



**HAL**  
open science

# Normative conflict and the limits of self-governance in heterogeneous populations

Lata Gangadharan, Nikos Nikiforakis, Marie Claire Villeval

► **To cite this version:**

Lata Gangadharan, Nikos Nikiforakis, Marie Claire Villeval. Normative conflict and the limits of self-governance in heterogeneous populations. 2017. halshs-01208230v2

**HAL Id: halshs-01208230**

**<https://shs.hal.science/halshs-01208230v2>**

Preprint submitted on 24 Jul 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

WP 1525 – September 2015, revised July 2017

## **Normative conflict and the limits of self-governance in heterogeneous populations**

Lata Gangadharan, Nikos Nikiforakis, Marie Claire Villeval

### **Abstract:**

Mechanisms to overcome social dilemmas provide incentives to maximize efficiency. However, often – such as when agents are heterogeneous – there is a trade-off between efficiency and equality – a normative conflict – which is overlooked. Agents' concerns for equality in such instances can limit the ability of mechanisms to promote efficiency. We provide evidence for this from a public good experiment using a simple mechanism allowing individuals to communicate periodically with other group members and reward them for their actions. We show that, in homogeneous populations – where there is no conflict between efficiency and equality – the mechanism permits groups to obtain maximum efficiency. This is not the case in heterogeneous populations despite the fact that individuals could use rewards to resolve the normative conflict. Although almost all heterogeneous groups agree to follow specific contribution rules with positive contributions, most of them either prioritize equality over efficiency or strike a compromise between the two. These findings suggest normative conflict can be difficult to overcome, imposing limits on the ability of heterogeneous populations to reach efficient outcomes through self-governance.

### **Keywords:**

Communication, rewards, cooperation, normative conflict, heterogeneity

### **JEL codes:**

C92, H41, D74

# Normative conflict and the limits of self-governance in heterogeneous populations\*

Lata Gangadharan,<sup>a</sup> Nikos Nikiforakis<sup>b</sup> and Marie Claire Villeval<sup>c</sup>

This version: June 11, 2017

## Abstract

Mechanisms to overcome social dilemmas provide incentives to maximize efficiency. However, often – such as when agents are heterogeneous – there is a trade-off between efficiency and equality – a *normative conflict* – which is overlooked. Agents’ concerns for equality in such instances can limit the ability of mechanisms to promote efficiency. We provide evidence for this from a public good experiment using a simple mechanism allowing individuals to communicate periodically with other group members and reward them for their actions. We show that, in homogeneous populations – where there is no conflict between efficiency and equality – the mechanism permits groups to obtain maximum efficiency. This is not the case in heterogeneous populations despite the fact that individuals could use rewards to resolve the normative conflict. Although almost all heterogeneous groups agree to follow specific contribution rules with positive contributions, most of them either prioritize equality over efficiency or strike a compromise between the two. These findings suggest normative conflict can be difficult to overcome, imposing limits on the ability of heterogeneous populations to reach efficient outcomes through self-governance.

*Keywords:* Communication, rewards, cooperation, normative conflict, heterogeneity

*JEL Codes:* C92, H41, D74

<sup>a</sup> Department of Economics, Monash University, Clayton, Australia. [lata.gangadharan@monash.edu](mailto:lata.gangadharan@monash.edu)

<sup>b</sup> Division of Social Science, New York University Abu Dhabi, P.O. Box 129188, Abu Dhabi, UAE. [nn30@nyu.edu](mailto:nn30@nyu.edu)

<sup>c</sup> Univ. Lyon, CNRS, GATE, 93 Chemin des Mouilles, F-69130, Ecully, France; IZA, Bonn, Germany; Department of Public Finance, University of Innsbruck. [villeval@gate.cnrs.fr](mailto:villeval@gate.cnrs.fr)

\* The paper was previously circulated under the title “Equality concerns and the limits of self-governance in heterogeneous populations”. We are grateful to R. Cautain and Q. Thévenet for programming the experiment, and to Z. Dai, J. Benistant and R. Suchon for valuable research assistance. We also thank A. Cappelen, D. Cooper, E. Fatas, H. Orzen, B. Tungodden, R. Zultan and participants at the “Social Norms and Institutions” conference in Ascona, the workshop on “Cooperation, Cultural Aspects and Norms” at the Hebrew University of Jerusalem, the Workshop on “Behavioral Public Economics” in Mannheim, the Florida State University Entrepreneurship Experimental Workshop in Tampa, the North-American ESA Conference in Dallas, the Workshop on Social Norms and Moral Norms in Lyon, the ASFEE conferences in Paris, the AFSE conference in Rennes, and in seminars at the ChoiceLab in Bergen, Beijing Normal University, University of East Anglia, New York University Abu Dhabi, University of Nice, Queensland University of Technology, Shandong University, the University of Fribourg, the University of Lausanne, Texas A&M, Hong Kong University of Science and Technology, Hitotsubashi University, Kochi University of Technology, Monash University and Tsinghua University for useful comments. This research has been supported by a grant of the French National Research Agency (ANR, EMCO program, HEIDI grant, ANR-11-EMCO-011-01) and funding from the Australian Cooperative Research Centre for Water Sensitive Cities (CRC grant number 20110044). It was performed within the framework of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the program Investissements d’Avenir (ANR-11-IDEX-007) operated by the French National Research Agency.

## 1. Introduction

Social dilemmas are a topic of continuing interest among economists due to the tension between private and social interest. If this tension is not resolved, outcomes are likely to be inefficient (*e.g.*, Samuelson, 1954). Accordingly, a large array of mechanisms has been proposed to provide incentives for individuals to act in accordance with the social interest (*e.g.*, Laffont, 1987; Myerson, 2008). By ‘social interest’ most economists mean efficiency; *i.e.*, efficiency is the ultimate normative criterion for evaluating the performance of a given mechanism. In many instances in daily life, however, efficiency is at odds with another normatively appealing rule: equality (*e.g.*, Okun, 1975; Browning and Johnson, 1984). If individuals care strongly about equality as some evidence suggests (*e.g.*, Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Charness and Rabin, 2002; Cappelen *et al.*, 2007; Almas *et al.*, 2011; Cooper and Kagel, 2016), the ensuing *normative conflict* may undermine the ability of standard mechanisms to promote efficiency.<sup>1</sup>

Despite decades of research on social dilemmas, surprisingly little is known about how severe the tension between efficiency and equality is. This however is an important empirical question as the tension between efficiency and equality arises naturally in many instances such as when individuals derive different benefits from the provision of public goods (*e.g.*, Chen and Plott, 1996; Nikiforakis *et al.*, 2012; Reuben and Riedl, 2013). When some agents enjoy greater benefits from the public good than others, contributions to the public good can increase both efficiency and inequality. This raises the question of how difficult it is to overcome this tension in practice. To what extent can equality concerns undermine the ability of mechanisms to promote efficiency in heterogeneous groups facing social dilemmas?

We address these questions in a laboratory experiment using a variant of the Voluntary Contribution Mechanism (VCM hereafter) for public goods provision as a testing ground. All group members are given the same endowment and must decide how much of it to contribute to a public good. In half of our treatments, we assign higher returns from the public good to some group members than their peers, based on their relative performance in a previous real-effort task. The heterogeneous returns create a stark trade-off between efficiency and equality: when efficiency is maximized, so is the inequality in the earnings of the different players’ types. To evaluate the severity of the normative conflict, we compare behavior to that in treatments in which all group

---

<sup>1</sup> A normative conflict is said to arise when multiple normatively appealing rules about how one ought to behave in a given situation coexist and prescribe different actions. For this to happen, these rules must have some properties that appeal to a large fraction of individuals. Considerable evidence suggests that equality and efficiency are such rules. For more general discussions on the topic, see Nikiforakis *et al.* (2012), and Reuben and Riedl (2013).

members receive the same benefit from the public good. Homogeneity implies that maximizing efficiency simultaneously leads to equality in earnings.

The mechanisms designed to support cooperation in the literature differ greatly. On one end of the spectrum are centralized mechanisms with clearly defined incentives for performance (*e.g.*, Groves and Ledyard, 1977), which are often complex and difficult to implement in practice (*e.g.*, Laffont, 1987; Jackson and Moulin, 1992).<sup>2</sup> On the other end are simple, easy to implement in practice, decentralized mechanisms, without well-defined incentives. These typically rely on peer-monitoring, peer-communication and peer-provided material incentives such as rewards (Rockenbach and Milinski, 2006; Sefton *et al.*, 2007; Rand *et al.*, 2009; Sutter *et al.*, 2010) or punishment (*e.g.*, Fehr and Gächter, 2000; Masclet *et al.* 2003; Nikiforakis, 2008). We chose a simple mechanism for our purposes.

Our mechanism allows for communication between group members and rewards at regular intervals. Rewards in our study take the form of monetary transfers from one group member to another.<sup>3</sup> In the case of heterogeneous populations, rewards are of prime interest as individuals deriving a larger benefit from the public good can agree to transfer funds to those deriving a smaller benefit to alleviate (or even fully resolve) the normative conflict *ex post*.<sup>4</sup> To facilitate such agreements, in some treatments, group members are invited periodically to a forum and are given the opportunity to communicate. Communication is anonymous, free form, and non-binding, but it can nonetheless assist groups in reaching an agreement both about how much each member ought

---

<sup>2</sup> Jackson and Moulin (1992, p.126) write: “The general results are technically impressive, but generally impractical for producing plausible mechanisms”. Laffont (1987, p. 567) similarly writes: “... any real application will be made with methods which are crude approximations to the mechanisms obtained here ... considerations such as simplicity and stability to encourage trust, goodwill and cooperation, will have to be taken into account.” Recently less complex, centralized mechanisms have been developed, *e.g.*, Andreoni and Gee (2012), Falkinger (1996) and Varian (1994).

<sup>3</sup> According to the Merriam-Webster dictionary, “reward” can be defined as the act of giving “money or another kind of payment to someone for something good that has been done” (Merriam-Webster). This definition implies that monetary transfers in our experiment can be thought of as “rewards” so long as they are *conditional* on another person’s behavior (*e.g.*, high contribution, honoring a group’s agreement). The vast majority of transfers in our experiment appear to fit this definition. We therefore follow the convention in the literature of referring to transfers between group members as rewards. What are some of the analogues of such rewards in daily life? Rand *et al.* (2009) provide the following examples: helping a friend move furniture, working extra shifts to cover an ill co-worker or being extra helpful to a neighbor who has just bought a low emission car. Of course, rewards can also take the form of actual monetary transfers, *e.g.*, compensating an individual for his/her cost in generating a positive externality.

<sup>4</sup> Formally, the opportunity to reward implies that an agreement in which all members contribute fully is not only Pareto efficient, but also Kaldor-Hicks efficient. An outcome is said to be Kaldor-Hicks efficient if those that are made better off could potentially compensate those worse off. The compensation does not have to occur.

to contribute to the public good and who should receive a monetary transfer by whom.<sup>5</sup> To that end, we explore behavior in control treatments, with reward but no communication opportunities.

Although communication and rewards can help resolve the normative conflict *ex post*, it is far from obvious this will happen in practice. First, rewarding is costly. Individuals may therefore be unwilling to personally compensate their low-return counterparts (*i.e.*, rewarding is a second order public good), even more so in an environment such as ours where the privilege of obtaining higher benefits from the public good has been earned (*e.g.*, Gee *et al.*, in press). Second, monetary transfers can be used for purposes other than eradicating inequality. For instance, one may wish to reward high contributors irrespective of their type, to engage in parochial behavior by rewarding participants of the same type, to establish a ‘reward-exchange’ relationship with others, or to increase others’ earnings without any regard for individual contributions and earnings equality. If heterogeneous groups cooperate less than homogeneous groups, it will be evidence that the tension between equality and efficiency is not simple to overcome in practice.

The fact that normative conflict can pose a significant challenge to cooperation in heterogeneous groups is perhaps best reflected in the talks for an international agreement to reduce greenhouse gas emissions. After years of deliberations, several countries declined to ratify the Kyoto Protocol. The agreement itself failed to curb the growth of global emissions (World Bank, 2010) that continue to increase at an accelerating pace (Guardian, 2014). The 2009 conference in Copenhagen also failed to find a replacement agreement for Kyoto (Guardian, 2013), and many rounds of pre-negotiation have had to be organized to be able to reach an agreement at the last minute at the Conference for climate change in Paris in 2015. Different accounts for the failure of the climate-change talks suggest that a contributing factor was normative conflict: developing countries argued in favor of *relative* reductions in emissions, emphasizing the smaller size of their economies and their historically lower rate of emissions. Rich, developed countries, on the other hand, reasoned that some of world’s largest emitters are developing countries and argued in favor of *absolute* emission reductions (World Bank, 2010, p. 235).<sup>6</sup>

---

<sup>5</sup> As we discuss in the next section, several studies have shown that non-binding communication can promote cooperation. However, with few exceptions (Isaac and Walker, 1988; Hackett *et al.*, 1994; Chan *et al.*, 1999; Tavoni *et al.*, 2011; Dekel *et al.*, 2017), all preceding studies allowing non-binding communication in social dilemmas used homogeneous groups.

<sup>6</sup> Another example in the domain of environment can be found in Wiggins and Libecap (1985) about the difficult unitization negotiations in oil fields in the U.S. between firms with large acreage and firms with small ones. Another, more recent example in a different domain involves the talks to overcome the debt crisis in Greece affecting the entire Eurozone. Throughout the five-year negotiation period, repeated appeals to fairness from both sides of the negotiation and to the need for structural reforms improving efficiency have led to tensions and costly delays.

The results from our experiment demonstrate the difficulty of overcoming normative conflict in heterogeneous populations and the limits posed by it on self-governance. While communication has a positive effect on cooperation rates and group earnings in both homogeneous and heterogeneous populations, in line with our hypothesis, the *size* of the effect is substantially smaller in the latter case where there is normative conflict. Only 23.50% of heterogeneous groups reach the fully efficient outcome when our mechanism is available, relative to 100% of the homogeneous groups. As a result, we find that the increase to cooperation rates in heterogeneous groups caused by enabling communication is only 38.17% of what it is in homogeneous groups. Similarly, the increase in group earnings from enabling communication in heterogeneous groups is only 57.22% of that in homogeneous groups.

The lower efficacy of the mechanism in heterogeneous populations is not the result of an inability on the part of group members to agree on a given strategy: all but one of the groups (16 out of 17) reach stable agreements about which contribution rule to follow. The reason for the relative inefficacy of the mechanism is the revealed preference for equality. Most of the groups either prioritize equality over efficiency or strike a compromise between the two. Interestingly, rewards are assigned mostly equally to members, irrespective of the contribution rule adopted or the returns obtained from the public good.

The paper proceeds as follows. Section 2 introduces our experiment. This section starts with a discussion of earlier studies informing our design, and then presents details about the experiment, and the experimental procedures. In Section 3 we present our empirical findings. Section 4 concludes.

## **2. The experiment**

### *2.1. Related literature*

Our experimental design builds on earlier studies that have used laboratory experiments to investigate the impact of rewards, communication, and group heterogeneity on cooperation in social dilemmas. Although there are numerous studies on these topics, as we will see below, few consider the impact of more than one of these elements simultaneously; we are aware of only one other study (Dekel *et al.* 2017) that considers all three, albeit in a different kind of social dilemma. The efficacy of the vast majority of centralized (e.g., Yamagishi, 1986; Falkinger *et al.*, 2000; Andreoni and Gee, 2012) and decentralized mechanisms for promoting cooperation (e.g., Egas and Riedl, 2008; Fehr and Gächter, 2000) has been evaluated in the context of homogeneous groups.

As a consequence, little is known about the severity of the normative conflict in heterogeneous populations. In this section, we offer a brief overview of this literature.

The impact of rewards in promoting cooperation in *homogeneous* groups has been the topic of several studies. Reward takes the form of monetary transfers between participants. The monetary transfer can either leave the total surplus unaffected (*i.e.*, a 1:1 technology), or increase it (*e.g.*, a 1:3 technology, in which the sender's earnings are reduced by 1 per reward unit and that of the recipient increase by 3 per reward unit), in which case they can be thought of as modelling instances where individuals derive different marginal utility from the same unit of a good. Two recent studies considering the impact of rewards on cooperation are those by Rand *et al.* (2009) and Sutter *et al.* (2010). Rand *et al.* (2009), use a 1:3 technology and find that in homogenous groups rewards are as effective as punishment in maintaining cooperation and lead to higher total earnings. Sutter *et al.* (2010) show that a 1:3 technology is more effective in promoting cooperation than a 1:1 technology. In the former case, rewards are found to be more effective than punishment at promoting cooperation. Sutter *et al.* (2010) also show that subjects tend to prefer using rewards than punishment. With the exception of Dekel *et al.* (2017), which we discuss below, we are not aware of any study which allows rewarding in heterogeneous groups. As mentioned earlier, in this case, rewards are of prime interest as they can alleviate the normative conflict by reducing inequalities *ex post*. Communication can facilitate such monetary transfers.

Several studies have shown that, when individuals derive the *same* benefits from cooperation, communication has a positive effect on cooperation and often leads to maximal contributions (*e.g.*, Ostrom *et al.*, 1992; Cason and Kahn, 1999; Brosig *et al.*, 2003; Bochet *et al.* 2006; Bochet and Putterman, 2009; Janssen *et al.*, 2010; Cason and Gangadharan, 2014; Oprea *et al.*, 2014).<sup>7</sup> Overall, the size of the effect – the largest of any other single factor in social dilemmas (such as marginal per capita return, peer punishment and fixed group matching) according to the meta-analyses of Sally (1995) on prisoners' dilemma experiments and of Zelmer (2003) on public good games – is such that it has led some scholars to state that “self-governance is possible” (*e.g.*, Ostrom *et al.*, 1992; Janssen *et al.*, 2010; Ostrom, 2012). However, as mentioned, the vast majority of these studies allow for communication between homogeneous agents.

---

<sup>7</sup> The form and timing of communication differs across studies. Early studies allowed for face-to-face communication, which took place once at the start of the experiment. More recently – with the advent of computer software – chatroom communication at different points in the experiment has become more common. An advantage of chatroom communication is that it can maintain subject anonymity. Bochet *et al.* (2006) show that periodic chatroom communication is as effective as face-to-face communication in promoting cooperation.



Heterogeneity has been introduced in the VCM in a variety of ways. In some studies individuals are given different endowments from which to contribute (*e.g.*, Buckley and Croson, 2006), while in others they are given different returns from the public account (*e.g.*, Fisher *et al.*, 1995; Reuben and Riedl, 2009). In some studies, the privilege to be a high ‘income’ participant is randomly assigned to participants (*e.g.*, Reuben and Riedl, 2013); in others, it is earned in a separate task (*e.g.*, Nikiforakis *et al.*, 2012; Weng and Carlsson, 2015). Overall, it is safe to say that the impact of heterogeneity remains unclear. Some studies find that heterogeneity has a negative effect on cooperation (Cherry *et al.*, 2005; Anderson *et al.*, 2008; Fischbacher *et al.*, 2014), and some find it has a positive effect (Chan *et al.*, 1996; Chan *et al.*, 1999; Buckley and Croson, 2006; Reuben and Riedl, 2013).<sup>8,9</sup>

Only a handful of studies have allowed communication between heterogeneous group members in social dilemmas. In Isaac and Walker (1988), subjects who receive different endowments (but the same benefits from the public good) can communicate in a VCM environment. The authors find an 18-percent reduction in cooperation rates relative to a homogeneous treatment. In contrast, Chan *et al.* (1999) show that communication helps heterogeneous groups more than homogeneous groups as long as subjects have complete information on others’ endowments or preferences. Two further studies allow for communication in heterogeneous groups, but do not compare behavior to that in homogeneous groups. In a Common Pool Resource game, Hackett *et al.* (1994) show that allowing heterogeneous agents to communicate increases cooperation compared to a no-communication environment (see also Tavoni *et al.*, 2011, who use a threshold public bad game). The second study by Dekel *et al.* (2017) is the one most closely related to ours.

In a recent study, Dekel *et al.* (2017) examine situations in which cooperation confers a benefit to the group on aggregate, but harms a small minority of group members (*i.e.*, one of the three group members). The authors explore the efficacy of punishment and reward (separately), with and without communication. They find that rewards are not used by majority players to compensate those harmed when communication is unavailable. Communication fosters compensation, but interestingly, it is only the majority players and not the harmed minority who increase their

---

<sup>8</sup> Fischbacher *et al.* (2014), in particular, investigate how heterogeneous benefits from a public good impact conditional as well as unconditional contributions. They find that unconditional contributions are negatively affected by the introduction of heterogeneity, but conditional contributions are not significantly affected, suggesting that heterogeneity impacts mainly subjects’ beliefs which become more pessimistic about the prospects of cooperation.

<sup>9</sup> Noussair and Tan (2011) find that when group members differ in the value of their contributions to other group members, groups fail to establish efficient punishment institutions. In a similar set up, Tan (2008) shows that punishment does not increase group earnings relative to a baseline treatment without punishment opportunities. However, neither of these two studies includes a treatment with homogeneous groups.

contributions.<sup>10</sup> Apart from the fact that Dekel *et al.* (2017) do not compare behavior in heterogeneous and homogeneous groups, the main difference from our study is clearly that minority players receive a negative return from the level of the ‘public good’. This changes fundamentally the incentives of the different types of players, their interaction, and the nature of normative conflict as different normatively appealing rules may be at play. For example, a rule of “do no harm unto others” could increase the willingness of subjects to compensate minority players.

In summary, it remains unclear from the literature whether the tension between equality and efficiency triggered by heterogeneous benefits from the public account poses a challenge when group members can communicate and reward each other.

## 2.2. *The basic game*

In all treatments, subjects participate in a six-player linear public good game. Group assignment is fixed for the 20 periods of the game. Each period is divided into two stages. At the start of the first stage, each subject is given an endowment in experimental currency units (ECU 20, with ECU 10 = Euro 0.25). Subjects have to allocate this endowment between a “private account” and a “public account”. The earnings of group member  $i$  at the end of the first stage are given by:

$$\pi_i^1 = \gamma(E - g_i) + \beta_i \sum_{j=1}^n g_j$$

The parameters  $E$ ,  $\gamma$ , and  $\beta_i$  denote, respectively, the endowment that subject  $i$  has to allocate in the two accounts, the return to the endowment allocated to the private account (equal to 1 in our game for simplicity), and the return to total allocations by the group towards the public account that, as explained below, may differ across subjects. The subject’s decision variable in this stage is  $g_i$ , the amount of the endowment to contribute to the public good. This payoff function represents a social dilemma if and only if  $\beta_i < \gamma < n\beta_i$ , where  $n$  is the number of members in the group. The first inequality shows that individual returns from the public good are lower than private returns, hence zero public good contribution is the dominant strategy Nash equilibrium. The second inequality implies that total returns received by the group from any public good allocation exceed the private returns; hence the social optimum is for every subject to contribute all resources to the public good. In our experiment,  $\gamma = 1$  and  $\beta_i$  varies depending on the treatment.

---

<sup>10</sup> A reason for this could be that the rewards in Dekel *et al.* (2017) do not suffice to fully compensate the minority player if she chooses to contribute fully, i.e., they cannot fully resolve the normative conflict ex post.

At the start of the second stage subjects are informed about how much each group member has contributed to the public good, and  $\beta_j$  (the returns from the public account) for all  $j$ . Subjects in this stage decide whether or not to reward other members in their group and increase their earnings by distributing “addition points”.<sup>11</sup> Each point that individual  $i$  assigns to group member  $j$ ,  $r_{ij}$ , increases  $j$ 's earnings by 1.5 ECU and reduces those of the sender by 0.50 ECU. The reason for using a 1:3 ratio vis-à-vis a 1:1 ratio is that it makes the eradication of inequality ex post easier (cheaper). The earnings of individual  $i$  at the end of stage 2 are given by:

$$\pi_i^2 = \pi_i^1 - 0.5 \sum_{\substack{j=1 \\ j \neq i}}^n r_{ij} + 1.5 \sum_{\substack{j=1 \\ j \neq i}}^n r_{ji}$$

By assigning reward points, subjects can generate extra income from the experiment. To prevent subjects from “abusing” the reward medium, we impose the following restrictions. Subject  $i$  can assign a maximum of 20 *points* in total to all other group members, *i.e.*,  $\sum_j r_{ij} \leq 20$ . Under this limit, there is no constraint on the number of points  $i$  can assign to a particular member. The advantage of this restriction is that it allows us to easily calculate the maximum possible earnings for participants. Another reason for putting such a cap was to limit differences in the rewarding power across subjects (these differences could have been substantial if for example, we had bounded the number of rewarding points to the amount earned in the first stage of the game).

The second restriction is that no subject can benefit from receiving more than 20 ECU in rewards – the equivalent of their initial endowment – *i.e.*,  $1.5 * \sum_j r_{ji} \leq 20 \leftrightarrow \sum_j r_{ji} \leq 13.33$  points. The value of points exceeding 20 ECU is not counted towards the final earnings, but the cost is incurred by the sender. This requires some coordination between group members to avoid wasting resources in assigning addition points that cannot be converted into money. Allowing the possibility to communicate should help subjects in this respect.

At the end of the period, each subject receives feedback on the individual number of points he received from each of his group members, together with each member's contribution, their identification number and their return rate from the public good ( $\beta_i$ ). The feedback also indicates the total cost of the points the subject assigned to others, the total benefit from the points that he received, and his net earnings in that period.

---

<sup>11</sup> To allow rewarding, at the start of each period group members are given a unique number between 1 and 6. To limit the extent of strategic reward-exchange over periods, the identifying numbers were randomly reassigned in each period.

### 2.3. *Experimental treatments*

The experiment consists of four treatments using a 2x2 design. The treatments differ in two dimensions: the homogeneity of group members with respect to their individual returns from the public account, and the ability to communicate. We discuss each next.

*Heterogeneity of returns from the public account:* In two of the treatments, referred to as “homogenous treatments”, subjects benefit equally from the public good ( $\beta_i = 0.3$ , for all  $i$ ). In the two other treatments, referred to as “heterogeneous treatments”, half of the subjects in a group receive a higher return rate from the public account ( $b_i = 0.4$ ), and the other half receives a lower return rate ( $b_i = 0.2$ ). Thus, the average return rate from the public account is the same as in the homogenous treatments.<sup>12</sup>

*Communication:* Our second treatment variable is communication. In the HOM-NC and the HET-NC treatments (HOM for Homogenous, HET for Heterogeneous, and NC for No Communication), subjects are not able to communicate with others in their group. In contrast, in the HET-C and HOM-C treatments, subjects have the opportunity to communicate with the other five participants in their group in chat rooms on their computers. Our operationalization of communication is inspired by Bochet *et al.* (2006). Communication is permitted four times during the experiment: at the start of periods 1, 6, 11 and 16. The chat rooms are open for 5 minutes at the start of periods 1 and 6, and for 3 minutes at the start of periods 11 and 16.<sup>13</sup> Messages are freeform. To increase the likelihood that communication helps groups overcome normative conflict and maximize efficiency, messages are visible to all group members.<sup>14</sup> All participants are requested to follow some simple

---

<sup>12</sup> Similar to Nikiforakis *et al.* (2012), the assignment of the return rates depends on the relative performance of the subjects in a previous encryption task (like in Erkal *et al.*, 2011, and Charness *et al.*, 2014). For ten minutes, subjects are given a table assigning a number to each letter of the alphabet and are asked to use it to encrypt a number of words. To minimize selection effects, when performing this task, subjects are not provided with details about the public good game. They are only informed that: (i) their relative performance will influence the maximum amount they can earn in the second part, and (ii) that they will be placed in groups of six, three of which will be “type A” and three of “type B”, these two types receiving different returns from their decisions. The top 50% performers in the session are assigned to the type that obtains a higher return from the public good ( $\beta = 0.4$ ) and the rest are assigned to the type that receives a lower return ( $\beta = 0.2$ ). At the end of the task, subjects are informed about their type, but not about the number of words other players encrypted. This information is provided at the end of the last period of the game. The mean performance of the subjects who received a high-return is 44.74 units (S.D.=5.02); for those who received a low-return, it is 34.86 units (S.D.= 4.98).

<sup>13</sup> We varied the duration of communication as we expected group members to need less time to reach an agreement in the second half of the experiment, and also wanted to avoid participants getting bored waiting for the end of the communication time.

<sup>14</sup> In the HET-C treatment, we explored two variants of communication. In one variant, subjects could only send messages to everyone. In the second variant, subjects could also send ‘private messages’ to the group members with the same return as them. This was done to examine the robustness of our conclusions with regards to different types

rules for this communication—they cannot identify themselves, they need to be civil to each other and cannot use any rude words. Apart from this, however, they can communicate about anything (messages were monitored randomly during the session and checked systematically *ex post*; all subjects respected the rules).

#### 2.4. The trade-off between equality and efficiency

In the homogeneous treatments, on aggregate, efficiency and equality are not at odds: when all group members contribute fully, efficiency is maximized and all earn the same amount. In the heterogeneous treatments, however, there is a tension between equality and efficiency: as contributions increase on average, so do group earnings and inequality. Keeping the option of rewarding others aside for a moment, for earnings to be equalized, high-return individuals have to contribute substantially more than their low-return peers. For our experimental parameters, high-return players must contribute *four times as much* as low-return players to equalize earnings prior to the reward stage. Efficiency, on the other hand, would prescribe that all group members contribute equally their entire endowment. However, maximizing efficiency in the absence of rewards implies maximizing inequality in the first stage. In our setup, a simultaneous 1 ECU increase in each group member’s contribution increases the net earnings of high-return players by 1.4 ECU and that of low-return players by 0.2 ECU.<sup>15</sup> Thus, the overall difference increases by 1.2 ECU with each additional ECU contributed by each group member to the public account. This implies that earning inequality in the first stage is maximized when group earnings are maximized, in which case  $\pi_H - \pi_L = 20 \cdot 1.2 = 24$  ECU, with  $\pi_L = 24$  ECU and  $\pi_H = 48$  ECU, where  $\pi_L$  denotes the earnings of each low-return agent, and  $\pi_H$  that of every high-return agent.

Of course, rewards could be used to eradicate any inequality in earnings at the end of the first stage. Recall from section 2.2 that we imposed restrictions on the number of reward points that can be assigned to prevent the exploitation of the reward medium by participants. A benefit of such restrictions is that it allows us to calculate the maximum earnings when all contribute fully to the public account and compare them to the *actual* earnings from the experiment, as a measure of efficiency. In homogeneous treatments, group earnings are maximized when  $g_i = g_{-i} = 20, r_{ij} =$

---

of communication. The figures (A1 for contributions, A2 for earnings, and A3 for reward points received) presented in Appendix 2 reveal that behavior is remarkably similar in the two variants (and always statistically insignificant). Hence, for brevity, we pool the data in our analysis.

<sup>15</sup> With a simultaneous 1 ECU increase by each group member, the total contribution to the public good increases by 6 ECU. The 1.4 ECU increase in the net earnings of high-return members is decomposed as follows: 0.4\* 6 ECU – 1 ECU (for the additional contribution). The 0.2 ECU increase in the net earnings of low-return members is decomposed as follows: 0.2\* 6 ECU – 1 ECU.

13.33 (which implies a cost of 6.67 ECU) and  $\sum_j r_{ji} = 13.33$  (which implies a gain of 20 ECU). These numbers imply that the maximum individual earnings in the experiment from both stages are  $36 - 6.67 + 20 = 49.33$  ECU and the maximum group earnings are 296 ECU.

In heterogeneous treatments, on average maximum earnings are the same as in homogeneous treatments, but given the different returns from the public account, when all contribute fully to the public account and assign the ‘optimal’ number of points (13.33), high-return subjects can earn up to 61.33 ECU ( $= 48 - 6.67 + 20$ ), and low-return subjects can earn up to 37.33 ECU ( $= 24 - 6.67 + 20$ ). Note that this results in inequality. Equality could be reached with minor loss of efficiency if low-return subjects assigned one reward point to a high-return subject and high-return subjects assigned the maximum of 13.33 points to a low-return subject. With such a strategy, both types of subjects would earn about 43 ECU (43.5 for the low-return and 42.85 for the high-return) each and the six-player groups would earn 259 ECU in total (instead of 296 ECU). Allowing individuals to reward each other is therefore a critical component of our mechanism and can reduce inequality substantially.

## 2.5. Procedures

The experiment was conducted at GATE-LAB, Lyon, France. 246 subjects (of which 53.66% were female) were drawn from the undergraduate student population from local engineering, business and osteopathic medical schools and were invited to participate using ORSEE (Greiner, 2015). Subjects were only invited if they had not previously participated in any similar social dilemma experiment. No subject participated in more than one session. We ran 13 sessions in total. While subjects interacted anonymously in six-person fixed groups, multiple groups under the same treatment conditions participated simultaneously in the laboratory. Sessions were conducted on computers using z-Tree software (Fischbacher, 2007). Table 1 provides a summary of the treatments and the number of independent groups in each treatment. We collected more data in the heterogeneous treatments as we expected to observe more variability in decision making in these treatments.

**Table 1** - Summary of experimental sessions

Treatment	Returns from public account	Communication	Rewards	Participants	Groups
HOM-NC	0.3	No	Yes	48	8
HOM-C	0.3	Yes	Yes	36	6
HET-NC	0.2 or 0.4	No	Yes	60	10

HET-C	0.2 or 0.4	Yes	Yes	102	17 <sup>(*)</sup>
Total	-	-	-	246	41

*Note:* HOM-NC: Homogenous groups without communication; HOM-C: Homogenous groups with communication; HET-NC: Heterogeneous groups without communication; HET-C: Heterogeneous groups with communication. <sup>(\*)</sup>Het-C includes 48 participants (8 groups) who had the possibility to send messages visible either to all group members or only to group members of the same type (Public-Private), and 54 participants (9 groups) whose messages were made visible to all group members systematically (Public).

Upon arrival, subjects drew a tag from an opaque bag, which randomly assigned them to a computer terminal. Sets of instructions were distributed at the beginning of each part and read aloud by the experimenter. The instructions and computer decision screens used neutral terminology, e.g., subjects “assigned” ECU to the public account instead of “contributing” to it; they received and assigned “addition” points, not “reward” points. Subjects answered a few control questions after reading the instructions to ensure that they understood the instructions and the process of decision-making. The instructions for the HET-C treatment are available in Appendix 1 (the instructions for the other treatments are available upon request). After the experiment, subjects filled out a survey intended to collect information about demographic characteristics and attitudes towards cooperation and rewarding others.

Subjects’ earnings were the sum of their payoffs in the 20 periods in ECU which were then converted to Euros at the end of the session. Subjects earned 22.93 Euro on average, including a show-up fee of 5 Euro. Payments were made individually in cash and in private in a separate room by a secretary who was not aware of the content of the experiment (this was made common knowledge in the instructions). Including the instructions and payment distribution time, sessions lasted between 90 and 120 minutes.

### 3. Results

We start this section by showing evidence for the existence of normative conflict in heterogeneous groups. Then, the presentation of the main experimental results is divided into two parts. In the first part, we present our main findings on how the normative conflict between equality and efficiency affects the efficacy of our mechanism in promoting cooperation and efficiency. In the second part, we explore the different contribution rules groups adopt and analyze the content of communication.

#### 3.1. *The existence of normative conflict in heterogeneous groups*

The existence of a normative conflict in our heterogeneous treatments is attested by the responses of our subjects in a post-experimental questionnaire. For the equal-earnings rule, using a scale

ranging from 1 (“I fully disagree”) to 7 (“I fully agree”), we asked them to tell us the extent to which they agree with the statement “Ideally, all group members should earn more or less the same from the experiment irrespective of whether they are of Type A [*i.e.*, high-return] or B [*i.e.*, low-return]”. For the equal-contributions rule, we asked them to tell us the extent to which they agree with the statement “Ideally, all group members should allocate the same number of ECU to the public account irrespective of whether they are of Type A or B and thus earn different amounts”.

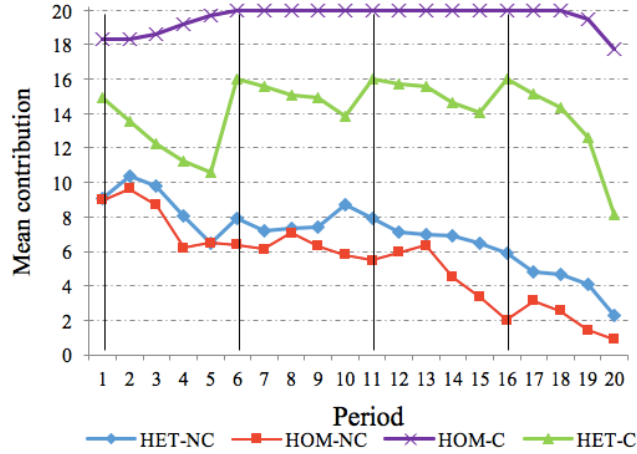
The evidence from the post-experiment questionnaire suggests that group heterogeneity was successful in triggering normative conflict in our experiment. The mean response on the equal-earnings rule are 4.09 (S.D.=1.96) for high-return subjects, and 5.11 (S.D.=1.99) for low-return subjects. Using a Mann-Whitney two-sided test and each individual as an independent observation we find that the difference is statistically significant ( $p=0.006$ ) suggesting that low-return subjects find the equal-earnings rule more appealing than their high-return counter-parts. The mean response on the norm of equal contribution is respectively 4.19 (S.D.=2.13) for high-return subjects and 3.49 (S.D.=2.37) for low-return subjects in heterogeneous groups. The difference is (weakly) statistically significant ( $p=0.086$ ) suggesting that high-return subjects find the equal-contributions rule more appealing than low-return subjects. Low-return individuals tend to agree more with the equal-earnings rule than the equal-contributions rule (two-tailed Wilcoxon tests,  $p=0.001$ ). While high-return participants agree more with the equal-contributions rule, the difference is not statistically significant ( $p=0.714$ ). Columns 6-9 of Table A3 in Appendix 3 give additional information about the subjects’ responses.

### *3.2. Cooperation and efficiency*

Figures 1 and 2 present respectively the evolution of contributions and earnings across treatments. Table 2 presents summary statistics on contributions and earnings in the various treatments, along with the results from non-parametric tests. We proceed to state our main results and provide support for them.

**Figure 1** – Mean contributions across treatments





Note: The vertical bars indicate the periods with communication (1, 6, 11 and 16) in the treatments with communication.

**Result 1:** Communication has a positive impact on contribution levels in all treatments, but the size of the effect in heterogeneous groups is only 38.17% of that in homogeneous groups.

**SUPPORT:** As can be seen in Figure 1 and Table 2, in the absence of communication, although group members could reward high contributors, average contributions unravel and are similar in homogeneous (HOM-NC:  $\bar{c} = 5.37$ ) and heterogeneous groups (HET-NC:  $\bar{c} = 6.97$ ) (two-sided Mann-Whitney rank-sum test,  $p=0.374$ ). Communication increases contributions in both homogeneous (HOM-C) and heterogeneous groups (HET-C), but its relative efficacy appears to be considerably limited in the presence of normative conflict: homogeneous groups contribute on average 97.85% of the maximum ( $=19.57/20$ ), which amounts to a 265% increase relative to contribution levels in HOM-NC. A Mann-Whitney test at the group level using the average contribution across periods as an independent observation indicates that the difference is statistically significant ( $p=0.002$ ). Heterogeneous groups on the other hand contribute on average 70.10% of the maximum when they can communicate ( $=14.02/20$ ) – an increase of 101.15% relative to contributions in HET-NC (Mann-Whitney,  $p<0.001$ ). Therefore, the increase in contributions due to the introduction of communication opportunities in heterogeneous groups is 38.17% of that in homogeneous groups. A Mann-Whitney test indicates that contributions are not significantly different in HOM-NC and HET-NC ( $p=0.374$ ), but are significantly lower in HET-C than in HOM-C ( $p<0.001$ ). ■

**Table 2** - Summary statistics on contributions and final earnings

Treatments	Homogeneous (1)	Heterogeneous			Ratio H/L (5)
		All Subjects (2)	High-return Subjects (3)	Low-return Subjects (4)	
<u>Mean contributions</u>					
Without communication (a)	5.37 (6.99)	6.97 (7.43) <sup>ns</sup>	9.53 (7.72)	4.40 (6.14) <sup>**</sup>	2.17
With communication (b)	19.57 (2.35)	14.02 (7.49) <sup>***</sup>	17.27 (6.18)	10.76 (7.27) <sup>***</sup>	1.60
(a) vs. (b)	$p=0.002$	$p<0.001$	$p<0.001$	$p=0.003$	
<u>Mean final earnings</u>					
Without communication (a)	25.63 (6.46)	27.02 (7.20) <sup>ns</sup>	29.25 (7.61)	24.79 (6.00) <sup>***</sup>	1.18
With communication (b)	46.35 (4.33)	39.52(11.04) <sup>***</sup>	44.66(11.00)	34.38 (8.37) <sup>***</sup>	1.30
(a) vs. (b)	$p=0.002$	$p<0.001$	$p<0.001$	$p<0.001$	

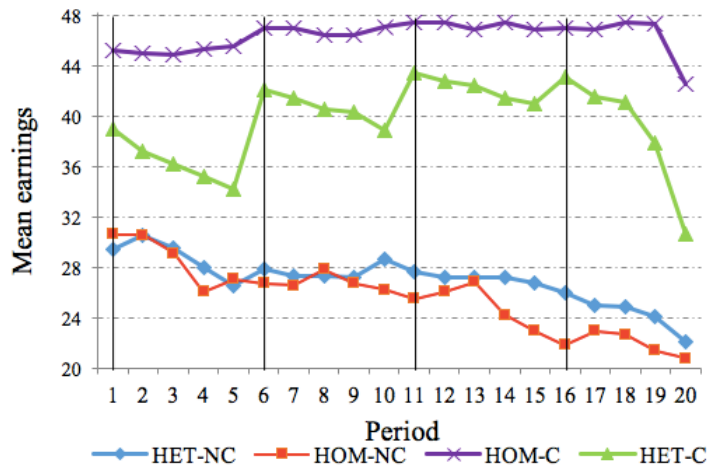
*Note:* Standard deviations in parentheses. Comparisons between conditions (a) vs. (b) report the  $p$ -values of Mann-Whitney tests in which each group averaged over all periods gives one independent observation. In column (2) (“All subjects”), <sup>\*\*\*</sup> and <sup>ns</sup> indicate significance at the 1% level and no significance ( $p>0.10$ ), respectively, in Mann-Whitney tests comparing the mean contribution and mean earnings, respectively, in heterogeneous (column (2)) and in homogeneous groups (column (1)). In column (4), <sup>\*\*\*</sup> and <sup>\*\*</sup> indicate significance at the 1% and 5% levels, respectively, in Wilcoxon tests comparing contributions (upper panel) or earnings (lower panel) of low-return subjects (column (4)) relative to high-return subjects (column (3)).

It is interesting to note that the effect of our mechanism on contributions in Figure 1, apart from being weaker in heterogeneous groups, also appears to be less *stable*: after an increase in periods of communication (*i.e.*, periods 1, 6, 11, and 16), contributions decline in the following periods. A linear regression with individual-level random effects and clustering at the group level shows that, on average, contributions in HET-C are 2.19 ECU higher in communication periods ( $p<0.001$ ) compared to non-communication periods. We explore the causes of this instability in heterogeneous groups in the following subsection. Next, we present our second main result regarding earnings in the experiment. This is a distinct question from how communication impacts on cooperation as subjects could reward each other.

**Result 2:** Communication has a positive impact on group earnings in all treatments, but the size of the effect in heterogeneous groups is only 57.22% of that in homogeneous groups.

**SUPPORT:** Figure 2 presents the evolution of earnings across treatments. On average, earnings are not significantly different, without communication, in homogeneous groups (HOM-NC:  $\bar{\pi}_t^2 = 25.63$ ) and in heterogeneous groups (HET-NC:  $\bar{\pi}_t^2 = 27.02$ ) (Mann-Whitney,  $p=0.286$ ). Earnings are higher with communication, but again we observe that the relative efficacy of communication appears to be considerably weakened by the presence of normative conflict: an individual in a homogeneous group earns on average 46.35 ECU, *i.e.*, 94% of the maximum. This amounts to an increase of 80.84% relative to HOM-NC, which is statistically significant (Mann-Whitney,  $p=0.002$ ). In contrast, individuals in heterogeneous groups earn on average 39.52 ECU, *i.e.*, 80.11% of the maximum. The increase in earnings relative to HET-NC is 46.26% and statistically significant (Mann-Whitney,  $p<0.001$ ). However, it is only 57.22% of the increase in earnings due to communication in homogeneous treatments. A Mann-Whitney test indicates that earnings are significantly lower in HET-C than in HOM-C ( $p<0.001$ ). ■

**Figure 2 – Mean earnings across treatments**



Note: The vertical bars indicate the periods with communication (1, 6, 11 and 16) in the treatments with communication.

The fact that earnings are 94% of the maximum when homogeneous groups can communicate is quite extraordinary. Unlike previous experiments where individuals had to agree only on how much to contribute, subjects in HOM-C also had to agree on their rewarding strategies. The design (see section 2.2) allowed for over-use of rewards (when  $13.33 < \bar{r}_{ij} \leq 20$ ) and under-use (when

$\bar{r}_{ij} < 13.33$ ). Avoiding wasteful over-use of the reward medium is not simple and requires coordination amongst group members. We conclude this subsection by investigating the aggregate patterns concerning the receipt of rewards across treatments.

**Result 3:** *Individuals in homogeneous groups receive on aggregate slightly more reward points than is optimal when they can communicate. In contrast, individuals in heterogeneous groups receive slightly less than the optimal amount of reward. Consequently, fewer reward points are received in heterogeneous versus homogenous groups.*

**SUPPORT:** Table 3 presents summary statistics concerning reward points received across treatments. Clearly, communication increases the number of reward points received substantially in both homogeneous groups ( $\bar{r}_{ij}$  [HOM-NC] = 1.34 vs.  $\bar{r}_{ij}$  [HOM-C] = 16.97; Mann-Whitney,  $p=0.002$ ), and heterogeneous groups ( $\bar{r}_{ij}$  [HET-NC] = 2.74 vs.  $\bar{r}_{ij}$  [HET-C] = 11.24;  $p<0.001$ ). More interestingly, the number of reward points is significantly higher than the optimal number of points ( $\bar{r}_{ij}=13.33$ ) in HOM-C (Wilcoxon sign rank,  $p=0.046$ ), while it is slightly (albeit significantly) lower in HET-C (Wilcoxon sign rank,  $p=0.024$ ). The number of reward points in HET-C is significantly lower than in HOM-C (Mann-Whitney,  $p=0.003$ ). For completeness, we report that the difference in reward points in HET-NC and HOM-NC does not reach a conventional level of statistical significance (Mann Whitney,  $p=0.197$ ). ■

**Table 3** - Summary statistics on the receipt of reward points

Treatments	Homogeneous	Heterogeneous		
		All subjects	High return subjects	Low return subjects
Without communication (1)	1.34 (2.32)	2.74 <sup>ns</sup> (4.53)	3.37 (5.14)	2.12 <sup>ns</sup> (3.72)
With communication (2)	16.97 (4.53)	11.24 <sup>***</sup> (7.16)	11.62 (7.31)	10.86 <sup>ns</sup> (6.99)
(1) vs. (2)	$p=0.002$	$p<0.001$	$p<0.001$	$p<0.001$

*Notes:* Entries are the mean number of reward points received. Standard deviations are in parentheses. Comparisons between conditions ((1) vs. (2) report the  $p$ -values of Mann-Whitney tests in which each group gives one independent observation. In column “All subjects”, \*\*\* and <sup>ns</sup> indicate significance at the 1% level and no significance ( $p>0.10$ ), respectively, in Mann-Whitney tests comparing the mean number of reward points received in heterogeneous and in homogeneous groups. In column “Low return”, <sup>ns</sup> indicates no significance ( $p>0.10$ ) in Wilcoxon tests comparing the mean number of points received by low-return and high-return subjects.

Result 3 provides two insights. The first is that the slight loss in efficiency relative to the maximum in HOM-C between periods 6 and 19 when they contribute their full endowment in the

public account discussed in the support for Result 2 (see Figures 1 and 2) stems from the fact that more reward points are used than what is required for maximum efficiency. We define “over-rewarding” as the case in which a subject receives more than 14 addition points (since in such cases, the additional points given generate a cost for the sender but no benefit to the receiver. Over-rewarding was never observed in HOM-NC, but it occurred in 483 cases out of 720 (67.08%) in HOM-C, and the average number of points received in the cases of over-rewarding is 19.46 (S.D.= 1.85). By comparison, over-rewarding occurred in 49 cases out of 1200 (4.08%) in HET-NC and in 768 cases out of 2040 (37.65%) in HET-C – which is significantly different from HOM-C at the 1% level.<sup>16</sup>

The second insight is that the inefficiency observed in heterogeneous groups is not only due to the low levels of contribution, but also to a relative under-use of rewards. If we define under-rewarding as the cases in which a subject receives less than 13 addition points, this occurred in 1019 cases out of 2040 in HET-C (49.95%) and in 1137 cases out of 1200 in HET-NC (94.75%) (in 5.00% and 99.79% in HOM-C and HOM-NC, respectively). In particular, the lower-than-optimal number of reward points in HET-C (11.24 on average *vs.* 13.33) translates into a mean efficiency loss of 2.09 ECU (=  $[13.33-11.24]*[1.5-0.5]$ ).

We can therefore calculate the extent of the inefficiency that is due to suboptimal contributions and suboptimal use of rewards. Recall that earnings in HET-C were on average 9.81 ECU lower than the maximum earnings (49.33 – 39.52). Therefore, the suboptimal number of reward points accounts for 21.30% of the loss in efficiency, whereas the remaining 79.70% is due to the relatively low levels of contribution. Of course, this conclusion differs if we consider each type separately. The suboptimal number of rewards received accounts for 83.73% of the loss of efficiency for the low-return subjects and for only 10.25% of the loss for the high-return subjects.<sup>17</sup> This suggests that the loss of efficiency comes both from a reluctance to reward low-return subjects and from the low contributions made by low-return subjects.

---

<sup>16</sup> The lower frequency of over-rewarding in the treatments without communication can be explained by the relatively lower level of contributions in these treatments compared to those with communication. But communication also helped some groups adjust their strategy when they realized that they were over-rewarding some subjects. In particular, subjects talked about using the identifier of the players to assign points. See, for example, the messages exchanged by one group in period 6: “Don’t forget: 14 points in the second part...14 points=20 UME... Ah!\*! I gave 20 points in each part... Give to the player below you: 1 gives to 2, 2 to 3, 3 to 4, 4 to 5 and 6 to 1... 14 to the guy below you...”

<sup>17</sup> The low-return people receive on average 10.86 reward points, which represents a mean efficiency loss of 2.47 ECU and their earnings are on average 2.95 ECU lower than the maximum (37.33-34.38). The high-return people receive 11.62 points, which represents an efficiency loss of 1.71 ECU and their earnings are on average 16.68 ECU lower than the maximum (61.34-44.66).

In order to better understand why our mechanism based on communication and rewards is unable to facilitate full cooperation in heterogeneous groups, we turn our attention to the behavior of high and low-return players in heterogeneous treatments.

### 3.3. Contribution rules in HET-C and rewarding behavior

Having found support for our hypothesis that the normative conflict in heterogeneous groups undermines the efficacy of communication, we turn our attention to individual groups. We are interested in understanding *why* normative conflict has this effect. Is it that groups fail to agree on a contribution rule or is it that they adopt rules which favor equality over efficiency? To answer this question, we focus on behavior in treatment HET-C since in HET-NC, it is clear that contribution unravels over time, whereas in HOM-C all groups agree to contribute fully for almost the entire experiment.

Table 4 presents median contributions, reward points and profits from periods 6 to period 19 (*i.e.*, from the second communication round to the end of the experiment, excluding the first block of periods where subjects learn the game and excluding the last period to avoid end-game effects). The use of medians provides a clearer picture of the contribution rules favored by groups as it is less sensitive to outliers.

Groups can be classified into one of three categories with regards to the contribution rule they adopted: groups that adopted the contribution rule favoring equality over efficiency; groups that adopted the rule favoring efficiency over equality; and groups that struck a compromise between these two rules. The classification of groups in Table 4 and the conclusions derived from it are robust to using medians for periods 11 to 19, or periods 16 to 19.

**Result 4:** *Most heterogeneous groups manage to adopt stable contribution rules with positive contributions to the public account.*

**SUPPORT:** As can be seen in Table 4, 15 of the 17 groups (*i.e.*, 88.2%) reach a *stable* agreement about which contribution rule to follow. This can be seen in the fifth column that presents the “Contribution ratio trend”. The contribution ratio is the ratio of contributions by high return subjects to low return subjects; the “trend” is the coefficient from an OLS regression with the contribution ratio as the dependent variable and period as the only independent variable. Except group 1 which experienced an increase in the contribution ratio between periods 6 and 19, the

contribution ratio did not change significantly over time in any of the other groups.<sup>18</sup> Although the contribution ratio remained constant over time in group 12, contributions collapsed to 0 for all group members in period 15 and following periods, except in period 16.<sup>19</sup> ■

**Result 5:** *Only a small minority of heterogeneous groups select the efficient contribution rule. Most groups choose either a contribution rule that strictly prioritizes equality over efficiency, or a rule that strikes a compromise between equality and efficiency concerns.*

**SUPPORT:** As can be seen in Table 4, only four of the 17 groups (*i.e.*, 23.5%) reached a stable agreement in which high and low-return group members contributed their entire endowment. Six of the 17 groups (*i.e.*, 35.3%) chose the contribution rule that maximizes group earnings but conditional on high and low-return members having equal (first-stage) earnings, *i.e.*,  $c_H = 20$  and  $c_L = 5$ . That is, these groups prioritize equality over efficiency. Six more groups eventually settled on rules satisfying  $c_H = \alpha c_L$ , where  $1 < \alpha < 4$ . These rules are in line with Young (1998) who argues that in cases of normative conflict individuals would eventually reach a compromise. It is also worth noting that a contribution ratio of two – observed in three groups – is in line with the “proportionality rule” according to which group members should contribute in proportion to their relative benefit from the public account (Reuben and Riedl, 2013).<sup>20</sup> In sum, of the 16 groups that adopted and maintained a contribution rule with positive contributions (*i.e.*, all groups except group 12), 12 groups (*i.e.*, 75%) agree on a rule consistent with a concern for equality. ■

Results 4 and 5 help explain the main findings in the previous section, *i.e.*, why efficiency is lower in heterogeneous groups. The reason cooperation rates and efficiency are lower in heterogeneous groups in HET-C compared to HOM-C is not because groups cannot find an agreement. It is because 75% of the groups in HET-C showed a concern for reducing inequality in earnings from the first stage and did not use rewards in the second stage to compensate those who get a lower return from cooperation. This concern places a limit to the ability of our mechanism to increase efficiency up to its maximum level. It is important to note that an analysis of the content of

---

<sup>18</sup> In periods 16-19, in group 1, the median contribution of the high-return players increased to 20 while that of low types remained unchanged at 15. This implies that, all together, 16 out of 17 groups (94.1%) adopted rules which prevented the decay of cooperation.

<sup>19</sup> Group 12 is solely responsible for the downward trend observed in contributions in Figure 1 in periods following communication. In none of the other groups do contributions decline significantly following communication.

<sup>20</sup> To evaluate whether differences in effort during the real-effort stage affected the selection of contribution rules, we calculated the ratio “performance of high-return participants / performance of low-return participants in the preliminary task” for each group, which is 1.26 in the groups that favor equality over efficiency, 1.18 in the groups favoring efficiency over equality, and 1.35 in the other groups.

communication, as well as the responses to a post-experiment survey at the group level fully supports our classification of groups and the conclusions above. This analysis can be found in Table A3 in Appendix 3. We conclude this section with an examination of rewarding behavior across groups adopting different contribution rules.

**Result 6:** *Overall, rewards are not used to alleviate earning equalities. Low-return group members do not receive on average more reward points than high-return members.*

**SUPPORT:** As can be seen in Table 3, for all periods, and in Table 4 for periods 6 to 19, on average, low and high-return subjects receive a similar amount of points. A Wilcoxon sign-rank test shows that the average number of reward points received by high-return subjects across all periods and groups is far from being significantly different from that of low-return subjects ( $p=0.981$ , two-tailed). Even if we concentrate our attention on the 11 groups that did not adopt the equal-earnings rule, *i.e.*, the groups in which one would expect that low-return members would be rewarded more, the difference is small and far from being statistically significant (high: 12.6, low: 11.9;  $p=0.929$ ). In two of the four groups prioritizing efficiency over equality, there is no difference between high and low type subjects. In “other groups” on average low types receive more reward points, but not significantly so ( $p=0.203$ ).<sup>21</sup> ■

The analysis of the content of communication shows that groups rarely discussed the idea of using rewards to compensate low return group members (see the analysis of the content of communication in Appendix 3). As a result, overall the ratio of the high-return subjects’ earnings to the low-return subjects’ earnings is higher in treatment with communication than without (1.30 *vs.* 1.18, see column (5) in Table 2).<sup>22</sup> Moreover, the degree of earnings inequality within groups, as measured by the mean standard deviation of total earnings, remains very different across these different categories of groups, regardless of the period (see Figure 3). It is obviously much lower in groups favoring equality over efficiency, compared to the two other categories.

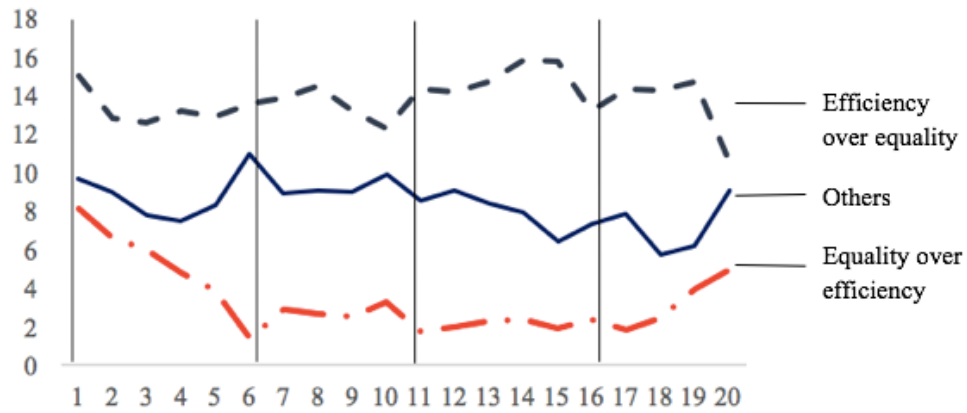
---

<sup>21</sup> As we show in Appendix 3, we find no evidence that addition points are assigned to reward norm compliance with the contribution rules followed by the groups.

<sup>22</sup> Since both types of subjects contribute more in HET-C than in HET-NC, this could suggest that communication increases tolerance of inequality. However, responses to a post-experimental questionnaire indicate that subjects in HET-NC tend to agree less than in HET-C with the statement that “ideally, members should earn more or less the same regardless of their type” (the lower panel of Table A3 in Appendix 3).



**Figure 3** – Evolution of within-group inequality over time, by category of groups (HET-C treatment)



*Notes:* The three curves indicate the evolution of the mean standard deviation of total earnings within groups. The vertical bars indicate the periods with communication (1, 6, 11 and 16) in the treatments with communication.

**Table 4** – Group-level analysis of contributions, rewards and earnings in the HET-C treatment

Group number	Contribution high-return	Contribution low-return	Contribution ratio	Contribution ratio trend	Reward received by high-return	Reward received by low-return	Earnings high-return	Earnings low-return
(1)	(2)	(3)	(4)	(5) <sup>+</sup>	(6)	(7)	(8)	(9)
<i>Equality-over-Efficiency Groups</i>								
4 <sup>#</sup>	20	5	4	-0.11	14	14	40	40
5 <sup>#</sup>	20	5	4	0.00	12.5	12	40.5	40.5
9	20	5	4	0.00	15	15	42.5	42.5
11	20	5	4	0.00	20	20	40	40
13	20	5	4	0.18	14	14	40	40
14	20	5	4	0.00	13	13	43	43
<i>Efficiency-over-Equality Groups</i>								
3 <sup>#</sup>	20	20	1	0.00	18	6.5	60.3	30.3
7 <sup>#</sup>	20	20	1	0.00	20	14	58	34.3
10	20	20	1	-0.01	10	10	58	34
17	20	20	1	0.00	15	15	60.5	36.5
<i>Other Groups</i>								
1 <sup>#</sup>	15	15	1	0.03***	20	20	51	34.7
2 <sup>#</sup>	20	10	2	-0.01	9.5	5	42.3	33
6 <sup>#</sup>	20	0	-	0.00	11	13	43.5	48.2
8 <sup>#</sup>	20	10	2	0.00	11	11.5	46.5	36
12 <sup>a</sup>	20	20	1	0.00	1.5	3.5	38.5	28
15	20	15	1.33	0.00	10	16	50.5	40
16	20	10	2	-0.03	12	16	41.8	40
Average	19.7	11.2	2.3	0.00	13.3	12.9	46.3	37.8

*Notes:* Entries are *medians* from periods 6 to 19. # indicates the groups in which subjects could use two chat boxes to send messages visible to everyone in their group and private messages to their group members with the same type (Public-Private). <sup>a</sup> In Group 12, although the contribution ratio does not change, cooperation collapses with all contributing 0 after period 14 (and except period 16). So, it is not consistent with efficiency. <sup>+</sup> Coefficients from an OLS regression for periods 6 to 19 with the “contribution ratio” as the dependent variable and “period” the only independent variable. \*\*\*  $p < 0.001$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$ .

#### 4. Discussion

In many instances when the need for cooperation arises, concerns for equality are at odds with efficiency. We presented evidence from a laboratory experiment investigating how such tension between equality and efficiency could affect the ability of groups to cooperate. The ‘news’ are both good and bad. All groups except one (16 out of 17; *i.e.*, 94.1%) managed to agree on specific contribution rules with positive contribution levels and follow them. This is a novel and promising finding which we did not anticipate. As Elster (1989) notes in a discussion of evidence from social psychology, individuals often try to select contribution rules as social norms in a self-serving manner. In line with this, Nikiforakis *et al.* (2012) found that high-return group members tried to enforce efficiency, while low-return group members tried to enforce equality in their groups. In contrast, in our experiment, even if the different returns from the public goods created a tension, our mechanism allowing groups to periodically communicate and reward group members, enabled virtually all of them to reach stable cooperation agreements with positive contributions. This finding therefore supports the view that, at least in some instances, “self-governance is possible” (Ostrom *et al.*, 1992). This is the good news.

The bad news is that concerns for equality appear to impose a limit on the efficacy of self-governance mechanisms: less than a quarter of groups (4 out of 17; *i.e.*, 23.5%) chose to follow the contribution rule that maximizes group earnings. This is another unexpected finding. First, the inequality of benefits in our experiment results from the participants’ relative performances in a preliminary task. Therefore, it might be viewed as a relatively “fair” inequality with participants having a relatively higher tolerance towards inequality and a lower demand for redistribution (see *e.g.*, Cappelen *et al.*, 2007; Krawczyk, 2010; Gee *et al.*, in press). Second, groups could have eradicated inequality through the use of rewards. But rewards are used by the majority of groups neither to minimize inequality nor to maximize overall efficiency across the two stages (*i.e.*, by contributing fully and assigning as many rewards as possible). Instead of incorporating the two stages in their strategy in combining efficiency and equality concerns, subjects seem to use rewards either to reciprocate cooperation in the first stage or to maximize earnings in the second stage when the return from action (*i.e.*, from rewarding) is the same for all the players.

Why do individuals fail to use rewards to minimize inequality when clearly they seem to care about equality? This is a puzzle and indicates that more empirical work is needed in order to establish the limits of our conclusions and better understand how to effectively promote cooperation in heterogeneous populations. One thing is clear from our experiment: although individuals appear to be far more able to cooperate than typically assumed by economists, the

problem of cooperation is also more complex than simply aligning private and ‘social’ interests. The latter includes not only concerns for efficiency, but also for equality, and possibly other kinds of normatively appealing rules such as those for equity.<sup>23</sup>

The fact that fairness concerns influence which contribution rules groups adopt suggests numerous interesting avenues for future research. Most experimental work examining the efficacy of mechanisms for promoting cooperation and efficiency has considered the special case of homogeneous groups. In heterogeneous populations, in addition to overcoming free riding incentives, group members have to resolve the conflict between efficiency and inequality. Some mechanisms will be better at dealing with this normative conflict than others. For example, mechanisms that rely solely on coercion and peer punishment may be less effective (see *e.g.*, Nikiforakis *et al.* 2012; Cason and Gangadharan, 2012; Cason and Gangadharan, 2014; Dekel *et al.*, 2017). Moreover, the efficacy of a given mechanism may depend on the type of heterogeneity (*e.g.*, different returns to cooperation *vs.* different abilities to contribute), the extent of the heterogeneity (*e.g.*, small *vs.* large differences) or the ability of individuals to opt out from mechanisms that lead to unequal outcomes (*e.g.*, Chen and Plott, 1996). Therefore, it is important to test the performance of mechanisms also in heterogeneous and not only in homogeneous populations as is often the case. The source of heterogeneity could be also important. What if heterogeneity was generated by luck instead of by merit, as in our study? This might affect which rule is chosen by a group to deal with the normative conflict. Testing this would be an interesting extension of our work.

To conclude, the main message from our experiment is the following: in many instances in daily life, the predicament with cooperation may not be individuals’ unwillingness to cooperate due to private incentives to free ride as often assumed (see our HOM-C treatment), but their inability to reach an agreement that will simultaneously satisfy the fairness concerns of *all* individuals and their heterogeneous preferences. Given the widespread need for cooperation in the modern world (*e.g.*, climate negotiations and the recent talks in the Eurozone) acknowledging the complexity of cooperation in heterogeneous populations and trying to understand how to engineer cooperation appears to be not only an interesting research topic, but also a topic of great practical significance.

---

<sup>23</sup> According to one common reading, the principle of equity prescribes that one is compensated in proportion to his or her input (*e.g.*, Konow, 2003). The real-effort stage in our experiment implies that equity is at odds with equality and makes the dominant role of the latter, even more striking.

## References

- Almås, I., Cappelen, A.W., Lind, J.T., Sørensen, E.Ø., Tungodden, B. (2011). Measuring unfair (in)equality. *Journal of Public Economics* 95(7), 488-499.
- Anderson, L. R., J. M. Meller, Milyo, J. (2008). Inequality and public good provision: an experimental analysis. *Journal of Socio-Economics* 37 (3), 1010-1028.
- Andreoni, J., Gee, L.K. (2012). Gun for hire: delegated enforcement and peer punishment in public goods provision. *Journal of Public Economics* 96 (11), 1036-1046.
- Bochet, O., Page, T., Putterman, L. (2006). Communication and punishment in voluntary contribution experiments. *Journal of Economic Behavior and Organization* 60, 11–26.
- Bochet O, Putterman L. (2009). Not just babble: A voluntary contribution experiment with iterative numerical messages. *European Economic Review* 3, 309-326.
- Bolton, G.E., Ockenfels, A. (2000). ERC: A Theory of Equity, Reciprocity, and Competition. *American Economic Review* 90(1), 166-193.
- Brosig, J., J. Weimann, Ockenfels, A. (2003). The effect of communication media on cooperation. *German Economic Review* 4(2), 217–241.
- Browning, E.K., Johnson, W.R. (1984). The trade-off between equality and efficiency. *Journal of Political Economy* 92 (2), 175-203.
- Buckley, E., Croson. R. (2006). Income and wealth heterogeneity in the voluntary provision of linear public goods. *Journal of Public Economics* 90(4-5), 935-955.
- Cappelen, A.W., Drange Hole, A., Sørensen, E.Ø., Tungodden, B. (2007). The Pluralism of Fairness Ideals: An Experimental Approach. *American Economic Review* 97(3), 818-827.
- Cason, T.N., Kahn, F.U. (1999). A laboratory study of voluntary public goods provision with imperfect monitoring and communication. *Journal of Development Economics* 58, 533-552.
- Cason, T.N., Gangadharan, L. (2013). Empowering neighbors versus imposing regulations: an experimental analysis of pollution reduction schemes. *Journal of Environmental Economics and Management*. 2013, 65, 469-484.
- Cason, T., Gangadharan, L. (2014). Swords without covenants do not lead to self-governance, *Journal of Theoretical Politics*, 1-30.
- Chan, K. S., Mestelman, S., Moir, R., Muller, R. A. (1996). The voluntary provision of public goods under varying income distributions. *The Canadian Journal of Economics* 29(1), 54-69.
- Chan, K. S., Mestelman, S., Moir, R., Muller, R. A. (1999). Heterogeneity and the voluntary provision of public goods. *Experimental Economics* 2(1), 5-30.
- Charness G., Rabin, R. (2002). Understanding Social Preferences with Simple Tests. *Quarterly Journal of Economics* 117(3), 817-869.
- Chen, Y., Plott, C. R. (1996). The Groves-Ledyard mechanism: An experimental study of institutional design. *Journal of Public Economics* 59 (3), 335-364.
- Cherry, T., Kroll, S., Shogren, J. (2005). The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab. *Journal of Economic Behaviour and Organization* 57, 357-365.

- Charness, G., Masclet, D., Villeval M.C. (2014). The dark side of competition for status. *Management Science* 60(1), 38–55.
- Cooper, D.J., Kagel, J.H. (2016). Other regarding preferences: A selective survey of experimental results, in J.H. Kagel and A.E. Roth (eds.), *Handbook of Experimental Economics*, Vol. 2, Princeton: Princeton University Press, 217-274.
- Dekel, S., Fischer, S., Zultan, R. (2017). Punishment and reward institutions with harmed minorities. *Journal of Public Economics*, 146(c), 87-96.
- Egas, M., Riedl, A. (2008). The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of the Royal Society B - Biological Sciences* 275 (1637), 871-878.
- Erkal, N., Gangadharan, L., Nikiforakis, N. (2011). Relative Earnings and Giving in a Real Effort Experiment. *American Economic Review* 101(7): 3330-3348.
- Falkinger, J. (1996). Efficient Private Provision of Public Goods by Rewarding Deviations from Average. *Journal of Public Economics* 62(3), 413–422.
- Falkinger, J., Fehr, E., Gächter, S., Winter-Ebmer, R. (2000). A Simple Mechanism for the Efficient Provision of Public Goods: Experimental Evidence. *American Economic Review* 90(1), 247-264.
- Fehr, E., Schmidt, K.M., (1999). A Theory of Fairness, Competition and Cooperation. *Quarterly Journal of Economics* 114, 817-868.
- Fehr, E., Gächter, S. (2000). Cooperation and Punishment in Public Goods Experiments. *American Economic Review* 90 (4), 980-994.
- Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments, *Experimental Economics* 10(2), 171-178.
- Fischbacher, U., Schudy, S., Teyssier, S. (2014). Heterogeneous Reactions to Heterogeneity in Returns from Public Goods. *Social Choice and Welfare* 43(1), 195-217.
- Fisher, J., Isaac, R., Shatzberg, J., Walker, J. (1995). Heterogeneous demand for public goods: behavior in the voluntary contributions mechanism. *Public Choice* 85(3), 249-266.
- Gee, L.K., Migueis, M, Parsa S. (in press). Redistributive choices and increasing income inequality: experimental evidence for income as a signal of deservingness. *Experimental Economics*.
- Greiner, B. (2015). Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE. *Journal of the Economic Science Association* 1, 114-125.
- Groves, T., Ledyard, J.O. (1977). Optimal allocation of public goods: A solution to the ‘free-rider’ problem, *Econometrica* 4, 783–809.
- Hackett, S., Schlager, E., Walker, J. (1994). The Role of Communication in Resolving Commons Dilemmas: Experimental Evidence with Heterogeneous Appropriators. *Journal of Environmental Economics and Management* 27, 99-126.
- Isaac, R., Walker, J. (1988). Communication and free-riding behavior: the voluntary contribution mechanism. *Economic Inquiry* 26, 585-608.
- Jackson, M., Moulin, H. (1992). Implementing a public project and distributing its cost, *Journal of Economic Theory* 57, 125-140.

- Janssen, M., Holahan, R., Lee, A., Ostrom E. (2010). Lab Experiments for the Study of Social-Ecological Systems. *Science* 328 (5978), 613-617.
- Konow, J. (2003). Which Is the Fairest One of All? A Positive Analysis of Justice Theories. *Journal of Economic Literature* Vol. XLI, 1188–1239.
- Krawczyk, M. (2010). A glimpse through the veil of ignorance: Equality of opportunity and support for redistribution. *Journal of Public Economics* 94 (1), 131-141.
- Laffont, J.J. (1987). Incentives and the allocation of public goods. In the *Handbook of Public Economics*, Vol. 2, 537–69, Eds. A. J. Auerbach and M. Feldstein, Amsterdam: North-Holland.
- Ledyard, J. O., (1995). *Public goods: a survey of experimental research*. In the *Handbook of Experimental Economics* 2, 111-194, eds: Kagel J., and A. Roth, Princeton University Press, Princeton, New Jersey.
- Markussen, T., Putterman, L., Tyran, J-R. (2014) Self-Organization for Collective Action: An Experimental Study of Voting on Sanction Regimes. *Review of Economic Studies* 81 (1), 301-324.
- Masclet, D., Noussair, C., Tucker, S., Villeval, M.C. (2003). Monetary and Non-Monetary Punishment in the Voluntary Contributions Mechanism. *American Economic Review* 93 (1), 366-380.
- Myerson, R. B. (2008). Mechanism design. *The New Palgrave Dictionary of Economics*. Second Edition. Eds. S. N. Durlauf and L. E. Blume. Palgrave, Macmillan.
- Nikiforakis, N. (2008) Punishment and counter-punishment in public good games: Can we really govern ourselves? *Journal of Public Economics* 92 (1-2), 91–112.
- Nikiforakis, N., Noussair, C.N., Wilkening, T. (2012). Normative Conflict and Feuds: The Limits of Self-Enforcement. *Journal of Public Economics* 96 (9-10), 797-807.
- Noussair, C.N., Tan, F. (2011) Voting on punishment systems within a heterogeneous group. *Journal of Public Economic Theory* 13 (5), 661-693.
- Okun, A.M (1975). *Equality and efficiency: The big tradeoff*. Brookings Institution, Washington D.C.
- Oprea, R., Charness, G., Friedman, D. (2014). Continuous time and communication in public-goods experiments. *Journal of Economic Behavior & Organization* 108, 212-223.
- Ostrom, E., Walker, J., Gardner, R. (1992). Covenants with and without a sword: self-governance is possible. *American Political Science Review* 86(2), 404-417.
- Ostrom, E. (2012). Experiments Combining Communication with Punishment Options Demonstrate How Individuals Can Overcome Social Dilemmas. *Behavioral and Brain Sciences* 35(1), 33–34.
- Rand, D.G., Dreber, A., Ellingsen, T., Fudenberg, D., Nowak, M.A. (2009). Positive interactions promote public cooperation. *Science* 325 (5945), 1272-1275.
- Reuben E., Riedl, A. (2009). Public goods provision and sanctioning in privileged groups. *Journal of Conflict Resolution* 53: 72-93.
- Reuben, E., Riedl, A. (2013). Enforcement of Contribution Norms in Public Good Games with Heterogeneous Populations. *Games and Economic Behavior* 77, 122-137.

- Rockenbach, B., Milinski (2006). The efficient interaction of indirect reciprocity and costly punishment. *Nature* 444, 718-723.
- Sally, D. (1995). Conversation and cooperation in social dilemmas: a meta-analysis of experiments from 1958 to 1992. *Rationality and Society* 7, 58-92.
- Samuelson, P.A. (1954). The pure theory of public expenditure. *Review of Economics and Statistics* 36 (4), 387-389.
- Sefton, M., Shupp, R., Walker, J.M. (2007). The effect of rewards and sanctions in provision of public goods. *Economic Inquiry* 45 (4), 671-690.
- Sutter, M., Haigner, S., Kocher, M.G. (2010). Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations. *Review of Economic Studies* 77 (4), 1540-1566
- Tan, F. (2008). Punishment in a linear public good game with productivity heterogeneity. *De Economist* 156 (3), 269-293.
- Tavoni, A., Dannenberg, A., Kallis, G., Löschel, A. (2011). Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proceedings of the National Academy of Sciences of the United States of America* 108 (29), 11825-11829.
- Varian, H.R. (1994). Sequential provision of public goods. *Journal of Public Economics* 53, 165-186.
- Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology* 51 (1), 110-116.
- Young, H.P., 1998. Conventional contracts. *Review of Economic Studies* 65, 773-792.
- Zelmer J. (2003). Linear Public Goods Experiments: A Meta-Analysis. *Experimental Economics* 6, 299-310.
- Weng, Q., Carlsson, F. (2015). Cooperation in teams: The role of identity, punishment, and endowment distribution. *Journal of Public Economics* 126, 25-38.
- Wiggins, S.N., Libecap, G.D. (1985). Oil field Unitization; Contractual Failure in the Presence of Imperfect Information. *American Economic Review* 75 (3), 368-385.
- World Bank (2010). [Integrating development into a global climate regime](#).



# ONLINE SUPPLEMENTARY MATERIAL

## APPENDIX 1. Instructions for the Het-C treatment (translated from French)

### General information

You are now taking part in an experiment in economics. Your earnings depend on your decisions and the decisions of other participants. It is therefore important that you read the following instructions carefully. The instructions which we have distributed to you are for your private information. Please do not communicate with the other participants during the experiment.

During the experiment we shall not speak of Euro, but of Experimental Currency Units (ECU). Your entire earnings will be calculated in ECUs. At the end of the experiment the total amount of ECUs you have earned will be converted to Euro at the rate of **10 ECU = 25 cents** and will be immediately paid to you in cash. In addition, we will give you a one-off payment for showing up on time of €5. All payments will be made in private by somebody who is not aware of the content of the experiment and none of the other participants will ever know how much you earned in the experiment.

The experiment is divided into two parts. Here, we explain the first part of the experiment. Once the first part is finished you will receive detailed information about the second part of the experiment.

### Part 1

#### *The task*

In the first part, all participants will perform an encoding task. The task is the same for everyone. You will be presented with a set of letters that form “words” with no specific meaning and your task will be to encode these letters by substituting them with numbers using the Table located permanently at the top of your computer screen, as seen in Figure 1.

1 of 2

Temps restant (en secondes): 90

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
8	12	14	10	9	6	24	22	7	5	11	3	18	1	21	16	23	2	13	19	25	4	26	17	20	15

Vous codez actuellement le mot numéro 1

MOT: I A N I V H X I

CODE: 7 8 7 1 4 22 17 7

Notes:

- Utilisez TAB pour passer à la boîte suivante rapidement.
- Utilisez Shift+TAB pour revenir à la boîte précédente rapidement.
- Une fois tous les codes donnés, appuyez sur le bouton "OK" afin de vérifier le codage et passer au mot suivant.

OK

Figure 1. Example of task on the screen

Example: You are given the word LFAT. The Table shows that L=3, F=6, A=8, and T=19.

Once you encode a word correctly, the computer will prompt you with another word to encode. Once you encode that word, you will be given another word and so on. **This process will continue for 10 minutes.** All group members will be given the same words to encode in the same sequence.

*The purpose of the task*

The relative performance of each individual in the task will influence the maximum amount they can earn in the second part of the experiment. This will happen in the following way. In the second part, participants will be divided into groups of six. You will therefore be in a group with five other participants. Each group member will be given an endowment (20 ECU) and will be asked to divide it between two accounts. There will be two types of participants in each group. We will refer to them as “Type A” and “Type B”. Of the six individuals in each group, three will be of type A and three of type B.

The difference between the two types of participants is the return they get from accounts in which they put their endowment. In particular, **Type A participants will obtain a higher return than Type B participants.** Information about the exact returns will be given in the second part.

*Relative Performance in Part 1 and Type Allocation*

The allocation of types depends on the relative performance of each person in the task compared with that of the other participants in this session. At the end of the Part 1, the computer will rank the participants based on the number of words they encoded. The top 50% of the individuals in the session will be assigned the role of Type A participants. The bottom 50% of the individuals will be assigned the role of Type B participants. If two or more participants tie, the computer will determine their type randomly.

You will be informed as to whether you are of Type A or Type B at the beginning of Part 2 of the experiment. Your type will remain the same throughout the session. You will not be informed about the precise number of words encoded by each of your group members until the end of the experiment. At that time, this will be done using a table indicating the number of words decoded by each member of your group.

Please read again these instructions. If you have any question, we will answer your questions in private.

**Control Questions**

Please answer the following questions. If you have any questions or have answered all questions, please raise your hand and one of the experimenters will come to you. 1. What does the task in the first part determine?

.....

2. Consider the following example. Assume that there are twelve participants in the session. The following Table indicates the number of words decoded by each participant (These numbers are just an example and should not be taken as an indication about the number of words to encode.)

Participant ID	Number of words	Participant ID	Number of words
Participant 1	5	Participant 7	20
Participant 2	10	Participant 8	8

Participant 3	12	Participant 9	6
Participant 4	7	Participant 10	22
Participant 5	9	Participant 11	10
Participant 6	15	Participant 12	4

In the boxes below tick the participant(s) who will be assigned the role of a Type-A participant in part 2.

<input type="checkbox"/> Participant 1	<input type="checkbox"/> Participant 7
<input type="checkbox"/> Participant 2	<input type="checkbox"/> Participant 8
<input type="checkbox"/> Participant 3	<input type="checkbox"/> Participant 9
<input type="checkbox"/> Participant 4	<input type="checkbox"/> Participant 10
<input type="checkbox"/> Participant 5	<input type="checkbox"/> Participant 11
<input type="checkbox"/> Participant 6	<input type="checkbox"/> Participant 12

3. Please state whether the following sentences are true or false.

a. Type-A participants will have a higher return from the account in Part 2.

True       False

b. Type-B participants will have a higher return from the account in Part 2

True       False

## Part 2

This part of the experiment will consist of 20 periods. Your final gain in this experiment will be the sum of your gains in each of these 20 periods. At the start of this part of the experiment, you will be randomly divided into groups of six. You will therefore be in a group with 5 other participants. **The composition of each group will remain the same throughout the experiment.**

In the beginning of each period, each participant in your group will be randomly assigned a number for identification (i.e. Participant 1, Participant 2, Participant 3, Participant 4, Participant 5, and Participant 6). Identification numbers will be randomly re-assigned in every period. For example, an individual who is Participant 2 in the 1st period, might be Participant 1 in the 2nd period, Player 4 in the 3rd period and so on. **This means that although the participants in your group will remain the same throughout the experiment, you will not be able to link the actions of any given participant across periods.** Once the experiment is over, the identities of each participant will be kept anonymous.

Based on the number of words that you and the members of your group encoded in Part 1, you have been assigned the role of either a “Type A” or a “Type B” participant. At the beginning of Part 2, the computer will inform you about your type: Type A or Type B. **Your type will remain the same throughout the experiment** and will influence the return that you receive from the *public* account as explained below. Each of the groups will have three Type A and three Type B participants.

Once the experiment is over, you will be paid in private and at no point will your group or participant number be revealed. No one will know who was in their group or what actions were carried out by each individual.

Each of the 20 periods is divided into two stages. Below we discuss the two stages in detail.

### ***First Stage***

At the beginning of each of the 20 periods, each participant will receive 20 ECU. In the following, we shall refer to this amount as the “endowment”. Your task in the first stage is to decide how to use your endowment. **You have to decide how many of the 20 ECU you want to allocate to a public account (from 0 to 20) and how many of them to keep for yourself.** You will be able to make your decision by using a screen as the one in Figure 2 (shown for a participant of Type A). The consequences of your decision are explained in detail on the next page.

Once all the group members will have chosen their allocation to the public account, you will be informed about each group member’s type and allocation to the public account, your income from the public account and your payoff in this period through a screen as the one seen in Figure 3.

The screenshot shows a web interface for the first stage of the experiment. At the top, it indicates 'Period 1 of 2' and 'Temps restant (en secondes): 0'. The main text reads: 'Vous disposez de 20 UME' and 'Vous pouvez allouer autant d'UME que vous souhaitez dans le compte public.' Below this, it states 'Vous êtes le joueur 1' and 'Votre TYPE est A'. The interest rate is shown as 'Votre taux de rendement du compte public sera 0.40'. At the bottom, there is a prompt 'Indiquez SVP le nombre d'UME alloués au compte public' followed by a text input field. An 'OK' button is located in the bottom right corner.

Figure 2. Example of screen-shot in the first stage ( Type A participant)

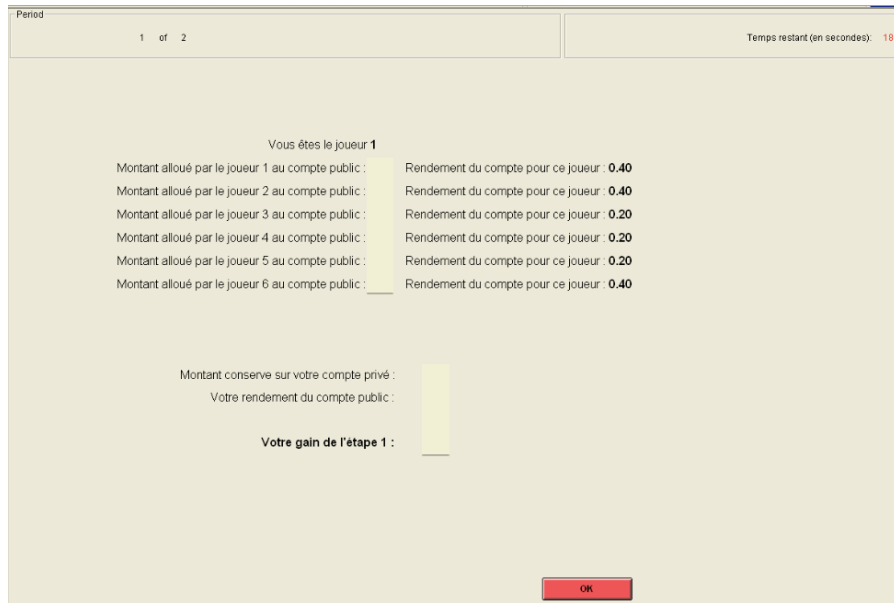


Figure 3. Example of screen-shot with feedback at the end of the first stage

Your earnings for the first stage are calculated using a formula that differs based on whether you are a Type-A participant or a Type-B participant.

The earnings of a Type-A participant are calculated using the following formula (If you have any difficulties understanding these formulas do not hesitate to ask us.)

$$\text{Earnings at stage 1} = \text{Endowment of ECU} - \text{your allocation to the public account} + 0.4 * \text{Total allocation to the public account}$$

The earnings of a Type-B participant are calculated using the following formula:

$$\text{Earnings at stage 1} = \text{Endowment of ECU} - \text{your allocation to the public account} + 0.2 * \text{Total allocation to the public account}$$

This formula shows that your earnings at the end of the first stage consist of two parts:

- 1) The ECU which you have kept for yourself (endowment – allocation)
- 2) The income from the public account, which equals 40% of the group's total allocation if you are of Type A and 20% of group's total allocations if you are of Type B.

**Example:**

Suppose the sum of the allocations of all group members is 90 ECU.

In this case, a Type-A member of the group receives an income from the public account of:

$0.4 * 90 = 36$  ECU. A Type-B member of the group receives an income from the public account of:  $0.2 * 90 = 18$  ECU.

If the total allocation to the public account is 15 points, then Type-A members receive  $0.4 * 15 = 6$  ECU from

the public account while Type-B members receive  $0.2 * 15 = 3$  ECU from the public account.

You always have the option of keeping the ECU for yourself or allocating them to the public account. Each ECU that you keep raises your end of period income by 1 ECU. Supposing you allocated this point to the public account instead, then the total allocation to the public account would rise by 1 ECU. Your income from the public account would thus rise by  $0.4 * 1 = 0.4$  ECU if you are a Type-A member and  $0.2 * 1 = 0.2$  ECU if you are a Type-B member. However, the income of the other group members would also rise by 0.4 or 0.2 ECU each, so that the total income of the group from the public account would be increased by  $3 * 0.4 + 3 * 0.2 = 1.8$  ECU. Your allocation to the public account therefore also raises the income of the other group members. On the other hand, you also earn an income for each ECU allocated by the other members to the public account. In particular, for each ECU allocated by any member you earn either 0.2 or 0.4 ECU depending on your type.

### Communication

In certain periods, you will have the possibility to exchange messages with the other members of your group, **before choosing the number of ECU** you would like to allocate to the public account.

A dialogue box will appear on the screen in which you will be able to type your messages, as shown in Figure 4. In order to send your message, press the “ENTER” key.

During your messages exchange with the others, please follow two simple rules: 1) remain respectful and do not write rude words, 2) do not identify yourself.

All other group members will be able to read your messages. The id of the participant *in the current period* will appear next to each message.



Figure 4. Example of screen-shot with communication

The duration of the communication varies according to the period, as indicated below:

Period 1: 5 minutes

Period 6: 5 minutes

Period 11: 3 minutes

Period 16: 3 minutes.

During this time, you can send as many messages as you want. During other periods you will not be able to exchange messages.

### Second Stage

In the second stage of each period, you will be informed about how much each group member of each Type allocated individually to the public account in the first stage. In this stage you can **increase or leave unchanged the earnings of each member of your group by distributing addition points**. The other group members can also increase your income if they wish to.

You will have to decide how many addition points to assign to each of the other five group members. You must enter a number for each of them. If you do not wish to change the earnings of a specific group member, then you must enter 0. You will be invited to confirm your decisions by entering again these numbers.

Figure 5 below gives an example of a screenshot of decisions in stage 2.

Joueur	Rendement du compte public	Allocation au compte public	Gains de la 1ère étape	Points d'addition que vous attribuez	Confirmez ces points d'addition
Vous	0.20				
Joueur 2	0.20			<input type="text"/>	<input type="text"/>
Joueur 3	0.40			<input type="text"/>	<input type="text"/>
Joueur 4	0.20			<input type="text"/>	<input type="text"/>
Joueur 5	0.20			<input type="text"/>	<input type="text"/>
Joueur 6	0.40			<input type="text"/>	<input type="text"/>

Points disponibles : 20

OK

Figure 5. Example of a screenshot of decisions in stage 2

Assigning addition points is costly. Every addition point you assign costs you 0.50 ECU. **For example**, if you assign 2 addition points to another member, this costs you 1 ECU; if, in addition, you assign 4 addition points to another member this costs you an additional 2 ECU. In total for this example you have assigned 6 points and your total costs therefore amount to 3 ECU.

Costs of assigned addition points = 0.50 ECU x total number of addition points you have assigned

You can assign a maximum of 20 points in total to the other group members. Within this limit, you may distribute as many points as you wish to a given participant. However, the gain of the total number of points assigned by all group members to a given participant cannot exceed that participant's endowment of 20

ECU.

If you assign 0 addition point to a particular group member, you will not modify his or her earnings. If you assign 1 addition point to a group member, you will increase the earnings of this group member by 1.5 ECU. If you assign a group member 2 addition points, you will increase the group member's earnings by 3 ECU, and so on. Each addition point that you assign to another group member will increase his or her earnings by 1.5 ECU. Similarly, each addition point assigned to you by another group member will increase your allocation stage earnings by 1.5 ECU.

$$\text{Earnings of received addition points} = 1.5 \text{ ECU} \times \text{total number of addition points you have received}$$

Your total earnings from the two stages are therefore calculated as follows:

$$\text{Total earnings (in ECU) at the end of the second stage} = \text{Earnings from the 1st stage} + 1.5 * (\text{total number of addition points you have received}) - 0.50 * (\text{total number of addition points you have assigned})$$

If the value of points that you receive exceeds your initial endowment of 20 ECU, the points in excess assigned by the other participants cannot increase your earnings any more. **The value of the points exceeding your initial endowment will not be counted in determining your earnings from the period.**

After all participants have made their decision, the summary of decisions from the period as well as the associated payoffs will be displayed on a screen such as the one shown in figure 6.



Figure 6. Example of screenshot with information at the end of the period

After you have viewed the earnings screen the period is over and the next period starts automatically. Please read these instructions again. If you have any further questions please raise your hand and one of the experimenters will come to help you. Otherwise, please answer the control questions.



## Control Questions

Please answer the following questions:

1. Each group member has an endowment of 20 ECU. If nobody (including you) allocates any ECU to the public account. What are:
  - a. The earnings of Type-A participants after the first stage? .....
  - b. The earnings of Type-B participants after the first stage? .....
  
2. Each group member has an endowment of 20 ECU. You allocate 20 ECU to the public account. All other group members allocate 20 ECU each to the public account. Suppose that you are a participant of Type-A. What are:
  - a. Your earnings after the first stage? .....
  - b. The earnings of the other Type-A group members?.....
  - c. The earnings of the Type-B group members? .....
  
3. Each group member has an endowment of 20 ECU. Suppose you and the other Type-A participants allocate 20 ECU each to the public account and the three Type-B participants allocate 5 ECU each to the public account. What are:
  - a. Your earnings after the first stage? .....
  - b. The earnings of the other Type-A group members?.....
  - c. The earnings of the Type-B group members? .....
  
4. Each group member has an endowment of 20 ECU. The other five group members allocate together a total of 30 ECU to the public account. Suppose that you are a participant of Type-A. What are:
  - a. Your earnings after the first stage if you allocate 0 ECU to the public account? .....
  - b. Your earnings at the end of the first stage if you allocate 15 ECU to the public account? .....
  
5. Each group member has an endowment of 20 ECU. You allocate 8 ECU to the public account. Suppose that you are a participant of Type-B. What are:
  - a. Your earnings after the first stage if the other group members together allocate a further total of 7 ECU to the public account?.....
  - b. Your earnings at the end of the first stage if the other group members together allocate a further total of 22 ECU to the public account?.....
  
6. Your earnings from the first stage are 22 ECU. How much will your earnings at the end of the second stage be if:
  - a. You receive 2 points, but do not assign any yourself? .....
  - b. You receive 2 points and assign 4 points to a group member?.....
  
7. Please indicate whether the following statements are true or false:
  - a. Each participant keeps the same type (A or B) throughout the 20 periods.  
 True             False
  - b. Each participant keeps the same id number (1,2,3,4,5 or 6) throughout the 20 periods.  
 True             False

## APPENDIX 2 – Comparison across the two variants of communication in HET-C

In Figures A1 and A2, we pool the data of the two variants of the HET-C treatment: in one variant subjects could only send messages to everyone (HET-C Public); in the second variant, subjects could send both messages to everyone and ‘private messages’ to the group members with the same return as them (HET-C Public-Private). As can be seen, the aggregate patterns across the two conditions are remarkably similar.

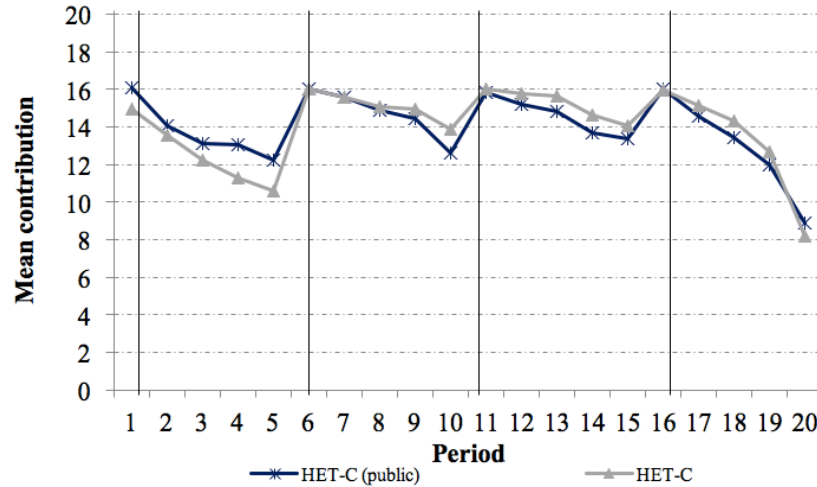


Figure A1 – Mean contributions across treatments

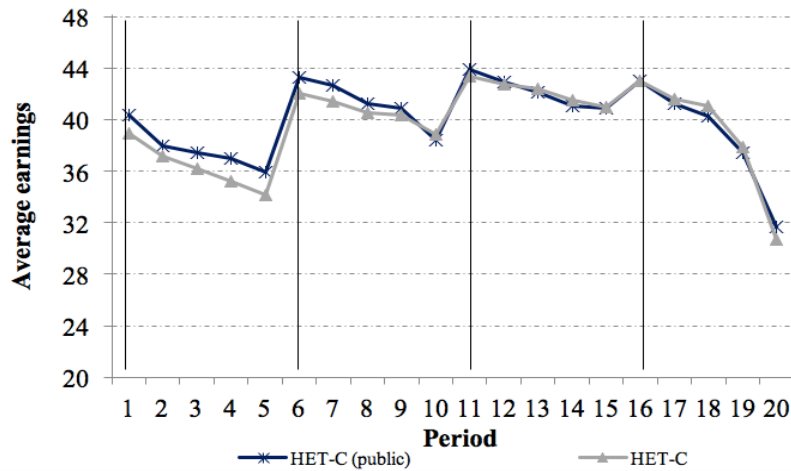
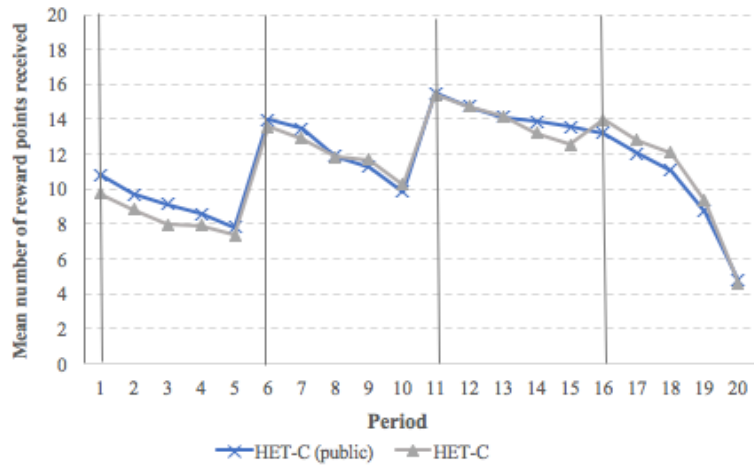


Figure A2 – Mean earnings across treatments



**Figure A3** – Mean number of reward points received across treatments

In Table A1, we compare the mean contributions and mean final earnings in the two variants of the HET-C treatment. Non-parametric tests indicate minor difference between these two variants in contribution behavior which are far from being statistically significant. The only significant difference is found for the low return subjects' total earnings.

**Table A1** - Summary statistics on contributions and final earnings in the two variants of the HET-C treatment

Treatments	All subjects	High-return subjects	Low-return subjects	Ratio H/L
<i>Mean contributions</i>				
HET-C Public (1)	13.98 (7.66)	17.57 (6.28)	10.40 (7.23)**	1.69
HET-C Public-Private (2)	14.06 (7.30)	16.94 (6.07)	11.17 (7.29)**	1.52
(1) vs. (2)	$p=0.847$	$p=0.335$	$p=0.847$	
<i>Mean final earnings</i>				
HET-C Public (3)	39.99 (10.16)	44.04 (10.94)	35.95 (7.37)**	1.22
HET-C Public-Private (4)	38.98 (11.94)	45.36 (11.04)	32.61 (9.06)**	1.39
(3) vs. (4)	$p=0.630$	$p=0.700$	$p=0.043$	

*Note:* Standard deviations in parentheses. Comparisons between conditions ((1) vs. (2) and (3) vs. (4)) report the p-values of Mann-Whitney tests in which each group averaged over all periods gives one independent observation (N=9 in HET-C Public and N=8 in HET-C Public-Private). In the column “Low-return subjects”, \*\* indicate significance at the 5% level in Wilcoxon tests comparing contributions and earnings relative to low-return vs. high-return subjects.

## APPENDIX 3 – Analysis of communication in HET-C

### 1. Methodology

We asked three independent coders who were blind to our research hypotheses to code the content of messages exchanged between subjects in treatment HET-C. Then, we requested two additional coders to check the consistency between the initial classification of messages. We present the results from this exercise in this appendix. In summary, this analysis supports the conclusions concerning group contribution norms in Section 3.2.

Using the communication data, we created several variables for each communication period and each group. We provided the list of variables to the coders and this is this codebook that the coders used to enter the data. The variables contain the following information:

- total number of messages, number of messages from high-return subjects, number of messages from low-return subjects, number of public messages, number of public messages from high-return subjects, number of public messages from low-return subjects, number of private messages, number of private messages from high-return subjects, number of private messages from low-return subjects
- whether group members talk about both stages, only about stage 1, only about stage 2, about nothing related to the experiment
- regarding stage 1, whether the group talks about: a norm of equal earnings, a norm of proportional contributions, a norm of equal contribution, maximization of the surplus, no norm at all, several possible norms
- regarding stage 2, whether the group talks about: maximization of the number of rewards assigned, assigning the same rewards to every group member, using rewards to compensate the low-return group members, assigning rewards only to in-groups with the same type
- whether the messages include the notions of: sanction, trust
- whether the messages relate to the calculation of the number of rewards to assign
- for the private messages: messages confirming the public messages, messages reexplaining the strategy or the calculation of payoffs, messages expressing some distrust of out-groups or preference for in-groups.

### 2. Communication on the assignment of rewards

In 44 cases out of 68 (65%), groups explicitly agree on a rule regarding at least one stage.

Regarding the content of messages, people discuss about the calculation of rewards in 34 cases (periods\*groups) out of 68 (17 groups, 4 communication periods), which amounts to 50% of the cases. More precisely, they discuss how to avoid buying too many reward points or about how to take the cap into account to maximize payoffs in 26 cases (38%). In 52 cases (76%), they mention the same number of reward points regardless of the group members' type. In 14 cases (20%), they discuss using rewards to compensate the low-return group members. Subjects do not plan explicitly to reward those adhering to a rule that was agreed upon. But sometimes, groups evoke the possibility to sanction those who do not contribute as much as decided earlier by reducing the number of reward points in the next periods (18 cases out of 68, 26%).

We found only one case in which group members explicitly mention reward points as a way to reward high contributors (reciprocity), in the private chat discussion, but this is combined with discrimination based on types (Group 2, period 11: “- I propose to give (*reward points*) to all of those who contribute 20 and we keep the rest for the other two of us. (...) – OK, so 4 to us and 2 to the B (*low-return*) members who

contribute 20”). Overall, they rarely talk of differentiating the assignment of rewards by type of players. They never evoke rewards as a means to encourage other group members to increase their contribution in the next periods (incentives).

### 3. Private and public communication

All nine groups that were offered the possibility to send both public messages (visible to all group members, regardless of their type) and private messages (only to the other members with the same type) use the opportunity to send private messages: there are only 4 periods in which a group did not use the private chat box. But this opportunity does not mean that group members do not send public messages at the same time. On average, among the 54.97 messages sent by a group in a period, 34.78 are public and 20.19 are private. In comparison, in the sessions in which communication was only public, groups send on average 43.92 messages per period. The private messages are mainly sent by the high-return members. The high-return members send on average 18.09 public messages and 16.03 private messages per group and period (to be compared with 23.17 messages when communication is only public). These numbers are respectively 16.69 and 4.16 for the low-return members per group and period (20.75 when communication is only public). The analysis of the content of the private messages indicate that in 34% of the cases (a case being defined as one period x one group) private communication simply confirms what has been decided in public communication. In 18% of the cases, subjects reexplain the strategy or calculations. In 56% of the cases, subjects express a willingness to give more rewards to their in-groups with the same type or only to those who fully contribute; they discuss how to persuade the low-return players to contribute more since it would increase the payoffs of the high-return players; or they express a willingness to ensure cooperation at least between players of the same type.

### 4. Initial round of communication, assignment of rewards and contributions

To test whether addition points are assigned to reward compliance with the initial norm, the following Table A2 reports the mean number of reward points received by the subjects in periods 1 to 5 (included), depending on their type (high- or low-return) and on whether they conform to the norm agreed upon during the first round of communication in the HET-C treatment. In this Table, we ignore the groups that failed to agree on either the efficiency-over-equality norm or the equality-over-efficiency norm.

**Table A2** – Communication in period 1 and rewarding behavior in periods 1 to 5

HET-C, periods 1-5	High-return subjects		Low-return subjects			
	$c_i=20$	$c_i<20$	$c_i=20$	$c_i<20$	$c_i=5$	$c_i<5$
Norm agreed upon in period 1						
Equality-over-efficiency (Groups 4, 9, 13,14)	16.29 (N=35)	1.44 (N=25)	- (N=0)	8.30 (N=60)	9.22 (N=54)	0 (N=6)
Efficiency-over- equality (Groups 3,6,7,10,12,16,17)	12.64 (N=69)	7.61 (N=36)	12.15 (N=60)	6.6 (N=45)	10.4 (N=5)	2.91 (N=11)

Notes: Table A2 reports the mean number of reward points received by the subjects in periods 1 to 5 (included), depending on their type (high- or low-return) and on whether they conform to the norm agreed upon during the first round of communication in the HET-C treatment.

The Table shows that full contribution by high-return subjects is not more rewarded in groups that agreed on efficiency-over-equality than in groups that agreed on equality-over-efficiency (12.64 vs. 16.29 points, respectively). Similarly, contributions of 5 ECU by low-return subjects do not trigger higher rewards

in groups that agreed on equality-over-efficiency compared to groups that agreed on efficiency-over-equality (9.22 and 10.4 points, respectively). Thus, there is no evidence that addition points are assigned to reward norm compliance.

This Table shows that in this first block of communication groups that favor efficiency over equality reward norm compliance by high-return subjects and by low-return subjects to the same extent (on average, 12.64 and 12.15 reward points, respectively). In contrast, groups that favor equality over efficiency reward norm compliance by high-return subjects more than norm compliance by low-return subjects (16.29 reward points on average for a contribution of 20 ECU by high-return subjects and 9.22 points for a contribution of 5 by the low-return subjects). The Table indicates that high-return subjects are more rewarded in groups that favor equality than in groups that favor efficiency: in these preliminary periods, points are assigned also to reward higher relative contributions. Table 4 in the main text indicates that this is no longer the case in the next blocks of periods.

Of course, these findings should be considered with care since they only apply to the first round of communication and the first five periods of contribution and reward, where a lot of learning is happening. The following sub-section considers the classification of groups based on median behaviour in periods 6 to 19.

## 5. Group-level analysis of communication within a period

Table A3 below complements Table 4 by reporting information on the content of messages for each group category, where each group in each period with communication provides one observation. Columns (1) to (3) report the percentage of observations relating to the efficient contribution rule, the equality of earnings rule, and other contribution rules, respectively. Column (4) indicates the percentage of observations in which conversations in a period evoke more than one contribution rule and column (5) reports the percentage of observations in which rewards are evoked as a solution to compensate low-return players from the inequality deriving from the first stage of the game. Columns (6) and (7) report the mean responses to a post-experiment survey about the appeal of different contribution rules.<sup>24</sup> Finally, columns (8) and (9) report the mean responses to the post-experiment survey about the player's motivation when assigning reward points.<sup>25</sup>

Table A3 indicates that the content of conversations differs strongly across groups. Column (1) indicates that the rule prescribing equal contributions is evoked in only 20.83% of the observations when groups behave according to the equality-over-efficiency rule, compared to 62.50% in groups behaving according to the efficiency-over-equality rule and 71.43% in groups following other rules. Conversely, column (2) shows that the rule of equal earnings is evoked in 79.17%, 6.25% and 17.86% of the corresponding observations. Table A2 also indicates that members of groups adopting the equality-over-efficiency contribution rule are significantly more likely than members of other groups to agree with the statement that group members should earn the same regardless of their type (column (6)); conversely, members of groups favoring the efficiency-over-equality rule are significantly more likely than members of

---

<sup>24</sup> For the equal-earnings rules, using a scale ranging from 1 ("I fully disagree") to 7 ("I fully agree"), we asked subjects to tell us the extent to which they agree with the statement "Ideally, all group members should earn more or less the same from the experiment irrespective of whether they are of type A [i.e., high-return] or B [i.e., low-return]". For the equal-contributions rule (which is consistent with efficiency), we asked them to tell us the extent to which they agree with the statement "Ideally, all group members should allocate the same number of ECU to the public account irrespective of whether they are of type A or B and thus earn different amounts".

<sup>25</sup> Using also a scale ranging from 1 ("I fully disagree") to 7 ("I fully agree"), subjects who assigned points had to assess the following two statements: "I assigned points to ensure that earnings of all players will be more or less the same" and "I assigned points to reward those who allocated more ECU to the public account than the others".

other groups to agree with the statement that group members should contribute the same regardless of their type (column (7)).

The entries in column (5) indicate that discussions about using rewards as a mean for compensating low-return players for the inequality deriving from the first stage of the game occur very rarely, regardless of the contribution rule adopted by the groups. Most discussions about rewards focus on how many points to assign. The post-experiment survey is also informative. Players from groups who prioritize equality are significantly more likely than the players from the other group categories to report that they assigned points to ensure that earnings of all players would be more or less the same (column (8)). A Wilcoxon test indicates that these players are also significantly more likely to report this reason for assigning points than the willingness to reward those who allocated more ECU to the public account than the others.



**Table A3** – Group-level analysis of communication within a period

Categories of groups	Rules discussed in a period			Multiple rules discussed in a period (4)	Rewards as compensation in a period (5)	Post-experimental questionnaire (scale 1 to 7)			
	Efficient contribution (1)	Equal earnings (2)	Other (3)			Members should earn the same (6)	Members should contribute the same (7)	Rewards to equalize earnings (8)	Rewards to reciprocate high contrib. (9)
<b>HET-C</b>									
Equality over efficiency N=6 (A) (4 <sup>#</sup> ,5 <sup>#</sup> ,9,11,13,14)	20.83% (41.49)	79.17% ** (41.49)	16.67% <sup>ns</sup> (38.07)	29.17% (46.43)	8.33% (28.23)	6.06 (1.57)	2.61 ** (2.05)	5.94 (1.71)	2.67 ** (2.27)
Efficiency over equality N=4 (B) (3 <sup>#</sup> ,7 <sup>#</sup> ,10,17)	62.50% (50)	6.25% * (25)	6.25% * (25)	12.50% (34.16)	12.50% (34.15)	4.42 (2.16)	5.04 <sup>ns</sup> (2.27)	3.12 (2.23)	3.25 <sup>ns</sup> (2.23)
Other groups N=7(C) (1 <sup>#</sup> ,2 <sup>#</sup> ,6 <sup>#</sup> ,8 <sup>#</sup> ,12,15,16)	71.43% (46)	17.86% ** (39)	39.29% ** (49.73)	39.29% (49.73)	35.71% (48.80)	4.67 (2.07)	4.09 <sup>ns</sup> (2.29)	4.02 (2.12)	4.07 <sup>ns</sup> (2.13)
(A) vs. (B)	<i>p</i> =0.006	<i>p</i> =0.008	<i>p</i> =0.598	<i>p</i> =0.165	<i>p</i> =0.617	<i>p</i> =0.024	<i>p</i> =0.010	<i>p</i> =0.010	<i>p</i> =0.198
(A) vs. (C)	<i>p</i> =0.004	<i>p</i> =0.007	<i>p</i> =0.169	<i>p</i> =0.544	<i>p</i> =0.414	<i>p</i> =0.018	<i>p</i> =0.012	<i>p</i> =0.012	<i>p</i> =0.045
(B) vs. (C)	<i>p</i> =0.366	<i>p</i> =0.506	<i>p</i> =0.090	<i>p</i> =0.108	<i>p</i> =0.679	<i>p</i> =0.703	<i>p</i> =0.058	<i>p</i> =0.254	<i>p</i> =0.086
<b>HET-NC</b>									
	-	-	-	-	-	4.05 (1.87)	3.70 (2.17)	2.90 (2.03)	5.20 (2.06)

Notes: Columns (1) to (5) report the content of messages where each group in each period with communication gives one observation. Columns (6) to (9) report the mean responses to a post-experiment survey using a scale ranging from 1 (“I fully disagree”) to 7 (“I fully agree”). Comparisons between categories of groups ((A) vs. (B), (A) vs. (C), and (B) vs. (C)) report the p-values of Mann-Whitney tests in which each group gives one independent observation. # indicates the groups in which subjects could use two chat boxes to send messages visible to everyone in their group and private messages to their group members with the same identity. \*\* and \* indicate significance at the 5% and 10% levels, respectively and *ns* is for no significance in Wilcoxon tests comparing (2) and (3) to the reference (1), comparing (7) to reference (6) and comparing (9) to reference (8). All tests are two-tailed.