The trade off between welfare and equality in a public good experiment : do the poor need Robin Hood or the Sheriff of Nottingham?

Agathe Rouaix, Charles C. Figuieres, Marc Willinger

▶ To cite this version:

Agathe Rouaix, Charles C. Figuieres, Marc Willinger. The trade off between welfare and equality in a public good experiment : do the poor need Robin Hood or the Sheriff of Nottingham?. 2013. hal-02804245

HAL Id: hal-02804245
https://hal.inrae.fr/hal-02804245v1

Preprint submitted on 5 Jun 2020

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.

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

Do the poor need Robin Hood or the Sheriff of Nottingham?

Agathe Rouaix, Charles Figuières, Marc Willinger

May 21, 2012

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, De Meza & Myles, 1997), because the larger contribution of the rich subjects overcompensates the lower contribution of the poor subjects. Last but not least, agents' behavior in situations of inequality of income, depends on initial conditions. In particular, the decisions of those who become poor seem to express a form of protest againts the new distribution of incomes: being poor is not the same thing as becoming poor.

1 Introduction

This article reports the results of a public good experiment where subjects' incomes are redistributed in a lump sum manner. We study the impact of such redistribution on voluntary contributions to the public good and, more importantly, on welfare.

We rely on this experimental setting in order to contribute to an old debate, the "Big trade-off between efficiency and equity" (Okun, 1975), with a new angle. In most developed countries the pursuit of social justice is confronted to the difficulty that cutting the pie less fairly can increase it, with the possibility that even the poorest in relative terms might be better-off in absolute terms. There are a number of reasons underlying this trade-off, for instance the risk of evasion of mobile resources and the disincentive effects of taxes on skilled labor. Not surprisingly, these reasons are echoed in a vast theoretical literature concentrating on optimal taxation (e.g. Mirrlees, 1971, Atkinson & Stiglitz, 1976, Piketty, 1993), and in an equally extensive empirical literature on the issue of efficient income redistribution (e.g., Gardner 1983, Bullock 1995)

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 provision of public goods[1]. The reason why the efficiency-equity tradeoff might be affected by the provision of public goods is independent of those already mentioned and remains valid even if, in the considered economy, resources are immobile and individuals all have the same productivity or skills. It suffices that individuals differ with respect to their propensity to contribute to the provision of public goods. More precisely, if rich individuals are willing to contribute larger amounts to the provision of public goods than poor 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.

[1]It is interesting here to situate this "public goods" argument in the academic debate. There are reasons to believe that the alleged compromise does not exist, that the objectives of equity and efficiency can be treated separately. In essence this is the message of the famous Second Welfare Theorem. However, the set of conditions under which this theorem is valid is restrictive. In particular, it does not hold any longer in the presence of public goods.

Itaya, de Meza and Myles (1997, IdMM hereafter) showed, in a symmetric public good framework, that non cooperative contributions result in the equalization of utilities over agents, even if the distribution of incomes is unequal. Therefore if the focus is on individual utility rather than on income *per se*, there seems to be no role for equalizing redistributions. In addition, the authors also proved that creating incomes inequality can increase social welfare when some agents cease to contribute. They conclude that state interventions to increase income inequalities, in their environment, can be desirable.

Their analysis, as ours, belongs to a stream of litterature that examines the policy implications of the famous *neutrality theorem* of income redistribution, presented first by Warr (1983) and developed by Bergström, Blume and Varian (1986, henceforth BBV). We are interested in the violation of the conditions of this theorem and the ensuing *non-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).

IdMM's finding, that a 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 to ensure arts patronage or various kinds of public works.

[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, 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.

Our experimental setting follows IdMM's theoretical framework. We run a public good game experiment in which we implement a reallocation of endowments among group members. We chose the set of parameters in such a way that the predictions are as follows: poor agents should not contribute any token to the public good, social welfare (group payoff) increases after an unequalizing redistribution, and a equalizing redistribution generates a Pareto degradation (poor and rich have a lower payoff). Our experimental data do not reject these predictions with one important exception: poor subjects still contribute to the public good. Last but not least, agents' behavior in situations of inequality of income, depends on initial conditions. In particular, the decisions of those who become poor seem to express a form of protest against the new distribution of incomes: being poor is not the same thing as becoming poor.

Section 2 offers a brief overview of previous public good experiments dealing with inequality of endowments or with the neutrality theorem. Section 3 presents our experimental design. Our experimental results are given in Section 4. The last section concludes.

2 Previous experiments

The *"Equity versus Efficiency"* trade-off has been analyzed in various types of experiments: gift exchanges games (Güth et al., 2003), generosity games (Güth et al., 2010), elicitation of preferences for redistribution (Durante & Putterman, 2009, Alesina & Giuliano, 2009 who provide survey data). One of the findings of this litterature is that subjects have a lower taste for equity when the price of equity increases (see e.g. Andreoni & Miller, 2002).

Public good games with endowment heterogeneity addresses the issue of the equity versus efficiency puzzle indirectly. Does endowment heterogeneity lead to lower or higher public good provision, or is endowment distribution irrelevant? This question was initially studied in the case of step level public good games. The main findings was that endowment heterogeneity is detrimental for efficiency: groups with heterogenous endowments are less successful in reaching the provision point than groups with equal endowments (Bagnoli and MacKee, 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 good games it was found that 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. (2008) found that it had no impact on the level of provision. However, based on a meta-analysis of 27 experimental studies, Zelmer (2003) found that endowment heterogeneity had a significantly negative impact on contributions. The existing experimental literature on linear public goods seems therefore to suggest that the tradeoff between equity and efficiency is irrelevant in the case of public good provision since moving towards a more equal endowment distribution enhances efficiency through increased public good provision. However, Georgantzis and Proestakis (2011) showed that such conclusion is dependent on how equity is measured. They elicited participants's real disposal income before constituting heterogenous income and endowment groups. 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.

Furthermore, experiments based on non-linear public good games (e.g. Chan et al., 1996,1999) provide also evidence about the relevance of the tradeoff. 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) are compatible with 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 more public good, clearly in 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., (2012) who implemented a real income redistribution in their experiment. Subjects played a public good game in two consecutive sequences of 10 periods, with two different endowment distributions: equal versus unequal (or the reverse order). They found that the quantity of public good provided is not affected by an equalizing or unequalizing redistribution. We will rely on the same design than in Maurice et al., (2012) to address the non-neutrality issue raised by IdMM that welfare increases when inequalities increase because of increased public good provision.

3 Experimental Design

3.1 The contribution game

Subjects are endowed with quasi-linear quadratic monetary payoffs, as in Keser (1996) or Bracht, Figuières & Ratto (2008):

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

where each subject i has an endowment w_i, $x_i = w_i - g_i$ is the consumption of the private good by subject i, 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 $n = 4$. With such monetary payoffs, we end up with one of the simplest public good games where the Nash equilibrium is a unique profile of interior decisions, a feature that is useful to test BBV's predictions. 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 $w_i = 20, i = 1, ..., 4$. The amount of public good will be $G_E = 4 \times 7 = 28$, where the index 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 the the group total is unchanged. The amount of public good is now $G_U = 2 \times 22 + 2 \times 0 = 44$, where U refers to an "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 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 that by the rich. Final wealth of the poor is therefore equal to 830 units ($170 = 41 * 5 - 5^2$, from their private consumption and $44 * 15 = 660$ from the public good) while the final wealth of the rich is 1024 units ($364 = 41 * 13 - 13^2$, from their private consumption and $44 * 15 = 660$ from the public good). The aggregate wealth is therefore larger in the inequality situation (3708 units) than in the equality situation (3136) [4]. This property is in

[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 of Theorem 5 in BBV: (i) *Any change in the wealth distribution that leave unchanged the aggregate wealth of current contributors*

accordance with the predictions of IdMM (1997): although they considered the case of 2 agents, they contend that their predictions can easily be extended to a larger number of agents. In our framework 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

will either increase or leave unchanged the equilibrium supply of public good.

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

for it. We therefore add two control treatments, EE and UU, to our design. 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 characteristics of the control treatments and table 2 does the same for our test treatments.

[Table 1 about here.]

[Table 2 about here.]

3.3 Practical procedures

We conducted the experiment at LEEM, the computerized laboratory of the University of Montpellier I, with the software z-Tree (Fischbascher, 2007). We run 8 sessions involving 16 subjects each. The 128 subjects were randomly selected from a pool of student-subjects containing more than 5 000 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 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 (20 tokens per player) or in an unequal way (two subjects received 5 tokens and two subjects received 35 tokens) 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 accross 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. Note that this procedure differs from other experiments having a restart of the game (Andreoni, 1988 ; Croson, 1996), but it is used in some other latest experiments as in Andersen et al. (2008)[6].

4 Results

Our main goal is to provide experimental evidence about the welfare-equality trade-off within the context of a social dilemma. A secondary objective is to verify if

[6]They chose this payment method because it reduces experimentator costs, and it "may induce subjects to think carefully about each decision since only one matters in determining their earnings". But they also admit that we do not know if subjects are affected or not by this method.

individual reactions to redistribution, as predicted by standard theory or by usual behavioral conjectures in public good experiments, are indeed those observed in the lab. Checks will be both at group level and individual level. In addition to contribution levels, we will also focus on variations of contributions caused by redistribution. ¿From a theoretical perpective there is no issue in investigating those variations since they can be directly deduced from the predicted contribution levels. However, subjects in the lab may react differently, depending on their status (rich or poor) and depending on the direction of redistribution (unequalizing or equalizing). Furthermore, observed variations might be compatible with predicted variations even though observed levels are not. A common finding from experiments on voluntary contributions is that subjects over-contribute on average with respect to their Nash-contribution. We therefore expect that it will also be the case in our experiment. But the behavioral conjecture about variations of contributions could still hold even if all contribution levels are shifted upwards.

We run non-parametric tests and, unless otherwise specified, all of our tests are two-sided at the 5% significance level. We mainly run tests at the group level by taking, for each variable of interest, the average value of the four subjects of groups over the 10 periods of a sequence. For some results we need to distinguish between poor and rich subjects. In these cases we run the same tests but group averages are taken on each kind of subject (poor and rich).

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. Those tests are useful for the sake of rigor, but of an indirect interest regarding the questions asked in this paper. Therefore they have been relegated to the Appendix. What is important to note at this stage is that these preliminary tests indicate that data are homogeneous across sessions, and there is no significant restart effect in the baseline treatments. Because of the latter results, the effects of redistribution can be tested both with between tests (comparing equal and unequal treatments) and within tests *i.e.* by comparing the two sequences of the same treatment. Since across tests, which compare second sequences having the same first sequence, provide the same results, we will present mainly results from the "within" tests.

4.1 The welfare / equality trade-off

Can we expect that an increase (decrease) of endowments inequality be associated to higher (smaller) social welfare? We first check with our data whether this is true at the group level. Second, we examine the payoffs of the rich and poor separately to test whether individual well-being is also modified as predicted in our game specification. In E and U-sequences we calculate the average individual payoff for each group, and in U-sequences we also compute the average payoff for poor and rich subjects.

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

Support 1 *In U-sequences subjects earn on average significantly more than in E-sequences. We reject the null hypothesis that payoffs are equal in both sequences (p-value = 0.026 for EU, and 0.004 for UE, Wilcoxon signed-rank test on-sided).*

Our data corroborates IdMM's prediction regarding welfare. We run two complementary tests connected with this result.

Firstly, we test whether the average payoff difference between the two sequences is equal to the predicted difference. Table 3 summarizes the data by treatment and by subject-status for unequal-sequences. The null hypothesis is accepted at group level (lines entitled 'all') for all treatments according to Wilcoxon Mann-Withney tests: average payoff difference between the two sequences of a treatment is equal to the predicted difference. The same tests are made for baseline treatments and reported in the appendix (table 8). We find no difference between the observed difference and the predicted one (*i.e.* 0).

[Table 3 about here.]

Secondly we test whether the average payoff levels differ from the predicted levels (see table 4). The null hypothesis is accepted in almost all cases (Wilcoxon Mann-Withney tests), with one exception: the first sequence of the EU treatment.

[Table 4 about here.]

While the predictions about the effects of redistribution on welfare are satisfied at the group level we observe that the payoffs of poor subjects do not vary as expected. This is stated as result 2:

Result 2 *The welfare level of rich subjects is higher in U-sequences than in E-sequences while it is identical for poor subjects in both sequences.*

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 equality and unequality sequences for the rich (p-value = 0.004, Wilcoxon signed-rank test, one-sided). The null hypothesis of payoff equality for the poor in E-sequences and U-sequences cannot be rejected (p-value = 0.727 for the EU treatment, p-value= 0.098 for the UE treatment; Wilcoxon signed-rank test).*

Note that at the group level the average payoff difference between the two sequences is equal to the predicted difference (Wilcoxon test, see table 3, *in the lines entitled "rich" and "poor"[7]).*

[7]Note that for poor subjects we do not find any significant difference in payoffs for poor subjects before and after the redistribution but the observed differences between payoffs of both sequences do not differ from 56 (or -56), the theoretical difference. So we also have compare the observed difference to zero, and we do not find any significant difference ($p - value = 0.641$ in EU the treatment and $p - value = 0.195$ 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 be put forward to explain these contrary results. First the difference to observe is very small in comparison to the absolute amount 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 represent 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.

Furthermore, rich and poor's average payoffs do not differ from the predicted payoffs (see table 4 which reports the test results).

As a consequence rich earn more than poor in all U sequences irrespective of the treatment: there is a significant payoff difference between rich and poor in the inegalitarian sequences of EU and UE (p-value = 0.004, Wilcoxon signed-rank test) which corresponds to the predicted difference (see table 5).

[Table 5 about here.]

Our main conclusion is that at the group level observed payoffs do not contradict IMM's predictions, both in magnitude and in difference. However at the individual level we observe that in inequality sequences rich subjects obtain a higher payoff while poor subjects earn the same payoff as in egalitarian sequences. This suggests that individual reactions to redistribution in the lab might differ from those predicted by BBV's theory. It is therefore useful to further investigate the accuracy of the theoretical predictions related to income redistribution with respect to our experimental data. This is done in the next section.

4.2 Understanding the reactions to redistribution

Increases and decreases of subjects' well-being after redistribution is due to the re-allocation of tokens between the private and the public goods which induces a variation of the amount of public good provided and, in theory, to a modification of the set of contributors. In this section we investigate whether our data fits these predictions. We will first compare the observed contributions, both at the group and at the individual levels, to the Nash predictions. Then, for rich and poor agents we compare the observed and the predicted variations of contributions. Finally, we compare the observed and predicted sets of contributing agents.

4.2.1 Non-neutrality: variation of the quantity of public good after a redistribution

At the group level, the model based on the assumption of selfish rational subjects predicts correctly the amount of public goods provided under the different income distributions.

Result 3 *The quantity of public provided is higher in U-sequences than in E-sequences in the two test treatments (EU and UE).*

Support 3 *At the group level we reject the null hypothesis of equal average contribution for the two sequences of the EU treatment ($p-value = .004$, one-sided Wilcoxon signed-rank test) and the two sequences of the UE treatment ($p-value = .026$, one-sided Wilcoxon signed-rank test).*

While the average group contributions are compatible with the standard predictions, the underlying reasons might differ from those which sustain the theory. It is therefore important to analyze individual contributions by taking care to distinguish the contributions by poor subjects from those of rich subjects.

First we check if the variations - not the levels - of the amount of public good and of contributions by type are in accordance with the Nash predictions. It turns out that the differences in group contributions between the U-sequence and the E-sequence of the test treatments are consistent with the prediction. For the EU treatment, the observed difference between both sequences is -8.412, and the predicted one is -16. But with the Wilcoxon test, we can not reject the null hypothesis of equality between the observed and the predicted difference ($p - value = 0.078$). The same conclusion can be drawn for the UE treatment: the observed difference is 15.66 and the predicted one is 16 ($p - value = 1$).

Elaborating on the same idea, but at individual levels and distinguishing rich and poor subjects, we find:

Result 4 *Following an unequalizing (EU) or equalizing (UE) redistribution both the subjects who become poorer and those who become richer adjust their contribution in the predicted direction.*

Support 4 *At individual level, we compare the average contribution of each type of subjects (poor or rich) before and after redistribution. The difference between the two sequences is always significant for the test treatments with one-sided Wilcoxon signed rank tests. In the EU treatment the* $p - value$ *is* .001 *for poor subjects and* .000 *for rich subjects. In the UE treatment the corresponding p-values are* .000 *for poor subjects and* .001 *for rich subjects. We conclude that subjects adjust in the right direction after a variation of their endowment.*

It remains to check whether the magnitude of the adjustment for each type is consistent with the theoretical prediction. As noted previously, subjects tend to over-contribute with respect to their Nash-contribution, but this fact does not preclude that they adjust their contribution after an income shock by the predicted amount. Let us define the Nash-adjustment as the difference between Nash-contributions after and before redistribution. Agents who become poor should adjust by 7 tokens. Precisely, in the EU treatment subjects who become poor should lower their contribution by -7 tokens. In contrast, for the UE treatment poor subjects who become "richer" should increase their contribution by $+7$ tokens. We observe respectively average adjustments of -8.025 and 6.075. The corresponding absolute values are not significantly different from the predicted ones ($p - value = 0.547$ for EU and $p - value = 0.353$ or UE, Wilcoxon test). Subjects who become richer in the EU treatment should increase their contribution by $+15$ (EU) while in treatment UE rich subjects who become "poorer" should lower it by -15 tokens. The observed average values are 12.225 for the EU treatment and -13.906 for the UE treatment. The corresponding values are not significantly different from the prediction for the UE treatment ($p - value = 0.076$, Wilcoxon test), but not for the EU treatment($p - value = 0.049$, Wilcoxon test). When they become richer, subjects under-adjust to the redistribution but not when they become poorer. The direction of the redistribution matters for them and not for poor subjects. But when we compare the absolute amount of adjustment of rich subjects becoming poorer to the adjustment of subjects becoming rich, we do not find any difference. Perhaps this asymmetry is not so robust.

We have also to mention that adjustments of rich subjects in UE are not very close from the one predicted, and that the p-value would be significant at 10%, while adjustments of poor subjects are very similar to the one predicted. It seems that it is more difficult for rich subjects to modify their contribution, upward or downward, than for poor subjects. To complete this idea, we have to compare the value of adjustments of subjects who become poorer (poor in EU and rich in UE) and of subjects who become richer (rich in EU and poor in UE). As these values are not equal, we compare the proportion of the adjustment made by subjects relatively to the one predicted. We find that the adjusted proportion is significantly larger for subjects who become poor in EU than for subjects who are no more rich in UE ($p-value = 0.001$, Mann-Witney Wilcoxon test) and for subjects who become rich in EU than for subject who are no more poor in UE($p-value = 0.001$, Mann-Witney Wilcoxon test).

Finally are the observed contributions consistent with the predicted ones? As we reminded above, one can expect that the answer will be negative, since over-contribution is a widespread phenomenon in public goods experiments. However, over-contribution is less pronounced in public good games with an interior Nash equilibrium such as 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 should expect only moderate or no over-contributions in E-sequences and the same for rich subjects in U-sequences, whereas there should be significant over-contributions by poor subjects in U-sequences.

At group levels, over-contribution is observed in all sequences but it is not always significant. Lines labelled 'all' in table 6 show that the average level of contribution is larger than the Nash prediction. The last column of table 6 provides the results

[8]However, Keser (1996) and van Dijk et al. (1997) 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 goods experiments see Laury and Holt (1998).

of one-sided Wilcoxon tests for the null hypothesis that the average contribution is equal to the predicted level of group contribution. In two out of four sequences of the tests treatments, groups over-contribute significantly to the public good. Over-contribution is frequent but not always significant in our experiment.

[Table 6 about here.]

Finally we check if poor and rich subjects over-contribute to the public good. For poor subjects the Nash equilibrium contribution is 0 token: a strictly positive contribution can therefore be considered as an "over-contribution". Poor subjects always over-contribute significantly in the U-sequences of the tests treatments, but contributions by rich subjects do not differ from the predicted level *(i.e.* 22 tokens), so there is no evidence of over-contribution for that kind of subject. Results of Wilcoxon one-sided tests are reported in table 6. Those results are in line with previous observations in public good experiments.

Let us summarize our findings: our data are consistent with the theoretical predictions concerning the amount of public good provided at the group level, but at the individual level we observe that, despite subjects adjust in the right direction after redistribution, poor over-contribute while rich Nash-contribute. Over-contribution by the poor subjects is inconsistent with theory but is consistent with behavioral conjectures in public goods experiments. It will probably have a consequence on the part of BBV's theorem concerning the modification of the 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 whose income is strictly larger than 13 are expected to be contributors. Therefore we should observe 4 contributors in E-sequences and 2 contributors in U-sequences. Sticking to this strict definition of contributors, only one group in the EU treatment has two contributors and two non-contributors, consistently with the theory (other groups have less than two or zero contributors). Data clearly reject 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 experimental subjects tend to avoid forcefull. The fact that it is not chosen may not be considered an absolute contradiction of the logic of free-riding, but rather 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 occurences of zero contributions during an U-sequence. We set, more or less arbitrary, at 5 the minimal number of periods 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 define a non contributor.

Table 7 reports the number of non-contributors and the number of groups with two non-contributors for different definitions of the non-contributors. As reported in Table 7, with the first definition, less than half subjects are called non-contributors, and only a quarter of groups have a number of contributors equal to the one predicted. Theoretical predictions are also rejected in this case. Under the second definition a non-contributor is a subject who contributes 0 or 1 in the 10 periods of a sequence. Here too, we do not have a number of non-contributors sufficient to accept the theory. We tried other combinations of the number of periods and amount of contributions, but none, that still make sense, confirmed the prediction of BBV of two non-contributors and two contributors per groups in U-sequences.

[Table 7 about here.]

4.2.2 Aversion to endowment inequalities

So far we have focused on within tests in EU and UE treatments. Results found with these tests are the same as those obtained with between tests comparing second sequences having the same first sequence. But testing identical second sequences having different first sequences can also be very instructive to understand a relatively ignored aspect of behaviors. Do subjects in the second sequence of EU have the same perception of their endowment as subjects in the second sequence of UU? Could subjects' reactions depend not only for their new environment in the absolute but also reveal a concern for the creation of inequalities in relative terms[9]? In the EU treatment, poor subjects may feel aggrieved, compared to subjects in UU treatment where they do not have experienced an egalitarian situation in the first place. When we compare contributions of poor players in these two sequences (1.24 for EU and 2.56 for UU) we find that they differ significantly (Mann-Withney Wilcoxon, $p-value = 0.024$). Contributions of rich subjects (22.94 in EU and 26.68 in UU) do not differ (Mann-Withney Wilcoxon, $p-value = 0.105$). On group level we also find that average group contributions are higher in the second sequence of UU than in the second sequence of EU (Mann-Withney Wilcoxon one-sided, $p-value = 0.033$). Perhaps these observations are due to the existence of a potential aversion to the creation of inequality. This effect is highlighted only for poor subjects; rich subjects might be also concerned by this aversion, but to a lesser extent since they benefit rather than suffer from the inequalities. In any case, no potential effect was detected.

These results have to be taken with caution. This experiment was not precisely designed to isolate this reaction and, at least, further tests or replications of this design are needed to have more data. But those results seem to be in line with another related experiment by Maurice et al (2012). Their experiment has exactly the same design but they have implemented a "small" redistribution that theory predicts to be neutral. In their results, all subjects under-react to the redistribution: they adjust their contribution in the right direction, but of an amount smaller than the one predicted. There is just one exception: poor subjects in the EU treatment adjust their contribution by the predicted amount. Here too there is an evidence of the existence of a potential aversion to the creation of an inequality, perhaps because of the perception of the redistribution in terms of fairness. We find a parallel result in

[9]That seems to be more interesting than the reaction to the disappearing of the inequalities, that could be tested by comparing second sequences of EE and UE.

our experiment concerning behavior of rich subjects: in EU treatment, rich subjects under-adjust, whereas they adjust as predicted in UE. Perhaps 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 contribute a larger amount to the public good than in E-sequences, thanks to the adjustment of contribution by subjects in the predicted direction. The amounts of the adjustments are also in conformity with predictions.

But we also found systematic departures from theory:

a) The 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.

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 artifacts indeed occurs in our experiment, as in the other experiments in the literature on public goods. But, although this behavioural regularity could explain abnormality a), it cannot account for abnormality b). Usual theories of aversion inequality 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 mixed up with the usual logic of self-centered behaviors.

Finally, in a broader perspective, what lesson can we draw from these results on the functionning of the welfare state? A response like "reduction of income inequality is useless, and may even lose opportunities for greater well-being" is of course too bold, based on this single article and the particular logic of voluntary contributions to public goods. It is better seen as an invitation to continue the debate, perhaps with empirical analysis on the relationship between income redistribution and voluntary contributions to some public goods. It is also an illustration of the interest to clearly distinguish between income inequality and inequality of well-being.

References

[1] Alesina, Alberto F. and Giuliano, Paola, 2009, "Preferences for Redistribution". NBER Working Paper Series, Vol. w14825. Available at SSRN: http://ssrn.com/abstract=1369061

[2] Anderson, Simon, Jacob Goeree and Charles Holt, 1998. "A Theoretical Analysis of Altruism and Decision Error in Public Goods Games", *Journal of Public Economics*, 70 (2): 297-323.

[3] Andreoni, James, 1988. "Why Free Ride? Strategies and Learning in Public Goods Experiments", *Journal of Public Economics*, 37 (3): 291-304.

[4] Andreoni, James, 1995. "Cooperation in Public-Goods Experiments: Kindness or Confusion?", *The American Economic Review*, 85 (4): 891-904.

[5] Andreoni, James and Miller, John, 2002. "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism," Econometrica, vol. 70(2), 737-753.

[6] Bagnoli, Mark. and McKee, Michael, 1991. "Voluntary Contribution Games: Efficient Private Provision of Public Goods". *Economic Inquiry*, 29 (2), 351-366.

[7] Bergstrom T.C., Blume L.E. and Varian H.R., 1986, "On the private provision of public goods", *Journal of Public Economics*, 29, 25-49.

[8] Bracht, Juergen, Figuières, Charles, and Ratto, Marisa, 2008. Relative performance of two simple incentive mechanisms in a public goods experiment, *Journal of Public Economics*, 92(1-2), 54-90.

[9] Cherry, Todd L., Kroll, Stephan and Shogren, Jason F. , 2005, "The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab", *Journal of Economic Behavior and Organization*, 57(3), 357-365.

[10] De Cremer, David, 2007. "When the rich contribute more in public good dilemmas: The effects of provision point and voice procedures". *European Journal of Social Psychology*, 37(3), 536-546.

[11] Durante, Ruben and Putterman, Louis G., 2009, Preferences for Redistribution and Perception of Fairness: An Experimental Study, Available at SSRN: http://ssrn.com/abstract=1004573 or http://dx.doi.org/10.2139/ssrn.1004573.

[12] Georgantzís, Nikolaos & Proestakis, Antonios, 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.

[13] Itaya, Jun-ichi, de Meza, David, and Myles, Gareth D., 1997. In praise of inequality: public good provision and income distribution, *Economics Letters*, 57(3), 289-296.

[14] Güth, Werner, Kliemt, Hartmut and Ockenfels, Axel, 2003, "Fairness Versus Efficiency - An Experimental Study of (Mutual) Gift Giving", *Journal of Economic Behavior and Organization*, vol. 50(4), 456-475.

[15] Güth, Werner, Pull, Kerstin, Stadler, Manfred and Stribeck, Agnes. 2010. "Equity versus Efficiency? Evidence from Three-Person Generosity Experiments." *Games 1*, no. 2: 89-102.

[16] Hofmeyr, Andre, Burns, Justine and Visser, Martine, 2008. "Income Inequality, Reciprocity and Public Good Provision: An Experimental Analysis," Working Papers 77, Economic Research Southern Africa.

[17] Maurice, Jonathan, Rouaix, Agathe and Willinger, Marc, 2011. "Income Redistribution and Public Good Provision: an Experiment," Working Papers 09-12, LAMETA, Universtiy of Montpellier.

[18] Mirrlees, James A., 1971, "An exploration in the theory of optimum income taxation", *Review of Economic Studies*, 38, 175-208.

[19] Musgrave, Richard A., 1959. The Theory of Public Finance. A study in public economy, New York, McGraw-Hill.

[20] Okun, A.M., 1975, Equality and Efficiency: The Big Trade-off, Washington (DC): The Brookings Institution.

[21] Piketty, T., 1993, "Implementation of first-best allocations via generalized tax schedules", *Journal of Economic Theory*, 61(1), 23-41.

[22] Suleiman, Ramzi, Budescu, David V. and Rapoport, Amnon, 2001, "Provision of Step-Level Public Goods with Uncertain Provision Threshold and Continuous Contribution", Group Decision and Negotiation, 253-274.

[23] Van Dijk, Eric and Wilke, Henk. 1994. "Asymmetry of wealth and public good provision", *Social Psychology Quarterly*, 57, 352–359.

[24] Van Dijk, Eric and Henk Wilke, Henk, 1995, "Coordination Rules in Asymmetric Social Dilemmas: A Comparison between Public Good Dilemmas and Resource Dilemmas", *Journal of Experimental Social Psychology*, 31(1), 1-27.

[25] Zelmer, Jennifer. 2003. "Linear Public Goods Experiments: A Meta-Analysis". *Experimental Economics*, 6(3), 299-310.

Appendix

A Preliminary results

A.1 Homogeneity

First we have to compare globally the data among treatments to discard the possibility that our results are attributable to sample specificities. We make this test on contributions to the public good and on payoffs of subjects. We compare the first identical sequences to control an eventual effect of the session. We find no significant differences on average contributions between first sequences of EE and EU ($p - value = 0.382$), and between first sequences of UE and UU ($p - value = 0.105$) with a Wilcoxon Mann-Withney test.

And according to Wilcoxon Mann-Withney tests, in sequences with same distribution of endowments, players earn the same payoffs.We find a $p - value = 0.442$ by comparing the first sequences of EE and EU, and a $p - value = 0.195$ by comparing the first sequences of UE and UU.

A.2 Restart effect

We have to check the presence of a restart effect in our experiment. If we don't observe a restart effect in benchmark treatments, we can then make the hypothesis that this effect will also not be present in test treatments and so remaining effect detected in the data after redistribution is attributable to the redistribution as such. This will allow us to perform within tests by comparing both sequences of a same test treatment (before/after redistribution). To test the presence of a restart effect, we compare both sequences of benchmark treatments. We found that restart effect does not exist in our data. By comparing both sequences of the EE treatment, we found a p-value=0.640 with a Wilcoxon signed-rank test, and we found 0.400 for the UU treatment.

But as the restart effect is an increase in period 11, we also compare period 10 and 11 to have a more robust result. With a Wilcoxon signed-rank test, we compare the vector of the 8 groups average contributions in period 10 with the vector of 8 groups average contributions in period 11, and this for our both benchmark treatments. We find no difference for the both treatments with Wilcoxon signed-rank tests, for the treatment EE, we find a p-value of 0.313 and a p-value of 0.106 for the treatment UU.

B baseline treatments

This section complete the results presented in the article. We follow exactly the same structure and we run the same tests on baseline treatments. As these tests are not essential for answering our main questions, we do not present them in the main part of the paper.

Most of the results are in line with previous findings. There is one recurrent exception with the second sequence of UU.

B.1 payoffs

First we make tests on payoffs. We find no payoffs difference between the two sequences of baseline treatments ($p - value = 0.945$ in EE and $p - value = 0.195$ in UU). Then we test if payoffs differences between the both sequences are equal to zero at group level and according to the endowment of players. Our data does not reject the null hypothesis that payoffs difference is equal to zero. Results of these tests are reported in table 9.

[Table 8 about here.]

Then we check if observed payoffs are equal to the predicted values, at group level and for poor and rich subjects. According to our data we accept the null hypothesis except for the second sequence of UU. Results of Wilcoxon tests are reported in table

10. This difference in the second sequence of UU will be explain by the contributions of subjects.

[Table 9 about here.]

In the UU treatment, we find, as in U-sequences of test treatments, that rich subjects earn more than poor subjects ($p-value = 0.004$ in the first sequence and $p-value = 0.011$ in the second sequence) and that this difference do not differ from the predicted one as reported in table 11.

[Table 10 about here.]

B.2 Contributions

In baseline treatment we should not observe any modification of contribution between sequence 1 and sequence 2. As already tested in the restart section, there in no difference at group level. We have compared the observed difference (1.276 in EE and -4.1 in UU on average) to zero and we can not reject the null hypothesis of the equality of the observed difference to zero ($p-value = 0.641$ in EE and $p-value = 0.400$ in UU with Wilcoxon tests). The same results are obtained when we test poor and rich subjects separately: for poor subjects, $p-value = 1$ and for rich subjects, $p-value = 0.293$ when we compare average contributions in the first sequence with average contributions in the second sequence. Then we compare the difference between average contribution of th both sequences to zero. We can not reject the null hypothesis that the difference of contribution is equal to zero for poor ($p-value = 1$) and rich ($p-value = 0.293$) subjects with Wilcoxon tests.

Finally we test the over-contribution: are the observed contributions equal to the one predicted by Nash or higher? In the EE treatment, over-contribution is observed in the both sequences but is significant only in the second one. In the UU treatment, over-contribution is significant in both sequences at group level. These results are in line with the literature and with results in test treatments. But in UU treatment we find that rich over-contribute significantly to the public good, and we did not observe this behavior in U-sequences of test treatments.

[Table 11 about here.]

List of Tables

Table I: : Equality-Equality (EE) and Inequality-Inequality (UU) treatments

	Benchmark 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 II: : Equality-Inequality (EU) and Inequality-Equality (UE) 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)[a]	28=(7,7,7,7)[b]
	Numbers of contributors	2	4
	payoffs	3728=(840,840,1024,1024)	3136=(784,784,784,784)

[a]Consistent with the neutrality result of Warr and the BBV model.
[b]Consistent with the proposition (i) of the theorem 4 of the BBV model.

Table III: average difference of payoffs between both sequences

null hypothesis: observed and predicted payoffs differences between the two sequences are equal
p-values from Wilcoxon tests are in brackets

treatment	**players**	**first sequence**	**second sequence**	**difference**	**predicted difference**	**differences are equal?**
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)

Table IV: average payoffs

null hypothesis: observed and predicted payoff differences between the two sequences are equal
p-values from Wilcoxon tests are in brackets

treatment	players	sequence	observed payoffs	Nash payoffs	observed payoffs are those predicted?
EU	**all**	1st	886.11	784	No (0.040)
		2nd	951.86	932	Yes (0.461)
	poor	1st	897.4	-	-
		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)
		2nd	814.15	-	-
	rich	1st	1051.56	1024	Yes (0.109)
		2nd	872.67	-	-

Table V: payoffs differences of poor and rich subjects in the same sequence

null hypothesis: observed and predicted payoff differences between poor and rich subjects are equal
p-values of Wilcoxon tests are in brackets

sequence[a]	**poor players**	**rich players**	**difference**	**predicted difference**	**equality?**
eU	863.56	1040.15	176.59	184	Yes (0.742)
Ue	849.34	1051.56	202.22	184	Yes (0.400)

[a]the uppercase represents the tested sequence

Table VI: average contribution - over-contribution

null hypothesis: the observed and predicted contributions are equal, if we reject the null hypothesis, we have over-contribution.
p-values of one-sided Wicoxon tests are in brackets

treatment	**sequence**	**players**	**observed contribution**	**Nash contribution**	**over-contribution**
EU	**all**	1st	9.99	7	Yes (0.008)
		2nd	12.09	11	No (0.071)
	poor	1st	-	-	-
		2nd	1.24	0	Yes (0.011)
	rich	1st	-	-	-
		2nd	22.94	22	No (0.578)
UE	**all**	1st	12.03	11	Yes (0.034)
		2nd	8.11	7	No (0.062)
	poor	1st	1.51	0	Yes (0.004)
		2nd	-	-	-
	rich	1st	22.54	22	No (0.466)
		2nd	-	-	-

Table VII: non-contributors

treatment		theoretical non-contributors [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]a non-contributor contributes 0 during the 10 periods of a sequence
[b]a non-contributor contributes 0 at least in 5 periods of a sequence
[c]a non-contributor contributes 0 or 1 in the 10 periods of a sequence

Table VIII: average difference of payoff between both sequences of each treatment

Can we accept the null hypothesis that observed and predicted payoffs differences between the two sequences are equal?
p-values from Wilcoxon tests are in brackets

treatment	**players**	**first sequence**	**second sequence**	**difference**	**predicted difference**	**test**
EE	all	836.8	833.29	-3.51	0	Yes (0.945)
UU	all	987.4	1040.18	52.78	0	Yes (0.195)
	poor	905.32	967.57	62.25	0	Yes (0.25)
	rich	1069.47	1112.79	43.32	0	Yes (0.313)

Table IX: average payoffs at each period

Can we accept the null hypothesis that observed and predicted payoff differences between the two sequences are equal?
p-values from Wilcoxon tests are in brackets

treatment	players	sequence	observed payoffs	Nash payoffs	test
EE	**all**	1st	836.8	784	Yes (0.313)
		2nd	833.29	784	Yes (0.109)
UU	**all**	1st	987.4	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)

Table X: payoffs difference of poor and rich players in the same sequence

Are the observed and predicted payoff differences between poor and rich players equal?
p-values of Wilcoxon tests are in brackets

sequence[a]	**poor players**	**rich players**	**difference**	**predicted difference**	**test**
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

Table XI: average contribution for each sequence

Can we accept the hypothesis that the observed and predicted contributions are equal?
p-values of one-sided Wicoxon tests are in brackets

treatment	sequence	players	observed contribution	Nash contribution	test
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	24.6	22	No (0.029)
		2nd	26.68	22	No (0.008)