different ways. The fact that high-benefit members punish low-benefit members but low-benefit members only reluctantly increase contributions is consistent with self-serving interpretations of the "correct" fairness norm (see, e.g., Babcock, Camerer, and Samuel 1995; Gächter and Riedl 2005). That diverging perceptions of what constitutes fair behavior can have detrimental effects on cooperation is also suggested by recent research on emotions and social behavior. For instance, Hopfensitz and Reuben (2007) and Reuben and van Winden (2007) find that sanctions enforce cooperative behavior only if people feel they have transgressed a social norm, and hence, deserve punishment (see also Bowles and Gintis 2005). If low-benefit members in privileged groups conceive that they do not deserve sanctioning, a weak response, as observed in our experiment, may be the consequence. In privileged groups, such diverging perceptions of fair behavior are the likely reason for costly punishment's being relatively ineffective and wasteful.

In an early study, Johnson and Libecap (1982) argue theoretically that group heterogeneity considerably complicates the private enforcement of cooperative behavior. The results of our article add to this finding the experimental evidence that the behavioral effectiveness of an institution in overcoming social dilemmas can crucially depend on group characteristics, an observation also in line with results showing that sanctioning institutions can have very different effects on members' behavior with diverging cultural backgrounds (Carpenter and Cardenas 2007; Gächter, Herrmann, and Thöni 2008). This implies that the design of effective—formal as well as informal—enforcement institutions asks for a proper understanding of the interplay of heterogeneous group characteristics, as in privileged groups, behavioral inclinations and biases, and incentives. Therefore, a promising line for future research is the endogenous formation of groups and institutions in the presence of heterogeneity. For example, if free association to groups is possible, low-benefit members might be more reluctant to punish high-benefit members if doing so induces their leaving the group. Additionally, if punishment is governed endogenously through voting, low-benefit members might respond to punishment more strongly with increasing contributions if punishment from high-benefit members can be prohibited.

## Notes

1. One of the best known examples is the Linux operating system. Benkler (2002) discusses additional examples.

2. Underprovision is a particular problem for the mostly assumed summation and weakest-links technologies of publicness. For public goods provided with a best-shot technology, privileged groups would, in theory, overcome to a large extent the underprovision problem (see Sandler 1992).

3. The reasons why individuals may value consumption of the public good differently are numerous. It could be simply a difference in monetary returns. For instance, an individual who owns a large plot of land would benefit much more from a regional irrigation system than an individual who owns only a small plot of land. It may also be that people perceive the importance of the public good differently. For instance, some neighbors more than others might enjoy the existence of a neighborhood swimming pool. Voluntary contributors in commons-based production organizations are another example. They likely receive (perceived or expected) benefits in excess of the costs from contributing to the project. The benefits may range from pure pleasure from contributing to the project to higher likelihood of getting a contract in traditional firms with experience in such commons-based organizations.

4. Evidence of voluntary contribution without punishment can be found in Ledyard (1995). Examples of studies on decentralized punishment include Ostrom et al. (1992), Fehr and Gächter (2000), Carpenter (2006), and Egas and Riedl (2008).

5. The same holds for various types of distributional preferences, such as a concern for the income of the least well off (Rawls 1971). On the other hand, if members care for overall efficiency, low contributions by low-benefit members are clearly undesirable. For more discussion on distributional preferences, see Charness and Rabin (2002) and Engelmann and Strobel (2004).

6. In the experiment, all participants are informed that they will interact for exactly ten rounds with the same individuals. Furthermore, they know the value of  $\alpha_i$  of all group members, but neither during nor after the experiment do they get to know their identities.

7. In line with other studies (e.g., Fehr and Gächter 2000), an upper limit for the amount of punishment points that could be allocated is introduced. The main reason is to avoid giving participants with

high earnings much more punishment power (as punishment is funded out of their own earnings). Note also that participants who are left with fewer than twenty tokens after the contribution stage can still use all their punishment points but would incur a loss in the respective period. Participants, however, cannot be punished below zero. See Note 8.

8. In case a participant makes losses, these are calculated as follows:  $\pi_i = \max[0, 20 - c_i + \alpha_i \sum_j c_j - 3 \sum_{j \neq i} p_{ji}] - \sum_{j \neq i} p_{ij}$ .

9. Certainly, as long as earnings end up equal, punishment can be used to enforce other cooperation levels. It would imply that in equilibrium, low-benefit individuals punish high-benefit individuals simply to reduce income differences. However, all such equilibria are Pareto-dominated by the equilibrium in which low-benefit individuals do not contribute.

10. Unless otherwise noted, for between-treatment comparisons, we use a Mann-Whitney test, which tests the hypothesis that two independent samples are drawn from the same population. For within-treatment comparisons, we use Wilcoxon's signed-ranks test for matched pairs. If not otherwise indicated, we apply one-sided test statistics and use group averages across all periods as independent observations. All reported  $p$  values are exact values.

11. We use Tobit regressions, censoring contributions at the maximum of twenty tokens. In the baseline treatment, the resulting coefficients are 0.75 (0.35) for the low-benefit participants' lagged contributions and 89.49 (8.49) for the last-period dummy ( $\chi^2 = 118.48, p < .001$ ). In the punishment treatment, they are  $-0.05$  (0.20) and 64.57 (12.52), respectively ( $\chi^2 = 34.12, p < .001$ ). Numbers in parentheses are robust standard errors (White 1980).

12. Another possibility for underprovision is spiteful behavior, as reported by Saijo and Nakamura (1995), in groups in which  $\alpha_i > 1$  for all group members. However, this fails to explain the observed relation with the contribution behavior of low-benefit participants. Brands and Schram (2001) also find a small degree of underprovision, which they attribute to errors. However, because of their design, they cannot explore the conditional-cooperation possibility.

13. As the amount of punishment points that could be allocated is constrained, we run the Tobit regressions censoring  $p_{ij}$  at 0 and 10 points. All reported standard errors are based on robust estimates, and dependence of observations within groups is accounted for by clustering on each period and group (White 1980). Note that endogenous censoring could occur for participants with earnings of fewer than twenty tokens after the contribution stage, because they might be not willing to punish (strongly), as they would incur losses. However, situations with potential (actual) losses occurred only in 1.94 percent (0.28 percent) of all cases.

14. If we drop these two explanatory variables concerning the high-benefit participant, all reported results stay quantitatively similar and qualitatively the same. The regression coefficients of the other explanatory variables change only marginally, and all reported (non)significances remain.

15. Denoting  $\Delta c_i$  as the change in contributions from period  $t$  to  $t + 1$ ,  $c_i$  as the contribution of member  $i$  in period  $t$ , and  $c_{-i}$  as the average contribution of the other two members in period  $t$ , the resulting regression equations are ( $t$ -statistics—in parenthesis—are calculated by clustering on each group and period): (1) for punished participants in normal groups,  $\Delta c_i = 9.39$  (2.81)  $- 0.63c_i$  (4.47)  $+ 0.02\hat{c}_{-i}$  (0.11); (2) for not-punished participants in normal groups,  $\Delta c_i = -4.07$  (1.76)  $+ 0.042c_i$  (0.28)  $+ 0.14\hat{c}_{-i}$  (1.28); (3) for punished participants in privileged groups,  $\Delta c_i = 0.30$  (0.08)  $- 0.37c_i$  (2.04)  $+ 0.36\hat{c}_{-i}$  (1.18); and (4) for not-punished participants in privileged groups,  $\Delta c_i = -1.75$  (1.28)  $- 0.23c_i$  (3.01)  $+ 0.20\hat{c}_{-i}$  (1.58). Figure 3 is drawn with mean values for  $\hat{c}_{-i}$ .

## References

- Babcock, Lina, Colin F. Camerer, and Samuel I. Samuel. 1995. Biased judgments of fairness in bargaining. *American Economic Review* 85:1337-43.
- Benkler, Yochai. 2002. Coase's Penguin, or, Linux and the nature of the firm. *Yale Law Journal* 112:369-446.

- Bolton, Gary E., and Axel Ockenfels. 2000. A theory of equity, reciprocity, and competition. *American Economic Review* 90:166-93.
- Bowles, Sam, and Herbert Gintis. 2005. Pro-social emotions. In *The economy as a complex evolving system III: Essays in honor of Kenneth Arrow*, ed. Lawrence E. Blume and Steven N. Durlauf, 337-67. Oxford, UK: Oxford University Press.
- Brandts, Jordi, and Arthur Schram. 2001. Cooperation and noise in public goods experiments: Applying the contribution function approach. *Journal of Public Economics* 79:399-427.
- Buckley, Edward, and Rachel Croson. 2006. Income and wealth heterogeneity in the voluntary provision of linear public goods. *Journal of Public Economics* 90:935-55.
- Cardenas, Juan C. 2003. Real wealth and experimental cooperation: Experiments in the field lab. *Journal of Development Economics* 70:263-89.
- Carpenter, Jeffrey P. 2006. The demand for punishment. *Journal of Economic Behavior & Organization* 62:522-42.
- Carpenter, Jeffrey, and Juan C. Cardenas. 2007. An inter-cultural examination of cooperation in the commons. Working paper, Middlebury College.
- Chan, Kenneth S., Stuart Mestelman, Robert Moir, and R. Andrew Muller. 1996. The voluntary provision of public goods under varying income distributions. *Canadian Journal of Economics* 29:54-59.
- . 1999. Heterogeneity and the voluntary provision of public goods. *Experimental Economics* 2:5-30.
- Charness, Gary, and Mathew Rabin. 2002. Understanding social preferences with simple tests. *Quarterly Journal of Economics* 117:817-69.
- Cherry, Todd L., Stephan Kroll, and Jason F. Shogren. 2005. The impact of endowment heterogeneity and origin on public good contributions: Evidence from the lab. *Journal of Economic Behavior & Organization* 57:357-65.
- Cinyabuguma, Matthias, Talbot Page, and Louis Putterman. 2006. Can second-order punishment deter perverse punishment? *Experimental Economics* 9:265-79.
- Egas, Martijn, and Arno Riedl. 2008. The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of the Royal Society B—Biological Sciences* 275:871-78.
- Engelmann, Dirk, and Martin Strobel. 2004. Inequality aversion, efficiency and maximum preferences in simple distribution experiments. *American Economic Review* 94:857-69.
- Falk, Armin, Ernst Fehr, and Urs Fischbacher. 2008. Testing theories of fairness: Intentions matter. *Games and Economic Behavior* 62:287-303.
- Falk, Armin, and Urs Fischbacher. 2006. A theory of reciprocity. *Games and Economic Behavior* 54:293-315.
- Fehr, Ernst, and Simon Gächter. 2000. Cooperation and punishment in public goods experiments. *American Economic Review* 90:980-94.
- Fehr, Ernst, and Klaus M. Schmidt. 1999. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114:817-68.
- Fischbacher, Urs. 2007. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10:171-78.
- Fisher, Joseph, Mark R. Isaac, Jeffrey W. Schatzberg, and James M. Walker. 1995. Heterogeneous demand for public goods: Effects on the voluntary contributions mechanism. *Public Choice* 85:249-66.
- Gächter, Simon, and Benedikt Herrmann. 2007. The limits of self-governance in the presence of spite: Experimental evidence from urban and rural Russia. Discussion paper 2007-11, CeDEx. University of Nottingham.
- Gächter, Simon, Benedikt Herrmann, and Christian Thöni. 2008. Antisocial punishment across societies. *Science* 319:1362-67.
- Gächter, Simon, and Arno Riedl. 2005. Moral property rights in bargaining with infeasible claims. *Management Science* 51:249-63.
- Hopfensitz, Astrid, and Ernesto Reuben. 2007. The importance of emotions for the effectiveness of social punishment. Discussion paper 05-075. Amsterdam, the Netherlands: Tinbergen Institute.

- Isaac, Mark R., James M. Walker, and Susan Thomas. 1984. Divergent evidence on free riding: An experimental examination of possible explanations. *Public Choice* 43:113-49.
- Johnson, Ronald N., and Gary D. Libecap. 1982. Contracting problems and regulation: The case of fishery. *American Economic Review* 72:1005-22.
- Ledyard, John O. 1995. Public goods: A survey of experimental research. In *The handbook of experimental economics*, ed. John H. Kagel and Alvin E. Roth, 111-94. Princeton, NJ: Princeton University Press.
- Olson, Mancur. 1965. *The logic of collective action*. Cambridge, MA: Harvard University Press.
- Ostrom, Elinor, Roy Gardner, and James Walker. 1994. *Rules, games, and common-pool resources*. Ann Arbor: University of Michigan Press.
- Ostrom, Elinor, James M. Walker, and Roy Gardner. 1992. Covenants with and without a sword: Self-governance is possible. *American Political Science Review* 86:404-17.
- Palfrey, Thomas R., and Jeffrey E. Prisbrey. 1996. Altruism, reputation and noise in linear public good experiments. *Journal of Public Economics* 61:409-27.
- . 1997. Anomalous behavior in public goods experiments: How much and why? *American Economic Review* 87:829-46.
- Rawls, John. 1971. *A theory of justice*. Cambridge, MA: Harvard University Press.
- Reuben, Ernesto, and Frans van Winden. 2007. Fairness and shame in the power to take. Discussion paper 05-014. Amsterdam, the Netherlands: Tinbergen Institute.
- Sadrieh, Abdolkarim, and Harrie A. A. Verbon. 2006. Inequality, cooperation, and growth: An experimental study. *European Economic Review* 50:1197-222.
- Saijo, Tatsuyoshi, and Hideki Nakamura. 1995. The spite dilemma in voluntary contribution mechanism experiments. *Journal of Conflict Resolution* 39:535-60.
- Sandler, Todd. 1992. *Collective action: Theory and application*. Ann Arbor: University of Michigan Press.
- van Dijk, Frans, Joep Sonnemans, and Frans van Winden. 2002. Social ties in a public good experiment. *Journal of Public Economics* 85:275-99.
- Visser, Martine, and Justine Burns. 2006. Bridging the great divide in South Africa: Inequality and punishment in the provision of public goods. Working paper. Göteborg, Sweden: Göteborg University.
- White, Halbert. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48:817-38.