Enforcement of Contribution Norms in Public Good Games with Heterogeneous Populations

ERNESTO REUBEN
ARNO RIEDL

CESIFO WORKING PAPER NO. 2725
CATEGORY 1: PUBLIC FINANCE
JULY 2009

CESifo *Working Paper No. 2725*

Enforcement of Contribution Norms in Public Good Games with Heterogeneous Populations

Abstract

Economic and social interaction takes place between individuals with heterogeneous characteristics. We investigate experimentally the emergence and informal enforcement of different contribution norms to a public good in homogeneous and different heterogeneous groups. When punishment is not allowed all groups converge towards free-riding. With punishment, contributions increase and differ distinctly across groups and individuals with different induced characteristics. We show econometrically that these differences are not accidentally but enforced by punishment. The enforced contribution norms are related to fairness ideas of equity regarding the contributions but not regarding the earnings. Individuals with different characteristics tacitly agree on the norm to be enforced, even if this leads to large payoff differences. Our results also emphasize the role of details of the environment that may alter focal contribution norms in an important way.

JEL Code: H41, C92, Z13.

Keywords: public good, heterogeneous groups, punishment, cooperation, social norms, norm enforcement.

Ernesto Reuben
Columbia University
ereuben@columbia.edu

Arno Riedl
Maastricht University
a.riedl@maastrichtuniversity.nl

This version: July 13, 2009
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.

1 Introduction

The need for cooperation among people with heterogeneous characteristics is an undeniable fact of social and economic life. At the work place, teams are composed of workers who may differ in their productivity, ability, and motivation (Hamilton et al., 2003). Irrigation systems are often jointly used and maintained by farmers of different sizes and water needs.[1] People also can derive very different benefits from public goods. For example, the elevation of dams along the Mississippi river gives very different benefits to individuals who live close to the river compared to those who live further away. In the international political and economical arena, countries that hugely differ in size and wealth are often confronted with situations that require them to find joint agreements in order to overcome social dilemmas. Sandler and Hartley (2001) discus this problem in the framework of international military alliances. Other prominent examples include the Kyoto protocol for the reduction of emissions of green house gases, fishing quotas among European Union members for the mitigation of over-fishing of open waters, and the Global Disease Detection Program spearheaded by the United States for the early detection of infectious diseases.

As diverse the above examples seem, they can all be viewed as special cases of a more general public goods problem where the enforcement of cooperation by third-parties is infeasible or very limited (e.g., due to actions being unobservable or the nonexistence of a supranational institution with coercive power). In such situations, cooperation has to be promoted through other mechanisms such as social norms that are informally enforced (Elster, 1989; Coleman, 1990).[2] The importance of social norms for sustaining cooperative behavior in public goods environments has been demonstrated in a number of controlled laboratory experiments (for recent reviews see, Fehr and Fischbacher, 2004; Gächter and Herrmann, 2009). However, this experimental evidence is based on homogeneous-group environments, and therefore, neglects the important fact that people differ. This is potentially a serious shortcoming because people who differ may also adhere to different norms, which may lead to conflicts and inefficiencies. In this paper, we experimentally investigate the emergence and informal enforcement of contribution

[1] For instance, in the western states of the United States, family farms dependent on irrigation vary in annual farm sales from below $100,000 to above $500,000 (U.S.D.A., 2004).

[2] Social norms are a widespread empirical phenomenon (Becker, 1996; Hechter and Opp, 2001; Posner, 2002). Numerous examples of the impact of norms on behavior have been meticulously documented (e.g., Roethlisberger and Dickson, 1947; Whyte, 1955; Hywel, 1985; Sober and Wilson, 1997; Gurven, 2004). For examples of the role of norms in the use of common resources see Ostrom (1990).

norms in the presence of heterogeneity.

Since the seminal paper of Fehr and Gächter (2000), a series of studies has contributed to our understanding of the emergence and enforcement of contribution norms in public good games with homogeneous groups. Generally, in homogeneous groups, people sanction those who contribute less than they do (and sometimes also those who contribute more). This behavior is consistent with the enforcement of a norm that prescribes equal contributions by all group members and is often successful in supporting relatively high levels of cooperation.[3] In homogeneous groups, given the symmetry of actors, such an equal-contributions norm is intuitively appealing and is in concordance with important general fairness principles: equality and equity (Konow, 2003).[4]

In heterogeneous groups it is much less obvious what contribution norm may emerge, if one emerges at all. Different and probably conflicting notions of fairness may be invoked by different people depending on their characteristics. If, for instance, people differ in their income, a norm of equal contributions may be appealing to those with more resources and a norm based on contributions proportional to income may be preferred by those with less resources. Similarly, if people are heterogeneous with respect to their preferences for the public good (or their productivity in producing it) equal contributions may be preferred by those who derive a lot of pleasure from the provision of the public good whereas those who enjoy the public good less may prefer a norm with asymmetric contributions. In contrast to homogeneous groups, the experimental evidence regarding contributions to public goods in heterogeneous groups is much less conclusive,[5] and evidence on the enforcement of contribution norms in heterogeneous groups is basically absent.[6]

[3]The success of informal sanctioning in supporting cooperative behavior has been shown to depend on the costs and effectiveness of the sanctioning (Anderson and Putterman, 2006; Carpenter, 2006; Egas and Riedl, 2008; Masclet and Villeval, 2008; Nikiforakis and Normann, 2008; Sutter et al., 2008), the possibility of taking revenge (Denant-Boemont et al., 2007; Nikiforakis, 2008), availability of information (Carpenter, 2007), communication opportunities (Bochet et al., 2006), and cultural factors (Gächter et al., 2008).

[4]Equality and equity considerations are commonly called upon in normative research and have been extensively discussed by numerous philosophers (e.g., Aristotle, 1925; Rawls, 1971; Corlett, 2003). Equality is also commonly invoked in social choice theory as axioms of symmetry and anonymity (e.g., Moulin, 1991; Gaertner, 2006).

[5]Experiments investigating endowment heterogeneity report mixed results. Ostrom et al. (1994), van Dijk et al. (2002), and Cherry et al. (2005) find that inequality leads to lower contributions, Chan et al. (1996) and Buckley and Croson (2006) report a positive effect, and Chan et al. (1999) and Sadrieh and Verbon (2006) no effect. With respect to heterogeneity in the marginal benefit from the public good, Fisher et al. (1995) find that individuals with a high marginal benefit contribute more than those with a low marginal benefit.

[6]To our knowledge, the only experiment that combines endowment heterogeneity and punishment possibilities

In this paper we provide experimental evidence for the emergence and informal enforcement of different contribution norms in a repeated linear public good game when people differ in their endowment or their preference for the public good. In total, we implement eight treatments consisting of four different heterogeneity conditions, each with and without punishment possibilities. In the unequal endowment treatments, heterogeneity is introduced by providing one person (out of three) with an endowment that is twice as high as the endowment of the other group members. To control for the effect of the extended contribution possibilities due to a larger endowment, we restrict the contribution possibilities to be the same for all group members in one pair of treatments, whereas in another pair we allow for contributions up to the entire endowment. In a third treatment pair, we keep endowments the same for all group members but induce a 50 percent higher marginal benefit from the public good for one of the three group members. As control treatments, we also examine behavior in homogeneous groups (with and without punishment). Our design allows us to isolate the effect of unequal endowments, unequal contribution possibilities, and unequal preferences for the public good on contribution behavior as well as their interaction with sanctioning possibilities within one experimental setting.

We find that without punishment possibilities, heterogeneity does not matter much. In all treatments free-riding is relatively frequent and steadily increases over time. In other words, we do not find evidence for a contribution norm other than free-riding to emerge. In the treatments with punishment the picture changes drastically. In homogeneous as well as heterogeneous groups, contributions are much higher than without punishment and they do not decrease over time. More importantly, the contribution pattern differs strongly across treatments. In the treatment with unequal endowments and unrestricted contribution possibilities, contributions are proportional to endowments. Similarly, in the treatment with unequal marginal benefits from the public good, contributions are almost perfectly proportional to the ratio of marginal benefits. In contrast, in the treatment with constrained contribution possibilities, group members with large endowments contribute not more than other group members despite the fact that they have an endowment that is twice as large. We show econometrically that contributions do not

is Visser and Burns (2006). They report that among South African fishermen punishment effectively promotes cooperation in both homogeneous and heterogeneous groups. In Reuben and Riedl (2009) we show that in privileged groups—that is, groups in which one player has a dominant strategy to contribute a positive amount—punishment does not promote contributions as effectively as in homogeneous groups. Tan (2008) finds a similar result for non-privileged heterogeneous groups and Noussair and Tan (2009) report that voting on punishment is not effectively increasing contributions in such groups.

differ accidently but are the result of informal enforcement of different contribution norms in the different treatments. Interestingly, irrespective of differences in endowments and marginal benefits from the public good, individuals within a group largely agree on which contribution norm to enforce—even when the norm implies that a group of individuals benefits relatively more.

The rest of the paper is organized as follows. Section 2 presents the experimental design and procedures. Section 3 discusses different focal and potentially conflicting contribution norms, given the heterogeneity among group members. Section 4 reports the contribution rates of the different treatments and presents the results regarding the enforcement of different contribution norms. Section 5 concludes and discusses our results.

2 Experimental Design

The basic game implemented in the experiment is a linear public good game that is played by the same group of three subjects for ten consecutive periods. The game consists of a contribution stage in which each subject i receives an endowment of y_i points. Subjects simultaneously decide how many points, c_i, they want to contribute to the public good, where $c_i \in [0, \bar{c}_i]$ and $\bar{c}_i$ is person i's maximum contribution. Every point contributed to the public good by any group member increases i's earnings by α_i points and every point not contributed by i increases i's earnings by one point. If $\alpha_i < 1$ for all i and $\sum_i \alpha_i > 1$, then each point contributed increases the sum of earnings in the group but decreases the earnings of the contributing subject, creating a tension between individual and group interest. Subject i's earnings at the end of the contribution stage are given by

$$\pi_i = y_i - c_i + \alpha_i \sum_j c_j.$$

Each subject takes part in one of eight treatments, which vary along two dimensions: first, in the degree of heterogeneity of endowments and marginal benefits from the public good, and second, in the possibility to punish other group members or not (see Fehr and Gächter, 2000).

Two treatments correspond to the standard public good game with homogeneous groups played with and without punishment, respectively. In these treatments, each group member i has the same endowment $y_i = 20$ points, and receives the same marginal benefit from the public good $\alpha_i = 0.50$. We call these treatments Equal (see Table 1). In the remaining six treatments, groups are heterogeneous: group members differ either in their endowment or in their marginal benefit from the public good. Specifically, in each group one member receives a

Table 1: Experimental treatments

Group type	Subject's type	y_i	α_i	$\bar{c}_i$	Number of groups without/with punishment
EQUAL	low	20 points	0.50	20 points	7 / 6
URE	low	20 points	0.50	20 points	7 / 7
	high	40 points	0.50	20 points	
UUE	low	20 points	0.50	20 points	6 / 6
	high	40 points	0.50	40 points	
UMB	low	20 points	0.50	20 points	7 / 6
	high	20 points	0.75	20 points	

higher endowment or a higher marginal benefit from the public good than the other two group members. For convenience we refer to the former as the *high* type and to the latter as the *low* type.[7]

In two treatments, high types receive an endowment of $y_H = 40$ points whereas low types get $y_L = 20$ points. Importantly, in these treatments contributions of both high and low types are restricted to a maximum of 20 points. We refer to these as the unequal-restricted-endowments treatments or URE. In two further treatments, high types again receive $y_H = 40$ points and low types $y_L = 20$ points. However, in contrast to URE the contributions of high types are unrestricted (i.e., $\bar{c}_H = 40$ points). We refer to these treatments as the unequal-unrestricted-endowments treatments or UUE. In the final two treatments, both types receive the same endowment of 20 points but high types earn a marginal benefit from the public good equal to $\alpha_H = 0.75$ while low types earn $\alpha_L = 0.50$. Correspondingly, we refer to them as the unequal-marginal-benefit treatments or UMB. The eight treatments are summarized in Table 1 along with the number of independent groups in each.

As mentioned, in half of the treatments subjects have no possibility to punish. In these treatments subjects' earnings at the end of a period correspond to their earnings after the contribution stage (see above). In the remaining half, subjects can punish each other as in Fehr and Gächter (2002). In these treatments, the contribution stage is followed by a punishment stage, in which each individual is informed of the contributions of the other group members.[8]

[7] In each group, subjects are randomly assigned to high and low types at the beginning of the experiment, and they stay in their role throughout the ten periods. This procedure is known to all participants.

[8] Subjects know the values of y_i, α_i, $\bar{c}_i$ of all group members. Hence, they can identify the contribution of high

Each subject i simultaneously decides how many punishment points, $p_{ij} \in [0, 10]$, to assign to each subject $j \neq i$ in the group. Each punishment point costs the punisher one point and reduces the earnings of the punished subject by three points.[9] After the punishment stage, subjects are informed of the total number of punishment points assigned to them. As in Fehr and Gächter (2000, 2002), subjects do not receive specific information concerning who punished whom. In the treatments with punishment, at the end of a period, earnings of a subject i are given by[10]

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

Treatments Equal and URE differ only in the higher endowment of one group member. Thus, comparing these treatments allows us to isolate the effect of endowment heterogeneity on contributions and punishment behavior of both high and low types. Due to the restriction on the contributions of high types in URE, we can be sure that any differences in behavior are solely driven by the fact that high types possess a high endowment and not because they can contribute more to the public good. The effect of higher contribution possibilities can be examined by comparing URE with UUE. Finally, comparing Equal with UMB allows us to investigate the effect of differences in the marginal benefits from the public good on contributions and punishment for given equal endowments.

Experimental Procedures

The computerized experiment was conducted in the CREED laboratory of the University of Amsterdam using the typical procedures of anonymity, neutrally-worded instructions, and monetary incentives. In total, 156 subjects participated in the one-hour long experiment. About half of the subjects were female. Also around half were students of economics (the other half came from other fields such as biology, engineering, political science, and law). Average earnings equaled €13.45 (≈US$17.50).

and low types.

[9] In line with Fehr and Gächter (2000) and others, we impose an upper limit on the amount of punishment i can assign to each j. The reason for this restriction is to prevent subjects with higher earnings from having the capability to punish more than subjects with lower earnings (as punishment is funded through their own earnings).

[10] To avoid subjects making losses during the experiment solely by the actions of others, if punished below zero points a subject i earns: $\pi_i = \max[0, y_i - c_i + \alpha_i \sum_j c_j - 3 \sum_{j \neq i} p_{ji}] - \sum_{j \neq i} p_{ij}$. Subjects may accept to incur a loss through the punishment points they deal out, in case they have less than twenty points after the contribution stage (Fehr and Gächter, 2000).

After arrival in the lab's reception room, each subject drew a card to be randomly assigned to a seat in the laboratory. Once everyone was seated, the instructions for the experiment were read aloud (a translation of the instructions, which are originally in Dutch, can be found in the online appendix at http://www.ereuben.net/). Thereafter, subjects answered a few questions to ensure their understanding of the instructions. When all subjects had correctly answered the questions, the computerized experiment (programmed with z-Tree, Fischbacher, 2007) was started. After the ten periods, subjects had to answer a short debriefing questionnaire and were confidentially paid their earnings in cash.

3 Focal and Conflicting Contribution Norms

If all subjects are rational and maximize solely their monetary earnings, all individuals in all treatments are predicted not to contribute to the public good. However, previous experimental evidence from homogeneous groups suggests that: (i) without punishment there is some initial cooperation that decreases to low levels over time (Ledyard, 1995), and (ii) with punishment, sanctions are used to enforce high contribution levels that do not decline with repetition (Fehr and Gächter, 2000; Gächter and Herrmann, 2009).[11]

Our main interest is the possible emergence of contribution norms in homogeneous groups and in different types of heterogeneous groups. In homogeneous groups, where all group members are symmetric at the outset and equal contributions imply equal earnings, it is natural to think that the ensuing contribution norm is one in which everyone contributes an equal amount. In this respect, it is interesting to note that many researchers implicitly assume such an equal-contributions norm when they analyze punishment behavior in public good games. For example, it is commonly assumed that punishment is motivated by deviations from either the average contribution (Fehr and Gächter, 2000, 2002; Anderson and Putterman, 2006; Sefton et al., 2007), the punisher's contribution (Gächter et al., 2008; Egas and Riedl, 2008; Sutter et al., 2008), or both (Masclet et al., 2003; Masclet and Villeval, 2008; Nikiforakis, 2008). In all these cases an equal-contributions norm is assumed.[12] In heterogeneous groups, it is much less obvious what

[11] Models of social preferences have been proposed to explain these deviations from standard economic theory (for a review see, Fehr and Schmidt, 2006). For the treatments without punishment these models predict low contribution levels in all treatments. For the treatments with punishment they predict a large number of equilibria, including some with high contribution levels (see e.g., Boyd and Richerson, 1992; Fehr and Schmidt, 1999).

[12] A notable exception to the common assumption of an equal-contributions norm in homogeneous groups is Carpenter and Matthews (2008). They explicitly look for different contribution norms and find, in a setting

Table 2: Focal contribution norms

Note: Contribution norms implied by the fairness concepts of equality and equity applied to both contributions and earnings. Equity can be interpreted as proportionality with respect to endowments or to marginal benefits (y or α), or proportionality with respect to the capacity to contribute ($\bar{c}$).

	Equality	*Equity to y or α*	*Equity to $\bar{c}$*	*Equality*	*Equity to y or α*	*Equity to $\bar{c}$*
	applied to contributions			applied to earnings		
EQUAL	$c_i = c_j$	$c_i = c_j$	$c_i = c_j$	$c_i = c_j$	$c_i = c_j$	$c_i = c_j$
URE	$c_H = c_L$	$c_H = 2c_L$	$c_H = c_L$	$c_H = 20, c_L = 0$	$c_H = \frac{2}{3}c_L$	$c_H = 20, c_L = 0$
UUE	$c_H = c_L$	$c_H = 2c_L$	$c_H = 2c_L$	$c_H = 20 + c_L$	$c_H = \frac{2}{3}c_L$	$c_H = \frac{2}{3}c_L$
UMB	$c_H = c_L$	$c_H = \frac{3}{2}c_L$	$c_H = c_L$	$c_H = 2c_L$	$c_H = \frac{3}{4}c_L$	$c_H = 2c_L$

the contribution norm would be.

The literature on fair allocation rules provides two prominent fairness concepts that can be used to predict the contribution norms that might emerge in heterogeneous groups: equality and equity (Konow, 2003; Konow et al., 2009). Equality is generally thought of as the equalization of output or outcomes with no necessary link to individual input or capacity. In contrast, equity is mostly interpreted as the dependence of fair outcomes—in a proportional way—on individual effort or ability. In Table 2 we summarize the contribution norms implied by the various interpretations of these two fairness concepts in the framework of our experiment. In the following paragraphs we discuss them in turn.

If subjects interpret equality as equality in contributions, then trivially the equal-contributions norm will emerge in both homogeneous and heterogeneous groups (i.e., everyone contributes equal amounts irrespective of differences in endowments or marginal benefits). Alternatively, if subjects apply the concept of equity, then contributions ought to be proportional. For example, in UUE proportionality implies a contribution norm in which high types contribute twice as much as low types. In the other heterogeneous treatments, however, it is possible to have conflicting interpretations of proportionality. In URE, on the one hand, proportionality can be related to unequal endowments, in which case high types should contribute twice the amount of low types. On the other hand, consistent with Major and Deaux (1982), proportionality can

with incomplete information and in-group and out-group punishment possibilities, that the decision to punish is triggered by deviations from almost full contributions, and that the amount of punishment depends on deviations from contribution rates of 36 percent.

be applied to the equal capacity to contribute, which implies that both types should contribute the same. In UMB, contributions proportional to marginal benefits entail a contribution norm in which high types contribute 50 percent more than low types, whereas proportionality with respect to the capacity to contribute translates into equal contributions for both types. In light of the evidence that individuals often resort to fairness concepts in a self-serving manner (Babcock and Loewenstein, 1997), the possible different interpretations of proportionality allow for low and high types subscribing to focal but conflicting contribution norms.

If subjects interpret equality as equality in earnings, then in both URE and UUE, high types should contribute 20 points more than low types, which in URE implies that low types do not contribute at all. In UMB equality in earnings means that high types contribute twice as much as low types. Applied to earnings, the concept of equity—in the sense of maintaining proportionality to endowments or marginal benefits from the public good—is somewhat counterintuitive. It implies that low types ought to contribute more than high types: 50 percent more in URE and UUE and 33 percent more in UMB. Proportionality of earnings to the capacity to contribute implies again that low types ought to contribute more than high types in UUE, but has the opposite implication in URE and UMB. In URE, low types should not contribute at all, and in UMB, they should contribute half as much as high types.

The application of the discussed fairness concepts considerably narrows down the set of contribution norms that might emerge. However, only in the homogeneous case all concepts lead to a coinciding focal norm. In heterogeneous groups the different conflicting norms make it impossible to tell theoretically which contribution norm will emerge, if a unique norm emerges at all. In the next section we investigate empirically which norm emerges in the different treatments.

4 Experimental results

In this section we first report the contributions to the public good in all treatments with and without punishment. We concentrate on the behavior of high and low types and on whether different types display different contribution patterns. Thereafter, we investigate econometrically the enforcement of contribution norms through punishment.

4.1 Contribution rates

Without punishment, behavior in Equal shows the commonly-observed pattern of initially positive contributions that decrease over time. The Spearman rank-order correlation between mean group contributions and periods is significantly negative ($\rho = -0.584$, $p \leq 0.001$). In each of the unequal treatments, we observe a similar decreasing pattern for both low ($\rho \leq -0.373$, $p \leq 0.003$) and high types ($\rho \leq -0.315$, $p \leq 0.008$).

Table 3 reports the average absolute contribution levels and average contributions relative to endowments for all treatments divided by type. Since we are interested in how emerging contribution norms might lead to persistent differences in the behavior of high and low types, we concentrate on the second half of the game to account for potential learning and experience effects.[13]

Given that high types have twice the endowment of low types in URE and UUE, and a 50 percent higher marginal benefit from the public good in UMB, it is reasonable to expect that their absolute contributions will be higher than those of low types. Surprisingly, the descriptive statistics for treatments without punishment (see first two columns in Table 3) show only small differences in the average absolute contributions of the two types. This impression is corroborated by statistical tests. Wilcoxon signed-rank tests do not find that high types contribute significantly more than low types in any of the unequal treatments (one-sided tests, $p > 0.118$).[14] Similarly, if we use Kruskal-Wallis tests to compare the contributions of *each type* across all treatments, we cannot reject the hypothesis that absolute contributions in all treatments are drawn from the same distribution (low types: $p = 0.374$; high types: $p = 0.711$). Lastly, for all types in all treatments, last-period contributions are very close to zero.[15]

In summary, without punishment there is a surprisingly little difference in contributions across the different types and treatments. Neither a 100 percent larger endowment nor a 50 percent higher benefit from the public good by one group member leads to significantly different contributions. As in homogeneous groups, contributions in heterogeneous groups decrease over time towards full free-riding. This shows that in these treatments, the emerging contribution norm is independent of within-group heterogeneity and the contribution possibilities of the high type. In particular, without punishment opportunities, in all treatments full free-riding emerges as the prevalent behavior.

[13]Descriptive statistics of average contributions for each period can be found in the online appendix (http://www.ereuben.net/). There, we also provide a statistical analysis using data from all periods and pooling across types.

[14]Given the similarity of average absolute contributions, it is evident that average relative contributions are lower for high types than for low types in URE and UUE.

[15]The percentage of subjects that contributed two points or less in the last period are: Equal 85.71%, URE: 90.48%, UUE: 88.89%, and UMB 76.19%.

Table 3: Average contributions

Note: Average contribution to the public good, by each type and treatment, in absolute terms and relative to the endowment. Data corresponds to the last five periods. Standard deviations are in parentheses.

	WITHOUT PUNISHMENT				WITH PUNISHMENT			
	absolute		relative		absolute		relative	
	low	high	low	high	low	high	low	high
EQUAL	2.23	–	0.11	–	16.38	–	0.82	–
	(2.11)	–	(0.11)	–	(4.02)	–	(0.20)	–
URE	4.24	5.63	0.21	0.14	14.36	13.66	0.72	0.34
	(2.52)	(4.22)	(0.13)	(0.11)	(5.47)	(5.97)	(0.27)	(0.15)
UUE	5.18	4.10	0.26	0.10	15.32	28.27	0.77	0.71
	(5.50)	(4.29)	(0.27)	(0.11)	(5.96)	(11.53)	(0.30)	(0.29)
UMB	6.71	7.77	0.34	0.39	11.43	14.30	0.57	0.72
	(5.08)	(7.18)	(0.25)	(0.36)	(5.46)	(5.61)	(0.27)	(0.28)

In the EQUAL treatment, the introduction of punishment leads to a significant increase in contributions. In the last five periods, contributions are 14.15 points higher with punishment than without punishment. Furthermore, with punishment, contributions do not show a statistically significant decreasing trend (Spearman's $\rho = 0.248$, $p = 0.056$).

In the unequal treatments, punishment has the same qualitative effect. However, the size of the effect varies considerably across treatments and types (see Table 3, four rightmost columns). For low types, the increase in contributions ranges from 10.13 points in UUE to only 4.72 points in UMB. For high types, it ranges from 24.17 points in UUE to 6.53 points in UMB.[16] It is also the case that, with punishment, contributions do not display a statistically significant decrease by any type in any of the unequal treatments (low types: $\rho \geq -0.042$, $p \geq 0.728$; high types: $\rho \geq -0.003$; $p \geq 0.984$).

Introducing punishment opportunities has a strong differential effect on the contributions of low and high types, with interesting differences across treatments. In URE, low and high types contribute almost the same amount in absolute terms (14.36 and 13.66 points on average).

[16] To test the statistical significance of these increases we use Mann-Whitney U tests. As we have a clear directional hypothesis, we use one-sided tests. We test separately each type and treatment using group averages across the last five periods as independent observations. The p-values for low types are: $p = 0.001$ for EQUAL, $p = 0.002$ for URE, $p = 0.012$ for UUE, and $p = 0.058$ for UMB. The p-values for high types are: $p = 0.009$ for URE, $p = 0.003$ for UUE, and $p = 0.087$ for UMB.

Consequently, relative to their endowment, high types contribute only half as much as low types. This stands in stark contrast to contribution levels of low and high types in UUE. There, the mean absolute contributions of high types (28.27 points) are almost twice as high as that of low types (15.32 points), implying that relative contributions are very similar (on average, 0.71 for high types and 0.77 for low types). In UMB, we also observe a difference in average contributions between low and high types. High types contribute 14.30 points whereas low types contribute only 11.43 points. Hence, high types contribute 25.11 percent more.[17]

The varying effect of punishment is also reflected in a clear difference in the behavior of high types across treatments. Kruskal-Wallis tests reject the null hypothesis that the contributions of high types are drawn form the same distribution ($p = 0.031$ for absolute contributions and $p = 0.032$ for relative contributions). Pair-wise comparisons reveal that the absolute contributions of high types in UUE, which are 28.27 points on average, are significantly different from the absolute contributions of high types in URE and UMB, which are on average 13.66 and 14.30 points, respectively (two-sided Mann-Whitney U tests $p \leq 0.050$).[18] Analogously, relative contributions are significantly lower in the URE treatment when compared to UUE and UMB (two-sided Mann-Whitney tests, $p \leq 0.050$). In other words, punishment has the strongest effect on behavior of high types in UUE. In this treatment high types contribute about six times more with punishment than without punishment, whereas in the other two treatments there is 'only' a two- to threefold increase in contributions. In contrast, the contributions of low types are very similar across all treatments (including the EQUAL treatment). For low types, a Kruskal-Wallis test does not reject the null hypothesis that contributions come from the same distribution ($p = 0.319$).

In summary, in keeping with existing studies of homogeneous groups, punishment also increases contributions and eliminates the decreasing trend in contributions in heterogeneous groups. Importantly, punishment induces recognizable quantitative differences in the contributions of different types within and across treatments. In URE, where the maximum contribution of high types is bound to be the same as that of low types, both types contribute equally in spite of the fact that high types have twice the endowment of low types and that their contributions

[17]Using Wilcoxon signed-ranks tests to see whether the absolute contributions of high types are significantly higher than those of low types gives the following p-values: $p = 0.877$ for URE, $p = 0.014$ for UUE, and $p = 0.056$ for UMB (one-sided tests). Applying the same tests to relative contributions gives: $p = 0.018$ for URE, $p = 0.206$ for UUE, and $p = 0.116$ for UMB (two-sided tests).

[18]Throughout the paper, whenever we carry out multiple pair-wise comparisons, we correct p-values using the Benjamini-Hochberg method—which reduces the risk of false positives and controls for the rate of false negatives (Benjamini and Hochberg, 1995).

are well below the maximum. In contrast, in UUE high types contribute twice as much as low types, and in UMB they contribute about 25 percent more.

These results clearly suggest that subjects are following different contribution norms in the different treatments. In URE contributions are consistent with a norm in which both types contribute the same in absolute terms. In UUE, behavior is consistent with a norm in which contributions differ and are proportional to the endowment. In UMB, contributions are roughly in line with contributions being proportional to the relative marginal benefits from the public good (i.e., low types contribute 33 percent less). Note that all these norms, if applied to the homogeneous case, imply that everyone should contribute the same amount. Given that in treatments without punishment contributions are very similar across treatments and types, it is likely that the differences we observe in treatments with punishment are the result of differences in punishment behavior. In particular, subjects might be using punishment to enforce different contribution norms in the different treatments. In the following section we explore precisely this conjecture.

4.2 Punishment and the enforcement of heterogeneous norms

In homogeneous groups, given that everyone is in the same position, it is reasonable to assume that when individuals decide who and how much to punish, they treat differences in contributions of different people in the same way. In heterogeneous groups, it is harder to know *a priori* how individuals compare differences in contributions. In principle, one could assume a specific motivation for punishment (e.g., equalize earnings) and then use it to make treatment comparisons. However, given the numerous ways interpersonal comparisons can be made (see Table 2) and our limited knowledge of how subjects perceive contributions in heterogeneous situations, we opt for a more flexible empirical approach. In particular, we aim at insights about the type of interpersonal comparisons individuals are making in the different treatments, and at how they depend on the individual's type.

For our analysis, we only assume that subjects have some idea—based on a contribution norm—about what the contribution of others compared to their own should be, and that at least some subjects are willing to punish individuals who deviate from this contribution. This is a relatively weak assumption that is consistent with existing evidence from experiments of public good games with punishment in homogeneous groups.[19]

[19]For example, Gächter et al. (2008) find this pattern in sixteen different countries across the world. Other

We do not assume a specific contribution norm but elicit the norm that is most consistent with the observed punishment data. More specifically, we estimate the following model:

$$p_{ijt} = \beta_{neg} \max[\mu c_{it} - (1-\mu)c_{jt}, 0] + \beta_{pos} \max[(1-\mu)c_{jt} - \mu c_{it}, 0] + v_i + \epsilon_{ijt}, \tag{1}$$

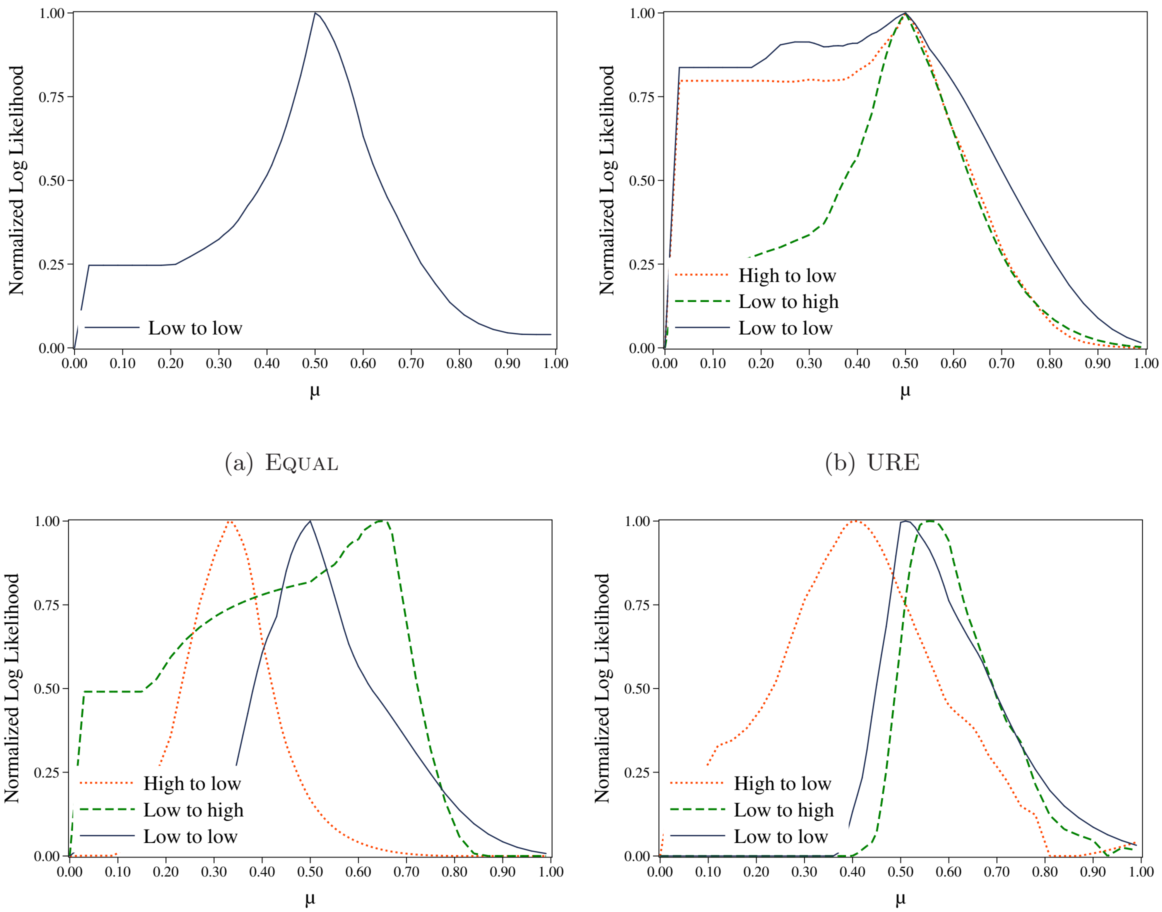
where p_{ijt} is the amount of punishment that i allots to j in period t. The term $\mu \in [0,1]$ captures the norm of how much subjects expect others to contribute in comparison to their own contribution. The first term in the model corresponds to negative deviations from this relative contribution norm. For example, if $\mu = 0.50$ and therefore $1-\mu = 0.50$ then the first term in (1) is positive whenever $c_{jt} < c_{it}$. In other words, subjects expect others to contribute as much as they do, and if someone contributes less, they consider this to be a negative deviation. Alternatively, if $\mu = 0.75$ and therefore $1-\mu = 0.25$ then subjects consider that a negative deviation occurs when $c_{jt} < 3c_{it}$, which implies that subjects expect others to contribute three times as much as they do. In the extremes, if $\mu = 0$, subjects expect to contribute everything themselves and others to contribute nothing, and if $\mu = 1$, subjects expect to contribute nothing themselves and others to contribute everything. The second term in (1) corresponds to positive deviations, which are evaluated using the same μ. The variable v_i captures unobserved individual characteristics, and ϵ_{ijt} is the error term.

To find the value of μ that best explains the data, we estimate (1) using values of μ between zero and one in steps of 0.01. For consistency reasons, we restrict the values of β_{neg} and β_{pos} to be greater than or equal to zero. Given that punishment is bounded by zero and ten, we use Tobit estimates. Furthermore, we treat the unobserved individual characteristics as random effects. Lastly, we use punishment data from all periods (as opposed to only the last five) because punishment occurs more often at the beginning of the game, and it is then when contribution norms should start to emerge.

In Equal, as subjects are in symmetric positions, we look for one value for μ. In the heterogeneous treatments, we distinguish between types and look for a value of μ in each of the following cases: low types punishing high types, low types punishing low types, and high types punishing low types. This allows us to detect not only differences in punishment across treatments, but also between types. Below we present the results of the estimation just described.[20]

examples include Gächter and Herrmann (2007), Egas and Riedl (2008), Nikiforakis (2008), and Reuben and Riedl (2009).

[20] In order to check the robustness of our results, we also estimate (1) using the following variations: (i) adding the period t and the total group contributions as additional dependent variables, (ii) considering only subjects



(a) EQUAL (b) URE

(c) UUE (d) UMB

Figure 1: Goodness of fit for different values of μ

Note: Log likelihood obtained when estimating (1) for values of $\mu \in [0,1]$. The log likelihood is normalized such that 0 equals the log likelihood of the worst fitting regression and 1 that of the best fitting regression.

The relative fits of regressions for the various values of μ in each treatment are depicted in Figure 1. The figure shows, for values of $\mu \in [0,1]$, the value of the log likelihood obtained when estimating (1). For convenience and the sake of comparison, we normalized the log likelihood such that 0 equals the log likelihood of a regression where we set $\beta_{neg} = \beta_{pos} = 0$ (i.e., only with the constant), and 1 equals the log likelihood of the regression with the value of μ that gives the best fit. Henceforth, we refer to the μ of the best-fitting regression as μ^*. Furthermore, we use

who punish at least once, (iii) using Logit estimates and treating punishment as a binary decision (either punish or not), and (iv) treating the unobserved individual characteristics v_i as unconditional fixed effects. These variations give very similar results, which are available in the online appendix (http://www.ereuben.net/).

a subscript to indicate the type of the punisher and punished (e.g., $\mu^*_{L\to H}$ indicates the case of low types punishing high types).

In EQUAL, one can see that the fit of the model has a clear maximum at the focal $\mu^*_{L\to L} = 0.50$. In other words, the best fit is obtained for the μ that implies that subjects punish those who deviate from their own contribution. Moreover, deviations from this value of μ monotonically worsen the model's performance.

A unique global maximum of μ is also observed for both types in the heterogeneous groups treatments. In URE we find that $\mu^*_{L\to H} = \mu^*_{L\to L} = \mu^*_{H\to L}$. That is, low and high types enforce on (other) low types the same contribution norm, and low types do not differentiate between types. For UUE and UMB, we find that subjects do make a distinction between types as $\mu^*_{L\to H} > \mu^*_{L\to L} > \mu^*_{H\to L}$. Hence, subjects' punishment behavior reveals that they expect high types to contribute more than low types, and judging from the differences between the values of μ^*, more so in the UUE treatment. To see this more clearly, we present in Table 4 the values of μ^* in each treatment—along with the estimated coefficients of the corresponding regression.

In all treatments, $\mu^*_{L\to L}$ is very close to or exactly 0.50, which reveals that low types expect other low types to contribute as much as they do, independent of the group heterogeneity. Differences between treatments occur when low types punish high types and vice versa. Remarkably, in URE low types expect high types to contribute as much as they and other low types do since $\mu^*_{L\to L} = \mu^*_{L\to H} = 0.50$. In UUE, in contrast, $\mu^*_{L\to H} = 0.65$, which implies that low types expect high types to contribute roughly twice as much as they do. Hence, it is the high types' ability to contribute more and not simply their higher endowment that makes low types demand that high types contribute more than low types. In UMB, low types also demand that high types contribute more but by a smaller amount: $\mu^*_{L\to H} = 0.56$ translates into roughly 25 percent higher contributions.

Consistent with the low types' contribution norm, high types in URE expect low types to contribute as much as they do (i.e., $\mu^*_{H\to L} = 0.50$). In contrast, in UUE, high types expect low types to contribute less. In this treatment, in agreement with the low types' contribution norm, a $\mu^*_{H\to L} = 0.33$ indicates that high types expect low types to contribute roughly half of their own contribution. The fact that in both URE and UUE high and low types enforce the same contribution norm on each other indicates that there is consensus on what the contribution of each type should be. In UMB the agreement seems somewhat weaker. $\mu^*_{H\to L} = 0.40$ reveals that high types expect low types to contribute 33 percent less than they do whereas low types expect high types to contribute 25 percent more than they do. Interestingly, this implies that

Table 4: Best-fitting contribution norms

Note: Value μ^*, which gives the best fit when estimating (1) with $\mu \in [0,1]$, and the corresponding estimated coefficients. Tobit estimates with subject level random effects and censoring at $p_{ij} = 0$ and $p_{ij} = 10$. Standard errors are in parentheses. The symbols *,**,*** indicate statistical significance at the 10, 5, and 1 percent level.

	Equal	URE			UUE			UMB		
	low→low	low→low	low→high	high→low	low→low	low→high	high→low	low→low	low→high	high→low
μ^*	0.50	0.50	0.50	0.50	0.50	0.65	0.33	0.51	0.56	0.40
β_{pos}	0.46***	0.15	0.00	0.32**	0.00	0.00	0.11	0.00	0.00	0.09
	(0.11)	(0.25)	(0.00)	(0.14)	(0.00)	(0.00)	(0.23)	(0.00)	(0.00)	(0.13)
β_{neg}	1.01***	1.25***	1.10***	0.92***	1.28***	0.85***	0.95***	0.67***	0.95**	1.06***
	(0.11)	(0.23)	(0.19)	(0.14)	(0.41)	(0.32)	(0.28)	(0.25)	(0.47)	(0.19)
cons.	-3.92***	-6.21***	-4.16***	-3.47***	-7.39***	-8.36***	-5.81**	-3.77***	-6.22**	-1.70**
	(0.53)	(1.47)	(1.24)	(1.02)	(2.40)	(2.93)	(2.22)	(1.27)	(2.67)	(0.83)
# obs.	360	140	140	140	120	120	120	120	120	120
χ^2	90.79	32.32	33.65	45.81	9.94	7.08	11.91	7.35	4.18	33.27

the disagreement in contribution norms is in favor of high types, in the sense that high types ask for higher own relative contributions than low types actually enforce.

In principle, in addition to enforcing different contribution norms, subjects in different treatments could differ in the severity with which they punish deviations from a given norm. We can see whether this is the case by looking at the magnitude of the estimated coefficients (available in Table 4). Interestingly, negative deviations from the contribution norm are punished similarly across treatments.[21] This suggests that the motivation to punish negative deviations from a contribution norm is largely independent of the exact norm and individuals' types.[22]

In summary, the observed differences in contributions of high types in the different treatments with punishment can be attributed to the informal enforcement of different contribution norms based on considerations of equality and equity. Interestingly, both high and low types largely agree on the norm that is enforced. Therefore, it is not simply a matter of one type using punishment to coerce the other type towards higher contributions. Of the possible contribution norms (see Table 2), the one that is actually enforced depends on the form of heterogeneity. In URE, where low and high types face the same maximum contribution, both types apply a norm consistent with equal contributions, despite the fact that high types earnings are (almost) twice as high as low types earnings. In UUE, where high types can contribute twice as much as low types, the enforced contributions are proportional to endowments. This pattern suggests an equity based contribution norm where contributions are proportional to the capacity to contribute. In UMB, with equal endowments but unequal marginal benefits from the public good, the enforcement behavior of high types is consistent with a contribution norm that is proportional to the ratio of marginal benefits.

---

[21] For example, if we test whether the coefficient for negative deviations of each regression equals 1.01—which is the value of the coefficient of the EQUAL treatment—we cannot reject the null hypothesis in any of the regressions (Wald tests, $p \geq 0.172$).

[22] Unlike negative deviations, punishment of positive deviations from the contribution norm (so-called antisocial punishment) is less common. It occurs under EQUAL and in the punishment of low types by high types in URE. In fact, as can be seen in Table 4, in some cases the coefficient for positive deviations is restricted to zero. In regressions without this restriction, the coefficient's value is close to zero in all cases and it is never statistically significant ($p \geq 0.467$).

5 Conclusions

In this paper, we provide evidence for the emergence and informal enforcement of different contribution norms in public good games with homogeneous and heterogeneous groups. We find that, in the absence of punishment, contributions steadily decline in all treatments, which results in similar behavior in homogeneous and heterogeneous groups. The trend towards free-riding also dissipates potential differences between individuals with different induced characteristics. Indeed, the behavior that prevails in all treatments is full free-riding. In stark contrast, when punishment is possible contributions do not only increase, but also exhibit considerable differences across treatments and between different types of individuals. We show that the differences in the individuals' contribution and sanctioning behavior are consistent with the enforcement of different contribution norms.

In treatments with unequal endowments, we find that the enforced contribution norm prescribes contributions that are proportional to the maximum feasible contribution. This implies that if contributions are bounded by the endowment (as in our treatment UUE), subjects with large endowments (high types) are expected to contribute more than subjects with small endowments (low types). At the same time, if the contribution possibilities are equal between types (as in our treatment URE) then both high and low types are expected to contribute the same amount. In the treatment with unequal marginal benefits from the public good (UMB), we find that the emerging norm prescribes contributions that are proportional to the ratio of marginal benefits. The identified contribution norms can be readily reconciled with considerations invoking ideas of equality and equity regarding contributions to the public good. In contrast, fairness ideas with respect to earnings fail to account for the differences across and within treatments. This is particularly evident from the relative earnings of high and low types in UUE and URE. Compared to UUE, in URE low types earn considerably less than high types. Hence, the enforced contribution norms are nonconsequentialist in the sense of Elster (1989).

The emergence of different norms among different people in different environments has a silver lining and a demerit. On the one hand, it shows that people are willing and able to informally enforce norms in heterogeneous environments, and that in spite of multiple and conflicting focal norms they can tacitly agree on a unique contribution norm. This leads to relatively high contributions to the public good in both homogeneous and heterogeneous groups. On the other hand, it also shows that the contribution norm that is actually enforced may hinge on details of the environment, which may make it difficult to predict.

In this paper, we concentrate on the enforcement of one norm, which is based on deviations from a subject's own contribution. Although this is in line with the common assumption in the literature, it is conceivable that other fairness ideas might also be at play. In this case, the question whether various norms are simultaneously enforced arises.[23] A promising first step in this direction is set by the study of Carpenter and Matthews (2008), who aim at identifying different types of norms in public good games with homogeneous groups. The investigation of the simultaneous enforcement of multiple norms in heterogeneous groups could build on this work, but calls for a much larger variation of treatments and considerably more data. We leave this for future research.

Recent theoretical and empirical studies point at the importance of norms in diverse areas of the economy and society. On the empirical side, Kim et al. (2006) find that norms of departmental productivity strongly influences the individual productivity of academics, and Goette et al. (2006) report that group membership increases cooperation and the willingness to enforce cooperative norms in platoons of the Swiss Army. Norms have been found to influence behavior even after individuals have moved across societies. For example, Fisman and Miguel (2007) show that norms strongly influence illegal parking behavior of U.N. diplomats in New York, and Guiso et al. (2006) demonstrate how the level of trust exhibited by decedents of immigrants to the United States correlates with the level of trust of the country from which their ancestors emigrated. On the theoretical side, Fischer and Huddart (2008) show that norms can put restrictions on optimal organizational design. Our study confirms that important differences in behavior between groups can be attributed to the informal enforcement of different norms, and are not necessarily due to differences in the preferences of group members. Moreover, we add to this literature the insight that heterogeneity and subtle variations in the environment can shift attention from one focal norm to another, resulting in considerably different outcomes.

References

Anderson, C. M. 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*, 1–24.

Aristotle (1925). *Ethica Nicomachea.* Oxford: Clarendon Press. Translation W.D. Ross.

[23]For example, the model of Charness and Rabin (2002) allows for individuals to simultaneously enforce a norm for efficiency and a norm for the welfare of the poorest individual. Konow (2003) also mentions efficiency principles as important fairness concepts that might be relevant for norm enforcement.

Babcock, L. and G. Loewenstein (1997). Explaining bargaining impasse: The role of self-serving biases. *Journal of Economic Perspectives 11*, 109–126.

Becker, G. S. (1996). *Accounting for Tastes*. Cambridge: Harvard University Press.

Benjamini, Y. and Y. Hochberg (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological) 57*, 289–300.

Bochet, O., T. Page, and L. Putterman (2006). Communication and punishment in voluntary contribution experiments. *Journal of Economic Behavior & Organization 60*, 11–26.

Boyd, R. and P. J. Richerson (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology 13*, 171–195.

Buckley, E. and R. Croson (2006). Income and wealth heterogeneity in the voluntary provision of linear public goods. *Journal of Public Economics 90*, 935–955.

Carpenter, J. P. (2006). The demand for punishment. *Journal of Economic Behavior & Organization 62*, 522–542.

Carpenter, J. P. (2007). Punishing free-riders: How group size affects mutual monitoring and the provision of public goods. *Games and Economic Behavior 60*, 31–52.

Carpenter, J. P. and P. H. Matthews (2008). What norms trigger punishment? Working paper, Middlebury College.

Chan, K. S., S. Mestelman, R. Moir, and R. A. Muller (1996). The voluntary provision of public goods under varying income distributions. *The Canadian Journal of Economics 29*, 54–59.

Chan, K. S., S. Mestelman, R. Moir, and R. A. Muller (1999). Heterogeneity and the voluntary provision of public goods. *Experimental Economics 2*, 5–30.

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

Cherry, T. L., 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 & Organization 57*, 357–365.

Coleman, J. S. (1990). *Foundations of Social Theory*. Cambridge: Harvard University Press.

Corlett, J. A. (2003). Making more sense of retributivism: Desert as responsibility and proportionality. *Philosophy 78*, 279–287.

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

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*(05-065/1), 871–878.

Elster, J. (1989). *The Cement of Society – A Study of Social Order*. Cambridge: Cambridge University Press.

Fehr, E. and U. Fischbacher (2004). Social norms and human cooperation. *Trends in Cognitive Sciences 8*, 185–190.

Fehr, E. and S. Gächter (2000). 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. M. Schmidt (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics 114*, 817–868.

Fehr, E. and K. M. Schmidt (2006). The economics of fairness, reciprocity and altruism–experimental evidence and new theories. In S. C. Kolm and J. M. Ythier (Eds.), *Handbook on the Economics of Giving, Reciprocity and Altruism*, pp. 615–691. Amsterdam: Elsevier.

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

Fischer, P. and S. Huddart (2008). Optimal contracting with endogenous social norms. *American Economic Review 98*(4), 1459–1475.

Fisher, J., M. R. Isaac, J. W. Schatzberg, and J. M. Walker (1995). Heterogeneous demand for public goods: Effects on the voluntary contributions mechanism. *Public Choice 85*, 249–266.

Fisman, R. and E. Miguel (2007). Corruption, norms and legal enforcement: Evidence from diplomatic parking tickets. *Journal of Political Economy 115*, 1020–1048.

Gächter, S. and B. Herrmann (2007). The limits of self-governance in the presence of spite: Experimental evidence from urban and rural russia. Discussion paper no. 2007-11, CeDEx.

Gächter, S. and B. Herrmann (2009). Reciprocity, culture, and human cooperation: Previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society B – Biological Sciences 364*, 791–806.

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

Gaertner, W. (2006, July). *A primer in social choice theory.* Oxford U.K.: Oxford University Press.

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 – Papers and Proceedings 96*, 212–216.

Guiso, L., P. Sapienza, and L. Zingales (2006). Does culture affect economic outcomes? *Journal of Economic Perspectives 20*, 23–48.

Gurven, M. (2004). To give or not to give: An evolutionary ecology of human food transfers. *Behavioral and Brain Sciences 27*, 543–583.

Hamilton, B. H., J. A. Nickerson, and H. Owan (2003). Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy 111*, 465–497.

Hechter, M. and K.-D. Opp (2001). *Social Norms.* New York: Russell Sage Foundation.

Hywel, F. (1985). The law, oral tradition and the mining community. *Journal of Law and Society 12*, 267–269.

Kim, E. H., A. Morse, and L. Zingales (2006). Are elite universities losing their competitive edge? *Journal of Financial Economics*, forthcoming.

Konow, J. (2003). Which is the fairest one of all? a positive analysis of justice theories. *Journal of Economic Literature 41*, 1186–1237.

Konow, J., T. Saijo, and K. Akai (2009). Morals and mores: Experimental evidence on equity and equality. Working paper, Loyola Marymount University.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. H. Kagel and A. E. Roth (Eds.), *The Handbook of Experimental Economics*, pp. 111–194. New Jersey: Princeton University Press.

Major, B. and K. Deaux (1982). Individual differences in justice behavior. In J. Greenberg and R. L. Cohen (Eds.), *Equity and Justice in Social Behavior*, pp. 43–76. New York: Academic Press.

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

Masclet, D. and M.-C. Villeval (2008). Punishment, inequality, and welfare: A public good experiment. *Social Choice and Welfare 31*, 475–502.

Moulin, H. (1991, July). *Axioms of cooperative decision making.* Econometric Society Monographs. Cambridge U.K.: Cambridge University Press.

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.

Noussair, C. N. and F. Tan (2009). Voting on punishment systems within a heterogeneous group. Discussion Paper 2009-19, CentER, Tilburg University.

Ostrom, E. (1990). *Governing the Commons – The Evolution of Institutions for Collective Action.* New York: Cambridge University Press.

Ostrom, E., R. Gardner, and J. Walker (1994). *Rules, Games, and Common-pool Resources.* Ann Arbor: University of Michigan Press.

Posner, E. A. (2002). *Law and Social Norms.* Cambridge: Harvard University Press.

Rawls, J. (1971). *A Theory of Justice.* Cambridge: Harvard University Press.

Reuben, E. and A. Riedl (2009). Public goods provision and sanctioning in privileged groups. *Journal of Conflict Resolution 53*, 72–93.

Roethlisberger, F. F. and W. J. Dickson (1947). *Management and the Worker: An account of a research program conducted by the Western Electric Company, Hawthorne Works, Chicago.* Cambridge: Harvard University Press.

Sadrieh, A. and H. Verbon (2006). Inequality, cooperation, and growth: An experimental study. *European Economic Review 50*, 1197–1222.

Sandler, T. and K. Hartley (2001). Economics of alliances: The lessons for collective action. *Journal of Economic Literature 39*(3), 869–896.

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

Sober, E. and D. S. Wilson (1997). *Unto others—The evolution and psychology of unselfish behavior*. Cambridge: Harvard University Press.

Sutter, M., S. Haigner, and M. Kocher (2008). Choosing the carrot or the stick? endogenous institutional choice in social dilemma situations. Working paper 2008-07, University of Innsbruck.

Tan, F. (2008). Punishment in a linear public good game with productivity heterogeneity. *De Economist 156*(3), 269–293.

U.S.D.A. (2004). United states department of agriculture western irrigated agriculture dataset. http://www.ers.usda.gov/data/westernirrigation/methods.htm.

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, Göteborg University.

Whyte, W. (1955). *Money and Motivation*. New York: Harper and Brothers.

CESifo Working Paper Series

for full list see www.cesifo-group.org/wp
(address: Poschingerstr. 5, 81679 Munich, Germany, office@cesifo.de)

2661 Oliver Falck, Stephan Heblich and Elke Luedemann, Identity and Entrepreneurship, May 2009

2662 Christian Lessmann and Gunther Markwardt, One Size Fits All? Decentralization, Corruption, and the Monitoring of Bureaucrats, May 2009

2663 Felix Bierbrauer, On the Legitimacy of Coercion for the Financing of Public Goods, May 2009

2664 Alessandro Cigno, Agency in Family Policy: A Survey, May 2009

2665 Claudia M. Buch and Christian Pierdzioch, Low Skill but High Volatility?, May 2009

2666 Hendrik Jürges, Kerstin Schneider, Martin Senkbeil and Claus H. Carstensen, Assessment Drives Learning: The Effect of Central Exit Exams on Curricular Knowledge and Mathematical Literacy, June 2009

2667 Eric A. Hanushek and Ludger Woessmann, Schooling, Cognitive Skills, and the Latin American Growth Puzzle, June 2009

2668 Ourania Karakosta, Christos Kotsogiannis and Miguel-Angel Lopez-Garcia, Does Indirect Tax Harmonization Deliver Pareto Improvements in the Presence of Global Public Goods?, June 2009

2669 Aleksandra Riedl and Silvia Rocha-Akis, Testing the Tax Competition Theory: How Elastic are National Tax Bases in OECD Countries?, June 2009

2670 Dominique Demougin and Carsten Helm, Incentive Contracts and Efficient Unemployment Benefits, June 2009

2671 Guglielmo Maria Caporale and Luis A. Gil-Alana, Long Memory in US Real Output per Capita, June 2009

2672 Jim Malley and Ulrich Woitek, Productivity Shocks and Aggregate Cycles in an Estimated Endogenous Growth Model, June 2009

2673 Vivek Ghosal, Business Strategy and Firm Reorganization under Changing Market Conditions, June 2009

2674 Francesco Menoncin and Paolo M. Panteghini, Retrospective Capital Gains Taxation in the Real World, June 2009

2675 Thomas Hemmelgarn and Gaëtan Nicodème, Tax Co-ordination in Europe: Assessing the First Years of the EU-Savings Taxation Directive, June 2009

2676 Oliver Himmler, The Effects of School Competition on Academic Achievement and Grading Standards, June 2009

2677 Rolf Golombek and Michael Hoel, International Cooperation on Climate-Friendly Technologies, June 2009

2678 Martin Cave and Matthew Corkery, Regulation and Barriers to Trade in Telecommunications Services in the European Union, June 2009

2679 Costas Arkolakis, A Unified Theory of Firm Selection and Growth, June 2009

2680 Michelle R. Garfinkel, Stergios Skaperdas and Constantinos Syropoulos, International Trade and Transnational Insecurity: How Comparative Advantage and Power are Jointly Determined, June 2009

2681 Marcelo Resende, Capital Structure and Regulation in U.S. Local Telephony: An Exploratory Econometric Study; June 2009

2682 Marc Gronwald and Janina Ketterer, Evaluating Emission Trading as a Policy Tool – Evidence from Conditional Jump Models, June 2009

2683 Stephan O. Hornig, Horst Rottmann and Rüdiger Wapler, Information Asymmetry, Education Signals and the Case of Ethnic and Native Germans, June 2009

2684 Benoit Dostie and Rajshri Jayaraman, The Effect of Adversity on Process Innovations and Managerial Incentives, June 2009

2685 Peter Egger, Christian Keuschnigg and Hannes Winner, Incorporation and Taxation: Theory and Firm-level Evidence, June 2009

2686 Chrysovalantou Milliou and Emmanuel Petrakis, Timing of Technology Adoption and Product Market Competition, June 2009

2687 Hans Degryse, Frank de Jong and Jérémie Lefebvre, An Empirical Analysis of Legal Insider Trading in the Netherlands, June 2009

2688 Subhasish M. Chowdhury, Dan Kovenock and Roman M. Sheremeta, An Experimental Investigation of Colonel Blotto Games, June 2009

2689 Alexander Chudik, M. Hashem Pesaran and Elisa Tosetti, Weak and Strong Cross Section Dependence and Estimation of Large Panels, June 2009

2690 Mohamed El Hedi Arouri and Christophe Rault, On the Influence of Oil Prices on Stock Markets: Evidence from Panel Analysis in GCC Countries, June 2009

2691 Lars P. Feld and Christoph A. Schaltegger, Political Stability and Fiscal Policy – Time Series Evidence for the Swiss Federal Level since 1849, June 2009

2692 Michael Funke and Marc Gronwald, A Convex Hull Approach to Counterfactual Analysis of Trade Openness and Growth, June 2009

2693 Patricia Funk and Christina Gathmann, Does Direct Democracy Reduce the Size of Government? New Evidence from Historical Data, 1890-2000, June 2009

2694 Kirsten Wandschneider and Nikolaus Wolf, Shooting on a Moving Target: Explaining European Bank Rates during the Interwar Period, June 2009

2695 J. Atsu Amegashie, Third-Party Intervention in Conflicts and the Indirect Samaritan's Dilemma, June 2009

2696 Enrico Spolaore and Romain Wacziarg, War and Relatedness, June 2009

2697 Steven Brakman, Charles van Marrewijk and Arjen van Witteloostuijn, Market Liberalization in the European Natural Gas Market – the Importance of Capacity Constraints and Efficiency Differences, July 2009

2698 Huifang Tian, John Whalley and Yuezhou Cai, Trade Sanctions, Financial Transfers and BRIC's Participation in Global Climate Change Negotiations, July 2009

2699 Axel Dreher and Justina A. V. Fischer, Government Decentralization as a Disincentive for Transnational Terror? An Empirical Analysis, July 2009

2700 Balázs Égert, Tomasz Koźluk and Douglas Sutherland, Infrastructure and Growth: Empirical Evidence, July 2009

2701 Felix Bierbrauer, Optimal Income Taxation and Public Goods Provision in a Large Economy with Aggregate Uncertainty, July 2009

2702 Marc Gronwald, Investigating the U.S. Oil-Macroeconomy Nexus using Rolling Impulse Responses, July 2009

2703 Ali Bayar and Bram Smeets, Government Deficits in the European Union: An Analysis of Entry and Exit Dynamics, July 2009

2704 Stergios Skaperdas, The Costs of Organized Violence: A Review of the Evidence, July 2009

2705 António Afonso and Christophe Rault, Spend-and-tax: A Panel Data Investigation for the EU, July 2009

2706 Bruno S. Frey, Punishment – and beyond, July 2009

2707 Michael Melvin and Mark P. Taylor, The Crisis in the Foreign Exchange Market, July 2009

2708 Firouz Gahvari, Friedman Rule in a Model with Endogenous Growth and Cash-in-advance Constraint, July 2009

2709 Jon H. Fiva and Gisle James Natvik, Do Re-election Probabilities Influence Public Investment?, July 2009

2710 Jarko Fidrmuc and Iikka Korhonen, The Impact of the Global Financial Crisis on Business Cycles in Asian Emerging Economies, July 2009

2711 J. Atsu Amegashie, Incomplete Property Rights and Overinvestment, July 2009

2712 Frank R. Lichtenberg, Response to Baker and Fugh-Berman's Critique of my Paper, "Why has Longevity Increased more in some States than in others?", July 2009

2713 Hans Jarle Kind, Tore Nilssen and Lars Sørgard, Business Models for Media Firms: Does Competition Matter for how they Raise Revenue?, July 2009

2714 Beatrix Brügger, Rafael Lalive and Josef Zweimüller, Does Culture Affect Unemployment? Evidence from the Röstigraben, July 2009

2715 Oliver Falck, Michael Fritsch and Stephan Heblich, Bohemians, Human Capital, and Regional Economic Growth, July 2009

2716 Wladimir Raymond, Pierre Mohnen, Franz Palm and Sybrand Schim van der Loeff, Innovative Sales, R&D and Total Innovation Expenditures: Panel Evidence on their Dynamics, July 2009

2717 Ben J. Heijdra and Jochen O. Mierau, Annuity Market Imperfection, Retirement and Economic Growth, July 2009

2718 Kai Carstensen, Oliver Hülsewig and Timo Wollmershäuser, Price Dispersion in the Euro Area: The Case of a Symmetric Oil Price Shock, July 2009

2719 Katri Kosonen and Gaëtan Nicodème, The Role of Fiscal Instruments in Environmental Policy, July 2009

2720 Guglielmo Maria Caporale, Luca Onorante and Paolo Paesani, Inflation and Inflation Uncertainty in the Euro Area, July 2009

2721 Thushyanthan Baskaran and Lars P. Feld, Fiscal Decentralization and Economic Growth in OECD Countries: Is there a Relationship?, July 2009

2722 Nadia Fiorino and Roberto Ricciuti, Interest Groups and Government Spending in Italy, 1876-1913, July 2009

2723 Andreas Wagener, Tax Competition, Relative Performance and Policy Imitation, July 2009

2724 Hans Fehr and Fabian Kindermann, Pension Funding and Individual Accounts in Economies with Life-cyclers and Myopes, July 2009

2725 Ernesto Reuben and Arno Riedl, Enforcement of Contribution Norms in Public Good Games with Heterogeneous Populations, July 2009