How Unconditional is Conditional Cooperation?

Abstract

Previous experimental research suggests that a non-negligible number of subjects do not resort to free-riding in the linear public goods game but rather are willing to contribute as long as other group members do so. We improve upon the classic design by Fischbacher et al. (2001) to enable the subjects to condition their behavior on the minimum, median, average or maximum contribution as well as on the full contribution profile of other group members. We find that conditional cooperation is a stable phenomenon that is not affected by the choice of a conditioning statistic used to classify the subjects as exhibiting such behavior. The presence of the self-serving bias, however, depends on the choice of the statistic. The subjects tend to reciprocate the average rather than anything else albeit without actually realizing that. Finally, we provide guidelines on how the total contribution level can be increased by varying the amount and type of information available to the subjects.

Keywords: public goods, conditional cooperation, self-serving bias, full information, summary statistic, strategy method, experiment

1 Introduction

Previous experimental research suggests that a non-negligible number of subjects do not exhibit free-riding behavior in the voluntary contribution public goods game, which is at odds with what rational choice theory predicts (Ledyard, 1995). According to Sugden (1984), it may be due to other-regarding preferences, or concerns for the well-being of other members of the group. The cooperative approach (Sonnemans et al, 1999) considers the possibility that an individual might contribute because he expects that some of other agents will contribute, suggesting that contributions may be driven by reciprocity to expected contributions by others.

Fischbacher et al. (2001) coin the term "conditional cooperation" to refer to the act of contributing more in response to a higher average contribution by others. In the process of documenting such tendency, they also develop a measurement tool based on the strategy method (Selten, 1967), in which subjects can condition their contribution on various average contribution levels of others. Much of subsequent experimental literature on public goods games utilizes this tool to measure conditional cooperation. Recent studies suggest that approximately half of the subjects display "conditionally cooperative" behavior when measured this way.

Follow-up experimental work investigates the influence of other types of information (apart from the average contribution of others) in the linear public good game. For instance, Bigoni and Suetens (2012), in a 10-period fixed-group setup, provide subjects with costless information about others' contributions from the previous period. In one treatment, subjects are given the average level of others' contribution in the previous period. In another treatment, they are given full information about others' contributions in the previous period. The authors conclude that differences in these types of information significantly affect the average contribution level in a given period.

A similar experimental design is used in Kurzban and Descioli (2008) where three summary statistics, such as the minimum, median and maximum, are available to the subjects. The authors conclude that conditional cooperators are mostly interested in the median contribution of others, while free-riders prefer information about the maximum contribution.

Zetland and Della Giusta (2012) provide subjects, in one treatment, with the average contribution of other group members in the previous period, and, in another treatment, with the total contribution in the previous period. Even though these two pieces of information are isomorphic, the shares of free-riders and conditional cooperators are significantly different between the treatments.

As a behavioral phenomenon, conditional cooperation is usually accounted for by reciprocity that subjects may exhibit toward other contributors, and the existing studies suggest that it may matter what kind of information is given to subjects as a guideline for such behavior. Indeed, the average contribution of others hides what the true distribution of the three contributions is.

The agenda of this paper is twofold. First, we want to have a better picture of preferences that drive contributions toward the public good. Barring subject confusion, cooperation involves a private sacrifice for social gain. One may think that willingness to substitute private for social welfare interacts not just with how much others contribute on average, but also on where exactly one would end up in the resulting distribution of contributions. Theories of both advantageous and disadvantageous inequality aversion (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000) speak exactly of that. As a result, besides the average, potential contributors may want to condition on other summary statistics, such as the maximum, median or minimum.

The second goal is on the methodological side. Existing literature uses the average as a conditioning statistic. But is it the statistic that subjects want to know? Or would they rather condition on some other statistic if they could choose? How would they behave if observing some alternative statistic? Say, a subject is reciprocal but desires to avoid disadvantageous inequality vis-a-vis all other contributors. He would like to match the lowest contribution then. However, it is not so easy to do if he only knows the average. As a result, some subjects may be exhibiting what is known in the literature as the self-serving bias (Ockenfels, 1999; Reuben and Riedl, 2013) or even appear to be free-riding when they would otherwise cooperate. By offering alternative statistics as well as the entire distribution of the contributions of others to condition upon, we want to examine the robustness of existing empirical findings regarding the incidence and intensity of conditional cooperation in public goods games.

We build upon the methodology of Fischbacher et al. (2001) to create an experimental playground, where subjects are able to condition their behavior on various pieces of information. Among other things, using a single-period design allows us to ignore dynamic feedback concerns inherently present in repeated settings[1].

The rest of the paper is structured as follows. Section 2 explains the experimental procedure. Section 3 introduces the research hypotheses and presents the results. Section 4 concludes.

2 Experimental Design

The experimental design is built around the standard linear public goods game. Part of a group of four, each individual has 10 indivisible tokens to distribute between his or her private account and what we call in the experiment, the "group project". The final individual profit is then calculated

[1]Even though one may argue that it is possible to mitigate such concerns by means of the stranger matching protocol, learning about the population, dynamic reciprocity etc. can still cause a problem.

using the following function[2]:

$$\pi_i = 10 - g_i + 0.75 \sum_{j=1}^{4} g_j, \tag{1}$$

where g_i is one's own contribution to the public good, and g_j denotes individual contributions from each group member.

Following Fischbacher et al. (2001), we started with explaining the basic structure of the public goods game to the subjects and testing their understanding of it with the help of ten control questions that all had to be answered correctly in order to proceed[3].

The actual experiment was framed as made of three stages that represented various types of contribution decisions that the subjects had to make. The stage specific instructions were distributed immediately before each respective stage[4]. There was no feedback in-between the stages, and only one stage would be randomly[5] chosen to be payoff relevant for a given subject. It was emphasized that all decisions could be relevant in the end and hence one should think carefully about all of them.

The actual stages represented various information treatments, which we further will be referring to as "unconditional contribution" (UC), "conditional contribution under full information" (CF) and "conditional contribution w.r.t. a summary statistic" (CS).

Just like its name suggests, UC is to elicit the subjects' willingness to contribute under no information about the decisions of other group members.

In CF, we use the strategy method and present 23 different scenarios of what the contributions from the other three group members can possibly be. These scenarios are shown to each subject one by one in a random order. One of them, varying from subject to subject, represents the actual contributions from stage 1 so that: (i) we know which scenario to use for payoff calculation should CF be chosen to be payoff relevant for a particular subject; and (ii) we can truthfully state in the instructions that one of the scenarios is a real one. The other 22 scenarios (see Table 5) are fixed across the subjects and

[2]A relatively high marginal per capita return coefficient was chosen in order to further incentivize cooperative behavior as we were not happy with the results of the pilot session that used a value of 0.4.

[3]The complete instruction sheets are available upon request.

[4]We had reasons to believe that presenting the subjects with, e.g., instructions to stage 3 could affect their behavior in stage 2, and hence opted for a piece-wise delivery of those.

[5]Within each group, 3 subjects would have their final payoff determined by their contribution decision in the first stage, whereas one subject would have his or her final payoff determined via stage 2 or stage 3 with equal probability.

come from the set of 286[6]. Forming this sample and/or its size is a very non-trivial task on its own[7]. Not denying openness for interpretation, we opted for a design that is both minimalistic and capable of keeping the values of the four statistics of interest (i.e., minimum, median, maximum, average) as far away from each other as possible. Some of the more specific considerations important to us were:

- spanning the whole value range for each statistic (except for scenarios with the minimum equal to 10 or maximum equal to 0);
- reducing the incidence of focal points (i.e., scenarios with two or three identical contributions);
- ensuring the least correlation between the medium and average contributions.

In CS, we use the strategy method again and ask the subjects to fill in four "contribution tables", each representing a particular type of statistic of what the contributions from other three group members can possibly be. Each such table contains eleven values from 0 to 10, and the order in which the tables are shown varies across the subjects to control for possible order effects. The subjects are informed that only one of the tables will be randomly selected to be payoff relevant should stage 2 be chosen for them. The details of the selection mechanism are disclosed only after all four tables have been processed, and that is also when the subjects have to make their final decision in the experiment. The decision is whether or not one would like to pay a minor fee to increase the probability of a particular table to be selected by the random mechanism. By default, each table is equally likely to be selected but the subject can pay 1 or 2 tokens to boost the probability of one to 0.5 or 0.75, respectively (the remainder of the probability mass is then equally distributed across the remaining three tables). Of course, the fee is only levied if stage 3 turns out to be the one that is payoff relevant for the subject.

The experiment was conducted at the Laboratory of Experimental Economics at the University of Economics in Prague using z-Tree (Fischbacher,

[6]The total number of possible scenarios is given by a standard combinatoric of unordered sampling with replacement, which counts the number of multisets of length 3 on 11 symbols.

[7]An immediate way of doing that, despite being well grounded theoretically, cannot be implemented for practical reasons. It is quite intuitive that we want to have a sample with more variation across the payoffs rather than less. One could then define a value function on a sample of a particular size and go over all possible combinations but that is not feasible due to the dimensionality of the problem.

2007). The majority of the subjects were Bachelor and Master students (44% and 39%, respectively) of various fields. There were 4 sessions with either 20 or 24 subjects each. Altogether, there were 88 subjects that formed 22 groups. There was no show up fee for those who actually participated in the experiment. The average earnings were CZK 345 (about EUR 12.6 or USD 17.3).

3 Results

3.1 Conditioning Conditional Cooperation

Starting with Fischbacher et al. (2001), there has been plenty of evidence that there exist certain behavioral archetypes as far as contribution strategies in linear public goods games are concerned. Similarly, we classify 33% of our subjects as free riders, 42% as conditional cooperators and 9% as ones with "hump-shaped" contribution patterns when using the average of the contribution profile of other players as a conditioning statistic (see Figure 6).

Another common finding in the public goods domain is the so-called "self-serving bias". Again, the aforementioned study reports that it occurs only 11.9% of the time that subjects classified as conditional cooperators contribute strictly more than the average of other players in their group when knowing said average. In our experiment, the exact same measure is equal to 16.0% when using the average as a conditioning statistic but goes up to 39.6% when using the minimum as such.

In this section, we would like to put the de-facto standard methodology of Fischbacher et al. (2001) to the test in order to see how unconditional conditional cooperation really is. Having an experimental design richer in the number of summary statistics (including the complete contribution profile of other players) that the subjects can condition their own contribution strategy upon, enables us to answer the following questions as far as such behavior is concerned.

First, we are interested in knowing whether or not conditional cooperation as such is a stable concept that can be used effectively when speaking of the subject behavioral archetypes. While the choice of the average as a conditioning statistic is a fairly intuitive one, would it make sense to speak of the various demographics that result if they were subject to change because of one?

Second, we would like to know if the self-serving bias is a mere artifact of imperfect information available to the subjects. Can it, perhaps, be mitigated by letting the subjects know the minimum contribution of other players as opposed to the average?

We formulate two research hypotheses to answer these questions.

Hypothesis one: behavioral archetypes in a linear public goods game are persistent phenomena and do not depend on the summary statistic used to elicit them.

Hypothesis two: subjects do not exhibit the self-serving bias when able to condition on the minimum contribution of other players.

In CS, the subjects are asked to fill in four contribution tables, each representing a particular type of statistic of what the contributions from the other three group members can possibly be. This enables us to compare the results of the de-facto standard classification method between any pair of such tables. The first research hypothesis is then addressed by conducting a series of tests for independence within each pair of classification outcomes that are the results of using either the minimum, median, average or maximum as a summary statistic that the subjects can condition their behavior upon. In statistical terms, it is equivalent to testing the following hypothesis for each such pair:

$$\begin{aligned} H_0: &\quad X^i \text{ and } X^j \text{ are independent,} \\ H_A: &\quad X^i \text{ and } X^j \text{ are not independent,} \end{aligned} \tag{2}$$

where X^i and X^j represent the outcomes of classifying the subjects into the free riders, conditional cooperators or "hump-shaped" cooperators using either summary statistic, $\{i, j : i \neq j\} \in \{minimum,\ medium,\ maximum,\ average\}$ as a conditioning one.

We conduct 6 such tests and conclude that the choice of a conditioning statistic does not affect the classification results significantly[8]. Table 1 summarizes the overall classification results for each conditioning statistic as well as the final classification results, where the subjects get assigned into a specific category only if their assignment is invariant to the choice of statistic. We deem the latter the definitive classification and use it throughout the rest of the paper.

We consider hypothesis one confirmed. Out of 88 subjects in our sample, 24 (27%) can be classified as genuine free riders, 29 (33%) as genuine conditional cooperators, 3 (3%) as genuine "hump-shaped" cooperators, 6 (7%) as random (or other) and 26 (30%) as inconsistent or varying w.r.t. the conditioning statistic.

[8]We use Fisher's exact test to obtain the individual p-values and then the Holm-Bonferroni method to adjust the error rate for the multiple comparisons. The actual results are suppressed due to the fact that all individual p-values are less than 0.0001 whereas rejecting the null at the significance level of 0.05 would require a minimum value of approximately 0.008.

Table 1: Subject classification results as per choice of conditioning statistic

Cond. Statistic	Behavioral Archetype				
	Free Riders	Cond. Coop.	HS[9] Coop.	Other	Inconsist.
Minimum	29 / 33%	39 / 44%	5 / 6%	15 / 17%	.
Median	27 / 31%	43 / 49%	6 / 7%	12 / 14%	.
Maximum	27 / 31%	39 / 44%	4 / 5%	18 / 20%	.
Average	29 / 33%	37 / 42%	8 / 9%	14 / 16%	.
Final	24 / 27%	29 / 33%	3 / 3%	6 / 7%	26 / 30%

Overall classification results when using either summary statistic as base, as well as the final classification, where the subjects get assigned into a specific category only if their assignment is invariant to the choice of base.

[9]The so-called "hump-shaped" cooperators (Fischbacher et al., 2001).

Now, the self-serving bias. Figure 1 below replicates the main graph from Fischbacher et al. (2001) that shows the average contributions of different behavioral archetypes, while using the minimum as a conditioning statistic. The conditional contributors as a group seem to be very close to the 45-degree line (compare to Figure 6 in Appendix).

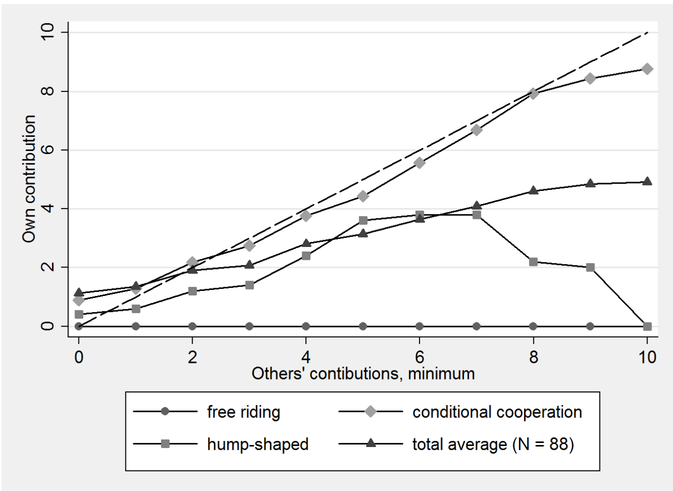


Figure 1: Own contribution as function of minimum contribution among other players

Replication of the famous graph from Fischbacher et al. (2001). Average contributions of different behavioral archetypes while using the minimum as a conditioning statistic.

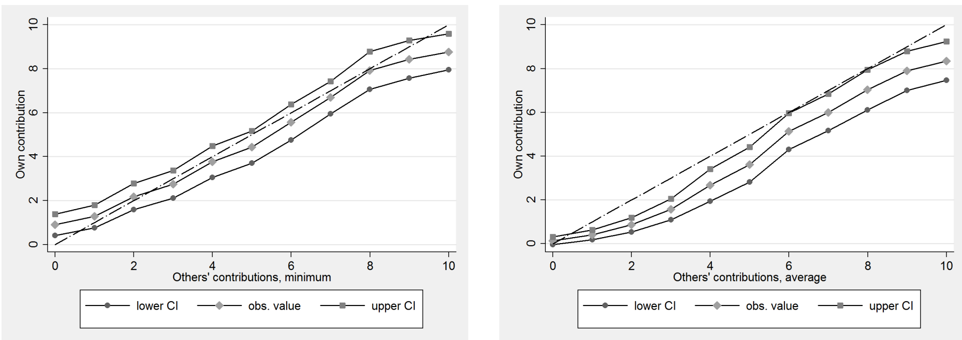
Our second research hypothesis is exactly about that. The statistical null and alternative are defined as follows[10]:

$$
\begin{array}{ll}
H_0: & M(X^{MIN}) \text{ lies on the diagonal,} \\
H_A: & M(X^{MIN}) \text{ lies below or above the diagonal,}
\end{array}
\tag{3}
$$

where $M(\cdot)$ denotes the mathematical mean and X^{MIN} represents individual contributions of conditional cooperators as a function of the minimum contribution of other players.

To test for significance, we will be relying on the notion of confidence intervals (0.05 significance level, bias-corrected), constructing them by resampling observations (with replacement) from the collected sample 100 thousand times (i.e., bootstrapping). Figure 2 below plots the test results, both for the minimum and average (as a reference) as conditioning statistics. In the former case, the contribution schedule indeed lies on the diagonal; whereas in the latter case, it is clearly below it[11].

We take this as evidence in favor of hypothesis two and conclude that the self-serving bias usually observed in public goods games is a mere artifact of using the average as a conditioning statistic[12].



(a) Minimum as cond. statistic

(b) Average as cond. statistic

Figure 2: Self-serving bias as function of conditioning statistic

Average contribution schedule of conditional cooperators as a function of either miimum (a) or average (b) contribution of other players, with 95% confidence intervals.

[10]The two-sided alternative here is to allow for the opposite of the self-serving bias.

[11]Obviously, the end points are meaningless here.

[12]Figure 7 plots the relative incidence of subjects meeting, falling short of or exceeding the minimum contribution level of other group members.

3.2 Preferences over Summary Statistics

Existing literature suggests that subjects have heterogeneous preferences over information about the contribution profiles of others and that these preferences correlate with their behavioral archetypes (Kurzban and Descioli, 2008; and Croson, 2006). We believe that there could be two sides to the story. On the one hand, conditional cooperators could be more likely to have such preferences to begin with. On the other hand, they could have a different distribution over the summary statistics as a group[13]. This allows us to formulate the following two hypotheses.

Hypothesis three: conditional cooperators are more willing to pay to increase the odds of some conditioning statistic to be payoff relevant than the rest of the subjects.

Hypothesis four: the actual distributions of the preferred conditioning statistic differ between conditional cooperators and the rest of the subjects.

In both cases, the statistical null and alternative are defined as in (2) above whereas X^i and X^j represent the behavioral data from either i, conditional cooperators, or j, the rest of the subjects. The actual experimental outcomes are provided in Tables 2-3.

Table 2: Tokens paid for increasing odds of some statistic to be payoff relevant

	Tokens Paid			
Behavioral Archetype	0	1	2	Total
Conditional cooperators	12 / 41%	13 / 45%	4 / 14%	29
Everyone else	39 / 66%	12 / 20%	8 / 14%	59
Free riders	17 / 71%	4 / 17%	3 / 13%	24
"Hump-shaped" cooperators	3 / 100%	.	.	3
Other	5 / 83%	.	1 / 17%	6
Inconsistent	14 / 54%	8 / 31%	4 / 15%	26
Total	51 / 58%	25 / 28%	12 / 14%	88

The p-values of Fisher's exact statistic for the two tests are 0.042 and 0.855, which allows us to reject the null at the significance level of 0.05 for hypothesis three but not for hypothesis four. Between our two demographics of interest, we therefore find no evidence for systematic differences in preferences over the conditioning statistics[14] provided such preferences are not degenerate in the first place. The latter condition, however, is significantly

[13]The distinction between the two is analogous to the popular distinction between the external and internal margins when speaking of labor force participation in labor economics.

[14]Even though the preference distributions seem to be identical, we believe that different demographics could be using very different motives to get there.

Table 3: Distribution of non-degenerate preferences over summary statistics

Behavioral Archetype	Preferred Statistic Minimum	Median	Average	Total
Conditional cooperators	10 / 59%	1 / 6%	6 / 35%	17
Everyone else	14 / 70%	1 / 5%	5 / 25%	20
Free riders	5 / 71%	.	2 / 29%	7
“Hump-shaped” cooperators	.	.	.	.
Other	1 / 100%	.	.	1
Inconsistent	8 / 67%	1 / 8%	3 / 25%	12
Total	24 / 65%	2 / 5%	11 / 30%	37

more likely to be satisfied for conditional cooperators than for the rest of the subjects. The minimum and average appear to be the all-around first and second choices, while nobody seems to be interested in knowing the maximum of the contribution profile of other group members. It is also noteworthy that 7 out 29 free riders in our experiment seem to have a specific preference over the conditioning statistic, which is at odds with what is usually found in the literature.

Now that we know which statistics different subject demographics tend to like the most according to their own reports, one may wonder which statistics tend to actually drive the behavior. In the terminology of Croson (2006), we want to know what kind of reciprocity the subjects could be motivated by while contributing to the public good. Or in technical terms, we want to see which summary statistic over the contribution profile of the other three members in the group — i.e., the minimum, median, maximum or average — is a better predictor of one's own contribution. Unlike other papers in the literature that consider at most three statistics at a time[15], we distinguish between all four of them by using the mean square deviation (MSD) as a measure of accuracy.

Effectively, we take the individual data from treatments CF and CS and compute four MSD scores, one for each summary statistic, for each subject using the following formula:

$$MSD^i = \frac{1}{23}\sum_{j=1}^{23}(X_j^{FULL} - X_j^i)^2, \tag{4}$$

where X^{FULL} represents the individual contribution data from treatment CF; $\{X^i : i \in \langle minimum,\ medium,\ maximum,\ average\rangle\}$ represents the

[15]The usual approach includes running a linear regression of one's own contribution on all statistics considered (and other covariates perhaps), which limits the set to three due to perfect collinearity in case the group size is 4.

individual contribution data from when using either summary statistic for conditioning; and $j \in \{1, \ldots, 23\}$ denotes a particular scenario from treatment CF.

We then consider the statistic with the lowest MSD score as the better predictor for said subject. If there happens to be a draw between two or more such scores, we consider none of the statistics as the better one. The results of this exercise are presented in Table 4.

Table 4: Summary statistics implied by MSD score ranking

Behavioral	Implied Statistic					
Archetype	Minimum	Median	Maximum	Average	Neither	Total
Cond. cooperators	4 / 14%	8 / 28%	2 / 7%	15 / 52%	.	29
Everyone else	8 / 14%	6 / 10%	8 / 14%	16 / 27%	21 / 36%	59
Free riders	.	.	.	7 / 29%	17 / 71%	24
HS cooperators	1 / 33%	.	.	2 / 67%	.	3
Other	1 / 17%	1 / 17%	.	2 / 33%	2 / 33%	6
Inconsistent	6 / 23%	5 / 19%	8 / 31%	5 / 19%	2 / 8%	26
Total	12 / 14%	14 / 16%	10 / 11%	31 / 35%	21 / 24%	88

Our findings suggest that there is evidence for the subjects reciprocating the average rather than any other summary statistic. The results of this classification, both for conditional cooperators and the rest of the subjects, were tested against the null of random uniform allocation and found to be statistically significant at the 0.05 level[16].

As a benchmark, we repeat the exercise but without the average among the competing statistics, which results in the median found to be the better predictor of one's own contribution[17], just like, e.g., in Croson (2006).

Finally, we are interested in whether the subjects are actually aware of which summary statistic they tend to reciprocate. To answer this question, we contrast the results of the subject classification according to the MSD criterion above with the stated preference for the conditioning statistic. Hence hypothesis five.

Hypothesis five: preferences over the summary statistics need not be evident to the subjects themselves[18].

We test the hypothesis separately for conditional cooperators and for the rest of the subjects. In both cases, the statistical null and alternative

[16] The corresponding p-values of the Chi-Squared test statistic were 0.0035 and 0.0074, respectively.

[17] As one might expect, the majority of the subjects found to be reciprocating the average would be considered reciprocating the median if the former were not considered.

[18] Obviously, the analysis only applies to those subjects that actually state such a preference (see Tables 2-3).

are again defined as in (2) above while X^i and X^j represent the preferred summary statistic as reported by i, the subject, and j, the outcome of the MSD score ranking, respectively.

We find no statistical evidence in favor of rejecting the null since the p-values from Fisher's exact tests are equal to 0.268 and 0.777 for the conditional cooperators and for the rest of the subjects, respectively. This implies that the subjects do not realize what actually drives their contribution behavior as far as the four summary statistics over the contribution profile of other group members are concerned.

3.3 Impact of Information Type on Overall Contribution Level

The final logical part of the analysis revolves around using information as a means to maximize the overall level of contributions.

In CF, the subjects are asked to submit their contributions while observing the full contribution profile of the other three members in their group. In CS, the subjects are asked to submit their contributions while observing only a summary statistic over the contribution profile of the other three members in their group. Exploiting the natural links between the scenarios in the two treatments allows for a great opportunity to tap into the area of potential policy implications.

Indeed, all individual differences across the subjects aside, one may also be interested in manipulating the total contribution level. Our design enables us to answer the question of what should be revealed if one's goal is such.

The following two hypotheses are postulated to address the issue.

Hypothesis six: providing the subjects with full information about the contribution profile of others results in the highest level of contributions.

Hypothesis seven: it is possible to ensure the highest level of contributions by selecting a particular summary statistic over the contribution profile of others.

In terms of motivation, the difference between the two hypotheses boils down to whether or not it is even feasible to disclose full information about the contribution profile of others. While the former assumes that it is and then it is just a matter of whether such disclosure pays off or not, the latter assumes that disclosing full information is not feasible and one has to opt for some summary statistic instead.

Overall, we have 22 scenarios to consider and hence the analysis is going to be done on the per-scenario basis. We like to think of them as of various worlds one could be living in. As such, any generalizations over all of them would be of no particular value to a policy maker from either. For the sake of exposition, though, we will be mapping the scenarios into a two-dimensional space, with the dimensions measuring the distance from the medium contribution to either lowest or highest contribution.

To address research hypothesis six, the statistical null and alternative are formulated as follows:

$$\begin{aligned} H_0: &\quad M(X^{FULL}) - max\{M(X^i), M(X^j), M(X^k), M(X^l)\} = 0, \\ H_A: &\quad M(X^{FULL}) - max\{M(X^i), M(X^j), M(X^k), M(X^l)\} > 0, \end{aligned} \tag{5}$$

where $M(\cdot)$ denotes the mathematical mean while X^{FULL} and $\{X^i, X^j, X^k, X^l : i = minimum, j = medium, k = maximum, l = average\}$ represent the individual contribution data from treatment CF and from when using either summary statistic for conditioning.

Since the test statistic does not fall into any family with known distributional properties, we will be relying, again, on the notion of confidence intervals, or rather lower confidence bounds, since the alternative hypothesis is one-sided.

To test for significance, we bootstrap the test statistic by resampling observations (with replacement) from the sample 100 thousand times. The complete test results are available in Table 6. Figure 3 below plots the observed and estimated values of the test statistic as well as the lower confidence bound (bias-corrected) at the 0.05 significance level for each scenario.

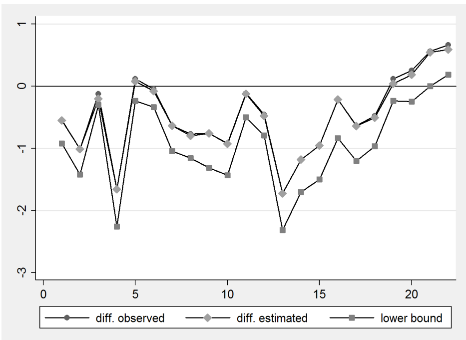


Figure 3: Full information vis-a-vis all four summary statistics

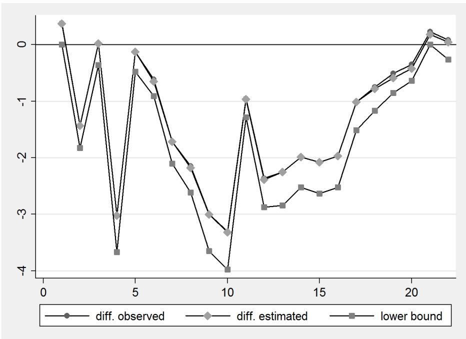
The results of bootstrapping the test statistic as defined in (5). Observed and esimated values as well as the lower 95% conidence bound for each of the 22 contribution scenarios considered (horizontal axis).

To address research hypothesis seven, we run four statistical tests, each with the following statistical null and alternative:

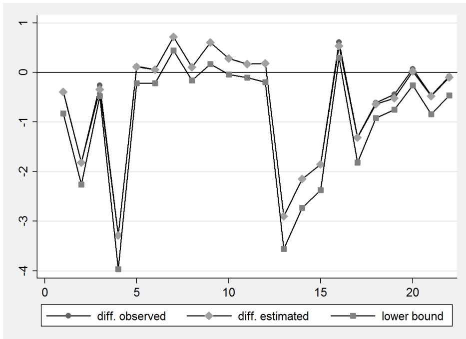
$$\begin{aligned} H_0: &\quad M(X^i) - max\{M(X^j), M(X^k), M(X^l)\} = 0, \\ H_A: &\quad M(X^i) - max\{M(X^j), M(X^k), M(X^l)\} > 0, \end{aligned} \tag{6}$$

where $M(\cdot)$ denotes the mathematical mean and $\{X^i, X^j, X^k, X^l : (i \neq j \neq k \neq l) \cap (i, j, k, l \in \langle minimum,\ medium,\ maximum,\ average \rangle)\}$ represent the individual contribution data from when using either summary statistic for conditioning.

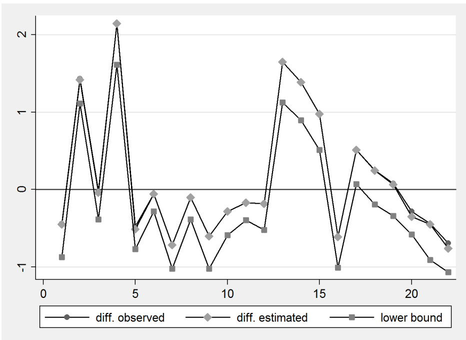
Again, the complete results of each test are available in Table 6. Figure 4 below plots the observed and estimated values of the test statistics as well as the lower confidence bound (bias-corrected) at the 0.05 significance level for each summary statistic and each scenario considered.



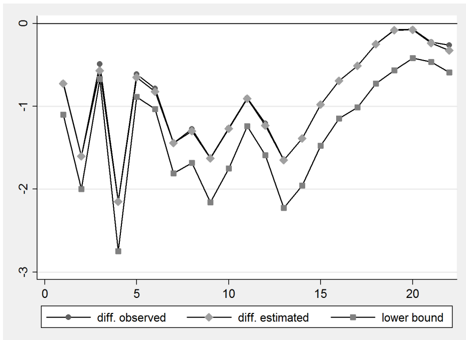
(a) Minimum vis-a-vis other three summary statistics



(b) Median vis-a-vis other three summary statistics



(c) Maximum vis-a-vis other three summary statistics



(d) Average vis-a-vis other three summary statistics

Figure 4: Imperfect information, average contribution

The results of bootstrapping the test statistic as defined in (6). Observed and estimated values as well as the lower 95% confidence bound for each of the 22 contribution scenarios considered (horizontal axis).

As one can see from Figures 3–5, the impact of various information types, be it the complete contribution profile of other group members or some summary statistic over it, is highly dependent on the actual scenario the hypothetical policy maker may be facing. Among the scenarios covered by

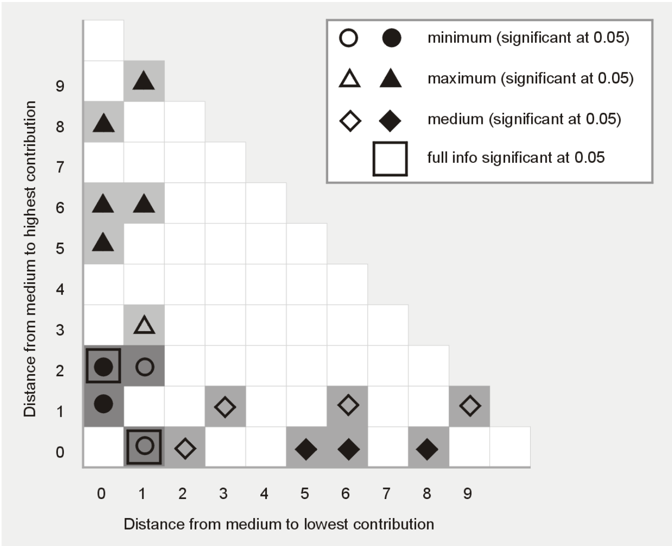


Figure 5: Impact of information type on overall contribution level

The effect of presenting either summary statistic or complete contribution profile on the total level of contributions. For each hypothetical scenario considered, the information type resulting in the highest level is marked (solid for statistical significance at the 0.05 level). The scenarios are arranged on a two-dimensional grid measuring the distances from the medium to the lowest and highest contributions of other group members.

our analysis, there seem to be only a handful of cases where providing the participants with full information significantly boosts the total level of contributions. As far as selecting a particular summary statistic from the set of four is concerned, there is no evident leader either. We will elaborate more on that in the next section.

4 Concluding Remarks

We improve upon the original experimental design of Fischbacher et al. (2001) to enable the subjects to condition their contribution behavior in a linear public goods game, whilst free from any dynamic concerns, on the minimum, median, average or maximum contribution as well as on the full contribution profile of other group members.

Our findings suggest that conditional cooperation is a stable phenomenon that is not affected by the choice of a conditioning statistic used to classify the subjects as exhibiting such behavior. Approximately two thirds of subjects in our experiment exhibit persistent behavior, with roughly half of those classified as free riders and another half as conditional cooperators.

Never the less, the choice of a conditioning statistic is not a mute point at all. Letting conditional cooperators know the minimum contribution among other group members effectively mitigates the self-serving bias usually found in the literature.

We obtain further evidence that conditional cooperators differ from other behavioral archetypes in their preferences over the summary statistics over the contribution profile of other group members. However, unlike previous literature, we can also identify subjects who have no such preference. Our results suggest that subjects, regardless of the demographic, prefer the minimum, with the median being their second choice. In the end, it is then the share of persons that care about any statistic at all that separates conditional contributors from the rest. It is also noteworthy that 7 out 29 free riders in our experiment seem to have a specific preference over the summary statistics, which is at odds with what is usually found in the literature.

As the next step, we contrast contribution behavior in the full information treatment with that under limited information to determine which summary statistic is a better predictor of one's own contribution. Using the mean square deviation as a measure of accuracy enables us to consider all four statistics at once. Conditional cooperators and approximately two thirds of the rest of the subjects can be thought of as reciprocating one particular summary statistic over the complete contribution profile of other group members. For about half of the former and more than one-third of the latter, this statistic is the average, which is quite peculiar in light of what the subjects actually state that they prefer. If we exclude the average and repeat the analysis, we observe that the majority of the subjects tend to reciprocate the median, which is on par with what is usually found in the literature. We thus conclude that, as far as the minimum, median, maximum and average are concerned, it is the latter that actually drives one's contribution behavior, albeit the subjects do not seem to fully realize that.

Finally, we investigate the effect of information available to the subjects on the total contribution level. Our findings suggest that providing full information is hardly ever the best way to go from this point of view. Selecting a particular statistic from the notorious set of four is not an easy task either as the clear victor can be declared only in half of the scenarios considered. Even then, making generalizations is not that straightforward as merely grouping such scenarios is a challenge in its own right. Some interesting observations can be made none the less. Please consider Figure 5 (the complete test results, for each information protocol, are available in Table 6).

First of all, one would want to avoid revealing the average, arguably the most informative summary statistic of the four, at all costs. For fairly even distributions, the minimum should be considered, albeit it is also the region where obtaining statistical significance is most difficult. As the distributions get more and more skewed, the median or maximum, depending on the direction, begin to stand out. Unfortunately, any attempt to explain these findings would be pure speculation at this stage. Perhaps some further studies could look into whether or not the subjects are actually trying to back engineer the original distributions from given values of a summary statistic, and how successful they are at that.

5 Acknowledgements

This project was made possible through the generous grant support of the Global Development Network (GDN), a leading international organization of developing and transition country researchers, policy and research institutes promoting the generation, sharing, and application to policy of multidisciplinary knowledge for the purpose of development.

All opinions expressed in the paper are those of the authors, and have not been officially endorsed by the GDN.

6 References

1. Bigoni, M., Suetens, S., 2012. Feedback and dynamics in public good experiments. J. Econ. Behav. Organ. 82, 86–95.

2. Bolton, G.E., Ockenfels A., 2000. ERC: A Theory of Equity, Reciprocity, and Competition. Am. Econ. Rev. 90:1, 166–193.

3. Croson, R., 2006. Theories of commitment, altruism, and reciprocity: Evidence from linear public goods games. Econ. Inq. 45:2, 199–216.

4. Fehr, E., Schmidt, K.M., 1999. A Theory of Fairness, Competition and Cooperation. Q. J. Econ. 114:3, 817–868.

5. Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10:2, 171–178.

6. Fischbacher, U., Gachter, S., Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. Econ. Lett. 71, 397–404.

7. Kurzban, R., Descioli, P., 2008. Reciprocity in groups: Information-seeking in a public goods game. Eur. J. Soc. Psychol. 38, 139–158.

8. Ledyard, J.O., 1995. Public Goods: A Survey of Experimental Research, Princeton University Press, Princeton, NJ.

9. Ockenfels, A., 1999. Fairness, Reziprozität und Eigennutz. Ökonomische Theorie und experimentelle Evidenz. Bd. 108. Mohr Siebeck, Tübingen.

10. Reuben, E., Riedl A., 2013. Enforcement of contribution norms in public good games with heterogeneous populations. Games Econ. Behav. 77:1, 122–137.

11. Selten, R., 1967. Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments. In: Sauermann, H. (Ed.), Beiträge zur experimentellen Wirtschaftsforschung Vol. I, 136–168.

12. Sonnemans, J., Schram, A., Offerman, T., 1999. Strategic behavior in public good games: when partners drift apart. Econ. Lett. 62, 35–41.

13. Sugden, R., 1984. Reciprocity: The Supply of Public Goods through Voluntary Contributions. Econ. J. 94, 772–787.

14. Zetland, D., Della Giusta, M., 2012. I See What You're Doing: The Role of Gender in Cooperation, SSRN Work. Pap. 1122144.

Appendix. Auxiliary Tables and Figures

Table 5: Contribution scenarios used in experiment

Scenario	Individual Contribution Levels			Average (ref.)
	Lowest	Medium	Highest	
1	0	0	1	0
2	0	0	6	2
3	0	1	3	1
4	0	1	10	4
5	0	2	2	1
6	0	3	4	2
7	0	5	5	3
8	0	6	7	4
9	0	8	8	5
10	0	9	10	6
11	1	4	5	3
12	1	7	8	5
13	2	2	10	5
14	2	3	9	5
15	3	4	10	6
16	4	10	10	8
17	5	5	10	7
18	5	6	9	7
19	6	7	10	8
20	7	9	10	9
21	8	8	10	9
22	9	10	10	10

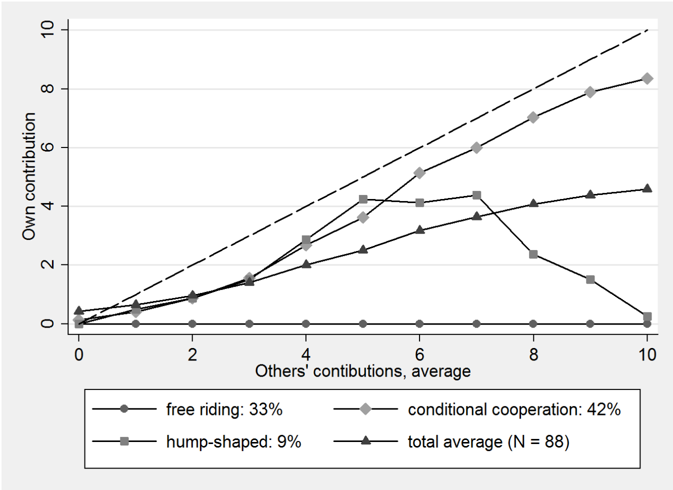


Figure 6: Own contribution as function of average contribution among other players

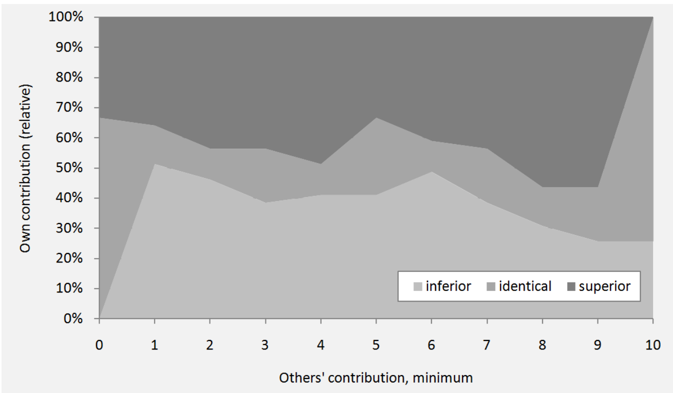


Figure 7: Own contribution vis-a-vis minimum contribution of others. Relative incidence of inferior, identical and superior contributions

Table 6: Impact of information type on overall contribution level. Bootstrapping results

Contribution Levels			Full Info vs. All Stats			Minimum vs. Other Stats			Median vs. Other Stats			Maximum vs. Other Stats			Average vs. Other Stats			Distance from Med. to	
			Difference		Lower	Difference		Lower	Difference		Lower	Difference		Lower	Difference		Lower		
Low.	Med.	High.	Obs.	Est.	Bound	Obs.	Est.	Bound	Obs.	Est.	Bound	Obs.	Est.	Bound	Obs.	Est.	Bound	Low.	High.
0	0	1	-0.5455	-0.5542	-0.9205	*0.3864*	*0.3693*	*0.0000*	-0.3864	-0.3905	-0.8295	-0.4432	-0.4502	-0.8750	-0.7159	-0.7244	-1.1023	0	1
0	0	6	-1.0114	-1.0121	-1.4205	-1.4318	-1.4331	-1.8295	-1.8182	-1.8195	-2.2614	*1.4318*	*1.4182*	*1.1136*	-1.6023	-1.6035	-2.0000	0	6
0	1	3	-0.1250	-0.2070	-0.3068	0.0227	0.0212	-0.3636	-0.2614	-0.3432	-0.4773	-0.0227	-0.0392	-0.3864	-0.4886	-0.5709	-0.6705	1	2
0	1	10	-1.6591	-1.6599	-2.2614	-3.0227	-3.0236	-3.6705	-3.2841	-3.2848	-3.9659	*2.1477*	*2.1488*	*1.6136*	-2.1477	-2.1488	-2.7500	1	9
0	2	2	0.1136	0.0774	-0.2386	-0.1250	-0.1253	-0.4773	0.1250	0.1157	-0.2159	-0.4773	-0.5127	-0.7727	-0.6136	-0.6495	-0.8864	2	0
0	3	4	-0.0455	-0.0820	-0.3409	-0.6136	-0.6508	-0.9091	0.0568	0.0567	-0.2159	-0.0568	-0.0569	-0.2841	-0.7841	-0.8212	-1.0341	3	1
0	5	5	-0.6364	-0.6366	-1.0455	-1.7159	-1.7163	-2.1023	*0.7159*	*0.7158*	*0.4432*	-0.7159	-0.7158	-1.0227	-1.4432	-1.4437	-1.8068	5	0
0	6	7	-0.7727	-0.8003	-1.1591	-2.1477	-2.1756	-2.6136	0.1023	0.1018	-0.1591	-0.1023	-0.1018	-0.3864	-1.2727	-1.3008	-1.6818	6	1
0	8	8	-0.7614	-0.7630	-1.3182	-3.0000	-3.0015	-3.6477	*0.6023*	*0.6024*	*0.1705*	-0.6023	-0.6024	-1.0227	-1.6250	-1.6262	-2.1591	8	0
0	9	10	-0.9205	-0.9286	-1.4318	-3.3068	-3.3151	-3.9773	0.2841	0.2840	-0.0455	-0.2841	-0.2840	-0.5909	-1.2614	-1.2694	-1.7500	9	1
1	4	5	-0.1136	-0.1230	-0.5000	-0.9545	-0.9647	-1.2841	0.1705	0.1702	-0.1023	-0.1705	-0.1702	-0.3977	-0.8977	-0.9081	-1.2386	3	1
1	7	8	-0.4545	-0.4807	-0.7955	-2.3636	-2.3894	-2.8750	0.1818	0.1818	-0.1932	-0.1818	-0.1818	-0.5227	-1.2045	-1.2302	-1.5909	6	1
2	2	10	-1.7273	-1.7279	-2.3182	-2.2500	-2.2505	-2.8409	-2.8977	-2.8983	-3.5568	*1.6477*	*1.6483*	*1.1250*	-1.6477	-1.6483	-2.2273	0	8
2	3	9	-1.1818	-1.1828	-1.7045	-1.9886	-1.9887	-2.5227	-2.1477	-2.1479	-2.7273	*1.3864*	*1.3864*	*0.8977*	-1.3864	-1.3865	-1.9545	1	6
3	4	10	-0.9545	-0.9555	-1.5000	-2.0795	-2.0800	-2.6364	-1.8523	-1.8528	-2.3750	*0.9773*	*0.9779*	*0.5114*	-0.9773	-0.9779	-1.4773	1	6
4	10	10	-0.2159	-0.2175	-0.8409	-1.9659	-1.9683	-2.5227	*0.6136*	*0.5366*	*0.2955*	-0.6136	-0.6148	-1.0114	-0.6932	-0.6941	-1.1477	6	0
5	5	10	-0.6364	-0.6403	-1.2045	-1.0114	-1.0159	-1.5114	-1.3068	-1.3110	-1.8182	*0.5114*	*0.5119*	*0.0682*	-0.5114	-0.5122	-1.0114	0	5
5	6	9	-0.4773	-0.5049	-0.9659	-0.7500	-0.7782	-1.1705	-0.6136	-0.6418	-0.9205	0.2500	0.2465	-0.1932	-0.2500	-0.2505	-0.7273	1	3
6	7	10	0.1136	0.0358	-0.2386	-0.5114	-0.5902	-0.8523	-0.4432	-0.5216	-0.7500	0.0795	0.0648	-0.3409	-0.0795	-0.0820	-0.5682	1	3
7	9	10	0.2500	0.1806	-0.2500	-0.3523	-0.4207	-0.6364	0.0682	0.0235	-0.2614	-0.2841	-0.3490	-0.5795	-0.0682	-0.0789	-0.4205	2	1
8	8	10	*0.5568*	*0.5433*	*0.0000*	*0.2273*	*0.1794*	*0.0000*	-0.4659	-0.4781	-0.8409	-0.4432	-0.4505	-0.9091	-0.2273	-0.2377	-0.4659	0	2
9	10	10	*0.6591*	*0.5898*	*0.1818*	0.0795	0.0450	-0.2614	-0.0795	-0.0963	-0.4659	-0.6932	-0.7635	-1.0682	-0.2614	-0.3258	-0.5909	1	0

The actual values of the observed and estimated differences as well as the lower 95% confidence bounds of the test statistics defined in (5) and (6) for each of the 22 contribution scenarios considered.