universität innsbruck

Faculty of Economics and Statistics

No response to changes in marginal incentives in one-shot public good experiments

Natalie Struwe, Esther Blanco, James M. Walker

Working Papers in Economics and Statistics

2023-08

University of Innsbruck
https://www.uibk.ac.at/EconStat/

University of Innsbruck
Working Papers in Economics and Statistics

The series is jointly edited and published by

- Department of Banking and Finance
- Department of Economics
- Department of Public Finance
- Department of Statistics

Contact address of the editor:
Faculty of Economics and Statistics
University of Innsbruck
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
Tel: + 43 512 507 96136
E-mail: Dean-EconStat@uibk.ac.at

The most recent version of all working papers can be downloaded at
https://www.uibk.ac.at/fakultaeten/volkswirtschaft_und_statistik/forschung/wopec/

For a list of recent papers see the backpages of this paper.

No response to changes in marginal incentives in one-shot public good experiments

Natalie Struwe [a], Esther Blanco [a,b], James M. Walker [b,c]

July 26th, 2023

Abstract: We report novel results from changes in the marginal per capita return (MPCR) in a one-shot public good game where participants make a single provision decision. Data was collected using three data collection processes: an online experiment conducted on Prolific, an online experiment conducted with a subject pool of university students, and an experiment implemented following the conventional procedures of the economic laboratory with university students. In three between-subject treatment conditions, we confront participants from each of these three samples with either a low MPCR of 0.4, a high MPCR of 0.8 holding constant the individual endowment, or a high MPCR of 0.8 reducing the individual endowment to hold constant maximum possible group earnings. Based on a total sample size of 952 participants, we find that, unlike results from previous experiments where subjects make multiple contribution decisions in varying experimental designs, contributions to the public good are not significantly different for the different MPCR conditions we study. We consider these results to be highly relevant in highlighting the limits to our understanding of cooperative behavior for settings without repeated interactions.

JEL classification: C91, C92, H41

Keywords: Voluntary contribution mechanism, Public goods, Marginal per capita return, Social dilemma, Experiments

Acknowledgements: We thank members of the Austrian Science Fund Grant SFB F-63 "Credence Goods: Incentives and Behavior", Ala Avoyan, Tim Cason, Raphael Epperson, Lata Gangadharan, Mark Isaac, Michael Kirchler, Daniela Puzzelo and Andreas Steinmayr for discussions and helpful comments related to this study. We are thankful for the generous support by the Austrian Science Fund, Grant P 32859-G.

Disclosure of interest: none

Corresponding author: James M. Walker, e-mail: walkerj@iu.edu

[a] Department of Public Finance, University of Innsbruck, Universitaetsstrasse 15, 6020 Innsbruck, Austria.

[b] The Ostrom Workshop, Indiana University, USA

[c] Department of Economics, Indiana University, Wylie Hall 105, Bloomington, IN 47405, USA

INTRODUCTION

In the now over forty-year-long tradition of economists experimentally investigating cooperation in laboratory public good settings, one of the robust findings is that individuals respond to changes in marginal contribution incentives. The previous experimental evidence from linear public good environments shows that the provision of public goods increases in the Marginal Per Capita Return (MPCR) first established by Isaac et al., (1984) and since replicated numerous times (see, for example, Isaac & Walker, 1988; Isaac et al., 1994; Fisher et al., 1995; Goeree et al., 2002; Gunnthorsdottir et al., 2007; Carpenter et al., 2009; Reuben & Riedl, 2009; Fischbacher et al., 2014; Nosenzo et al., 2015; ; Lugovskyy et al., 2017; Weimann et al., 2019; Goeschl et al., 2020; for reviews see also Ledyard, 1995; Zelmer, 2003; and Chaudhuri, 2011).[1] This finding has been shown to be robust to numerous alternative experimental designs, including within- and between-subject changes of the MPCR, repeated partner and stranger matching protocols, small or large groups, homogeneous or heterogenous MPCRs within groups, provision or appropriation frames (for the latter, see Blanco, Haller, et al., 2016; Blanco, Lopez, et al., 2016 and Stoddard, 2017), and online populations (van den Berg et al. 2020). In this study we test the limits of the relevance of the MPCR in the provision of public goods by analyzing single contribution decisions.

In pre-registered studies using a linear Voluntary Contribution Mechanism (VCM) (n=952; n= 716 in the main experiments, and n=236 in robustness conditions), we examine decisions in a one-shot decision setting where subjects make a single contribution decision[2], without feedback or interaction among group members. To the best of our knowledge, this study is the first to examine responses to changes in the MPCR in a standard linear public good game where participants make a single contribution decision.[3] We provide novel results for three alternative data collection processes (herein, *samples*),

[1] Related evidence exists from studies considering non-repeated Prisoners' Dilemmas with changes in the monetary payoffs from mutual cooperation, or similarly, the benefit-to-cost ratios, e.g. Charness et al. (2016) show in their between-subjects comparison that higher payoffs for cooperation significantly increase cooperation rates. Further, in their modified PD with continuous transfers between 0 and 10 token, Capraro et al. (2014) find that a substantial amount of participants transfer 50% in all benefit-to-cost-to ratios considered, but there is a shift towards less 0% transfers and more full (100%) transfers for higher multiplication factors. Finally, Gupta et al. (2021) provide a direct comparison of behavior in a laboratory sample (University of Pittsburg) and two online samples (Mturk and Prolific) in Prisoners' Dilemmas varying the marginal incentives to cooperate (similar to the experimental design of Charness et al. (2016), but with a within-subject comparison). One of the main findings is that as opposed to the laboratory sample, the response of Prolific participants to changes in marginal cooperation incentives is "near negligible".

[2] Given that our experiment is conducted in a one-shot/one-decision context, we add to the – relatively small - literature on studying social dilemmas in non-repeated one-shot public good experiments with subjects making a single provision decision (as opposed to experiments using the strategy method). These studies show that for different contexts of one-shot interactions a substantial share of subjects choose to contribute positive amounts to the public good (see, among others, Rondeau, Schulze, & Poe, 1999; Cherry, Kroll, & Shogren, 2005; Kroll, Cherry, & Shogren, 2007; Barcelo & Capraro, 2015; Bilancini et al., 2022).

[3] In a recent and closely related online experiment run via Amazon Mechanical Turk (MTurk), van den Berg et al. (2020) examined the effect of incremental changes in the MPCR on public good provision in a between-subjects design. Notice that while the authors refer to this as one-shot decisions, participants made repeated decisions in randomly re-matched groups of three. The authors find that, on average, contributions increase substantially for MPCRs between 0.4 and 0.7 (from below 40%

including an online experiment using the Prolific platform, an online experiment with a subject pool of university students, and a conventional laboratory experiment with university students. Based on prior studies, our initial conjecture was that for all subject groups we would observe an increase in cooperation with the MPCR.[4] Surprisingly, certainly to us and we believe possibly to other scholars studying behavior in social dilemma situations, we find an *insignificant* effect of changing the MPCR within each of the samples that we study. We consider three treatment conditions varying the MPCR, namely, a *Low MPCR* of 0.4, a *High MPCR* of 0.8 holding constant the individual endowment, and a *High MPCR(lowE)* with MPCR of 0.8 reducing the individual endowment to hold constant maximum possible group earnings' and thus efficiency.

Our sample size was calibrated to detect a difference of 15%-points (0.45 std. deviations) between treatments and control conditions (calculation is based on experimental data from previous linear public good experiments, with conventional significance level of 0.5 and power of 80%). The 15%-points differences across treatments is well within effect sizes to be expected from the previous literature analyzing changes in the MPCR in public goods games. For example, this is similar to the first-period effect size for groups of 4 and MPCR values of 0.3 and 0.75 reported in Isaac et al. (1984) and Isaac & Walker (1988), and less than half of the effect size that van den Berg et al. (2020) find in the first-period comparisons of contributions in their lowest MPCR condition (0.367) to their highest MPCR condition (0.833). Based on our collected data, the minimum detectable treatment effect size for the pooled data across all samples is an 8.19%-points difference between treatment and control conditions, and for each of the three samples (Prolific, students online and lab samples) separately, they are 13.44%-points, 14.31%-points and 14.68%-points, respectively. We assess treatment effects within each sample in terms of average contributions, average expectations of others' contributions, and the distributions of contributions and expectations. We find no significant differences at the conventional 5% significance level in any of these outcome variables.

The lack of a significant and consistent response to variations in the MPCR is therefore not driven by the specific subject population under consideration (students vs. general population), nor the physical proximity of participants (students in the lab vs. students online). Further, when comparing behavior

of endowments to almost 70% of endowments). Their results are qualitatively stable when considering only the first period decisions.

Further, there is recent evidence on non-repeated multi-level public good experiments. Gallier et al. (2019) provide evidence from an artefactual field experiment where participants decided on contributions to a local (including only members from same neighborhood) vs a regional (including members from the same metropolitan area) public good. Increasing the MPCR to the regional public good (while keeping constant the MPCR for the local public good) resulted in a substitution effect between both public goods and no overall increase in efficiency. That is, the authors observe an increase in contributions to the regional public good, but at the expense of contributions to the local one. Catola et al. (2023) confirm such a substitution effect in their online experiment conducted in Prolific.

[4] Pre-registrations available at https://aspredicted.org/HSX_P32 (for the Prolific data collection) and https://aspredicted.org/82N_8DJ (for both data collections at the University of Innsbruck). The one-shot experiments on Prolific were conducted during the pandemic and motivated by the ongoing inability / uncertainty to conduct conventional experiments in the economic laboratory at that time. The results from those experiments motivated collecting the additional data with students at the University of Innsbruck, both online and in the economic laboratory.

within a given treatment and between samples (that is comparing the Prolific sample composed of a general population, the online experiment with same population as in the lab, and the conventional laboratory) contributions are also on average not significantly different. In summary, we provide evidence on the *limits* to the relevance of marginal private costs and marginal social benefits in the voluntary provision of public goods. We consider these results to be central to developing an understanding of the drivers of cooperative behavior in situations where repeated interactions among a group of individuals is not a characteristic of the decision environment.

EXPERIMENTAL DESIGN & PROCEDURES

Participants in our experiment face a one-shot linear public good decision environment. An experimental group consists of $n = 4$ members, where each member receives an endowment of w that can be used to make contributions $g_i \in [0, w]$ to a Group Account of size $G = \sum_{i=1}^{n_i} g_i$. The Group Account constitutes a public good with an equal marginal return (MPCR) of a for all group members, where $\frac{1}{n} < a < 1$, so that the cumulative value of a contribution across all exceeds the marginal cost of a contribution. The payoff function for each group member is given by equation (1).

$$\pi_i = w - g_i + aG \qquad (1)$$

We implement three different treatment conditions – *Low MPCR*, *High MPCR* and *High MPCR (lowE)*, varying the level of the parameter a and the individual endowment of the group members. In the *Low MPCR* treatment $a_{low} = 0.4$; in both *High MPCR* and *High MPCR (lowE)* the MPCR is $a_{high} = 0.8$.[5] Between *Low MPCR* and *High MPCR* we hold constant the endowment $w = 100$ Experimental Currency Units (ECUs). This implies a potential wealth effect at the efficient outcome in *High MPCR* compared to the *Low MPCR* condition, as the maximum (social optimum) payoff in *High MPCR* is 320 ECUs for each participant and in the *Low MPCR* it is 160 ECUs. To account for this, we also implement *High MPCR (lowE)* where all participants receive an endowment of $w = 50$ ECUs, thus holding constant the maximum payoff, as compared to *Low MPCR.*

In all treatments, $a < 1$, such that contributions to the public good constitute a social dilemma in each condition and free-riding incentives exist. Following the *symmetric binary choice approach* as developed in Isaac et al. (1994), one can compute the monetary gains from cooperation comparing either the social optimum outcome (all members contribute fully to the public good) to the Nash Equilibrium outcome assuming fully self-regarded, payoff-maximizing individuals (where all members contribute nothing to the public good), given by equation (2):

[5] Based on findings of van den Berg et al. (2020), we chose the interval for the MPCR of $[0.4, 0.8]$ to allow for a substantial amount of variation in behavior.

$$\pi^{SO} - \pi^{NE} = anw - w = w(an - 1) \quad (2)$$

Table 1 displays the numerical values for the three different MPCR conditions and shows the gains from full cooperation are more than three times higher in the High MPCR condition, than in the Low MPCR condition.

Table 1: Gains from cooperation in the different MPCR conditions.

	Low MPCR	*High MPCR*	*High MPCR (lowE)*
$\boldsymbol{\pi^{SO} - \pi^{NE}}$	60	220	110

Conjecture 1 below is based on these observations and two primary results from the previous literature: i) a well-established positive relationship between contribution rates and MPCR in previous experimental studies (see references in the introduction), and ii) evidence that average contributions in one-shot public good games (holding constant the MPCR) are not significantly different for online vs lab settings (e.g., Buso et al., 2021; Hergueux & Jacquemet, 2015).

> **Conjecture 1:** In all samples, average contributions (in percent of endowment) to the public good will be higher for high MPCR environments of 0.8 (*High MPCR* and *High MPCR (lowE)*) compared to the low MPCR environment of 0.4.

Procedures

The experiment was programmed in oTree (Chen et al., 2016). In all sessions, all participants received the same (treatment specific) instructions, numerical examples and were asked to answer a series of comprehension questions (see the online supplementary materials for the instructions, comprehension questions and screenshots of the experimental program). Two of the comprehension questions required participants to type numerical responses regarding the private value of an ECU and the MPCR ("*Each ECU a participant moves to the Group Account reduces the value of his/her Private Account by ECUs:___*" and "*Each ECU a participant moves to the Group Account generates earnings from the Group Account for each member of his/her group of ECUs:___*"). All participants had to answer the comprehension questions correctly before they could move forward in the experiment. Thereafter, each participant took part in three tasks: (1) an incentivized estimation of the expected behavior of the other members with whom they would be grouped to calculate payoffs [6], (2) their individual contribution to the public good, and (3) a questionnaire containing questions about the main motivations for their decisions and on their donation and volunteering in their day-to-day lives.

[6] Participants were asked to estimate in integer values the average amount of ECUs they expected each of the other participants in their group to contribute to the Group Account. To incentivize informed estimates, participants could earn £1.5 divided by the absolute difference between the actual value and their estimate, up to a maximum of £1.5. This entails a difference of 1 ECU would give the maximum payoff, and an estimate that is exactly the actual value would also give the maximum payoff. Participants were informed in the instructions that they would be asked to give this estimate before decision making.

We chose to elicit beliefs before decision making to provide further assurance that participants thought carefully about how others might interpret the decision opportunities within this particular decision environment and to encourage behavior motivated by expected conditional reciprocity to unfold (given the lack of repeated interactions). Notice that existing studies concerned with the effect of eliciting beliefs of others on own contributions in repeated versions of the public good game with partner matching protocols find conflicting results, depending on the order in which beliefs and decisions are elicited. Gächter & Renner (2010) elicit beliefs after decision making and find that it increases contributions compared to a control treatment without belief elicitation. On the other hand, Croson (2000 and 2007) compare contributions in a treatment with belief elicitation before decision making, to a control of no belief elicitation and find contributions are lower with belief elicitation. Notice though, we are not primarily interested in the *level* of contributions to the public goods for a given treatment, but in the *difference* in the level of contributions between treatments (that is, across different MPCRs). Because we hold constant the order of belief elicitation and contribution decisions in each treatment, our protocol might impact all observed contribution levels, but not have a differential impact on contributions to the public good across our different MPCR treatments (Table A2 in the appendix provides support by the non-significant interaction effects between treatments and expectations on individual contributions). For the purpose of examining the robustness of our results, we collected additional data via Prolific for each of the three MPCR treatments for a sample of n=236 participants (pre-registered here: https://aspredicted.org/SD2_L1N) where we ceteris-paribus changed the order of decision making, such that participants first made their contribution decision and next provide an (incentivized) expectation of the behavior of others. We provide the full set of analysis to these robustness treatments in Appendix B and a short summary of the main results in the Results section.

Each individual made decisions independent of the others in their session. On the decision-making screen, individuals were reminded about the value of the MPCR ("*Remember, for each ECU allocated to the Group Account, all 4 participants in your group will earn 0.4 (0.8) ECUs*"). Participants received no feedback regarding the decisions of the other participants in the group in which they were randomly matched, and only knew that their final payment would be a composite of their earnings from the estimation and decision task as well as the base payment for completing the study. Since each participant made a single contribution decision, without information about others in their group, the individual decisions are independent contribution decisions by each participant.

- Prolific online experiment

Data collection via Prolific took place online in October 2022. A total of 232 participants were recruited in three sessions on two consecutive days. Participants were recruited from the U.K. Prolific population with the requirements that they were fluent in English and had a minimum approval rate of 95% from previous studies. The session was open until the necessary number of participants that finished the experiment was reached. Groups were matched into a subgroup of four according to the arrival time of participants to the waiting page after the control questions. Thereafter, the sum of decisions made by four participants in a matched group was calculated to determine individual payoffs that were sent to participants as bonus rewards via Prolific. The experiment was designed to take about 20 minutes and participants who took longer than 90 minutes were timed out via Prolific automatically.[7] Participants earned on average £5, which included a base payment of £2. Participants could only participate once.

The data collection for the robustness treatments mentioned above took place via Prolific on May 17th, 2023 with a total of 236 participants. Participants first made a contribution decision and then were confronted with the (previously unmentioned) incentivized expectation task. In these treatments, to reduce waiting time for participants, we deleted the waiting-pages after control questions and matched participants into groups of four ex-post, based on their registration time (the time when they entered the experiment). Because matching into groups is only relevant to calculate payments, differences in this technical detail of the procedures cannot impact individual decisions. All other procedures related to our initial experiments were held constant. One participant only gave partial data and therefore this participant and their matched group members had to be dropped from the final sample. Average earnings were £4.80.

– ***Student online experiment***

For the online experiments with the university students, participants from the student subject pool of the EconLab of the University of Innsbruck were recruited into simultaneously running sessions during March 2023 (using the same recruitment platform as for the lab sessions). On the day of the experiment, they received a participant-specific link that allowed them to participate only once, at any time during the day in a pre-specified time window. Participants were made aware that once they started the experiment, they needed to finish it within 90 minutes, or otherwise would be timed out (to follow the same protocol as in Prolific). The matching of four participants into groups was based on their registration time to sign up for the experiment. Payments were made online via PayPal. On average, participants earned 5.30 Euro, including a base payment of 2€.

– Student Laboratory experiment

For the laboratory sessions, participants from the student subject pool of the EconLab of the University of Innsbruck were recruited into 12 sessions during March of 2023. Upon arrival, participants were randomly seated in the lab and instructed to proceed through the experiment on the computer (instructions were not read out loud in order to follow the same protocol as in the online experiments). Participants could raise their hand to ask questions which were answered privately by the experimenter.

[7] Across the three treatments, between the welcome screen and finishing the comprehension questions, 87 participants returned their submission via Prolific. In addition, 9 participants were timed out via the system (did not finish within 90 minutes). These 96 participants were dropped from the study.
Two participants provided only partial data (that is, they did not provide both their belief of others' behavior and own contribution decision) and therefore these participants as well as their matched group members had to be dropped from the final sample.

After all participants finished the experiment, they were matched into groups of four based on their randomly assigned seat number. Payments were then made in cash. On average, participants earned 5.40 Euro, including a base payment of 2€.

RESULTS

Table 2 presents overview statistics of average individual contributions and standard deviations (in % of endowment) for all main treatments pooled across all samples, and for each sample separately. Average contributions are positive (between 39% and 50% of endowments) in all three treatment conditions for all three samples, and within the typical range of values previously reported in the literature.

Table 2: Summary statistics of average individual contributions, in % of endowments, by treatment and sample. Std. deviations in parentheses.

	Pooled	*n*	**Prolific**	*n*	**online**	*n*	**Lab**	*n*
LOW MPCR	**42.28** *(30.60)*	244	**39.58** *(30.05)*	80	**41.9** *(27.66)*	80	**45.23** *(33.74)*	84
HIGH MPCR	**41.63** *(33.46)*	240	**41.57** *(29.48)*	76	**42.36** *(36.91)*	80	**41** *(33.77)*	84
HIGH MPCR (low E)	**47.01** *(31.80)*	232	**42.32** *(26.12)*	76	**50.38** *(35.99)*	80	**48.16** *(32.14)*	76

Note: In the following, all reported p-values for comparisons in this note are from two-sided t-tests. For the pooled sample, there is a 0.65%-point decrease in contributions moving from *Low MPCR* to *High MPCR* (p-value = 0.82) and a 4.73%-points increase moving to *High MPCR (lowE)* (p-value = 0.1). For the Prolific sample, there is a 1.99%-point increase in contributions moving from *Low MPCR* to *High MPCR* (p-value = 0.68), and a 2.74%-points increase moving to *High MPCR (lowE)* (p-value is 0.55). For the student online sample, the difference between *Low MPCR* and *High MPCR* is 0.46%-points (p-value = 0.93) and *High MPCR (lowE)* it is 8.48%-points (p-value = 0.097). Finally, for the laboratory sample, the *Low MPCR* results in slightly higher contributions than the *High MPCR*, with a difference of 4.23%-points (p-value = 0.42), while the *High MPCR (lowE)* results in 2.93%-points higher contributions than *Low MPCR* (p-value = 0.58). In summary, none of these differences are significant at a conventional p-value < 0.05.

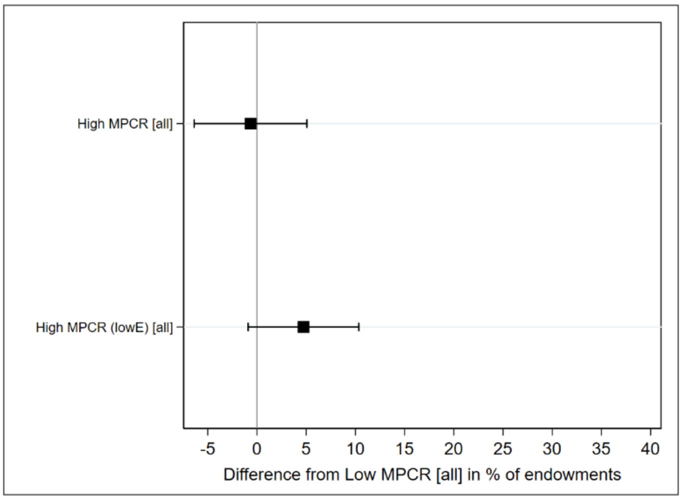

Fig 1: Treatment effects for average individual contributions comparing both *High MPCR* and *High MPCR (low E)* with *Low MPCR*, pooled across all three samples. Point estimates and 95% confidence intervals from OLS regression with clustered standard errors at the group level.

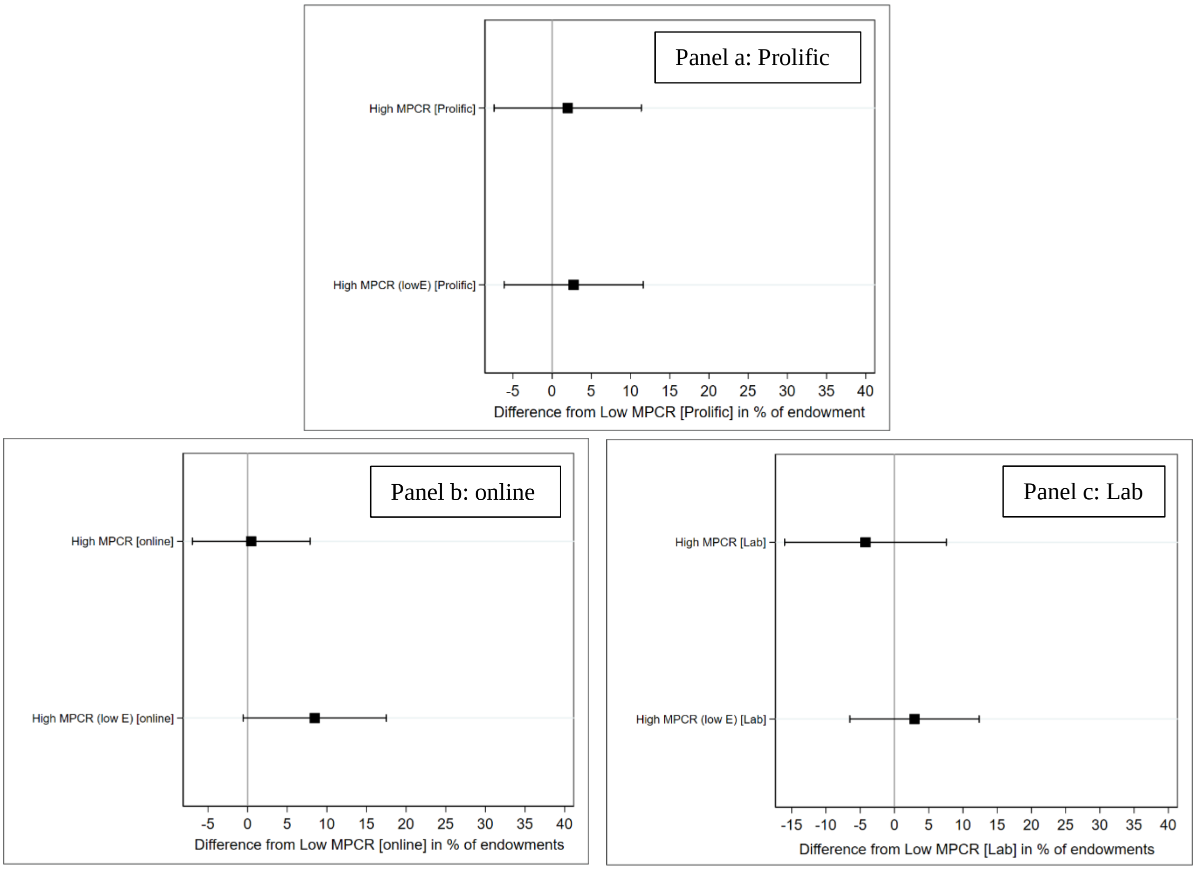

Fig 2: Treatment effects for average individual contributions comparing both *High MPCR* and *High MPCR (low E)* with *Low MPCR*. Point estimates and 95% confidence intervals from OLS regression with clustered standard errors at the group level. **Panel a:** Prolific data collection. **Panel b:** Online data collection with Innsbruck EconLab students. **Panel c:** Data collection in the Innsbruck EconLab.

To test Conjecture 1, we conduct OLS regressions for each sample, with robust standard errors on the subject level.[8] The independent variable is an individual's contribution to the public good in percentage of endowment and the explanatory variable is a dummy variable indicating the treatment condition. Figure 1 presents the coefficient plots for these regressions on treatment effects, pooled across all samples; while Figure 2 presents the same analysis for each sample separately. We find that the differences between contributions are not significant for all treatment comparison under consideration at conventional p-values > 0.05.[9]

[8] Notice that, in Prolific, participants were grouped by arrival time on the wait page. Therefore, in this sample there might be some unobserved differences between participants that are matched faster than others, apart from participants starting the experiment at different points in time. Nonetheless, the grouping was only relevant to calculate payoffs and we do not report analysis on group-level outcomes, which is why we report robust standard errors on the individual level (as opposed to the group level in all samples). Importantly, all regression results are robust to clustering on either group or individual level.

[9] For the pooled sample (Fig 1), the p-value comparing *Low MPCR* and *High MPCR* is 0.82, and for the comparison of *Low MPCR* with *High MPCR (lowE)* it is 0.1. Moving to Fig 2, for Prolific, the p-value comparing *Low MPCR* and *High MPCR* is 0.68, and for the comparison of *Low MPCR* with *High MPCR (lowE)* it is 0.54. For the online sample, the p-value comparing *Low MPCR* and *High MPCR* is 0.93, and for the comparison of *Low MPCR* with *High MPCR (lowE)* it is 0.096. For the lab sample, the p-value comparing *Low MPCR* and *High MPCR* is 0.42, and for the comparison of *Low MPCR* with *High MPCR (lowE)* it is 0.57.

Result 1: *There is no significant increase in cooperation in the single-decision public good game when increasing the MPCR from 0.4 to 0.8 for neither the Prolific data collection, students online, nor students in the economic laboratory for p-values > 0.05.*

In Appendix A, we present further analyses regarding the distribution of individual contributions (Figure A1), the level, distribution and impact of individual's expectations about the behavior of others (Table A1, Figure A2 and Table A2), the distribution of deviations between individual expectations and contributions (Figure A3), as well as the frequency and the impact of the self-reported motivations on individual contributions (Figures A4 and A5). These results show that across treatment conditions and samples, there are no systematic differences in the distributions of individual contributions, the levels of average expectations of others, nor for the distributions of expectations of others. Moreover, paired t-tests show that, on average, individual expectations are not significantly different from individual contributions within each treatment (all respective p-values > 0.05), highlighting that, on average across participants, one's contribution does not differ from ones' expectations of others. Finally, there is no systematic difference in the frequency of self-reported motivation and how motivations affect individual contribution decisions across treatments and samples. Mistrust generally has a negative impact on contributions, while social efficiency considerations increase contributions throughout. The egoistic motives, feeling of responsibility or social norms to contribute, are generally insignificant.

Result 2: *The distributions of contributions, average expectations of others' contributions, and the distributions of expectations of others' contributions are not significantly different across the different treatments or samples for p-values > 0.05.*

These results do not seem motivated by a general inattention by participants resulting in random behavior. First, at the individual level, for all treatment conditions and samples, higher expected contributions of others significantly correlate with higher own contributions (Table A2), consistent with conditional cooperation. The correlation between contributions and expectations is not significantly different when comparing the high and low MPCR environments, and since average expectations are similar in these conditions, contributions to the public good are also similar across conditions. Second, as opposed to purely random choices, the distributions of contributions and expectations are not uniformly distributed (that is, all p-values < 0.0001 from one-sided Kolmogorov-Smirnov tests against uniformly distributed random integer variates on the interval [0,100]). Third, only 2.8% of participants reported not having fully understood the decision task. The results do not change when excluding these participants.

Notice, to have 80% power to estimate such small effect sizes as we see in our data, we would have needed a sample size ranging between 2,300 – 76,426 participants per treatment condition, depending on the sample under consideration (with the exception of the comparison between *Low MPCR* and *High MPCR (lowE)* in the online sample, where 226 participants per treatment would be sufficient to have 80% for the moderate effect size of 8.48%-points).

Finally, the results from our robustness treatments to changes in the order of decisions, using Prolific (see analysis in Appendix B) can be summarized as follows: (i) across the two Prolific samples, there are no significant differences in contribution levels when comparing the treatments with expectations made first and the ones with contributions made first, and (ii) within the additional sample experiments, the comparisons across different MPCR treatments with contribution decisions first, there is a significant difference only between the *Low MPCR* and the *High MPCR (lowE)* treatment at p-value < 0.05. The differences in other comparisons (*Low MPCR* vs *High MPCR,* and *High MPCR* vs *High MPCR (lowE)*) are insignificant. Given the multiplicity of hypotheses we are testing simultaneously in this study, it is appropriate to correct our analysis for multiple hypothesis testing: The adjusted p-values show that indeed no comparison is significantly different at p-value < 0.05 (see analysis and discussion in Table B3 in Appendix B). We conclude from this additional analysis that the order in which participants are confronted with expectations and contribution decisions does not systematically affect either contribution levels or treatment effects. Thus, there is no robust and consistent evidence for a response to changes in the MPCR in the one-decision public good decision environment.

DISCUSSION & CONCLUSION

The striking main finding in this study is that in linear public goods games where participants make a single provision decision, average contributions do not reflect a significant response to changes in the marginal benefit to contribute. Across the three samples (Prolific, students online, or students in the Lab), neither individual contributions nor expectations of others' behavior vary systematically with a change in the MPCR, a change in endowment (holding constant the level of the MPCR), a change in the subject population (general population vs students), or a change in physical distance to group members (online vs lab). These results are in sharp contrast to previous evidence from repeated public good studies and from studies using within-subject designs where participants make multiple decisions across treatment parameters. Thus, we perceive our novel results as highly relevant to the literature studying cooperative behavior in public good settings, providing evidence of the limits to the positive relationship between voluntary public good provision and the MPCR. As such, it advances our understanding of the fundamental components of cooperative behavior.

It is worth emphasizing that the lack of treatment effects is not associated with low cooperation levels. Average contributions to the public good range between 40-50% of endowment across treatment conditions and samples. It is also worth emphasizing that the beliefs of others' contributions were on average largely accurate, with estimates of 39-49% of endowment and on average 93-112% of the actual contribution values. Moreover, expectations positively and robustly predict contributions in all treatments and samples in a similar way. So, across variations in the marginal benefits to contribute and across samples, subjects are conditionally cooperative and have relatively accurate expectations of others' cooperation. Since expectations are similar across treatment conditions and subjects conditionally define their contributions to expectations, they are not found to systematically respond to changes in the MPCR.

The fact that average participant responses in all three samples are highly *insensitive* to changes in the MPCR in the one-decision public goods game that we studied, thus indeed not responding to the higher potential gains from cooperation in the high MPCR environments, calls for additional research aimed at understanding the mechanisms behind this lack of differences in responses. Disentangling such effects is not straightforward. The mechanism behind these results could include, first, that a one-shot interaction does not allow for the development of long-term payoff considerations that Isaac et al. (1994) and Dreber et al. (2014) suggest as a motivation of cooperation in repeated games. Specifically, Isaac et al. (1994) develop a model for expected intertemporal gains from cooperation where individuals are forward-looking and perceive their contributions as meaningful signals to the other members of their group. In every period, each group member contemplates whether to contribute to the public good or to keep their endowment to themselves. Contributions are considered a "successful signal", if payoffs equal at least the safe option of contributing nothing. Isaac et al. (1994) show that the aggregate contributions of the other group members necessary to make an individual's contribution a successful signal decrease with increasing MPCRs (that is, $G_{n-1} = g_i[\frac{1-a}{a}]$). Given our parameterization, in the *Low MPCR* condition, the average contributions of the other three group members must therefore exceed 0.5 times the individual's contribution, whereas in both *High MPCR* conditions, the average contributions of the other three group members must only exceed 0.083 times the individual's contribution. Importantly, though, our non-repeated decision environment resembles the final period of a repeated decision environment, in the sense that there is no future period of interaction. Thus, an individual's contribution does not have any signal value and indeed, in the one-shot game, individuals cannot be forward-looking. Our results can therefore be interpreted as validating the predictions of the model in Isaac et al. (1994). Of course, in the repeated version of the game, subjects have an opportunity to observe decisions of group members prior to the end round, allowing for additional motives such as reciprocity or fairness. In fact, a robust result in repeated versions of the public good game is that there is often a tendency for contributions to decline as the end period nears. However, zero contributions are generally not observed.

Interestingly though, while previous evidence from studies with repeated decision settings suggests that repetition amplifies treatment differences over time, systematic differences in public good provision for different MPCRs can be observed already in the first decision period. For example, Isaac et al., (1984) find a 17%-points difference in the first period between the low (0.3) and high (0.75) MPCR (with adjusted endowments, as in our *High MPCR (lowE)* condition). Further, holding endowments constant between different MPCR levels, Nosenzo et al., 2015 and Lugovskyy et al., 2017 show first-period treatment effects of 30%-points and 20%-points, respectively, for their low (0.3) and high (0.75 and 0.6 respectively) MPCR conditions. Finally, even with a repeated-strangers matching protocol, van den

Berg et al., (2020) present evidence of significant first-round differences between their lowest (0.367) and highest (0.833) MPCR of 30%-points, while also holding constant endowments between treatments.

Second, a related but different mechanism could be associated with the development of reciprocal strategies that support cooperation in repeated public goods games. Participants motivated by reciprocal preferences would only contribute more with higher MPCRs if they expect others to do so. Indeed, a recent study by Gächter & Fages (2023) provide *causal* evidence that higher MPCRs induce higher beliefs about the behavior of others and thus differences in cooperation levels observed are driven by differences in beliefs.[10] The one-shot decision settings we studied did not generate systematic and consistent differences in expectations to develop, but these differences in expectations might appear in a repeated decision setting or even in one-shot settings where subjects are overtly subjected to information or framing that leads them to reflect on how others might respond to differences in the marginal value of the public good, ceteris paribus.[11]

From a broader perspective, the broad group of public good experiments provide us with evidence on the fundamentals of human decision making, and how these fundamentals interact with changes in the decision environment, including the institutional setting. Our novel results presented here provide a critical complement to prior research, much of which focused on repeated settings where group dynamics could be impacted by forward-looking behavior. Our single-decision setting represents a "stark" reality related to when humans find themselves in settings where there is little if any information that helps to shape their expectations and decisions other than what they have internally developed. That is, we capture individual decision-making that is not incentivized to consider future decisions of others. Further, the decision maker is not "prompted" to reflect on how their decisions might differ if they were faced with alternative incentives, where such incentives could impact their expectations of others, which in turn could affect their decision via a social norm of anticipated reciprocity. Those with strong reciprocal norms must therefore turn to what seems reasonable given their expectations of others based on the information set at hand. This leaves the important question of how information provided by

[10] The experiment of Gächter & Fages (2023) was run via Prolific. The authors used a two-stage design where participants had to first decide on contributions to different MPCRs (0.4 and 0.8) via the strategy method. Thereafter participants played a one-shot public good game with direct responses for one of the MPCRs. The authors find that both conditional (stage 1) and direct-response contributions are higher with the higher MPCR. Notice that while stage 2 is indeed the same decision that participants in our experimental setting make, in our study participants do not face (and reflect on) a stage 1 decision where they face a menu of different MPCRs.

[11] Additionally, in repeated games experimenter demand effects and social image concerns might play a larger role, if participants are systematically confronted with feed-back screens reporting to all group members the (aggregate or disaggregated) behavior of its members.

Finally, a note regarding common information in this study. In our lab sessions, instructions were presented to participants on the computer screen and not read aloud by the experimenter. This approach was to make the instructions comparable to the online samples where reading aloud to all participants was not feasible. Note, however, the instructions included the sentence *"All 4 members of a group will receive the same instructions."* Both Isaac & Walker (1988) and Isaac et al. (1994) did not read out instructions aloud. The latter study included additional experiments giving participants much more detailed information regarding maximum and minimum possible earnings than what is commonly presented in public good games. The results were not affected by this change in information presented. Even further, Isaac & Walker (1998) examined a complete information treatment on the symmetry of payoff conditions for all individuals in a group. They found that this change did not affect contributions.

institutions can help shape such expectations. In sum, the results of this study call for future research designed to advance our understanding of the foundations of cooperative behavior and the interrelation with expectations of other's cooperativeness, with an emphasis on settings void of repeated interactions.

REFERENCES

Barcelo, H., & Capraro, V. (2015). Group size effect on cooperation in one-shot social dilemmas. *Scientific Reports*, *5*. https://doi.org/10.1038/srep07937

Bilancini, E., Boncinelli, L., & Celadin, T. (2022). Social value orientation and conditional cooperation in the online one-shot public goods game. *Journal of Economic Behavior and Organization*, *200*, 243–272. https://doi.org/10.1016/j.jebo.2022.05.021

Blanco, E., Haller, T., Lopez, M. C., & Walker, J. M. (2016). The tension between private benefits and degradation externalities from appropriation in the commons. *Journal of Economic Behavior and Organization*, *125*, 136–147. https://doi.org/10.1016/j.jebo.2016.02.008

Blanco, E., Lopez, M. C., & Walker, J. M. (2016). The Opportunity Costs of Conservation with Deterministic and Probabilistic Degradation Externalities. *Environmental and Resource Economics*, *64*(2), 255–273. https://doi.org/10.1007/s10640-014-9868-7

Carpenter, J., Bowles, S., Gintis, H., & Hwang, S. H. (2009). Strong reciprocity and team production: Theory and evidence. *Journal of Economic Behavior and Organization*, *71*(2), 221–232. https://doi.org/10.1016/j.jebo.2009.03.011

Catola, M., D'Alessandro, S., Guarnieri, P., & Pizziol, V. (2023). Multilevel public goods game: Levelling up, substitution and crowding-in effects. *Journal of Economic Psychology*, *97*. https://doi.org/10.1016/j.joep.2023.102626

Charness, G., Rigotti, L., & Rustichini, A. (2016). Social surplus determines cooperation rates in the one-shot Prisoner's Dilemma. *Games and Economic Behavior*, *100*, 113–124. https://doi.org/10.1016/j.geb.2016.08.010

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Experimental Economics*, *14*(1), 47–83. https://doi.org/10.1007/s10683-010-9257-1

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. https://doi.org/10.1016/j.jbef.2015.12.001

Cherry, T. L., Kroll, S., & Shogren, J. F. (2005). The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab. *Journal of Economic Behavior & Organization*, *57*(3), 357–365.

Croson, R. T. A. (2000). Thinking like a game theorist: factors affecting the frequency of equilibrium play. In *Journal of Economic Behavior & Organization* (Vol. 41).

Croson, R. T. A. (2007). Theories of commitment, altruism and reciprocity: Evidence from linear public goods games. *Economic Inquiry*, *45*(2), 199–216. https://doi.org/10.1111/j.1465-7295.2006.00006.x

Dreber, A., Fudenberg, D., & Rand, D. G. (2014). Who cooperates in repeated games: The role of altruism, inequity aversion, and demographics. *Journal of Economic Behavior and Organization*, *98*, 41–55. https://doi.org/10.1016/j.jebo.2013.12.007

Fischbacher, U., Schudy, S., & Teyssier, S. (2014). Heterogeneous reactions to heterogeneity in returns from public goods. *Social Choice and Welfare*, *43*(1), 195–217. https://doi.org/10.1007/s00355-013-0763-x

Fisher, J., Isaac, R. M., Schatzberg, J. W., & Walker, J. M. (1995). Heterogenous Demand for Public Goods: Behavior in the Voluntary Contributions Mechanism. In *Source: Public Choice* (Vol. 85, Issue 4). https://about.jstor.org/terms

Gächter, S., & Fages, D. M. (2023). *Using the Strategy Method and Beliefs to Explain Group Size and MPCR Effects in Public Good Experiments*.

Gächter, S., & Renner, E. (2010). The effects of (incentivized) belief elicitation in public goods experiments. *Experimental Economics*, *13*(3), 364–377. https://doi.org/10.1007/s10683-010-9246-4

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 Behavior and Organization*, *164*, 500–517. https://doi.org/10.1016/j.jebo.2019.05.028

Goeree, J. K., Holt, C. A., & Laury, S. K. (2002). Private costs and public benefits: Unraveling the effects of altruism and noisy behavior. *Journal of Public Economics*, *83*(2), 255–276. https://doi.org/10.1016/S0047-2727(00)00160-2

Goeschl, T., Kettner, S. E., Lohse, J., & Schwieren, C. (2020). How much can we learn about voluntary climate action from behavior in public goods games? *Ecological Economics*, *171*. https://doi.org/10.1016/j.ecolecon.2020.106591

Gunnthorsdottir, A., Houser, D., & McCabe, K. (2007). Disposition, history and contributions in public goods experiments. *Journal of Economic Behavior and Organization*, *62*(2), 304–315. https://doi.org/10.1016/j.jebo.2005.03.008

Gupta, N., Rigotti, L., & Wilson, A. (2021). The Experimenters' Dilemma: Inferential Preferences over populations. *ArXiv Preprint ArXiv:2107.05064.*

Isaac, R. M., & Walker, J. M. (1988). Group size effects in public goods provision: The voluntary contributions mechanism. *Quarterly Journal of Economics*, *103*(1), 179–199. https://doi.org/10.2307/1882648

Isaac, R. M., Walker, J. M., & Thomas, S. H. (1984). Divergent Evidence on Free Riding : An Experimental Examination of Possible Explanations. *Public Choice*, *43*(2), 113–149.

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. https://doi.org/10.1016/0047-2727(94)90068-X

Kroll, S., Cherry, T. L., & Shogren, J. F. (2007). The impact of endowment heterogeneity and origin on contributions in best-shot public good games. *Experimental Economics*, *10*(4), 411–428.

Ledyard, J. O. (1995). Public Goods: A Survey of Experimental Research. *The Handbook of Experimental Economics, Ed. by J. H. Kagel, A. Roth, Chapter 2, Princeton University Press.*, 111–194. https://doi.org/10.1021/ed024p574

List, J. A., Shaikh, A. M., & Xu, Y. (2019). Multiple hypothesis testing in experimental economics. *Experimental Economics*, *22*(4), 773–793. https://doi.org/10.1007/s10683-018-09597-5

Lugovskyy, V., Puzzello, D., Sorensen, A., Walker, J., & Williams, A. (2017). An experimental study of finitely and infinitely repeated linear public goods games. *Games and Economic Behavior*, *102*, 286–302. https://doi.org/10.1016/j.geb.2017.01.004

Nosenzo, D., Quercia, S., & Sefton, M. (2015). Cooperation in small groups: the effect of group size. *Experimental Economics*, *18*(1), 4–14. https://doi.org/10.1007/s10683-013-9382-8

Reuben, E., & Riedl, A. (2009). Public goods provision and sanctioning in privileged groups. *Journal of Conflict Resolution*, *53*(1), 72–93. https://doi.org/10.1177/0022002708322361

Rondeau, D., Schulze, W. D., & Poe, G. L. (1999). Voluntary revelation of the demand for public goods using a provision point mechanism. *Journal of Public Economics*, *72*(3), 455–470. https://doi.org/10.1016/S0047-2727(98)00104-2

Stoddard, B. (2017). Risk in payoff-equivalent appropriation and provision games. *Journal of Behavioral and Experimental Economics* , *69*, 78–82. https://doi.org/10.1016/j.socec.2017.06.002

van den Berg, P., Dewitte, P., Aertgeerts, I., & Wenseleers, T. (2020). How the incentive to contribute affects contributions in the one-shot public goods game. *Scientific Reports*, *10*(1), 8–12. https://doi.org/10.1038/s41598-020-75729-8

Weimann, J., Brosig-Koch, J., Heinrich, T., Hennig-Schmidt, H., & Keser, C. (2019). Public good provision by large groups – the logic of collective action revisited. *European Economic Review*, *118*, 348–363. https://doi.org/10.1016/j.euroecorev.2019.05.019

Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics*, *6*(3), 299–310. https://doi.org/10.1023/A:1026277420119

APPENDIX A: Additional Analysis

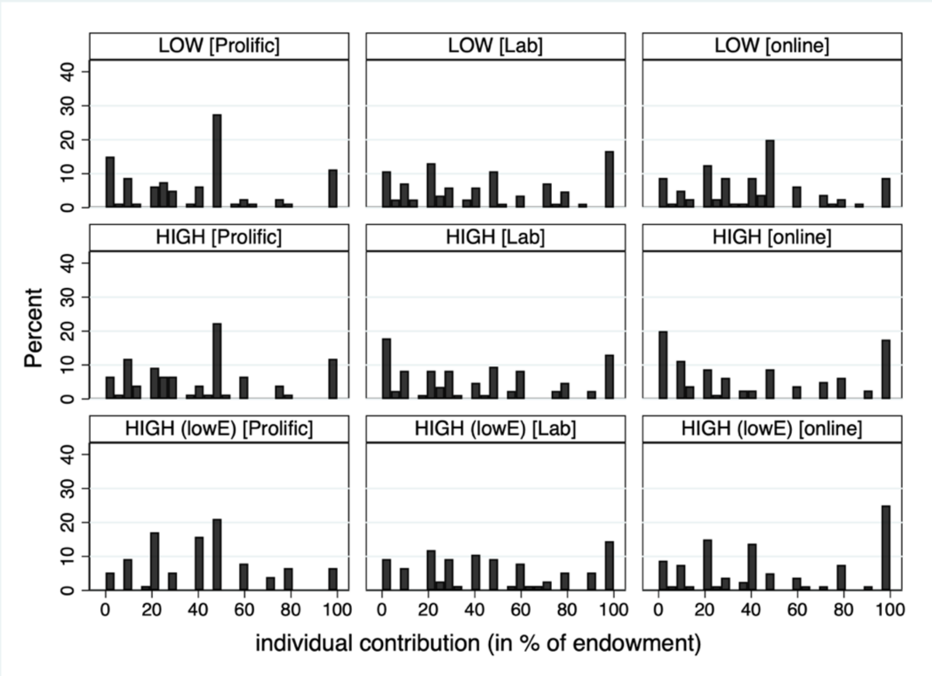

Fig A1: Histogram of individual contributions (in percent of endowment), for each treatment and each data collection process separately.

Note: Kolmogorov-Smirnov tests for equality of distributions suggest that none of the comparisons in the distribution of expectations are significant at a p-value < 0.05 (**Prolific**: p-values for comparison of *Low MPCR* with *High MPCR* and with *High MPCR (lowE)* respectively are 0.95 and 0.65. **Lab**: p-values for comparison of *Low MPCR* with *High MPCR* and with *High MPCR (lowE)* respectively are 0.93 and 0.79. **Online**: p-values for comparison of *Low MPCR* with *High MPCR* and with *High MPCR (lowE)* respectively are 0.17 and 0.054).

Table A1: Summary statistics of average individual expectations, in % of endowments, by treatment and sample. Std. deviations in parentheses.

	Prolific	# *observations*	Lab	# *observations*	online	# *observations*
LOW MPCR	**42** *(22.61)*	80	**47.55** *(26.89)*	84	**43.75** *(20.02)*	80
HIGH MPCR	**41.03** *(24.65)*	76	**41.42** *(26.87)*	84	**47.45** *(27.28)*	80
HIGH MPCR (low E)	**39.66** *(18.87)*	76	**49.39** *(26.04)*	76	**49.08** *(28.67)*	80

Note: For comparison of expectations in *Low MPCR* with *High MPCR* and respectively with *High MPCR (lowE)* we report p-values from two-sample t-tests for each sample separately. P-values for the Prolific comparisons are 0.80 & 0.48. P-values for the online comparisons are 0.33 & 0.18. P-values for the lab comparisons are 0.14 & 0.66.

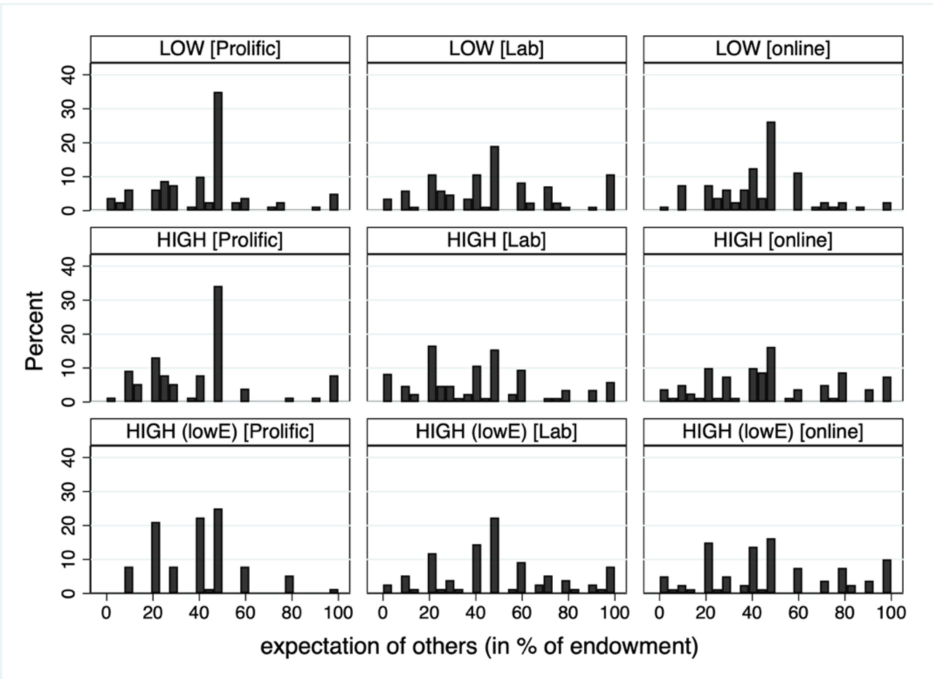

Fig A2: Histogram of individual expectations of behavior of others (in percent of endowment), for each treatment and each data collection process separately.

Note: Kolmogorov-Smirnov tests for equality of distributions suggest that none of the comparisons in the distribution of expectations are significant at a p-value < 0.05. (**Prolific**: p-values for comparison of *Low MPCR* with *High MPCR* and with *High MPCR (lowE)* respectively are 0.81 and 0.41. **Lab**: p-values for comparison of *Low MPCR* with *High MPCR* and with *High MPCR (lowE)* respectively are 0.59 and 0.96. **Online**: p-values for comparison of *Low MPCR* with *High MPCR* and with *High MPCR (lowE)* respectively are 0.24 and 0.12).

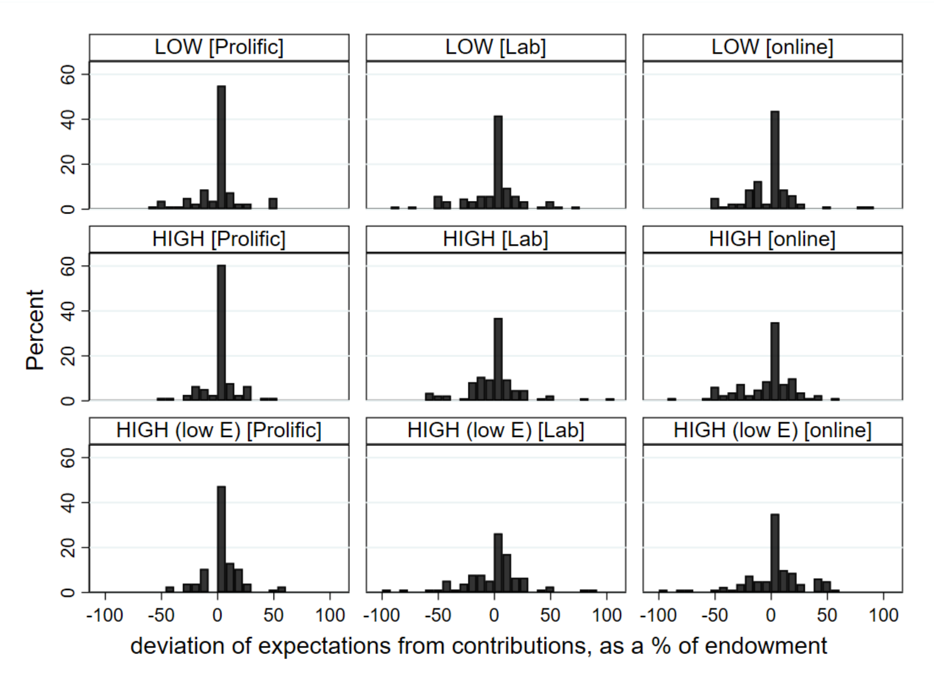

Fig A3: Histogram of difference between individual expectation of behavior of others (in percent of endowment) from individual contribution (in percent of endowment), for each treatment and each data collection process separately.

Note: Kolmogorov-Smirnov tests for equality of distributions suggest that none of the comparisons in the distribution of expectations are significant at a p-value < 0.05. (**Prolific**: p-values for comparison of *Low MPCR* with *High MPCR* and with *High MPCR (lowE)* respectively are 0.92 and 0.42. **Lab**: p-values for comparison of *Low MPCR* with *High MPCR* and with *High MPCR (lowE)* respectively are 0.98 and 0.76. **Online**: p-values for comparison of *Low MPCR* with *High MPCR* and with *High MPCR (lowE)* respectively are 0.56 and 0.33).

Table A2: Correlations of expectations and contributions. Data from OLS regression with robust standard errors at the individual level. ** p<0.005, * p<0.05

	(1)	(3)	(2)
Dep. Var: individual contribution (in % of endowment)	Prolific	Online	Lab
Expectation	0.964**	0.857**	0.774**
	(0.0508)	(0.0740)	(0.0831)
Expectation**High MPCR*	0.0632	-0.00319	0.0236
	(0.0583)	(0.0758)	(0.0795)
Expectation**High MPCR (lowE)*	0.103	0.0575	-0.0169
	(0.0693)	(0.0784)	(0.0869)
Constant	-0.493	3.933	9.003*
	(1.533)	(3.073)	(3.322)
Observations	*232*	*240*	*244*
R-squared	*0.618*	*0.447*	*0.387*

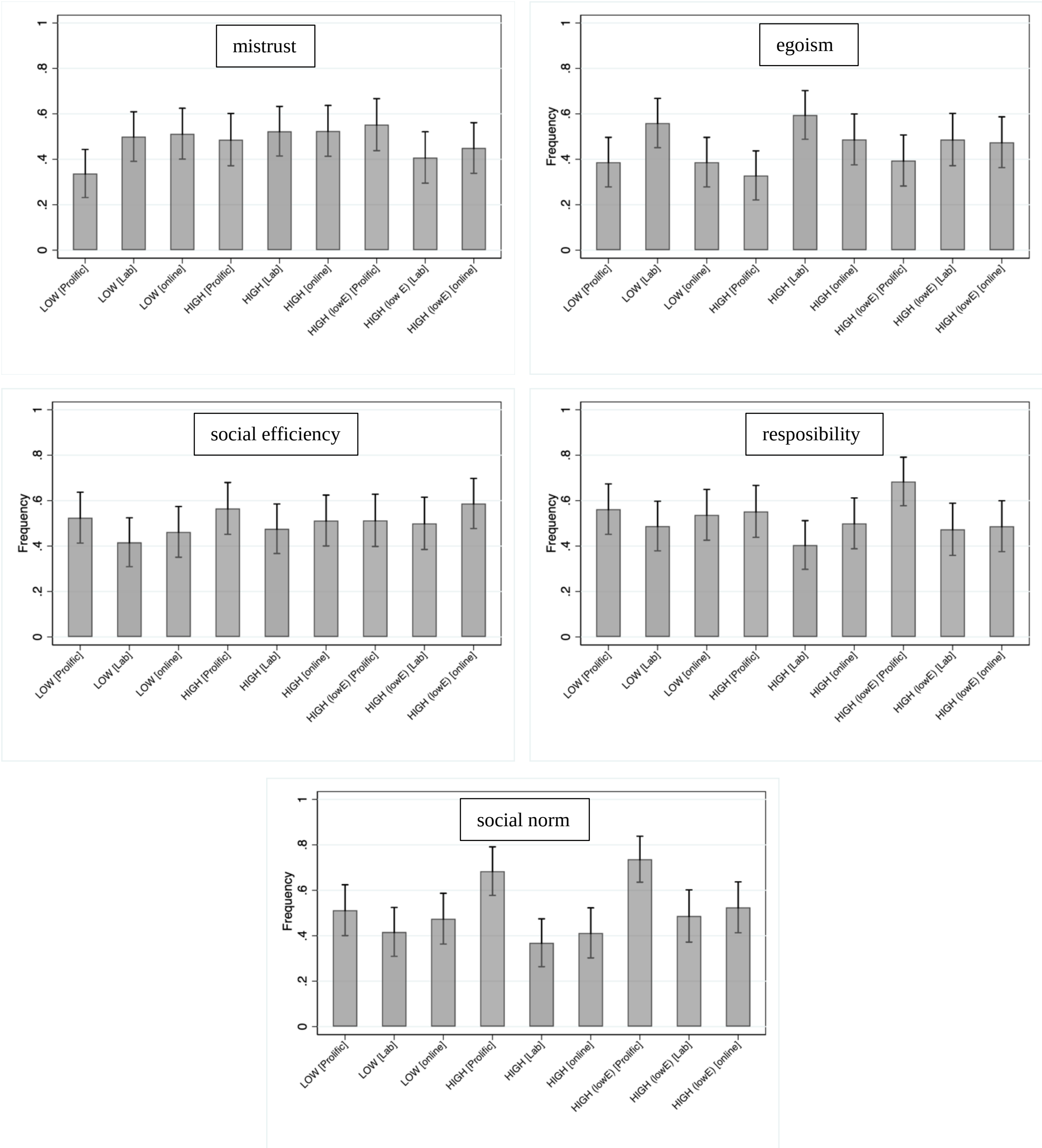

Figure A4: Frequency of self-reported motivations for each treatment and each sample.

Here, we make use of the self-reported motivations from the post-experimental questionnaire. These are measured in 5-Iikert-scale questions, with answers ranging from “I fully agree” … to … “I fully disagree”. For the analysis, these were coded as dummy variables, with individuals receiving a 1 for a given motivation if they answered the question with either “I fully agree” or “I agree”, and 0 otherwise. See the screenshots of the questionnaire in the online SM for the questions related to these motivations.

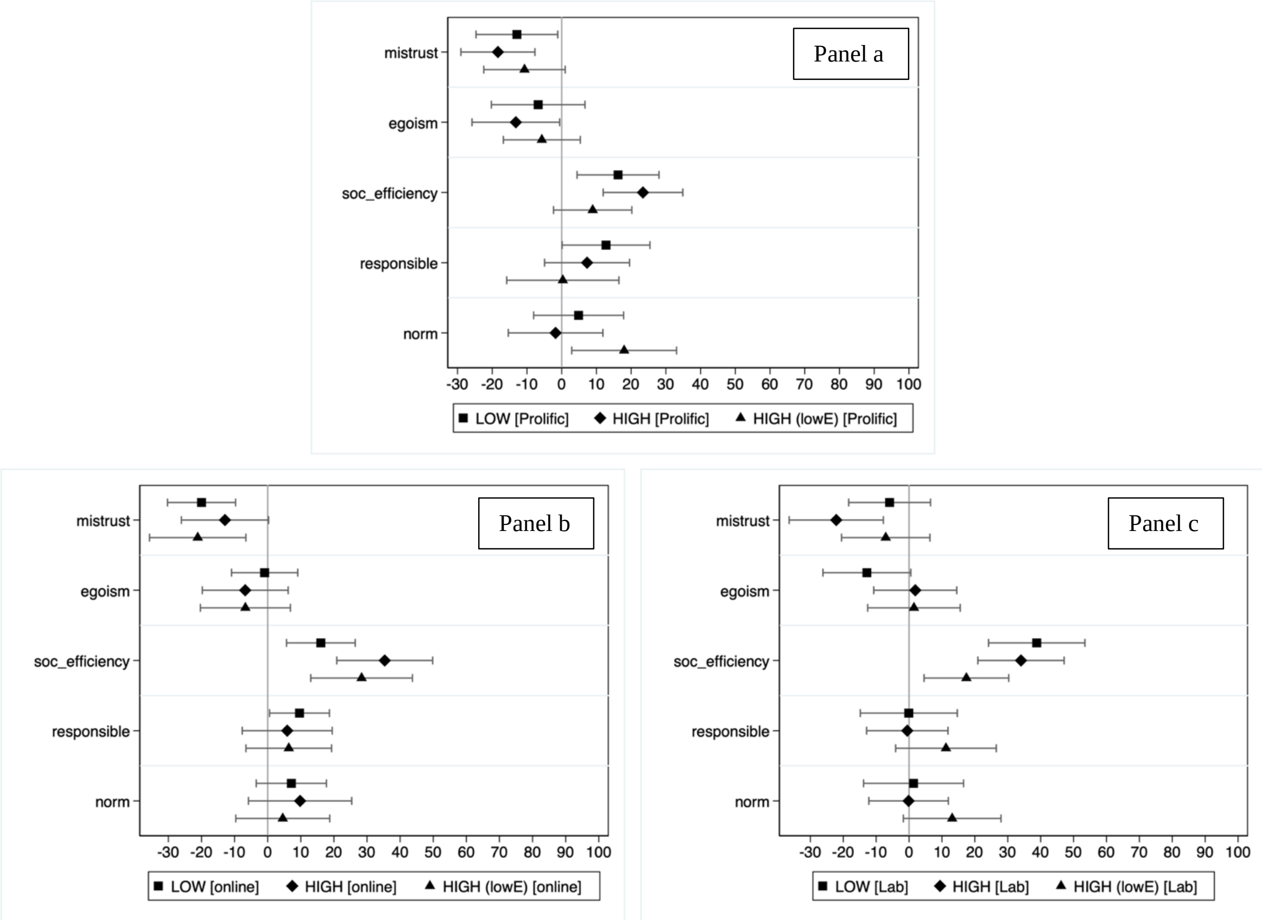

Figure A5: Correlations of self-reported motivations and contributions. Point estimates and 95% confidence intervals from OLS regression with clustered standard errors at the subject level. **Panel a:** Prolific, **Panel B:** students online, **Panel c:** students in laboratory.

APPENDIX B: Robustness tests – Analysis of Reversed Order Treatments

Table B1: Summary statistics of average individual contributions in the Prolific and Prolific *Reversed Order* data collections, in % of endowments. Std. deviations in parentheses.

	Prolific Reversed Order	*# observations*	**Prolific**	*# observations*
LOW MPCR	**32** *(26.25)*	80	**39.58** *(30.05)*	80
HIGH MPCR	**36.71** *(31.47)*	76	**41.57** *(29.48)*	76
HIGH MPCR (low E)	**44.12** *(37.04)*	80	**42.32** *(26.12)*	76

Horizontal comparisons: The difference between the *Low MPCR [RevOrd]* and the *Low MPCR* is 7.58%-points (p-value is 0.092; all reported p-values for comparisons regarding Table B1 are from two-sided t-tests). There is a 4.86%-points difference between *High MPCR [RevOrd]* and *High MPCR* (p-value is 0.33) and a 1.81%-points difference between *High MPCR (lowE) [RevOrd]* and *High MPCR (lowE)* (p-value is 0.71). **Vertical comparisons:** For the Prolific Reversed Order treatments there is a 4.71%-points increase in contributions moving from from *Low MPCR* to *High MPCR* (p-value is 0.31), and a 12.13%-points increase moving to *High MPCR (lowE)* (p-value is 0.018). There ist a 7.41%-points difference between *High MPCR* and *High MPCR (lowE)* (p-value is 0.18).

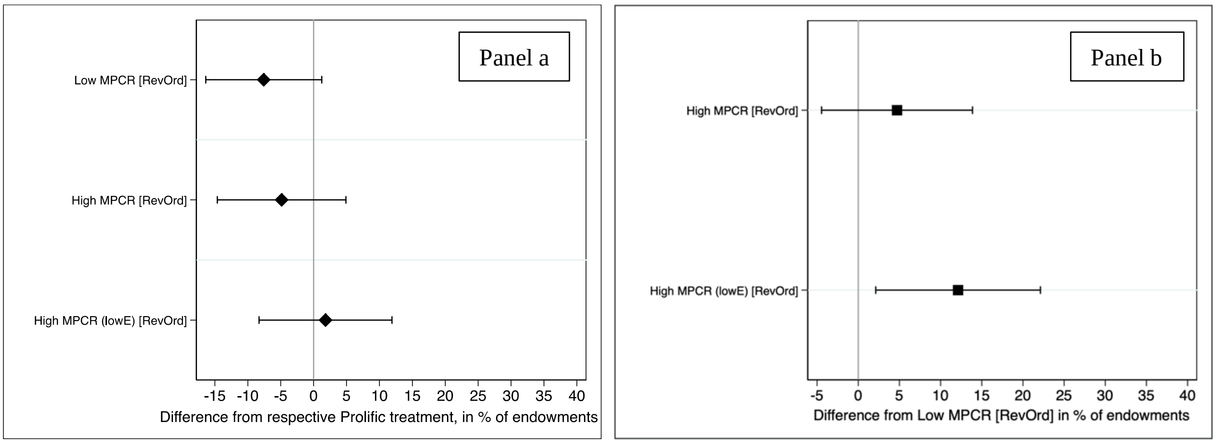

Fig B1. Panel a: Treatment effects from changing order of contributions and expectations; comparing each MPCR treatment with decisions first (marked as *RevOrd*) with its respective counterpart with expectations first. **Panel b:** Treatment effects of the *Reversed Order treatments* for average individual contributions comparing both *High MPCR [RevOrd]* and *High MPCR (low E) [RevOrd]* with *Low MPCR[RevOrd]*. Point estimates and 95% confidence intervals from OLS regression with clustered standard errors at the group level.

Result B1: (i) Within the Prolific population, it does not make a difference whether expectations are elicited before or after contribution decisions: contributions in the treatments with contribution first are not significantly different from the treatments with expectations first. **(ii) There is no systematic and consistent evidence that a higher MPCR with contributions first significantly increases public good provision:** *High MPCR [RevOrd]* does not result in significantly higher contributions as compared to *Low MPCR [RevOrd]*. The difference in contributions between *High MPCR (lowE) [RevOrd]* and *Low MPCR [RevOrd]* is significant at p-value < 0.05, however, contributions in *HighMPCR (lowE) [RevOrd]* and *High MPCR[RevOrd]* are not significantly different from each other (p-value = 0.19).

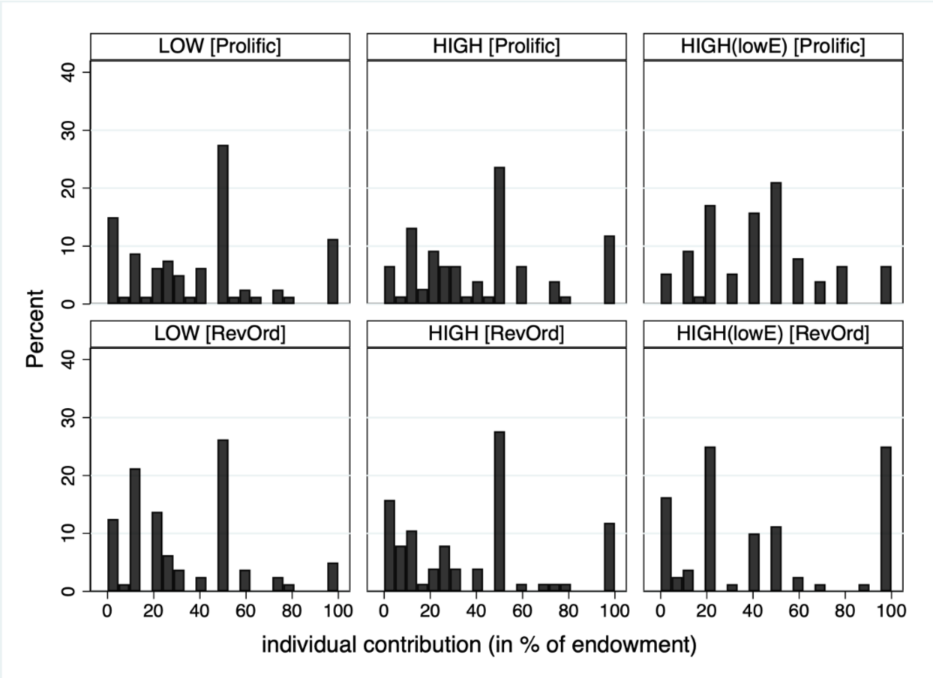

Fig B2: Histogram of individual contributions (in percent of endowment), for each treatment from the Prolific and the Prolific *Reversed Order* data collections separately.

Note: Kolmogorov-Smirnov tests for equality of distributions suggest that none of the comparisons in the distribution of contributions are significant at a p-value < 0.05. (Reversed Order treatments: p-values for comparison of *Low MPCR* with *High MPCR* and with *High MPCR (lowE)* respectively are 0.84 and 0.054. Comparison of each reversed order treatment with its Prolific counterpart of expectations first: p-values for comparison of *Low MPCR [RevOrd]* with *Low MPRC [Prolific]* is 0.24, for the comparison of *High MPCR [RevOrd]* with *High MPCR [Prolific]* it is 0.30 and for the comparison of *High MPCR (lowE) [RevOrd]* with *High MPCR (lowE) [Prolific]* it is 0.098).

Table B2: Summary statistics of average individual expectations in the Prolific and Prolific *Reversed Order* data collections, in % of endowments. Std. deviations in parentheses.

	Prolific Reversed Order	*# observations*	**Prolific**	*# observations*
LOW MPCR	**30.21** *(17.48)*	80	**42** *(22.61)*	80
HIGH MPCR	**35.74** *(21.50)*	76	**41.03** *(24.65)*	76
HIGH MPCR (low E)	**43.65** *(29.00)*	80	**39.66** *(18.87)*	76

Horizontal comparisons: The difference between the *Low MPCR [RevOrd]* and the *Low MPCR* is 11.78%-points (p-value is 0.0003; all reported p-values for comparisons regarding Table B1 are from two-sided t-tests). There is a 5.29%-points difference between *High MPCR [RevOrd]* and *High MPCR* (p-value is 0.16) and a 3.99%-points difference between *High MPCR (lowE) [RevOrd]* and *High MPCR (lowE)* (p-value is 0.31). **Vertical comparisons:** For the Prolific Reversed Order treatments there is a 5.52%-points increase in expectations moving from *Low MPCR* to *High MPCR* (p-value is 0.08), and a 13.44%-points increase moving to *High MPCR (lowE)* (p-value is 0.0005). There is a 7.91%-points difference between *High MPCR* and *High MPCR (lowE)* (p-value is 0.056).

Finally, paired t-tests show that within a treatment, on average, individual contributions are not significantly different from individual expectations (all p-values from paired t-tests > 0.05).

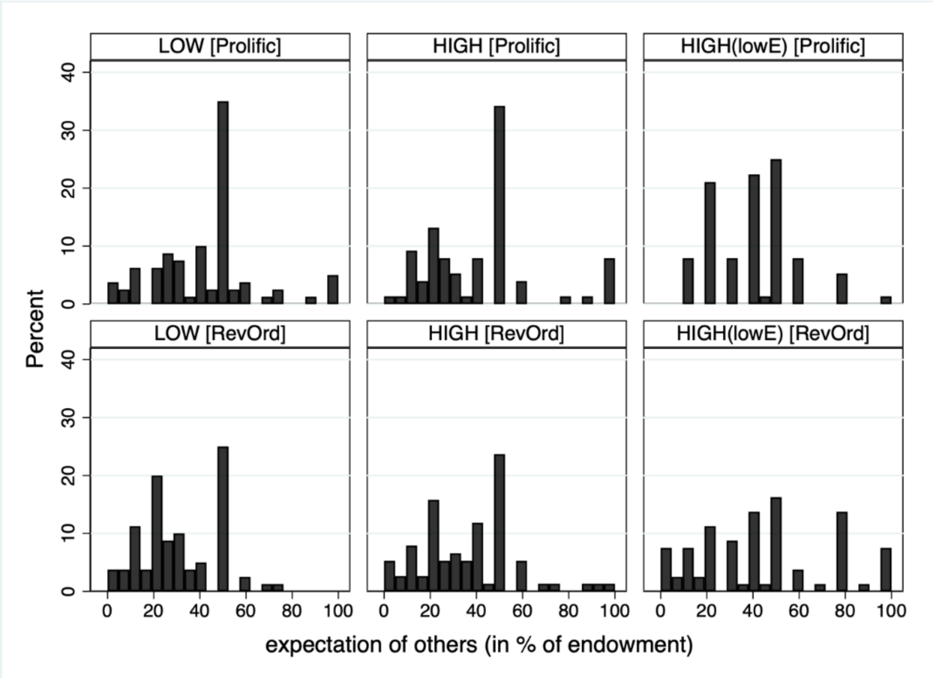

Fig B3: Histogram of expectations of others' contributions (in percent of endowment), for each treatment from the Prolific data collection and the Prolific data collection with *Reversed Order* separately.

Note: Kolmogorov-Smirnov tests for equality of distributions suggest only the comparisons between *Low MPCR* with *Low MPCR [RevOrd]* and with *High MPCR (lowE) [RevOrd]* are significant at a p-value < 0.05. (Prolific *Reversed Order* data collection: p-values for comparison of *Low MPCR [RevOrd]* with *High MPCR [RevOrd]* and with *High MPCR (lowE) [RevOrd]* respectively are 0.24 and 0.022. Comparison of each reversed order treatment with its Prolific counterpart of expectations first: p-values for comparison of *Low MPCR [RevOrd]* with *Low MPRC [Prolific]* is 0.003, for the comparison of *High MPCR [RevOrd]* with *High MPCR [Prolific]* it is 0.53 and for the comparison of *High MPCR (lowE) [RevOrd]* with *High MPCR (lowE) [Prolific]* it is 0.20).

Result B2: (i) No differences in the distribution of contributions between treatments: The distributions of contributions are not significantly different comparing the different MPCR treatments with the reversed order; as well as comparing for each MPCR treatment the different order of decision makings (expectations first vs contributions first). **(ii) Some, but no systematic differences in expectations about the average behavior of others**: average expectations as well as the distribution of expectation in the *Low MPCR [RevOrd]* treatment are significantly different from its counterpart *Low MPCR* in Prolific and from *High MPCR (lowE) [RevOrd]*. All other pair-wise comparisons of expectations and distributions of expectations are insignificant.

Multiple hypothesis correction

Since we are interested in the differences in contributions between the three low and high MPCR treatments for each sample, this gives eight main comparisons that we are testing simultaneously. This relatively large number of simultaneous hypotheses tests increases the probability of false rejections of the null hypothesis and justifies correcting our analysis for the multiplicity of tested null hypotheses. Table B3 presents results for the p-value adjustments due to multiple hypothesis testing.

In column II we give for each sample and each MPCR comparison the difference in contributions in percentage of endowments. Column III shows the unadjusted (original) p-values. The last three columns show the adjusted p-values using different testing procedures for multiple hypothesis testing. The first is the procedure introduced in List et al. (2019), which asymptotically controls the familywise error rate (which gives the probability of one or more false rejections when testing multiple null hypotheses). We additionally report p-values adjusted according to the classical Bonferroni and Holm corrections.

Table B3: Correction for multiple hypothesis testing.

				Adjusted p-values		
		Difference (in % of endowment)	***Unadjusted p-value***	***List et al (2019)***	***Bonf***	***Holm***
Prolific	*Low vs High MPCR*	1.99	0.68	0.90	1	1
	Low vs High MPCR (lowE)	2.74	0.54	0.95	1	1
Online	*Low vs High MPCR*	0.46	0.93	0.93	1	0.92
	Low vs High MPCR (lowE)	8.48	0.098	0.49	0.78	0.69
Lab	*Low vs High MPCR*	4.23	0.42	0.93	1	1
	Low vs High MPCR (lowE)	2.93	0.58	0.93	1	1
Prolific (RevOrd Robustness)	*Low vs High MPCR*	4.71	0.31	0.87	1	1
	Low vs High MPCR (lowE)	12.13	0.02	0.11	0.12	0.12

University of Innsbruck - Working Papers in Economics and Statistics
Recent Papers can be accessed on the following webpage:

https://www.uibk.ac.at/eeecon/wopec/

2023-08 **Natalie Struwe, Esther Blanco, James M. Walker:** No response to changes in marginal incentives in one-shot public good experiments

2023-07 **Sebastian Bachler, Sarah Lynn Flecke, Jürgen Huber, Michael Kirchler, Rene Schwaiger:** Carbon Pricing, Carbon Dividends and Cooperation: Experimental Evidence

2023-06 **Elisabeth Gsottbauer, Michael Kirchler, and Christian König-Kersting:** Climate Crisis Attitudes among Financial Professionals and Climate Experts

2023-05 **Xiaogeng Xu, Satu Metsälampi, Michael Kirchler, Kaisa Kotakorpi, Peter Hans Matthews, Topi Miettinen:** Which income comparisons matter to people, and how? Evidence from a large field experiment

2023-04 **Tanja Hoertnagl, Rudolf Kerschbamer, Regine Oexl, Rudi Stracke, and Uwe Sunde:** Heterogeneity in Rent-Seeking Contests with Multiple Stages: Theory and Experimental Evidence

2023-03 **Melissa Newham, Marica Valente:** The Cost of Influence: How Gifts to Physicians Shape Prescriptions and Drug Costs

2023-02 **Natalie Struwe, Esther Blanco, James M. Walker:** Determinants of Financial Literacy and Behavioral Bias among Adolescents

2023-01 **Marco Aschenwald, Armando Holzknecht, Michael Kirchler, Michael Razen:** Determinants of Financial Literacy and Behavioral Bias among Adolescents

2022-20 **Silvia Angerer, Daniela Glätzle-Rützler, Philipp Lergetporer, and Thomas Rittmannsberger:** Beliefs about social norms and (the polarization of) COVID-19 vaccination readiness

2022-19 **Edward I. Altman, Marco Balzano, Alessandro Giannozzi, Stjepan Srhoj:** Revisiting SME default predictors: The Omega Score

2022-18 **Johannes Diederich, Raphael Epperson, Timo Goeschl:** How to Design the Ask? Funding Units vs. Giving Money

2022-17 **Toman Barsbai, Vojtěch Bartoš, Victoria Licuanan, Andreas Steinmayr, Erwin Tiongson, and Dean Yang:** Picture This: Social Distance and the Mistreatment of Migrant Workers

2022-16 **Andreas Steinmayr, Manuel Rossi:** Vaccine-skeptic physicians and COVID-19 vaccination rates

2022-15 **Stjepan Srhoj, Alex Coad, Janette Walde:** HGX: The Anatomy of High Growth Exporters

2022-14 **Martin Obradovits, Philipp Plaickner** Price-Directed Search, Product Differentiation and Competition

2022-13 **Utz Weitzel, Michael Kirchler** The Banker's Oath And Financial Advice

2022-12 **Julian Granna, Wolfgan Brunauer, Stefan Lang:** Proposing a global model to manage the bias-variance tradeoff in the context of hedonic house price models

2022-11 **Christoph Baumgartner, Stjepan Srhoj and Janette Walde:** Harmonization of product classifications: A consistent time series of economic trade activities

2022-10 **Katharina Momsen, Markus Ohndorf:** Seller Opportunism in Credence Good Markets ? The Role of Market Conditions

2022-09 **Christoph Huber, Michael Kirchler:** Experiments in Finance ? A Survey of Historical Trends

2022-08 **Tri Vi Dang, Xiaoxi Liu, Florian Morath:** Taxation, Information Acquisition, and Trade in Decentralized Markets: Theory and Test

2022-07 **Christoph Huber, Christian König-Kersting:** Experimenting with Financial Professionals

2022-06 **Martin Gächter, Martin Geiger, Elias Hasler:** On the structural determinants of growth-at-risk

2022-05 **Katharina Momsen, Sebastian O. Schneider:** Motivated Reasoning, Information Avoidance, and Default Bias

2022-04 **Silvia Angerer, Daniela Glätzle-Rützler, Philipp Lergetporer, Thomas Rittmannsberger:** How does the vaccine approval procedure affect COVID-19 vaccination intentions?

2022-03 **Robert Böhm, Cornelia Betsch, Yana Litovsky, Philipp Sprengholz, Noel Brewer, Gretchen Chapman, Julie Leask, George Loewenstein, Martha Scherzer, Cass R. Sunstein, Michael Kirchler:** Crowdsourcing interventions to promote uptake of COVID-19 booster vaccines

2022-02 **Matthias Stefan, Martin Holmén, Felix Holzmeister, Michael Kirchler, Erik Wengström:** You can't always get what you want–An experiment on finance professionals' decisions for others

2022-01 **Toman Barsbai, Andreas Steinmayr, Christoph Winter:** Immigrating into a Recession: Evidence from Family Migrants to the U.S.

2021-32 **Fanny Dellinger:** Housing Support Policies and Refugees' Labor Market Integration in Austria

2021-31 **Albert J. Menkveld, Anna Dreber, Felix Holzmeister, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Sebastian Neusüss, Michael Razen, Utz Weitzel and et al:** Non-Standard Errors

2021-30 **Toman Barsbai, Victoria Licuanan, Andreas Steinmayr, Erwin Tiongson, Dean Yang:** Information and Immigrant Settlement

2021-29 **Natalie Struwe, Esther Blanco, James M. Walker:** Competition Among Public Good Providers for Donor Rewards

2021-28 **Stjepan Srhoj, Melko Dragojević:** Public procurement and supplier job creation: Insights from auctions

2021-27 **Rudolf Kerschbamer, Regine Oexl:** The effect of random shocks on reciprocal behavior in dynamic principal-agent settings

2021-26 **Glenn E. Dutcher, Regine Oexl, Dmitry Ryvkin, Tim Salmon:** Competitive versus cooperative incentives in team production with heterogeneous agents

2021-25 **Anita Gantner, Regine Oexl:** Respecting Entitlements in Legislative Bargaining - A Matter of Preference or Necessity?

2021-24 **Silvia Angerer, E. Glenn Dutcher, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** The formation of risk preferences throughsmall-scale events

2021-23 **Stjepan Srhoj, Dejan Kovač, Jacob N. Shapiro, Randall K. Filer:** The Impact of Delay: Evidence from Formal Out-of-Court Restructuring

2021-22 **Octavio Fernández-Amador, Joseph F. Francois, Doris A. Oberdabernig, Patrick Tomberger:** Energy footprints and the international trade network: A new dataset. Is the European Union doing it better?

2021-21 **Felix Holzmeister, Jürgen Huber, Michael Kirchler, Rene Schwaiger:** Nudging Debtors to Pay Their Debt: Two Randomized Controlled Trials

2021-20 **Daniel Müller, Elisabeth Gsottbauer:** Why Do People Demand Rent Control?

2021-19 **Alexandra Baier, Loukas Balafoutas, Tarek Jaber-Lopez:** Ostracism and Theft in Heterogeneous Groups

2021-18 **Zvonimir Bašić, Parampreet C. Bindra, Daniela Glätzle-Rützler, Angelo Romano, Matthias Sutter, Claudia Zoller:** The roots of cooperation

2021-17 **Silvia Angerer, Jana Bolvashenkova, Daniela Glätzle-Rützler, Philipp Lergetporer, Matthias Sutter:** Children's patience and school-track choices several years later: Linking experimental and field data

2021-16 **Daniel Gründler, Eric Mayer, Johann Scharler:** Monetary Policy Announcements, Information Schocks, and Exchange Rate Dynamics

2021-15 **Sebastian Bachler, Felix Holzmeister, Michael Razen, Matthias Stefan:** The Impact of Presentation Format and Choice Architecture on Portfolio Allocations: Experimental Evidence

2021-14 **Jeppe Christoffersen, Felix Holzmeister, Thomas Plenborg:** What is Risk to Managers?

2021-13 **Silvia Angerer, Daniela Glätzle-Rützler, Christian Waibel:** Trust in health care credence goods: Experimental evidence on framing andsubject pool effects

2021-12 **Rene Schwaiger, Laura Hueber:** Do MTurkers Exhibit Myopic Loss Aversion?

2021-11 **Felix Holzmeister, Christoph Huber, Stefan Palan:** A Critical Perspective on the Conceptualization of Risk in Behavioral and Experimental Finance

2021-10 **Michael Razen, Alexander Kupfer:** Can increased tax transparency curb corporate tax avoidance?

2021-09 **Changxia Ke, Florian Morath, Anthony Newell, Lionel Page:** Too big to prevail: The paradox of power in coalition formation

2021-08 **Marco Haan, Pim Heijnen, Martin Obradovits:** Competition with List Prices

2021-07 **Martin Dufwenberg, Olof Johansson-Stenman, Michael Kirchler, Florian Lindner, Rene Schwaiger:** Mean Markets or Kind Commerce?

2021-06 **Christoph Huber, Jürgen Huber, and Michael Kirchler:** Volatility Shocks and Investment Behavior

2021-05 **Max Breitenlechner, Georgios Georgiadis, Ben Schumann:** What goes around comes around: How large are spillbacks from US monetary policy?

2021-04 **Utz Weitzel, Michael Kirchler:** The Banker's Oath And Financial Advice

2021-03 **Martin Holmen, Felix Holzmeister, Michael Kirchler, Matthias Stefan, Erik Wengström:** Economic Preferences and Personality Traits Among Finance Professionals and the General Population

2021-02 **Christian König-Kersting:** On the Robustness of Social Norm Elicitation

2021-01 **Laura Hueber, Rene Schwaiger:** Debiasing Through Experience Sampling: The Case of Myopic Loss Aversion.

University of Innsbruck

Working Papers in Economics and Statistics

2023-08

Natalie Struwe, Esther Blanco, James M. Walker

No response to changes in marginal incentives in one-shot public good experiments

Abstract
We report novel results from changes in the marginal per capita return (MPCR) in a one-shot public good game where participants make a single provision decision. Data was collected using three data collection processes: an online experiment conducted on Prolific, an online experiment conducted with a subject pool of university students, and an experiment implemented following the conventional procedures of the economic laboratory with university students. In three between-subject treatment conditions, we confront participants from each of these three samples with either a low MPCR of 0.4, a high MPCR of 0.8 holding constant the individual endowment, or a high MPCR of 0.8 reducing the individual endowment to hold constant maximum possible group earnings. Based on a total sample size of 952 participants, we find that, unlike results from previous experiments where subjects make multiple contribution decisions in varying experimental designs, contributions to the public good are not significantly different for the different MPCR conditions we study. We consider these results to be highly relevant in highlighting the limits to our understanding of cooperative behavior for settings without repeated interactions.

ISSN 1993-4378 (Print)
ISSN 1993-6885 (Online)