Enforcement of Contribution Norms in Public Good Games with Heterogeneous Populations[*]

Ernesto Reuben

IZA and Columbia University, e-mail: ereuben@columbia.edu

Arno Riedl

CESifo, IZA, and Maastricht University, e-mail: a.riedl@maastrichtuniversity.nl

Abstract

We investigate the emergence and enforcement of contribution norms to public goods in homogeneous and heterogeneous groups. With survey data we demonstrate that uninvolved individuals hold well defined yet conflicting normative views of fair contribution rules related to efficiency, equality, and equity. In the experiment, in the absence of punishment no positive contribution norm is observed and all groups converge towards free-riding. With punishment, strong and stable differences in contributions emerge across group types and individuals in different roles. In some cases these differences result from the emergence of an efficiency norm where all fully contribute. In the cases where full efficiency is not attained, these differences result from the enforcement of different relative contribution norms. Hence, our experimental data show that, even in heterogeneous groups, individuals can overcome the collective action problem inherent in public good games by agreeing on and enforcing a contribution norm.

JEL Codes: H41, C92, D63

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

---

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 frequently differ in their productivity, ability, and motivation (Hamilton et al. 2003). Irrigation systems are often jointly used and maintained by farmers with different plot sizes and water needs.[1] Also, people can derive very different benefits from public goods. For example, the elevation of dams along the Mississippi River are of different value to individuals who live close to the river compared to those who live further away. In the international political and economic arena, countries that greatly 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) discuss this problem in the framework of international military alliances. Other prominent examples of international cooperation include the Kyoto Protocol, which aims to reduce emissions of greenhouse gases, fishing quotas by European Union members to mitigate the overfishing of open waters, and the Global Disease Detection Program spearheaded by the United States that seeks early detection of infectious diseases.

As diverse as the above examples seem, they can all be viewed as special cases of a more general public good problem where the formal enforcement of cooperation by third-parties is infeasible or very limited (e.g., due to high monitoring costs or the absence of a supranational institution with coercive power). For such situations scholars suggest that informal sanctions are used to enforce a social norm of cooperation (Elster 1989; Coleman 1990). Ostrom (1990) describes, among others, the case of fishermen in Alanya, Turkey, who overcame the commons problem through informal rules that are defended violently if necessary. The private lobster management in Maine provides another case where it is reported that free-riders of the informal self-regulatory system were "discouraged by surreptitious violence" (p. 92, Berkes et al. 1989).

The question of contribution norms supporting cooperation becomes especially interesting when taking into account that at least some of the involved parties are motivated by social preferences (Rabin 1993; Fehr and Schmidt 1999). Social preferences transform the social dilemma into a coordination problem with many Nash equilibria, where existing theoretical models give little guidance as to what outcomes to expect. However, if players can (tac-

---

[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 ("Western Irrigated Agriculture Dataset." *United States Department of Agriculture*, July 20, 2004. Retrieved from http://www.ers.usda.gov/data/westernirrigation/methods.htm on September 19, 2011).

itly) agree and, if necessary, punish deviations from specific contribution norms this could help solve the equilibrium selection problem. Indeed, for homogeneous groups, controlled laboratory experiments have shown that cooperative behavior in public good problems can be supported through punishment, and it has been suggested that cooperation is achieved through punishment of deviations from an equal contributions norm. Such evidence is lacking for heterogeneous groups.[2]

In this paper we study, for homogeneous and different heterogeneous groups, whether contribution norms are observed in public goods problems, whether and how they are enforced, and whether they help groups overcome the coordination problem inherent in public goods problems with social preferences. To this end we define a norm as a rule of behavior that is only observed if people are aware that a normatively appealing behavioral rule exists and sufficiently many people follow that rule, either because it is internalized or because of the threat of sanctions (Bicchieri 2006; Young 2008). Specifically, we investigate (i) whether people hold specific normative views regarding rules of contribution behavior (e.g., equal and/or efficient contributions) that could serve as a basis for a contribution norm, and (ii) whether these normative views differ across different types of group heterogeneity. In addition, we examine (iii) whether individuals interacting in homogeneous and heterogeneous groups follow a *specific* rule of contribution behavior, (iv) whether sanctioning is needed for them to follow such a rule, and (v) whether the followed rule depends on the type of group heterogeneity.

---

[2]For recent reviews on homogeneous groups, see Fehr and Fischbacher (2004) and Gächter and Herrmann (2009). In contrast to homogeneous groups, the experimental evidence regarding contributions to public goods in heterogeneous groups is less conclusive. 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. Furthermore, evidence on the effect of sanctioning in heterogeneous groups is sparse. To our knowledge, the only experiment that combines endowment heterogeneity and punishment possibilities is Visser and Burns (2006). They report that punishment effectively promotes cooperation in both homogeneous and heterogeneous groups. Tan (2008) and Noussair and Tan (2011) study groups with heterogeneous productivities regarding the public good and find that punishment does not increase contributions as much in heterogeneous as in homogeneous groups. Nikiforakis et al. (2010) report that heterogeneity in punishment effectiveness, controlling for the average effectiveness, does not affect cooperation levels. Lastly, Reuben and Riedl (2009) and Nikiforakis et al. (2012) investigate groups with heterogeneous benefits from the public good. The former show that punishment is less effective in heterogeneous groups due to the reluctance of subjects with low benefits to increase their contribution after being punished. The latter report lower earnings in heterogeneous groups compared to homogeneous groups due to increased counter-punishment, which they attribute to conflicting normative expectations in heterogeneous groups.

For homogeneous groups it has been shown that high contributions to the public good can be sustained because some high contributors reliably sanction deviations from their own (or the group's average) contribution. In such groups, such behavior is intuitively appealing and consistent with the enforcement of a contribution norm grounded in fairness principles such as efficiency, equality, and, given that individuals are symmetric, also equity.[3]

In heterogeneous groups, however, it is not clear what contribution norm may emerge, if one emerges at all. If people differ, fairness principles of equality, equity, and efficiency will often stipulate different normatively appealing rules of behavior. In such a case, even uninvolved individuals might find it hard to unambiguously answer questions such as: should high-income individuals contribute more to the public good even though they benefit equally from it? Such ambiguity in the normative appeal of fairness principles is also well known in the public finance literature where it is discussed as the benefit-received versus the ability-to-pay principles (see, e.g. Musgrave 2008). In addition, even if individuals have a clear view about the rule of behavior they find normatively appealing, self-serving interpretations of fairness principles (Roth and Murnighan 1982; Babcock and Loewenstein 1997) may make it difficult for individuals in heterogeneous groups to agree on a specific rule that may constitute the basis of a contribution norm. For instance, if people have different tastes for the public good, a rule of equal contributions may be seen as normatively desirable by those who derive a higher benefit from the public good, whereas those who enjoy the public good less may think that such a rule is normatively unappealing.

We investigate three types of group heterogeneity in a linear public good game where people can differ in their endowment, contribution capacity, and the benefits they receive from the public good. Specifically, we look at groups where one person (out of three) receives an endowment that is twice as high as the endowment of the other group members. To isolate the effect of extended contribution possibilities due to a higher endowment, in some groups, we restrict the contribution capacities to be the same for all group members, whereas in others, we allow for contributions up to the entire endowment. Further, we investigate groups where everyone has the same endowment but one group member receives a 50 percent higher marginal benefit from the public good. In addition, we also study homogeneous groups.

---

[3]The principles of efficiency, equality, and equity are commonly called upon in normative research and have been extensively discussed by 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). Frohlich et al. (1987) and Frohlich and Oppenheimer (1990) are seminal experimental investigations into these justice principles. Konow (2003) provides an excellent survey on normative and positive views of justice.

Conceptually, two dimensions of normatively appealing rules of contribution behavior can be distinguished: efficiency and fairness. Accordingly, an *efficiency rule* asks for maximal contributions by all, and a *relative contribution rule* prescribes a contribution that is 'fair' relative to contributions of others.[4] Using a questionnaire study, we find that for both homogeneous and heterogeneous groups, uninvolved people indeed regard contribution behavior that enhances efficiency as normatively appealing. This result is in line with Carpenter and Matthews (2012) who find that third-parties are willing to sanction deviations from efficiency. In addition, normative views related to the fairness principles of equality and equity are very prominent. In homogeneous groups, questionnaire respondents largely agree that contribution behavior that adheres to efficiency and equality is normatively desirable. In heterogeneous groups, however, there is considerable disagreement on the specific contribution rule that is deemed as normatively desirable. Furthermore, these normative views change with the type of group heterogeneity.

In the laboratory experiment, we find that without punishment possibilities, heterogeneity does not matter much. In all groups free-riding is relatively frequent and steadily increases over time. In other words, despite the normative appeal of contribution rules involving positive contributions, we do not find evidence that such a norm emerges. By contrast, with punishment, we observe contribution behavior consistent with adherence to the *efficiency norm* in a significant subset of groups, irrespective of the type of group heterogeneity. Furthermore, in groups that do not achieve full efficiency, we observe behavior consistent with adherence to *relative contribution norms* based on equality and equity. Hence, we find evidence that emerging contribution norms indeed induce coordination on socially preferable outcomes. Interestingly, contribution patterns consistent with different relative contribution norms are observed depending on the type of group heterogeneity. In groups with unequal endowments and unrestricted contribution capacities, contributions are proportional to endowments. By contrast, in groups with unequal endowments but constrained contribution capacities, group members with twice the endowment of other group members do not contribute more. Likewise, in groups with unequal marginal benefits from the public good, contributions are similar irrespective of the marginal benefits.

---

[4]In the field, relative contribution rules to a public good often come in the form of proportionality. For instance, as reported by Gardner et al. (2000), fishing cutbacks enacted by the European Union are proportional to fleet size and CO_2 emissions cutbacks proposed in the Kyoto protocol are proportional to 1990 emission levels. A more recent example is the Eurozone rescue fund in which each member state's share is proportional to their capital in the European Central Bank ("Framework agreement." *European Financial Stability Facility*, June 07, 2010. Retrieved from http://www.efsf.europa.eu/about/legal-documents/ on September 19, 2011).

We show econometrically that deviations from the efficiency norm are sanctioned in all types of groups and that the observed differences in contribution patterns across different types of heterogeneity do not occur accidentally but are due to the *enforcement* of different relative contribution norms. Interestingly, with unequal endowments, individuals largely agree on which contribution norm to enforce—even when the norm implies that some individuals benefit relatively more—whereas with unequal benefits from the public good individuals tend to enforce relative contribution norms in a self-serving manner.

2   Design and Procedures

The game used in both the questionnaire study and the experiment is a linear public good game with groups of three players. The game in its basic form consists of a contribution stage in which each player i receives an endowment of y_i points. Players simultaneously decide how many points they want to contribute to the public good, $c_i \in [0, \bar{c}_i]$ where $\bar{c}_i$ is 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. In other words, i's earnings at the end of the contribution stage are equal to

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

Note that, if $\alpha_i < 1$ and $\sum_i \alpha_i > 1$, then each contributed point strictly increases the sum of earnings in the group but strictly decreases the earnings of the contributing player, which creates a tension between individual and group interest.

We implement four types of groups. The first type is the homogeneous case where 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 refer to this group type as EQUAL. In the remaining group types we introduce heterogeneity. Specifically, in each group, one player receives either a higher endowment or a higher marginal benefit from the public good than the other two players. For convenience we refer to the former as the *high* player and to the two latter ones as *low* players. In the second group type, the high player receives an endowment of $y_H = 40$ points whereas low players get $y_L = 20$ points. Importantly, in these groups contributions are restricted to a maximum of 20 points for all players (i.e., $\bar{c}_L = \bar{c}_H = 20$ points). We refer to this group type as the Unequal Restricted Endowment type or URE. In the third group type, the high player again receives $y_H = 40$ points and low players $y_L = 20$ points. However, in contrast to URE, the contributions of the high player is only restricted

Table 1: Group types

Group type	Player's role	Parameters			Respondents to the questionnaire	Subjects in treatments with	
		y_i	α_i	$\bar{c}_i$		no punishment	punishment
EQUAL	low	20 points	0.50	20 points	39	21	33
URE	low	20 points	0.50	20 points	60	21	33
	high	40 points	0.50	20 points			
UUE	low	20 points	0.50	20 points	64	18	33
	high	40 points	0.50	40 points			
UMB	low	20 points	0.50	20 points	62	21	30
	high	20 points	0.75	20 points			

by its endowment (i.e., $\bar{c}_L = 20$ points and $\bar{c}_H = 40$ points). We refer to this group type as Unequal Unrestricted Endowment or UUE. In the fourth group type, all players receive the same endowment of 20 points but the high player earns a marginal benefit from the public good equal to $\alpha_H = 0.75$ while low players earn $\alpha_L = 0.50$. Correspondingly, we refer to it as the Unequal Marginal Benefit group type or UMB. The four types of groups are summarized in Table 1.

The Questionnaire Study

In order to elicit normatively appealing rules of behavior in the public good game, we conducted an online questionnaire study with uninvolved individuals. Specifically, we asked students from the subject pool of the University of Amsterdam's CREED laboratory to take part in a 15-minute questionnaire. Importantly, the respondents had experience reading experimental instructions but had not participated in any of our experimental sessions (described below). To increase the response rate, students who completed the questionnaire had a 2% chance of receiving €100,-.

Each respondent took part in one of four versions of the questionnaire. Each version corresponds to one of the four group types (i.e., either EQUAL, URE, UUE, or UMB). At the beginning of the questionnaire, respondents were told that they will be answering normative questions concerning an experiment that had been run in the CREED laboratory. To ensure their understanding, respondents read the instructions seen by the subjects that participated in the experiment and answered a series of control questions (the experiment is described below, only instructions of treatments without punishment where used). Respondents who answered the control questions incorrectly were not allowed to continue. In total 225 students

completed the questionnaire. Table 1 contains the number of respondents for each version of the questionnaire.

To effectively elicit normative rules of behavior that are related to the principles of efficiency, equality, and equity, the questionnaire contains two sets of questions. The first set consists of one question asking respondents to specify "what is the *fair* amount that each of the group members should contribute to the group project? [the public good]". Respondents indicated a contribution for each group member for one period of play and could choose any contribution between 0 and $\bar{c}_i$. The second set corresponds to a series of questions of the form: "what is the *fair* amount that group member i and group member j *should* contribute if group member k contributes x tokens to the group project? [the public good]", where x varies across different questions. Respondents indicated a contribution for group members i and j and could choose contributions between 0 and respectively $\bar{c}_i$ or $\bar{c}_j$. In all questions, respondents are asked to put themselves in the position of a *neutral uninvolved arbitrator*. Moreover, in the heterogeneous treatments, respondents are reminded of the characteristics of each group member (i.e., their endowment, maximum contribution, and marginal benefit from the public good). The precise wording of the questions can be found in the online appendix.

We opted for the direct elicitation of normative views with respondents that did not participate in the experiment for two main reasons. First, by using uninvolved respondents we avoid affecting behavior in the experiment with the elicitation of normative views or vice versa (depending on which task precedes the other). Second, since uninvolved respondents have no incentive to misreport their answers to the questionnaire, we can elicit normative views by simply asking respondents for their first-order normative expectations. Third, direct elicitation has the advantage of simplicity over incentivized elicitation of normative views.[5]

---

[5]As any hypothetical measure, asking for normative views has the disadvantage that noise might be introduced due to respondents being uninterested. However, unlike with many hypothetical measures, in our case we possess theoretical benchmarks that can help us determine whether respondents are answering arbitrarily (Section 3). Krupka and Weber (2012) develop a clever method to elicit normative expectations that would be more appropriate than our questionnaire if our goal was to elicit normative second-order expectations. Also, note that we elicit the unbiased *normative* views of respondents. That is, we do not ask them to put themselves into the role of a high or a low player. We do this to stay close to the theoretical frameworks of social norms, which are based on unbiased normative views (e.g., Bicchieri 2006; Young 2008). The drawback of this approach is that we do not capture normative disagreement due to self-serving biases, which might make the emergence of a contribution norm less likely in heterogeneous groups. See Bernard et al. (2012) for a study that elicits the biased normative views of involved participants in step-level public good games.

The Laboratory Experiment

To study the emergence and enforcement of contribution norms when individuals actually interact with each other, we conducted the laboratory experiment. Each subject took part in one of eight treatments, which vary along two dimensions: (i) the type of group they are in (either EQUAL, URE, UUE, or UMB), and (ii) whether or not they have the option to punish other group members. In all our treatments, subjects interacted in the same group of three subjects for ten consecutive periods. Moreover, in heterogeneous groups, subjects were randomly assigned to roles (high or low) at the beginning of the first period and kept the same role throughout the experiment.

In treatments without punishment, subjects simply play the public good game previously described. In treatments with punishment, 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 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.[6] In each period, at the end of the contribution stage, every player i sees the value of c_j and π_j of each other group member j. The values of y_j, α_j, $\bar{c}_j$ are also shown. In treatments with punishment, subjects are informed of the total number of punishment points assigned to them after they play the punishment stage. As in Fehr and Gächter (2000, 2002), subjects do not receive specific information concerning who punished whom. In treatments without punishment, earnings at the end of a period correspond to earnings after the contribution stage. In treatments with punishment, the earnings of a subject i at the end of a period are given by:[7]

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

---

[6]Subjects can identify other group members through a randomly assigned ID number that remained constant throughout the experiment. We impose an upper limit on the amount of punishment i can assign to each j (as Fehr and Gächter 2002, and others). This restriction prevents subjects with higher earnings from having the capacity to punish more than subjects with lower earnings. In the experiment, this restriction does not seem to have censored behavior (only 22 out of 2580 punishment decisions equaled 10 points).

[7]In order to avoid excessive losses, subject i in fact earns: $\pi_i = \max[0, y_i - c_i + \alpha_i \sum_j c_j - 3 \sum_{j \neq i} p_{ji}] - \sum_{j \neq i} p_{ij}$. This way, i cannot be punished below zero points by others, but must always pay to punish others (we always allow i to punish fully). That is, even if i expects that $y_i - c_i + \alpha_i \sum_j c_j - 3 \sum_{j \neq i} p_{ji} \leq 0$, i can pay the cost of punishing others by incurring negative earnings in that period (see also, Fehr and Gächter 2002). In the experiment, negative earnings in a period occurred in only 6 (out of 1290) instances.

The computerized experiment was conducted in the CREED laboratory using the typical procedures of anonymity, neutrally worded instructions, and monetary incentives. In total, 210 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 fields such as biology, engineering, political science, and law). Mean earnings equaled €13.83 (≈US$17.50).

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). 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 in z-Tree, Fischbacher 2007) started. After the ten periods, subjects answered a short debriefing questionnaire and were confidentially paid their earnings in cash.

3   Focal and Conflicting Contribution Norms

The experimental literature has led to the development of models of social preferences.[8] While some predictions from these models are indeed consistent with observed behavior in public good games, most models will in fact generate a large number of equilibria most of which are not empirically observed. Specifically, it has been shown that models of social preferences transform public goods games into coordination games with multiple equilibria, some with and some without positive contributions, even for homogeneous groups (e.g., Rabin 1993 and Propositions 4 and 5 in Fehr and Schmidt 1999).

Therefore, our main interest is not in the role of other-regarding preferences *per se*, but instead concerns the possible emergence of *contribution norms* in the homogeneous and heterogeneous groups as a coordination device. This is in concordance with Young (2008), who argues that the term "norm" can only apply to an equilibrium in games with multiple equilibria.[9]

For a norm to exist and to be observed the following has to be satisfied (Bicchieri 2006; Young 2008). A social norm is a behavioral rule, for which it must hold for sufficiently many

---

[8]For example, Levine (1998), Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), Dufwenberg and Kirchsteiger (2004), Falk and Fischbacher (2006), and Cox et al. (2007).

[9]For an alternative view in the tradition of Sudgen (1986) and Coleman (1990), see Bicchieri (2006) who argues that norms enforce non-equilibrium behavior in situations where there is a tension between individual and collective material welfare.

people that: first, they know that such a (normatively desirable) behavioral rule exists; second, they are motivated to follow the rule *under the condition* that sufficiently many others will (a) also conform to the rule, and (b) expect them to follow the rule and are willing to sanction deviations from the rule.[10]

Importantly, this definition allows a social norm to exist while not always being followed and, hence, not necessarily emerge as observable behavior. Only if sufficiently many people have the appropriate expectations about others' behavior and, if necessary, are able and ready to punish transgressions, an existing social norm will be consistently followed. In addition, the existence of a social norm does not imply that all people care with the same strength about norm compliance.[11] Together, this implies that even if all involved people have the 'right' empirical or normative expectations, actual sanctioning may have to take place in order to make people with a weak inclination toward norm obedience follow the norm. Lastly, note that this definition of a social norm does not specify the motivation behind individuals' willingness to sanction deviations from the rule. In other words, it allows for sanctioning that is due to a combination of intrinsic motivations (e.g., social preferences) and more instrumental motivations such as maximizing future payoffs in games with repeated interaction.

To investigate the emergence of contribution norms in the public goods game, we use the laboratory experiment. First, we examine whether subjects follow a *specific* rule of contribution behavior. Second, we test whether sanctioning is needed for them to follow such a rule, and third, we discern whether contribution behavior differs across homogeneous and (the various forms of) heterogeneous groups. The questionnaire study informs us of the existence of normatively appealing rules of behavior and, hence, of likely candidates of contribution norms. Moreover, it also indicates whether people enter the game as a "blank slate" and learn a rule of behavior as they interact, or whether there is a plurality of predetermined normatively appealing rules and people arrive at a commonly acceptable one. In the following we refer to a contribution pattern as a contribution *norm* if it conforms to a normatively appealing rule of behavior and is observed in the experiment.

---

[10]Elster (1989) also defines a social norm in a similar way when he writes that "[for] norms to be *social*, they must be (a) shared by other people and (b) partly sustained by their approval and disapproval" and "norms are *social* [in] that other people are important for enforcing them" (p. 99, emphasis in original).

[11]Ostrom (2000) observes that "social norms may lead individuals to behave differently in the same objective situation depending on how strongly they value conformance with (or deviance from) a norm" (p. 144) and Andreoni and Bernheim (2009) find that subjects in a dictator game experiment adhere with different intensities to the 50-50 division norm (see also, Krupka and Weber 2012).

Hypotheses

We hypothesize that normatively appealing rules of contribution behavior have two dimensions related to motivations that are normatively and behaviorally important. The first dimension relates to the maximization of collective welfare (for evidence see, Charness and Rabin 2002; Engelmann and Strobel 2004), which can be thought of as an *efficiency rule* that prescribes that one ought to contribute as much as possible to the public good.[12] The second dimension relates to reciprocity, which in public good games translates into conditional cooperation (for evidence see, Keser and van Winden 2000; Fischbacher et al. 2001; Fischbacher and Gächter 2010) or, in our context, a *relative contribution rule*, meaning that one should contribute a 'fair amount' in relation to what others contribute. What such a fair amount is may depend on the circumstances, which we discuss in detail below.[13]

Since collective welfare increases with contributions in all investigated group types, we expect to find support for the desirability of an *efficiency rule* in homogeneous as well as heterogeneous groups. However, since social norms are not absolute prescriptions of behavior, such as 'always contribute everything to the public good', but instead are to be followed only if sufficiently many others do so as well, we do not expect maximal contributions to be the (only) normatively desirable rule under all circumstances. In particular, this might be the case in situations where maximum material efficiency is unattainable or when contributing maximally strongly violates conditional cooperation.

In all group types, we also expect support for the desirability of a *relative contribution rule*. Given the symmetry of homogeneous groups (EQUAL), we expect questionnaire respondents to largely agree that the most attractive relative contribution rule is for everyone to contribute an equal amount. In heterogeneous groups, it is less obvious what the most appealing relative contribution rule will be. Since, an important characteristic of social norms is that they are "local and context dependent" (Bicchieri 2008, p. 229), in heterogeneous groups, the normative desirability of the relative contributions of high and low players may depend on particularities that may be interpreted differently by people in different roles. Yet, the literature on fair allocation rules provides two prominent principles that can be used to discuss the appeal of different relative contribution rules: equality and equity (Konow 2003; Konow et al. 2009).

---

[12]Throughout the paper, the term "efficiency" refers to the sum of material payoffs and not to Pareto efficiency.

[13]In homogeneous groups reciprocally fair contributions lead to fair outcomes in the sense that equal contributions imply equal earnings. As we will discuss below this one-to-one relation between fairness in contributions and fairness in earnings does not necessarily hold for heterogeneous groups.

Table 2: Normatively appealing relative contribution rules

	Equality of contributions	Equality of earnings	Proportional to endowments y_i	Proportional to capacities $\bar{c}$	Proportional to benefits α
EQUAL	$c_i = c_j$	$c_i = c_j$	$c_i = c_j$	$c_i = c_j$	$c_i = c_j$
URE	$c_H = c_L$	$c_H = 20, c_L = 0$	$c_H = 2c_L$	$c_H = c_L$	$c_H = c_L$
UUE	$c_H = c_L$	$c_H = 20 + c_L$	$c_H = 2c_L$	$c_H = 2c_L$	$c_H = c_L$
UMB	$c_H = c_L$	$c_H = 2c_L$	$c_H = c_L$	$c_H = c_L$	$c_H = 1.5c_L$

Equality is generally thought of as the equalization of outcomes with no necessary link to individual characteristics such as capacity. By contrast, equity is mostly interpreted as the dependence of fair outcomes—in a proportional way—on effort or individual characteristics such as ability.

If people apply the principle of equality to contributions, then it trivially follows that the equal contributions rule will be the one that is viewed as normatively most attractive in all our studied group types. However, if instead they apply equality of earnings or proportionality to endowments, capacities, or marginal benefits than the implied contributions by the different player roles differ substantially across group types. For instance, in UUE, proportionality applied to capacities $\bar{c}_i$ implies a relative contribution rule in which high players contribute twice as much as low players, whereas in URE it implies equal contributions of both player roles. Table 2 summarizes the relative contribution rules implied by the various interpretations of equality and equity when applied to the group types under investigation. It is obvious from the table that these fairness principles considerably narrow down the set of relative contribution rules that respondents might find normatively attractive. However, only in homogeneous groups do all principles lead to one focal relative contribution rule. In heterogeneous groups, there is still a variety of normatively appealing rules of behavior, making it difficult to say *a priori* which one will be the most preferred.

The existence of normatively appealing rules of behavior in people's minds does not imply that a contribution norm will emerge in an experiment because it also has to be followed. In the absence of punishment, this may happen at positive contribution levels if the norm is fully internalized (see, Elster 1989; Bicchieri 2006; Young 2008). However, in the absence of a fully internalized contribution norm, the nonexistence of an effective tool to sanction deviations can easily lead to noncompliance and low contribution levels.[14] Since there is no *a priori* reason

---

[14]The existence of a normatively desirable contribution rule does not imply that everyone is equally willing to adhere to it (see footnote 11). Conversely, given the conditional character of following norms, not observing

to assume that the general willingness to comply varies with the type of group heterogeneity, we expect low and decreasing contribution levels in all groups without punishment.

If it is possible to punish others, existing experimental evidence from homogeneous groups leads us to expect high contribution levels that do not decline with repetition (Fehr and Gächter 2000; Gächter and Herrmann 2009). In addition to higher contributions, we can make predictions concerning punishment patterns based on the discussion above.[15] Specifically, we expect that (some) subjects will use sanctions to (try to) enforce the efficiency rule. Since higher contribution levels increase collective welfare in homogeneous and heterogeneous groups, deviations from the maximum contribution will be punished similarly in all group types. Furthermore, in homogeneous groups we expect that punishment will be used to (try to) enforce a relative contribution rule of equal contributions by all group members. In heterogeneous groups, we expect punishment to be in accordance with relative contribution rules based on equality or equity. However, due to the multiplicity of normatively attractive relative contribution rules, it is difficult to tell *a priori* which specific rule will be used.

4   Empirical results

This section consists of two parts. In the first part (Section 4.1), we use our questionnaire study to analyze the rules of behavior that individuals uninvolved in the public good game find normatively appealing. We seek an answer to the following three questions. Do respondents' answers conform to the fairness principles discussed in the previous section? Do they agree on the same rule of behavior? And, do they coincide or differ across the different group types? In the second part (Sections 4.2 and 4.3), we study the behavior of people involved in the public good game. We first analyze the contributions to the public good in all treatments with and without punishment. We focus on the behavior of high and low players. In particular, we analyze whether players in different roles display different contribution patterns in the different treatments and whether these patterns are related to the discussed and elicited normatively appealing rules of behavior. In other words, whether observed behavior is consistent with the emergence of (different) contribution norms. Lastly, we investigate econometrically whether observed contribution norms are enforced through punishment. Throughout the paper we report p-values of two-sided tests.

compliance to a normatively desirable rule does neither imply that people do not have normative views in mind nor that they are principally unwilling to follow such a rule.

[15]Predicting precise contributions levels is impossible without knowing the subjects' precise willingness to comply with norms as well as their willingness to sanction norm violators.

4.1 Normative rules of behavior

To investigate whether respondents select the efficiency rule, we analyze the first question of the questionnaire. In this question, respondents state, for each of the three group members, what the fair contribution to the public good is (see Section 2). We find that in all group types the modal answer stipulates for all group members to contribute the maximum amount $\bar{c}_i$. Specifically, in EQUAL 59 percent of the respondents say that contributing fully is the fair rule. Interestingly, in the heterogeneous groups this percentage decreases to 50, 39, and 31 percent in URE, UUE, and UMB, respectively. A likelihood ratio χ^2-test shows that the differences across group types are statistically significant ($p = 0.022$).[16]

In all group types, almost all deviations from the efficiency rule are consistent with one of the relative contribution rules described in Section 3. In EQUAL only 3 percent of all choices are inconsistent with either the efficiency rule or one of the discussed relative contribution rules. In heterogeneous groups, this percentage increases slightly but is still quite low: 8 percent in URE, 9 percent in UUE, and 15 percent in UMB. Hence, the data confirm that the efficiency rule is a normatively appealing rule of behavior for many respondents and that an overwhelming majority of the answers that imply inefficiency are consistent with the relative contribution rules discussed in Section 3. It is important to note that the efficiency rule coincides with different relative contribution rules in the different group types, which makes it impossible to disentangle the weights given by respondents to the various relative contribution rules. To deal with this problem we use the answers to the second set of questions, which exclude fully efficient outcomes by design. As described in Section 2, we asked respondents to indicate the fair contributions of i and j given that k contributes $c_k < \bar{c}_k$ points to the public good. We varied k's type (either high or low) and the amount c_k so that the discussed relative contribution rules can be identified as clearly as possible in each group type (see the online appendix for the specific questions and answers used to identify each rule).

Table 3 presents for each group type the fraction of answers that coincides with the discussed relative contribution rules.[17] For completeness, it also reports the fraction of answers consistent with the efficiency rule (interpreted as choosing $c_i = \bar{c}_i$ and $c_j = \bar{c}_j$ irrespective of c_k) as well as the fraction that is not consistent with any of the discussed rules. Lastly, in cases where various rules inevitably overlap (as they do in EQUAL), we assign the answers to the equal contributions column and leave the others blank.

---

[16] Pair-wise comparisons reveal that, compared to EQUAL, the reported percentages are significantly smaller in UUE and UMB ($p \leq 0.049$). This percentage is also smaller in UMB compared to URE ($p = 0.029$).

[17] We do not differentiate relative contribution rules based on proportionality to capacities, $\bar{c}$, because they coincide with other contribution rules (see Table 2).

Table 3: Fraction of answers coinciding with selected contribution rules

Note: Fraction of answers to the second set of questions in the questionnaire study that coincides with the contribution rules discussed in Section 3. In group types where various rules inevitably overlap answers are assigned to the equal contributions rule.

Group type	Equality of contributions	Proportional to endowments	Proportional to benefits	Equality of earnings	Efficiency	Other
EQUAL	0.74	–	–	–	0.13	0.13
URE	0.32	0.24	–	0.14	0.03	0.27
UUE	0.12	0.36	–	0.26	0.02	0.24
UMB	0.16	–	0.24	0.29	0.00	0.31

As can be seen in the table, when fully efficient outcomes are unattainable due to individual k contributing $c_k < \bar{c}_k$, there is a strong decrease in normative appeal of the efficiency rule. By contrast, a large majority of respondents indicate contributions that are consistent with at least one of the discussed relative contribution rules. This is clear evidence that the relative contribution rules based on equality and equity are normatively appealing rules of behavior.[18] However, as hypothesized in Section 3, only in homogeneous groups there is widespread agreement on how one should behave (i.e., in accordance with the equal contribution rule). In heterogeneous groups there is no consensus on a normatively appealing way to behave, and the appeal of the different relative contribution rules varies with the type of group heterogeneity. In URE the modal normative rule is equal contributions while in UUE it is contributions proportional to the (unequal) endowments. Compared to unequal endowments, the diversity of normatively appealing rules is highest in UMB. Moreover, a relatively large fraction of respondents in this group type report contributions that do not fit with any of the discussed relative contribution rules. Hence, it seems that unequal benefits from the public good engender more disagreement on the normative appeal of the discussed rules of behavior. A likelihood ratio χ^2-test rejects the hypothesis that the distribution of relative contribution rules is the same across the three group types ($p = 0.049$). Given this plurality of normatively attractive behavioral rules in heterogeneous groups, it is unclear whether behavior consistent with any of the discussed contribution rules will be observed and enforced when people have stakes to gain or lose. We explore this next by examining behavior in the experiment.

---

[18]We use a strict categorization rule that does not allow for errors. If we allow respondents to make an error of 1 point when answering each question then, overall, the category "other" shrinks from 26 percent to 17 percent (the results with this categorization is presented in the online appendix). Of the remaining choices, another important component are prominent numbers such as $c_i = c_j = 10$, $c_i = c_j = 5$, or $c_i = c_j = 0$.

Table 4: Descriptive statistics

Note: Average contributions and punishment points given depending on the subjects' role and treatment. Averages are calculated for all periods and for the last five periods. Standard deviations using group averages as the unit of observation are in parentheses.

		All periods				*Last 5 periods*			
		EQUAL	URE	UUE	UMB	EQUAL	URE	UUE	UMB
		WITHOUT PUNISHMENT							
Contributions	low	4.21	6.81	7.48	8.63	2.23	4.24	5.18	6.71
		(2.00)	(2.97)	(5.94)	(4.96)	(2.11)	(2.52)	(5.50)	(5.08)
	high	–	9.41	9.05	9.87	–	5.63	4.10	7.77
		–	(3.94)	(7.42)	(5.18)	–	(4.22)	(4.29)	(7.18)
		WITH PUNISHMENT							
Contributions	low	16.22	15.68	15.39	12.21	16.87	15.58	16.44	12.24
		(2.88)	(3.26)	(4.63)	(5.16)	(3.62)	(4.87)	(4.78)	(5.95)
	high	–	15.32	28.31	14.59	–	15.96	30.22	14.34
		–	(5.31)	(10.27)	(5.61)	–	(5.62)	(10.14)	(6.47)
Punishment	low→low	0.51	0.40	0.30	0.45	0.40	0.42	0.19	0.53
		(0.36)	(0.50)	(0.41)	(0.37)	(0.40)	(0.72)	(0.33)	(0.47)
	low→high	–	0.88	0.66	0.69	–	0.86	0.42	0.71
		–	(0.96)	(0.61)	(0.79)	–	(1.15)	(0.61)	(0.87)
	high→low	–	0.48	0.72	0.71	–	0.41	0.45	0.46
		–	(0.55)	(0.70)	(0.60)	–	(0.64)	(0.74)	(0.39)

4.2 Contribution and punishment behavior

Table 4 reports the average contributions across all periods depending on the treatment and subjects' role. Any emerging contribution norm should be reflected by persistent differences in behavior. Therefore, we also report the averages for the second half of the game. When we use nonparametric tests to analyze the experiment's data, we use the groups' averages for the relevant subject role and periods as the observations for the tests.[19]

Table 4 reveals that, without punishment, the difference in average contributions between high and low players is small and disappears over time. Taking all periods into account, Wilcoxon signed-rank tests (henceforth WSR tests) do not detect a statistically significant

---

[19]For example, if we compare the contributions of different player roles in a treatment then the average contribution of subjects in a role over all periods in each group represents an observation. Similarly, if we test whether there is a time trend in contributions in a treatment then the average contribution in each period in each group is considered as an observation.

difference in contributions between high and low players in URE and UUE ($p > 0.128$), and a difference in UMB that is almost significant at the 5 percent level ($p = 0.051$). In the last five periods, the contributions of high and low players are statistically indistinguishable from each other (WSR tests, $p > 0.205$). The main reason for this lack of significant differences is that, with repetition, contributions decrease towards zero in all group types. Using Page's trend test we can accept the alternative hypothesis that average group contributions decrease across periods in all group types for both low ($\rho \leq -0.474$, $p \leq 0.001$) and high players ($\rho \leq -0.301$, $p \leq 0.010$).[20] In other words, despite the fact that most uninvolved individuals think that high types should contribute more and indicate that the efficiency rule is normatively appealing (Section 4.1), without punishment opportunities (almost) full free-riding by both player types emerges as the prevalent behavior.

By contrast, when punishment is possible no significant decline in the contributions of either high or low players in any group type is observed (Page's correlation coefficients $\rho \geq -0.060$, $p > 0.569$). Consequently, overall contributions in all group types are higher with punishment than without.[21] However, the size of the increase caused by the opportunity to punish varies considerably across group types. For instance, the contributions of low players increase by 12.01 points in EQUAL but only by 3.58 points in UMB. Similarly, the contributions of high players increase by 19.26 points in UUE but only by 4.72 points in UMB. Using Fligner-Policello robust rank order tests (henceforth RRO tests), we find that the introduction of punishment leads to a significant increase in the contributions of high and low players in all group types ($p \leq 0.030$) except for low players in UMB ($p = 0.074$).[22]

In addition, introducing punishment opportunities has a strong differential effect on the contributions of high and low players, which also differs across group types. In URE, high and low players contribute almost the same amount: 15.32 vs. 15.68 points on average (WSR test, $p = 0.789$). This stands in stark contrast to UUE where the average contribution of high players is almost twice as high as that of low players: 28.31 points vs. 15.39 points

---

[20]Throughout the paper we use Page's trend test to test whether there are significant time trends. This test is similar to Spearman's rank correlation coefficient, but it allows for dependence across periods within groups (see Page 1963). To provide a measure of magnitude, we also report Page's correlation coefficient ρ.

[21]As in many studies, we do not find a significantly positive effect of punishment on earnings. However, while earnings decrease over time in groups without punishment, they increase in groups with punishment. The online appendix contains a detailed analysis of earnings.

[22]The RRO test is similar to the the commonly used Mann-Whitney U test but it is more appropriate when the samples come from populations with different higher-order moments (see Fligner and Policello 1981), which is the case in our data (see the unequal variances reported in Table 4.)

(WSR test, $p = 0.003$). In UMB, we observe that high players contribute a bit more than low players: 14.59 vs. 12.21 points (WSR test, $p = 0.114$). In other words, introducing punishment possibilities has the strongest effect on the behavior of high players in UUE. In this group type, high players contribute about six times more with punishment than without punishment, whereas in the other two group types there is 'only' a two- to threefold increase in contributions. By contrast, the contributions of low players are very similar across all group types (including EQUAL).[23]

In Table 4 (bottom panel), we also show the average number of punishment points assigned to each other subject depending on the role of the punishing and the punished subject. We do not observe any significant differences in the amount of punishment given or received by high and low players within each group type (WSR tests, $p > 0.140$). Hence, the observed differential contribution patterns across roles are not accompanied by significant differences in the amount of punishment.

Notably, despite the multiplicity of equilibria and normatively appealing contribution rules in heterogeneous groups, there are clear contribution patterns observed. First, the relatively high contributions suggest a role for the efficiency rule in all group types.[24] Second, the distinctive differences in contributions between group types and player roles point toward the importance of different relative contribution rules. The latter also indicates coordination on different equilibria depending on the kind of heterogeneity. In URE, contributions are consistent with a rule of equal contributions. In UUE, behavior is in line with a rule of contributions proportional to endowments. Only in UMB, the relative contributions of high and low players cannot be clearly attributed to one of the normative relative contribution rules: high players contribute only slightly more (around 20%) than low players. In the following section we explore in detail whether differences in the way punishment is applied are the basis for the emergent contribution patterns and, hence, whether we can speak of punishment being used to enforce contribution norms.

---

[23]All the reported significant differences also hold if we concentrate only on the last five periods. The same is true for differences that are not statistically significant except that high players in UMB contribute significantly more than low players (WSR test, $p = 0.047$).

[24]High contributions can be the result of people anticipating sanctions if they do not follow the efficiency rule or the result of people holding higher expectations of compliance, which alone can increase the willingness to comply.

4.3 Emergence and enforcement of contribution norms

In order to analyze the enforcement of contribution norms, we build on our discussion in Section 3 and the results of the questionnaire study (Section 4.1), which point to the importance of an efficiency rule and relative contribution rules as the basis for contribution norms in public good games.

We define *efficient* groups as those that attain maximal contributions by all group members in at least five periods and the rest as *inefficient* groups. There are 5 (out of 11) efficient groups in EQUAL and 4 (out of 11 groups in URE and UUE and 10 in UMB) in each heterogeneous group type.[25] In these groups, deviations from full efficiency are severely punished, which is consistent with the emergence of the efficiency norm. In inefficient groups, although overall average punishment is not lower, deviations from full efficiency are punished less.[26] Hence, with punishment about 40% of all the homogeneous and heterogeneous groups solve the coordination problem by successfully enforcing the efficiency norm.

The results from the questionnaire show that relative contribution rules are particularly noticeable in normative terms when full efficiency is not attained. Further, in efficient groups there is a behavioral equivalence between the efficiency rule and some relative contribution rules, which makes the identification of the enforcement of these rules problematic. Indeed, in efficient groups, the contributions of high and low players are approximately equal in URE (19.13 vs. 18.99 points) and UMB (18.98 vs. 17.45 points), whereas in UUE they are approximately proportional to endowments (38.75 vs. 18.21 points for high and low players, respectively). Therefore, in the following, we concentrate on *inefficient* groups, which allows a clean analysis of the emergence and enforcement of relative contribution norms. For completeness, we redid the subsequent analysis using data from all groups in the online appendix. The results are similar and are discussed later in the paper.

---

[25] In all group types, efficient groups account for 85% of the periods with maximal contributions (see the online appendix for the complete distribution). Although the mean number of periods in which groups attain efficiency varies between treatments, the differences are not statistically significant (Kruskal-Wallis test, $p = 0.837$).

[26] Regressing the number of punishment points received from others on the subjects' deviation from their maximal contribution indicates that a deviation of 1 point in EQUAL significantly increases punishment by 0.58 points, which decreases earnings by 1.74 points ($p \leq 0.001$, linear estimates of Tobit regressions with subject random effects and clustering on groups). The same calculation yields an earnings reduction of 1.91 points in URE, 1.82 points in UUE, and 1.13 points in UMB ($p \leq 0.001$ for all). Compared to efficient groups, deviations from maximal contributions are punished less in inefficient groups (from 0.34 points less in UUE to 0.24 points less in URE, $p \leq 0.044$ in all except UMB where $p = 0.075$). Note that although a 1-point deviation is punished less in inefficient groups, such groups have larger deviations, which is why the overall amount of punishment is similar.

The average contributions in inefficient groups are well below the maximum (68 percent in URE, 65 percent in UUE, and 48 percent in UMB), which implies that high players are not artificially constrained and could in principle contribute substantially more than low players. Nevertheless, we see relative contribution patterns similar to those in efficient groups. In URE the contributions of high and low players are roughly equal (13.14 vs. 13.79 points) and not significantly different from each other (WSR test, $p = 0.735$). By contrast, in UUE the contributions of high players are only slightly less than twice of those of low players (22.34 vs. 13.78 points) and significantly higher (WSR test, $p = 0.018$). In UMB the difference between high and low players is in between that of URE and UUE (11.67 vs. 8.71 points, WSR test, $p = 0.249$). Inefficient groups also display stable contributions, which implies that the differences between high and low players are also stable over time.[27] In the remainder of this section, we explore in detail whether the difference in contribution behavior across group types and roles can be explained by subjects punishing according to different relative contribution rules.

Compared to the efficiency rule, where any contribution below the maximum constitutes a deviation, evaluating whether subjects punish deviations from a relative contribution rule is not straightforward. In homogeneous groups, where there is consensus on the normatively desirable rule of behavior, it seems reasonable that individuals punish deviations from equal contributions. In heterogeneous groups, where a variety of rules are seen as normatively appealing, it is *a priori* unclear which rule, if any, will be enforced. Therefore, we opt for a flexible econometric approach and elicit the relative contribution rule that is most consistent with the punishment data without imposing strong restrictions on the potentially enforced norm.[28]

Specifically, we assume that, *ceteris paribus*, subject i punishes j depending on how much their contributions differ according to the following expression: $(1-\mu)c_j - \mu c_i$, where the term $\mu \in [0,1]$ captures the relative contribution rule used to compare one's contribution to those

---

[27]For all group types, Page's trend test cannot reject the null hypothesis that average contributions are nondecreasing across periods (for low players $p \geq 0.858$ and for high players $p \geq 0.114$). Moreover, if we focus on the last five periods, the contributions of high and low players are: 13.94 and 13.06 points in URE (WSR test, $p = 0.445$), 24.63 and 15.26 points in UUE (WSR test, $p = 0.018$), and 10.73 and 8.48 points in UMB (WSR test, $p = 0.249$).

[28]Our approach is similar to that of Carpenter and Matthews (2009), who elicit norms in public good games with homogeneous groups. They find that the best description of the data is attained with a norm akin to the enforcement of our efficiency rule. However, they do not consider, as we do, that subjects might be punishing deviations from both efficiency and the contributions of others relative to their own.

of others. For example, if $\mu = 0.50 = 1 - \mu$ then the above expression is negative if (and only if) $c_j < c_i$ (i.e., the rule prescribes equal contributions), which from i's perspective implies a negative deviation from the rule by j. Alternatively, if $\mu = 0.75$ and therefore $1 - \mu = 0.25$ then subject i considers that a negative deviation from the rule occurs when $c_j < 3c_i$ (i.e., the rule prescribes that j should contribute three times as much as i). In the extremes, if $\mu = 0$, the rule prescribes that i contributes everything and j contributes nothing, and vice versa if $\mu = 1$. In EQUAL, as subjects are in symmetric positions, we estimate one value for μ. In the heterogeneous treatments, we distinguish between roles and estimate a value for μ in each of the following cases: high players punishing low players ($\mu^{H \to L}$), low players punishing high players ($\mu^{L \to H}$), and low players punishing low players ($\mu^{L \to L}$). Specifically, we estimate the following model in each group type:

$$p_{ijkt}^{H \to L} = \beta_{pos}^{H \to L} \max\left[(1 - \mu^{H \to L})c_{jt} - \mu^{H \to L}c_{it}, 0\right] + \beta_{neg}^{H \to L} \max\left[\mu^{H \to L}c_{it} - (1 - \mu^{H \to L})c_{jt}, 0\right]$$
$$+ \lambda(\bar{c}_j - c_j) + \gamma^p t + \gamma^r \sum_{j \neq i} p_{jikt-1} + \alpha^{H \to L} + \eta_k + \upsilon_i + \epsilon_{ijt}$$

$$p_{ijkt}^{L \to H} = \beta_{pos}^{L \to H} \max\left[(1 - \mu^{L \to H})c_{jt} - \mu^{L \to H}c_{it}, 0\right] + \beta_{neg}^{L \to H} \max\left[\mu^{L \to H}c_{it} - (1 - \mu^{L \to H})c_{jt}, 0\right]$$
$$+ \lambda(\bar{c}_j - c_j) + \gamma^p t + \gamma^r \sum_{j \neq i} p_{jikt-1} + \alpha^{L \to H} + \eta_k + \upsilon_i + \epsilon_{ijt}$$

$$p_{ijkt}^{L \to L} = \beta_{pos}^{L \to L} \max\left[(1 - \mu^{L \to L})c_{jt} - \mu^{L \to L}c_{it}, 0\right] + \beta_{neg}^{L \to L} \max\left[\mu^{L \to L}c_{it} - (1 - \mu^{L \to L})c_{jt}, 0\right]$$
$$+ \lambda(\bar{c}_j - c_j) + \gamma^p t + \gamma^r \sum_{j \neq i} p_{jikt-1} + \alpha^{L \to L} + \eta_k + \upsilon_i + \epsilon_{ijt}.$$

The variable $p_{ijkt}^{H \to L}$ is the amount of punishment points assigned (not inflicted) by a subject i in the high role to a subject j in the low role in group k and period t. The variables $p_{ijkt}^{L \to H}$ and $p_{ijkt}^{L \to L}$ are analogously defined. The first two terms in each equation capture the effect of deviations from the relative contribution rule $\mu^{H \to L}$, $\mu^{L \to H}$, or $\mu^{L \to L}$. As can be seen, by estimating two coefficients for each rule, we allow for the possibility that positive and negative deviations from the rule are punished with different intensities. The third term in the equations, $(\bar{c}_j - c_j)$, is simply the amount of points j did not contribute to the public good, which captures the effect of deviations from the efficiency rule. We estimate the same coefficient λ in all equations because the efficiency rule ought to be enforced equally by high and low players. That is, unlike relative contribution rules, the efficiency rule does not prescribe different behavior depending on a player's type. The next two terms control for motivations for punishment that are generally thought to be unrelated to the enforcement of norms. Specifically, we include the period number, t, to capture strategic motivations such as punishing in early periods to elicit cooperation in latter periods, and a lagged variable for

21

the punishment received in the previous period, $\sum_{j \neq i} p_{jikt-1}$, to control for punishment that might be motivated by revenge (Denant-Boemont et al. 2007; Nikiforakis 2008; Hopfensitz and Reuben 2009).[29] Finally, $\alpha^{H \to L}$, $\alpha^{L \to H}$, and $\alpha^{L \to L}$, respectively, corresponds to the constant in each equation, η_k to group dummy variables, v_i to unobserved individual characteristics, and ϵ_{ijt} is the error term.

To find the combination of μ's that best explains the data, we repeatedly estimate the model—using the simplex method developed by Nelder and Mead (1965)—until we find the values of $\mu^{H \to L}$, $\mu^{L \to H}$, and $\mu^{L \to L}$ (rounded to two decimal points) that give the best fit.[30] Given that the amount of punishment is bounded, we use Tobit estimates. Furthermore, we treat the unobserved individual characteristics as random effects and we assume that deviations from the enforced rules have a nonnegative effect on punishment (i.e., we restrict the values of all β's as well as that of λ to be greater than or equal to zero). Note that we do not restrict the values of the different μ's to be consistent with the enforcement of a mutually shared contribution rule. In other words, we allow for cases where low players are punishing high players according to one rule, whereas high players are punishing low players according to another rule.

In all group types, we find a unique combination of μ's ($\mu^{*H \to L}$, $\mu^{*L \to H}$, and $\mu^{*L \to L}$) that globally maximizes the log likelihood function. The result is visualized in Figure 1, which shows the log likelihood of the estimated model as we vary one of the μ's keeping the other two μ's constant at their optimal value. For convenience and the sake of comparison, we normalize the log likelihood such that zero equals the log likelihood of a regression where there is no enforcement of the corresponding μ (i.e., one where the β_{neg} and β_{pos} are equal to zero), and one equals the log likelihood of the regression with the optimal μ. The precise values of $\mu^{*H \to L}$, $\mu^{*L \to H}$, and $\mu^{*L \to L}$ are reported in Table 5.

In EQUAL the fit of the model has a clear maximum almost exactly at $\mu^{L \to L} = 0.50$, and deviations from this value monotonically worsen the model's performance. In other words, the best fit is obtained for the μ that implies that subjects enforce equal contributions by all. In URE we find values for $\mu^{*H \to L}$, $\mu^{*L \to H}$, and $\mu^{*L \to L}$ that are also very close to 0.50. That is, low players enforce that high players contribute as much as they do and vice versa, which is consistent with the enforcement of an equal contribution norm. In stark contrast, in

---

[29]There are other plausible interpretations for these two control variables. For example, being punished might teach subjects that the use of punishment is appropriate and, hence, lead to more future punishment.

[30]We estimate the model using one regression and interaction variables to separate the coefficients that vary across the three equations. We use the log likelihood of this regression to measure the fit of a set of μ's.
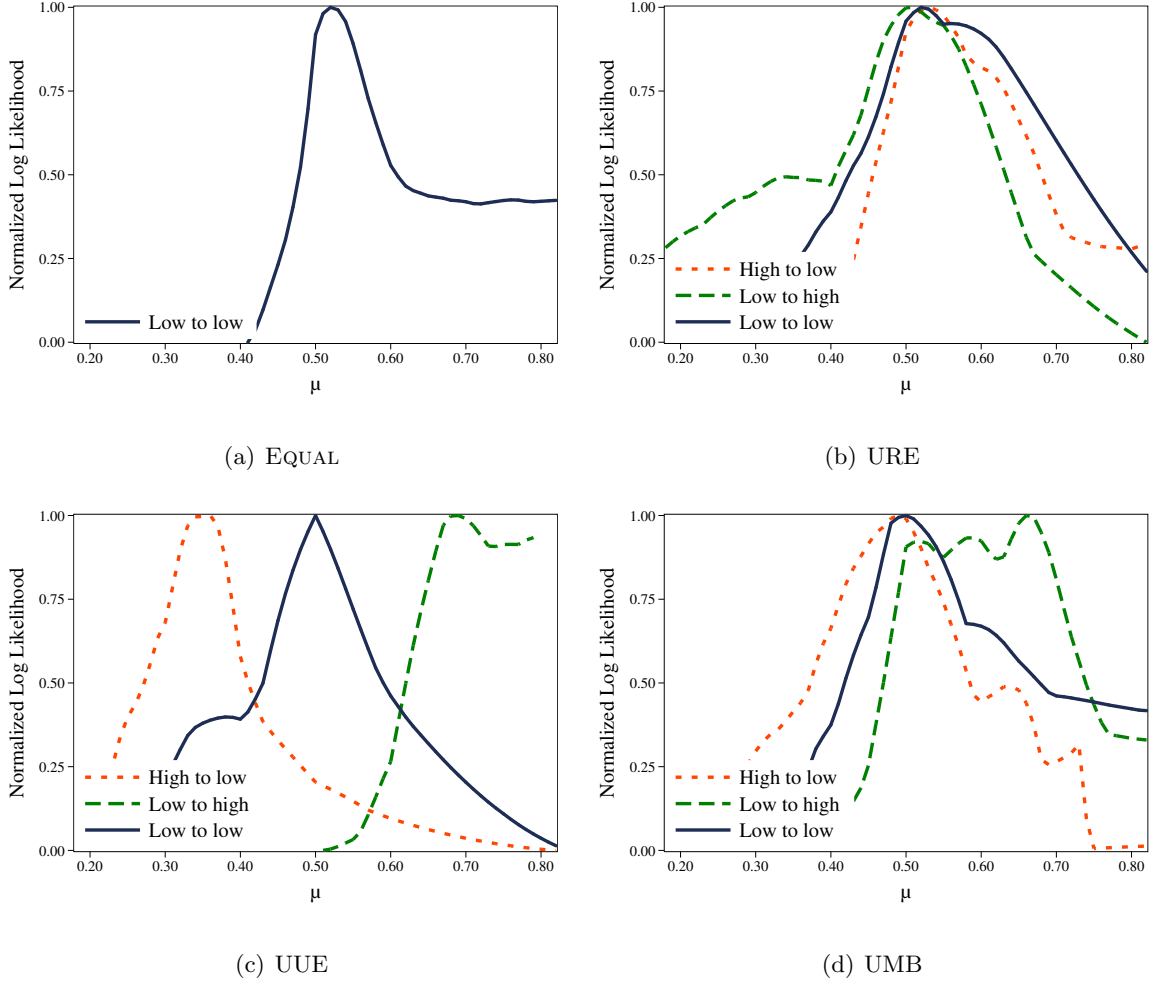
Figure 1: Goodness of fit for different values of μ

Note: Log likelihood obtained using different values of $\mu^{H \to L}$, $\mu^{L \to H}$, and $\mu^{L \to L}$. The log likelihood is normalized such that, for each μ, 0 equals the log likelihood of the worst fitting regression and 1 that of the best fitting regression.

UUE both high and low players clearly sanction so that high players contribute more than low players. Specifically, according to the values $\mu^{*H \to L} = 0.36$ and $\mu^{*L \to H} = 0.69$, low players enforce that high players contribute roughly twice as much as they do and high players enforce that low players contribute half as much as they do. This is consistent with the enforcement of a contribution norm proportional to endowments (and capacities). In UMB we observe some disagreement between player types. High players punish as if they expect low players to contribute as much as they do, whereas low players punish high players as if they expect them to contribute twice as much as they do. The former is consistent with the enforcement of an equal contribution norm whereas the latter is consistent with a norm equalizing earnings.

The fact that the values of the μ^*'s coincide with normatively attractive rules of behavior (see Section 4.1) indicates that in different group types different relative contribution norms are indeed enforced, which explains the observed differences in the contributions across group types and roles in the treatments with punishment. Additional support for this interpretation comes from the observation that in URE and UUE the optimal values $\mu^{*H\to L}$ and $\mu^{*L\to H}$ show that both high and low players enforce the same relative contribution.

As previously mentioned, we also looked for the values of $\mu^{*H\to L}$, $\mu^{*L\to H}$, and $\mu^{*L\to L}$ using the data from all groups (inefficient and efficient). We find that in both URE and UUE the results remain practically identical. The only group type for which we see a difference is UMB where the value of $\mu^{*L\to H}$ differs and is more in line with the enforcement of an equal contributions norm. Note that this is also the case where we see a flatter likelihood function (see Figure 1d), which, as we discuss in the conclusions, might be the result of different rules being enforced in different groups.[31]

In addition to establishing which contribution norm is enforced, it is interesting to see how severely subjects punish deviations from the norm. Table 5 reports the estimated coefficients of the regressions using the best fitting μ's.[32] The regressions reveal three important facts. First, the introduction of the variables capturing the sanctioning of deviations from the efficiency rule and a relative contribution rule improves the fit in all regressions (Wald tests, $p \leq 0.001$).[33] Second, the coefficients for negative deviations from the efficiency rule (λ) and the elicited relative contribution rule (the β_{neg}'s) are statistically significant in all regressions.[34] Third,

---

[31]These regressions are available in the online appendix. As a robustness check, in the online appendix we also present the results of an analysis treating punishment as a binary decision (using probit estimates). The results are very close to those obtained when using data from all groups.

[32]It is also interesting to analyze how the subjects' contribution decision depends on punishment and deviations from the estimated relative contribution rules. Regressions performing this analysis are provided in the online appendix. The results are as expected in the sense that punishment increases the contributions of subjects who deviated negatively from the contribution rule and decreases the contributions of subjects who deviated positively.

[33]With OLS regressions and the easily interpretable R^2 statistic, norm enforcement variables improve the R^2 from 0.07 to 0.28 in EQUAL, 0.12 to 0.32 in URE, 0.07 to 0.23 in UUE, and 0.10 to 0.18 in UMB.

[34]Unlike for negative deviations, the coefficients for punishment of positive deviations from the enforced relative contribution rule are rarely statistically significant. Albeit, in various cases this is due to a large standard error and not due to the size of the coefficient. Hence, it would be premature to dismiss the importance of this type of punishment even if it is more irregular than punishment of negative deviations. It should be noted, however, that punishment of positive deviations is not necessarily 'antisocial' (see, Herrmann et al. 2008). It is consistent with the interpretation of enforcement of a relative contribution norm, which may sanction deviations in any direction.

Table 5: Regressions for the best fitting contribution norms

Note: Set of μ's that best describe the punishment data and the coefficients of the corresponding Tobit regressions. In addition to the shown coefficients, all regression have group dummies, a constant, and dummies indicating the role of i and j. Standard errors are in parentheses. Statistical significance at 5 and 1 percent are indicated by * and **.

	Optimal μ's			
	EQUAL	URE	UUE	UMB
$\mu^{*H\rightarrow L}$		0.53	0.36	0.49
$\mu^{*L\rightarrow H}$		0.51	0.69	0.67
$\mu^{*L\rightarrow L}$	0.52	0.52	0.50	0.50

	Regressions			
	EQUAL	URE	UUE	UMB
$\beta_{pos}^{H\rightarrow L}$		0.42	0.19	1.02**
		(0.23)	(0.20)	(0.31)
$\beta_{neg}^{H\rightarrow L}$		0.44*	0.60*	0.67**
		(0.18)	(0.29)	(0.22)
$\beta_{pos}^{L\rightarrow H}$		0.30	0.30	0.43
		(0.21)	(0.64)	(0.36)
$\beta_{neg}^{L\rightarrow H}$		0.76**	0.46*	0.39*
		(0.21)	(0.21)	(0.18)
$\beta_{pos}^{L\rightarrow L}$	0.04	0.46	0.07	0.49
	(0.12)	(0.25)	(0.51)	(0.33)
$\beta_{neg}^{L\rightarrow L}$	0.38**	0.92**	0.98**	0.71*
	(0.14)	(0.21)	(0.31)	(0.33)
λ	0.13*	0.23**	0.25**	0.27**
	(0.06)	(0.07)	(0.06)	(0.07)
γ^r	0.30**	0.15	0.14	0.21
	(0.10)	(0.10)	(0.11)	(0.11)
γ^p	-0.10	-0.04	-0.20	-0.03
	(0.07)	(0.09)	(0.12)	(0.08)
# obs.	324	378	378	324
log likelihood	-325.97	-307.39	-325.06	-351.07

the values of the estimated coefficients are such that in all treatments, free-riding is not a profit maximizing action.[35]

---

[35]For example, the linear estimates indicate that reducing one's contribution by 1 point in EQUAL increases the punishment received from others by 0.31 points and therefore reduces one's earnings by 1.38 points. In the heterogeneous groups, the same calculation yields a reduction that varies between 1.89 points (for high players in UUE) and 3.93 points (for low players also in UUE). If one separates the effect of sanctioning deviations

We briefly summarize our findings in this section. In each treatment we find a subset of efficient groups where all subjects adhere to the efficiency norm, which for these groups is also consistent with adherence to a corresponding relative contribution norm. In the remaining groups, which fail to successfully enforce full efficiency, the observed differences in the contributions across player roles and group types can be attributed to the enforcement of different relative contribution norms. Moreover, in URE and UUE, both high and low players enforce the contribution norm and spend similar amounts of effort punishing deviations from it. Of the possible relative contribution rules (see Table 2), the one that is enforced depends on the type of heterogeneity in the group. In URE, despite the fact that high players' earnings are (almost) twice as high as low players' earnings, both roles enforce an equal contribution norm. In UUE subjects enforce a relative contribution norm where contributions are proportional to endowments. Finally, in UMB behavior in inefficient groups is consistent with low players attempting to enforce relative contributions leading to equal earnings and high players attempting to enforce equal contributions.

5 Conclusions

With social preferences, public good problems are transformed into coordination problems with a large number of equilibria, where existing theoretical models do not have much predictive power. We investigate the existence and enforcement of different contribution norms and whether they help overcome these coordination problems, in homogeneous as well as heterogeneous groups. According to scholars in social norms research (Bicchieri 2006; Young 2008), a norm is a rule of behavior that is observed if people are aware that a (normatively appealing) behavioral rule exists and sufficiently many people follow the rule, where the latter can come about through internalized beliefs or through (expected) sanctions in case of a norm violation. To investigate the existence of normatively appealing contribution rules we use data from a questionnaire study that elicits individuals' normative views of how much one ought to contribute in the different roles and types of heterogeneous groups. In order to explore the actual existence and enforcement of contribution norms we use data from a laboratory public good experiment with and without punishment opportunities.

We define two types of behavioral rules that may form the basis of a contribution norm. The *efficiency* rule prescribes to contribute as much as possible whereas the *relative contri-*

---

from the efficiency and relative contribution rules, in all cases the efficiency rule accounts for around 40% of the increase in punishment and the relative contribution rule for the remaining 60%.

bution rule is defined relative to the contributions of others. The questionnaire study reveals that uninvolved individuals have clear views of what constitutes normatively desirable behavior. For homogeneous as well as heterogeneous groups we find support for the normative appeal of both types of rules. The efficiency rule is favored for all group types, especially when full efficiency is feasible. In addition, we find that the normatively appealing relative contribution rules are those that are based on the fairness principles of equality and equity. For homogeneous groups, where equality and equity both lead to the relative contribution rule of equal contributions, we observe a consensus on the normative appeal of this rule. For heterogeneous groups, the various possible interpretations of equality and equity lead to less agreement on the specific rule of behavior that is considered normatively appealing.

In the laboratory experiment, we find that, in the absence of punishment possibilities, contributions steadily decline in all group types to the point where the prevalent behavior is (almost) full free-riding. In other words, in spite of there being strong normative support for the efficiency rule, we do not see this rule being followed. Importantly, this does not imply that efficiency and relative contribution rules do not exist in subjects' minds. In fact, the observed behavior is consistent with the existence of such rules, given that the willingness to follow them differs across people and sufficiently many have not or only weakly internalized them. The ubiquitous trend towards free-riding also dissipates any potential differences between the various forms of group heterogeneity or between individuals with different induced characteristics.

In stark contrast, when punishment is possible, contributions are high in all group types and also exhibit considerable differences in contributions across group types and between players with different roles. We show that deviations from full efficiency are similarly punished in all group types and that in a subset of groups the efficiency norm is successfully established. In addition, we provide evidence that the observed differences in contributions of players in different roles and their sanctioning behavior are consistent with the enforcement of different relative contribution norms. Specifically, in groups with unequal endowments, we find that punishment behavior is consistent with the enforcement of a relative contribution norm that prescribes contributions that are proportional to the maximum feasible contribution. This punishment behavior can be readily reconciled with the normative judgments elicited in the questionnaire. Indeed, despite the fact that there exists a plurality of normatively appealing rules of behavior, the norm that is enforced coincides with the rule that is favored by the (relative) majority of questionnaire respondents.

When marginal benefits from the public good are unequal the picture is different. In groups that do not successfully enforce the efficiency rule we find some disagreement between high and low players regarding the relative contribution rule to be enforced. This result is consistent with the finding of Nikiforakis et al. (2012) who observe that, in comparison to homogeneous groups, punished subjects in heterogeneous groups are more likely to counter-punish. A possible albeit speculative reason for the disagreement could be that some subjects interpret the high marginal benefits rather as 'deservingness' than as an obligation to contribute more to the public good. For instance, those who benefit more from cancer research may not be expected to donate more to it than those who benefit less. Further, the observed disagreement also points to a self-serving interpretation of desirable contribution rules and suggests an interesting avenue for future research on contribution norms. Namely, to explore whether different players with the same characteristics will (try to) enforce different contribution norms. Such an investigation of norm enforcement at the individual group level could build on our work, but it calls for a different design where there is considerably more punishment data at the group level.

A main message of our study is that in all group types contribution norms enforced by punishment emerge. In a subset of groups this leads to successful coordination on the efficiency norm in all group types. Further, in spite of the observed plurality of normatively appealing relative contribution rules, subjects involved in the public good game are mostly able to coordinate on following a unique relative contribution norm, provided that deviations from it can be sanctioned. The enforcement of this relative contribution norm is based on views of equality and equity and clearly related to the environment players are immersed in. In this respect, it is important to improve our understanding of how the saliency of particular contribution rules depends on features of the environment (e.g., the experimental parameters) in order to make behavior in public good games with heterogeneous groups no less predictable than the behavior of homogeneous groups.

Regarding the effect of unequal benefits from the public good, a remark has to be made that is suggested by both the larger variety of answers in our questionnaire study and the more ambiguous behavior of experimental groups in that treatment: when the features of the environment allow for too many different normatively appealing contribution rules then there might be increased conflict and inefficiencies (e.g., Nikiforakis et al. 2012) and the emerging contribution norms of specific groups may be more difficult to predict.

Two other important results are, first, that with heterogeneity there exists a plurality of normatively appealing rules of behavior that are potential candidates for emerging contri-

bution norms. In this respect, our paper contributes to the recent empirical literature on fairness ideals (e.g. Konow 2003; Konow et al. 2009). However, while the existing studies mainly examine allocation tasks (Cappelen et al. 2007) or bargaining situations (Gächter and Riedl 2006), we extend this research to the more complex decision problem of contributions to a public good. Second, despite the documented existence of normatively appealing contribution rules no such rule is observed when sanctioning is impossible. Hence, the eventual internalization of a contribution norm is too weak for its emergence, an observation also made by Schram and Charness (2011) in the context of allocation tasks.

Our study shows in a controlled environment that differences in behavior of individuals in different groups can be attributed to the enforcement of different norms, and are not necessarily due to collectively different preferences of group members. Moreover, we add the insight that in heterogeneous groups institutional variations can easily shift attention from one detail to another. This in turn anchors norm guided behavior on different salient features of the environment, resulting in considerably different outcomes, including fully efficient outcomes but also added conflict. In consequence, optimal regulation and organizational design, that traditionally focuses on the effects of direct material incentives, should not underestimate the importance of institutional details for collective norm guided behavior, especially when people differ.

References

Andreoni, J. and Bernheim, B. D. (2009). Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, 77(5):1607–1636.

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

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

Berkes, F., Feeny, D., McCay, B. J., and Acheson, J. M. (1989). The benefits of the commons. *Nature*, 340:91–93.

Bernard, M., Reuben, E., and Riedl, A. (2012). Coordination and the focality of (multiple) fairness ideals. Working paper, Stockholm School of Economics.

Bicchieri, C. (2006). *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge University Press, New York.

Bicchieri, C. (2008). The fragility of fairness: An experimental investigation on the conditional status of pro-social norms. In Sosa, E. and Villanueva, E., editors, *Interdisciplinary Core Philosophy*, volume 18 of *Philosphical Issues*, pages 229–248. Wiley-Balckwell, Boston MA & Oxford UK.

Bolton, G. E. and Ockenfels, A. (2000). A theory of equity, reciprocity, and competition. *American*

Economic Review, 90:166–193.

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

Cappelen, A. W., Hole, A. D., and Sørensen, E. Ø. (2007). The pluralism of fairness ideals: An experimental approach. *American Economic Review*, 97(3):818–827.

Carpenter, J. P. and Matthews, P. H. (2009). What norms trigger punishment? *Experimental Economics*, 12(3):272–288.

Carpenter, J. P. and Matthews, P. H. (2012). Norm enforcement: Anger, indignation or reciprocity. *Journal of the European Economic Association*, 10:555–572.

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

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

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

Cherry, T. L., Kroll, S., and Shogren, J. (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*. Harvard University Press, Cambridge.

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

Cox, J. C., Friedman, D., and Gjerstad, S. (2007). A tractable model of reciprocity and fairness. *Games and Economic Behavior*, 59(1):17–45.

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

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

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

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

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

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

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

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

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly*

Journal of Economics, 114:817–868.

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

Fischbacher, U. and Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods. *American Economic Review*, 100(1):541–556.

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

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

Fligner, M. A. and Policello, G. E. (1981). Robust rank procedures for the behrens-fisher problem. *Journal of the American Statistical Association*, 76:162–168.

Frohlich, N. and Oppenheimer, J. A. (1990). Choosing justice in experimental democracies with production. *American Political Science Review*, 84:461–477.

Frohlich, N., Oppenheimer, J. A., and Eavey, C. L. (1987). Choices of principles of distributive justice in experimental groups. *American Journal of Political Science*, 31:606–637.

Gächter, S. and Herrmann, B. (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. and Riedl, A. (2006). Dividing justly in bargaining problems with claims: Normative judgments and actual negotiations. *Social Choice and Welfare*, 27(3):571–594.

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

Gardner, R., Herr, A., Ostrom, E., and Walker, J. A. (2000). The power and limitations of proportional cutbacks in common-pool resources. *Journal of Development Economics*, 62:515–533.

Hamilton, B. H., Nickerson, J. A., and Owan, H. (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.

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

Hopfensitz, A. and Reuben, E. (2009). The importance of emotions for the effectiveness of social punishment. *Economic Journal*, 119:1534–1559.

Keser, C. and van Winden, F. (2000). Conditional cooperation and voluntary contributions to public goods. *Scandinavian Journal of Economics*, 102:23–39.

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

Krupka, E. L. and Weber, R. A. (2012). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, forthcoming.

Levine, D. K. (1998). Modeling altruism and spitefulness in experiments. *Review of Economic Dy-*

namics, 1:593–622.

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

Musgrave, R. A. (2008). Public finance. In Durlauf, S. N. and Blume, L. E., editors, *The New Palgrave Dictionary of Economics.* Palgrave Macmillan, Basingstoke.

Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, 7:308–313.

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

Nikiforakis, N., Normann, H.-T., and Wallace, B. (2010). Asymmetric enforcement of cooperation in a social dilemma. *Southern Economic Journal*, 76:638–659.

Nikiforakis, N., Noussair, C. N., and Wilkening, T. (2012). Normative conflict and feuds: The limits of self-enforcement. *Journal of Public Economics*, 96:797–807.

Noussair, C. N. and Tan, F. (2011). Voting on punishment systems within a heterogeneous group. *Journal of Public Economic Theory*, 13:661–693.

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

Ostrom, E. (2000). Collective action and the evolution of social norms. *Journal of Economic Perspectives*, 14(3):137–158.

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

Page, E. B. (1963). Ordered hypotheses for multiple treatments: A significance test for linear ranks. *Journal of the American Statistical Association*, 58:216–230.

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

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

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

Roth, A. E. and Murnighan, K. J. (1982). The role of information in bargaining: An experimental study. *Econometrica*, 50:1123–1142.

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

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

Schram, A. and Charness, G. (2011). Social and moral norms in the laboratory. Working paper, University of Amsterdam.

Sudgen, R. (1986). *The Economics of Rights, Cooperation and Welfare.* Basil Blackwell, Oxford.

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

van Dijk, F., Sonnemans, J., and van Winden, F. (2002). Social ties in a public good experiment. *Journal of Public Economics*, 85:275–299.

Visser, M. and Burns, J. (2006). Bridging the great divide in south africa: Inequality and punishment in the provision of public goods. Working paper, Göteborg University.

Young, H. P. (2008). Social norms. In Durlauf, S. N. and Blume, L. E., editors, *The New Palgrave Dictionary of Economics*, volume 7, pages 647–651. Palgrave Macmillan, Basingstoke.