UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

WORKING PAPERS IN ECONOMICS

No 551

Cooperation in teams: the role of identity, punishment and endowment distribution

Qian Weng
Fredrik Carlsson

January 2013

ISSN 1403-2473 (print)
ISSN 1403-2465 (online)

Department of Economics
School of Business, Economics and Law at University of Gothenburg
Vasagatan 1, PO Box 640, SE 405 30 Göteborg, Sweden
+46 31 786 0000, +46 31 786 1326 (fax)
www.handels.gu.se info@handels.gu.se

Cooperation in teams: the role of identity, punishment and endowment distribution

Qian Weng[*], Fredrik Carlsson[†]

Abstract

Common identity and peer punishment have been identified as important means to reduce free riding and to promote cooperation in teamwork settings. This paper examines the relative importance of these two mechanisms, as well as the importance of income distribution in team cooperation. In a repeated public good experiment, conditions vary among different combinations of homogenous or heterogeneous endowment, strong or weak identity, and absence or presence of peer punishment. We find that without punishment, strong identity can counteract the negative impact of endowment heterogeneity on cooperation. Moreover, punishment increases cooperation irrespective of income distribution and identity strength, and cooperation is similar across all treatments with punishment. These findings provide important implications for management policy makers in organizations: implementing *ex ante* income heterogeneity within teams should be done with caution, and a very strong peer punishment mechanism is more effective in enhancing cooperation over common identity when both are viable.

Keywords: Endowment distribution; identity; punishment; cooperation; public goods experiment

JEL classification: C91; D63; H41; M54

[*] Department of Economics, School of Business, Economics and Law, University of Gothenburg, Sweden. Vasagatan 1, 405 30 Gothenburg, Sweden. Tel: +46 31 786 1351, Fax: +46 31 786 1326, Email: qian.weng@economics.gu.se.

[†] Department of Economics, School of Business, Economics and Law, University of Gothenburg, Sweden. Vasagatan 1, 405 30 Gothenburg, Sweden. Tel: +46 31 786 4174, Fax: +46 31 786 1043, Email: fredrik.carlsson@economics.gu.se.

Acknowledgements: We thank Maria Bigoni, Martin Dufwenberg, Haoran He, Martin Kocher, Peter Martinsson, Matthew Rabin, Amrish Patel, Matthias Sutter, Måns Söderbom, and seminar and conference participants at the University of Gothenburg, the 2012 Asia-Pacific ESA Conference in Xiamen, and the 2012 International ESA Conference in New York for helpful discussions and comments. We are grateful to Xiujuan Zhang for assistance in conducting the experiment. Financial support from Formas through the COMMONS program and from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged. Any views expressed in this paper are those of the authors.

1. Introduction

Teams have been increasingly viewed as an important way to enhance the efficiency of organizations and firms. One common underlying philosophy of successful teams is to foster cooperation among their members (Che and Yoo, 2001). However, organizations face several challenges to efficient teamwork. The benefits of working as a team may be undercut by the incentives to free ride, which cannot be completely controlled through formal contracts if compensation is based on team output rather than personal input (Alchian and Demsetz, 1972). Experiments have shown that cooperation typically cannot be sustained by intrinsic altruistic motives alone (e.g., Andreoni, 1995; Fischbacher et al., 2001; Fischbacher and Gächter, 2010). Rather, (centrally) building a common identity among employees and allowing (decentralized) mutual monitoring and sanctioning of team members have been considered effective attempts to discipline free riding and to promote cooperation in teamwork settings. Social identity theory (Tajfel and Turner, 1979, 1985) has received growing interest in the organizational literature (see, e.g., Akerlof and Kranton, 2000, 2005, 2008). A number of experiments have shown that salient identification with an organization or a team can increase cooperation (e.g., Eckel and Grossman, 2005; McLeish and Oxoby, 2008).[1] Punishment, in terms of both pecuniary consequences such as reduced salaries and non-pecuniary ones such as social pressure and disapproval, has also been shown to be an important means to increase cooperation (Fehr and Gächter, 2000b, 2002; Masclet et al., 2003; Kandel and Lazear, 1992; Mas and Moretti, 2009).[2]

An additional aspect of teams is that they are often composed of individuals who are unequal in productivity, ability, and motivation. Payments tend to be differentiated partly to induce greater individual effort and partly to incentivize employees contributing to team output to stay away from distinct outside options (Balafoutas et al., 2010). Previous experiments investigating the role of income distribution in cooperation have shown mixed results – Cherry et al. (2005) report a negative effect of heterogeneity on aggregate cooperation, Chan et al. (1996), Visser and Burns (2006), and Prediger (2011) find the opposite, and Hofmeyr et al. (2007) find no significant difference. However, when it comes to individual behavior in unequal income teams, low-income people are ubiquitously found to cooperate relatively more than their high-income counterparts (e.g., Buckley and Croson, 2006; van Dijk et al., 2002). Some studies further explore whether the power of punishment in norm enforcement in symmetric settings can carry over to asymmetric settings, and obtain an affirmative answer that punishment in heterogeneous populations shows similar or even higher effectiveness (e.g., Nikiforakis et al., 2010; Visser and Burns, 2006; Prediger, 2011). Reuben and Riedl (2013) examine

[1] A closely related strand of literature focusing on identity conflict between two groups in general find favoritism toward ingroup members and discrimination against outgroup ones in terms of cooperation (e.g., Charness et al., 2007; McLeish and Oxoby, 2007), coordination (e.g., Chen and Chen, 2011; Chen et al., 2010), social preferences (e.g., Chen and Li, 2009), and norm enforcement (e.g., Ruffle and Sosis, 2006; Bernhard et al., 2006; Goette et al., 2006; Goette et al., 2012).

[2] However, some other studies question the beneficial effects of punishment (Egas and Riedl, 2008; Houser et al., 2008; Abbink et al., 2010), and some even find anti-social punishment directed at relatively cooperative people (e.g., Herrmann et al., 2008; Nikiforakis, 2008; Cinyabuguma et al., 2006).

the emergence and enforcement of contribution norms (i.e., an absolute efficiency rule and several relative contribution rules) to public goods in homogeneous and heterogeneous groups, and attain different results depending on the availability of punishment.

In this paper we study the three dimensions affecting team cooperation: identity, punishment, and income distribution. While identity and punishment have been shown to increase cooperation, the potential interaction and relative importance of these two means have not, to the best of our knowledge, been investigated. Clearly, when deciding on team incentives and organization, the relative importance and interaction between identity and punishment is central. One goal of the present paper is to provide some evidence on this issue. Our experimental design allows us to study all three dimensions in isolation and the interaction between them. In addition, there are only a few studies looking at the interaction between identity and punishment, but the results are inconclusive. Chen and Li (2009) find that individuals are less likely to punish an ingroup member for misbehavior, whereas McLeish and Oxoby (2007) find that unfair offers to ingroup members incur greater use of costly punishment than those to outgroup members.

Moreover, although the effect of income distribution on team cooperation both in the absence and presence of punishment has been investigated, whether and how income distribution affects the role of identity has not. Thus, an additional goal of this paper is to demonstrate whether and how income distribution affects the function of identity and punishment, respectively, as a cooperation norm enforcement mechanism. We use a three-stage laboratory experiment to address these questions. At the first stage, a common identity is induced in the strong identity treatments via a face-to-face identity-building activity involving all subjects; this stage is absent in the weak identity treatments. At the second stage, subjects individually solve a quiz whose results indicate their differences in productivity. At the third stage, a two sub-stage repeated linear public good game framed as a team production problem following the design of Fehr and Gächter (2002) is played in fixed teams of four. At the first sub-stage, subjects are randomly assigned into teams. Each team member decides individually how to allocate an endowment between individual work and team work We distinguish two team income distribution environments: in one, endowment is homogenously distributed among team members irrespective of their differences in productivity; in the other, each member is given a different endowment according to her productivity ranking within the team, yet the total team endowment is the same as with homogenous endowment teams. To compare the difference in behavior with and without punishment, we add a second sub-stage in half of the treatments where the subjects are informed of each team member's contribution to the team work and can reduce their earnings by assigning costly punishment points.

The main results from the experiment are as follows. At the team level, when punishment is not possible, endowment heterogeneity negatively affects cooperation. However, strong identity can counteract this negative impact. The introduction of punishment successfully raises and sustains cooperation in all treatments, i.e., in both homogenous and heterogeneous teams and with weak and strong identities. Cooperation is similar in all treatments with punishment, and so is the punishment inflicted. Within the heterogeneous teams, low endowment individuals always show the greatest degree of cooperation relative to endowment, and punish least intensively. Strong identity increases contribution rates at all endowment levels without punishment, but not with punishment. Nor does strong identity have any impact on punishment assignment at each endowment level.

2. Experimental design

The experiment uses a 2×2×2 design. In one dimension we vary the endowment distribution by giving subjects in a team the same or different endowment in order to create homogenous or heterogeneous teams. In the second, we make the strength of identity strong or weak by conducting or not conducting an identity-building activity. The third dimension concerns whether or not subjects have the opportunity to punish other team members. This generates eight different combinations of conditions, each of which is a treatment of the experiment as summarized in Table 1. The experiment is conducted in three stages. The first stage is an identity-building stage. The second stage is an endowment determination stage. The third stage is a repeated linear public good game.

<Table 1 about here>

The identity-building stage was included only in the four treatments with strong identity. A "human knot" game was played with all subjects in one session in another room before they entered the laboratory. Subjects stood shoulder to shoulder, in a circle, facing each other. First they were asked to form a knot by lifting both hands and reaching across the circle to hold the hands of two other subjects who were not standing directly beside them, left hand to left hand and right hand to right hand. After ensuring that a knot had been constructed, subjects were asked to untangle the knot to form one or a couple of circles without letting go of any hands. Anyone who let go of a hand was required to immediately grab the same hand again. The game lasted for approximately ten minutes regardless of whether or not the knot was successfully untangled. The reason for choosing such an identity-building activity was that it is a typical activity conducted in orientation or training programs in real-world organizations to promote mutual understanding, raise common objectives, and yield organizational belongingness among new members or members from different departments. Communication was allowed during the course of the game. The experimenters observed that the game sparked extensive communication among team members. After finishing the identity-building activity, the subjects were led to the laboratory. In the four treatments with weak identity, subjects entered the laboratory directly once everybody had arrived, yet they did have a chance to meet each other while waiting for the experiment to start.

The rest of the experiment was conducted in the laboratory, where subjects were first seated in partitioned computer terminals and then given written instructions while the experimenter read the instructions aloud. At the second stage, subjects individually solved a six-minute quiz consisting of 20 general knowledge questions. The quiz performance determined the endowment levels of subjects in the heterogeneous teams for the public good game. That is, the more questions that were answered correctly, the higher the endowment level. The quiz was used to create feelings of entitlement over the endowment (see, e.g., Hoffman and Spitzer, 1985; Gächter and Riedl, 2005) and to justify the fairness of inequalities within the heterogeneous teams. To enable comparison across treatments, this stage was also conducted in the homogenous endowment treatments, although in this treatment the endowment levels were not affected by quiz performance.

At the third stage, four subjects out of 24 in one session were randomly assigned to a team and played a public good game framed as a team production problem for 10 periods. The reason for using partner rather than stranger matching was that wanted to mimic the situation where people usually worked in relatively fixed teams and interacted repeatedly over a period of time.[3] The subjects knew that their teams consisted of themselves and three other individuals, whereas their identities were kept anonymous throughout the experiment.

At the beginning of each period, each subject was endowed with a fixed amount of experimental currency units (ECUs), E_i. They decided simultaneously and without communication how to allocate the endowment between individual work and team work (i.e., the public good). By freely choosing an amount to contribute to the team work, c_i, where $0 \leq c_i \leq E_i$, the remaining endowment, $E_i - c_i$, was automatically considered the allocation to the individual work. Each ECU that a subject kept for individual work generated one ECU for herself, whereas the payoff from the team work was 50% of the team's total contribution. That is, the marginal per capita return (MPCR) from a contribution to the public good was equal to 0.5. In the heterogeneous teams, members were endowed with 80, 60, 40, and 20 ECUs, respectively, according to their quiz performance ranking within a team. In the homogenous teams, each member was endowed with 50 ECUs. Subject i's period payoff was given by

$$\pi_i^c = (E_i - c_i) + 0.5 \sum_{h=1}^{4} c_h \qquad (1)$$

In the treatments with punishment, a second sub-stage was added. Subjects were informed of the other team members' proportion of endowment contributed, i.e., contribution rate, and were given the opportunity to punish each other.[4] To punish, member i could assign punishment points to member j

[3] See Botelho et al. (2009) for a critical review of the experimental literature on partner and stranger matching. The authors further compare behavior under random strangers and perfect strangers matching (where subjects meet only once) in a public good experiment, and find a significantly lower proportion of subjects contributing in a random strangers than in a perfect strangers protocol.

[4] We reveal relative contribution rather than absolute contribution amount to preserve the anonymity of endowment levels and to prevent individual reputation building. We are aware of the possible different impacts posed by different feedback formats on cooperation and efficacy of punishment as pointed out by Nikiforakis (2010). The author considers three feedback formats – subjects receive information about each team member's contribution, earnings, or both contribution and earnings before making punishment decisions – and finds that earnings feedback leads to significantly less cooperation and lower efficiency than contribution feedback. Nevertheless, this paper follows the most common format used in public goods experiments with peer punishment to adopt the contribution feedback. A potential drawback is that a relative contribution norm is exogenously imposed. Brekke et al. (2012) compare the cooperation effect of three ways of framing the decision variable in a multi-period threshold public goods experiment with unequally endowed participants: absolute

within the same team, p_{ij}, $i \neq j$. The punishment decisions were made simultaneously and without communication. However, punishment points were not costless. Each assigned punishment point cost the punished member 3 ECUs and the punishing member 1 ECU. Hence, subject i's payoff at the end of the period was given by

$$\pi_i^p = \pi_i^c - \sum_{\substack{j=1 \\ j \neq i}}^{4} p_{ij} - 3 \sum_{\substack{j=1 \\ j \neq i}}^{4} p_{ji} \quad (2)$$

Equation (2) implies that a subject could have a negative payoff in a given period. To reduce the probability of this, we constrained the income reduction associated with received punishment to not exceed the income from the contribution sub-stage, i.e., $3\sum_{\substack{j=1 \\ j \neq i}}^{4} p_{ji} \leq \pi_i^c$. In addition, a subject could at most distribute 25 points to each other team member, i.e., $p_{ij} \leq 25, j = 1,2,3,4, j \neq i$. Despite the restrictions, negative payoff could still occur in some extreme cases where subjects had little income from the contribution sub-stage, attracted considerable punishment, and also decided to punish heavily. Negative period payoff occurred in three out of 1,920 possible cases (192 subjects ×10 periods); these losses were covered by cumulative payments from previous periods. As is common in public goods experiments with punishment, each subject was also given a one-off lump-sum payment of 50 ECUs to pay for any total loss that might be incurred during the experiment. In our experiment, however, nobody incurred such a loss.

The endowment distribution, the payoff functions, the duration of the experiment (10 periods), and the instructions were common knowledge to all participants in each treatment. Before the commencement of actual decision making, the subjects were required to answer control questions to ensure that they had understood the features of the game correctly. In the treatment without punishment, at the end of each period the subjects were informed of their team's total contribution, their own income, and the contribution rates of other team members in the current period. In the treatments with punishment, at the end of each period the subjects were reminded of the income from the contribution sub-stage and the associated cost of the punishment points they had assigned. They were also informed of the punishment they received in total, the associated income reduction, as well as their final income from that period as given by (2). Each of the four subjects was randomly assigned an identification number from 1 to 4 to identify her actions in a given period, but to prevent the possibility of individual reputation formation the numbers were randomly shuffled across periods.

The experiment was conducted using z-Tree (Fischbacher, 2007) in the experimental laboratory at Beijing Normal University in May and June 2011. This university is located in the center of Beijing

contributions, contributions relative to endowments, and amounts of endowments kept (i.e., in terms of the effects of contributions on final payoffs). They find no significant difference in absolute contribution amounts between the absolute and relative framings for both high and low endowment subjects at conventional levels. Their finding to some extent mitigates the norm imposing concern in our experiment. Moreover, we are aware of the different views on fair contribution rules. See Reuben and Riedl (2013) and Brekke et al. (2012) for detailed discussions.

and has approximately 20,000 full-time students. The subjects were recruited via announcements on a bulletin board system (BBS) and bulletin boards in teaching and accommodation buildings at the university. In total, we had observations from 384 subjects[5], 48 for each treatment. All subjects were allowed to participate in only one session, and they did not know about any treatments other than the one in which they participated. To control for experimenter effect, the same two individuals, who were unknown to the participants, ran all sessions. To keep the outcome of the experiment anonymous, subjects were informed at the beginning that they would be paid confidentially and individually in another room and that they would leave the laboratory successively so that they would not meet and communicate with other subjects after completing the session. The final earnings from the experiment totaled the sum of the period payoffs at an exchange rate of 1 ECU to 0.1 Chinese yuan (CNY) plus a show-up fee of 10 CNY. The experiment lasted an average of about 76 (104) minutes in the treatments without (with) punishment, including above-described stages and a post-experimental survey covering questions on demographics, academic background, past donation behavior, and perceptions about their team in the experiment. The subjects on average earned 80.9 (94.6) CNY[6] in the treatments without (with) punishment, including the show-up fee in all treatments and the lump-sum payment in the treatments with punishment.

3. Behavioral hypotheses

This section develops behavioral hypotheses on how income distribution and identity strength affect cooperation and punishment behavior based on theories and existing empirical evidence. Assuming that all people are rational and self-interested exclusively in their material payoffs, the standard economic model predicts that people will not contribute in a linear public good game, irrespective of the income distribution, saliency of identity or punishment opportunities. However, there is considerable experimental evidence that such a model fails to predict actual behavior under many circumstances, suggesting that people are motivated by other-regarding preferences and that concerns for fairness and reciprocity cannot be overlooked in social interactions.

3.1 Contributions when punishment is not possible

It is clear from the game theoretic prediction that endowment heterogeneity[7] is completely irrelevant for contribution levels in a linear public good game, and previous experiments testing this prediction have obtained mixed results. The first hypothesis, which has been tested several times before, is thus:

[5] All subjects were Chinese citizens and university students with various academic majors.

[6] The average exchange rate in May and June 2011 was 1USD=6.48 CNY. The average hourly wage for university students in Beijing at the time of the experiment was approximately 50 CNY.

[7] Apart from endowment heterogeneity, heterogeneity can also be represented by different demands for a public good, i.e., different MPCR of group members (e.g., Isaac and Walker, 1988; Fisher et al., 1995; Carpenter et al., 2009), or different fixed lump-sum payments such as show-up fees (e.g., Anderson et al., 2008).

Hypothesis 1.1: *Team average contribution rates will be the same in homogenous (Homo-Weak-NoPunish) and heterogeneous (Hetero-Weak-NoPunish) teams when identity is weak and there is no punishment.*

One implication from social identity theory is that once an individual has gone through a cognitive change and emotional investment process to categorize herself as part of a unit with shared goals, values, and norms, her behavior tends to conform to the norms of that unit, which could lead to a higher degree of team cohesion and more effective teamwork (Lembke and Wilson, 1998). Recent research shows that inducing a salient common organizational or team identity has a positive impact on pro-social behavior. Hence, we expect the induced strong identity to enhance cooperation in homogeneous teams:

Hypothesis 1.2: *Team average contribution rate will be higher in homogenous teams with strong identity (Homo-Strong-NoPunish) than in homogenous teams with weak identity (Homo-Weak-NoPunish) when there is no punishment.*

The effect of strong identity on cooperation has not been explored so far in a heterogeneously endowed team setting. However, provided that individual average endowment is the same in heterogeneous teams as in homogenous teams, we expect a similar effect of strong identity in heterogeneous teams:

Hypothesis 1.3: *Team average contribution rate will be higher in heterogeneous teams with strong identity (Hetero-Strong-NoPunish) than in heterogeneous teams with weak identity (Hetero-Weak-NoPunish) when there is no punishment.*

A related question is if the inducement of a strong identity has different effects in homogenous and heterogeneous teams. There is no theory or evidence suggesting that the increase in contribution rates should be the same or different; hence, conditional upon that we cannot reject Hypothesis 1.1, we state the following hypothesis:

Hypothesis 1.4: *Team average contribution rates will not differ between homogenous (Homo-Strong-NoPunish) and heterogeneous (Hetero-Strong-NoPunish) teams when identity is strong and there is no punishment.*

In heterogeneous teams, the question is as well if low and high endowment subject contributes the same (in absolute or relative terms) or not. If subjects were inequity averse (Fehr and Schmidt, 1999), then subjects with higher endowment would contribute a larger share of their endowment to the public good than low endowment subjects (Buckley and Croson, 2006). However, if individuals are trying to maximize the net benefits from mutual cooperation, lower endowment individuals have stronger incentives to contribute more proportionally, at least initially, to signal their intent to commit to cooperative behavior, hoping higher endowment individuals will reciprocate (Prediger, 2011). This is because the MPCR is the same across endowment levels; lower endowment subjects can realize greater net gains if everyone contributes. Existing evidence provides support for the latter argument (e.g., Buckley and Croson, 2006; Prediger, 2011). Hence,

Hypothesis 1.5: *In heterogeneous teams, subjects with lower endowment will give more in relative terms than subjects with a higher endowment when identity is weak (Hetero-Weak-NoPunish) and there is no punishment.*

Are there reasons to believe that this would change with the inducement of a strong identity? Inducing a strong identity could strengthen the inequity aversion and hence increase the contributions of higher endowment subjects. At the same time, the potential envy from lower endowment subjects could be reduced, hence increasing their contributions; for example, Chen and Li (2009) show that participants show a 93% decrease in envy when matched with an in-group member than with an out-group member. Since there are two potential effects that go in opposite directions, we do not have a definite hypothesis regarding contribution rates of low and high endowment subjects with strong identity.

Concerning the effect of strong identity on contribution rates of subjects at each endowment level in heterogeneous teams, we would expect the positive impact at team average to carry over to the individual level:

Hypothesis 1.6: *Contribution rate of subjects at each endowment level will be higher in heterogeneous teams with strong identity (Hetero-Strong-NoPunish) than in heterogeneous teams with weak identity (Hetero-Weak-NoPunish) when there is no punishment.*

3.2 Contributions when punishment is possible

A well-established finding from repeated public goods experiments is that the existence of peer punishment increases and sustains cooperation. Inequity-averse subjects who cooperate could be sufficiently upset by the payoff inequality so that they are willing to sanction the free riders even at their own cost (Fehr and Schmidt, 1999). Free rider could perceive the threat of punishment to be credible and thus would tend to cooperate (Fehr and Gächter, 2000a). We hence hypothesize:

Hypothesis 2.1: *The introduction of peer punishment will increase team average contribution rates in both homogenous (Homo-Weak-NoPunish vs. Homo-Weak-Punish) and heterogeneous teams (Hetero-Weak-NoPunish vs. Hetero-Weak-Punish) when identity is weak.*

What if identity is strong? The answer depends on the relative strengths of strong identity and punishment on contribution, and the potential interaction between the two. If strong identity increases contribution rates substantially, there is little room for an additional effect of introducing punishment. And vice versa, if the existence of punishment opportunities increases contribution rates substantially, there will be little effect of identity on contribution behavior. At the same time, there could be reinforcement between the two. Our hypotheses are that both punishment and strong identity affect contribution rates even in the presence of each other, thus

Hypothesis 2.2: *The introduction of peer punishment will increase team average contribution rates in both homogenous (Homo-Strong-NoPunish vs. Homo-Strong-Punish) and heterogeneous teams (Hetero-Strong-NoPunish vs. Hetero-Strong-Punish) when identity is strong.*

Hypothesis 2.3: *Team average contribution rates in teams with strong identity (Homo-Strong-Punish and Hetero-Strong-Punish) will be higher than in teams with weak identity (Homo-Weak-Punish and Hetero-Weak-Punish) even with the introduction of peer punishment irrespective of endowment distribution.*

The empirical evidence on the difference in effects of punishment on contributions between homogenous and heterogeneous teams is scarce. Visser and Burns (2006) and Prediger (2011) do find that with punishment, relative cooperation is significantly higher in heterogeneously endowed groups, but there is no clear theoretical prediction. We therefore do not make any conjecture about team average contribution rates between homogenous and heterogeneous teams irrespective of identity strength.

What differences regarding contribution rates of heterogeneously endowed individuals could one expect when punishment is introduced? Multiple underlying motives outlined previously, such as inequity aversion, reciprocity, net benefit maximization from mutual cooperation, envy, altruism, etc. may be at work, and how people punish norm violators will also affect their reactions in terms of contributions in the next period. It is difficult to conjecture whether and in which direction contribution rates will differ within the *Hetero-Weak-Punish* or *Hetero-Strong-Punish* treatments. Thus, we again do not derive any definite hypothesis. Since it is unclear how identity and punishment interplay, we do not hypothesize about the contribution rates of subjects at each endowment level in heterogeneous teams between strong and weak identities.

3.3 Punishment behavior

It has been well documented that a substantial fraction of subjects are willing to engage in costly punishment of free riders (e.g., Fehr and Gächter, 2000b; Nikiforakis and Normann, 2008; Anderson and Putterman, 2006; Carpenter, 2007). Negative emotions toward free riders triggered by payoff inequality (i.e., inequity aversion) is the main motive behind this altruistic punishment (Fehr and Gächter, 2002; Fuster and Meier, 2010). Will punishment behavior change if endowment becomes unequal? Based on the hypothesis that contributions will not differ between homogenous and heterogeneous teams, we hypothesize that there will be no difference in punishment assignment either:

Hypothesis 3.1: *Team average punishment assignment will not differ between homogenous and heterogeneous teams when identity is weak (Homo-Weak-Punish vs. Hetero-Weak-Punish) or strong (Homo-Strong-Punish vs. Hetero-Strong-Punish).*

The effect of identity on punishment behavior reflects how negative reciprocity works. Chen and Li (2009) find that individuals are forgiving to ingroup members for misbehavior, whereas McLeish and Oxoby (2007) find that unfair offers to ingroup members incur greater use of punishment than those to outgroup members. The contradictory directions do not lead us to a definite hypothesis regarding the difference in team average punishment point assignment between teams with weak and strong identity irrespective of endowment distribution.

How will punishment behavior differ among individuals with heterogeneous endowments? On the one hand, lower endowment subjects face higher costs of sanctioning (including the direct cost of awarding punishment and the possible additional cost of retaliation) relative to their endowment than their higher endowment counterparts, which may lead to lower frequency and intensity of punishment (Visser and Burns, 2006; Prediger, 2011). On the other hand, if negative emotions are invoked by perceived unfairness of endowment inequality, although endowment distribution is justified by quiz performance, lower endowment subjects may punish higher endowment members more vehemently even when higher endowment subjects have contributed more (Prediger, 2011). An additional incentive for lower endowment subjects to assign more punishment is to discipline free riders and promote mutual cooperation through which they can obtain greater net gains (Visser and Burns, 2006; Prediger, 2011). Nevertheless, previous research shows similar punishment intensity among low and high endowment subjects (Visser and Burns, 2006; Prediger, 2011). We hence do not have a definite hypothesis. Given the indefinite conjecture of the effect of identity on team average punishment behavior, it is also hard to predict how strong identity will affect punishment behavior of subjects at each endowment level.

4. Results

In this section, we analyze the impact of endowment distribution and identity strength on contributions to the team work when punishment is absent and present, and on punishment behavior.

4.1 Contributions when punishment is not possible

Figure 1 depicts the evolution of average contribution rates over the 10 periods for all treatments. For the four treatments without punishment, consistent with previous experimental findings, average contributions start from 30% to 50% of subjects' endowment. They all rise in the early periods and then decline, although the peaks appear at different points in time and the rates of change differ across treatments. As the experiment progresses, contribution rates in the Hetero-Weak treatment becomes substantially lower than those of the other three treatments without punishment.

<Figure 1 about here>

Table 2 reports the average contribution rates over all 10 periods depending on treatment (first row) and endowment level (last four rows). Throughout the paper, for team average, the unit of observation is team mean over all periods; for subject average, the unit of observation is subject mean over all periods. *High, Second, Third*, and *Low* refer to endowment levels with 80, 60, 40, and 20 ECUs, respectively. In the four treatments without punishment, team average contribution rates in *Homo-Weak*, *Homo-Strong*, and *Hetero-Strong* are at least 50% higher than that in the *Hetero-Weak* treatment (see left panel first row). Comparisons among the last four entries within columns (2) and (4) suggest that average contribution rates vary among subjects with different endowment levels. Friedman two-way analysis of variance by ranks tests reject the null hypothesis that contribution rates of different endowment levels are from the same population under either identity strength (p-values<0.01).

<Table 2 about here>

Since individual cross-period differences and the data structure are not taken into consideration in the summary statistics, we turn to a formal statistical analysis by regressing individual contribution rate on treatment variables of the experiment. Since contribution rates range between zero and one in each period, i.e., truncated from both above and below, we use a double-censored tobit model. We also include team random effects to account for the interaction of team members across periods. We construct one dummy variable for each endowment distribution and identity strength combination, i.e., *Hetero-Weak*, *Homo-Weak*, *Hetero-Strong*, and *Homo-Strong*, equal to one if the observation comes from the respective treatment and zero otherwise. Period dummies are also included to control for time order effects. To investigate how contribution rates differ among subjects with different endowment levels and identity strengths in the heterogeneous teams, we use one separate binary dummy variable for each endowment level and identity strength combination, i.e., *Weak-High*, *Weak-Second*, *Weak-Third*, *Weak-Low*, *Strong-High*, *Strong-Second*, *Strong-Third*, and *Strong-Low*. *Hetero-Weak* and *Weak-Low* are excluded from the regressions as the reference groups.

Table 3 presents the regression results. Models (1) and (2) are estimated for the four treatments without punishment. Model (1) includes both homogenous and heterogeneous teams to investigate the aggregate treatment effect, and model (2) includes only heterogeneous teams to study the endowment effect. The top panel reports the average marginal effects of the independent variables.[8] In model (1), when identity is weak, homogenous teams on average contribute 13.4 percentage points more than heterogeneous teams. This significant difference rejects Hypothesis 1.1, but is in line with the finding in Cherry et al. (2005). It might be explained by the perceived unfairness of endowment heterogeneity, which reduces the possibility for a team contribution norm to emerge. When identity becomes strong, the significant difference between homogenous and heterogeneous teams disappears, which means that we cannot reject Hypothesis 1.4 (see bottom panel (i)). This suggests that building a strong identity can counteract the negative impact of endowment heterogeneity on contributions. The bridging of the difference is because strong identity significantly and substantially increases contribution rates in heterogeneous teams (15.0 percentage points), but it does not have a significant effect on contributions in homogenous teams (see bottom panel (ii)). Thus we reject Hypothesis 1.2 but not 1.3.

<Table 3 about here>

[8] Using McDonald and Moffitt (1980) decomposition, the marginal effect of contribution rates, $\left(\frac{c}{E}\right)_{i,t}$, is calculated as
$$\frac{\partial E(\left(\frac{c}{E}\right)_i|\mathbf{x})}{\partial x_j} = \frac{\partial \Pr\left(0<\left(\frac{c}{E}\right)_i<1|\mathbf{x}\right)}{\partial x_j} \cdot E\left(\left(\frac{c}{E}\right)_i \middle| \mathbf{x}, 0 < \left(\frac{c}{E}\right)_i < 1\right) + \Pr\left(0 < \left(\frac{c}{E}\right)_i < 1|\mathbf{x}\right) \cdot \frac{\partial E(\left(\frac{c}{E}\right)_i|\mathbf{x},0<\left(\frac{c}{E}\right)_i<1)}{\partial x_j} + \frac{\partial \Pr\left(\left(\frac{c}{E}\right)_i=1|\mathbf{x}\right)}{\partial x_j} \cdot 1.$$

When breaking heterogeneous teams down to various endowment levels (model (2)), we observe that the marginal effects of the endowment level dummies under weak identity are negative and statistically significant, indicating that low endowment subjects on average always contribute the largest proportion of endowment compared to their team members with higher endowments under weak identity. Thus, we cannot reject Hypothesis 1.5.[9] The result is in line with previous findings (e.g., Buckley and Croson, 2006; Prediger, 2011) and supports the net benefit maximization from mutual cooperation argument. That low endowment subjects also contribute relatively more in the *Hetero-Strong* treatment is interesting (see bottom panel (iii)-(v)). It is primarily this group that increases contribution rates when identity becomes strong compared to when it is weak (see marginal effect of *Strong-Low*). If we compare contribution rates of subjects at the three other endowment levels when identity is weak and strong (see bottom panel (vi)-(viii)), we only find a significant effect for the second endowment level, i.e., for this group the contribution rate is higher when identity is strong. This means that we cannot reject Hypothesis 1.6 for two out of four endowment levels.

4.2 Contributions when punishment is possible

In this section, we examine whether and how contribution behavior changes when peer punishment is introduced. Comparing team average contribution rates in each column between the left and right panel of Table 2 (first row), we find that contribution rates are drastically and significantly higher in the treatments with punishment for all endowment distribution and identity strength combinations (Mann-Whitney U test, *p*-value=0.024 for *Homo-Weak*; *p*-value=0.002 for *Hetero-Weak*; *p*-value=0.002 for *Homo-Strong*; *p*-value=0.043 for *Hetero-Strong*). Consequently, we cannot reject Hypotheses 2.1 and 2.2. The same pattern can be found for subjects at the same endowment level when we compare the last four entries of column (2) with (6) and (4) with (8) (Mann-Whitney U test, all *p*-values<0.1). However, the magnitude of the increase varies considerably across treatments and endowment levels. The strong effect of punishment is not unique to our experiment. Other studies using partner matching with similar MPCR and punishment cost-effectiveness as ours obtain a similar increase in contribution rates when punishment is introduced (e.g., Herrmann et al., 2008; Reuben and Riedl, 2013). As shown in Figure 1, average contribution rates in the treatments with punishment are all at a higher level after a similar starting point as in the treatments without punishment, and overall appear to be increasing over time. The evolution of contribution rates follows a similar pattern among the four treatments with punishment except *Homo-Strong*, which outstands the others from the beginning of the experiment. The divergence between treatments with and without punishment over time confirms the general finding from the existing literature that the presence of punishment opportunities is effective in improving and sustaining cooperation. However, the average contribution rates do not reach the maximum possible level in any of the four treatments with punishment. The

[9] However, it should be noted that the pattern is different if we look at absolute contribution amounts: higher endowment subjects always contribute a greater absolute amount.

proportion of full contributions are 40%, 35%, 47%, and 33% of the total observations in *Homo-Weak*, *Hetero-Weak*, *Homo-Strong*, and *Hetero-Strong*, respectively, suggesting that the contribution "ceiling" is not reached by the majority in any of the four treatments. Comparisons among the last four entries within columns (6) and (8) suggest that average contribution rates vary among subjects with different endowment levels. Friedman two-way analysis of variance by ranks tests reject the null hypothesis that contribution rates of different endowment levels come from the same population under either identity strength (p-values<0.01).

Models (3) and (4) in Table 3 present the regression results for the four treatments with punishment. Model (3) includes both homogenous and heterogeneous teams, and model (4) includes only heterogeneous teams. The same model and specification as in the treatments without punishment are applied. In model (3), there is no statistically significant difference between homogenous and heterogeneous teams, or between weak and strong identity (see marginal effects of *Homo-Weak*, *Hetero-Strong*, and bottom panel (i)-(ii)).[10] Thus, we can reject Hypothesis 2.3. These results are in sharp contrast to the findings for the treatments without punishment, where both endowment distribution and identity have significant effects. One possible explanation for why strong identity does not further raise contributions in either endowment distribution may be that peer punishment alone is effective enough to push contribution rates to a high level and a strong common identity will not exert any further influence. This finding suggests that under our experimental design, peer punishment dominates common identity when both are viable in the effect on cooperation enhancement.

Regarding various endowment levels within heterogeneous teams (model (4)), we find that like in absence of punishment, low endowment subjects on average always contribute a significantly greater proportion of the endowment than subjects with higher endowments, under both weak and strong identities (see marginal effects of *Weak-High*, *Weak-Second*, and *Weak-Third*, and bottom panel (iii)-(v), except (v), where the difference is insignificant at conventional levels).[11] These results could hence be interpreted by similar motives as those underlying behavior in heterogeneous teams without

[10] The finding of no difference in contribution rates among the four treatments with punishment may raise a concern that subjects contribute a high share anyway due to the presence of punishment and do not respond to different endowment distributions and identity strengths. Besides the proportion of full contributions in each treatment with punishment, we also look at a less restrictive concept of the "ceiling," which is an arbitrarily high contribution rate but not 1. To test the presence of such a "ceiling effect" in contribution rates, we split the observations in the treatments with punishment into two subsamples – one with team average contribution rate above the median of each treatment and one below. The average contribution rate in the above median subsample is 0.89, 0.88, 0.93, and 0.86 for *Homo-Weak*, *Hetero-Weak*, *Homo-Strong*, and *Hetero-Strong*, respectively. These are rather high rates. We also rerun model (3) of Table 3 for each subsample separately. We find that in the below median subsample, the regression results are qualitatively identical to those in the full sample, whereas in the above median subsample the team average contribution rate in the *Homo-Strong* treatment is significantly higher than in the *Hetero-Strong* and *Homo-Weak* treatments at conventional levels. This suggests that subjects in the above median subsample respond to the treatments and do not contribute anyway at a high level, and the finding of no difference in contribution rates among the four endowment distribution and identity strength combinations is actually driven by the result from the below median subsample.

[11] Absolute contribution amounts are also always higher from higher endowment subjects.

punishment. However, strong identity does not exert a significant impact at any endowment level (see marginal effect of *Strong-Low*, and bottom panel (vi)-(viii)). We hence can reject Hypothesis 2.6 but not 2.5.

4.3 Punishment behavior

We now turn to the analysis of punishment behavior. Table 4 reports the average number of punishment points assigned by subject *i* to *j* in the same team depending on treatment and endowment level. The first row shows that the average number of punishment points allocated is around 0.5 out of a maximum of 25 in all four treatments. Punishment occurs in 1,071 out of 5,760 possible cases (196 subjects times 3 targets per period times 10 periods), and boils down to 22% of 1,440 possible cases in *Homo-Weak*, 17% in *Hetero-Weak*, 19% in *Homo-Strong*, and 16% in *Hetero-Strong*. The last four entries in columns (2) and (4) demonstrate that there are some variations in punishment assignment across endowment levels within heterogeneous teams. Friedman two-way analysis of variance by ranks tests reject the null hypothesis that punishment points assigned by subjects of different endowment levels are from the same population under either identity strength (p-values<0.01).

<Table 4 about here>

Some regularities regarding punishment behavior have been identified from previous public goods experiments (see, e.g., Fehr and Gächter, 2000b; Carpenter and Matthews, 2009; Nikiforakis et al., 2010). In particular, punishment is mostly directed toward team members contributing less than the team average, and the severity of punishment increases with the difference between the contribution of the target and the team average. In order to investigate this, we conduct a regression analysis of punishment assignment behavior. To account for the large number of zero punishment and a handful of full punishment, we again apply the double-censored tobit model with team random effects. In addition to the treatment variables and period dummies, we include the following three independent variables in some of the regressions to capture the regularities in punishment behavior: others' average contribution rate, absolute negative deviation, and positive deviation. *Others' average contribution rate* is the average value of the team members' contribution rates of subject *j* (i.e., $\sum_{h\neq j}\left(\frac{c}{E}\right)_{h,t}/3$), excluding that of subject *j*. *Absolute negative deviation* is the absolute value of the actual deviation of subject *j*'s contribution rate from the others' average in case her own contribution is below the average (i.e., $\max\{0, \frac{\sum_{h\neq j}\left(\frac{c}{E}\right)_{h,t}}{3} - \left(\frac{c}{E}\right)_{j,t}\}$). This variable is zero if the subject's own contribution rate is equal or above the others' average. *Positive deviation* (i.e., $\max\{0, \left(\frac{c}{E}\right)_{j,t} - (\sum_{h\neq j}\left(\frac{c}{E}\right)_{h,t})/3\}$) is constructed analogously. [12]

[12] We are aware of other possible punishment regularities within one's own group such as that based on individual contribution comparison between the punisher and the target. That is, individuals often punish team members who contribute proportionally less than they do. Although we choose to follow the literature and use

Table 5 reports the regression results. Models (1) and (2) are estimated using both homogenous and heterogeneous teams, whereas models (3) and (4) only using heterogeneous teams. Models (1) and (3) only include treatment variables, whereas models (2) and (4) also account for the punishment regularities. The top panel reports the average marginal effects of the independent variables.[13] In model (1), the model results indicate that there is no statistically significant difference in punishment assignment at team level between homogenous and heterogeneous teams (see marginal effect of *Homo-Weak*, and bottom panel (i)). Similar punishment levels across different endowment distributions support Hypothesis 3.1, and this is consistent with our previous finding that average contribution rates are similar as well. That punishment does not vary with identity strength (see marginal effect of *Hetero-Weak* and bottom panel (ii)) is at odds with the findings of both Chen and Li (2009) and McLeish and Oxoby (2007), indicating that negative reciprocity is not a motive affecting behavior in our sample. In model (2), when punishment regularities are accounted for, the treatment effects are generally the same as in model (1) (except bottom panel (i), which becomes marginally significant). The three regularity variables are all statistically significant with expected signs. The negative marginal effect of *Others' average contribution rate* suggests that less punishment is used when a high common team contribution standard has already been established. The positive marginal effect of *Absolute negative deviation* and negative marginal effect of *Positive deviation* show that the extent of punishment increases (decreases) with the size of absolute negative (positive) deviation of the target's from the others' average contribution rate. Figure 2 provides some visual evidence for the above findings.

<Table 5 about here>

The patterns in punishment behavior discussed above are at an aggregate level for all four treatments with punishment. In order to check whether these patterns are common across treatments, we examine them separately for each treatment. Table 6 reports the regression results for tobit models with both upper and lower censoring and team random effects. Entries in the top panel are the average marginal effects of the independent variables. Following Goette et al. (2012), we test the equality of marginal effects across treatments in the bottom panel using two-sided z-tests for single parameter

---

the most commonly assumed punishment regularity since Fehr and Gächter (2000b) as based on team average contribution comparison, qualitatively similar results are obtained when we instead control for *individual absolute negative deviation* (i.e., $\max\{0, \left(\frac{c}{E}\right)_{i,t} - \left(\frac{c}{E}\right)_{j,t}\}$) and *individual positive deviation* (i.e., $\max\{0, \left(\frac{c}{E}\right)_{j,t} - \left(\frac{c}{E}\right)_{i,t}\}$).

[13] Using McDonald and Moffitt (1980) decomposition, the marginal effect of punishment, $p_{i,t}$, is calculated as

$$\frac{\partial E(p_i|\mathbf{x})}{\partial x_j} = \frac{\partial \Pr(0 < p_i < 25|\mathbf{x})}{\partial x_j} \cdot E(p_i|\mathbf{x}, 0 < p_i < 25) + \Pr(0 < p_i < 25|\mathbf{x}) \cdot \frac{\partial E(p_i|\mathbf{x}, 0 < p_i < 25)}{\partial x_j} + \frac{\partial \Pr(p_i = 25|\mathbf{x})}{\partial x_j} \cdot 25$$

comparison and χ^2-tests for parameter vector comparison.[14] In all treatments, the marginal effect of *Absolute negative deviation* is positive and highly significant, i.e., the more an individual's contribution rate falls below the others' average, the more she gets punished. The tests comparing two marginal effects show no significant difference across treatments (see bottom panel (ii)). *Others' average contribution rate* exerts a negative and significant effect in all treatments only except in *Hetero-Weak*, where the effect is not significant at conventional levels. However, the marginal effects do not differ between any endowment distribution and identity strength combinations (see bottom panel (i)). In contrast, *Positive deviation* has a significant negative impact only in the *Homo-Strong* treatment. It may hence be this result that drives the variable significance at the aggregate level in Table 5 column (2). Tests on the equality of γ_3 show that subjects' reaction to positive deviation differs significantly between the *Homo-Strong* and *Hetero-Strong* and between the *Homo-Strong* and *Homo-Weak* treatments (see bottom panel (iii)). This confirms the visual finding from Figure 2 that punishment level in *Homo-Strong* reduces to zero when positive deviation exceeds 0.4, whereas in the other three treatments punishment keeps at a low level for all positive deviation intervals. The difference in *Positive deviation* may also contribute to the overall difference in punishment patterns between the *Homo-Strong* and *Hetero-Strong* treatments (see bottom panel (iv)).

<Table 6 about here>

Models (3) and (4) of Table 5 present the regression results on punishment assignment by subjects with different endowment levels in the heterogeneous teams. In model (3), the marginal effects of the endowment level dummies under both weak and strong identities are positive and statistically significant, except the second level under weak identity (see marginal effects of *Weak-High*, *Weak-Second*, and *Weak-Third* and bottom panel (iii)-(v)). This indicates that higher endowment subjects tend to use punishment more intensively than low endowment subjects. Our results contrast those in Visser and Burns (2006) and Prediger (2011), who find no significant difference between low and high endowment subjects. Thus, punishment decreases with the relative cost of sanctioning (see, e.g., Anderson and Putterman, 2006; Nikiforakis and Normann, 2008), and is not income inelastic which is in contrast with the findings in Carpenter (2007). Neither perceived unfairness of endowment heterogeneity nor expected net gains from mutual cooperation is a motivation underlying our results. Similar to at team average level, strong identity does not have a significant impact on punishment at any endowment level (see bottom panel (vi)-(viii)). When punishment regularities are accounted for (model (4)), punishment assignment responds to *Others' average contribution rate* and to *Absolute*

[14] For single parameter tests, we calculate $= \frac{\gamma_j - \tilde{\gamma}_j}{\sqrt{\Sigma_{jj} + \tilde{\Sigma}_{jj}}}$, where γ_j and $\tilde{\gamma}_j$ are the two parameters of interest from the two regressions, and Σ_{jj} and $\tilde{\Sigma}_{jj}$ are the corresponding main diagonal elements in the variance-covariance matrix. Since the two parameters come from two separate regressions, their covariance by construction is zero. z follows a standard normal distribution under the null of equality. For parameter vector tests, we calculate the analogous test statistic $\chi^2 = (\boldsymbol{\gamma} - \tilde{\boldsymbol{\gamma}})'(\boldsymbol{\Sigma} + \tilde{\boldsymbol{\Sigma}})^{-1}(\boldsymbol{\gamma} - \tilde{\boldsymbol{\gamma}})$, where $\boldsymbol{\gamma}$ is a column vector. χ^2 follows a Chi-squared distribution with *k* degrees of freedom, where *k* is the number of variables in $\boldsymbol{\gamma}$.

negative deviation in a similar fashion as that in the pooled sample with both homogenous and heterogeneous teams. However, *Positive deviation* does not have a significant impact on the size of punishment in heterogeneous teams, which is consistent with the results (γ_3) in columns (2) and (4) of Table 6.

5. Conclusions

How to foster cooperation in organizations when free-riding incentives prevail and individual members are diverse in for example ability and motivation is an important economic problem. In this paper, we have investigated the relative importance of identity and punishment under homogenous and heterogeneous income distributions on contribution rates to a team public good. There are three key findings. First, when punishment is not possible, endowment heterogeneity negatively affects cooperation, yet strong identity can counteract this negative impact. However, strong identity does not have any effect on cooperation in homogenous teams. One possible explanation for this difference is that contribution rates are already very high in homogenous teams, and therefore the impact of a strong identity is weakened. This finding complements the literature of induced identity by showing that not only the saliency of identity matters for its effect on cooperation, but only when the level of cooperation is sufficiently low. Second, the introduction of punishment successfully raises and sustains cooperation in both homogenous and heterogeneous teams, and under both weak and strong identities. Moreover, cooperation behavior is similar across all treatments with punishment, and so is the punishment inflicted. Thus, the effect of punishment is so strong that identity and income distribution play no role under our experimental design. Third, within the heterogeneous teams, lower endowment individuals contribute more in relative terms but less in absolute terms, and they punish less and hence receive greater net gains, which to some extent equalizes the *ex ante* endowment inequality through repeated interactions.

This study should be viewed as a first step toward considering the interactive effects of income distribution, identity, and punishment on cooperation. A natural extension would be to conduct the same experiment with real employees from real-world workplaces to test the external validity of our results. Moreover, our findings on identity strength might depend on the design method used, i.e., induced identity, and on the activity used to induce identity. It would be interesting to use natural identities within existing social groups or primed natural social identities, such as gender, ethnicity, and different household registration types to study the same issues.

References

Abbink, K., J. Brandts, B. Herrmann, and H. Orzen. 2010. "Intergroup Conflict and Intra-group Punishment in an Experimental Contest Game." American Economic Review 100 (1): 420–447.

Akerlof, G.A., and R.E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics* 115(3): 715–753.

Akerlof, G.A., and R.E. Kranton. 2005. "Identity and the Economics of Organizations." *Journal of Economic Perspectives* 19(1): 9–32.

Akerlof, G. A., and R.E. Kranton. 2008. "Identity, Supervision, and Work Groups." *American Economic Review* 98 (2): 212–217.

Alchian, A.A., and H. Demsetz. 1972. "Production, Information Costs, and Economic Organization." *American Economic Review* 62: 777–795.

Anderson, C., and L. Putterman. 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): 1–24.

Anderson, L.R., J.M. Mellor, and J. Milyo. 2008. "Inequality and Public Good Provision: An Experimental Analysis." *Journal of Socio-Economics* 37: 1010–1028.

Andreoni, J. 1995. "Cooperation in Public-goods Experiments: Kindness or Confusion?" *American Economic Review* 85(4): 891–904.

Balafoutas, L., M.G. Kocher, L., Putterman, and M. Sutter. 2010. "Equality, Equity and Incentives: An Experiment." Working Papers in Economics and Statistics 2010–2026, University of Innsbruck.

Bernhard, H., E. Fehr, and U. Fischbacher. 2006. "Group Affiliation and Altruistic Norm Enforcement." *American Economic Review* 96(2): 217–221.

Botelho, A, G.W. Harrison, L.M. Costa Pinto, and E.E. Rutström. 2009. Testing Static Game Theory with Dynamic Experiments: A Case Study of Public Goods. *Game and Economic Behavior* 67: 253–265.

Brekke, K.A., J. Konow, and K. Nyborg. 2012. "Cooperation is Relative: Income and Framing Effects with Public Goods." Memorandum 16/2012, Department of Economics, University of Oslo.

Buckley, E., and R. Croson. 2006. "Income and Wealth Heterogeneity in the Voluntary Provision of Linear Public Goods." *Journal of Public Economics* 90(4): 935–955.

Carpenter, J. 2007. "The Demand for Punishment." *Journal of Economic Behavior and Organization* 62 (4): 522–542.

Carpenter, J., S. Bowles, H. Gintis, and S. Hwang. 2009. "Strong Reciprocity and Team Production: Theory and Evidence." *Journal of Economic Behavior and Organization* 71: 221–232.

Carpenter, J., and P.H. Matthews. 2009. What Norms Trigger Punishment? *Experimental Economics* 12: 272–288.

Chan, K.S., S. Mestelman, R. Moir, and R.A. Muller. 1996. "The Voluntary Provision of Public Goods under Varying Income Distributions." *Canadian Journal of Economics* 29(1): 54–69.

Charness, G., L.Rigotti, and A. Rustichini. 2007. “Individual Behavior and Group Membership.” *American Economic Review* 97(4): 1340–1352.

Che, Y.-K., and S.-W. Yoo. 2001. “Optimal Incentives for Teams.” *American Economic Review* 91(3): 525–541.

Chen, R., and Y. Chen. 2011. “The Potential of Social Identity for Equilibrium Selection.” *American Economic Review* 101(6): 2562–2589.

Chen, Y., and S.X. Li. 2009. “Group Identity and Social Preferences.” *American Economic Review* 99(1): 431–457.

Chen, Y., S.X. Li, T.X. Liu, and M. Shih. 2010. “Which Hat to Wear? Impact of Natural Identities on Coordination and Cooperation. ” Working Paper.

Cherry, T., S. Kroll, and J. Shogren. 2005. “The Impact of Endowment Heterogeneity and Origin on Public Good Contributions: Evidence from the Lab.” *Journal of Economic Behavior and Organization* 57(3): 357–365.

Cinyabuguma, M., T. Page, and L. Putterman. 2006. “Can Second-order Punishment Deter Perverse Punishment? *Experimental Economics* 9: 265–279.

Eckel, C.C., and P.J. Grossman. 2005. “Managing Diversity by Creating Team Identity.” *Journal of Economic Behavior and Organization* 58(3): 371–92.

Egas, M., and A. Riedl. 2008. “The Economics of Altruistic Punishment and the Maintenance of Cooperation.” *Proceedings of the Royal Society B: Biological Sciences* 275: 871–878.

Fehr, E., and S. Gächter. 2000a. “Fairness and Retaliation: The Economics of Reciprocity.” *Journal of Economic Perspectives* 14(3): 159–181.

Fehr, E., and S. Gächter. 2000b. “Cooperation and Punishment in Public Goods Experiments.” *American Economic Review* 90: 980–994.

Fehr, E., and S. Gächter. 2002. “Altruistic Punishment in Humans.” *Nature* 415: 137–140.

Fehr, E., and K. Schmidt. 1999. “A Theory of Fairness, Competition, and Cooperation.” *The Quarterly Journal of Economics* 114(3): 817–868.

Fischbacher, U., 2007. “z-Tree: Zurich Toolbox for Ready-made Economic Experiments.” *Experimental Economics* 10 (2): 171–178.

Fischbacher, U., and S. Gächter. 2010. “Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments.” *American Economic Review* 100 (1): 541–556.

Fischbacher, U., S. Gächter, and E. Fehr. 2001. “Are People Conditionally Cooperative? Evidence from a Public Goods Experiment.” *Economics Letters* 71(3): 397–404.

Fisher, J., R.M. Isaac, J.W. Schatzberg, and J.M. Walker. 1995. “Heterogeneous Demand for Public Goods: Behavior in the Voluntary Contributions Mechanism.” *Public Choice* 85: 249–266.

Fuster, A., and S. Meier. 2010. “Another Hidden Cost of Incentives: The Detrimental Effect on Norm Enforcement.” *Management Science* 56 (1): 57–70.

Gächter, S., and A. Riedl. 2005. “Moral Property Rights in Bargaining with Infeasible Claims,” *Management Science* 51: 249–263.

Goette, L., D. Huffman, and S. Meier. 2006. “The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence Using Random Assignment to Real Social Groups.” *American Economic Review* 96(2): 212–216.

Goette, L., D. Huffman, S. Meier, M. Sutter. 2012. “Competition between Organizational Groups: its Impact on Altruistic and Anti-social Motivations.” *Management Science* 58(5): 948–960.

Herrmann, B, C. Thöni, and S. Gächter. 2008. “Antisocial Punishment across Societies.” *Science* 319: 1362–1367.

Hofmeyr, A., J. Burns, and M. Visser. 2007. “Income Inequality, Reciprocity and Public Good Provision: An Experimental Analysis.” *South African Journal of Economics* 75(3): 508–520.

Hoffman, E., and M.L. Spitzer. 1985. “Entitlements, Rights, and Fairness: An Experimental Examination of Subjects’ Concepts of Distributive Justice.” *Journal of Legal Studies* 14: 259–297.

Houser, D., E. Xiao, K. McCabe, and V. Smith. 2008. “When Punishment Fails: Research on Sanctions, Intentions and Non-cooperation.” *Games and Economic Behavior* 62 (2): 509–532.

Isaac, Mark R., and James M. Walker. 1988. “Group Size Effects in Public Goods Provision: The Voluntary Contribution Mechanism.” *Quarterly Journal of Economics* 103(1): 179–199.

Kandel, E., and E.P. Lazear. 1992. “Peer Pressure and Partnerships.” *Journal of Political Economy* 100(4): 801–817.

Lembke, S., and M.G. Wilson. 1998. “Putting the ‘Team’ into Teamwork: Alternative Theoretical Contributions for Contemporary Management Practice.” *Human Relations* 51: 927–944.

Mas, A., and E. Moretti. 2009. “Peers at Work.” *American Economic Review* 99(1): 112–145.

Masclet, D., C. Noussair, S. Tucker, and M.C., Villeval. 2003. “Monetary and Non-monetary Punishment in the Voluntary Contributions Mechanism.” *American Economic Review* 93: 366–380.

McDonald, J. F, and R.A. Moffitt. 1980. “The Uses of Tobit Analysis.” *Review of Economics and Statistics* 62: 318–321.

McLeish, K.N., and R.J. Oxoby. 2007. “Identity, Cooperation, and Punishment.” IZA Discussion Paper No. 2572.

McLeish, K.N., and R.J. Oxoby. 2008. “Social Interactions and the Salience of Social Identity.” IZA Discussion Paper No. 3554.

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., and H.T. Normann. 2008. “A Comparative Statics Analysis of Punishment in Public-Good Experiments.” *Experimental Economics* 11: 358–369.

Nikiforakis, N. 2010. “Feedback, Punishment and Cooperation in Public Good Experiments.” *Games and Economic Behavior* 68: 689–702.

Nikiforakis, N., H.T. Normann, and B. Wallace. 2010. “Asymmetric Enforcement of Cooperation in a Social Dilemma.” *Southern Economic Journal* 76 (3): 638–659.

Prediger, S. 2011. “How Does Income Inequality Affect Cooperation and Punishment in Public Good Settings.” MAGKS Joint Discussion Paper Series in Economics No. 38–2011.

Reuben, E., and A. Riedl. 2013. “Enforcement of Contribution Norms in Public Good Games with Heterogeneous Populations.” *Games and Economic Behavior* 77(1): 122–137.

Ruffle, B., and R. Sosis. 2006. “Cooperation and the In-group-Out-group Bias: A Field Test on Israeli Kibbutz Members and City Residents.” *Journal of Economic Behavior and Organization* 60(2): 147–163.

Tajfel, H., and J. Turner. 1979. “An Integrative Theory of Intergroup Conflict.” In S. Worchel and W. Austin (eds.), *The Social Psychology of Intergroup Relations*.. Monterey, CA: Brooks/Cole, pp. 33–47.

Tajfel, H., and J. Turner. 1985. “The Social Identity Theory of Intergroup Behavior.” In S. Worchel and W. Austin (eds.), *The Psychology of Intergroup Relations*.. Chicago: Nelson-Hall, pp. 7–24.

Van Dijk, F., J. Sonnemans, and F. van Winden. 2002. “Social Ties in a Public Good Experiment.” *Journal of Public Economics* 85: 275–299.

Visser, M., and J. Burns. 2006. “Bridging the Great Divide in South Africa: Inequality and Punishment in the Provision of Public Goods.” Working Paper in Economics No.219. University of Gothenburg, Gothenburg.

Table 1. Experimental treatments

Treatment	Endowment distribution	Identity	Punishment
Homo-Weak-NoPunish	Homogenous	Weak	No
Hetero-Weak-NoPunish	Heterogeneous	Weak	No
Homo-Strong-NoPunish	Homogenous	Strong	No
Hetero-Strong-NoPunish	Heterogeneous	Strong	No
Homo-Weak-Punish	Homogenous	Weak	Yes
Hetero-Weak-Punish	Heterogeneous	Weak	Yes
Homo-Strong-Punish	Homogenous	Strong	Yes
Hetero-Strong-Punish	Heterogeneous	Strong	Yes

Table 2. Average contribution rates across treatments

	Without punishment				With punishment			
	Homo-Weak	Hetero-Weak	Homo-Strong	Hetero-Strong	Homo-Weak	Hetero-Weak	Homo-Strong	Hetero-Strong
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average	0.46	0.31	0.50	0.46	0.69	0.67	0.79	0.67
	(0.22)	(0.16)	(0.16)	(0.22)	(0.25)	(0.25)	(0.17)	(0.25)
High		0.27		0.38		0.65		0.63
		(0.21)		(0.26)		(0.27)		(0.26)
Second		0.25		0.42		0.63		0.61
		(0.14)		(0.27)		(0.27)		(0.30)
Third		0.32		0.46		0.68		0.72
		(0.20)		(0.24)		(0.27)		(0.26)
Low		0.39		0.59		0.74		0.73
		(0.29)		(0.26)		(0.27)		(0.28)

Notes: The table reports the average contribution rates depending on treatment (first row) and endowment level (last four rows) in heterogeneous endowment treatments. Standard deviations in parentheses.

Table 3. Determinants of contribution rates

Dependent variable: contribution rate of subject i in period t, $\left(\frac{c}{E}\right)_{i,t}$				
	Without punishment		With punishment	
	Homo & Hetero	Hetero only	Homo & Hetero	Hetero only
	(1)	(2)	(3)	(4)
Homo-Weak	0.134*		0.031	
	(0.069)		(0.084)	
Hetero-Strong	0.150**		0.00008	
	(0.069)		(0.086)	
Homo-Strong	0.175**		0.116	
	(0.069)		(0.078)	
Weak-High		-0.118***		-0.099***
		(0.030)		(0.027)
Weak-Second		-0.141***		-0.121***
		(0.031)		(0.028)
Weak-Third		0.068**		-0.059**
		(0.031)		(0.026)
Strong-High		-0.008		-0.115
		(0.078)		(0.088)
Strong-Second		0.022		-0.132
		(0.079)		(0.088)
Strong-Third		0.056		-0.034
		(0.079)		(0.083)
Strong-Low		0.200**		0.002
		(0.078)		(0.080)
Observations	1920	960	1920	960
Wald Chi2	211.05***	155.48***	338.09***	228.94***
Log-likelihood	-843.76	-405.60	-758.10	-354.82
Left censored observations	262	145	37	27
Right censored observations	185	93	746	325
Linear combination of the model marginal effects:				
(i) (Homo-Strong) - (Hetero-Strong)	0.025		0.116	
	(0.072)		(0.078)	
(ii) (Homo-Strong) - (Homo-Weak)	0.041		0.085	
	(0.072)		(0.076)	
(iii) (Strong-High) - (Strong-Low)		-0.208***		-0.117***
		(0.031)		(0.027)
(iv) (Strong-Second) - (Strong-Low)		-0.178***		-0.134***
		(0.032)		(0.028)
(v) (Strong-Third) - (Strong-Low)		-0.144***		-0.036
		(0.032)		(0.025)
(vi) (Strong-High) - (Weak-High)		0.110		-0.016
		(0.074)		(0.093)
(vii) (Strong-Second) - (Weak-Second)		0.163**		-0.010
		(0.074)		(0.095)
(viii) (Strong-Third) - (Weak-Third)		0.124		0.025
		(0.077)		(0.087)

Notes: The table reports the regression results for a tobit model with both upper and lower censoring and team random effects. Models (1) and (2) are estimated for the treatments without punishment, and models (3) and (4) with punishment. Models (1) and (3) are run on both homogenous and heterogeneous treatments, whereas models (2) and (4) are run on heterogeneous treatments only. Entries in the topmost panel are the average marginal effects of the independent variables. Period dummies are controlled in the regressions, but the results are not shown here. The bottom panel shows the linear combination of the model marginal effects. Standard errors in parentheses. *** indicates significance at the 1% level, ** at the 5% level, * at the 10% level.

Table 4. Average number of punishment points assigned across treatments

	Homo-Weak	Hetero-Weak	Homo-Strong	Hetero-Strong
	(1)	(2)	(3)	(4)
Average	0.53	0.47	0.46	0.36
	(0.36)	(0.44)	(0.32)	(0.41)
High		0.68		0.30
		(1.06)		(0.34)
Second		0.33		0.57
		(0.24)		(0.80)
Third		0.46		0.37
		(0.50)		(0.69)
Low		0.41		0.20
		(0.67)		(0.19)

Note: The table reports the average punishment points assigned by subject *i* to *j* depending on treatment (first row) and endowment level (last four rows) in heterogeneous endowment treatments. Standard deviations in parentheses.

Table 5. Determinants of punishment

Dependent variable: punishment points assigned from subject i to j in period t, $p_{ij,t}$				
	Homo & Hetero		Hetero only	
	(1)	(2)	(3)	(4)
Homo-Weak	0.158	0.157		
	(0.161)	(0.111)		
Hetero-Strong	-0.051	-0.068		
	(0.135)	(0.090)		
Homo-Strong	0.049	0.126		
	(0.148)	(0.110)		
Weak-High			0.292***	0.374***
			(0.101)	(0.087)
Weak-Second			0.053	0.092
			(0.073)	(0.062)
Weak-Third			0.153*	0.160**
			(0.083)	(0.067)
Strong-High			0.036	0.053
			(0.135)	(0.091)
Strong-Second			0.287*	0.373***
			(0.173)	(0.126)
Strong-Third			0.075	0.033
			(0.141)	(0.089)
Strong-Low			-0.094	-0.108
			(0.116)	(0.074)
Others' average contribution rate		-0.595***		-0.525***
		(0.111)		(0.140)
Absolute negative deviation		2.478***		2.289***
		(0.192)		(0.251)
Positive deviation		-0.257**		-0.145
		(0.119)		(0.146)
Observations	5760	5760	2880	2880
Wald Chi2	156.66***	934.42***	91.23***	402.97***
Log-likelihood	-4538.22	-4033.86	-2047.19	-1821.40
Left censored observations	4689		2406	
Right censored observations	2		1	
Linear combination of the model marginal effects:				
(i) (Homo-Strong) - (Hetero-Strong)	0.100	0.194*		
	(0.142)	(0.105)		
(ii) (Homo-Strong) - (Homo-Weak)	-0.109	-0.031		
	(0.167)	(0.123)		
(iii) (Strong-High) - (Strong-Low)			0.130*	0.160***
			(0.071)	(0.058)
(iv) (Strong-Second) - (Strong-Low)			0.381***	0.480***
			(0.108)	(0.093)
(v) (Strong-Third) - (Strong-Low)			0.169**	0.141**
			(0.077)	(0.055)
(vi) (Strong-High) - (Weak-High)			-0.256	-0.321**
			(0.177)	(0.130)
(vii) (Strong-Second) - (Weak-Second)			0.234	0.281**
			(0.178)	(0.132)
(viii) (Strong-Third) - (Weak-Third)			-0.078	-0.127
			(0.162)	(0.105)

Notes: The table reports the regression results for a tobit model with both upper and lower censoring and team random effects. Models (1) and (2) are estimated on both homogenous and heterogeneous treatments, whereas models (3) and (4) on heterogeneous treatments only. Models (1) and (3) only include treatment variables, whereas models (2) and (4) also include punishment regularity variables. Entries in the top panel are the average marginal effects of the independent variables. Period dummies are controlled in the regressions, but the results are not shown here. The bottom panel shows the linear combination of the model marginal effects. Standard errors in parentheses. *** indicates significance at the 1% level, ** at the 5% level, * at the 10% level.

Table 6. Response to punishment regularities by treatment

Dependent variable: punishment points assigned from subject i to j in period t, $p_{ij,t}$				
	Homo-Weak	Hetero-Weak	Homo-Strong	Hetero-Strong
	(1)	(2)	(3)	(4)
Others' average contribution rate (γ_1)	-0.572***	-0.342	-0.807***	-0.720***
	(0.206)	(0.218)	(0.313)	(0.226)
Absolute negative deviation (γ_2)	2.546***	2.520***	2.813***	2.004***
	(0.349)	(0.457)	(0.466)	(0.317)
Positive deviation (γ_3)	-0.271	-0.044	-1.252***	-0.142
	(0.236)	(0.234)	(0.415)	(0.188)
Observations	1440	1440	1440	1440
Wald Chi2	293.19***	174.87***	304.63***	196.16***
Log-likelihood	-1142.52	-991.01	-997.65	-859.81
Left censored observations	1120	1196	1163	1210
Right censored observations	0	0	1	1
Tests across treatments:				
	Homo-Weak = Hetero-Weak	Homo-Strong = Hetero-Strong	Homo-Weak = Homo-Strong	Hetero-Weak = Hetero-Strong
(i) Test that γ_1 differs	*p*= 0.442	*p*= 0.821	*p*= 0.531	*p*= 0.229
(ii) Test that γ_2 differs	*p*= 0.963	*p*= 0.151	*p*= 0.647	*p*= 0.354
(iii) Test that γ_3 differs	*p*= 0.495	*p*= 0.015	*p*= 0.040	*p*= 0.744
(iv) Test that γ_1, γ_2, and γ_3 differ	*P*= 0.832	*p*= 0.024	*p*= 0.226	*p*= 0.228

Notes: The table reports the regression results for a tobit model with both upper and lower censoring and team random effects. Each model is estimated for one treatment with the treatment name specified as the column heading. Entries in the top panel are the average marginal effects of the independent variables. Period dummies are controlled in the regressions, but the results are not shown here. The bottom panel shows cross-treatment test results. *p*-values in cross-treatment tests are all two-sided. Standard errors in parentheses. *** indicates significance at the 1% level, ** at the 5% level, * at the 10% level.

Figure 1. Evolution of average contribution rates across treatments

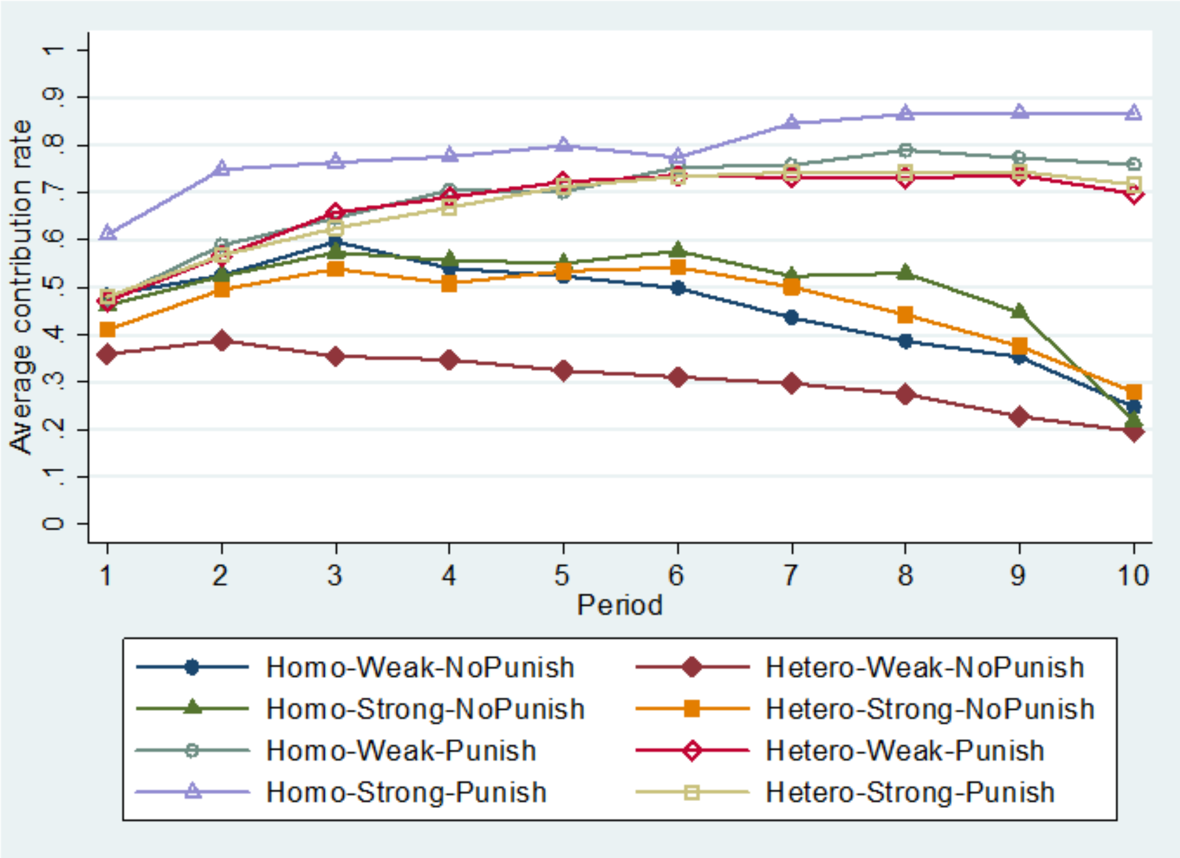

Figure 2. Average punishment points assigned to subject *j* by treatment and by category of deviation in contribution rate of subject *j* from the average of other team members