CrossMark

The trade-off between welfare and equality in a public good experiment

Agathe Rouaix[1] · **Charles Figuières**[2] ·
Marc Willinger[3]

Abstract We report the results of an experiment on voluntary contributions to a public good in which we implement a redistribution of the group endowment among group members in a lump sum manner. We study the impact of redistribution on group contribution, on individuals' contributions according to their endowment and on welfare. Our experimental results show that welfare increases when equality is broken, as predicted by theory (Itaya et al. in, Econ Lett 57:289–296, 1997), because the larger contribution of the rich subjects overcompensates the lower contribution of the poor subjects. However, our data suggest that the adjustment of individual contributions after redistribution is not always compatible with the predictions. In particular, subjects who become poor contribute much less than subjects who were poor since the beginning.

1 Introduction

Social justice is usually thought to be the key reason for income redistribution. However, from an economic point of view it is often the consideration of efficiency that matters. In many cases the pursuit of social justice is antagonistic with the efficiency objective: impoverishing the rich for increasing the income of the poor can be detrimental to overall welfare. Why such trade-off between equity and efficiency occurs

---

✉ Marc Willinger
willinger@lameta.univ-montp1.fr

1 Department of Economics, Business School, University of Aberdeen, AB24 3QY Scotland, UK

2 INRA, LAMETA, UFR d'Economie, Avenue Raymond Dugrand, CS79606, 34 960 Montpellier Cedex 2, France

3 Université de Montpellier, LAMETA, UFR d'Economie, Avenue Raymond Dugrand, CS79606, 34 960 Montpellier Cedex 2, France

can be due to many reasons: for instance, there may be a risk of evasion of mobile resources, or a disincentive effect of taxes on skilled labor. A local government which undertakes a generous redistributive policy would attract the poor from the other regions—an expenditure externality (Wildasin 1991)—and the required tax increase would drive the rich away to other regions—a tax externality. Overall, because of these externalities there is a loss of efficiency associated with the pursuit of fairness. Not surprisingly, the reasons mentioned above are echoed in a vast theoretical literature concentrating on optimal taxation (e.g. Mirrlees 1971; Atkinson and Stiglitz 1976; Piketty 1993), and in an equally extensive empirical literature on the issue of efficient income redistribution (e.g., Gardner 1983; Bullock 1995; Brueckner 2008; Dahlberg and Edmark 2008)[1].

In this paper we focus on yet another reason for the existence of an efficiency-equity trade-off, which is largely under-documented and rests on the impact of a redistribution of incomes on the voluntary provision of public goods. Imagine a group of individuals who differ with respect to their propensity to contribute to the provision of a public good. A redistribution of the group income will therefore generally affect the amounts of voluntarily provided public goods by such a group. If one further assumes that rich individuals have a higher propensity to contribute to the provision of public goods than poor individuals, any redistribution from rich to poor will lower the amounts of voluntarily provided public goods while a redistribution in the opposite direction, which enriches the rich, will increase them. In other words, increasing the income of the rich at the expense of the poor's income may increase the group's welfare, and may even in some cases lead to a Pareto-improvement, i.e. increasing both the utility *of the rich and the poor*.

Itaya et al. (1997, IdMM hereafter) showed that even if the income distribution is unequal, the larger contributions to the public good by the rich tends to equalize utilities ex post. Therefore if the focus is on individual utility rather than on income *per se*, there seems to be no role for equalizing incomes. In contrast, the authors also proved that creating more pronounced inequalities of income can increase social welfare when some agents cease to contribute. The conclusion is that state interventions to increase income inequalities, in IdMM's environment, may be desirable.

The analysis of IdMM belongs to a stream of literature that examines the policy implications of the famous *neutrality theorem* of income redistribution, first presented by Warr (1983) and developed by Bergstrom et al. (1986, henceforth BBV). We are interested in the violation of the conditions of this theorem and the ensuing *non-*

---

[1] It is worth noting that the inefficiency caused by redistribution, because of the discouraging effects of taxes on the effort of skilled workers, or the erosion of the tax base due to the mobility of factors, is quite robust to assumptions about the information possessed by the agents. Whether or not the characteristics of redistribution—who receives, who gives and how much?—are known does not remove the efficiency-equity trade-off. Often, information on the redistribution is not public. For example, Mirrlees (1971) and Diamond and Mirrlees (1971a, b), offer general equilibrium models: agents take prices (and in particular taxes) as exogenous variables. This is the same assumption in Atkinson and Stiglitz (1976). The formal framework of Piketty (1993) is a bit different: the statistical distribution of productivity parameters is common knowledge, and the outcome of social interactions is conceived as a Bayesian Nash equilibrium. Finally, in Wildasin (1991), behaviors are predicted using the concept of Nash equilibrium, presumably with complete and perfect information.

neutrality, i.e. situations where redistribution alters agents' well-being[2]. In particular, the set of contributors and the amount of public good are supposed to be affected by the income redistribution. According to IdMM the amount of public good provided increases (respectively decreases) when income inequalities increase (decrease). If richer individuals are willing to contribute larger amounts to the provision of public goods than poorer individuals, any redistribution from rich to poor may deteriorate welfare even though the increased private consumption by the poor compensates the reduced private consumption of the rich in utility terms. In contrast, redistribution from the poor to the rich might exactly have the opposite effect: by increasing the amount of public goods available through increased spending by the rich, it may simultaneously increase the welfare of the poor and of the rich.

IdMM's finding, that an unequalizing redistribution leads to an increase of society's welfare, is a provocative ethical conclusion. It should be taken with caution, for it potentially justifies policies that increase income inequalities, and gives arguments to people who defend privileges, on the basis that such policies would be more efficient for various kinds of public works or patronage. In addition, under the assumption of standard (selfish-oriented) preferences, redistribution from rich to poor does not necessarily lead to a Pareto-improvement, although it can often lead to a welfare-improvement but at the social cost of a higher level of inequality. The effectiveness of unequalizing redistribution policies for increasing welfare remains an empirical issue. There are many reasons, including unpopularity, why they cannot be implemented, at least in democratic economies. Several attempts have been made to empirically test IdMM's conjecture, especially in the case of the provision of household public goods [see e.g., Yamada and Naito 2014]. However these studies are limited and suffer from the usual critique that many factors can explain the observed variability in contributions, not only income heterogeneity. Furthermore, even if natural data about changes in income distribution is sometimes available, the counterfactual is generally missing. For these reasons we chose to investigate the issue of redistribution and voluntary provision of public goods experimentally. The strong advantage of economic experiments is that the experimenter can control for redistribution within a group of individuals and observe their decisions.

Our experimental setting follows IdMM's theoretical framework. In contrast to previous experiments that studied the impact of income inequality on voluntary contributions (Chan et al. 1996, 1999; Cherry et al. 2005; Hofmeyr et al. 2007; Anderson et al. 2008; Georgantzís and Proestakis 2011) we run a public good game experiment in which we implement a reallocation of the group endowment among group members: after a first sequence of rounds in which endowment equality prevails, we

---

[2] The neutrality result is the property that, after an income redistribution, donors cut their contribution to the public good by the same amount as their budget reduction, while benefactors contribute their whole income increment to the public good. As a result: (i) the aggregate level of public good remains unchanged, (ii) private consumptions are the same as before and, (iii) utility levels are kept constant. The remarkable and counter-intuitive property about those offsetting changes is that they are unilateral best decisions or, put differently, constitute equilibrium reactions. This property holds as long as redistribution is small enough to allow each agent to maintain the level of consumption of the private good he enjoyed before redistribution. But if redistribution is too large, in the sense that for some agents consuming the private good as before becomes incompatible with their new income, individual decisions are necessarily modified.

implement an unequalizing redistribution at the beginning of the second sequence of rounds. Symmetrically, we also study the case of an equalizing redistribution: after an initial sequence in which an unequal endowment distribution prevails, we implement an equalizing distribution before the second sequence. These two test-treatments are compared to baseline treatments for which there is no change in endowment distribution between the two sequences of rounds. Our within-subjects setting allows us both to study, as in previous literature, the impact of endowment heterogeneity on group contributions by comparing the contributions in the first sequences across treatments, and to isolate the impact of redistribution (unequalizing or equalizing) on contributions. In addition, we rely on a partner design, because we want to implement a real redistribution of group endowment. To that end we need at least two rounds, which differ with respect to the distribution of group endowment. However, more rounds are typically desirable because, as illustrated by many previous public goods experiments, the dynamics of group interactions matters and generates heterogeneous group outcomes.

We chose the set of parameters in a way to allow to test the following predictions: (1) poor agents do not contribute any token to the public good after redistribution, (2) social welfare (i.e. group payoff) increases after an unequalizing redistribution, and (3) an equalizing redistribution generates a Pareto degradation (i.e. both poor and rich have a lower payoff).

While at the group level our experimental results are consistent with predictions (2) and (3), at the individual level we observe that subjects' behavior contradicts the Nash predictions. First, poor subjects contribute to the public good although they should not, in violation of prediction (1). Second, subjects who become poor after redistribution contribute less than those who were poor at the outset: *being* poor is not the same thing as *becoming* poor. Such behavioral asymmetry is specific to poor. It is not observed for the rich who contribute always the same amount whether they became rich after redistribution or whether they were rich at the outset. Finally, the fact that subjects who became poor after redistribution contribute less than subjects who were initially poor affects negatively group performance.

Section 2 offers a brief overview of previous public good experiments that addressed the issue of inequality of endowments or the neutrality issue. Section 3 presents our experimental design. The results are presented in Sect. 4. Section 5 concludes.

2 Previous experiments

The *"Equity versus Efficiency"* trade-off has been investigated in various types of experiments: gift exchanges games (Güth et al. 2003), generosity games (Güth et al. 2010), elicitation of preferences for redistribution (Durante et al. 2014; Alesina and Giuliano 2009). A common findings of this literature is that subjects have a lower taste for equity when the price of equity increases (see e.g. Andreoni and Miller 2002).

In public good games with endowment heterogeneity the issue was addressed indirectly through the following question: "Does endowment heterogeneity lead to lower or higher provision of public goods, or is endowment distribution irrelevant?" The question was initially studied in the case of step level public good games. The main finding is that endowment heterogeneity is detrimental for efficiency: groups with heterogeneous endowments are less successful in reaching the provision point than

groups with equal endowments (Bagnoli and McKee 1991; Rapoport and Suleiman 1993) despite rich and poor subjects tend to contribute the same proportion of their endowment (van Dijk and Grodzka 1992; Rapoport and Suleiman 1993; van Dijk and Wilke 1995; de Cremer 2007).

In the case of linear public goods games, Zelmer (2003) showed in a meta-analysis of 27 experimental studies, that endowment heterogeneity had a significantly negative impact on contributions. However, according to more recent studies the impact of endowment heterogeneity on the level of public good provision is ambiguous: Cherry et al. (2005) found that endowment heterogeneity decreases the amount of public good while Hofmeyr et al. (2007) found that it had no impact on the level of provision. The existing experimental literature on linear public goods does not give a clear answer to the equity versus efficiency trade-off with respect to public good provision. However, Georgantzís and Proestakis (2011) showed that the way inequality is defined matters. In their experiment they constituted groups of players according to participants' real disposal income. In contrast to previous experiments on public goods they observed that group heterogeneity increased the level of contributions, i.e. there is an equity versus efficiency trade-off. Participants who are both "rich" inside and outside the lab contribute the highest fraction of their endowment while those who are both "poor" inside and outside the lab contribute the lowest fraction.

Experiments based on non-linear public good games (e.g. Chan et al. 1996, 1999) provide also evidence about the relevance of the trade-off. Both papers relied on a non-linear public good game for investigating the predictions of BBV's theory about the effect of income redistribution within a group of contributors. In their seminal theoretical paper BBV showed that a "small" unequalizing redistribution is neutral with respect to the amount of voluntarily provided public good while a "large" unequalizing redistribution increases the amount of public good. The experimental findings by Chan et al. (1996) do not reject this prediction: groups with small endowment heterogeneity provide the same amount of public good than groups with equal endowments but groups with large endowment disparity provide a greater amount of public good, in clear contradiction with the equity versus efficiency trade-off. The neutrality prediction, i.e. a small unequalizing or equalizing redistribution has no effect on group contribution - was also reported by Maurice et al. (2013) who implemented a real income redistribution in their experiment. Subjects played a public good game in two consecutive sequences of 10 rounds, with two different endowment distributions: equal versus unequal (or unequal versus equal). They found that the amount of public good provided is not affected by an equalizing or unequalizing redistribution. In our experiment we rely on the same design than in Maurice et al. (2013) to address the non-neutrality issue raised by IdMM: welfare increases when inequalities increase because of the increased level of provision of the public good.

3 Experimental design

3.1 The contribution game

We rely on a quasi-linear quadratic utility function, as in Keser (1996) or Bracht et al. (2008), in order to derive an interior Nash equilibrium solution for individual

contributions. The parametric setting is given by:

$$u_i (x_i, G) = 41 (w_i - g_i) - (w_i - g_i)^2 + 15G,$$

where w_i is subject's i endowment, $x_i = w_i - g_i$ is his consumption of the private good, g_i stands for his contribution to the public good and $G = \sum_{i=1}^n g_i$ is the amount of public good. The number of subjects in a group is $n = 4$. Under standard behavioral assumptions, i.e. fully rational and self-centered agents, the dominant strategy for subject i is to contribute $g_i = \max(w_i - 13; 0)$. In our game, subjects can be endowed with either 5, 20 or 35 tokens. Hence according to the standard prediction subjects endowed with 5 tokens contribute $g_i^* = 0$, subjects with an endowment of 20 tokens choose $g_i^* = 7$, and subjects endowed with 35 tokens contribute $g_i^* = 22$.

Consider now the case where the 4 members of a group have the same endowment $w_i = 20, i = 1, ..., 4$. At equilibrium, the amount of public good is $G_E = 4 \times 7 = 28$, where the subscript E refers to an "equal distribution of endowments". In contrast consider the case where two group members are endowed with 35 tokens and the other two with 5 tokens. Note that the group's total endowment is unchanged. The amount of public good is now $G_U = (2 \times 22) + (2 \times 0) = 44$, where the subscript U refers to the "unequal distribution of endowments". Compared to the egalitarian case, the situation is as if the agents who became rich had increased their contribution by an amount that is equal to the quantity of private good that is no longer consumed by the agents who became poor.

In the inegalitarian situation the set of contributors is restricted to the rich agents, since poor agents no longer contribute. This particular setting allows to test some of the statements of theorems 4 and 5 of BBV[3]. Note that despite the *ex ante* income inequality among group members, there is an *ex post* redistribution through the provision of the public good by the rich. The final wealth of the poor is therefore equal to 840 units: $180(41 \times 5 - 5^2)$, from their private consumption and $660(44 \times 15)$ from the public good. The final wealth of the rich is 1024 units: $364(41 \times 13 - 13^2)$, from their private consumption and $660(44 \times 15)$ from the public good. The aggregate wealth is therefore larger in the inequality situation (3728 units) than in the equality situation (3136)[4]. This property is in accordance with the predictions of IdMM (1997): although they considered the case of 2 agents, they contend that their predictions can

---

[3] Theorem 4 in BBV contains comparative statics results. Those that are of direct concern with our experiment are as follows: "(iii) If a redistribution of income among current contributors increases the equilibrium supply of the public good, then the set of contributing consumers after the redistribution must be a proper subset of the original set of contributors. (iv) Any simple transfer of income from another consumer to a currently contributing consumer will either increase or leave constant the equilibrium supply of the public good." Theorem 5 presents extensions if preferences are identical, as it will be the case in our experiment. It offers in particular the following statements: "(i) All contributors will have greater wealth than all non-contributors. (ii) All contributors will consume the same amount of the private good as well as of the public good. (iii) An equalizing wealth redistribution will never increase the voluntary equilibrium supply of the public good. (v) Equalizing income redistributions that involve any transfers from contributors to non-contributors will decrease the equilibrium supply of the public good."

[4] That allows us to test experimentaly one of the statements in BBV: (i) *Any change in the wealth distribution that leave unchanged the aggregate wealth of current contributors will either increase or leave unchanged the equilibrium supply of public good.*

easily be extended to a larger number of agents. Note that with our specification the unequalizing redistribution is not only welfare-improving but is also Pareto-improving since each agent has a larger utility level than under equality[5].

3.2 Treatments

Our experimental setting is designed to allow within comparisons of group contributions when the distribution of endowments is altered while keeping the aggregate group endowment unchanged. Therefore each of our four treatments consists of two sequences of 10 periods each. In each period the group endowment is 80 tokens. Sequences have either an *equal* or an *unequal* distribution of endowment among subjects. We shall use the unambiguous notation E and U to refer to them. In E-sequences each group member has the same endowment (20 tokens) at the beginning of each period. In an U-sequence there are two poor group members with an endowment of 5 tokens and two rich group members with an endowment of 35 tokens. Endowment distribution is common knowledge but anonymity is maintained in U-sequences. Our two test treatments are noted EU, an equal sequence followed by an unequal sequence, and UE, an unequal sequence followed by an equal sequence. This design allows us to make both within comparisons by comparing the contributions of each group in the first sequence and in the second sequence, and between comparisons by comparing the first sequence of EU to the first sequence of UE. By moving from sequence 1 to sequence 2 two changes arise: first the endowment distribution is altered and second the repeated contribution game re-starts from the beginning. Restarting a new sequence can affect contributions. Andreoni (1988) and Croson (1996) observed for linear public goods games that an unexpected restart of a new sequence in fixed groups increases sharply contributions in the beginning of the new sequence. Since such an effect may also be present in our data we need to control for it. We therefore add to our design, two control treatments, EE and UU, called baseline treatments thereafter. The EE treatment consists of two E-sequences and the UU treatment of two U-sequences. An increase in group contributions in the early periods of the second sequence of the EE treatment can only be due to the restart effect, and similarly for UU. In contrast in the EU and UE treatments, a restart effect might be mixed up with a redistribution effect. By comparing the second sequence of the control treatment to the second sequence of the test treatment, both having the same first sequence, the restart effect, if any, is wiped away by taking the difference of the group contributions. After controlling for the restart effect, the EU treatment allows us to study the effect of an unequalizing redistribution, while the UE treatment allows to study the effects of an equalizing redistribution.

Table 1 summarizes the features of the control treatments and Table 2 does the same for the test treatments.

---

[5] This reasoning relies on agents who have standard preferences. Things would become more complicated with preferences featuring an aversion to inequality.

Table 1 Baseline treatments

Baseline treatments	EE treatment	UU treatment
Sequence 1		
Endowments	80 = (20, 20, 20, 20)	80 = (5, 5, 35, 35)
Nash-contribution	28 = (7, 7, 7, 7)	44 = (0, 0, 22, 22)
Numbers of contributors	4	2
Payoffs	3136 = (784, 784, 784, 784)	3728 = (840, 840, 1024, 1024)
Sequence 2		
Endowments	80 = (20, 20, 20, 20)	80 = (5, 5, 35, 35)
Nash-contribution	28 = (7, 7, 7, 7)	44 = (0, 0, 22, 22)
Numbers of contributors	4	2
Payoffs	3136 = (784, 784, 784, 784)	3728 = (840, 840, 1024, 1024)

Table 2 Test treatments

Test treatments	EU treatment	UE treatment
Sequence 1		
Endowments	80 = (20, 20, 20, 20)	80 = (5, 5, 35, 35)
Nash-contribution	28 = (7, 7, 7, 7)	44 = (0, 0, 22, 22)
Numbers of contributors	4	2
Payoffs	3136 = (784, 784, 784, 784)	3728 = (840, 840, 1024, 1024)
The redistribution is	Inequalizing	Equalizing
Sequence 2		
Endowments	80 = (5, 5, 35, 35)	80 = (20, 20, 20, 20)
Nash-contribution	44 = (0, 0, 22, 22)	28 = (7, 7, 7, 7)
Numbers of contributors	2	4
Payoffs	3728 = (840, 840, 1024, 1024)	3136 = (784, 784, 784, 784)

3.3 Practical procedures

The experiment was conducted at LEEM, the computerized laboratory of the University of Montpellier I, with the software z-Tree (Fischbacher 2007). We run 8 sessions involving 16 subjects each. The 128 subjects were randomly selected from a pool of student-subjects containing more than 5000 volunteers from the Universities of Montpellier. In each session groups of 4 anonymous participants were randomly formed and remained fixed for the whole session. The experiment consisted in 20 periods of play of the constituent game (with or without equal endowments) divided into 2 sequences. In each period subjects were asked to invest each of their tokens in a private account or in a public account. At the end of each period the following data were displayed on each subject's computer screen : his invested amount in each of the two accounts, the total contribution to the public account by the group, his earning from the private

account, his earning from the public account and his total earnings for that round. Furthermore, the record of previous periods was also on display.

Written instructions were provided for the first ten periods only. Subjects were unaware that a second sequence of 10 periods would be played after the first 10-periods sequence which was announced in the instructions. At the end of the tenth period, a new sequence of 10 periods was publicly announced. Subjects were given a new set of instructions, which emphasized the changes with respect to the first sequence, namely the new income distribution among the group members in the test treatments. Each independent group was endowed with 80 tokens at each period. The 80 tokens were split between the four subjects in an egalitarian way or in an unequal way and this distribution was common knowledge.

We chose not to announce the redistribution at the beginning of the experiment, in order to avoid uncontrolled effects that could have been generated by differing expectations across subjects about future endowment after the redistribution. If subjects are more or less optimistic (or pessimistic) about their future income, their contribution to the public good in the first sequence could have been affected. Subjects were paid according to their accumulated number of points in one of the two sequences, which was randomly chosen at the end of the session to be paid out for real. This procedure differs from other experiments that involved a restart of the game (Andreoni 1988; Croson 1996), but it is similar to the procedure implemented by Anderson et al. (2008).

4 Results

Let us recall that our main objective is to provide experimental evidence about how subjects solve the welfare-equality trade-off in a social dilemma context. Our secondary objective is to analyze subjects' reactions to the redistribution of their group endowment according to their status, i.e. becoming richer or poorer. Depending on their status (rich or poor) and depending on the direction of redistribution (unequalizing or equalizing), subjects may react differently from what theory predicts. A common finding from experiments on voluntary contributions is that subjects tend to over-contribute with respect to their Nash contribution. We therefore expect to observe also over-contribution in our experimental data. However, behavioral conjectures about variations of contributions can still hold even if all contribution levels are shifted upwards.

We rely both on non-parametric tests and on panel regressions. The significance level is set at 5 % for all of our statistical analyzes. Non-parametric tests are performed at the group level by taking, for each variable of interest, the average value of the four subjects of a group over the 10 periods of a sequence. For tests which distinguish between poor and rich subjects we refine these tests for each endowment category (rich and poor).

Before stating our main results, two preliminary sets of tests are necessary: (1) since we sometimes compare different treatments involving different subjects, it is first necessary to check whether subjects are randomly sampled from the same population; *i.e.* we need to check the homogeneity of our data; (2) as previously mentioned, we also need to control for the presence of an eventual restart effect between period

10 and 11. While these tests are useful for the sake of rigor, they are of secondary interest regarding the research questions. We therefore summarize them as follows: the null hypotheses of homogeneity of the data and absence of a restart effect cannot be rejected for the baseline treatments (see Appendix 1: Preliminary results). Because of the latter results, the effects of redistribution can be tested both across treatments (by comparing equal and unequal sequences of different treatments) and within treatments (by comparing the two sequences of the same treatment). Because the between tests which compare the second sequences which have the same first sequence provide consistent results, we mainly concentrate on the results of the "within" tests.

4.1 Is there a "welfare versus equality" trade-off?

Does a variation of endowment inequality affect social welfare? We answer this question by comparing the average group-payoff under equality and inequality of group members' endowments. First we check whether endowment inequality leads to higher or lower group payoff by analyzing our data at the group level. Second, we examine the payoffs of the rich and the poor separately to test whether individual well-being is modified as predicted in our game specification. For all sequences we rely on the average individual payoff for each group. Additionally, for U-sequences we also take into account the average payoff of the poor and of the rich separately.

Result 1 *An unequalizing (equalizing) redistribution of the group endowment increases (decreases) the group payoff.*

Support 1. In U-sequences subjects earn significantly more on average than in E-sequences. We reject the null hypothesis that payoffs are equal in both sequences ($p = 0.026$ for EU, and $p = 0.004$ for UE, Wilcoxon signed rank test, one-sided). Based on panel regressions for treatments EU and UE we find that the dependent variable *group payoff* is significantly and positively related to endowment inequality (the coefficient of the dummy variable *unequal* which is equal to 1 for inequality rounds has $p < 0.01$). The dummy variable that accounts for the ordering of the sequences does not affect the group payoff and the time variable has an insignificant negative impact on group contributions.

Under the assumption that players choose their Nash contribution, the predicted group payoff difference between sequence 2 and sequence 1 is equal to 148 points for EU (-148 for UE). The average payoff differences for treatments EU and UE are summarized in Table 3, both at the group level and for each player category (*rich* and *poor*). The null hypothesis is accepted at the group level: the average payoff-difference between the two sequences is equal to the predicted difference for the two treatments (see lines entitled "All" in Table 3). We perform the same tests for the baseline treatments. The results are reported in Appendix 2a (Payoffs, Table 9). The observed difference in contributions between Sequence 1 and Sequence 2 are not significantly different from zero. Finally, we test whether the observed payoff levels differ from the predicted ones (see Table 4). While in all sequences of all treatments we observe that group payoffs are larger than predicted, the null hypothesis of equality with the predicted level is rejected only in one case: in the first sequence of EU. We

Table 3 Payoffs difference between sequences 1 and 2 for test treatments

Treatment	Subjects	Sequence 1	Sequence 2	Observed difference	Predicted difference	H0 accepted[a]
EU	All	886.11	951.86	65.75	148	Yes (0.078)
	Poor	897.40	863.56	−33.84	56	Yes (0.055)
	Rich	874.81	1040.15	165.34	240	Yes (0.078)
UE	All	950.45	813.41	−137.04	−148	Yes (0.844)
	Poor	849.34	814.15	−35.19	−56	Yes (0.641)
	Rich	1051.56	872.67	−178.89	−240	Yes (1.000)

[a] Wilcoxon Mann–Whitney (*p*-value)

Table 4 Average payoffs for test treatments

Treatment	Subjects	Sequence	Observed payoffs	Nash payoffs	H0 accepted[a]
EU	All	1st	886.11	784	No (0.040)
		2nd	951.86	932	Yes (0.461)
	Poor	1st	897.40	–	–
		2nd	863.56	840	Yes (0.945)
	Rich	1st	874.81	–	–
		2nd	1040.15	1024	Yes (0.313)
UE	All	1st	950.45	932	Yes (0.547)
		2nd	813.41	784	Yes (0.313)
	Poor	1st	849.34	840	Yes (1.000)
		2nd	814.15	–	–
	Rich	1st	1051.56	1024	Yes (0.109)
		2nd	872.67	–	–

[a] Wilcoxon Mann–Whitney (*p*-value)

run the same tests for the baseline treatment (Appendix 2a (Payoffs), Table 10) and we reject the null hypothesis only in the second sequence if UU.

Taking group payoff as a measure of welfare, our data is consistent with IdMM's prediction. However, if we consider players' types (rich and poor) we observe that the difference in average payoff between equal and unequal sequences does not always have the expected sign for the poor. This is stated as result 2:

Result 2 *The welfare level of rich subjects is higher in U-sequences than in E-sequences. There is no welfare-change between sequences for poor subjects.*

Support 2. Rich subjects earn significantly more in the inegalitarian sequence of treatments EU and UE. We reject the null hypothesis of equal payoffs in E-sequences and U-sequences for the rich ($p = 0.004$, Wilcoxon signed-rank test, one-sided). On the other hand, the null hypothesis of payoff equality for the poor in E-sequences and U-sequences cannot be rejected ($p = 0.727$ for the EU treatment, $p = 0.098$ for the UE treatment, Wilcoxon signed-rank test). Further support is provided by panel

Table 5 Payoffs differences of poor and rich subjects in the same sequence for test treatments

Sequence[a]	Poor subjects	Rich subjects	Difference	Predicted difference	H0 accepted[b]
eU	863.56	1040.15	176.59	184	Yes (0.742)
Ue	849.34	1051.56	202.22	184	Yes (0.400)

[a] Wilcoxon Mann–Whitney (p-value)

[b] The uppercase represents the tested sequence

regressions on individual data. The dependent variable is the per period individual payoff. The coefficient of the dummy variable *unequal* is highly significant for the rich: the estimated increase of their payoff is equal to 202 points in U-sequences with respect to E-sequences. In contrast for the poor, the corresponding coefficient is insignificant. These effects are independent of the ordering of the sequences.

Note that the average payoff difference between the two sequences is equal to the predicted difference (see Table 3, in the lines entitled "rich" and "poor"[6]). Finally, even if both the average payoff of the rich and of the poor are larger than predicted, the difference with the prediction is not significant (see Table 4 which reports the test results).

As a consequence rich subjects earn more than poor subjects in all U-sequences irrespective of the treatment. In the U-sequences of EU and UE, the significant payoff difference between rich and poor ($p = 0.004$, Wilcoxon signed-rank test) corresponds to the predicted difference (see Table 5 above). Note that we observe the same difference in the UU treatment as reported in Table 11 (see Appendix 2a (Payoffs)).

Our main conclusion is that at the group level observed payoffs do not contradict IdMM's predictions, both in magnitude and in difference. However at the individual level, we observe that subjects who become richer in U-sequences get higher payoffs (as predicted) but this does not happen for subjects who become poorer in contradiction with the Nash prediction. This suggests that subjects' reactions to redistribution might differ from those predicted by BBV's theory. In the next section we further investigate this issue.

4.2 Contributions to the public good after redistribution

Variations of subjects' well-being after redistribution is due to the players' re-allocation of tokens between their private and their public accounts. The redistribution is predicted to have two effects: first, it expands or contracts the set of contributors, and second, it

---

[6] Note that for poor subjects there is no significant difference in payoffs before and after the redistribution. The observed differences between payoffs of both sequences do not differ from 56 (or -56), the predicted difference. We also compare the observed difference to zero, and we do not find any significant difference ($p = 0.641$ in the EU treatment and $p = 0.195$ in the UE treatment). Payoffs differences for poor subjects between both sequences of the test treatments are neither different from 0 nor from 56. Two reasons can explain this fact. First, there is a very small difference for the poor, in comparison to the absolute level of payoffs. In EU treatment, it represents an increase of 7.14 % of the payoffs of poor subjects. While for rich subject, the increase of payoff in the same treatment represents 28.58 % of their theoretical payoffs in the first sequence. It will be easier to observe the difference for rich subjects than for poor with so few data. Secondly the variability of our data is certainly too important to conclude about variations of payoffs of poor subjects.

increases or decreases the amount of voluntarily provided public good. In this section we investigate whether our data fits these predictions. We first analyze whether and how the set of contributors is affected by redistribution. Then, we compare the observed contributions to the Nash contributions, both at the group level and at the individual level. Finally, we compare the observed and the predicted variations of contributions for the rich and the poor subjects.

4.2.1 Set of contributors

Strictly speaking, a contributor is a subject who contributes at least one token to the public good. In our experimental setting, only agents endowed with an income strictly larger than 13 are expected to be contributors. Therefore the expected number of contributors is equal to four in E-sequences and to two in U-sequences. Sticking to this strict definition of contributors, for the three treatments with inequality sequences, there is only one group out of 24 which is consistent with the prediction about the set of contributors. All other groups have more than two contributors. Therefore, our data clearly rejects the predictions about the number of contributors in a group.

One might nevertheless think that this refutation is more an experimental artifact than a flaw in the theory. Indeed, the zero corner contribution has something special, that subjects tend to avoid forcefully. The fact that it is not chosen may not be considered an absolute contradiction of the logic of free-riding, but rather as a limit of its range of validity. This point of view suggests considering more pragmatic definitions of the status of "non-contributor". This can be done along two dimensions.

One possibility is, for any particular subject, to be less stringent about the number of occurrences of zero contributions during an U-sequence. We set, more or less arbitrary, at 5 the minimal number of periods per sequence without contributions that gives a subject the status of non-contributor. A subject who contributes zero in five out of 10 periods is a non-contributor half of the time. In other words if we take randomly a contribution of this subject, we have a probability of 1/2 to draw a non-contributor. The other possibility is to accept that not only zero but also a unitary contribution defines a non contributor.

Table 6 reports the number of non-contributors and the number of groups with two non-contributors for these weaker definitions of a non-contributor. As reported in the table, whatever the definition, the number of groups with two non-contributors remains too small to accept the theoretical predictions about the set of contributors.[7]

One reason why the set of contributors differs from the predicted one, is that subjects tend to over-contribute whatever their level of endowment. This reason is examined in the next subsection.

---

[7] We tried even weaker criteria concerning the number of periods with zero contribution or concerning the level of the amount contributed. However, none of these more permissive definitions, that still make sense, confirmed the prediction of BBV of two non-contributors and two contributors per groups in U-sequences.

Table 6 Non-contributors in U sequences

Treatment		Predicted[a]	Definition 1[b]	Definition 2[c]
EU	Number of non-contributors	3	6	5
	Number of groups with 2 non-contributors	1	2	1
UE	Number of non-contributors	3	7	4
	Number of groups with 2 non-contributors	0	2	0

[a] Contributes always 0
[b] Contributes 0 at least half of the time
[c] Contributes 0 or 1

Table 7 Average contribution for the test treatments

Treatment	Subjects	Sequence	Observed contribution	Nash contribution	H0 accepted[a]
EU	All	1st	9.99	7	No (0.008)
		2nd	12.09	11	Yes (0.071)
	Poor	1st	–	–	–
		2nd	1.24	0	No (0.011)
	Rich	1st	–	–	–
		2nd	22.94	22	Yes (0.578)
UE	All	1st	12.03	11	No (0.034)
		2nd	8.11	7	Yes (0.062)
	Poor	1st	1.51	0	No (0.004)
		2nd	–	–	–
	Rich	1st	22.54	22	Yes (0.466)
		2nd	–	–	–

[a] Wilcoxon Mann–Whitney (p-value)

4.2.2 Overcontribution by endowment-category

Over-contribution is a widespread phenomenon in public goods experiments, but which is generally less pronounced in public good games with an interior Nash equilibrium (e.g. public goods games with a quadratic payoff function[8]), compared to public good games with a corner solution (i.e. linear public good experiments). Accordingly, in our experiment we expect to see only moderate or no over-contributions in E-sequences. However, it remains an open question whether average over-contributions will be large or not in U-sequences, and whether it will be different for poor and rich.

At the group level, over-contribution is observed in all sequences but is not always significant. Lines labelled 'All' in Table 7 show that the average level of contribution is larger than the Nash prediction. The last column of Table 7 provides the results of

---

[8] However, Keser (1996) found significant rates of over-contribution for payoff functions which are quadratic in the private good. For a survey on the evidence from interior-Nash public good experiments see Holt and Laury (1998).

one-sided Wilcoxon tests for the null hypothesis that the average group contribution is equal to the predicted level. Over-contribution is frequent but not always significant in our data.

Apparently the rate of over-contribution depends on the level of endowment and the nature of redistribution. For poor subjects the Nash equilibrium contribution is 0 token: any strictly positive contribution is therefore an "over-contribution".

Result 3 *In the U-sequences of the test treatments, poor subjects over-contribute significantly while the contribution of rich subjects is consistent with the predicted level.*

Support 3: Results of Wilcoxon one-sided tests are reported in Table 7. We also analyze overcontributions in relative terms by defining $R_i(w_i) = \frac{g_i(w_i) - g^*(w_i)}{w_i}$. $R_i(w_i)$ is individual i's rate of overcontribution with respect to his endowment w_i. $g_i(w_i)$ is his contribution when his endowment is w_i and $g^*(w_i)$ is the Nash-contribution for endowment level w_i. For all treatments and for each independent group we observe that in inequality sequences the average rate of overcontribution of the poor is always larger than the rate of overcontribution of the rich. This of course implies that the rates of overcontribution of the poor are significantly larger at any standard level of significance and for any type of test.

The fact that observed overcontributions are negatively correlated with the level of the endowment is compatible with the idea that, at the individual level, contributions are driven by the Relative Strength of the Social Dilemma (RSSD), i.e. the relative gap between the socially optimum contribution and the Nash contribution, as shown in Maurice et al. (2013). The RSSD hypothesis implies that agent's i relative overcontribution $R_i(w) = \frac{g_i(w) - g^*(w)}{w}$ decreases with w. Maurice et al. (2013) showed that if individual i is sensitive to the RSSD, his rate of overcontribution decreases with his endowment ($R_i'(w) < 0$), a property that they observed in their data. Our data exhibits the same pattern: on average $R(5) > R(20) > R(35)$. For instance in sequence 2 of the EU treatment we observe average values $R(5) = 0.25$ and $R(35) = 0.03$ and in sequence 1 an average value $R(20) = 0.15$. Similarly in the case of the UE treatment we observe $R(5) = 0.30 > R(20) = 0.06 > R(35) = 0.02$.

Therefore individuals who "feel" a relatively stronger social dilemma have a tendency to overcontribute more. This happens more likely as individuals get poorer and becomes exacerbated for the poorest who are expected to contribute 0 % of their endowment.

In order to provide further support for this conjecture we need to examine more closely how subjects adjust their contributions after redistribution according to their type.

4.2.3 Contributions adjustments after redistribution

As predicted, after an unequalizing (equalizing) redistribution we observe a significant increase (decrease) of the group contribution. The null hypothesis of equal average contributions in the two sequences is rejected for the two test-treatments ($p = 0.004$ for EU and $p = 0.026$ for UE, Wilcoxon signed-rank test, one-sided). We also observe

that subjects adjust in the right direction after a variation of their endowment. There is a significant difference in average contribution before and after redistribution for each income-category ($p = 0.001(0.000)$ for poor (rich) subjects in the EU treatment and $p = 0.000$ (0.001) for poor (rich) subjects in the UE treatment, Wilcoxon signed rank tests, one-sided).

For the test treatments, the observed difference in group contribution between the U-sequence and the E-sequence is consistent with the predicted one. These differences are equal to -8.412 for the EU treatment and 15.66 for the UE treatment, respectively, and do not differ significantly from $+/-16$ ($p = 0.078$ for EU and $p = 1.000$ for UE, Wilcoxon Mann–Whitney test). It is striking however that the magnitude of the adjustment is almost twice as large under UE than under EU. In fact such differential adjustment in mainly due to the reaction of subjects who become richer in the EU treatment as explained below. The observed adjustments in contributions at the group level actually hide the fact that individual adjustments per income-category differ from the predicted ones.

Result 4 *After redistribution, subjects who become richer in the EU treatment under-adjust their contribution while rich subjects who become poorer in the UE treatment adjust their contribution as predicted.*

Support 4. We define the *Nash-adjustment* as the difference between the Nash contribution before and after redistribution. The Nash-adjustment depends on the agent's type. Poor adjust by 7 tokens: precisely, in the EU treatment subjects who become poor should lower their contribution by -7 tokens while in the UE treatment poor subjects who become "richer" should increase their contribution by $+7$ tokens. Similarly, rich adjust by 15 tokens: subjects who become richer in the EU treatment should increase their contribution by $+15$ and subjects who become "poorer" in treatment UE should lower it by -15 tokens. The average adjustment of the poor is -8.025 in EU and 6.075 in UE. The null hypothesis is not rejected for both treatments ($p = 0.547$ for EU and $p = 0.353$ or UE, Wilcoxon Mann–Whitney test). For the rich, the average adjustment is 12.225 in the EU treatment and -13.906 in the UE treatment. The null hypothesis is accepted for the UE treatment ($p = 0.076$, Wilcoxon Mann–Whitney test) but rejected for the EU treatment ($p = 0.049$, Wilcoxon test). Subjects who become richer in the EU treatment under-adjust after redistribution while rich subjects who become poorer adjust as predicted.

4.2.4 Equalizing versus unequalizing redistribution

The direction of the redistribution seems to matter for the rich but not for the poor, but a more careful investigation is required. We need to compare the levels of the adjustments of subjects who become poorer to those who become richer. There are two types of subjects who become poorer: those who had an endowment of 20 tokens in sequence 1 of EU and those who had an endowment of 35 tokens in sequence 1 of UE. Equivalently, there are two types of subjects who become richer: those who had an endowment of 20 tokens in sequence 1 of EU and those who had an endowment of 5 tokens in sequence 1 of UE. The average values of these adjustments are reported in Table 8. To make

Table 8 Contribution adjustments

Treatment	Average adjustment		Relative adjustment	
	Poor	Rich	Poor	Rich
EU	−8.03	12.23	1.15	0.82
UE	6.08	−13.91	0.87	0.93

them comparable we also report the relative adjustments, i.e. the ratios of the observed adjustment over the predicted adjustment. The ratio is significantly larger for subjects who become poor in sequence 2 of EU compared to subjects who are no longer rich in sequence 2 of UE ($p = 0.001$, Wilcoxon Mann-Whitney test). Similarly the ratio is larger for subjects who become rich in sequence 2 of EU compared to subjects who are no longer poor in sequence 2 of UE ($p = 0.001$, Wilcoxon Mann-Whitney test).

These observations suggest that becoming poorer in an unequalizing society does not trigger the same reactions than becoming poorer in an equalizing society. Similarly, becoming richer in an unequalizing society arouses a different reaction than becoming richer in an equalizing society. Such differential behavioural reactions, which are at odds with the theoretical predictions, require further analysis.

We do this by comparing treatments which have identical second sequences but different first sequences, for instance EU and UU. In the EU treatment, poor subjects may feel aggrieved, compared to subjects in the UU treatment who did not experience an egalitarian situation in the first place. To some extent individuals who are "born" in an unequal society find it more natural to be poor or rich than individuals who were born in an equal society and who experienced a change in status by becoming poor or rich with respect to their former group mates. Therefore poor and rich subjects in the second sequence of EU may have a different perception of their status compared to poor and rich subjects in the second sequence of UU. Such difference in perception can therefore lead to different levels of contribution, both by rich and by poor, in the second sequence of EU compared to UU. In particular subjects are likely to express a stronger concern for inequalities by their contributions when inequalities are arbitrarily generated compared to a situation where they already preexisted.

Result 5 *Average group contributions are lower in the second sequence of EU than in the second sequence of UU. The lower group contribution is due only to a lower contribution of subjects who become poor in EU. Subjects who become rich in EU contribute as the rich in the second sequence of UU.*

Support 5. We compare contributions of poor subjects in the second sequence of UU and EU. The observed average contributions (1.24 for EU and 2.56 for UU) are significantly different ($p = 0.024$, Wilcoxon Mann–Whitney). Contributions by rich subjects (22.94 in EU and 26.68 in UU) do not differ ($p = 0.105$, Wilcoxon Mann–Whitney). We also run panel data regressions in which we take as the dependent variable the *average over-contribution* of the poor (rich) in the second sequence of UU and EU. The over-contribution is measured by $g_i(w) - g^*(w)$ where $g^*(w)$ is the Nash contribution for endowment level w. Since $g^*(5) = 0$ a poor subject's over-contribution is simply equal to his contribution. For a rich subject $g^*(35) = 22$ and

therefore his predicted overcontribution is equal to $g_i(w)$—22. The coefficient of the treatment dummy variable *EU* (equal to 1 for EU and 0 for UU) shows that after redistribution poor subjects lower their overcontribution by 1.35 points on average in comparison with poor subjects in UU ($p = 0.01$). In contrast for rich subjects the coefficient of the treatment dummy is not significant, indicating that becoming rich or being rich triggers the same reactions.

At the group level we also find that the average group contribution is higher in the second sequence of UU than in the second sequence of EU ($p = 0.033$, Wilcoxon Mann–Whitney one-sided).

Our interpretation of result 5 is that poor exhibit aversion to the creation of a detrimental inequality. This effect is highlighted only for poor subjects: rich subjects may also have a concern for inequality, but to a lesser extent since they benefit rather than suffer from the creation of inequalities. In any case, no effect is detected for the rich. Our interpretation must be taken with caution however because our experiment was not designed to isolate subjects' reactions towards the creation of inequalities. At least further tests or replications of this design are needed to provide robustness. Nevertheless, our findings about the differential reactions of poor and rich to redistribution are in line with those of Maurice et al. (2013). Their experiment had exactly the same design but they implemented a "small" redistribution which was predicted to be "neutral" by standard theory. They showed that in general subjects tended to under-react to redistribution, i.e. that the magnitude of their adjustment was smaller than predicted, except in one case: subjects who became poor in the EU treatment. The latter ones did not underact, but lowered their contribution exactly by the predicted amount. We believe that the absence of an adjustment-gap in contributions for subjects who became poor after a small redistribution in EU in the experiment of Maurice et al. (2013), is certainly due the same reason why in our experiment subjects who become poor after a large redistribution contribute less than those who were poor at the outset. These subjects probably perceived the inequity resulting from redistribution differently from those who were in the poor situation initially. Finally, we observe a parallel reaction for rich: in our experiment rich under-adjust in the EU treatment, whereas they adjust as predicted in the UE treatment. Again, this observation suggests that subjects react not only to the creation of an inequality, but also to the decrease of their own income.

5 Conclusion

Okun (1975) became famous by defining redistribution as a transfer of money from rich to poor in a leaky bucket. This metaphorical statement concisely draws attention on a major social issue, which many call the Big trade-off between equality and efficiency. In theory, there are many well identified reasons behind the existence of such a trade-off. The one we explore in the lab has been relatively neglected so far. It is related to the effect of income redistribution on the provision of public goods (Itaya et al. 1997). More precisely we test, in a stylized public good experiment, whether increasing income inequality also increases welfare.

Some of our experimental results support the theory: (i) an unequalizing (equalizing) redistribution of the group endowment increases (decreases) the utilitarian social

welfare; (ii) payoffs of rich subjects are higher in the inegalitarian sequences than in egalitarian ones; (iii) in the U-sequences of the test treatments groups make larger contributions to the public good than in E-sequences. The directions and magnitudes of the adjustments are also in conformity with the predictions, although under-adjustments are frequently observed.

But we also found two systematic departures from theory: (a) redistribution does not induce a modification of the set of contributors, as predicted by BBV, neither in the EU treatment nor in the UE treatment, and (b) poor subjects in the second sequence of EU do not behave as poor subjects in the second sequence of UU. This can be understood as a concern for the creation of a detrimental inequality.

It is worth recalling that, because in U-sequences theory predicts that poor subjects should contribute zero, noisy behaviors are necessarily biased towards positive contributions. This experimental artifact indeed occurred in our experiment, as in the other experiments in the literature on public goods. But, although this behavioral regularity could explain abnormality (a), it cannot account for abnormality (b). Most theories of inequality aversion cannot either, because they would predict the same behaviors in the second sequence of EU and in the second sequence of UU.

Another explanatory mechanism is necessary where the variations, and not just the levels, of inequalities enter in the utility functions. In such a theory, subjects' concerns for the variations of inequalities would be mixed up with the usual logic of self-centered behaviors. A candidate explanation is that subjects have an additional motivation to contribute which is related to their sensitivity to the social dilemma. Their sensitivity to the social dilemma is affected not only by the distribution of the endowment between group members, but also by the variation of this distribution with respect to a reference point.

We considered however two very specific types of redistributions in our experiment: one where the reference situation was an egalitarian endowment distribution preceding an unequal distribution and one where the reference situation was an unequal endowment distribution preceding an egalitarian one. In addition the unequalizing redistribution entailed a Pareto-improvement. It would be of interest to relax these two properties in order to test whether our findings are robust to other types of redistributions, in particular a redistribution which impoverishes the poor when the reference distribution is unequal.

Conflict of interest None.

Appendix 1: Preliminary results

1.a Homogeneity

Identical first sequences are compared across treatments to control for an eventual session effect. The null hypothesis of equal average contributions in Sequence 1

cannot be rejected for the comparison between EE and EU ($p = 0.382$, Wilcoxon Mann–Whitney test), and for the comparison between UE and UU ($p = 0.105$, Wilcoxon Mann–Whitney test). Similarly, the null hypothesis of equal group payoff in Sequence 1 is not rejected, neither for the comparison between EE and EU ($p = 0.442$, Wilcoxon Mann–Whitney test) nor for the comparison between UE and UU ($p = 0.195$, Wilcoxon Mann–Whitney test).

1.b Restart effect

To test for the presence of a restart effect, we compare the two sequences of the baseline treatments. The null hypothesis of equal group contributions in Sequence 1 and in Sequence 2 cannot be rejected ($p = 0.640$ for EE, and $p = 0.400$ for UU, Wilcoxon signed-rank test). We also compare period 10 of Sequence 1 to period 1 of Sequence 2 (period 11) to refine our test. The null hypothesis of equal group contributions in these two periods cannot be rejected ($p = 0.313$ for EE, and $p = 0.106$ for UU, Wilcoxon signed-rank tests).

Appendix 2: Comparison of sequences in baseline treatments

2.a Payoffs

The null hypothesis of equal group payoffs in Sequence 1 and in Sequence 2 cannot be rejected for the two baseline treatments ($p = 0.945$ for EE and $p = 0.195$ for UU, Wilcoxon signed-rank test). We also test if payoff differences between the two sequences are equal to zero at group level and depend on the endowment of players. Our data does not reject the null hypothesis that payoff differences are equal to zero. Results of these Wilcoxon Mann–Whitney tests are reported in Table 9.

We check wether observed payoffs are equal to the predicted ones, both at group level and for poor and rich subjects separately. The null hypothesis is accepted except for Sequence 2 of UU (Table 10, Wilcoxon Mann–Whitney tests).

In accordance with the findings of the test treatments, in the UU baseline treatment rich subjects earn more than poor subjects : the null hypothesis of equal payoff for rich and poor is rejected ($p = 0.004$ for Sequence 1 and $p = 0.011$ for Sequence

Table 9 Payoffs difference between sequences 1 and 2 for baseline treatments

Treatment	Subjects	Sequence 1	Sequence 2	H0 accepted[a]
EE	All	836.80	833.29	Yes (0.945)
UU	All	987.40	1040.18	Yes (0.195)
	Poor	905.32	967.57	Yes (0.250)
	Rich	1069.47	1112.79	Yes (0.313)

[a] Wilcoxon Mann–Whitney (p-value)

Table 10 Average payoffs for baseline treatments

Treatment	Subjects	Sequence	Observed payoffs	Nash payoffs	H0 accepted[a]
EE	All	1st	836.80	784	Yes (0.313)
		2nd	833.29	784	Yes (0.109)
UU	All	1st	987.40	932	Yes (0.109)
		2nd	1040.18	932	No (0.008)
	Poor	1st	905.32	840	Yes (0.109)
		2nd	967.57	840	No (0.016)
	Rich	1st	1069.47	1024	Yes (0.195)
		2nd	1112.79	1024	No (0.008)

[a] Wilcoxon Mann–Whitney (p-value)

Table 11 Payoffs difference of poor and rich subjects in the same sequence for baseline treatments

Sequence[a]	Poor subjects	Rich subjects	Difference	Predicted difference	H0 accepted[b]
Uu	905.32	1069.47	164.15	184	Yes (0.547)
uU	967.57	1112.79	145.22	184	Yes (0.547)

[a] The uppercase represents the tested sequence
[b] Wilcoxon Mann–Whitney (p-value)

2, Wilcoxon Mann–Whitney test). The observed difference agrees with the predicted one (see Table 11).

2.b Contributions

The restart test already showed that average contributions are the same in Sequence 1 and Sequence 2 for the baseline treatments. We further compare the average observed

Table 12 Average contribution—over-contribution for baseline treatments

Treatment	Subjects	Sequence	Observed contribution	Nash contribution	H0 accepted[a]
EE	All	1st	8.90	7	Yes (0.055)
		2nd	8.58	7	No (0.027)
UU	All	1st	13.60	11	No (0.004)
		2nd	14.62	11	No (0.004)
	Poor	1st	2.59	0	No (0.004)
		2nd	2.56	0	No (0.004)
	Rich	1st	24.60	22	No (0.029)
		2nd	26.68	22	No (0.008)

[a] Wilcoxon Mann–Whitney test (p-value)

difference (1.276 for EE and -4.1 for UU) to zero: we cannot reject the null hypothesis of equality of the observed difference with zero ($p = 0.641$ for EE and $p = 0.400$ for UU, Wilcoxon Mann–Whitney tests). The null hypothesis of equal average contributions is also accepted if we test separately for poor and for rich in the UU treatment. The average contribution of Sequence 1 is equal to the average contribution of Sequence 2 ($p = 1.000$ for poor subjects and $p = 0.293$ for rich subjects, Wilcoxon Mann–Whitney test). In comparing the average difference between Sequence 1 and Sequence 2, we cannot reject the null hypothesis that the difference of contribution is equal to zero for both types of subjects ($p = 1.000$ for poor and $p = 0.293$ for rich, Wilcoxon Mann–Whitney tests). Finally, a significant over-contribution—with respect to the Nash prediction - is observed in all sequences of the baseline treatments, except in the first sequence of EE. Also, rich over-contribute significantly in the two sequences UU in contrast to the inequality sequences of the test treatments (Table 12).

References

Alesina AF, Giuliano P (2009) Preferences for redistribution. NBER Working Paper Series, vol w14825. Available at SSRN: http://ssrn.com/abstract=1369061

Anderson LR, Mellor JM, Milyo J (2008) Inequality and public good provision: an experimental analysis. J Soc-Econ 37:1010–1028

Andreoni J (1988) Why free ride? Strategies and learning in public goods experiments. J Public Econ 37:291–304

Andreoni J, Miller J (2002) Giving according to GARP: an experimental test of the consistency of preferences for altruism. Econometrica 70:737–753

Atkinson AB, Stiglitz JE (1976) The design of tax structure: direct versus indirect taxation. J Public Econ 6:55–75

Bagnoli M, McKee M (1991) Voluntary contribution games: efficient private provision of public goods. Econ Inq 29:351–366

Bergstrom TC, Blume L, Varian HR (1986) On the private provision of public goods. J Public Econ 29:25–49

Bracht J, Figuières C, Ratto M (2008) Relative performance of two simple incentive mechanisms in a public goods experiment. J Public Econ 92:54–90

Brueckner JK (2008) Welfare reform and the race to the bottom: theory and evidence. South Econ J 66(2):505–525

Bullock DS (1995) Are government transfers efficient? An alternative test of the efficient redistribution hypothesis. J Polit Econ 103:1236–1274

Chan KS, Mestelman S, Moir R, Andrew Muller RA (1996) The voluntary provision of public goods under varying income distributions. Can J Econ 29:54–69

Chan KS, Mestelman S, Moir R, Andrew Muller RA (1999) Heterogeneity and the voluntary provision of public goods. Exp Econ 2:5–30

Cherry TL, Kroll S, Shogren JF (2005) The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab. J Econ Behav Org 57:357–365

Croson RTA (1996) Partners and strangers revisited. Econ Lett 53:25–32

Dahlberg M, Edmark K (2008) Is there a "race-to-the-bottom" in the setting of welfare benefit levels? Evidence from a policy intervention. J Public Econ 92(5–6):1193–1209

de Cremer D (2007) When the rich contribute more in public good dilemmas: the effects of provision point level. Eur J Soc Psychol 37:536–546

Diamond PA, Mirrlees JA (1971a) Optimal taxation and public production: I—Production efficiency. Am Econ Rev 61(1):8–27

Diamond PA, Mirrlees JA (1971b) Optimal taxation and public production: II—Tax rules. Am Econ Rev 61(3):261–278

Durante R, Putterman L, Weele J (2014) Preferences for redistribution and perception of fairness: an experimental study. J Eur Econ Assoc 12(4):1059–1086

Fischbacher U (2007) z-Tree: Zurich toolbox for ready-made economic experiments. Exp Econ 10:171–178

Gardner B (1983) Efficient redistribution through commodity markets. Am J Agric Econ 65:225–234

Georgantzís N, Proestakis A (2011) Accounting for real wealth in heterogeneous-endowment public good games. The Papers 10/20, Department of Economic Theory and Economic History of the University of Granada

Güth W, Kliemt H, Ockenfels A (2003) Fairness Versus efficiency—an experimental study of (mutual) gift giving. J Econ Behav Org 50:456–475

Güth W, Pull K, Stadler M, Stribeck A (2010) Equity versus efficiency? Evidence from Three-person generosity experiments. Games 1 2:89–102

Holt CA, Laury SK (1998) Theoretical explanations of treatment effects in voluntary contributions experiments. In: Plott C, Smith V (eds) Handbook of experimental economics results. Elsevier, New York

Hofmeyr A, Burns J, Visser M (2007) Income inequality, reciprocity and public good provision: an experimental analysis. South African J Econ 75(3):508–520

Itaya J-I, de Meza D, Myles GD (1997) In praise of inequality: public good provision and income distribution. Econ Lett 57:289–296

Keser C (1996) Voluntary contributions to a public good when partial contribution is a dominant strategy. Econ Lett 50:359–366

Maurice J, Rouaix A, Willinger M (2013) Income redistribution and public good provision: an experiment. Int Econ Rev 54(3):957–975

Mirrlees JA (1971) An exploration in the theory of optimum income taxation. Rev Econ Stud 38:175–208

Okun AM (1975) Equality and efficiency: the big trade-off. The Brookings Institution, Washington

Piketty T (1993) Implementation of first-best allocations via generalized tax schedules. J Econ Theory 61:23–41

Rapoport A, Suleiman R (1993) Incremental contribution in step-level public goods games with asymetric players. Org Behav Hum Decis Process 55:171–194

van Dijk E, Grodzka M (1992) The influence of endowments asymmetry and information level on the contribution to a public step good. J Econ Psychol 13:329–342

van Dijk E, Wilke H (1995) Coordination rules in asymmetric social dilemmas: a comparison between public good dilemmas and resource dilemmas. J Exp Soc Psychol 31:1–27

Warr PG (1983) The private provision of a public good is independent of the distribution of income. Econ Lett 13:207–211

Wildasin DE (1991) Income redistribution in a common labor market. Am Econ Rev 81(4):757–774

Yamada K, Naito H (2014) Neutrality theorem revisited: an empirical examination of household public goods provision. Tsukuba Economics Working Papers

Zelmer J (2003) Linear public goods experiments: a meta-analysis. Exp Econ 6:299–310