Journal of Economic Psychology 97 (2023) 102626

Contents lists available at ScienceDirect

Journal of Economic Psychology

journal homepage: www.elsevier.com/locate/joep

Multilevel public goods game: Levelling up, substitution and crowding-in effects☆

Marco Catola[a], Simone D'Alessandro[b,*], Pietro Guarnieri[b], Veronica Pizziol[c]

[a] School of Business and Economics, Maastricht University, The Netherlands
[b] Department of Economics and Management, University of Pisa, Italy
[c] IMT School for Advanced Studies Lucca, Italy

ARTICLE INFO

JEL classification:
C9
D71
H4

Keywords:
Multilevel public goods game
Online experiment
Efficiency
Social dilemma

ABSTRACT

In an online multilevel public goods experiment, we implement four treatments where we gradually increase the marginal per capita return of the global public good. First, we find evidence of an increase in the contribution to the global good (levelling-up effect). Secondly, subjects fund their higher contribution to the global good by reducing their contribution to the local good (substitution effect) rather than by increasing total contribution, i.e., the sum of their contributions to the local and the global good (marginal crowding-in effect). Moreover, we observe that total contribution increases as a consequence of the mere introduction of the global good (categorical crowding-in effect). Finally, we observe that subjects continue to contribute to both public goods even when they are dominated in terms of costs and returns.

1. Introduction

The Multilevel Public Goods Game (MLPGG) is an experimental design characterised by multiple public goods in a nested structure. Decision makers are assigned to one of several groups and asked to allocate their endowment among their private account, the public good provided only to their group (namely, the local public good), and the public good provided to all the subjects in the game (namely, the global public good).

This design has often been applied to investigate the tension between the individual tendency to favour their own groups (in-group favouritism) and the kind of pro-sociality that leads individuals to contribute to the overall social benefit. This line of research typically acts on group composition to elicit identity. Buchan et al. (2011, 2009) apply the MLPGG to investigate the impact of globalisation on the willingness of nationality-based groups to cooperate at the international level. Chakravarty and Fonseca (2017) study whether contributing to the local public good can be used to exclude members of other groups because of their lack of cooperation or to reward group members for their cooperation. Beekman et al. (2017) induce strong group identity by making groups conflict with each other in a pre-task. Gallier et al. (2019) measure in-group favouritism by eliciting group identity in subjects living in the same region of Germany. Finally, building on the established literature in the public good game (e.g., Martinangeli, 2021), Lange et al. (2022) differentiate between high- and low-endowment local groups to explore the effect of income heterogeneity on contributions.

Furthermore, the manipulation of the marginal per capita return (MPCR) – i.e., the return of a unitary contribution – has allowed scholars to study to what extent changes in the relative efficiency of the local and the global public goods affect contribution decisions

☆ All data and analyses for this study can be accessed at: https://osf.io/mjbc3/?view_only=2cdfef8212984e25b3794e57f4053512. Preregistration #45141 on AsPredicted.org is available upon request.

* Corresponding author.
E-mail address: simone.dalessandro@unipi.it (S. D'Alessandro).

https://doi.org/10.1016/j.joep.2023.102626
Received 18 January 2021; Received in revised form 3 February 2023; Accepted 28 March 2023
Available online 12 April 2023
0167-4870/

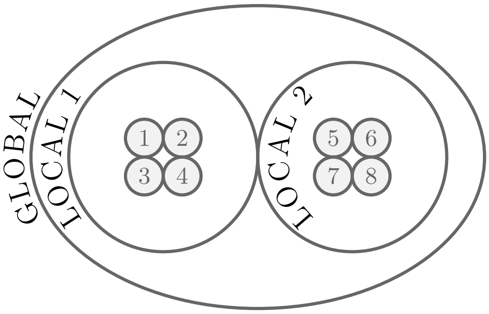


Fig. 1. Structure of a multilevel public goods game.

in the MLPGG. In fact, while it is an established result that an increase in the MPCR has a positive effect on contribution in the standard public good game (Chaudhuri, 2011; Isaac & Walker, 1988; Isaac et al., 1994; Ledyard, 1995; Zelmer, 2003), efficiency changes in the nested structure of the MLPGG entail additional trade-offs with several potential effects that make predictions on contributions less straightforward.

In this study, we exclusively focus on efficiency effects and sterilise group identity by running our experiment online, thus obtaining complete anonymity and excluding any feedback on group composition. The main objective is to add robustness to the evidence collected in the MLPGG literature and systematise the mixed and non-conclusive findings. To this end, we perform a set of treatments which investigate how subjects' allocation decisions are affected by the increase in the relative efficiency of the global public good. In particular, we investigate (i) to what extent this increase levels up the contribution to the global good itself (*levelling-up effect*), (ii) whether it decreases the contribution to the local public good – thus producing a substitution in the allocation between the local and the global goods – (*substitution effect*), (iii) or whether it crowds in the overall amount contributed to the two public goods (*marginal crowding-in effect*). Furthermore, we follow Bowles (2016) and Bowles and Polania-Reyes (2012) and investigate the presence of a *categorical crowding-in effect* by adding a treatment where only the local public good is provided in order to single out the impact on total contribution of the mere addition of the global good.

Our results provide robust evidence of a levelling-up effect. While we find no evidence of marginal crowding in, we observe a decrease in the contribution to the local public good that enables us to confirm the substitution effect. Moreover, the mere introduction of a global public good significantly increases total contribution, thus verifying the categorical crowding-in effect. Finally, we observe that subjects contribute to one of the public goods even when it is dominated by the other public good both in terms of costs and returns. This evidence reinforces the argument that in the context of the MLPGG, subjects' decisions can be inconsistent with the narrow preference for maximising either individual or group payoffs, and may be driven, for instance, by preferences for allocations revealing inequity aversion or fairness criteria.

This paper is organised as follows. Section 2 provides a review of the designs in the MLPGG literature and illustrates our experimental treatments and main hypotheses. Section 3 presents the main results, and Section 4 discusses them by positioning our findings within the context of the MLPGG literature and presents directions for further research based on the limitations of this study.

2. Methods

In the MLPGG framework, subjects are placed both in a local and a global group, the former being nested in the latter to form a hierarchical structure. In fact, the nested structure is what distinguishes the MLPGG from other multiple public goods designs (e.g., Bernasconi et al., 2009; Cherry & Dickinson, 2008; Falk et al., 2013; McCarter et al., 2014). Moreover, an alternative approach to MLPGG design consists in keeping the standard single public good set up while allowing for different spillovers between the local and the global groups (Engel & Rockenbach, 2011; Güth & Sääksvuori, 2012).

We illustrate the specific settings of our design to introduce the main features of the MLPGG structure. As depicted in Fig. 1, we set two local groups of 4 members each, forming a global group of 8.

Each subject has to decide how to allocate an initial endowment of 10 tokens among three alternatives: their private account, a local public good, and a global public good. Every token contributed to the local good is multiplied by a local-specific factor and then redistributed equally among all 4 members of the subject's group, while every token allocated to the global good is multiplied by a global-specific factor and then redistributed equally among the 8 subjects. Finally, the tokens allocated to the private account are simply retained by the subjects.

Given the structure of the game, the payoff of player i is equal to:

$$\pi_i = 10 - c_i - C_i + \alpha \sum_{j=1}^{M} c_j + \beta \sum_{k=1}^{N} C_k. \tag{1}$$

where c is the individual contribution to the local public good, and C is the individual contribution to the global good; α and β are the MPCRs of the local and global public goods, respectively; M and N represent the sizes of the local and global groups, respectively. Also, we will refer to T as the total contribution defined as the sum of c and C.

Table 1
Summary of experimental designs employed to explore changes in relative efficiency in the MLPGG literature. Type: whether the experiment was run in the field or in the lab; α: local MPCRs for each treatment; β: global MPCRs for each treatment; M: number of local group members; N: number of global group members; Group identity: strategy used to manipulate group identity (if present).

Authors	Type	Iterations	α	β	M, N	Group identity
Gallier et al. (2019)	Field	One-shot	0.5	0.25, 0.5	4, 8	Neighbourhood
Chakravarty and Fonseca (2017)	Lab	Repeated	0.4, 0.8	0.4	3, 6	Klee–Kandinsky task
Fellner and Lünser (2014)	Lab	Repeated	0.4	0.2, 0.3	4, 8	No manipulation
Blackwell and McKee (2003)	Lab	Repeated	0.3	0.1, 0.15,	4, 12	Group colours

2.1. Review of related studies

In recent years, several scholars have studied the efficiency effects in the MLPGG (Blackwell & McKee, 2003; Chakravarty & Fonseca, 2017; Fellner & Lünser, 2014; Gallier et al., 2019). However, while the levelling-up effect has been confirmed in all available studies, the debate concerning the substitution and marginal crowding-in effects is far from settled. On the one hand, Blackwell and McKee (2003) do not find any supporting evidence for the substitution effect and conclude that a rise in efficiency increases total contribution. On the other hand, more recent studies (Chakravarty & Fonseca, 2017; Fellner & Lünser, 2014; Gallier et al., 2019) obtain instead a strong substitution effect, which in the case of Chakravarty and Fonseca (2017) and Gallier et al. (2019) fully balances the levelling up, leaving the total contribution unchanged. The differences in the results are accompanied by a high degree of heterogeneity in the experimental designs, involving the manipulation of group identity and of the relative efficiency.[1]

Group identity manipulation serves the purpose of inducing in-group bias in the context of the MLPGG structure. While Blackwell and McKee (2003) apply a minimal identity approach (Tajfel, 1970, 1974, 1982) and Fellner and Lünser (2014) rely on random assignments of individuals to different groups, Chakravarty and Fonseca (2017) implement an endogenous reinforced procedure to form groups and make group identity more salient before subjects play the game. In contrast, Gallier et al. (2019) set up an artefactual field experiment exploiting the fact that participants belong to municipalities within the same region to bring out localism in a natural way. These differences are bound to impact on the efficiency effects as they affect the trade-off between the contribution to the subjects' own group and the global public good differently.

There are also significant differences in terms of efficiency manipulation. Indeed, while Blackwell and McKee (2003) employ four different efficiency treatments, the subsequent studies only rely on two. In particular, both Chakravarty and Fonseca (2017) and Gallier et al. (2019) rely on a simplified design where only two critical treatments are compared. In the first treatment, the MPCRs of the public goods are normalised for group size (i.e., $\beta = \frac{M}{N}\alpha$), while in the second the MPCRs are equal (i.e., $\alpha = \beta$). This experimental setting eliminates the trade-offs between returns, strategic risk and costs, and it is likely to work in favour of a levelling up and against the marginal crowding-in effect. Indeed, in the normalised case, the goods' total returns are equal (as $\alpha M = \beta N$) but the local public good is safer in terms of strategic uncertainty and less costly, thus undermining the incentive to contribute to the global good. Conversely, when $\alpha = \beta$, the two goods are equally costly for the player but the potential returns for the global good are higher, providing a strong incentive to choose the global good. Consider for example the setup of Gallier et al. (2019) with 2 local groups of 4 members. In the first treatment, where $\alpha = 0.5$ and $\beta = 0.25$, the revenue generated by a token contributed to the local public good is twice the revenue generated by a token contributed to the global good but only half of the players enjoys it. In the second treatment, where $\alpha = \beta = 0.5$, the revenue generated by the public goods is the same, but in the case of the global public good, it is enjoyed by all 8 players rather than just 4. Therefore, while the evidence of levelling up obtained by comparing only the two critical cases might be overestimated and hardly generalisable, Chakravarty and Fonseca (2017) and Gallier et al. (2019) are, nevertheless, the only two studies that do not find any evidence of marginal crowding-in.

Table 1 provides a summary of the differences in terms of efficiency treatments and group identity elicitation in the previously mentioned studies.

Scholars have exploited the characteristics of the normalised efficiency treatment mentioned above to test in-group favouritism in the MLPGG setup. Indeed, the two public goods produce the same expected gain (in the case of equal contribution by each local-group member) and, thus, the evidence that people tend to contribute more to the local public good than to the global public good has been interpreted as revealing a bias in favour of the local. This evidence was standard in the MLPGG experiments (Blackwell & McKee, 2003; Chakravarty & Fonseca, 2017; Fellner & Lünser, 2014), up until Gallier et al. (2019) who could not replicate it. However, despite the robustness of this effect across studies, its interpretation is still controversial since the normalised case maintains an imbalance between the two public goods in terms of strategic uncertainty and opportunity cost in the contribution. Chakravarty and Fonseca (2017), for instance, see it as a consequence of the lower degree of strategic uncertainty in cooperation at the local level due to the lower number of players (*size effect*). A similar conclusion is reached by Gallier et al. (2019) who, in reviewing the previous findings, point out that a larger contribution to the local public good in the normalised treatment is not *per se* evidence of parochialism since this may derive from the contribution being responsive to MPCR and irresponsive to group size. The role of strategic uncertainty might also explain why, in Fellner and Lünser (2014), higher returns alone are not sufficient to sustain contribution to the global public good unless they are combined with feedback on the contribution of others.

[1] While the MLPGG is usually implemented in lab setting and with repeated interactions, Gallier et al. (2019) rely on a one-shot field experiment. However this does not seem to account for the differences in the empirical results in terms of the impact of efficiency changes and in-group bias.

Table 2
Summary of treatments' parameters.

Treatment	Local PG			Global PG		
	M	α	TB	N	β	TB
T_0	4	0.6	2.4	–	–	–
T_1	4	0.6	2.4	8	0.15	1.2
T_2	4	0.6	2.4	8	0.3	2.4
T_3	4	0.6	2.4	8	0.45	3.6
T_4	4	0.6	2.4	8	0.6	4.8

Another common result in the literature is that, albeit lower, contribution to the local public good persists even when the MPCRs are equal. This result somehow questions the role of efficiency as the sole driver of contribution. Chakravarty and Fonseca (2017) interpret it as a sign that financial considerations do not totally overcome the effect of (local) group social identity. However, the literature has not yet tested whether the contribution to the global public good persists when there are no financial incentives.

Finally, in a standard PGG, Bernasconi et al. (2009) and Cherry and Dickinson (2008) show that the addition of an identical public good to the players' choice set leads to an increase in total contribution. More recently, Chakravarty and Fonseca (2017) document the presence of the same categorical crowding-in effect in the context of a MLPGG by showing that adding a local public good to an already available global one increases total contribution.

2.2. Our experimental design

The general objective pursued by our pre-registered design is to provide robust evidence of efficiency effects in the MLPGG. Firstly, we investigate the robustness of the levelling up by studying whether the contribution to the global public good always increases whenever its relative efficiency rises. Secondly, we investigate whether such an increase in efficiency produces a marginal crowding in that increases total contribution or induces a substitution with subjects simply shifting their contribution choice between the two public goods.

The review of experimental evidence suggests that results are sensitive to the specific characteristics of the designs. Namely, the variety of strategies adopted to induce group identity might condition the replication of stable tendencies in contribution decisions. Consequently, we opted to avoid any manipulation of group identity in order to minimise its effects on the allocation decisions between the local and the global public good. Accordingly, we provided participants with no group characterisations or feedback on group composition.

Moreover, since the experiment was run online, no other visual reference was available to subjects, thus making it possible to avoid other sources of potential identification. Finally, the decision to implement a one-shot game instead of a repeated one reduces the opportunity for the individuals in the local groups to learn and adopt strategic spillovers across rounds.

In a between-subjects design, we keep α at a fixed value of 0.6 across all treatments, whereas β takes values of 0.15, 0.30, 0.45 and 0.6. Table 2 provides a summary of all the parameters across treatments and, to better clarify the social efficiency of each public good, the value of the total benefit (TB), defined by Gallier et al. (2019) as the individual earnings from a good obtained when every group member makes a one-token contribution to it (i.e., αM and βN respectively).

In line with Blackwell and McKee (2003), treatments involve only the manipulation of β. Specifically, T_2 and T_4 represent the two commonly implemented special cases. On the one hand, T_2 corresponds to the situation where the returns of the public goods are normalised ($\alpha M = \beta N$), thus sterilising any efficiency effect due to scale. Consequently, the local good is less costly and hence less risky, given that the individual return from a token contributed to this public good is higher than the return of a token contributed to the global public good.

Conversely, T_4 corresponds to the opposite case in which marginal returns are equal ($\alpha = \beta$). Therefore, the public goods are equally costly, but the global public good is more efficient because of the scale effect. This feature has two main implications. Firstly, for the individual player, the two public goods are equally risky as the return from the contribution is the same. Secondly, while in T_2 the members of the local group are better off if their fellow member i contributes to the local account rather than to the global one (as $\alpha > \beta$), this is not the case for T_4 (given that α and β are equal). Therefore, contributing to the local public good in T_4 is neither less costly for the contributors nor does it provide higher payoffs for their fellow local group members. Thus, the only difference between the two public goods in T_4 is that contribution to the local public good excludes the members of the other group from the benefit of the public good provision.

Differently, in treatment T_1 – which is a specific novelty of our design – we introduce a global public good that is worse than the local one in all respects. It is more costly – β is lower – and the TB is lower as well. Hence, payoff-wise, there is no incentive to contribute to the global public good, and the decision to contribute may then be motivated by concerns about equity and fairness.

T_3, which is analogous to the treatment used by Fellner and Lünser (2014), is an intermediate case where both the trade-offs of cost and total benefit are present – $\alpha > \beta$ but $\alpha M < \beta N$ – and affect the decision in opposite directions, favouring contribution to the local and to the global public good, respectively. Finally, treatment T_0 is designed to test for the categorical crowding-in effect, given that subjects in this treatment can only contribute to a local public good. Thus, we test our hypothesis by adding a global good to a situation where only the local good is present, and not *viceversa* as in Chakravarty and Fonseca (2017).

Table 3
Sample sizes and participants' average characteristics by treatment. Education is coded as: 1 "no formal qualifications", 2 "secondary education", 3 "high school diploma", 4 "undergraduate degree", 5 "graduate degree", 6 "doctorate degree". Personal income is coded as: 1 "less than 10k", 2 "10–20k", 3 "20–30k", 4 "30–40k", 5 "40–50k", 6 "50–60k", 7 "60–70k", 8 "80–90k", 9 "greater than 90k". Socioeconomic status refers to participants self-reported place on a ladder representing society from 1 to 10.

	N	Age	Male	Income	Student	Soc. status	Edu.	Employed
T_0	164	36.28	0.32	2.59	0.23	5.39	3.68	0.70
T_1	160	35.01	0.31	2.42	0.20	5.31	3.79	0.74
T_2	164	33.89	0.30	2.27	0.26	5.36	3.64	0.70
T_3	160	34.28	0.37	2.59	0.18	5.46	3.72	0.68
T_4	154	34.16	0.30	2.64	0.20	5.32	3.65	0.76

2.3. Hypotheses

Our design enables us to single out three main hypotheses which address the main efficiency effects investigated in the MLPGG literature.

Hypothesis 1 (*Levelling Up*)**.** Average contribution to the global public good $\bar{C}$ is an increasing function of β; i.e. individuals tend to increase their contributions to the global good as its relative efficiency increases.

Hypothesis 2 (*Substitution Effect*)**.** Average contribution to the local public good $\bar{c}$ is a decreasing function of β; i.e. individuals tend to decrease their contributions to the local good as the relative efficiency of the global good increases.

Hypothesis 3 (*Marginal Crowding In*)**.** Average total contribution $\bar{T}$ is an increasing function of β; i.e. individuals tend to increase their overall contributions as the relative efficiency of the global good increases.

As for the categorical crowding in, we formulate the following hypothesis.

Hypothesis 4 (*Categorical Crowding In*)**.** Average total contribution $\bar{T}$ increases as a consequence of the addition of a global good *per se*.

2.4. Implementation

The experiment was implemented using oTree (Chen et al., 2016) and conducted online on the Prolific platform (Palan & Schitter, 2018), which allowed for the recruitment of a socio-demographically varied and well-powered sample with a guarantee of complete anonymity and full randomisation. A total of 802 UK nationals participated in two different sessions. 80 subjects participated in the first session (run as a pilot), and the remaining 722 in the second session.[2] Each subject was randomly assigned to one of the treatments and then to a local and a global group. We succeeded in obtaining sub-samples of almost the same size, although some dropouts led to slight imbalances due to the substitution procedure which randomly assigns new entrants to treatments. Table 3 reports on our sample's size and demographics and shows that the treatment sub-samples were homogeneous in terms of key individual-specific variables confirming that the randomisation of individuals across treatments worked successfully.[3] It is also worth noticing that compared to experiments in the lab, which are the standard in the MLPGG literature, the average age of our participants is notably higher, and the fraction of students is lower, thus making our sample more representative of the actual population.

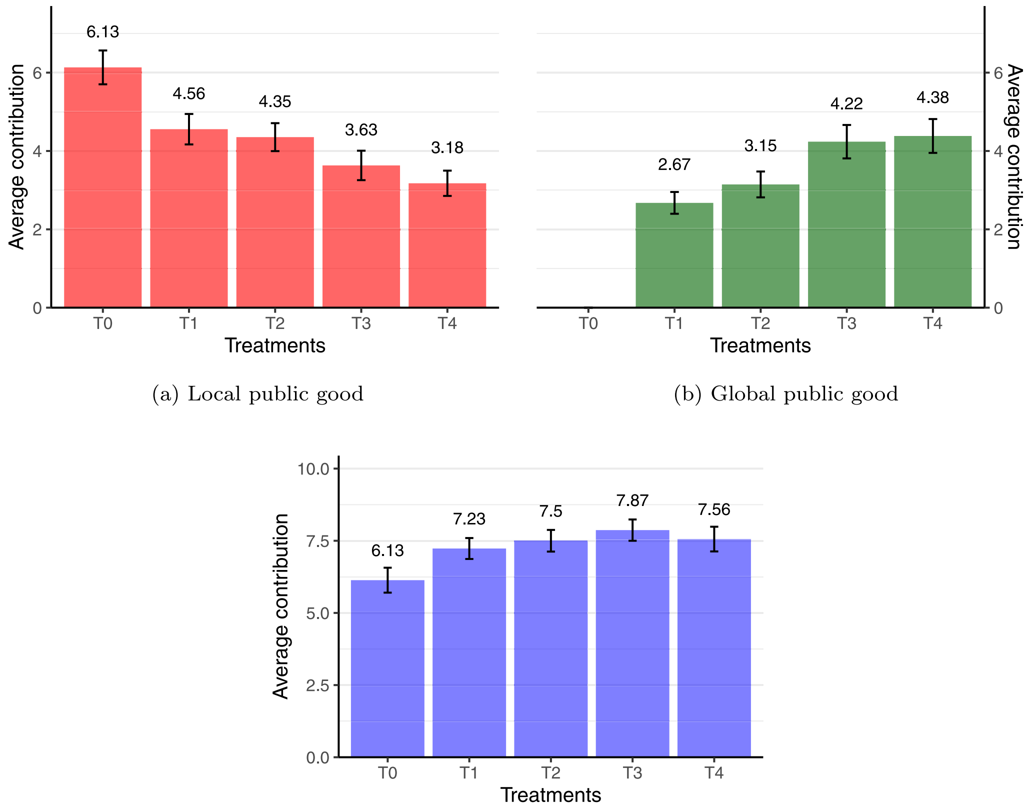
After going through the instructions – available in Online Appendix F –, subjects faced the decision on the main task, i.e., how to allocate their endowment between their personal account, the local public good and the global public good. After the decision task, participants answered questions to measure their empirical expectations, personal normative beliefs, and normative expectations (Bicchieri & Chavez, 2010; Bicchieri & Xiao, 2009).[4] At the end of the experimental questionnaire, subjects replied to three control questions and a 3-item Cognitive Reflection Test in the standard version proposed by Frederick (2005), followed by subjects' elicitation of their social and risk preferences using questions AF.1.2, AF.2.1, AF.3.2, AF.4.3, AF.5.1. and AF.6. from Falk et al. (2018).

Each participant was endowed with 10 tokens and advised in the instructions that tokens would be converted into pounds at the end of the experiment at a given rate (i.e., 1 point corresponding to £ 0.025). Overall, the average payment was £ 1.13 (out of which £ 0.50 were show-up fees).

[2] We aggregated the two sessions because no substantial changes occurred between sessions 1 and 2, and we have chosen out of caution the same time slots and days of the week to launch them.

[3] There is no statistically significant difference across treatments at any level of significance. We performed Kruskal–Wallis tests for the variables: age, income, socioeconomic status and education, and Fisher's tests for the dichotomous variables: gender, student status and employment status.

[4] The effect of norms in shaping contributions is a growing topic in the PGG literature (e.g., Bašić & Verrina, 2021; Catola et al., 2021; Engel & Kurschilgen, 2020; Kandul & Lanz, 2021; Otten et al., 2021). We will investigate the impact of norms in the MLPGG in further research.



(a) Local public good (b) Global public good

(c) Total contribution

Fig. 2. Average contributions per treatment. C.I. at the 95% level.

3. Results

Fig. 2 shows both the average contribution to the local and global goods for each treatment, and the average total contribution. Local and global contributions are always positive across treatments and show opposite trends as β increases (see panels a and b). In contrast, total contribution appears stable between T_1 and T_4, but lower in T_0 (see panel c).

These general trends are only partially confirmed by non-parametric tests of the differences between consecutive treatments. The difference in the global contribution is shown to be statistically significant only in the comparison between T_2 and T_3 (*MW-U* tests, $T_1 - T_2$, $p = 0.0502$; $T_2 - T_3$, $p = 0.0003$; $T_3 - T_4$, $p = 0.3700$). A similar result holds for the contribution to the local public good. Indeed, the decrease in contribution is only statistically significant when moving up from T_2 to T_3 (*MW-U* tests, $T_1 - T_2$, $p = 0.6124$; $T_2 - T_3$, $p = 0.0020$; $T_3 - T_4$, $p = 0.2135$). However, we must note that comparisons between non-consecutive treatments always provide statistically significant differences for contributions both to the local and global public goods. Comparisons across all treatments are provided in Online Appendix B.

Also the non-parametric tests confirm that there is no significant increase in total contribution as β increases from T_1 to T_4 (*MW-U* tests, $T_1 - T_2$, $p = 0.1974$; $T_2 - T_3$, $p = 0.1237$; $T_3 - T_4$, $p = 0.4479$). In contrast, when only a local good is present, the total contribution is lower than in all the other treatments (*MW-U* tests, $p < 0.001$ for each comparison between T_0 and other treatments). It is worth underlining the statistical significance of the comparison between T_0 and T_1, as it shows that the addition of an inefficient public good is enough to increase total contributions.

3.1. Contribution to the local and global public goods

In this subsection, we focus our analysis on the contributions to each public good and leave the study of the total contribution to the following subsection. Accordingly, we exclude the observations of T_0 from this analysis, given that subjects in that treatment do not face the decision on whether (and how much) to contribute to the local or the global good since there is no global public good in T_0.

To test our hypotheses, we perform a set of OLS regressions using β – i.e., the MPCR of the global public good – as the main regressor to estimate the average effect of changes in efficiency on the local and global contributions, respectively. We chose

Table 4
OLS regressions with robust standard errors in parentheses. Columns (1)–(2) show the results from regressions without controls. Columns (3)–(4) show the results from regressions that include control variables. Coefficients of the control variables are available in Online Appendix D. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

	(1) Local contribution	(2) Global contribution	(3) Local contribution	(4) Global contribution
β	−3.245*** (0.541)	4.158*** (0.550)	−3.206*** (0.609)	4.189*** (0.605)
Constant	5.147*** (0.231)	2.052*** (0.197)	2.961*** (0.755)	0.639 (0.782)
Controls	No	No	Yes	Yes
Observations	638	638	525	525
R^2	0.052	0.079	0.105	0.147

the OLS for comparability with the main studies in the literature (see Blackwell & McKee, 2003; Gallier et al., 2019), however, applying Tobit models provides consistent results (see Online Appendix C). Control variables include socio-demographic information collected through Prolific (age, gender, income, socioeconomic status, education, employment status and student status) and a set of individual-specific characteristics about preferences (altruism, patience, risk, trust, negative and positive reciprocity) collected in the post-task questionnaire. We also include a variable to measure the response time in the task, a score variable for correct answers in the Cognitive Reflection Test, and a measure of the performance in three comprehension questions. Given that the task, although simple, entails computational difficulties, we include the individual comprehension score as a control variable, thus allowing for some degree of miscalculation.[5]

Table 4 shows that, on average, the effect of β is positive for the contribution to the global good (levelling-up effect) and negative for the contribution to the local good (substitution effect), thus leading to our first two results.

Result 1 (*Levelling Up*). *Contribution to the global public good on average increases as β increases.*

Result 2 (*Substitution Effect*). *Contribution to the local public good on average decreases as β increases.*

Therefore, Results 1 and 2 indicate that individuals tend to substitute their contribution to the local public good with that to the global good as the relative efficiency of the latter increases. However, the robustness of these results might appear in contrast with the non-parametric tests on the differences in contribution between consecutive treatments presented above. To further investigate this potential limitation, we run an OLS analysis employing treatment dummy variables rather than regressor β (see Online Appendix D). The results confirm that levelling up and substitution do not always occur between successive steps across our treatments. However, we need to consider that the differences in the relative efficiency between consecutive treatments are very small, potentially reducing their impact on changes in contributions. Indeed, when bigger jumps are considered – i.e. comparisons between non-consecutive treatments – the differences in contributions both to the local and global public good are always statistically significant. Therefore, we cannot, in principle, exclude that the lack of statistical significance is due to a lack of power.

To further analyse the relationships between Results 1 and 2, we check, within each treatment, which public good receives the higher average contribution. Fig. 3 shows the difference between the average contribution to the local good and the average contribution to the global good by treatment (i.e., $\bar{c} - \bar{C}$). This difference is positive in T_1 and T_2, whereas in T_3 and T_4 it would appear to be negative. Indeed, in both T_1 and T_2 the average contribution to the local good is significantly higher than the average contribution to the global good (*Wilcoxon signed-rank* tests, $p < 0.001$ in both cases), while the opposite is true for T_3 and T_4, even though this difference is statistically significant only in T_4 (*Wilcoxon signed-rank* tests, $p = 0.1670$ for T_3; $p = 0.0007$ for T_4).

Therefore, we observe that subjects contribute more to the local public good unless the global one yields a higher total benefit. This analysis is connected to the debate concerning the interpretation of the treatment where the total benefits are equal (i.e., T_2). The existing literature finds the same positive difference as in our T_2, with the exception of Gallier et al. (2019). Even if the interpretation of this result, which relies on strategic risk and *size effect* (as proposed by Chakravarty & Fonseca, 2017), seems more suitable in our case, our design does not allow us to exclude that, indeed, in-group bias plays a role. Finally, the trade-off between opportunity cost and potential returns in T_3 may explain why our result differs from the literature. Indeed, while Fellner and Lünser (2014) obtain an average contribution to the global public good that is significantly higher than the average contribution to the local public good, we find that this difference is not statistically different from zero.

[5] In Online Appendix E we consider sub-samples based on the number of correct answers. We find no difference in the results, except for the categorical crowding-in effect, discussed in Section 3.2.

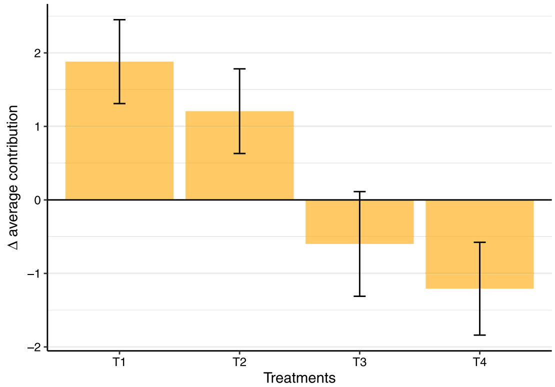


Fig. 3. Difference between average contributions to the local and global goods per treatment. C.I. at the 95% level.

Table 5
OLS regressions with robust standard errors in parentheses. Column (1) shows the results of the baseline specification. Column (2) shows the results of the regression that includes control variables. In T_0, we impute the value of 0 to β. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

	(1) Total contribution	(2) Total contribution
G	1.065***	1.160***
	(0.319)	(0.334)
β	0.914	0.745
	(0.594)	(0.642)
Constant	6.134***	2.168**
	(0.218)	(0.762)
Controls	No	Yes
Observations	802	658
R^2	0.051	0.164

3.2. Total contribution

Results 1 and 2, while questioning the possibility of an increase in the total contribution, cannot rule it out. Indeed, the presence of a reallocation of resources between the local and global public goods does not exclude the possibility of an overall increment in the total amount contributed. To investigate this possibility, we again use regressor β, representing the MPCR of the global public good – which we impute to 0 for T_0 –, and estimate its impact on total contribution to test for the existence of a marginal crowding-in effect. Differently from the analysis in Table 4, however, we add a distinct regressor, G, to identify, if present, a categorical crowding-in effect. G is a dummy variable that is equal to 1 if there is a global public good (hence, for observations in T_1, T_2, T_3, and T_4) and 0 otherwise (hence, for observations in T_0).

Table 5 reports on the results of the regression on total contribution of regressors G and β (Column 1), with the inclusion of control variables (Column 2). We can derive our third and fourth results from this analysis.

Result 3 (*Marginal Crowding In*)**.** *There is no statistically significant evidence of a marginal crowding-in effect.*

Result 4 (*Categorical Crowding In*)**.** *The introduction of an additional global public good produces per se a statistically significant increase in total contribution.*

While introducing an additional global good increases the overall level of contributions, the marginal increase in efficiency is completely ineffective in increasing total contribution. In fact, changes in relative efficiency have only redistributive effects and do not induce subjects to increase their overall contribution.

Our analysis of marginal crowding in does not consider the fact that the total efficiency of overall contribution varies across treatments. To provide further detail on total contribution, we compute an index of relative efficiency (REI) as the ratio between the actual generated public good per treatment and the maximum attainable level per treatment, that is:

$$REI = \frac{\bar{c} \cdot TB_c + \bar{C} \cdot TB_C}{10 \cdot \max\{TB_c, TB_C\}}, \tag{2}$$

Table 6
Relative Efficiency Index (REI) per treatment.

	T_0	T_1	T_2	T_3	T_4
REI	61.3%	58.93%	75.0%	66.6%	59.7%

Table 7
Summary of the main results in the MLPGG literature, including our study.

Authors	Levelling up	Substitution	Marginal crowd. in	Categorical crowd. in
Gallier et al. (2019)	Y	Y	N	–
Chakravarty and Fonseca (2017)	Y	Y	N	Y
Fellner and Lünser (2014)	Y	Y	Y	–
Blackwell and McKee (2003)	Y	N	Y	–
Our study	Y	Y	N	Y

Legend: Y = the effect is found; N = not found; – = not investigated.

where TB_c is the total benefit of the local public good, and TB_C is the total benefit of the global public good (see Table 2). Results are shown in Table 6.[6] By construction, the value of the index in T_0 and T_2 is equal to $1/10$ of the total contribution (as the total benefits cancel out), while it is lower for all other treatments (as $TB_c \neq TB_C$ and the contribution to both public goods is always positive). Therefore, for any given level of the total contribution, T_2 produces the highest relative efficiency because, in terms of efficiency, the two goods are perfect substitutes. As long as players contribute, it does not matter how they allocate their resources since there are no "wrong choices".

The sharp decline in the REI in T_2, T_3 and T_4 is caused specifically by the combination of a lack of marginal crowding in and the persistence of the contribution to the local public good. In other words, as the difference in total benefits between the global and the local public goods increases, subjects throw away the opportunity for a greater total benefit by keeping on contributing to the local good and, at the same time, by not increasing their total contribution. The same reasoning applies to the difference between T_0 and T_1. Subjects choose to partially contribute to the inefficient public good, thus obtaining a total benefit lower than the maximum attainable level even though the total contribution in T_1 is higher than in T_0.

Finally, it is worth mentioning that the presence of the categorical crowding-in effect is the only result that does not hold in the restricted analysis where we select only those participants who perform well in the comprehension questions. Indeed, introducing a relatively inefficient additional public good does not produce a statistically significant increase in total contributions for this category of people, even though their contribution to both public goods remains significantly positive (see the Online Appendix E).

4. Concluding remarks

In an online multilevel public goods experiment, we investigated the effects of changing the MPCR of the global public good on contribution decisions. The general objective was to systematise the evidence and interpretations provided in the literature while adding, at the same time, new insights on some aspects which have either been neglected or not well-understood. In particular, we aimed to shed light on whether and to what extent increasing the MPCR of the global public good induces the levelling up of contribution to the global good and, if this was the case, whether and to what extent this effect is accompanied by a decrease of contribution to the local good – i.e., by the substitution effect – or by an increase in total contributions – i.e., by the marginal crowding-in effect. Moreover, by adding a control treatment where only the local public good is provided, we were able to measure the effect of the mere addition of a global public good *per se* – i.e., the categorical crowding-in effect.

Table 7 summarises the evidence collected in previous studies for each of the effects analysed. We briefly discuss them in the summary of our main findings.

The levelling-up effect is the most robust evidence in the literature, as it has been repeatedly replicated, including in the recent papers by Chakravarty and Fonseca (2017) and Gallier et al. (2019). We confirm this effect with our Result 1 and provide a generalisation by extending the analysis to a series of efficiency increases of the global good, which allowed for an estimation of the average linear effect and many more pairwise comparisons than those usually referred to in the standard literature.

The evidence concerning the substitution effect and the marginal crowding-in effect is much more mixed. For Blackwell and McKee (2003) there is no substitution from the local to the global, but only an increase in the total contribution; Fellner and Lünser (2014) find that both the effects are jointly active following the rise in the productivity of the global good; only Chakravarty and Fonseca (2017) and Gallier et al. (2019) find that substitution cancels out any increase in total contribution. This latter finding is consistent with our Results 2 and 3 as we also observe that as the efficiency of the global public good increases, the levelling-up is financed out of a complete substitution of the contribution to the local public good, thus leaving total contribution unchanged. However, our design offers more robust evidence for both the decrease in the contribution to the local public good and the stability

[6] Non-parametric tests show that these values differ significantly by treatment (*Kruskal–Wallis* test, $p < 0.001$; *MW-U* tests for pairwise comparisons, all p's < 0.001).

of total contribution, which we test at several levels of relative and absolute efficiency. Notably, the decision to sterilise the group identity condition – usually manipulated in the standard multilevel design (one exception being the baseline condition in Gallier et al., 2019) – might have contributed to clearing these results.

With Result 4, we confirm the findings of Cherry and Dickinson (2008) for the standard PGG who show that adding the possibility to contribute to a larger number of public goods brings about a rise in total contribution. Moreover, we produce a new piece of evidence in the context of the MLPGG design. Differently from Chakravarty and Fonseca (2017) who add a local public good to a baseline condition with only a global good, we added a global good to the local good in the baseline. However, we do find the same positive effect on total contribution.

Finally, by looking at the within-treatments analyses, we also confirm several standard results in the literature and provide some novel insights. Firstly, the circumstance that subjects contribute more to the local good until the global good has a higher total benefit confirms a common finding in the MLPGG literature (Blackwell & McKee, 2003; Chakravarty & Fonseca, 2017; Fellner & Lünser, 2014). While we believe that in our anonymous and one-shot setting this result is likely to be explained by the lower strategic uncertainty of the local public good, this interpretation cannot be considered the only plausible one unless a disentanglement of the individual propensity to reduce strategic risk (by opting for the public good where fewer players are involved) is implemented by design. Secondly, we focused on two treatments in which one of the public goods is financially dominated by the other. In the case of T_4, it is the local public good that is (weakly) dominated; in accordance with the literature, we find that, despite the lack of incentives, subjects keep contributing to the local public good. This suggests that allocation criteria other than individual payoff maximisation are at stake. On the other hand, with our T_1, we provide a new test of a condition where the global public good is both riskier and less productive. In this case, the positive and significant contribution to the global public good cannot be justified on the grounds of individual payoff and depends on a specific willingness to contribute to a public good that benefits all players, such as, for example, fairness concerns or inequity aversion. In this sense, it is worth noting that the contribution to the global public good in T_1 and the local public good in T_4 remain significantly positive even for the restricted sample of those with a relatively higher comprehension of the task.

However, the explanation of this kind of decision in terms of some preferences which do not respond to individual or group utility maximisation is beyond the scope of our design and is left for further research. Likewise, additional investigation of the motivations that explain the stability of total contribution is required. It might be the case that a heuristic imposing a stable diversification between one's private account and the total contribution is at stake. However, the validity and robustness of this hypothesis require testing with a dedicated design (e.g., by comparing T_0 with a multilevel setting where more than one public good is added). Moreover, this hypothesis does not apply to the levelling up and substitution effects since, under such invariant automatic heuristics, changes in efficiency could not affect contribution decisions.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data are available in an OSF repository. Link is in the paper.

Acknowledgements

Prof. Simone D'Alessandro gratefully acknowledges financial support from the Italian Ministry of University and Research (PRIN 2017). We are thankful to the participants to the 2021 Annual Conference of the Italian Society of Law and Economics, the 2021 International Online Conference RExCon21 on "Social Preferences and Social Norms", and the University of Pisa and IMT School of Advanced Studies Lucca seminar participants.

We wish to thank two anonymous referees and the editor for their useful suggestions in improving this paper. All remaining errors are our own.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.joep.2023.102626.

References

Bašić, Z., & Verrina, E. (2021). Personal norms—and not only social norms—shape economic behavior. In *MPI collective goods discussion paper (2020/25)*.

Beekman, G., Cheung, S. L., & Levely, I. (2017). The effect of conflict history on cooperation within and between groups: Evidence from a laboratory experiment. *Journal of Economic Psychology*, *63*, 168–183.

Bernasconi, M., Corazzini, L., Kube, S., & Maréchal, M. A. (2009). Two are better than one!. Individuals' contributions to "unpacked" public goods. *Economics Letters*, *104*(1), 31–33.

Bicchieri, C., & Chavez, A. (2010). Behaving as expected: Public information and fairness norms. *Journal of Behavioral Decision Making*, *23*(2), 161–178.

Bicchieri, C., & Xiao, E. (2009). Do the right thing: But only if others do so. *Journal of Behavioral Decision Making*, *22*(2), 191–208.

Blackwell, C., & McKee, M. (2003). Only for my own neighborhood? Preferences and voluntary provision of local and global public goods. *Journal of Economic Behaviour and Organization*, *52*(1), 115–131.
Bowles, S. (2016). *The moral economy: Why good incentives are no substitute for good citizens*. Yale University Press.
Bowles, S., & Polania-Reyes, S. (2012). Economic incentives and social preferences: substitutes or complements? *Journal of Economic Literature*, *50*(2), 368–425.
Buchan, N. R., Brewer, M. B., Grimalda, G., Wilson, R. K., Fatas, E., & Foddy, M. (2011). Global social identity and global cooperation. *Psychological Science*, *22*(6), 821–828.
Buchan, N. R., Grimalda, G., Wilson, R., Brewer, M., Fatas, E., & Foddy, M. (2009). Globalization and human cooperation. *Proceedings of the National Academy of Sciences*, *106*(11), 4138–4142.
Catola, M., D'Alessandro, S., Guarnieri, P., & Pizziol, V. (2021). Personal norms in the online public good game. *Economics Letters*, *207*, Article 110024.
Chakravarty, S., & Fonseca, M. A. (2017). Discrimination via exclusion: An experiment on group identity and club goods. *Journal of Public Economic Theory*, *19*(1), 244–263.
Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Experimental Economics*, *14*(1), 47–83.
Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree-An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97.
Cherry, T. L., & Dickinson, D. L. (2008). Voluntary contributions with multiple public goods. In T. L. Cherry, S. Kroll, & J. F. Shorgen (Eds.), *Environmental economics, experimental methods* (pp. 184–193). London: Routledge.
Engel, C., & Kurschilgen, M. (2020). The Fragility of a Nudge: the power of self-set norms to contain a social dilemma. *Journal of Economic Psychology*, *81*, Article 102293.
Engel, C., & Rockenbach, B. (2011). We are not alone: the impact of externalities on public good provision. In *MPI collective goods preprint, Vol. 2009*. (29).
Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *Quarterly Journal of Economics*, *133*(4), 1645–1692.
Falk, A., Fischbacher, U., & Gächter, S. (2013). Living in two neighborhoods-social interaction effects in the laboratory. *Economic Inquiry*, *51*(1), 563–578.
Fellner, G., & Lünser, G. K. (2014). Cooperation in local and global groups. *Journal of Economic Behaviour and Organization*, *108*, 364–373.
Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42.
Gallier, C., Goeschl, T., Kesternich, M., Lohse, J., Reif, C., & Römer, D. (2019). Leveling up? An inter-neighborhood experiment on parochialism and the efficiency of multi-level public goods provision. *Journal of Economic Behaviour and Organization*, *164*, 500–517.
Güth, W., & Sääksvuori, L. (2012). Provision of multilevel public goods by positive externalities: Experimental evidence. *The B.E. Journal of Economic Analysis and Policy*, *12*(1).
Isaac, R. M., & Walker, J. M. (1988). Group size effects in public goods provision: The voluntary contributions mechanism. *Quarterly Journal of Economics*, *103*(1), 179.
Isaac, R. M., Walker, J. M., & Williams, A. W. (1994). Group size and the voluntary provision of public goods. Experimental evidence utilizing large groups. *Journal of Public Economics*, *54*(1), 1–36.
Kandul, S., & Lanz, B. (2021). Public good provision, in-group cooperation and out-group descriptive norms: A lab experiment. *Journal of Economic Psychology*, *85*, Article 102382.
Lange, A., Schmitz, J., & Schwirplies, C. (2022). Inequality, role reversal and cooperation in multiple group membership settings. *Experimental Economics*, *25*(1), 68–110.
Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. H. Kagel, & A. E. Roth (Eds.), *The handbook of experimental economics* (pp. 111–194). Princeton: Princeton University Press, chapter 2.
Martinangeli, A. F. (2021). Do what (you think) the rich will do: Inequality and belief heterogeneity in public good provision. *Journal of Economic Psychology*, *83*, Article 102364.
McCarter, M. W., Samek, A., & Sheremeta, R. M. (2014). Divided loyalists or conditional cooperators? Creating consensus about cooperation in multiple simultaneous social dilemmas. *Group & Organization Management*, *39*(6), 744–771.
Otten, K., Buskens, V., Przepiorka, W., & Ellemers, N. (2021). Cooperation between newcomers and incumbents: The role of normative disagreements. *Journal of Economic Psychology*, *87*, Article 102448.
Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27.
Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, *223*(5), 96–103.
Tajfel, H. (1974). Social identity and intergroup behaviour. *Social Science Information*, *13*(2), 65–93.
Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, *33*(1), 1–39.
Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics*, *6*(3), 299–310.