I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16605

Using the Strategy Method and Elicited Beliefs to Explain Group Size and MPCR Effects in Public Good Experiments

Simon Gächter
Diego Marino Fages

NOVEMBER 2023

DISCUSSION PAPER SERIES

IZA DP No. 16605

Using the Strategy Method and Elicited Beliefs to Explain Group Size and MPCR Effects in Public Good Experiments

Simon Gächter
University of Nottingham, IZA and CESifo

Diego Marino Fages
Durham University

NOVEMBER 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Using the Strategy Method and Elicited Beliefs to Explain Group Size and MPCR Effects in Public Good Experiments

In this paper we disentangle the role of cooperative preferences and beliefs for explaining MPCR and group size effects in public goods games. To achieve this, we use the ABC approach, which explains cooperation as a function of cooperative attitudes and beliefs. We measure cooperative attitudes using the incentive-compatible strategy method by Fischbacher et al. (2001, Economics Letters, 71-3, 397–404)(FGF). However, to keep the incentives in the strategy method equal across all group sizes (which FGF does not), we also compare FGF with a version of the strategy method that is scalable to any group size. We find that preference types are similar across strategy methods, group sizes of 3 and 9, and MPCRs of 0.4 and 0.8. Further experiments with group sizes of 3 and 30 again find similar distributions of conditional preferences. The ABC approach predicts actual cooperation in all conditions and for both strategy methods and reveals that, controlling for elicited preferences, differences in cooperation levels observed across the various games are mostly due to differences in beliefs.

JEL Classification: C92, H41

Keywords: public goods, group size, MPCR, strategy method, ABC approach, conditional cooperation, experiments

Corresponding author:
Diego Marino Fages
Durham University
The Palatine Centre
Stockton Rd
Durham DH1 3LE
United Kingdom

E-mail: diegomarinofages@gmail.com

1 Introduction

Public good games (PGGs) are one of the classic games to study cooperation in social dilemmas. PGGs have been used to investigate many aspects of voluntary cooperation (for surveys see Chaudhuri (2011); Fehr and Schurtenberger (2018); Gächter and Herrmann (2009); Ledyard (1995); Zelmer (2003); Drouvelis (2021)). In this paper, we revisit a classic question in PGG research by separately measuring preferences and beliefs: the role of group size and the marginal per capita return (MPCR) for cooperation.

One important insight that will be crucial for our analysis of MPCR and group size effects is that contributions are correlated with beliefs about others' contributions (e.g., Croson (2007); Fischbacher and Gächter (2010); Dufwenberg et al. (2011); Gächter and Renner (2018)). To go beyond mere correlations, Fischbacher et al. (2001) (henceforth FGF) introduced a tool based on the strategy method that measures the causal influence of beliefs about others' contributions on own contribution. The FGF method measures people's preferences for cooperation by eliciting contributions conditional on how much others contribute to the public good, which fixes beliefs. FGF also allows classifying people into preference types as conditional cooperators, free riders, and others.

FGF find that most people are conditional cooperators or free rider types. The distribution of types (in particular conditional cooperators) has been shown to be similar across comparable studies (Thöni and Volk, 2018); mostly robust to the (mis-)understanding of incentives (Gächter et al., 2022; Fosgaard et al., 2017); stable over time (Gächter et al., 2022; Volk et al., 2012); mostly similar across different cooperation games (Mullett et al., 2020; Eichenseer and Moser, 2020) and similar across different cultures (Weber et al., 2023). Moreover, the ABC approach - using people's conditional preferences (also called *a*ttitudes, dispositions) together with their *b*eliefs - can explain individual *c*ontributions to public goods fairly well (Fischbacher and Gächter, 2010; Fischbacher et al., 2012; Gächter et al., 2017, 2022; Isler et al., 2021; Weber et al., 2023). This explanatory power makes the ABC approach a promising methodology for analyzing MPCR and group size effects in cooperation.

A limitation of the existing evidence on the distribution of conditionally cooperative attitudes (preference types) as well as on the predictive power of the ABC approach to explain contributions is, however, that it mostly comes from very similar group sizes of 3 and 4 members, with one exception (Mill and Theelen, 2019) who also studied group sizes of 5 and 7.[1] As far as we are aware, there is no evidence on the distribution of cooperative types across substantially different group sizes. Similarly, not much is

[1]Relatedly, in the Cooperation Databank (a recent effort by Spadaro et al. (2022) to collect papers on cooperation in order to facilitate meta-analyses) only 8 out of 137 published papers related to continuous PGGs (with no deception) mention group size as a treatment variable.

known about how the MPCR affects conditionally cooperative preferences because most studies use MPCRs of 0.4 or 0.5. Finally, to our knowledge, there is no evidence on how well the ABC approach predicts cooperation in larger groups. Hence, the two main goals of our paper are, first, to measure how group size and MPCR affect attitudes, beliefs, and contributions; and, second, to apply the ABC approach for analyzing MPCR and group size effects in cooperation.

We present two studies based on one-shot games. In Study 1 (n=936), we use the strategy method and the ABC approach to study group sizes of 3 and 9 and MPCRs of 0.4 and 0.8, in a factorial design. In Study 2 (n=1,200) we extend group size to 30 members. We use the ABC approach to be able to disentangle preference and belief effects: if we observe different contributions across different group sizes or different MPCRs, it could be because (i) cooperative attitudes vary with group size (or MPCR) or (ii) because individuals respond to beliefs that change with group size (or MPCR) or (iii) both cooperative attitudes and beliefs are affected by group size (or MPCR).

Using the strategy method as developed by FGF to measure cooperative attitudes in different group sizes is not straightforward. The reason is as follows: The FGF method elicits cooperative attitudes incentive-compatibly. To achieve incentive compatibility, FGF asks for conditional and unconditional contributions from all participants and randomly selects one of the participants to have their conditional responses as payoff relevant; for the others in the group the unconditional contribution is used to calculate payoffs. Therefore, the probability of having the conditional response – which we are most interested in – selected as payoff relevant is $1/n$ for each subject. This is less problematic if all groups are of the same size. However, as group size increases, the incentives of the strategy method decisions change (since n increases) confounding the effect of the group size with the weights of the incentives, which become approximately hypothetical as n grows.

To overcome the problem of vanishing incentive compatibility as group size grows, we develop a scalable incentive-compatible strategy method to elicit conditionally cooperative preferences where the incentives for the conditional responses are independent of group size and thereby scalable to any group size. The essential idea of the scalable strategy method (henceforth, SSM) is to select the unconditional and the conditional strategies with equal probability. SSM ensures incentive scalability but might be conceptually problematic because the public good is not well defined. We describe the pros and cons in more detail in Section 3 (and the Appendix). By comparing FGF and SSM we treat the role of incentive scalability of FGF and the conceptual weakness of SSM for eliciting conditional contributions as empirical questions.

In Study 1, in a 2×2×2 design, we compare (i) groups sizes of 3 and 9; (ii) MPCRs of 0.4 and 0.8; and (iii) we elicit conditional preferences using either FGF or SSM.

In Study 2, we also run a 2×2×2 design: we compare groups of 3 and 30 members, with MPCRs of 0.04, 0.08, 0.4 and 0.8 (yielding two levels of marginal social return, 1.2 and 2.4), and we again use either FGF or SSM. Notice that with groups of 30, under FGF elicited conditional preferences become approximately hypothetical because the likelihood of the elicited conditional contribution to be payoff relevant reduces to 1/30. Thus, comparing FGF and SSM in groups of 30 reveals how important incentive scalability is for eliciting conditional preferences.

In both studies, our experiments have two phases: In Phase 1, we elicit people's conditionally cooperative preferences (either using FGF or SSM) in one of the four MPCR×GroupSize parameterizations. This is followed, in Phase 2, by a one-shot direct response PGG (with beliefs) keeping the same parameterization. As we will explain in more detail in Section 2.3, this setting allows us to apply the ABC method, i.e., using the elicited attitudes from Phase 1 and the beliefs from Phase 2 to predict contributions in Phase 2. We then compare these predictions to the actual contributions in Phase 2.

Our main results are as follows. Consistent with previous literature (see next section) we find in both studies that beliefs and contributions increase in MPCR but not in group size. By contrast, the distribution of elicited types of cooperative preferences in Phase 1 are very similar across MPCR and group sizes and in the two methods of how to elicit them (FGF or SSM). This is the case in both studies. We show that the ABC framework explains the observed contribution levels under both strategy methods. According to the ABC approach, the MPCR effect is driven by increased beliefs, and not cooperative preferences because preferences are not much affected by game parameters. Finally, we find evidence that the effect of beliefs on contributions interacts with group size and MPCRs. In particular, using the ABC approach, we see that the effect of beliefs becomes more positive with larger groups and higher MPCRs (or social returns).

This paper is organized as follows. Section 2 describes the standard PGG and the related literature on group size, MPCR effects and the ABC of cooperation. Section 3 describes the FGF and our new scalable strategy method, SSM. Section 4 contains the details of the experimental design of Study 1. Section 5 presents the main results of Study 1. Section 6 describes the design details of Study 2 and Section 7 presents its results. Section 8 investigates the role of beliefs in more detail. Section 9 concludes.

2 Related literature, and the ABC of cooperation

2.1 The public goods game

In the linear public goods game (PGG) we use, each of n group members receives an equal endowment, e tokens. Participants decide how to allocate the endowment between a private $(e - c_i)$ and a public account (c_i). The private account has a return of 1 whereas all the money allocated to the public account (by all participants) gets multiplied by $\alpha > 1$ and the resulting amount is split equally among the n participants so that each subject gets a return of α/n, also known as MPCR, from her investment in the public account. Formally, the material payoff function for subject i is the following:

$$U_i = e - c_i + \frac{\alpha}{n} \sum_{j=1}^{n} c_j, \tag{1}$$

where $e = 10$ and $\frac{1}{n} < \frac{\alpha}{n} < 1$ in all our experiments, which will vary α and n. Assuming monetary payoff maximization, the Nash equilibrium in all games states that all participants should contribute zero to the public account. However, the efficient allocation is achieved when all participants contribute $c_i = e = 10$ to the public account. The standard result in the experimental literature is that people contribute some amounts in between those extremes (e.g., Ledyard (1995); Zelmer (2003); Gächter and Herrmann (2009)). There are different theories for why this is the case, including altruism and warm glow (Croson, 2007; Palfrey and Prisbrey, 1997; Andreoni, 1995b); confusion (Andreoni, 1995a; Houser and Kurzban, 2002; Burton-Chellew et al., 2016; Bayer et al., 2013); reciprocity (Sugden, 1984; Weber et al., 2018; Isler et al., 2021); matching behavior (Guttman, 1986); inequality aversion (Fehr and Schmidt, 1999); and guilt aversion (Chang et al., 2011; Dufwenberg et al., 2011). See also Katuščák and Miklánek (2023) for a recent comparative analysis. In the following, we review the most relevant literature for our purposes, that is, the role of MPCR and group size for cooperation and the ABC approach to explaining cooperation.

2.2 MPCR and group size effects

Since the seminal paper by Isaac et al. (1984), MPCR effects - the higher the marginal per capita return in a public goods game, the higher the level of cooperation - have long been observed in the literature (see, e.g., van den Berg et al. (2022); Lugovskyy et al. (2017); Blanco et al. (2016); Fischbacher et al. (2014); Gunthorsdottir et al. (2007); Goeree et al. (2002); Brandts and Schram (2001); Palfrey and Prisbrey (1997); Fisher et al. (1995); Ledyard (1995) and Zelmer (2003)).[2] MPCR effects are also closely linked to the question of how group size affects voluntary cooperation. As can be seen from the payoff function shown in Eq. 1, if n increases, the MPCR (α/n) also diminishes. To isolate a pure group size effect therefore requires controlling for MPCR, which is what Isaac and Walker (1988) achieved in their seminal experiments.

Isaac and Walker (1988) found that group size increased cooperation but only if the MPCR changed with group size. When the MPCR was held constant they found no group size effects. Since then, a number of papers have studied the effect of group size on contributions to PGGs.[3] For instance, Pereda et al. (2019a) found that groups of 100 and 1000 are not significantly different in terms of contribution levels. This suggests that extrapolating conclusions from smaller group sizes might be reasonable. In a within-subjects design, with group sizes of 5, 10, ..., 40, Pereda et al. (2019b) observed a monotonic increase of the cooperation rate as a function of group size.

Diederich et al. (2016) also found a positive group size effect and concluded that the effect was driven by the intensive margin (i.e., those who contribute contributed more in larger group sizes). This is consistent with the beliefs of the subjects. They also show that the share of free riders remains constant in spite of the beliefs of the participants showing a larger share of free riders in larger groups. An older study (Kerr, 1989) argued that beliefs are also affected by an efficiency illusion (i.e., a diminished sense of self-efficacy in larger groups) even when the self-efficiency was kept constant.

Nosenzo et al. (2015) and Weimann et al. (2019) showed that the effect of group size depends on the MPCR level. When the MPCR is low, the effect of group size on contributions is positive, but the effect disappears or becomes negative when the MPCR is high. Weimann et al. (2019) argue that the reason is due to beliefs, which in turn are affected by the salience of the advantages of cooperation. They define salience as the distance between $\frac{1}{n}$ and the MPCR (that is, $d = \frac{\alpha}{n} - \frac{1}{n}$). The higher d, the more salient is the fact that cooperation is collectively beneficial. Weimann et al. (2019) showed that cooperation increases non-linealy with d.

In summary, beliefs take a central role in most recent studies, but how conditionally cooperative preferences are influenced by group size and MPCR remains unexplored. Moreover, it is unknown whether the ABC approach works for larger group sizes and high MPCRs. The aim of this paper is to provide the missing evidence.

[2]While MPCR effects have been observed very often, they do not always occur, as suggested by Struwe et al. (2023) based on experiments run on Prolific. As we will show below, we do find a MPCR effect of experiments also run on Prolific, so this difference in results is likely due to procedural differences in conducting the experiment.

[3]Further experiments on MPCR and group size include Marwell and Ames (1979); Isaac et al. (1994); Goeree et al. (2002); Carpenter (2007); see also Zelmer (2003) and Ledyard (1995) for overviews.

2.3 The ABC of cooperation

Our ambition in this paper is not only to provide evidence on conditionally cooperative preferences as a function of group size and MPCR but also to explore the explanatory power of conditionally cooperative preferences jointly with beliefs to explain cooperation levels across a variety of PGG parameter sets. The basic idea was provided by Fischbacher and Gächter (2010) who used the strategy method to explain the decline of cooperation in repeated public goods games. They elicited people's preferences for conditional cooperation (details follow in the next section) and had subjects play ten rounds of a repeated public goods game played in the Stranger matching protocol (that is, group composition is changed at random in each round). In each round, players made a contribution decision and reported their beliefs about how much their current group members contribute in the next round. Fischbacher and Gächter (2010) then showed that the elicited cooperative preferences evaluated at the beliefs of a given period predicted actual contributions round by round. Fischbacher et al. (2012) further showed that this method also predicts contributions well in one-shot public goods games.

Gächter et al. (2017) named this method the ABC approach to cooperation. It measures individual *a*ttitudes (a_i) to cooperation as a function of all possible rounded average contributions of other group members (sometimes the function a_i is also called dispositions or preferences), *b*eliefs (b_i) about others' contributions average contributions and actual *c*ontributions (c_i) separately and explains cooperation as $a_i(b_i) \rightarrow c_i$, that is, the function a_i evaluated at player $i's$ belief b_i predicts player $i's$ contribution c_i. Gächter et al. (2017) found that the ABC approach explains contributions in maintenance and provision versions of public goods (see also Gächter et al. (2022) for a related result).

Interestingly, Gächter et al. (2017) and Gächter et al. (2022) observed that the share of conditional cooperators was higher (and the share of free riders lower) in the provision version of the public good game than in the maintenance version, which suggests that attitudes to cooperation are influenced by the context of the game. Isler et al. (2021) replicated these findings and present a framework that generalizes the ABC approach to what they call the Contextualized Strong Reciprocity Approach (CSR).[4] CSR is a version of the ABC approach that allows for context effects: CSR explains cooperation as $a_i(f, b_i(f)) \rightarrow c_i$, where both a_i and b_i are potentially functions of the contextual features (f) of a given public goods game. In this paper, f is the MPCR or the group size, both of which were very similar in all previous studies that used the ABC approach (in Isler et al. (2021), f referred to maintenance or provision public good). Our goal

[4]Conditional cooperation is one important instance of "strong reciprocity". See Weber et al. (2018) and Isler et al. (2021) for discussions and analyses.

in this paper is to provide evidence on all elements of $a_i(f, b_i(f)) \rightarrow c_i$ where the experiments will manipulate f as group size and MPCR.

3 Strategy methods

The strategy method was introduced by Selten (1967) in the context of an oligopoly game but has now been used in all kinds of games (Brandts and Charness (2011); Keser and Kliemt (2021)). Here we focus on the application of the strategy method to PGG, where participants make contingent decisions for each possible average contribution of other members in the group (for a review, see Thöni and Volk (2018)). By eliciting a complete contribution profile, we get data on all possible situations including those that are only rarely reached (e.g., everyone in the group contributing nothing or their entire endowment). We will first review the standard method by Fischbacher et al. (2001) and present its problems in the context of our research questions in Section 3.1 and then introduce our new version, called the scalable strategy method, in Section 3.2.

3.1 Fischbacher, Gächter and Fehr (2001; FGF)

The first application of the strategy method in the one-shot public goods game is due to Fischbacher et al. (2001) (FGF). In the FGF method, participants are asked to make two decisions: (i) an unconditional contribution and, (ii) a contribution table (that is, a_i). The unconditional contribution is the amount they would like to contribute without any information on the contribution of the other group members. For the contribution table, participants need to answer how much they would like to contribute for each possible average contribution of the other group members (i.e., the contributions they would like to make conditional on the other group member's contributions as if they were last movers in a sequential game). Participants are told that after the decisions are made, $n-1$ of them will be selected randomly to have the unconditional contribution used as their payoff relevant decision and the remaining group member will have their contribution table used as their payoff relevant decision (computed by imputing the average of the other $n-1$ members in their contribution table). See the Online Appendix for the instructions.

The FGF method is an incentive-compatible way to elicit the conditional contribution decisions a_i (since all the choices are potentially payoff relevant). For our research purposes, it is important to note that the probability of a participant having their contribution table selected as payoff relevant depends on the number of participants in the group (i.e., $\frac{1}{n}\%$). Therefore, the larger the group, the more hypothetical the contribution table becomes. This posits a problem if we want to incentive-compatibly elicit cooperative dispositions as a function of group size. In the following subsection, we introduce a method that solves this problem.

3.2 A scalable strategy method (SSM)

Our research questions demand an incentive-compatible method that also keeps the probability of the contribution table being payoff relevant constant across group sizes ("scalable"). We achieve these goals by giving the contribution table and the unconditional contribution equal weights (50% each). So, instead of randomly selecting *one* group member (with probability $1/n$) for whom the contribution table is payoff relevant (like in FGF), SSM randomizes with equal probability for *each* group member whether their unconditional contribution or their contribution table is selected as payoff relevant. This feature achieves incentive scalability.

The SSM procedure involves two steps: First, we calculate the *computed conditional contribution* for group member i by plugging the average contribution of the remaining group members (rounded to the nearest integer) into group member i's contribution table. Second, for each group member i, we randomly select either their unconditional contribution or their contribution table with 50% probability each. If for i the unconditional contribution is selected, we use all other group members' computed conditional contributions to determine i's individual payoff; otherwise, if for i the contribution table is selected, we use the computed conditional contribution for player i and all other player's unconditional contributions to determine i's individual payoff (see the Appendix of this paper for further details and a concrete example of the procedure).

Notice that it is possible to have groups where either the unconditional contribution or the contribution table is selected for all players. The reason is that under SSM payoffs are independent between group members. By contrast, because FGF only selects one group member's contribution table to determine his/her contribution, whereas all other group members contribute their unconditional contribution, payoffs are interdependent like in any standard public good game. Hence, SSM gives up FGF's payoff interdependence to achieve the scalability that FGF lacks. Put differently, SSM has an undesirable feature: the size of the public good is individual-specific, as opposed to the standard PGG where the same size of the public good applies to everyone in the group.

However, there are several reasons why SSM's conceptual downside is not a major concern. First, we are still studying a social dilemma (i.e., the Nash equilibrium is to contribute zero, and efficiency demands to contribute everything). Second, this problem is independent of group size or MPCR. Third, and most importantly, we are only interested in the elicited conditional preferences and not the unconditional contributions. The conditional preferences are elicited incentive-compatibly and the incentives for their truthful revelation are independent of group size. The empirical comparison of the elicited a_i and the ABC predictions derived from them will reveal how important the respective disadvantages of FGF and SSM are.

4 Study 1: Design

The experiments consisted of a 2×2×2 between-subjects design varying the strategy method (FGF and SSM), group size (small=3 and large=9) and MPCR (low=0.4 and high=0.8). Participants (n=936) were randomized into one of the eight conditions (Table 1). Each group member was endowed with $e = 10$ tokens to either keep or invest into the public good (Section 2.1). The strategy methods therefore elicited conditional contributions a_i for each of the eleven (0 to 10) possible average contributions of other group members.

All experiments had two phases. In Phase 1, participants played the PGG in either the FGF or SSM strategy method. In Phase 2, they played the PGG in the direct-response mode. Apart from different parameters in the respective condition (see Table 1) the Phase 2 PGGs were the same regardless of the strategy method used in Phase 1.

Condition	Phase 1 Strat. method	MPCR	Group size (n)	Social return (α)	N
1	FGF	0.4	3	1.2	117
2	FGF	0.4	9	3.6	117
3	FGF	0.8	3	2.4	117
4	FGF	0.8	9	7.2	117
5	SSM	0.4	3	1.2	117
6	SSM	0.4	9	3.6	117
7	SSM	0.8	3	2.4	117
8	SSM	0.8	9	7.2	117

Table 1 Study 1 experimental conditions (between-subjects)

After explaining the game, we used three incentivized control questions: two questions about the PGG and one regarding the probability of having the contribution table as the payoff relevant decision (see the Online Appendix for instructions and control questions). Participants who responded incorrectly in the first attempt had to keep trying until they got it right before proceeding (but only those responding correctly in the first attempt were compensated for their correct answers). Participants then completed the respective strategy method task (FGF or SSM, depending on the condition they were randomly selected into). This completed Phase 1 of the experiment.

In Phase 2, participants made a contribution decision in a one-shot direct response PGG with the same parameters as relevant in the respective condition (see Table 1). We also elicited the beliefs (first and second order). We incentivized the belief elicitation by rewarding subjects if their prediction was exactly right. We will use these games to test for the predictions made by the ABC approach (see Section 2.3).

The experiment was programmed in Qualtrics and conducted online using the platform Prolific (Palan and Schitter (2018)) with participants from the UK in the 18-25 years old range (average age 21.8 years; 67.4% females; 60.0% students).[5] We restricted participation by age in order to minimize confounds caused by age and cohort effects.[6] Study 1 was run in early 2021. Participants were paid £12.71 per hour on average. The average participant took 8 minutes to complete the experiment.

5 Study 1: Results

5.1 Attitudes

Both strategy methods allow us to classify participants into types according to their cooperative preferences ("attitudes" a_i). Here we follow Thöni and Volk (2018) and classify people into attitude types (using their Stata command "cctype"). Conditional Cooperators increase their cooperation with the cooperation of others; Free Riders contribute zero for any contribution of others; Triangle Cooperators increase their contributions up to a certain point and then decrease with the contribution of others; Unconditional Cooperators contribute a non-zero constant value for any contribution of others.

Figure 1A shows that the distribution of types is remarkably similar in all eight conditions. Conditional Cooperators is the main category, which accounts for 80.9% of the subjects, followed by the Triangle Cooperators 6.1% and Free Riders with 4.9% (the remaining 5.7% are grouped as others). We also compare the attitude type distributions between FGF and SSM for each of the four parameterizations using Chi-squared tests and find that none of the pairwise comparisons is significant (smallest p-value = 0.338). This suggests that FGF and SSM are similar in identifying attitude types. For further illustration see the Online Appendix, Section C.

Figure 1B shows the average of the responses in the contribution tables (i.e., the

[5]See Hergueux and Jacquemet (2015) for the reliability of the online elicitation of social preferences and Arechar et al. (2018) for a methodologial discussion of online experiments.

[6]Age is known to influence cooperation positively, see, e.g., List (2004); Gächter and Herrmann (2011); Arechar et al. (2018). To maximize statistical power, we kept the age of our participants constant and similar to the vast majority of studies on MPCR and group size effects.

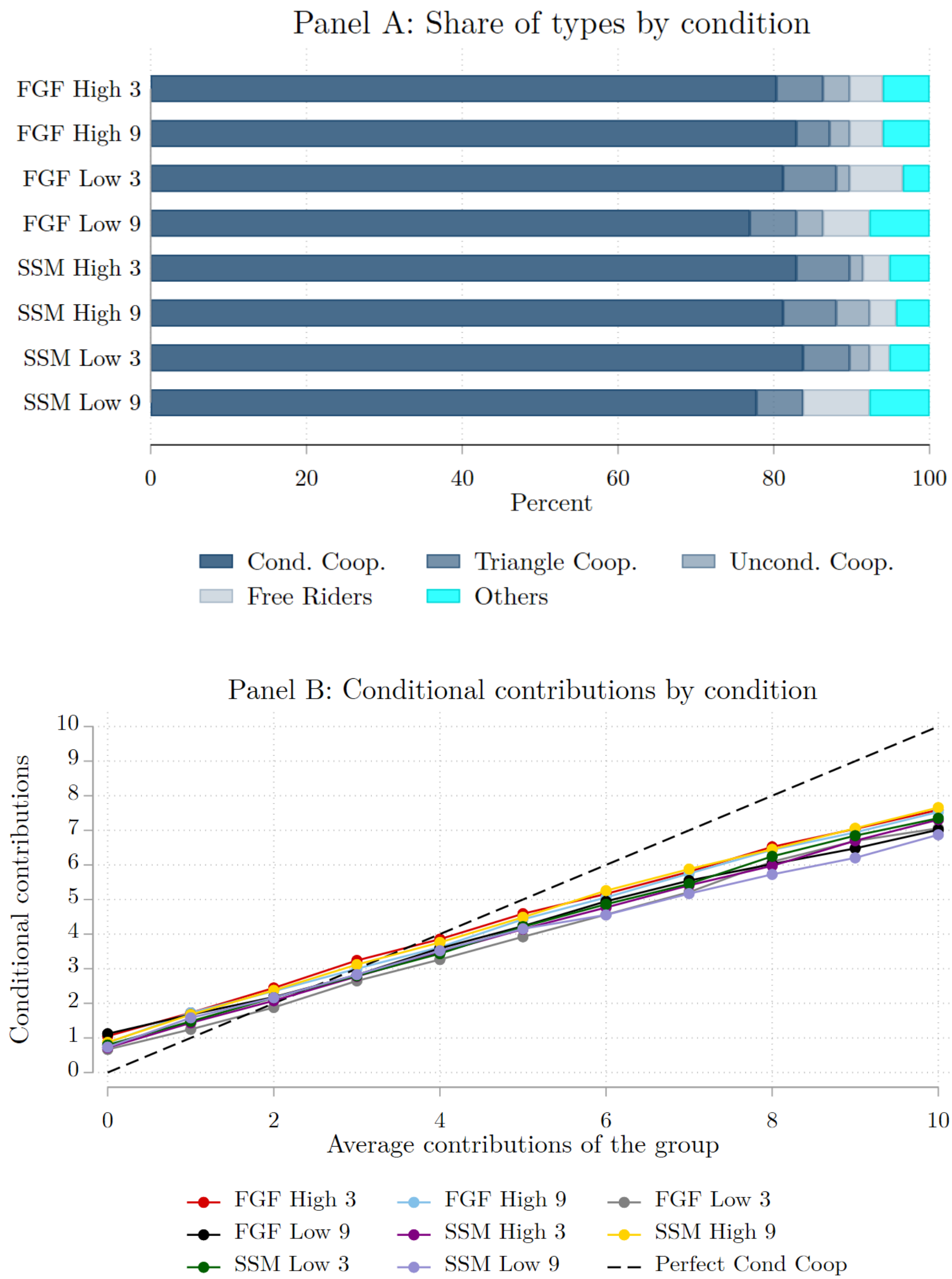


Figure 1 Cooperative attitudes a_i as elicited in Phase 1 in Study 1. Panel A: Distribution of conditionally cooperative types. Panel B: Average conditional contribution $\bar{a}_i$ across all preference types. High (Low) refers to an MPCR of 0.8 (0.4). Large (Small) refers to a group size of 9 (3).

average conditional contributions $\bar{a}_i$). As a benchmark, we also plot the line for perfect conditional cooperation (the diagonal). Average conditional cooperation $\bar{a}_i$ is very similar in all conditions (i.e., by method, MPCR, and group size).

	(1) SSM Av. cond. contrib.	(2) FGF Av cond. contrib.
large	-0.193	0.216
	(0.232)	(0.252)
high MPCR	-0.078	0.524**
	(0.226)	(0.244)
large x high MPCR	0.540	-0.336
	(0.333)	(0.351)
_cons	4.149***	3.932***
	(0.160)	(0.176)
R^2	0.009	0.011
N	468	468

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
OLS estimates. The dependent variable is the average of all 11 conditional contributions a_i in either FGF or SSM. *large* is an indicator that takes the value 1 for groups of 9 and 0 for groups of 3. *high MPCR* is an indicator that takes the value 1 for MPCR=0.8 and 0 for MPCR=0.4. Robust standard errors in parentheses.

Table 2 Average conditional contributions a_i as elicited in Phase 1 by SSM and FGF in Study 1 as a function of group size and MPCR

To test the link between game parameters and conditional cooperation formally, we averaged the conditional contributions for each subject and regress them on dummies for high MPCR, large group and their interaction. The results are shown in Table 2.[7] Although the results are, with one exception, not statistically significant, we see differences between the methods. In Column 1 (SSM), average conditional contributions are insignificantly lower in the large group and in MPCR, but the effect reverses insignificantly when the MPCR is high. In Column 2 (FGF), average conditional contributions increase insignificantly with group size and significantly with MPCR but decreases insignificantly with the interaction of the two.

[7] The results are virtually unchanged if we use a participant $i's$ 11 responses in their vector a_i (rather than the average of the 11 responses a_i as in Table 2) and cluster the standard errors by participant. We include these in the replication package.

5.2 Beliefs and contributions

After the strategy method PGG experiment in Phase 1, participants in Phase 2 took part in a one-shot direct-response PGG where we observed c_i and where we also elicited beliefs b_i. Recall that these one-shot PGG games in Phase 2 had the same parameter conditions as in Phase 1, regardless of whether we used FGF or SSM in Phase 1. Figure 2 shows the averages of beliefs and contributions by treatment and suggests a positive MPCR effect in both beliefs and contributions, but no group size effect.

It is also interesting to look at cooperation differences in terms of salience as proposed by Weimann et al. (2019) (see also Section 2.2). Recall that salience is defined as $d = \frac{\alpha}{n} - \frac{1}{n}$. For our MPCR and group size parameters (see Table 1), $d_{Low3} = 0.07$; $d_{High3} = 0.47$; $d_{Low9} = 0.29$; and $d_{High9} = 0.69$. The results are, to a large extent, consistent with the findings of Weimann et al. (2019). That is, the larger the salience, the higher the beliefs and contributions and the effect becomes smaller for high levels of d.

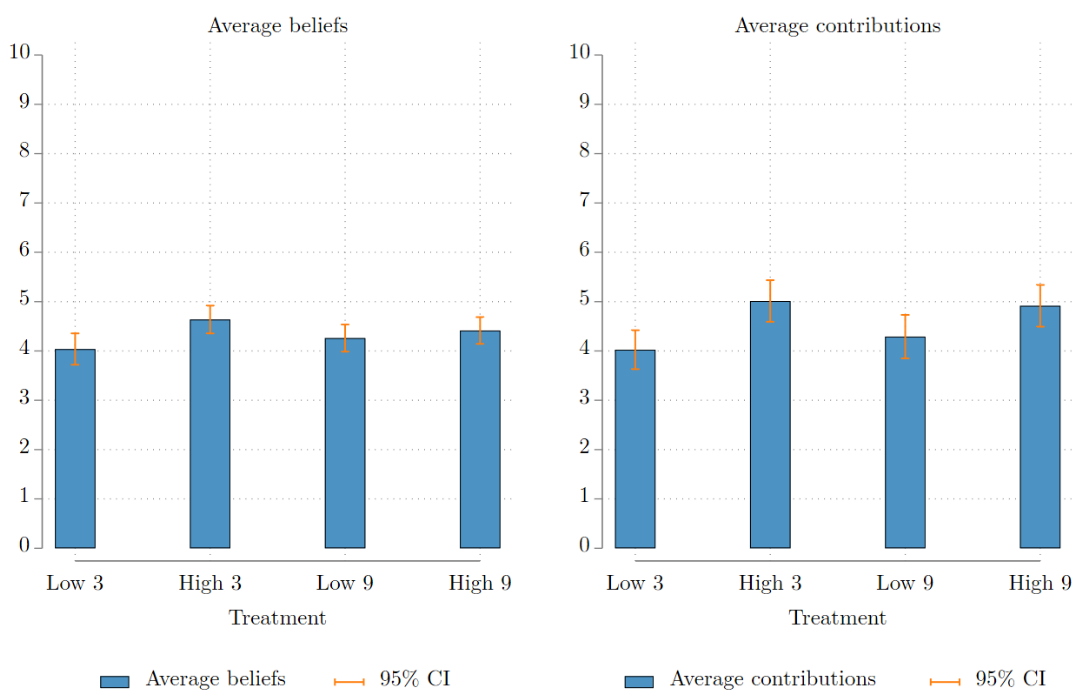


Figure 2 Average of beliefs b_i and contributions c_i by treatment in Phase 2 of Study 1. High (Low) refers to MPCR of 0.8 (0.4). Large (small) refers to group size of 9 (3).

To study the differences between the treatments formally, we regress beliefs and contributions on dummy variables for large group ($n = 9$) and high MPCR (0.8). For this, we pool the observations from both strategy methods since, for given game parameters, the direct response one-shot PGG in Phase 2 is the same regardless of the strategy method used in Phase 1. Table 3 presents the results. We find positive and significant effects of the MPCR on beliefs and contributions. These results confirm the MPCR effects reported in previous literature (see Section 2.2). Regarding group size, we find a positive effect but it is not statistically significant. Figure A.1 in the Online Appendix presents the full distributions.[8]

	(1) Pooled beliefs	(2) Pooled contrib.
large	0.222	0.265
	(0.213)	(0.300)
high MPCR	0.598***	0.987***
	(0.215)	(0.293)
large x high MPCR	-0.444	-0.363
	(0.291)	(0.426)
_cons	4.038***	4.026***
	(0.161)	(0.200)
R^2	0.010	0.016
N	936	936

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
OLS estimates. The dependent variables are the *beliefs* (column 1) and *contributions* (column 2) reported in the Phase 2 PGG. *large* is an indicator that takes the value 1 for groups of 9 and 0 for groups of 3. *high MPCR* is an indicator that takes the value 1 for MPCR=0.8 and 0 for MPCR=0.4. Robust standard errors in parentheses.

Table 3 Beliefs and contributions in Phase 2 of Study 1 as a function of group size and MPCR

5.3 Testing the accuracy of ABC predictions

Having shown the descriptive evidence, we proceed by testing the predictive accuracy of the ABC approach. We start by using the elicited beliefs b_i and check how well they can predict actual contributions c_i by using the attitudes a_i as described in subsection

[8]For completeness, Fig.A.2 (Online Appendix) presents the distributions of second-order beliefs. However, for the ABC approach only first-order beliefs matter, which is why we focus on them here.

2.3. We compute the prediction error as the difference between the actual contributions and the contributions predicted by the ABC approach.

Formally, the predicted contribution of individual i is calculated as $\hat{c}_i = a_i(b_i)$; the prediction error therefore is $\hat{c}_i - c_i$. Individual $i's$ $\hat{c}_i$ can be derived non-parametrically (by taking $i's$ contribution c_i as indicated in $i's$ table of conditional contributions a_i at the belief b_i) or parametrically (by estimating the slope of the conditional contributions $\hat{a}_i$ from all 11 entries in individual $i's$ vector a_i and using these parameters and the beliefs to calculate $\hat{c}_i = \hat{a}_i(b_i)$). Since most subjects are (imperfect) conditional cooperators, we follow the second approach to obtain smoother estimates, but the results are similar if we follow the non-parametric approach (see Online Appendix Section B).

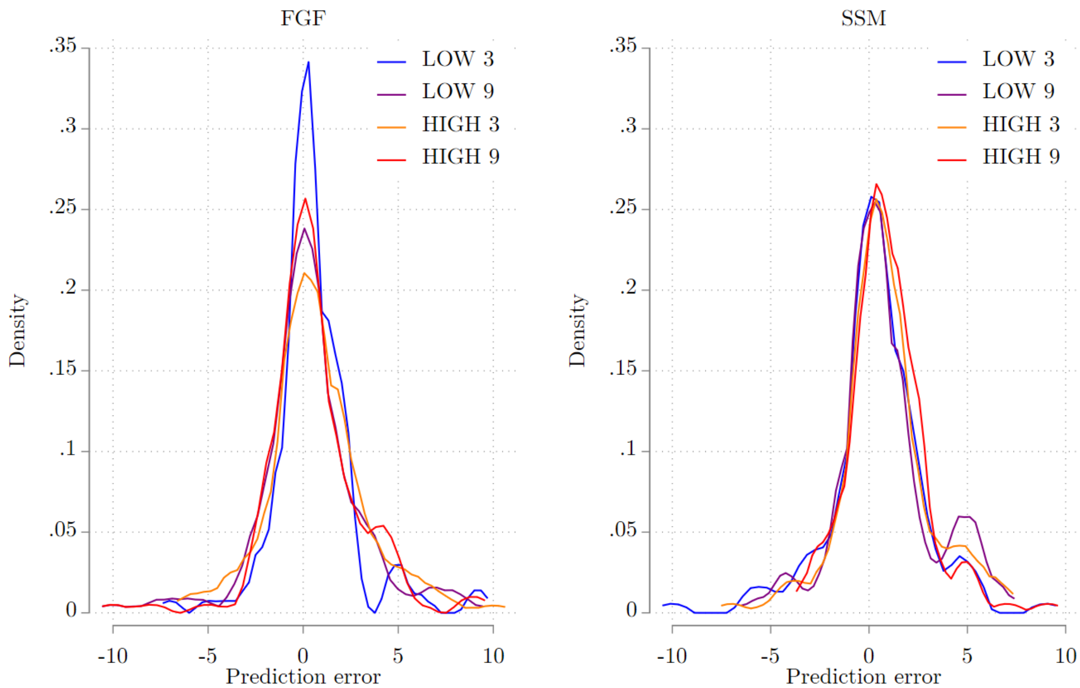


Figure 3 Distribution of ABC prediction errors ($\hat{c}_i - c_i$) in Study 1

Figure 3 shows that in all conditions the mode of the $\hat{c}_i - c_i$ distributions is at zero, which means the mode of actual contributions coincides with the contributions predicted by the ABC approach. Figure 3 also suggests that the distributions are similar across conditions. Note, however, that the distributions in different treatments overlap slightly more under SSM. Kolmogorov-Smirnoff tests comparing the prediction errors under FGF and SSM in each treatment only reject the null hypothesis of equal distributions in the High 9 condition (p-value=0.046), but this does not survive a Bonferroni correction for multiple comparisons (see the Online Appendix Figure A.4 for treatment-specific graphs).

	(1) Pooled FGF	(2) Pooled SSM	(3) Large (9) FGF	(4) Large (9) SSM	(5) Small FGF	(6) Small SSM
predicted	0.555***	0.615***	0.521***	0.670***	0.578***	0.560***
	(0.0919)	(0.0676)	(0.140)	(0.0911)	(0.124)	(0.101)
beliefs	0.405***	0.407***	0.467**	0.486***	0.362**	0.344***
	(0.118)	(0.0807)	(0.193)	(0.0996)	(0.152)	(0.124)
high MPCR	0.222	0.393*	0.223	0.268	0.233	0.510*
	(0.226)	(0.209)	(0.331)	(0.283)	(0.305)	(0.302)
large	-0.174	0.269				
	(0.220)	(0.202)				
_cons	0.528*	0.127	0.222	-0.108	0.614**	0.567
	(0.279)	(0.250)	(0.455)	(0.269)	(0.308)	(0.367)
R^2	0.489	0.547	0.454	0.615	0.528	0.484
$CVMSE$	5.938	4.807	6.650	4.306	5.482	5.303
N	468	468	234	234	234	234

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
OLS estimates. The dependent variable are the contributions c_i in the Phase 2 one-shot PGG. *beliefs* are the b_i elicted in Phase 2 and *predicted* is the ABC prediction using either FGF or SSM. *large* is an indicator that takes the value 1 for groups of 9 and 0 for groups of 3. *high MPCR* is an indicator that takes the value 1 for MPCR=0.8 and 0 for MPCR=0.4. Robust standard errors in parentheses.

Table 4 Regressions explaining contributions in Study 1 using the ABC method

Table 4 shows the regressions of the ABC model adapted from Fischbacher and Gächter (2010) (see their Table 2) with the contributions c_i in the Phase 2 one-shot PGG as dependent variable. This approach includes predicted contributions $\hat{c}_i$ and beliefs b_i as additional regressors because Fischbacher and Gächter (2010) have shown that $c_i = \beta\hat{c}_i + (1 - \beta)b_i$, that is, actual conntributions are a weighted average of predicted contributions and beliefs. In other words, beliefs matter on top of predicted contributions (presumably because of the enhanced salience of a specific belief in a direct response PGG). We also add a dummy for high MPCR and a dummy for large group size (in the pooled models of Columns 1 and 2). Columns 1, 3 and 5 show the results for FGF and columns 2, 4 and 6 for SSM for different samples.

The estimation results confirm that both beliefs and predicted contributions are highly significantly positively correlated with contributions in all models. High MPCR does not contribute significantly on top of beliefs and predicted contributions (except in Columns 2 and 6 where we find a weakly significantly positive effect).

Our next step is to compare the two methods in terms of in-sample and out-of-sample prediction power (as in, e.g., Andreoni et al. (2015)). The cross-validated Mean Squared Error (CVMSE) is a measure of out-of-sample prediction, in other words, how well the estimated model predicts the contributions of new observations. To compute this, we split the sample randomly in five sub-samples and train the data (i.e., estimate the parameters) using four of them to predict the remaining one.[9] After that, we compute the mean squared error of these predictions.

For the full sample comparison, we find that the SSM outperforms FGF because the CVMSE is smaller in SSM than in FGF (4.807 vs. 5.938). We should also expect this relative increase in performance to be higher when group size is larger. This is because the probability of the contribution table being payoff relevant changes with group size in FGF but not in the SSM: The probabilities are 0.33 (FGF) vs. 0.50 (SSM) in groups of 3 members and 0.11 (FGF) vs. 0.50 (SSM) in groups of 9 members (see Section 3.1). To see this, we split the data by group sizes in columns 3 to 6. The CVMSE in the large groups is 6.650 in FGF and 4.306 in SSM. Even in the small groups, SSM slightly outperforms FGF in out-of-sample predictions (5.303 vs. 5.482). Similarly, the R-squared (in-sample prediction) is 35% higher in the large group and 12% higher in the full sample, but 8% lower in the small group.[10] The main results are also robust to restricting the sample to conditional cooperator types only.

In summary, we find that neither MPCR nor group size affect conditional contributions a_i. By contrast, beliefs b_i and contributions c_i do significantly increase in MPCR but not group size, confirming previous literature. Because a_i are not affected by game parameters, but a high MPCR increases b_i, the ABC approach predicts higher contributions c_i when the MPCR is high. Our experiments confirm this prediction.

While our results are reassuring, some caution is warranted: although our large group was three times larger than the small group, we are still in a relatively small group size context. Next, we therefore explore significantly larger group sizes.

6 Study 2: Design

For Study 2, we compare group sizes of 3 and 30 (a ten-fold increase). This poses a new problem. Keeping the MPCR constant would imply a huge social return α. Thus,

[9] We used the Stata command "crossvalidate" (Schonlau, 2020).

[10] It is also possible that the SSM outperforms FGF because it is simpler to understand (i.e. 50% probability of each decision being payoff relevant as opposed to picking a participant at random with an ad-hoc probability). However, why this is the case is beyond the scope of this paper.

instead of holding the MPCR (α/n) constant, we hold α constant across group sizes but vary its level. The parameters are shown in Table 5.

In summary, the experiment in Study 2 consisted of a 2×2×2 between-subjects design using the SSM and FGF, varying group size (small = 3 and large = 30) and α (low = 1.2 and high = 2.4). Apart from these differences, the Study 2 experiment was very similar to the Study 1 experiment: like Study 1, Study 2 had two phases. In Phase 1 we elicited attitudes to cooperation using the SSM method, and in Phase 2 participants played a direct response PGG (with the parameterization of the respective condition). We also elicited first- and second-order beliefs.

Note that some of the conditions in Study 2 are exactly the same as in Study 1. In the Online Appendix (Section D) we show that the results are replicated.

Study 2 was run on Prolific with 1,200 participants (average age 21.6 years; 57.8% females; 64.8% students). Participants were paid the equivalent to £13.72 per hour on average. The average subject took around 8 minutes to complete the experiment. Study 1 participants could not take part in Study 2.

Condition	Phase 1 Strat. method	MPCR	Group size (n)	Social return (α)	N
1	FGF	0.4	3	1.2	150
2	FGF	0.04	30	1.2	150
3	FGF	0.8	3	2.4	150
4	FGF	0.08	30	2.4	150
5	SSM	0.4	3	1.2	150
6	SSM	0.04	30	1.2	150
7	SSM	0.8	3	2.4	150
8	SSM	0.08	30	2.4	150

Table 5 Study 2 experimental conditions (between-subjects)

7 Study 2: Results

7.1 Attitudes

Figure 4A reveals that the share of types is very similar in all conditions. In the low α treatments (with SSM), nonetheless, there is some evidence that the number of free riders increases (from 2% to 6.7%) and the number of conditional cooperators decreases (from 82.7% to 78.7%) with group size. This also happened in Study 1 (SSM) for the Low MPCR treatments, the number of free riders increases (from 2.6% to 8.6%) and the number of conditional cooperators decreases (from 83.8% to 77.8%) with group size.

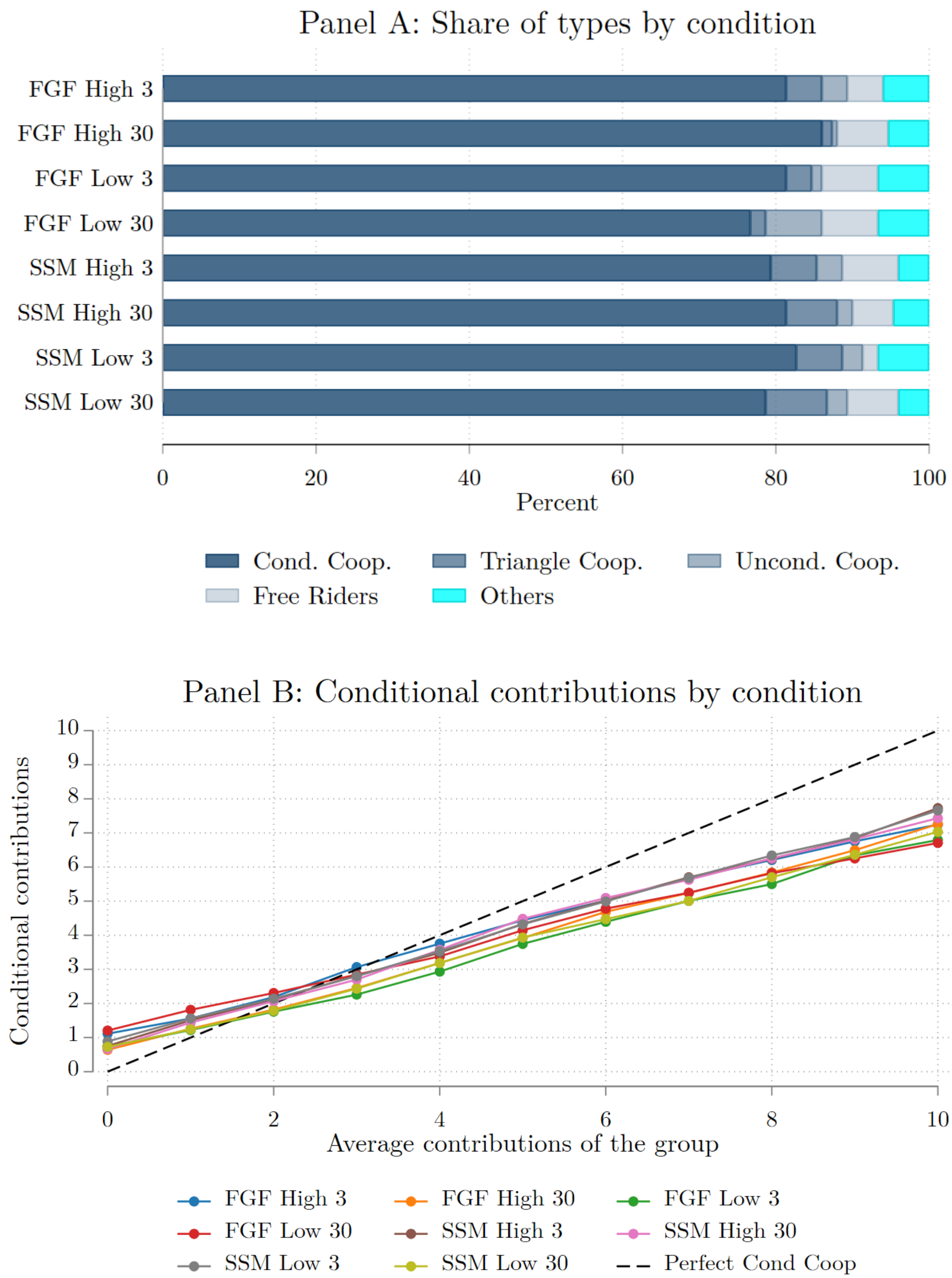


Figure 4 Cooperative attitudes a_i as elicited in Phase 1 of Study 2. Panel A: Distribution of conditionally cooperative types. Panel B: Average conditional contributions $\bar{a}_i$ across all preference types. High (Low) refers to a social return of 2.4 (1.2). Large (small) refers to group size of 30 (3).

Nosenzo et al. (2015) and Weimann et al. (2019) found that, in low MPCR treatments, cooperation increases with group size, but they do not disentangle the role of conditional preferences and beliefs for cooperation. Our results would be consistent with theirs if conditional cooperators in larger groups had a slope high enough to compensate for the increase in free riders and the reduction in conditional cooperators. This argument is consistent with Diederich et al. (2016), who found a positive group size effect driven by the intensive margin (see Section 2.2).

Figure 4B shows the average of the responses in the contribution tables by condition (with the dashed black line representing a hypothetical perfect conditional cooperator). Again, the slopes are similar in all eight conditions.[11]

	(1)	(2)	(3)	(4)
	SSM	FGF	Pooled	Pooled
	Av. cond. contrib.	Av. cond. contrib.	beliefs	contrib.
large	-0.447**	0.347	-0.037	-0.190
	(0.211)	(0.247)	(0.170)	(0.260)
high α	-0.023	0.575**	0.883***	0.917***
	(0.205)	(0.223)	(0.187)	(0.261)
large x high α	0.411	-0.730**	-0.587**	-0.447
	(0.299)	(0.327)	(0.252)	(0.378)
_cons	4.254***	3.699***	3.847***	3.797***
	(0.134)	(0.155)	(0.134)	(0.175)
R^2	0.010	0.011	0.028	0.012
N	600	600	1200	1200

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
OLS estimates. The dependent variable is the average of all 11 *conditional contributions* a_i as elicited in Phase 1 of Study 2 (model 1); the *beliefs* b_i (model 2); and *contributions* c_i (model 3) as elicited in the direct-response one-shot PGG in Phase 2 of Study 2. *large* is an indicator that takes the value 1 for groups of 30 and 0 for groups of 3. *high* α is an indicator that takes the value 1 when the social return = 2.4 and 0 when the social return = 1.2. Robust standard errors in parentheses.

Table 6 Average conditional contribution a_i; beliefs b_i; and contributions c_i, all as a function of group size, MPCR, and their interaction in Study 2

In the first two columns of Table 6 we test formally how game parameters affect the attitudes to cooperation. Column 1 shows that (in SSM) preferences for cooperation are lower in the large groups, but the interaction term with high α is positive and of the

[11]See also the Online Appendix Section C for another comparison of the slopes and average contributions.

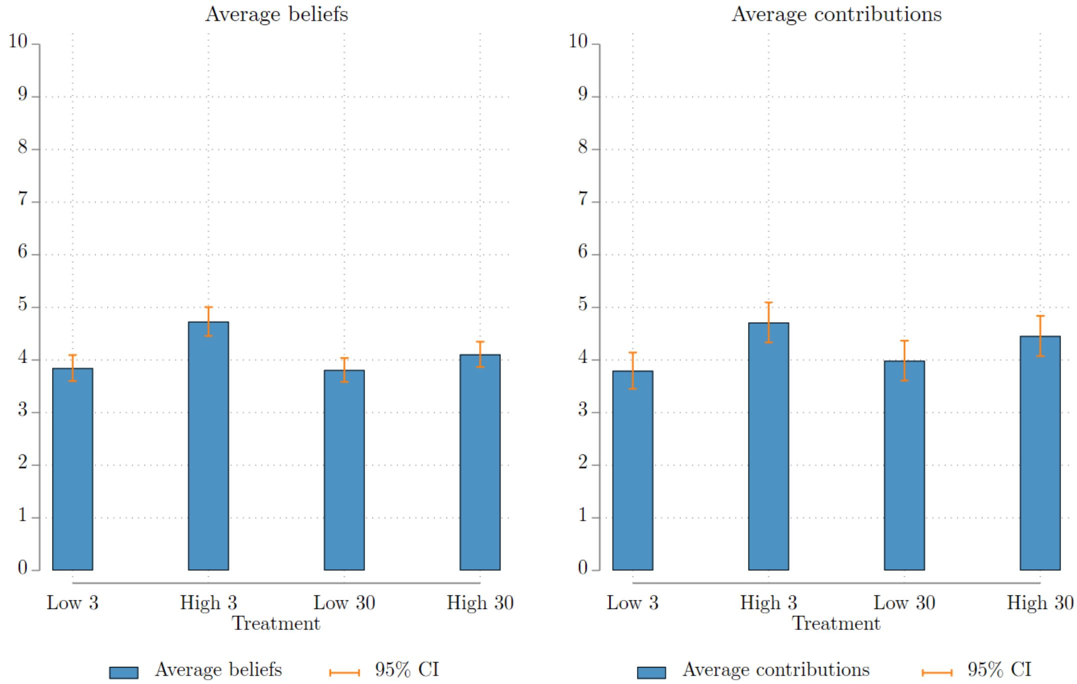


Figure 5 Average of beliefs b_i and contributions c_i by treatment in Phase 2 of Study 2. High (Low) refers to social returns of 2.4 (1.2). Large (small) refers to group size of 30 (3).

same magnitude, meaning that only large groups with low α have significantly lower preferences for cooperation. Column 2 shows (in FGF) a positive (but not statistically significant) effect of large groups with a negative interaction with α, meaning that large groups with high α have significantly weaker preferences for conditional cooperation. These differences between SSM and FGF are similar to the differences found in Study 1.

7.2 Beliefs, contributions, and ABC prediction errors

Figure 5 presents the average beliefs b_i (left) and contributions c_i (right) in each treatment of Study 2. Consistent with the results from Study 1 (see Fig. 2), beliefs and contributions increase in the social return (α) and suggest no group size effect. Again, the results are consistent with the salience argument by Weimann et al. (2019): for a given group size, contributions increase in salience and the increase is larger for larger groups (where the change in salience is smaller).

The formal analysis is presented in Columns 3 and 4 of Table 6, which shows the regressions of beliefs and contributions on dummy variables for groups of 30 and α =

2.4. We find a positive and statistically significant effect and similar in magnitude for α in both cases. The coefficients for group size (*large*) are negative for beliefs and contributions but not statistically significant. In both cases, the interaction term (*large x high* α) is negative but only statistically significant for beliefs.[12] Like in Study 1, the prediction errors are similar across conditions, as shown in Figure 6. Overall, the results are consistent with Study 1.

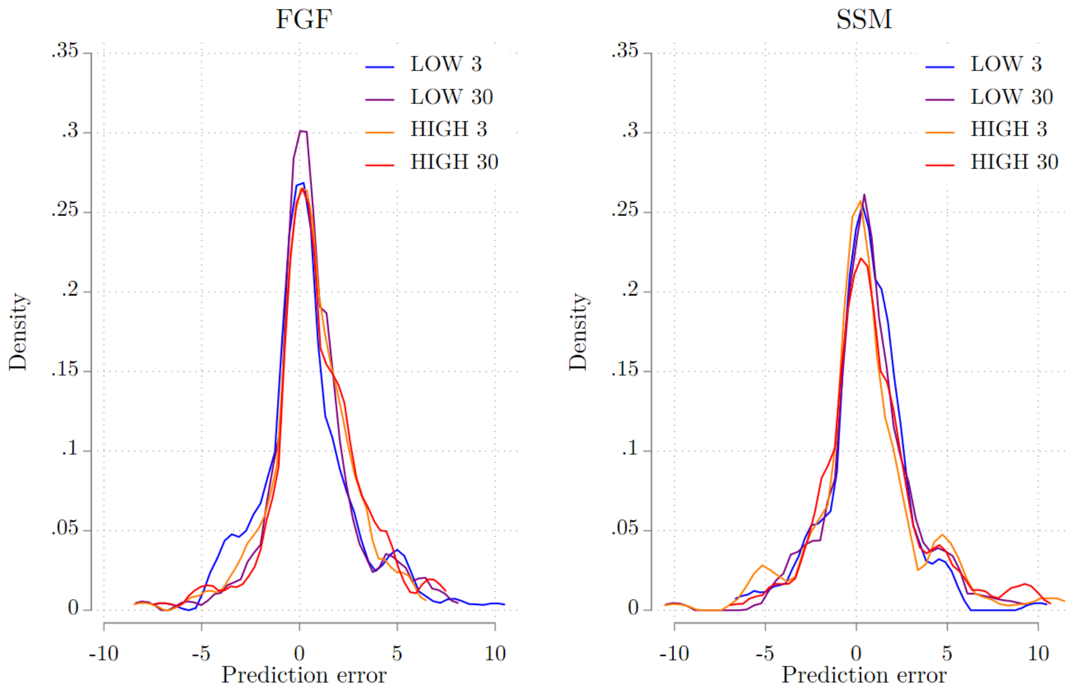


Figure 6 Distribution of ABC prediction errors ($\hat{c}_i - c_i$) in Study 2

Finally, Table 7 shows the regressions of the basic ABC model. Although we are holding α instead of MPCR constant, the results are qualitatively similar to Study 1: the coefficients of *predicted* and *beliefs* are economically and statistically significant for both FGF and SSM and the coefficients for *large* and *high* α are moderate and not statistically significant. Interestingly, the coefficients of *predicted* are higher under FGF than under SSM in all models, and for *beliefs* the opposite is the case. In contrast to Study 1, the R^2 is higher for models where a_i is elicited with FGF and the *CVMSE* is lower.

[12] Online Appendix Fig. A.5 separately presents the distributions of beliefs and contributions holding group size or MPCR constant. Fig. A.6 presents the distributions of second-order beliefs.

	(1) Pooled FGF	(2) Pooled SSM	(3) Large (30) FGF	(4) Large (30) SSM	(5) Small FGF	(6) Small SSM
predicted	0.784***	0.571***	0.902***	0.586***	0.671***	0.552***
	(0.0550)	(0.0801)	(0.0615)	(0.106)	(0.0836)	(0.121)
beliefs	0.180***	0.329***	0.111	0.379***	0.243**	0.297**
	(0.0689)	(0.101)	(0.0870)	(0.139)	(0.0987)	(0.146)
high α	0.190	0.147	0.190	0.244	0.258	0.0522
	(0.183)	(0.204)	(0.259)	(0.300)	(0.257)	(0.276)
large	0.301	0.217				
	(0.183)	(0.203)				
_cons	0.394*	0.713***	0.567**	0.627*	0.495*	0.977***
	(0.216)	(0.267)	(0.259)	(0.342)	(0.259)	(0.359)
R^2	0.551	0.421	0.571	0.429	0.537	0.415
$CVMSE$	5.158	6.524	5.603	6.460	4.865	6.441
N	600	600	300	300	300	300

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
OLS estimates. The dependent variable are the *contributions* c_i in the Phase 2 one-shot direct-response PGG. *predicted* are the ABC predictions using either FGF or SSM. *beliefs* are the b_i elicited in the Phase 2 PGG. *large* is an indicator that takes the value 1 for groups of 30 and 0 for groups of 3. *high* α is an indicator that takes the value 1 when the social return = 2.4 and 0 when the social return = 1.2. Column 1 uses the full sample and columns 2 and 3 split the sample by group size. CVMSE is the cross validated mean square error. Robust standard errors in parentheses.

Table 7 Regressions explaining Phase 2 contributions in Study 2 using the ABC approach

8 MPCR, group size, and the role of beliefs for cooperation

Confirming previous results in the literature (see Section 2.2), we find that, on average, MPCR (and α) has a positive effect on contributions (Figures 2 and 5, and Tables 3 and 6). Regarding group size the data show no clear pattern of the average effect on contributions (positive in Study 1 and negative in Study 2). [13]

The CSR approach (see Section 2.3) proposes that a_i and b_i can be functions of contextual features f (i.e., MPCR and group size). We found that f does not affect a_i but MPCR does affect beliefs b_i and contributions c_i. In theory, the CSR approach

[13]It might also be the case that the effect is non-linear, being positive for relatively small groups and negative for large increases in group size.

should internalize f in *predicted* (that is, $\hat{c}_i$), but, like Fischbacher and Gächter (2010), we found in both studies that *beliefs* matter on top of *predicted* to explain contributions (Tables 4 and 7). Although with large standard errors, Table 4 and Table 7 also show large coefficients for MPCR and group size. Here we explore whether f governs the effect of beliefs on contributions (e.g., beliefs might become a more important component in the ABC regressions as MPCR or group size increases). To test this formally, we re-run, for each study, the ABC regressions including beliefs interacted with group size and MPCR (or α).

Table 8 presents the results. Column 1 shows the results for Study 1 using SSM to elicit a_i. We find a strong negative effect of group size on contributions. The interaction of *beliefs* x *large* indicates that the effect size depends on the level of beliefs. Mathematically, $\frac{\partial c}{\partial \text{large}} = -0.803 + 0.245 * \textit{beliefs}$. The direct negative effect gets reversed if *beliefs* are high enough, that is, if $b_i > 0.803/0.245 = 3.28$ in our sample. These results are not replicated in the FGF sample (Column 2) but they are qualitatively similar in the pooled sample (Column 3).

Regarding MPCR, in column (1) we find a strong negative effect but with large standard errors. The interaction term *beliefs* x *MPCR* is positive, statistically significant and similar in magnitude to the interaction with group size. Mathematically, $\frac{\partial c}{\partial \text{MPCR}} = -0.636 + 0.236 * \textit{high MPCR}$. As in the group size case, the negative effect gets reversed if *beliefs* are high enough; here the threshold is $0.636/0.236 = 2.69$. Again, we do not reject any of these hypotheses in the FGF method but the results survive in the pooled sample.

Interestingly, we find that the beliefs threshold to make the overall group size effect positive (3.28) is higher than the beliefs threshold to make the overall MPCR effect positive (2.69). This could explain why there is more consensus in the effect of MPCR than on the group size effects in the literature.

Columns 4, 5 and 6 show the regression results for Study 2. Although the group size effect is not statistically significant, the results are qualitatively similar. We conjecture that this is due to the MPCR adjustment (i.e., holding the α constant instead of the MPCR lowers the incentives) offsetting the group size effect. Our interpretation is that group size (as well as MPCR effects) operate by affecting the weight of beliefs in the decision to contribute which is consistent with previous literature (e.g., Weimann et al. (2019); Kerr (1989)).

	(1) Study 1 SSM contrib.	(2) Study 1 FGF contrib.	(3) Study 1 Pooled contrib.	(4) Study 2 SSM contrib.	(5) Study 2 FGF contrib.	(6) Study 2 Pooled contrib.
predicted	0.601***	0.550***	0.570***	0.571***	0.783***	0.684***
	(0.069)	(0.092)	(0.0582)	(0.080)	(0.056)	(0.049)
beliefs	0.189*	0.364***	0.288***	0.263**	0.127	0.191**
	(0.114)	(0.126)	(0.0851)	(0.116)	(0.099)	(0.076)
large	-0.803**	-0.345	-0.542	-0.286	-0.025	-0.146
	(0.420)	(0.523)	(0.338)	(0.474)	(0.328)	(0.287)
beliefs x large	0.245***	0.043	0.139*	0.119	0.084	0.102
	(0.093)	(0.122)	(0.0776)	(0.112)	(0.082)	(0.070)
high MPCR	-0.636	-0.036	-0.316			
	(0.433)	(0.523)	(0.342)			
beliefs x high MPCR	0.236**	0.062	0.147*			
	(0.099)	(0.121)	(0.0786)			
high α				0.049	0.086	0.069
				(0.478)	(0.330)	(0.288)
beliefs x high α				0.024	0.030	0.027
				(0.110)	(0.084)	(0.070)
_cons	1.123	0.706	0.872***	0.996***	0.597*	0.775***
	(0.426)	(0.347)	(0.270)	(0.380)	(0.303)	(0.240)
R^2	0.560	0.490	0.520	0.422	0.551	0.486
N	468	468	936	600	600	1200

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
OLS estimates. The dependent variables are the *contributions* c_i in the Phase 2 one-shot PGG. *beliefs* are the b_i elicited in Phase 2. *predicted* is the ABC prediction using either FGF or SSM. *large* is an indicator that takes the value 1 for groups of 9 or 30 and 0 for groups of 3. *high MPCR* is an indicator that takes the value 1 for MPCR=0.8 and 0 for MPCR=0.4. *high* α is an indicator that takes the value 1 when the social return = 2.4 and 0 when the social return = 1.2. Robust standard errors in parentheses.

Table 8 The role of beliefs in explaining cooperation, controlling for ABC predictions, group size, and MPCR

9 Summary

In this paper we revisited a classic question in the experimental economics literature on public goods games: the role of group size and MPCR for voluntary cooperation. Our starting point was the observation that many people have a preference for conditional cooperation (Fischbacher et al., 2001; Fischbacher and Gächter, 2010; Chaudhuri, 2011; Thöni and Volk, 2018), which renders the beliefs about others' cooperation important. The specific question we addressed in this paper was how group size and MPCR affect preferences (that is, attitudes to cooperation) and beliefs, and how preferences and beliefs jointly explain cooperation.

For our analysis we used the ABC approach which measures an individual i's attitudes to cooperation (a_i) and beliefs (b_i) to explain i's cooperation (c_i): $a_i(b_i) \rightarrow c_i$, which has proven to be a good framework to explain, quantitatively, the level of cooperation we see in small-group experimental public goods games. The ABC approach we used in this paper $[a_i(f, b_i(f)) \rightarrow c_i]$ allows to go beyond simply observing resulting levels of cooperation as a function of the contextual features of the public good game: the contextual features f (group size and MPCR parameters in our case) might affect a_i or b_i or both, and it is their interactions $a_i(b_i)$ that explains the cooperation levels c_i we observe.

The ABC approach requires measuring the attitudes to cooperation a_i, for which (Fischbacher et al., 2001) (FGF) introduced an incentive-compatible method. The ABC approach has been successfully used in small groups of four players with very similar MPCR parameters (Fischbacher and Gächter, 2010; Fischbacher et al., 2012; Gächter et al., 2017, 2022; Isler et al., 2021). Until this paper, there has been no evidence on how the ABC approach fares in larger groups and with different MPCR parameters. Moreover, the FGF method to measure attitudes to cooperation becomes approximately hypothetical as group size increases. We therefore also introduced SSM - a new incentive-scalable version of the strategy method, which keeps the incentives for the elicited a_i constant for any group size. But SSM has the downside that the method to incentivize a_i implies that the public good is not well defined (see Section 3.2 and Appendix).

We conducted two studies. In Study 1 we applied the SSM and compared it with the traditional FGF method with group sizes 3 and 9 and MPCRs of 0.4 and 0.8. We found that the ABC approach on average correctly predicted actual contribution levels in all conditions. In terms of predictive success and using cross-validated mean squared error and R^2 as criteria, the SSM method did at least as well as FGF, in particular in large groups. In Study 2, however, with group sizes of 3 and 30, and MPCRs of 0.04, 0.08, 0.4 and 0.8, FGF performed better in terms of R^2 and cross-validated mean squared error (possibly due to the composition of the sample). Again, the ABC approach predicts contribution levels in all conditions.

Our results are striking: none of the game parameters in both studies, nor the elicitation method of cooperative attitudes (FGF or SSM) affects attitudes to cooperation

(a_i) in a significant way, with the exception of groups of 30 and low MPCR, where we find a_i to be slightly flatter than in other conditions. In terms of cooperation behavior, we find that higher MPCRs lead to higher beliefs and higher cooperation, confirming previous evidence. Holding MPCR constant, we find no statistically significant evidence for group size effects: group size neither affects a_i, nor b_i, nor c_i. However, group size magnifies the effects of beliefs on cooperation. Similarly, we conclude that MPCR effects work via increased beliefs about others' cooperativeness and not via changed attitudes to cooperation.

On a methodological note, our results on the two strategy methods show that the lack of incentive scalability of FGF does not matter much empirically (see also Figure A.3 and A.7 in the Online Appendix). This holds when comparing the elicited a_i under FGF and under SSM (see Figs. 1 and 4) and also for the ABC predictions (Figs. 3 and 6 and Tables 4 and 7). However, while SSM is incentive scalable by design it has the conceptual problem that the public good is not well defined but the fact that the elicited a_i are similar also shows that this conceptual problem is empirically not very important. Thus, it is up to the taste of researchers which method to use when studying group size effects in future research.

In summary, we conclude that the ABC approach works well to explain cooperation for a wide range of parameter sets and for either strategy method we use. Future research should therefore explore the applicability of the ABC approach in further public goods settings of economic interest.

Statements and Declarations

The authors declare that they have no conflict of interest.

Acknowledgements

This work was supported by the European Research Council [Grant Number ERC-AdG 101020453 PRINCIPLES] and the Centre of Decision Research and Experimental Economics (CeDEx) at the University of Nottingham. We are grateful for comments by Urs Fischbacher, Heike Hennig-Schmidt, Patrick Maus, Louis Putterman, Silvia Sonderegger, Kieran Stockley and Ro'i Zultan as well as those received at seminars at the University of Nottingham, the 8th Annual Workshop on Behavioral Game Theory at the University of East Anglia (2022), the International Conference on Social Dilemmas (2022), the 13th Conference of the French Experimental Economics Association (2023), the M-BEES conference in Maastricht (2023) and the ESA World Meeting in Lyon

(2023). S.G. gratefully acknowledges the hospitality of briq Bonn while working on this paper.

Data Availability

Data and analysis files are posted on https://osf.io/7psud/.

Online Appendix

The Online Appendix contains the experimental instructions and additional figures.

References

Andreoni, J., "Cooperation in public-goods experiments: kindness or confusion?," *American Economic Review*, 1995a, pp. 891–904.

_ , "Warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments," *The Quarterly Journal of Economics*, 1995b, *110* (1), 1–21.

_ **, M. A. Kuhn, and C. Sprenger**, "Measuring time preferences: A comparison of experimental methods," *Journal of Economic Behavior & Organization*, 2015, *116*, 451–464.

Arechar, A., S. Gächter, and L. Molleman, "Conducting interactive experiments online," *Experimental Economics*, 2018, *21* (1), 99–131.

Bayer, R., E. Renner, and R. Sausgruber, "Confusion and learning in the voluntary contributions game," *Experimental Economics*, 2013, *16* (4), 478–496.

Blanco, Esther, Tobias Haller, Maria Claudia Lopez, and James M Walker, "The tension between private benefits and degradation externalities from appropriation in the commons," *Journal of Economic Behavior & Organization*, 2016, *125*, 136–147.

Brandts, J. and A. Schram, "Cooperation and noise in public goods experiments: applying the contribution function approach," *Journal of Public Economics*, 2001, *79* (2), 399–427.

— **and G. Charness**, "The strategy versus the direct-response method: a first survey of experimental comparisons," *Experimental Economics*, 2011, *14* (3), 375–398.

Burton-Chellew, M., C. El Mouden, and S. West, "Conditional cooperation and confusion in public-goods experiments," *Proceedings of the National Academy of Sciences*, 2016, *113* (5), 1291–1296.

Carpenter, J. P., "Punishing free-riders: How group size affects mutual monitoring and the provision of public goods," *Games and Economic Behavior*, 2007, *60* (1), 31–51.

Chang, L., A. Smith, M. Dufwenberg, and A. Sanfey, "Triangulating the neural, psychological, and economic bases of guilt aversion," *Neuron*, 2011, *70* (3), 560–572.

Chaudhuri, A., "Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature," *Experimental Economics*, 2011, *14* (1), 47–83.

Croson, R., "Theories of commitment, altruism and reciprocity: Evidence from linear public goods games," *Economic Inquiry*, 2007, *45* (2), 199–216.

Diederich, J., T. Goeschl, and I. Waichman, "Group size and the (in) efficiency of pure public good provision," *European Economic Review*, 2016, *85*, 272–287.

Drouvelis, M., *Social preferences: an introduction to behavioural economics and experimental research*, Newcastle upon Tyne: Agenda Publishing, 2021.

Dufwenberg, M., S. Gächter, and H. Hennig-Schmidt, "The framing of games and the psychology of play," *Games and Economic Behavior*, 2011, *73* (2), 459–478.

Eichenseer, M. and J. Moser, "Conditional cooperation: Type stability across games," *Economics Letters*, 2020, *188*, 108941.

Fehr, E. and I. Schurtenberger, "Normative foundations of human cooperation," *Nature Human Behaviour*, 2018, *2* (7), 458–468.

— **and K. Schmidt**, "A theory of fairness, competition, and cooperation," *The Quarterly Journal of Economics*, 1999, *114* (3), 817–868.

Fischbacher, U. and S. Gächter, "Social preferences, beliefs, and the dynamics of free riding in public goods experiments," *American Economic Review*, 2010, *100* (1), 541–56.

— **, — , and E. Fehr**, "Are people conditionally cooperative? Evidence from a public goods experiment," *Economics Letters*, 2001, *71* (3), 397–404.

__ , __ , and S. Quercia, "The behavioral validity of the strategy method in public good experiments," *Journal of Economic Psychology*, 2012, *33* (4), 897–913.

Fischbacher, Urs, Simeon Schudy, and Sabrina Teyssier, "Heterogeneous reactions to heterogeneity in returns from public goods," *Social Choice and Welfare*, 2014, *43*, 195–217.

Fisher, Joseph, R Mark Isaac, Jeffrey W Schatzberg, and James M Walker, "Heterogenous demand for public goods: Behavior in the voluntary contributions mechanism," *Public Choice*, 1995, *85*, 249–266.

Fosgaard, T. R., L. G. Hansen, and E. Wengström, "Framing and misperception in public good experiments," *The Scandinavian Journal of Economics*, 2017, *119* (2), 435–456.

Gächter, S. and B. Herrmann, "Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment," *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2009, *364* (1518), 791–806.

__ and __ , "The limits of self-governance when cooperators get punished: Experimental evidence from urban and rural Russia," *European Economic Review*, 2011, *55* (2), 193–210.

__ , F. Kölle, and S. Quercia, "Reciprocity and the tragedies of maintaining and providing the commons," *Nature Human Behaviour*, 2017, *1* (9), 650–656.

__ , F. Kölle, and S. Quercia, "Preferences and perceptions in Provision and Maintenance public goods," *Games and Economic Behavior*, 2022, *135*, 338–355.

Gächter, Simon and Elke Renner, "Leaders as role models and 'belief managers' in social dilemmas," *Journal of Economic Behavior & Organization*, 2018, *154*, 321–334.

Goeree, J. K., C. A. Holt, and S. K. Laury, "Private costs and public benefits: unraveling the effects of altruism and noisy behavior," *Journal of Public Economics*, 2002, *83* (2), 255–276.

Gunthorsdottir, A, D Houser, K McCabe, and H Ameden, "Disposition, history and contributions in public good experiments," *Journal of Economic Behavior and Organization*, 2007, *62*, 304–315.

Guttman, J., "Matching behavior and collective action: Some experimental evidence," *Journal of Economic Behavior & Organization*, 1986, *7* (2), 171–198.

Hergueux, J. and N. Jacquemet, "Social preferences in the online laboratory: a randomized experiment," *Experimental Economics*, 2015, *18* (2), 251–283.

Houser, D. and R. Kurzban, "Revisiting kindness and confusion in public goods experiments," *American Economic Review*, 2002, *92* (4), 1062–1069.

Isaac, R. and J. Walker, "Group size effects in public goods provision: The voluntary contributions mechanism," *The Quarterly Journal of Economics*, 1988, *103* (1), 179–199.

_ , J. M Walker, and S. Thomas, "Divergent evidence on free riding: An experimental examination of possible explanations," *Public Choice*, 1984, *43* (2), 113–149.

Isaac, R. M., J. M. Walker, and A. W. Williams, "Group size and the voluntary provision of public goods: Experimental evidence utilizing large groups," *Journal of Public Economics*, 1994, *54* (1), 1–36.

Isler, O., S. Gächter, A. Maule, and C. Starmer, "Contextualised strong reciprocity explains selfless cooperation despite selfish intuitions and weak social heuristics," *Scientific Reports*, 2021, *11* (1), 1–17.

Katuščák, P. and T. Miklánek, "What drives conditional cooperation in public good games?," *Experimental Economics*, 2023, *26*, 435–467.

Kerr, N., "Illusions of efficacy: The effects of group size on perceived efficacy in social dilemmas," *Journal of Experimental Social Psychology*, 1989, *25* (4), 287–313.

Keser, C. and H. Kliemt, "The Strategy Method as an instrument for the exploration of limited rationality in oligopoly game behavior (Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperimentes)(by Reinhard Selten)," in "The Art of Experimental Economics," Routledge, 2021, pp. 30–36.

Ledyard, J. O., "Public goods: A survey of experimental research," *Handbook of Experimental Economics*, 1995, *883* (111-194), 884.

List, J. A., "Young, selfish and male: Field evidence of social preferences," *The Economic Journal*, 2004, *114* (492), 121–149.

Lugovskyy, Volodymyr, Daniela Puzzello, Andrea Sorensen, James Walker, and Arlington Williams, "An experimental study of finitely and infinitely repeated linear public goods games," *Games and Economic Behavior*, 2017, *102*, 286–302.

Marwell, G. and R. E. Ames, "Experiments on the provision of public goods. I. Resources, interest, group size, and the free-rider problem," *American Journal of Sociology*, 1979, *84* (6), 1335–1360.

Mill, W. and M. Theelen, "Social value orientation and group size uncertainty in public good dilemmas," *Journal of Behavioral and Experimental Economics*, 2019, *81*, 19–38.

Mullett, T., R. L. McDonald, and G. Brown, "Cooperation in Public Goods Games Predicts Behavior in Incentive-Matched Binary Dilemmas: Evidence for Stable Prosociality," *Economic Inquiry*, 2020, *58* (1), 67–85.

Nosenzo, D., S. Quercia, and M. Sefton, "Cooperation in small groups: the effect of group size," *Experimental Economics*, 2015, *18* (1), 4–14.

Palan, S. and C. Schitter, "Prolific.ac—A subject pool for online experiments," *Journal of Behavioral and Experimental Finance*, 2018, *17*, 22–27.

Palfrey, T. R. and J. E. Prisbrey, "Anomalous behavior in public goods experiments: How much and why?," *American Economic Review*, 1997, pp. 829–846.

Pereda, M., I. Tamarit, A. Antonioni, J. Cuesta, P. Hernández, and A. Sánchez, "Large scale and information effects on cooperation in public good games," *Scientific Reports*, 2019a, *9* (1), 1–10.

_ , V. Capraro, and A. Sánchez, "Group size effects and critical mass in public goods games," *Scientific Reports*, 2019b, *9* (1), 1–10.

Schonlau, M., "CROSSVALIDATE: Stata module to perform k-fold crossvalidation," Statistical Software Components, Boston College Department of Economics 2020.

Selten, R., "Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments," *Beiträge zur Experimentellen Wirtschaftsforschung, Tübingen*, 1967, pp. 136–168.

Spadaro, G., I. Tiddi, S. Columbus, S. Jin, A. Ten Teije, CoDa Team, and D. Balliet, "The Cooperation Databank: Machine-Readable Science Accelerates Research Synthesis," *Perspectives on Psychological Science*, 2022, *17*, 1472–1489.

Struwe, Natalie, Esther Blanco, and James M Walker, "No response to changes in marginal incentives in one-shot public good experiments," Technical Report 2023.

Sugden, R., "Reciprocity: the supply of public goods through voluntary contributions," *The Economic Journal*, 1984, *94* (376), 772–787.

Thöni, C. and S. Volk, "Conditional cooperation: Review and refinement," *Economics Letters*, 2018, *171*, 37–40.

van den Berg, Pieter, Siyuan Liu, Tom Wenseleers, and Jianlei Zhang, "Misrepresentation of group contributions undermines conditional cooperation in a human decision making experiment," *Scientific Reports*, 2022, *12* (1), 12320.

Volk, S., C. Thöni, and W. Ruigrok, "Temporal stability and psychological foundations of cooperation preferences," *Journal of Economic Behavior & Organization*, 2012, *81* (2), 664–676.

Weber, T., O. Weisel, and S. Gächter, "Dispositional free riders do not free ride on punishment," *Nature Communications*, 2018, *9* (1), 1–9.

Weber, Till O, Jonathan F Schulz, Benjamin Beranek, Fatima Lambarraa-Lehnhardt, and Simon Gächter, "The behavioral mechanisms of voluntary cooperation across culturally diverse societies: Evidence from the US, the UK, Morocco, and Turkey," *Journal of Economic Behavior & Organization*, 2023, *215*, 134–152.

Weimann, J., J. Brosig-Koch, T. Heinrich, H. Hennig-Schmidt, and C. Keser, "Public good provision by large groups–the logic of collective action revisited," *European Economic Review*, 2019, *118*, 348–363.

Zelmer, J., "Linear public goods experiments: A meta-analysis," *Experimental Economics*, 2003, *6* (3), 299–310.

Appendix: The Scalable Strategy Method (SSM)

Here we describe some further technical details of how SSM works and illustrate with an example. The instructions are available in the Online Appendix and a Qualtrics version (Qualtrics *qsf* format) is available at https://osf.io/7psud/.

Procedure

Consider a group of three people playing a PGG. Table 9 displays sample responses to the unconditional contribution and the contribution table (i.e., first 11 rows). For illustrative purposes, let's say we have a conditional cooperator, a free rider and a triangle cooperator. Each group member's payoff is randomly determined from either their unconditional contribution or from their contribution table, with 50% probability each. Since this is true for any group size, SSM guarantees the incentive scalability of the strategy method. Because both the unconditional contribution and the contribution table are potentially payoff relevant, participants have an incentive to take both the unconditional and the conditional contribution seriously.

The steps for calculating individual payoffs are the following:

- First: Calculate the *computed conditional contribution* for i by plugging in the average contribution of the remaining players into player i's contribution table. The contribution table is a function that depends on the average unconditional contributions (rounded to the nearest integer) of the other players. Example: Player 1 will have a computed conditional contribution of 2 because the average of the other two players' unconditional contributions is $(0 + 4)/2 = 2$. Then if we plug in 2 in Player 1's contribution table we get a computed conditional contribution of 2 (because Player 1 indicates in their contribution table that they will contribute 2 if others contribute 2 on average). With the same reasoning, Player 2's and Player 3's computed conditional contribution will be 0 and 2, respectively.

- Second: We proceed to compute the individual payoff for each player independently. For each player i, we randomize to use either their unconditional contribution or their contribution table (each with 50% probability).
  - If player i's *unconditional contribution* is randomly selected, we use all other player's computed conditional contributions to determine i's individual payoff (but not necessarily for the payoffs of the other players, as those are

computed independently using the same procedure). Example: If Player 1 is assigned to the unconditional contribution, he/she would contribute 5. To compute player i's payoff, we also need the contributions of the other group members. For those we will use their computed conditional contributions. Example of Player 1's payoff: $10 - 5 + \frac{\alpha(5+0+2)}{3}$. Note that player 1's payoff is not affected by the outcomes of the randomization for players 2 and 3.

- If player i's *contribution table* is randomly selected, we use player i's computed conditional contribution and all other player's unconditional contributions to determine i's individual payoff (but not necessarily for the payoffs of the other players, as those are computed independently using the same procedure). Example: Player 1 contributes 2. To compute player i's payoff we also need the contributions of the other players. For those, we will use the unconditional contributions. Example of Player 1's payoff: $10-2+\frac{\alpha(2+0+4)}{3}$. Note again that player 1's payoff is not affected by the outcomes of the randomization for players 2 and 3.

Type examples	Player 1 Conditional Cooperator	Player 2 Free Rider	Player 3 Triangle Cooperator
Unconditional cont.	5	0	4
Conditional table			
0	0	0	0
1	1	0	1
2	2	0	1
3	3	0	2
4	4	0	3
5	4	0	4
6	5	0	3
7	5	0	3
8	6	0	2
9	6	0	2
10	7	0	1
Computed conditional cont.	(2)	(0)	(2)

Table 9 Example of responses. The parenthesis in the last row indicate that the numbers are computed by the researcher using the other rows as explained in the procedure.

Remarks

- For each player, the probability that his/her contribution is determined from the unconditional contribution or the contribution table is guaranteed to be 50%

each. This is a result of the randomization in the first step of the procedure. Participants are told about this in the instructions (see Online Appendix).

- Incentive scalability (keeping incentive compatibility constant across group sizes) requires that the probability of selecting the unconditional contribution and the contribution table remains constant for any group size. We can only achieve incentive scalability by changing at what level the randomization happens:
  - FGF randomizes at the level of the group and randomly selects *one* group member to determine their contribution from their contribution table, while all other group members' contribute their unconditional contribution. Payoffs are therefore interdependent like in any standard public good game (e.g., if player 1's contribution is derived from player 1's contribution table, the others players automatically contribute their unconditional contribution). This feature of payoff interdependence renders FGF not scalable (see Section 3.1).
  - By contrast, SSM randomizes at the level of individual group members, that is, for *each* group member i, i's unconditional contribution or their contribution table is randomly selected. Hence, the outcome of the randomization for player i has no payoff consequences for the other players. Therefore, because the randomization between unconditional contribution and contribution table affects individual i only, payoffs under SSM are not interdependent but individual-specific (see section 3.2).

Online Appendix to "Using the Strategy Method and Beliefs to Explain Group Size and MPCR Effects in Public Good Experiments"

Simon Gächter and Diego Marino Fages

Contents

We present one version of the instructions as an example (all surveys in full are available in pdf and Qualtrics format at https://osf.io/7psud/).

A: Additional Figures

Figure A.1 (Panel A) presents the distribution of beliefs b_i holding constant the group size or the MPCR. We run Kolmogorov-Smirnoff (K-S) tests for each case and cannot reject the null hypothesis of equal distributions in any comparison at the 5% confidence level. Similarly, Figure A.1 (Panel B) presents the contributions c_i in the one-shot PGG for each combination of MPCR and group size. We only reject the K-S test for the small group at the 5% confidence level (p-value=0.011).

Figure A.2 presents the distribution of second order beliefs holding constant the group size or the MPCR. We run K-S tests for each case and we only reject the null in the small group (p-value=0.043).

Figure A.3 (Panel A) presents the distribution of beliefs in each treatment by strategy method. We run K-S tests by treatment and pooling by method and cannot reject any differences in the distributions (the lowest p-value is 0.224). Similarly, Figure A.3 (Panel B) presents the contributions in the one-shot PGG with no clear difference between the methods (the lowest p-value is 0.570).

Figure A.4 presents the distributions of the ABC prediction errors in Stage 2, $(\hat{c}_i - c_i)$, separated by the strategy method applied in Stage 1 (FGF vs. SSM). The four panels depict these distributions for high and low MPCR (HIGH = 0.8; LOW = 0.3) and small and large group sizes (3 and 9). K-S tests return p-value=0.046 for High 9; p-value=0.372 for High 3; p-value=0.947 for Low 0; and p-value=0.466 for Low 9. Figure A.5 presents the beliefs and contributions in Study 2. Panel A presents the distribution of beliefs holding constant the group size or the α. We run K-S tests for each case and only reject the null for α in the small group (p-value=0.002). Similarly, Panel B presents the contributions in the one-shot PGG in each case. The K-S tests are only rejected for the α effect in small groups (p-value=0.031).

Figure A.6 presents the distribution of second order beliefs holding constant the group size or the α in Study 2. We run K-S tests for each case and we only reject the null in the small group (p-value=0.007).

Figure A.7 presents the distribution of beliefs and contributions in each treatment by strategy method. For beliefs, we only reject the null in K-S tests when we pool all treatments (p-value is 0.050). For contributions, we only reject in small groups and low α (p-value is 0.059).

Figure A.8 presents the distributions of the ABC prediction errors in Stage 2, separated by the strategy method applied in Stage 1 (FGF vs. SSM). The four panels depict these distributions for high and low α (1.2 and 2.4) and small and large group sizes (3 and 30). K-S tests do not reject the null for any group (lowest p-value=0.106).

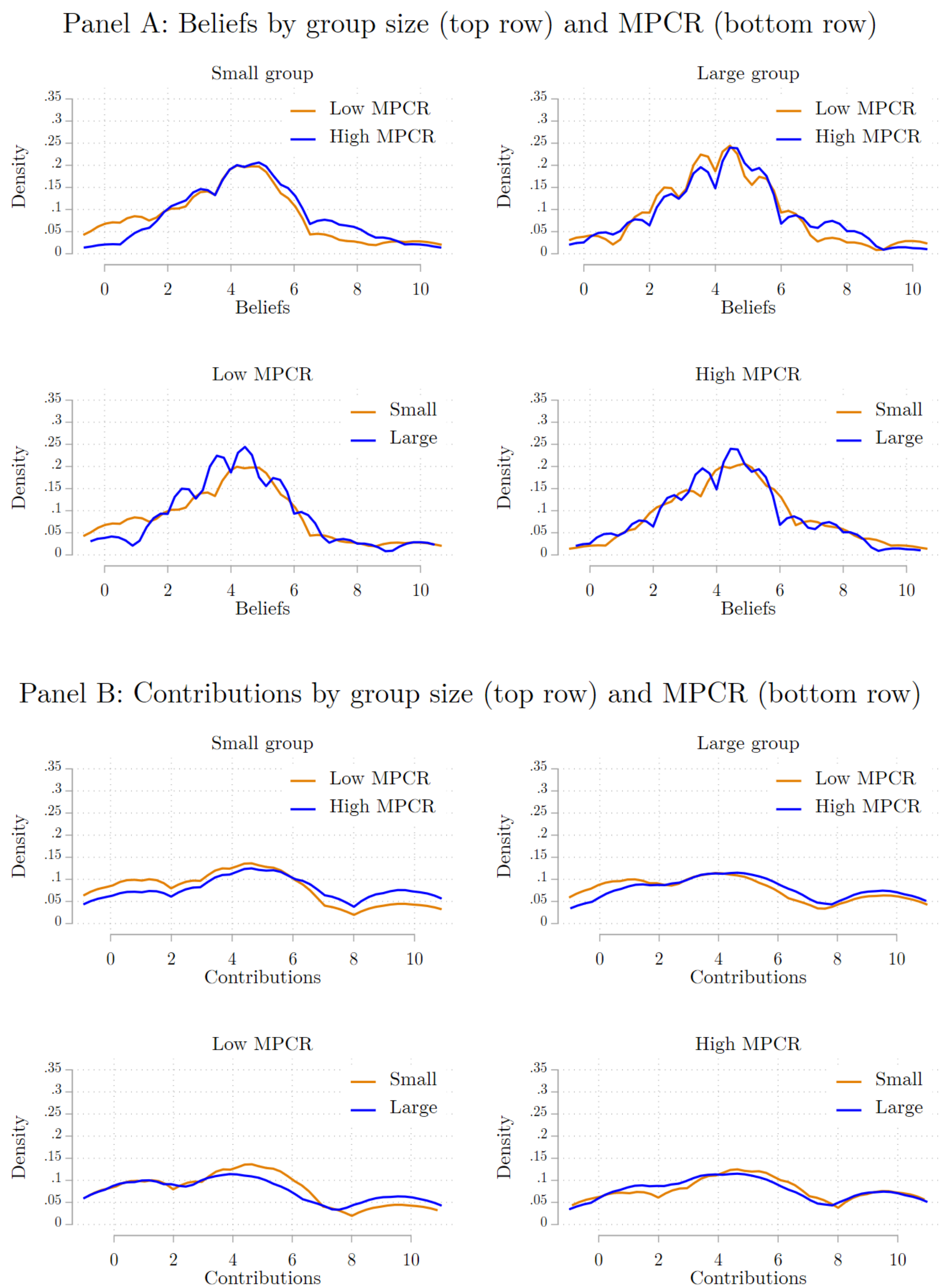


Figure A.1 Distributions of beliefs b_i (Panel A) and contributions b_i (Panel B) as elicited in direct response in Stage 2 of Study 1

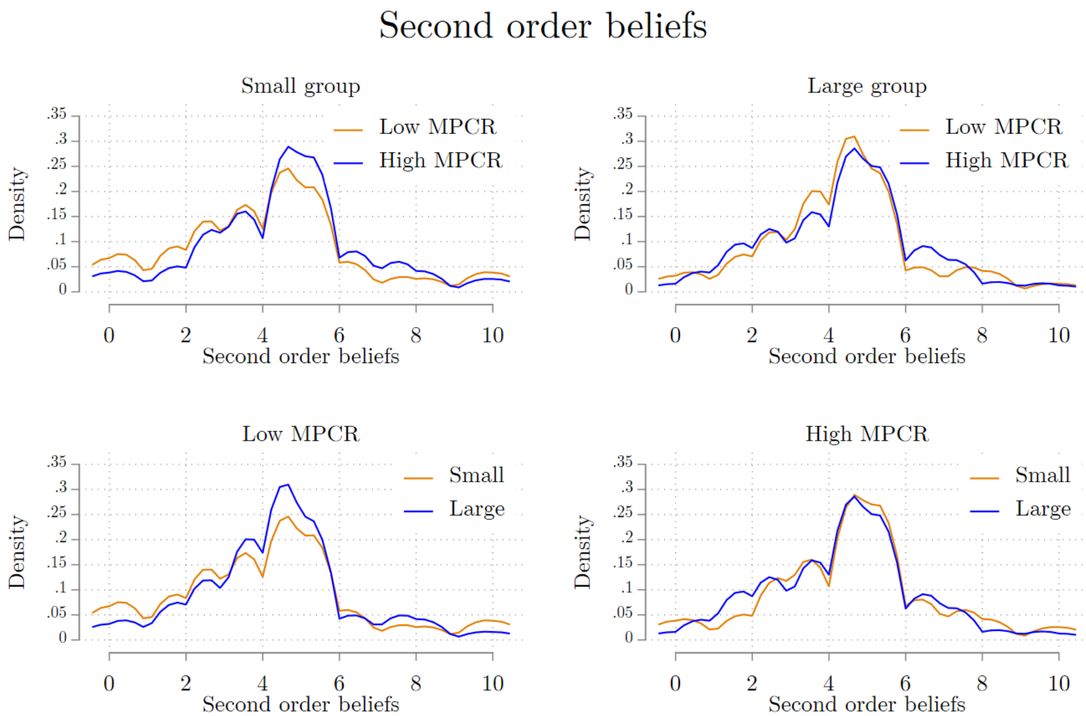
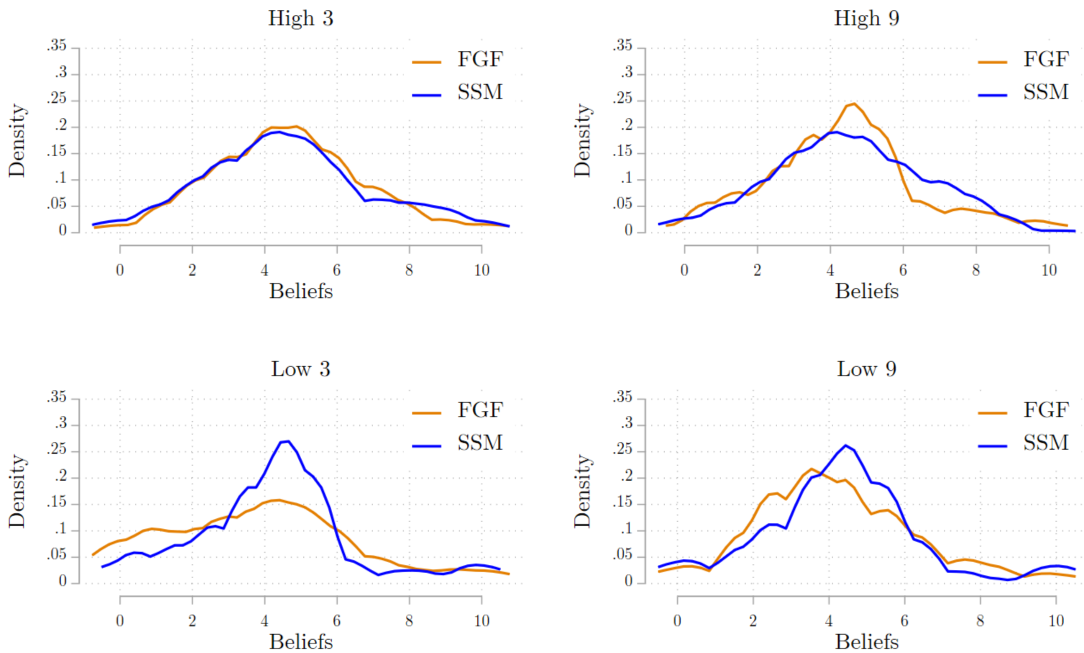


Figure A.2 Study 1: Second order beliefs by group size and MPCR

Panel A: Beliefs by group size (top row) and MPCR (bottom row)



Panel B: Contributions by group size (top row) and MPCR (bottom row)

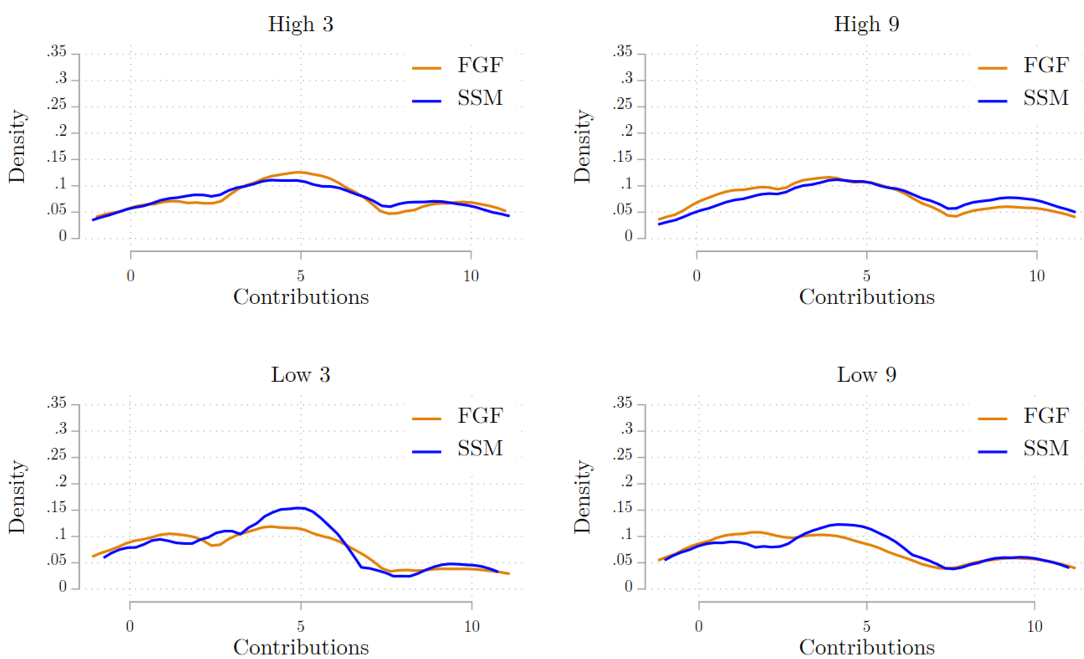


Figure A.3 Study 1: Beliefs and contributions by method

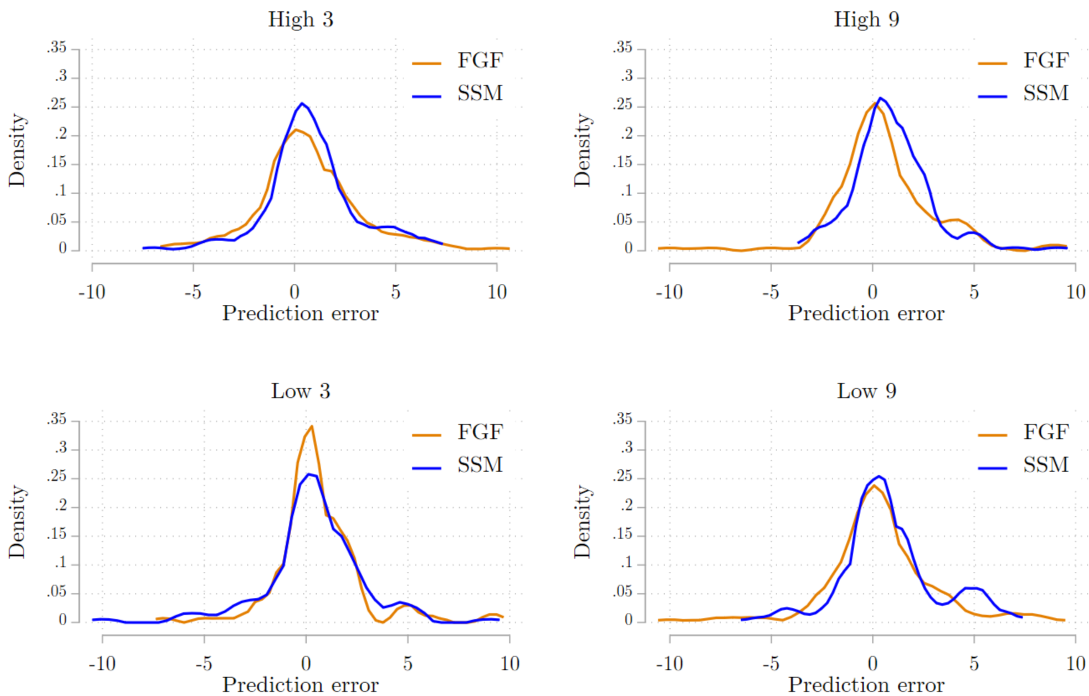
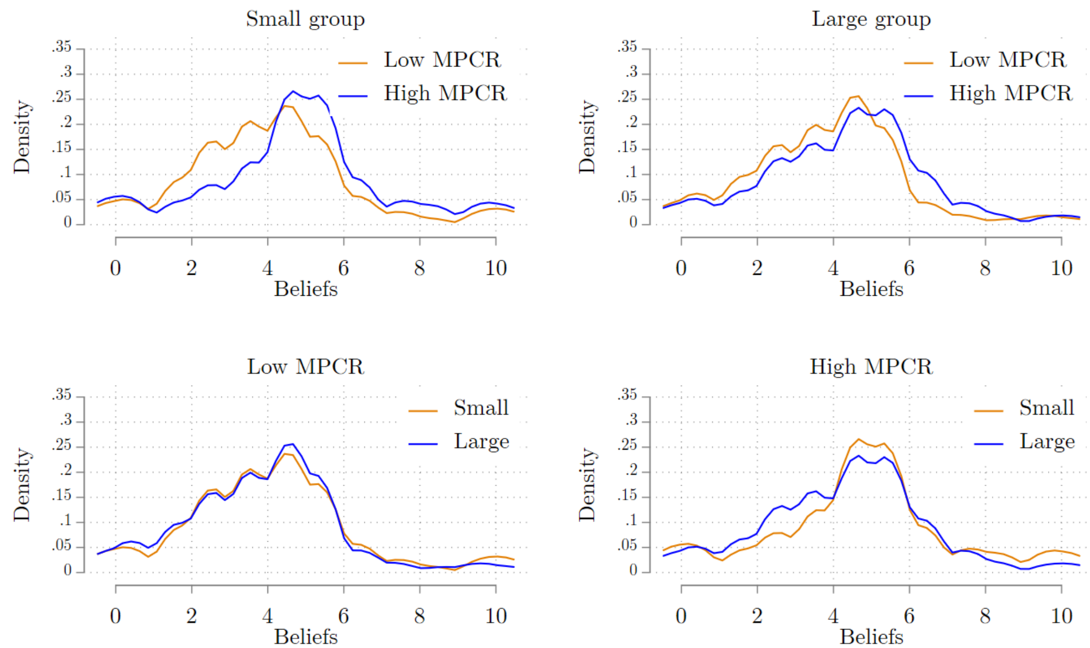


Figure A.4 Distribution of ABC prediction errors in Stage 2 ($\hat{c}_i - c_i$) in Study 1 by Stage 1 strategy method

Panel A: Beliefs by group size (top row) and MPCR (bottom row)



Panel B: Contributions by group size (top row) and MPCR (bottom row)

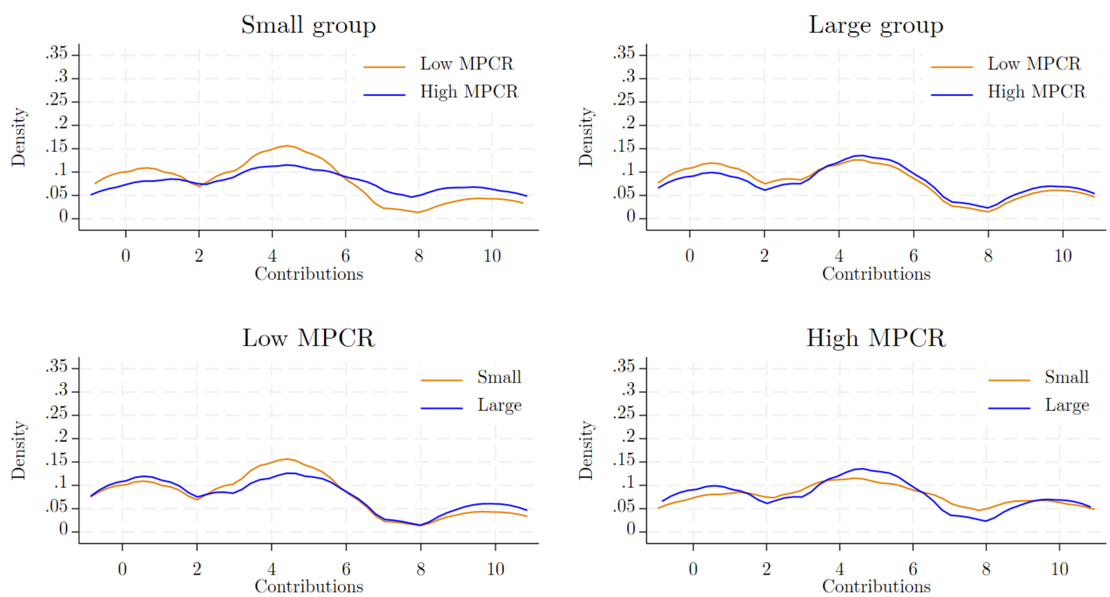


Figure A.5 Study 2: Beliefs and contributions by treatment

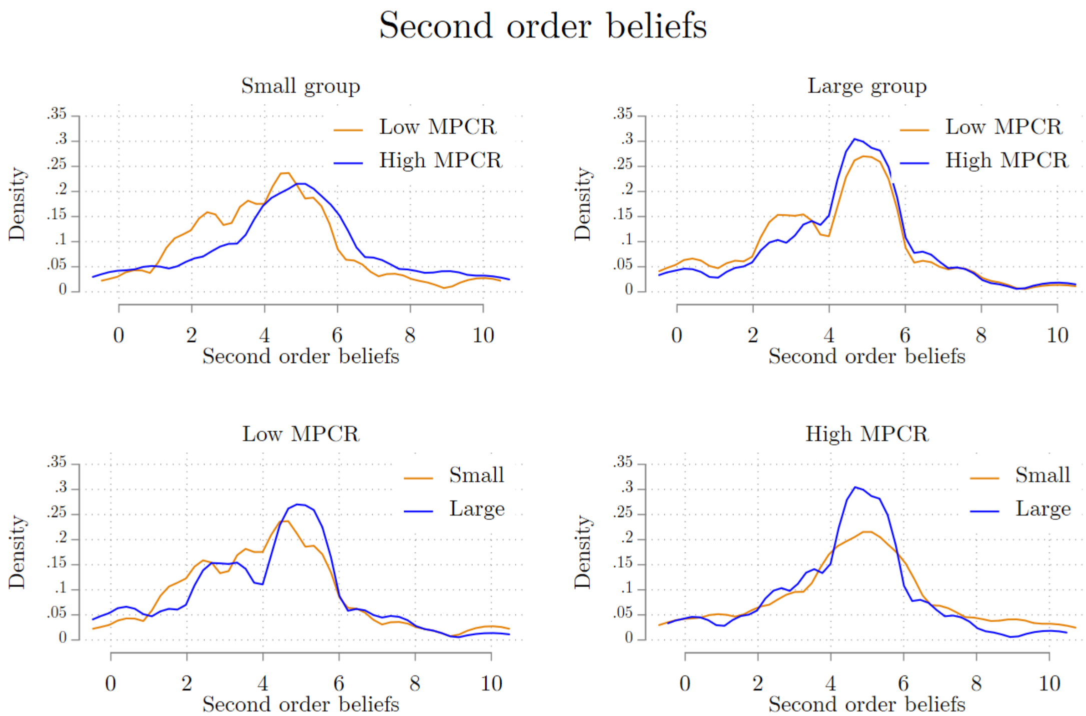


Figure A.6 Study 2: Second order beliefs by group size and MPCR

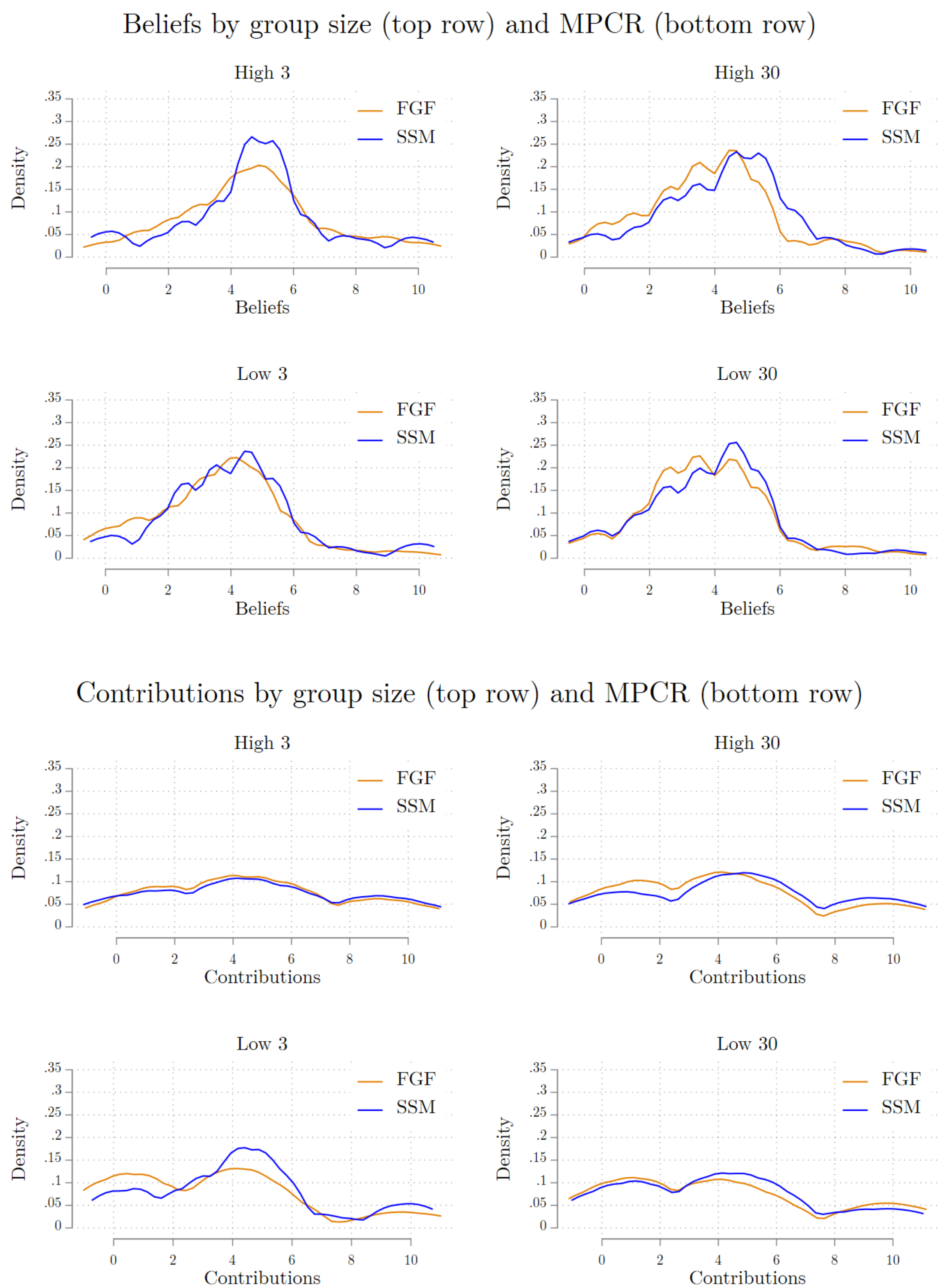


Figure A.7 Study 2: Beliefs and contributions by method

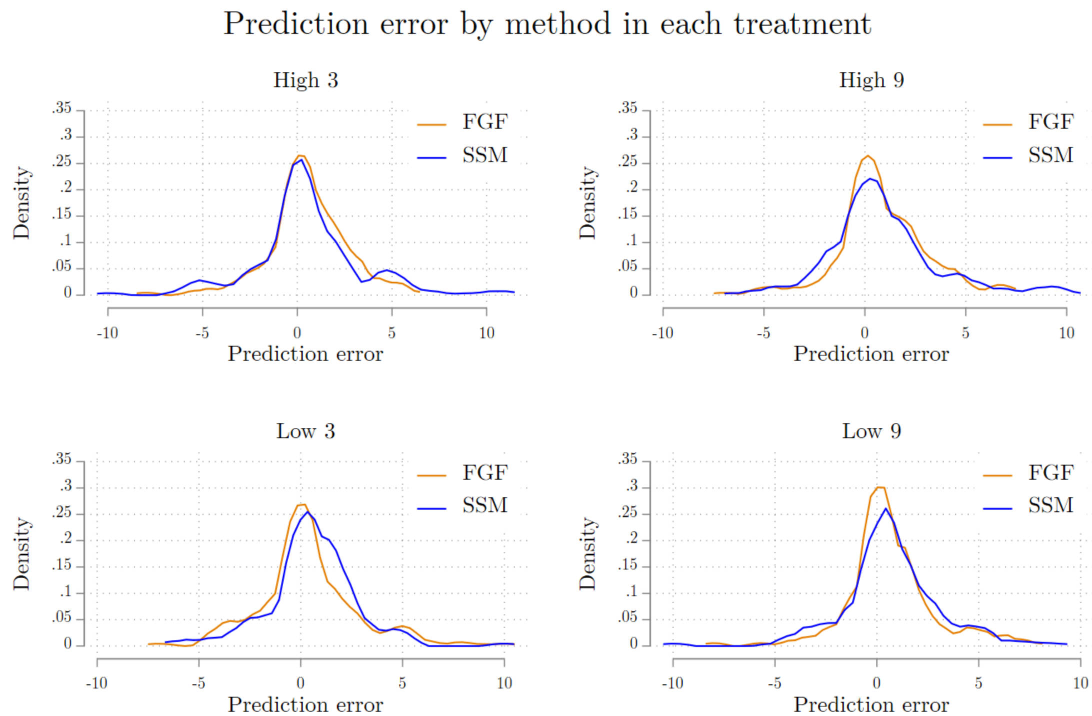


Figure A.8 Distribution of ABC prediction errors in Stage 2 ($\hat{c}_i - c_i$) in Study 2 by Stage 1 strategy method

B: Robustness using the non-parametric prediction

In this section we present the results of Figure 3 and Table 4 from the paper using the non parametric estimation of predicted (see Section 5.3 in the paper). To replicate other figures just replace predicted with predicted_abc in the corresponding dofile.

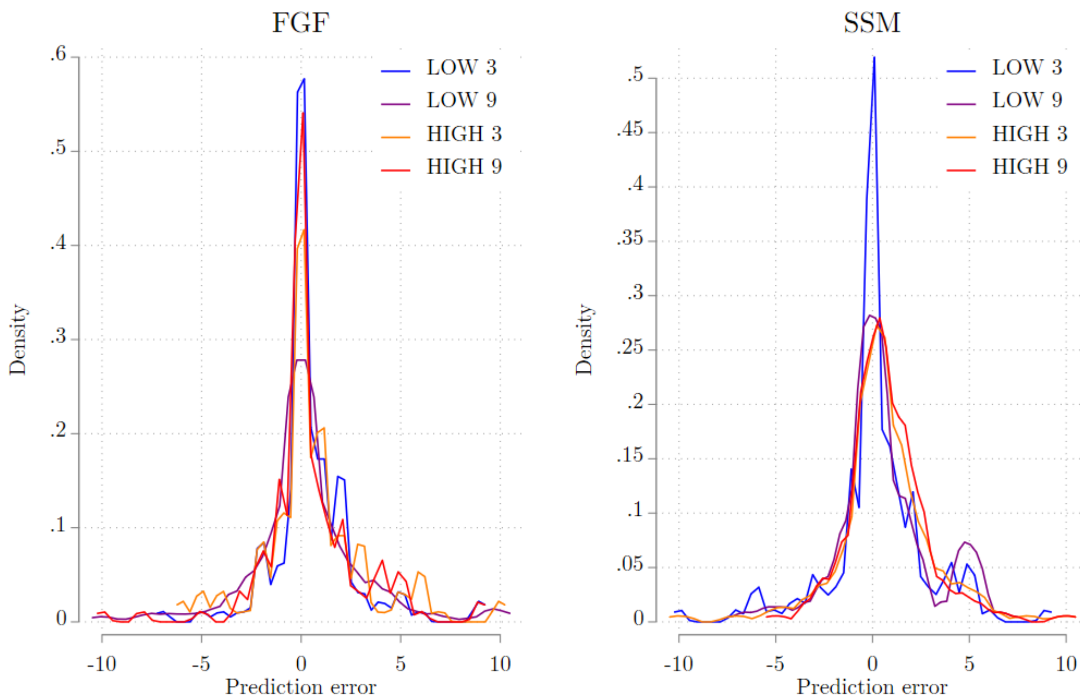


Figure B.1 Distribution of ABC prediction errors ($\hat{c}_i - c_i$) in Study 1 using the non parmetric estimation

	(1)	(2)	(3)	(4)	(5)	(6)
	All	All	Large (9)	Large (9)	Small	Small
	FGF	SSM	FGF	SSM	FGF	SSM
predicted_abc	0.502***	0.486***	0.454***	0.574***	0.536***	0.390***
	(0.0881)	(0.0696)	(0.141)	(0.0859)	(0.114)	(0.105)
beliefs	0.451***	0.513***	0.520***	0.534***	0.406***	0.505***
	(0.115)	(0.0877)	(0.198)	(0.0973)	(0.142)	(0.133)
high MPCR	0.232	0.442**	0.251	0.381	0.215	0.492
	(0.227)	(0.211)	(0.334)	(0.285)	(0.306)	(0.310)
large	-0.152	0.259				
	(0.222)	(0.206)				
_cons	0.486*	0.0598	0.219	-0.109	0.551*	0.465
	(0.276)	(0.251)	(0.453)	(0.255)	(0.303)	(0.362)
R^2	0.485	0.530	0.444	0.606	0.529	0.459
$CVMSE$	**5.989**	**5.022**	**6.764**	**4.421**	**5.458**	**5.559**
N	468	468	234	234	234	234

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
OLS estimates. The dependent variable are the contributions in the one shot public good game. The beliefs are the ones reported in the one shot public good game and predicted is the prediction using the either FGF or SSM (non paramtric version). Large is an indicator that takes the value 1 for groups of 9 and 0 for groups of 3. High MPCR is an indicator that takes the value 1 for MPCR=0.8 and 0 for MPCR=0.4. Robust standard errors in parentheses.

Table B.1 Regressions explaining contributions in Study 1 using the ABC method

C: Slope and average contributions in the schedules

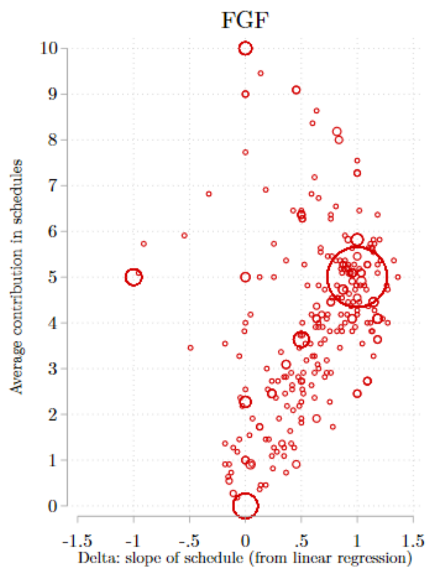


Figure C.1 Slope and average contributions in the schedules by method (Study 1)

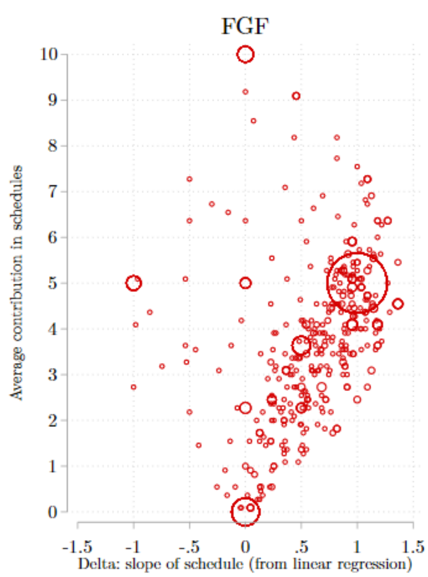


Figure C.2 Slope and average contributions in the schedules by method (Study 2)

D: Replication analysis

Note that some conditions in Study 2 are a replication of Study 1. In particular, the four conditions with groups of three (with the factorial design combining FGF/SSM and MPCR of 0.4/0.8).

Study 1 took place in January and February of 2021. Study 2 took place in two parts: the first one (all SSM conditions) in June 2021 and the second (all FGF conditions) in February 2023.

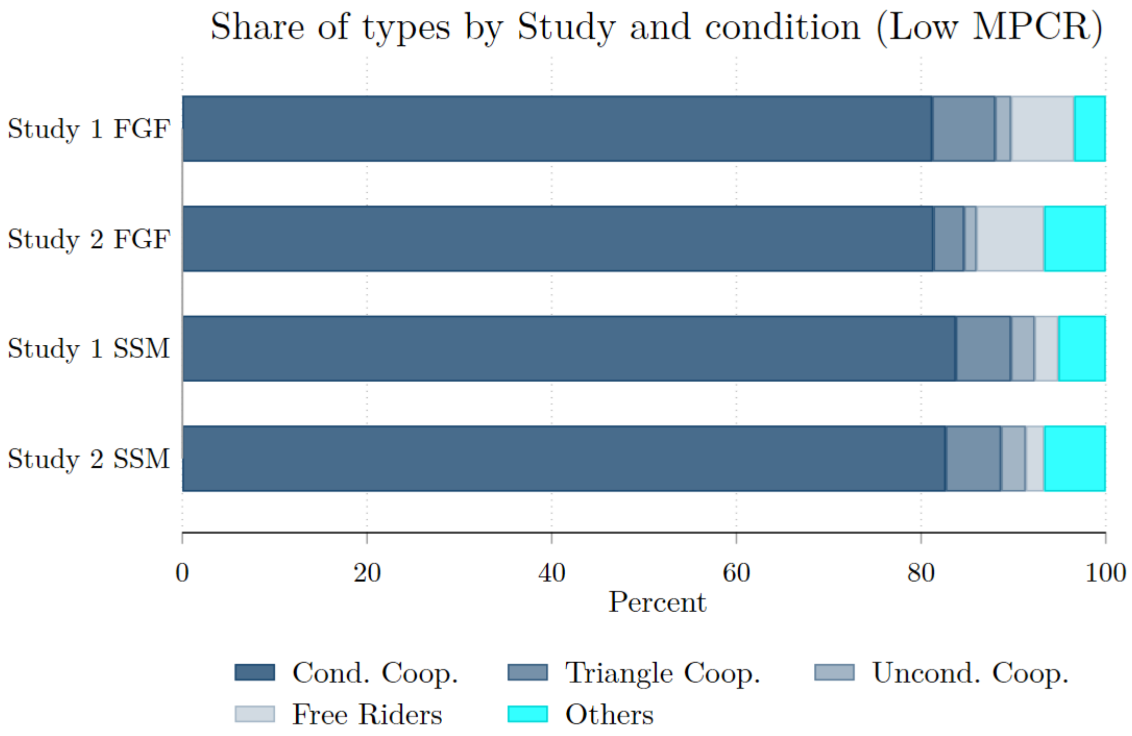
Demographics

	Study 1	Study 2 SSM	Study 2 FGF
Average age	21.8	21.6	21.6
% female	67.4	61.9	53.7
% of students	60.0	71.4	58.2

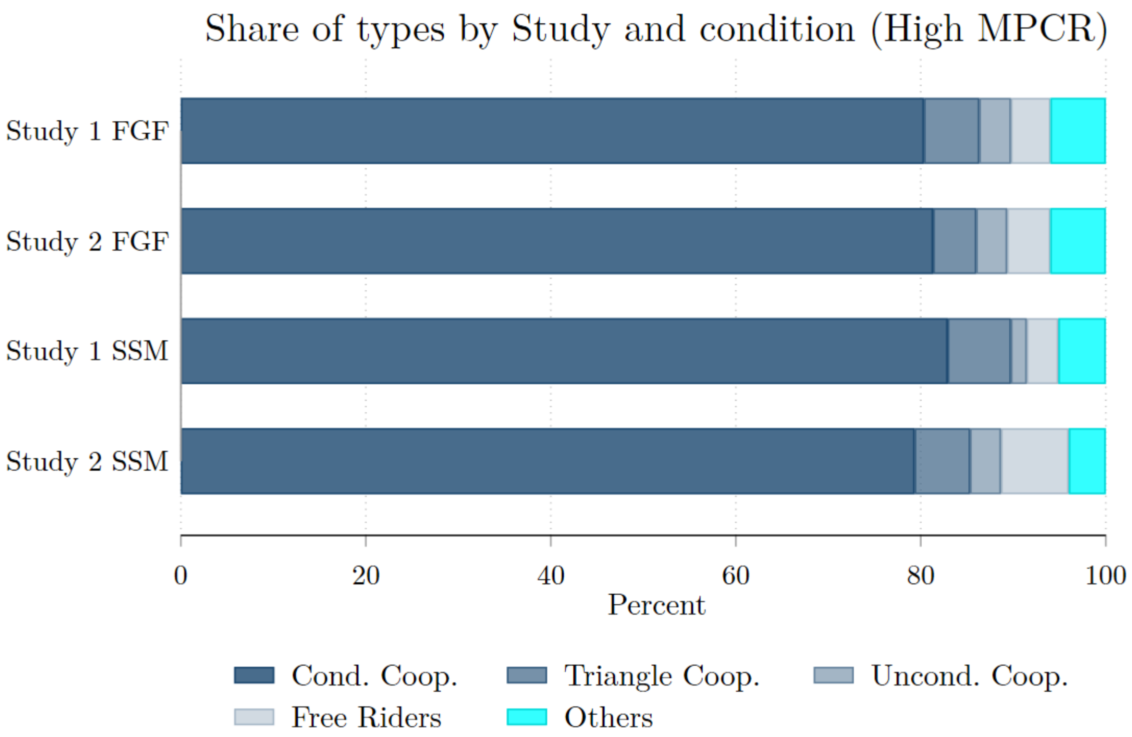
The following table compares t-tests of the main variables for each method in Study 1 vs Study 2. Columns 1 and 3 report the average of each variable in Study 1. Columns 2 and 4 report the difference between Study 2 and Study 1, the standard errors (in parenthesis) and the p-values (in brackets).

	1	2	3	4
	Study 1 SSM	Study 2 SSM	Study 1 FGF	Study 2 FGF
Average conditional contribution	4.110	-0.132 (0.153) [0.388]	4.195	0.208 (0.167) [0.215]
Unconditional contribution	4.590	-104 (0.268) [0.699]	4.936	0.669** (0.280) [0.017]
Beliefs	4.564	0.040 (0.200) [0.841]	4.235	0.058 (0.208) [0.779]
Contribution	4.440	0.077 (0.277) [0.780]	4.474	0.451 (0.284) [0.113]

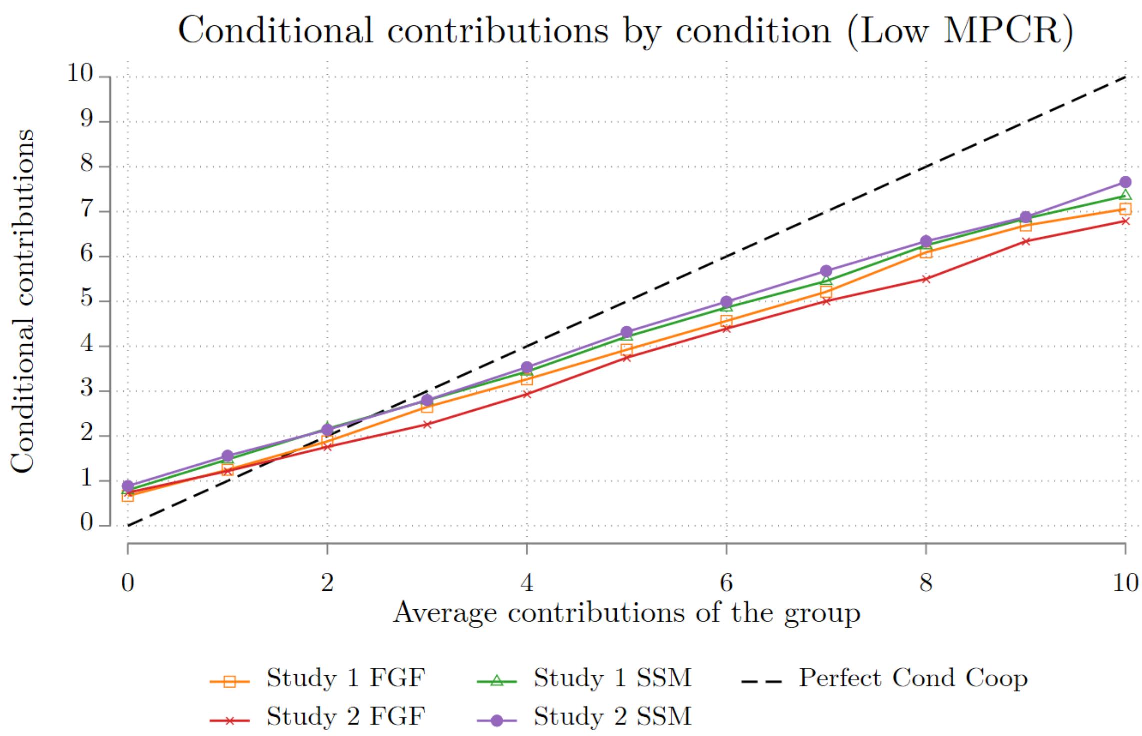
Next, we present the share of types and the slopes of the conditional contributions for each condition and study. In general, the results are replicated.

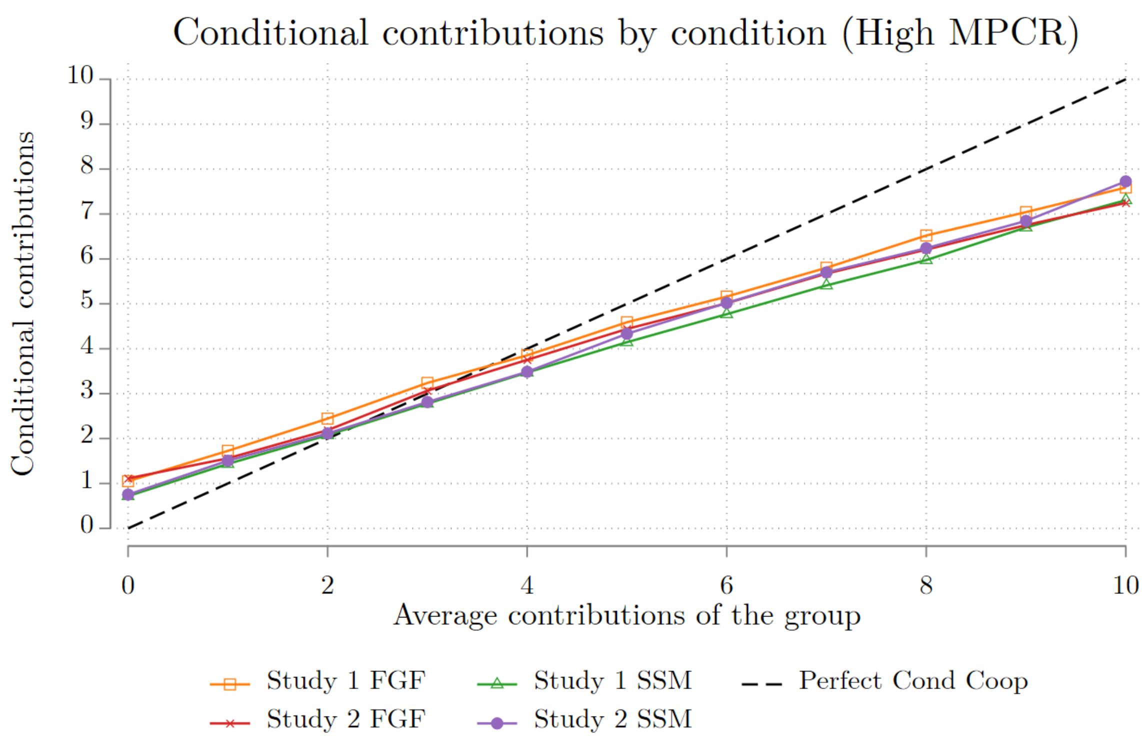


Share of types by Study and condition (Low MPCR)
Study 1 FGF
Study 2 FGF
Study 1 SSM
Study 2 SSM
0
20
40
60
80
100
Percent
Cond. Coop.
Triangle Coop.
Uncond. Coop.
Free Riders
Others



Share of types by Study and condition (High MPCR)
Study 1 FGF
Study 2 FGF
Study 1 SSM
Study 2 SSM
0
20
40
60
80
100
Percent
Cond. Coop.
Triangle Coop.
Uncond. Coop.
Free Riders
Others


Conditional contributions by condition (Low MPCR)
Conditional contributions
Average contributions of the group
Study 1 FGF
Study 1 SSM
Perfect Cond Coop
Study 2 FGF
Study 2 SSM


Conditional contributions by condition (High MPCR)
Conditional contributions
Average contributions of the group
Study 1 FGF
Study 1 SSM
Perfect Cond Coop
Study 2 FGF
Study 2 SSM

E: Online Surveys

The condition the example refers to Study 1, High MPCR, groups of 9, using SSM. The only modifications in the instructions for other conditions are the parameters for MPCR (0.04, 0.08, 0.4 and 0.8), group size (3, 9 and 30), and the strategy method used (SSM or FGF)(see the instructions below) and the figures that appear in the instructions (see below).

Instructions for each strategy method:

SSM:

<u>How your bonus will be determined</u>

We will randomly select either your **unconditional contribution** or your **contribution table** (with equal probability).

- If your **contribution table** is selected, we will use the other participants' unconditional contributions to select the number from your table.
- If your **unconditional contribution** is selected, we will use the other participants' contribution tables and select the entry by averaging everyone else's unconditional contributions.

This means that you should take both the **unconditional contribution** and the **contribution table** seriously because you don't know yet which one will be relevant for calculating your bonus.

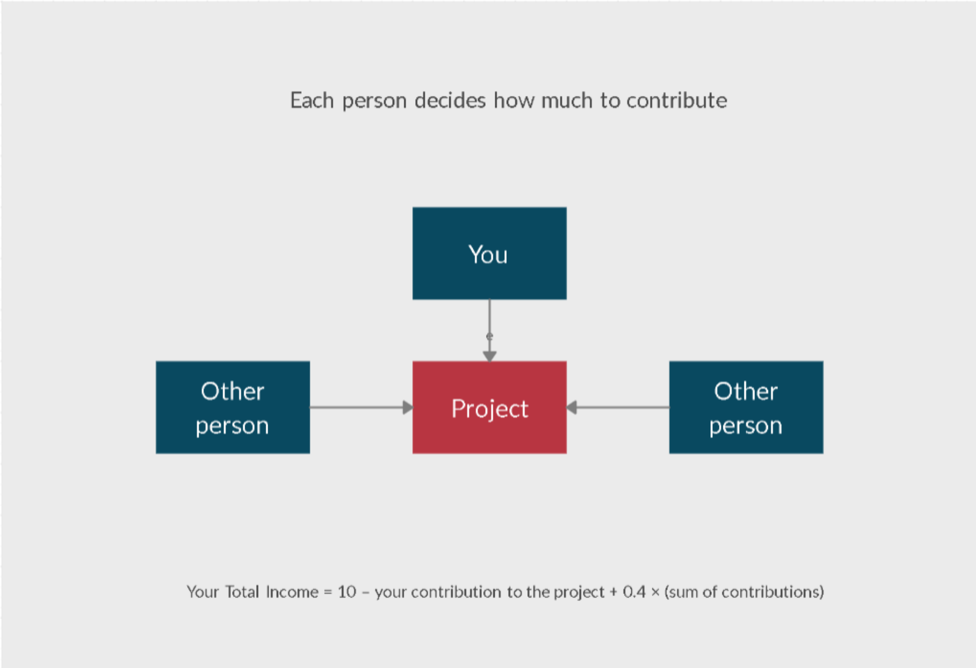
FGF:

<u>How your bonus will be determined</u>

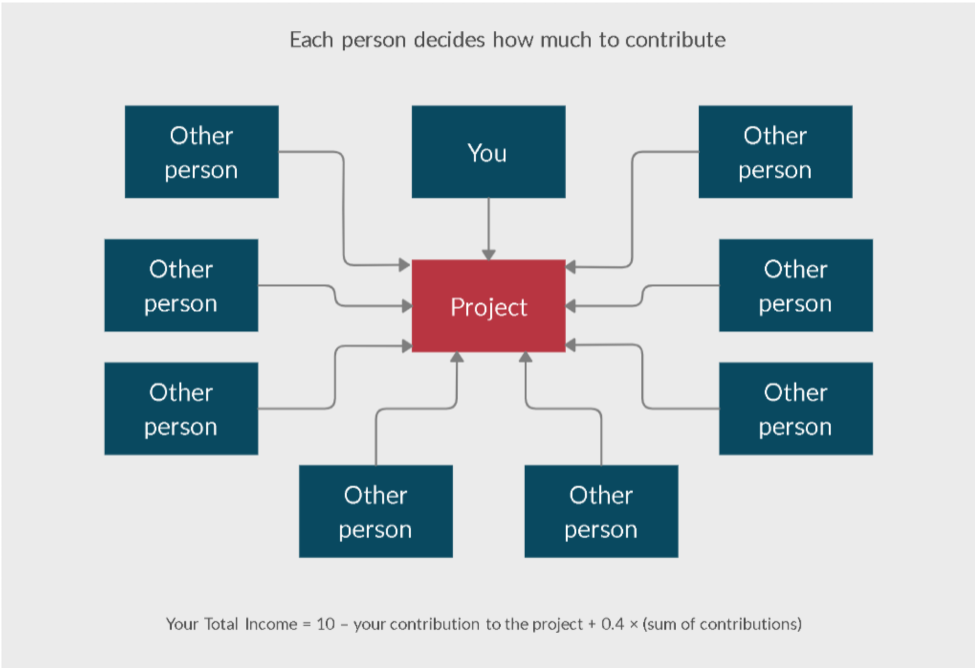
We will randomly select 1 group member for whom the **contribution table** will be payoff-relevant. For the remaining 8 group members, we will use the **unconditional contributions**.
This means that you should take both the **unconditional contribution** and the **contribution table** seriously because you don't know yet which one will be relevant for calculating your bonus.

Figures:

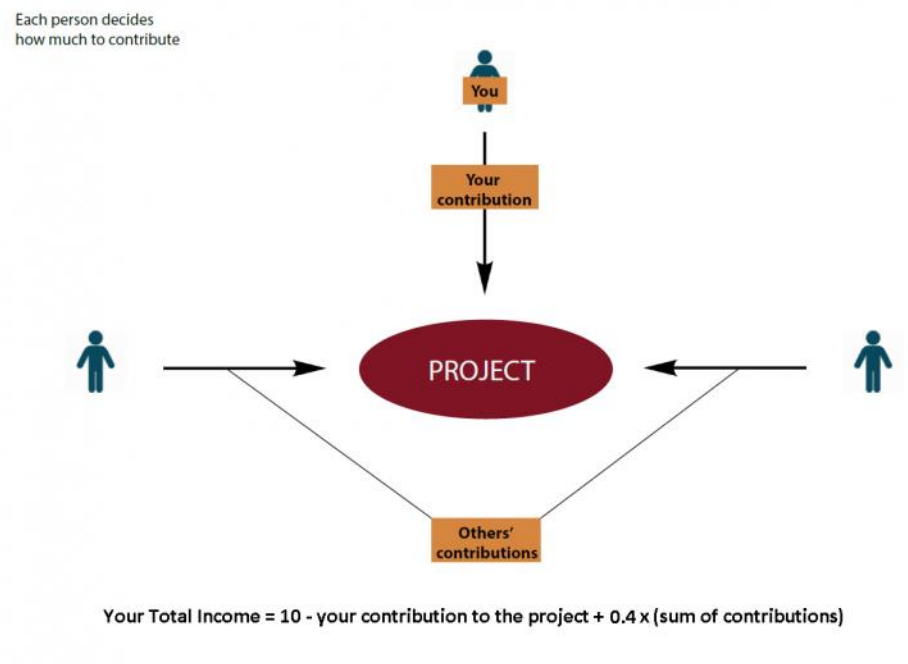
Low MPCR (0.4) – Groups of 3



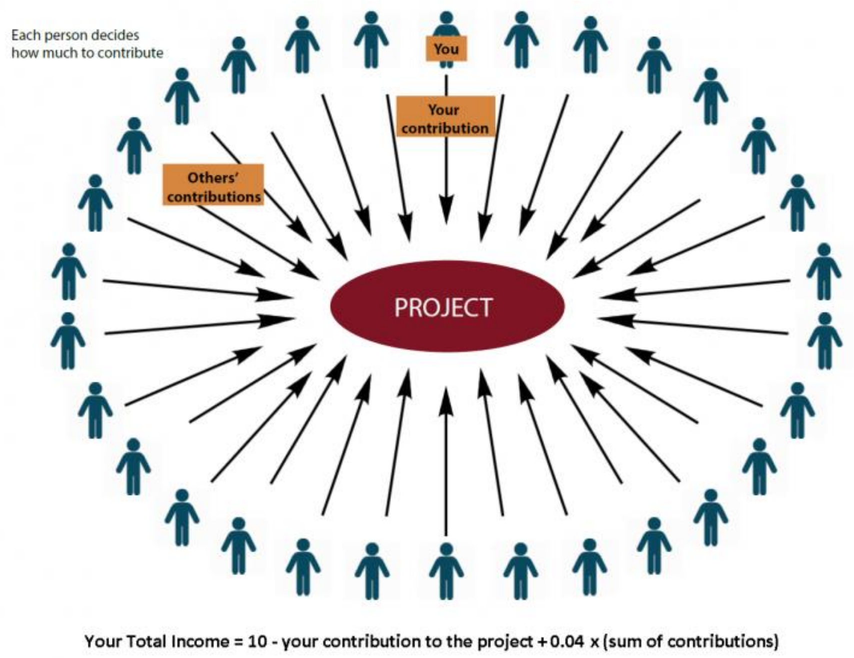
Low MPCR (0.4) – Groups of 9



Low MPCR (0.4) – Groups of 3



Low MPCR (0.04) – Groups of 30



Next, we present the full example of Study 1, High MPCR, groups of 9, using SSM.

Consent

Consent to participate in study

Duration: The study is expected to last about 7 minutes.
Risks: There are no physical or emotional risks involved.
Confidentiality: Your data will be recorded and stored as confidential and protected as such.
Your rights: You have the right to withdraw your consent or discontinue participation at any time. Your privacy will be maintained in all published and written data resulting from this study.

The study was approved by the Research Ethics Committe of the School of Economics at The University of Nottingham.

For more information you can email us at: diego.marinofages@nottingham.ac.uk

Please enter your Prolific ID here:

Intro

Thank you for participating in our Study

You are participating in an experiment in which you will earn some money. The amount will depend on the decisions taken by you and other participants.

Click >> to continue.

In this decision problem you will be randomly assigned to a group of other participants from Prolific. To determine your bo payment, we will first record your earnings in points and then exchange the points to Pounds.

Your **bonus** in Pounds will be determined as follows: **Earnings in Pounds = Earnings in Points x 0.02**

Decision situation: HIGH 9

You have been randomly assigned to interact with **8 other participants**. All of you receive this same set of instructions.

Each person in your group is given 10 tokens. You must each decide how many of these 10 tokens to keep for yourself a many to contribute to a group project.

A) You will earn 1 point from each token you keep for yourself. For example, if you put all 10 tokens into your private account, your income from your private account would be 10 points. If you put 6 tokens into your private account, your in from this account would be 6 points. No one except you earns anything from tokens you put in your private account.

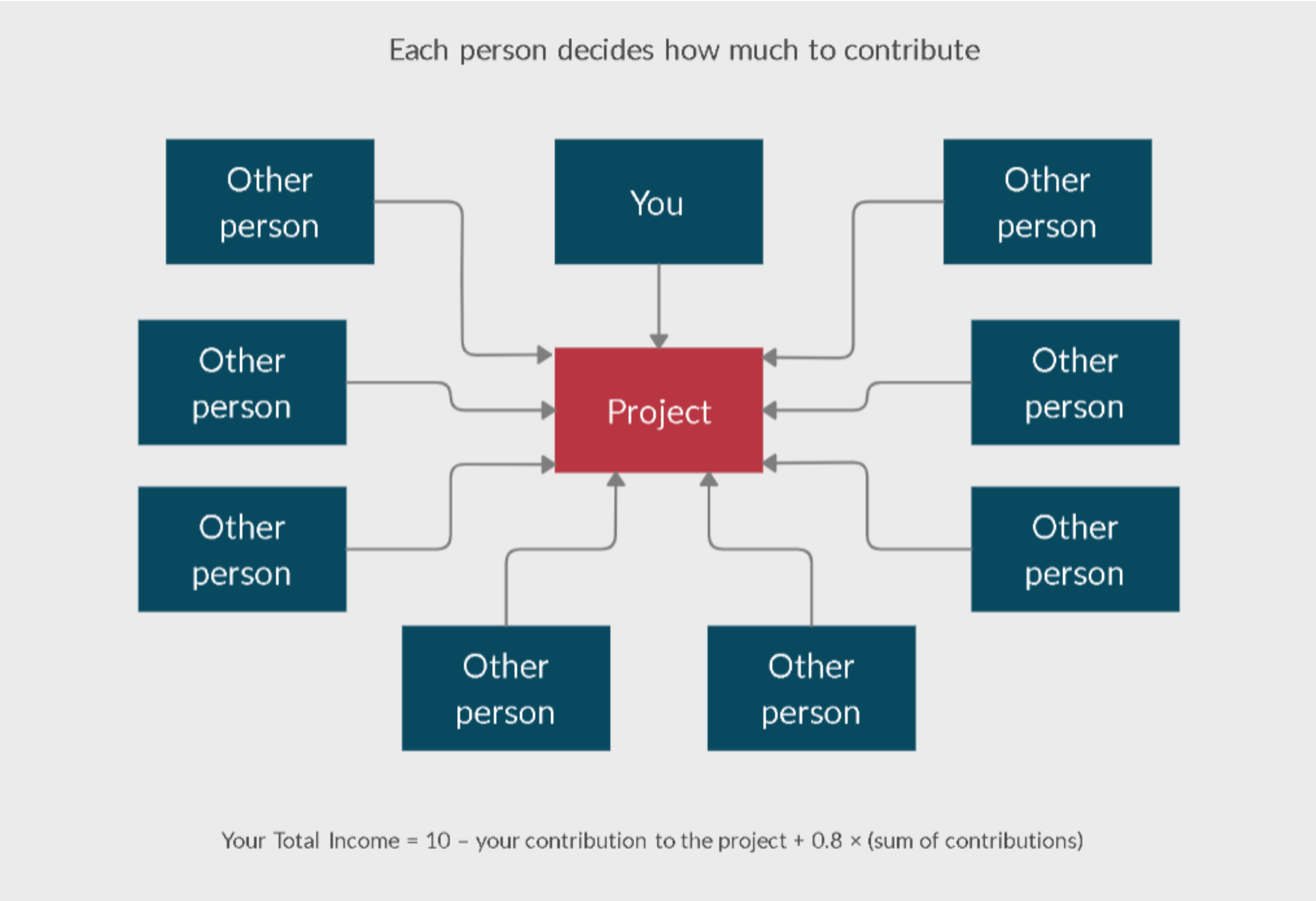
B) You will earn 0.8 points from each token that was contributed to the project by any participant including yous example, if the total contributions to the project is 10, you will receive 8 points. But in this case, each of the other particip will also get 8 points.

In summary,

Your Total Income = 10 – your contribution to the project + 0.8 × (sum of contributions)

Allocation: HIGH 9

The following diagram represents your group.



Control Questions1: HIGH 9

Please answer the following questions to check your understanding of the situation.
For each correct answer you will earn 2 points. To proceed you will need to answer all of them correctly but you ONLY earn money for the correct answers in the first attempt.

You can click the following button to review the instructions (opens a new tab).
Instructions

How many tokens would you contribute to the group project if you wanted to earn the most money for yourself personally

How many tokens would you contribute to the group project if you wanted to earn the most money for the group as a whc

Control Questions2: HIGH 9

YOUR ANSWER WAS INCORRECT, PLEASE TRY AGAIN.
You can click the following button to review the instructions (opens a new tab).
Instructions

How many tokens would you contribute to the group project if you wanted to earn the most money for yourself personally

How many tokens would you contribute to the group project if you wanted to earn the most money for the group as a whc

Instructions1: Strategy method II 9

All group members have two tasks:

- The **unconditional contribution** task: you simply decide how many tokens (up to 10) you want to contribute to the project.
- The **contribution table** task: you need to indicate how many tokens you want to contribute to the project for each p average contribution of the other group members (rounded to the next integer).

This is a one-off decision problem that is finished once you have made both decisions.

How your bonus will be determined

We will randomly select either your **unconditional contribution** or your **contribution table** (with equal probability).

- If your **contribution table** is selected, we will use the other participants' unconditional contributions to select the nur from your table.
- If your **unconditional contribution** is selected, we will use the other participants' contribution tables and select the by averaging everyone else's unconditional contributions.

This means that you should take both the **unconditional contribution** and the **contribution table** seriously because yo know yet which one will be relevant for calculating your bonus.

Please answer the following question. If you answer it correctly in your first attempt you will earn 2 additional points.
What is the probability that your contribution is determined from your contribution table?

- 1/3
- 1/9
- 1/2
- 1/5

Instructions2: Strategy method II 9

All group members have two tasks:

- The **unconditional contribution** task: you simply decide how many tokens (up to 10) you want to contribute to the project.
- The **contribution table** task: you need to indicate how many tokens you want to contribute to the project for each p average contribution of the other group members (rounded to the next integer).

This is a one-off decision problem that is finished once you have made both decisions.

How your bonus will be determined

We will randomly select either your **unconditional contribution** or your **contribution table** (with equal probability).

- If your **contribution table** is selected, we will use the other participants' unconditional contributions to select the nur from your table.
- If your **unconditional contribution** is selected, we will use the other participants' contribution tables and select the by averaging everyone else's unconditional contributions.

This means that you should take both the **unconditional contribution** and the **contribution table** seriously because yo know yet which one will be relevant for calculating your bonus.

YOUR ANSWER WAS INCORRECT. PLEASE TRY AGAIN.
What is the probability that your contribution is determined from your contribution table?

- 1/3
- 1/5
- 1/9
- 1/2

Strategy method play 9

We now ask you to make the **unconditional contribution** decision, followed by filling in the **contribution table**.

The unconditional contribution

How many tokens out of 10 do you contribute to the project?

The contribution table

Now we ask you to think about your contribution depending on how much the **other 8 group members** have contributed on average. Please indicate for *each* possible average contribution of others (rounded to an integer) how much you will contribute.

	I contribute
If each of the other 8 members contribute 0	
If each of the other 8 members contribute 1	
If each of the other 8 members contribute 2	
If each of the other 8 members contribute 3	
If each of the other 8 members contribute 4	
If each of the other 8 members contribute 5	
If each of the other 8 members contribute 6	
If each of the other 8 members contribute 7	
If each of the other 8 members contribute 8	
If each of the other 8 members contribute 9	
If each of the other 8 members contribute 10	

Instructions: One shot 9

Please now consider this NEW decision task

You are now taking part in a NEW one-off decision problem. The bonus you earn in this decision problem will be added to what you earned in the one y finished.

As in the previous decision problem you are in a **new group of 9** (i.e. you and **8 other participants** from Prolific).
Unlike in the previous decision problem, you will only make an **unconditional contribution**, that is, you will not have to fill out a contribution table.

How many tokens out of 10 do you contribute to the project?

Beliefs 9

Now we would like you to estimate a few values.

In this case you will be paid 2 points if your estimate is correct.

What is your estimate of the **average contributions** to the project of the **other 8 group members** (rounded to an intege

Beliefs 9 (second order)

Now we would like you to estimate what the other 8 group members believe the average contribution of the group will be (rounded to an integer). That i was the average response of the other members to the previous question?

Previous question: "What is your estimate of the **average contributions** to the project of the **other 8 group members** (roun an integer)?"

In this case you will be paid 2 points if your estimate is correct.

What is your estimate of what the other members BELIEVE the **average contributions** will be? (rounded to an integer)

Questionnaire

Thank you!

You're almost done, just fill out this brief survey before we finish.

If you were given the choice, would you prefer to be in a group of 3 or 9 people?

- 9
- I am indifferent
- 3

Why do you prefer that size?

Do you think a particular person would contribute MORE in the large or in the small group?

- Contribute more in the LARGE group
- Same
- Don't know
- Contribute more in the SMALL group

Assume people contribute the same amount no matter what group they are in: in which group size do you think people w make MORE money?

- LARGE group will make more money
- Don't know
- SMALL group will make more money
- Same

What is the probability (in percentages between 0 and 100 with no decimals) of drawing an odd number when rolling a fa