Reciprocity and the Tragedies of Maintaining and Providing the Commons

Simon Gächter\textsuperscript{1,4,5,*}, Felix Kölle\textsuperscript{2} & Simone Quercia\textsuperscript{3}

Affiliations:
\textsuperscript{1} School of Economics, University of Nottingham, UK.
\textsuperscript{2} Faculty of Management, Economics and Social Sciences, University of Cologne, Germany.
\textsuperscript{3} Institute for Applied Microeconomics, University of Bonn, Germany.
\textsuperscript{4} CESifo, Munich, Germany.
\textsuperscript{5} IZA, Bonn, Germany.
*Correspondence to: simon.gaechter@nottingham.ac.uk.

27 July 2017

Keywords: Tragedy of the Commons, public goods, strong and weak reciprocity, evolution of human cooperation

Social cooperation often requires collectively beneficial but individually costly restraint to maintain a public good\textsuperscript{1-4}, or it needs costly generosity to create one\textsuperscript{1,5}. Status quo effects\textsuperscript{6} predict that maintaining a public good is easier than providing a new one. Here we show experimentally and with simulations that even under identical incentives, low levels of cooperation (the ‘tragedy of the commons’\textsuperscript{2}) are systematically more likely in Maintenance than Provision. Across three series of experiments, we find that strong and weak positive reciprocity, known to be fundamental tendencies underpinning human cooperation\textsuperscript{2-10}, are substantially diminished under Maintenance compared to Provision. As we show in a fourth experiment, the opposite holds for negative reciprocity (‘punishment’). Our findings suggest that incentives to avoid the ‘tragedy of the commons’ need to contend with dilemma-specific reciprocity.
Humans are an exceptionally cooperative species able to collaborate for the creation of common benefit. Collective actions such as voting, participating in political movements, the provision of the welfare state, charity, volunteering and teamwork are examples of public goods that come into existence by the generosity of many people that puts the greater good before self-interest. Cooperation is, however, not always about providing collectively valuable resources, but often about maintaining existing ones. Limiting CO2 emissions, sustaining natural resources, or maintaining common pastures and biodiversity are important examples of cooperation problems that require restraint in exploiting existing socially beneficial public goods.

In this paper, we show experimentally and with simulations that cooperation for maintaining an initially existing public good is substantially and systematically weaker than cooperation for creating a new public good even if they are otherwise identical social dilemmas. This is unexpected, given that many people are biased towards the status quo and defaults, which should ease cooperation when the public good already exists compared to when it needs to be provided.

We show that the reason for lower cooperation in the maintenance dilemma is that reciprocity, a fundamental force behind the evolution of cooperation and human sociality, is substantially diminished in maintaining compared to providing a public good. Simulations show that, despite some variability, lower cooperation in Maintenance than Provision is a systematic effect to be expected with a likelihood of 70%. The simulation results also provide an explanation for the mixed findings in some related literature.

In our experiments, we focus sharply on the behavioural differences between initially existing and inexistent public goods (Fig. 1) and abstract from technological complexities, loss aversion, time discounting and institutional details relevant in real
world social dilemmas\textsuperscript{1,24-29}. In ‘Maintenance’, a group of four people possesses a
common pool of 80 tokens and each member can withdraw up to 20 tokens.
Upholding the status quo by withdrawing nothing earns each group member 32
money units (MU); if all withdraw maximally, everyone earns 20 MU. In ‘Provision’,
the common pool is initially empty and 80 tokens are distributed equally among group
members who decide simultaneously how many tokens (up to 20) to contribute to the
pool. In the status quo all earn 20 MU, and all contributing maximally earns each
member 32 MU.
Using the setup described in Fig. 1, we run three series of experiments with 704
participants who interact anonymously in three generic settings of social interaction
(see Methods). All experiments involve a between-subjects comparison of
cooperation in Maintenance and Provision. We also elicit beliefs about group
members’ contributions to (or withdrawals from) the public good. Participants need to
successfully complete a comprehension test before the experiment starts.
In the first experiment, called One-shot, participants \((n = 288)\) take a single
decision only. This experiment is a basic measure of people’s cooperativeness in the
absence of strategic incentives to cooperate. In a second experiment, called Strangers,
participants \((n = 256)\) play the games of Fig. 1 for 27 iterations with randomly
changing group composition in each round. This experiment is a sequence of one-shot
interactions that permit learning about cooperativeness in the population\textsuperscript{30-32}. The
third experiment \((n = 160)\), called Partners, keeps group composition constant across
the 27 iterations, which creates strategic incentives for cooperation\textsuperscript{32,33}.
The effective size of the public good (after withdrawals or contributions) is
smaller in Maintenance than Provision in all experiments (Fig. 2, Supplementary
Table 1). In One-shot, the public good in Maintenance is on average 27% smaller than
in Provision (Fig. 2, Panel 1; 23.8 vs. 32.6; two-sided t-test, $t = -2.51, P = 0.014$).

Low levels of the public good (less than 10% of the optimal size of 80), are more likely in Maintenance than Provision (23% vs. 0%; $\chi^2(1) = 9.51, P = 0.002$).

In Strangers, the public good starts out 23% lower in Maintenance than Provision (22.7 vs. 29.5; two-sided t-test, $t = -1.92, P = 0.059$) and decays on average to about 5% of the socially efficient level in both problems (Fig. 2, Panel 2). Thus, the tragedy of the commons is almost maximal in both Maintenance and Provision.

In Partners, the public good starts 33% smaller in Maintenance than Provision (27.7 vs. 41.3; two-sided t-test, $t = -2.96, P = 0.005$) and drops over time (Fig. 2, Panel 3). On average the public good is 37.3% smaller in Maintenance than Provision (10.6 vs.16.9; linear mixed effects model, $P = 0.035$).

Comparing Partners and Strangers reveals the extent to which strategic incentives help the provision of the public good. We find that in Maintenance the average size of the public good is only 3.6 units higher in Partners than Strangers (10.6 vs. 7.0; linear mixed-effects model, $P = 0.346$, Supplementary Table 2), while in Provision the public good is on average twice as large in Partners than Strangers (16.9 vs. 8.2; linear mixed effects model, $P = 0.004$). Thus, strategic incentives to increase cooperation are substantially weaker in Maintenance than Provision.

Taken together, these results show that high levels of the public good are harder to achieve in Maintenance than Provision in One-shot and in the first period of Partners and Strangers. This is surprising given that in Maintenance the public good enjoys a head start because it is already provided at the outset. Furthermore, while in Strangers the size of the public good converges to similar long-run equilibrium levels, in Partners the initial differences are persistent and lead to different long-run outcomes between Provision and Maintenance. The aim of our further analysis is to
understand the differences in cooperation outcomes by investigating whether initial
resource allocation affects reciprocity in response to restraint and generosity,
respectively.

Studying reciprocity is particularly interesting due to its fundamental role for
human sociality\(^7\)\(^-\)\(^10\). In our settings, reciprocity takes the form of conditional
coopera\tion: the willingness to cooperate provided others do the same\(^30\),\(^32\),\(^34\),\(^35\). Here, we distinguish between two forms of conditional cooperation, which are inspired by
the concepts of weak and strong reciprocity\(^9\),\(^10\),\(^36\). Weak reciprocity can occur in stable
relationships and means behaving conditionally cooperative for self-regarding
strategic reasons. By contrast, strong reciprocity entails non-selfish conditional
cooperation not only in repeated interactions but also in one-shot games. Strong
reciprocity is a preference for conditional cooperation, whereas weak reciprocity is a
behavioural strategy deployed for self-regarding reasons.

Studying reciprocity as a preference requires looking beyond cooperation
outcomes and to measure attitudes to cooperation separately from outcomes. The
reason why this is important is that people who differ in their \textit{ex ante} attitudes can \textit{ex
post} make the same cooperation decision. To see why, consider that a conditional
cooperator’s \textit{ex ante} attitude is to cooperate only if they believe their group members
do so too. But there may also be ‘free riders’, who never want to contribute to the
public good irrespective of their beliefs how much others contribute. A conditional
cooperator who believes that others do not contribute and a person with a free rider
attitude both contribute nothing: their \textit{ex post} behavior is observationally equivalent
despite different \textit{ex ante} attitudes. Thus, if cooperation is a function of attitudes and
beliefs, the challenge is to separate them empirically. Our approach, which we call the
‘ABC of cooperation’, achieves this separation. This also allows us to compare strong
reciprocity as measured by the ABC approach with reciprocity estimated from
observed behaviour.

The ABC approach measures individual attitudes ($a_i$), beliefs ($b_i$), and effective
contributions ($c_i$) separately and explains cooperation as $a_i(b_i) \rightarrow c_i$. It is inspired
by and implemented as follows. All three experiments start with an incentive-
compatible elicitation of attitudes without feedback in a one-shot version of either the
Maintenance or the Provision dilemma. The elicited attitudes are our main measure of
strong reciprocity. Eliciting attitudes involves specifying a vector $a_i$ of contributions
or withdrawals as a function of all possible average contributions or withdrawals of
other group members. We classify participants as conditional cooperators (that is,
strong reciprocators) if the entries in the vector $a_i$ are increasing in others’
contributions or withdrawals, or as a free rider if a participant’s $a_i$ consists of only
zero contributions or maximal withdrawals. We refer to the remaining participants as
‘others’. After attitude elicitation, the three experiments proceed as described above.

In all experiments, we elicit incentivized beliefs ($b_i$) about other group members’
average withdrawal or contribution and we observe effective contributions ($c_i$) to the
public good (see Methods).

In the repeated direct interactions of Strangers and Partners we measure
conditional cooperation in linear mixed-effects models by regressing individual
contributions or withdrawals on the average contribution or withdrawals of other
group members in the previous period (Supplementary Information). The relation
between these two variables, the coefficient $\beta_1$, is our measure of conditional
cooperation. We will call $\beta_1$ ‘estimated reciprocity’.

In Strangers, $\beta_1$ is an estimate of strong reciprocity because there are no
strategic incentives to pretend being a reciprocator. Because $\beta_1$ is estimated from
behavior only, it is a proxy for strong reciprocity. But we expect that participants with attitudes that classify them as conditional cooperators will have $\beta_1 > 0$, whereas people with a free rider attitude will display $\beta_1 \approx 0$.

In Partners, conditional cooperators will also have $\beta_1 > 0$, which may be larger than in Strangers due to added incentives for weak reciprocity. Free riders may therefore also display $\beta_1 > 0$. Furthermore, we will use the attitudes $a_i$ and $\beta_1$ to study the link between strong and weak reciprocity.

Elicited attitudes are significantly different in Maintenance and Provision ($\chi^2(2) = 31.03, P < 0.001$; Fig. 3a). In Maintenance, participants are significantly less likely to be conditional cooperators than in Provision (42% vs. 64%; $\chi^2(1) = 31.03, P < 0.001$); are significantly more likely to be free riders (28% vs. 17%; $\chi^2(1) = 10.46, P = 0.001$) and are also significantly more likely to display an unclassified attitude (‘others’: 30% vs. 19%; $\chi^2(1) = 11.08, P = 0.001$). Thus, in Maintenance 58% of participants do not reciprocate their group member’s effective contributions, which is almost the mirror image of the 64% in Provision who do reciprocate.

Estimated reciprocity $\beta_1$ in the repeated games is also significantly lower in Maintenance than Provision in both Strangers and Partners (Fig. 3b, panel 1; multilevel mixed-effects models, $P < 0.001$; Supplementary Table 3). The added strategic incentives for weak reciprocity significantly increase estimated reciprocity in both Maintenance and Provision (Fig. 3b, panel 1; multilevel mixed-effects models, $P < 0.001$; Supplementary Table 4).

Estimated reciprocity is also consistent with attitude types elicited prior to the repeated games (Supplementary Tables 5-6). Participants classified as conditional cooperators show high degrees of estimated reciprocity in Strangers and Partners, significantly above that of free riders in both Maintenance and Provision (Fig. 3b,
panels 2 and 3; multilevel mixed-effects models, $P < 0.001$). Conditional cooperators also display significantly higher $\beta_1$ in Partners than Strangers (multilevel mixed-effects models, $P < 0.001$). As predicted, free riders in Strangers display low estimated reciprocity but show increased $\beta_1$ in Partners compared to Strangers. Participants classified as ‘others’ do display a substantial $\beta_1$ but do not react to strategic incentives (Fig. 3b, panel 4; multilevel mixed-effects models, $P > 0.166$).

Our next step is to investigate whether the differences in reciprocity across Maintenance and Provision can explain the observed differences in cooperation outcomes (Fig. 2). We do this by applying our ABC framework that uses attitudes and beliefs to explain effective contributions. We calculate predicted effective contributions $[a_i(b_i) \rightarrow c_i^*]$ and compare them with actual effective contributions $c_i$ from One-shot as well as with the effective first-period contributions in the repeated experiments (Methods). Predicted and actual effective contributions are highly significantly positively correlated in One-shot as well as in all repeated games (all Spearman’s $\rho > 0.59; P < 0.001$).

We also calculate individual-level deviations from the predicted effective contribution, $c_i^* - c_i$. In One-shot, this measure lies within ± 2 tokens in 63% and 62% of the cases in Maintenance and Provision, respectively, with no differences between treatments ($\chi^2(1) = 0.01, P = 0.903$). We obtain similar results for first-period effective contributions in Strangers (66% and 63%; $\chi^2(1) = 0.43, P = 0.514$) and Partners (74% and 64%; $\chi^2(1) = 1.86, P = 0.172$). Finally, effective contributions differ significantly between attitude types: free riders contribute significantly less than conditional cooperators and ‘others’ in all conditions (Supplementary Fig. 1).

The fact that the ABC approach predicts equally well in Maintenance and Provision allows us to use the elicited attitudes and beliefs as a ‘population pool’ from
which we can sample at random to run ‘simulated experiments’ (Methods and Supplementary Information). The advantage of simulations is that we are not restricted to a specific laboratory sample we happen to draw at a given instance (with hitherto unobservable attitudes and beliefs); we can cost effectively perform a large number of identical experiments and therefore elicit a distribution of likely cooperation ratios of Maintenance relative to Provision. This also allows us to check how systematic the results are that we observe.

The results of 1000 simulated experiments (Fig. 4) show that effective cooperation levels in Maintenance are lower than in Provision in 70% of all simulated experiments. This result shows that our findings that cooperation in Maintenance is lower than in Provision are systematic.

Given that our results reveal important asymmetries in positive reciprocity between Maintenance and Provision, it is interesting to study whether initial resource allocation also affects negative reciprocity, which in our setting takes the form of punishment\(^9,36\). Furthermore, punishment is an expression of moral disapproval and social norms\(^37\) that are important in many real world public goods\(^38\). If the differences in positive reciprocity in Maintenance and Provision also translate into negative reciprocity, we should observe less punishment in Maintenance than Provision and, therefore, also a reduced effectiveness of punishment to stabilize cooperation in Maintenance compared to Provision.

We study punishment in a fourth experiment (‘Partners with Punishment’; \(n = 172\)), which is identical to Partners except for an added punishment stage in each period after group members have made their withdrawal or contribution decisions\(^39\). In the punishment stage, each group member can assign up to 5 punishment points to
each other member, where each punishment point costs one MU and reduces the
earnings of the punished group member by three MU (see Methods).

The attitudes elicited prior to the experiment replicate the results from Fig. 3a
(Supplementary Fig. 2). Contrary to expectations, negative reciprocity, estimated as
assigned punishment in reaction to negative deviations of others from own effective
contribution, is substantially and significantly higher in Maintenance than in
Provision. This effect is present both overall and for each attitude type (Fig. 5a, panels
1-4), and it is not driven by different frequencies of punishers (Supplementary Figure
3). There are no treatment differences for positive deviations (Fig. 5a, panel 1;
Supplementary Table 7). Interestingly, in contrast to estimated positive reciprocity,
estimated negative reciprocity does not differ between conditional cooperators and
free riders (Fig. 5a, panels 2-4; Supplementary Table 8).

As expected⁴⁰, punishment increases the public goods to substantially higher
levels compared to Partners (Fig. 5b; linear mixed effect models; Maintenance: 43.1
vs. 10.6, \( P < 0.001 \); Provision: 44.1 vs. 16.9, \( P < 0.001 \); Supplementary Tables 9-10).

Remarkably, the sizes of public goods are now very similar in Maintenance and
Provision (linear mixed effect models; Maintenance: 43.1, Provision: 44.1, \( P =
0.904 \)). Besides stronger negative reciprocity, a further reason for this result is that
reactions to received punishment (in terms of change in effective contributions) are
also stronger in Maintenance than Provision (Supplementary Table 11).

One way to reconcile the results on positive and negative reciprocity in Partners
and Partners with Punishment, respectively, is to argue that also in Partners people
engage in punishment by reducing their contributions in the current period as a
reaction to previous negative deviations of others from own effective contributions. If
such ‘implicit’ punishment is stronger in Maintenance than Provision, it could explain
why the decay in effective contributions is stronger in Maintenance than Provision.

However, this conjecture is not borne out by the data.

We find that participants in Partners significantly increase their contributions in round $t$ in response to positive deviations of others from own contributions in round $t-1$; the reverse holds for negative deviations. However, we find both of these reactions to be significantly more pronounced in Provision than in Maintenance (linear mixed effect models; both $P < 0.018$; Supplementary Table 12). This confirms once again stronger conditional cooperation in Provision compared to Maintenance in Partners. It also suggests another interpretation of the results of Partners with Punishment: because voluntary conditional cooperation is weaker in Maintenance than Provision, stronger extrinsic incentives are needed, here in the form of punishment, to stabilize cooperation in Maintenance at similar levels than in Provision.

Our analysis has revealed that the important principles of human cooperation of strong and weak reciprocity$^{7-10}$ are substantially diminished when cooperation requires restraint in exploiting a public good as opposed to when cooperation calls for generosity to provide a public good. Our findings are consistent with the observation that failing to contribute to a public good is judged more morally blameworthy than exploiting an existing public good.$^{41}$

Our results can also be explained by a model of revealed altruism$^{42,43}$, according to which initial resource allocation affects perceptions of generosity of actions and hence subsequent reciprocity. Because in Provision cooperation is the result of an act of commission (contributing), while in Maintenance cooperation is achieved by omission (not withdrawing), cooperation in Provision is perceived as more generous than in Maintenance and thus Provision triggers stronger positive reciprocity than
Maintenance. By contrast, our results suggest that negative reciprocity, as expressed by people’s costly punishment, does not follow this logic because punishment is more severe in Maintenance than Provision, likely to compensate for weaker voluntary cooperation in Maintenance.

Our findings from the experiments without punishment and the simulations also help explaining the mixed evidence from previous related literature, which, with a few exceptions\(^35,44,45\), only compared cooperation outcomes, that is, the effective size of the public good after contribution or withdrawal decisions. Some of these studies find higher cooperation in so-called ‘give-some’ vs. ‘take-some’ games\(^14-17\), some find the reverse\(^18\) and some find no significant differences\(^19-23\). The simulations based on our ABC approach can explain these mixed results (Fig. 4) but they also show that on average cooperation in Maintenance is generally expected to be lower than in Provision. The finding that Maintenance and Provision are systematically different also suggests that future research should choose the game (Maintenance or Provision) that comes closest to the social dilemma of interest.

Our results also have potential policy relevance.\(^46\) Recent policy proposals to foster cooperation build on the power of reciprocity in combination with economic incentives\(^47,48\). Policy makers who reckon with reciprocity should therefore consider that the extent of reciprocity that can be evoked is dilemma-specific. Moreover, a problem of incentives is that they might ‘crowd out’ strong reciprocity because incentives typically strengthen self-regarding motives to cooperate\(^49\). Our finding of higher reciprocity in Provision than Maintenance suggests that crowding out may be more problematic in provision problems than in maintenance problems, because in Maintenance more people display non-reciprocal attitudes. Future research will need to address these issues, including how reciprocity and incentives interact in non-linear
settings with thresholds, resource rivalry, discounting, and hybrid social dilemmas

where provision and exploitation can take place at the same time.

**Methods**

**Isomorphism of Maintenance and Provision under monetary incentives.** In *Maintenance*, each group of 4 members is initially endowed with 80 tokens placed in a “group project”; individual members have no endowment. Material incentives are described by equation (1):

\[
\pi_i = w_i + 0.4(80 - \sum_{j=1}^{4} w_j)
\]  

(1)

where \(0 \leq w_i \leq 20\) indicates the withdrawal of individual \(i\) from the project.

In *Provision*, the “group project” is initially empty and each group member has an endowment of 20 tokens instead. The material incentives for each individual \(i\) are described by equation (2):

\[
\pi_i = 20 - c_i + 0.4 \sum_{j=1}^{4} c_j
\]  

(2)

where \(0 \leq c_i \leq 20\) denotes the contribution of individual \(i\) to the project.

Hence, under rationality and money maximization, Maintenance and Provision are isomorphic social dilemmas. Using \(c_j = 20 - w_j\) for \(j = 1, \ldots, 4\) and substituting into eq. (2) we obtain (1). Analogously, using \(w_j = 20 - c_j\) for \(j = 1, \ldots, 4\) and substituting into (1) yields (2).

**Experimental design details.** The experiments were approved by the Research Ethics Committee in the School of Economics at the University of Nottingham. We conducted four series of experiments using the two decision situations described above and in Fig. 1. Each experiment was composed of three parts that allow to elicit the three components of our ABC framework: an individual \(i\)’s attitude \((a_i)\) towards cooperation \((i\)’s ‘type’), \(i\)’s beliefs \((b_i)\) about others’ contribution, and \(i\)’s contribution decision \((c_i)\). Participants knew that the experiment consisted of several parts but only received information about the relevant part upon progression of the experiment. To avoid spillover effects between different parts, information about decisions and payoffs were given only at the very end of the experiment. Experimental instructions are in the Supplementary Information.

In Part 1, participants were introduced to either the Maintenance or Provision problem. Before continuing, participants answered a set of computerized control questions.

In Part 2, we elicited cooperation attitudes \(a_i\) using a variant of the strategy method\textsuperscript{50}, which allows eliciting an individual’s willingness to cooperate as a function of the other group members’ cooperation decisions. Participants were asked to make an unconditional and a conditional cooperation decision. In the unconditional decision, participants were simply asked how much they want to withdraw from (contribute to) the common pool. In the conditional contribution participants had to fill a withdrawal (contribution) table in which they had to state
their withdrawal (contribution) decision for each possible (rounded) average withdrawal (contribution) of the other three group members. This gives us the vector $a_i$, our measure of strong reciprocity. To achieve incentive compatibility, in each group a random mechanism selected three members for which the unconditional decision was payoff-relevant and one member for whom the conditional decision for the (rounded) average unconditional withdrawal (contribution) of the three other group members was payoff-relevant.

Part 3 comprised a direct-response interaction that differed in its exact design protocol across the four experiments as described in the main text (One-shot, Strangers, Partners, and Partners with Punishment). This elicits component $c_i$ of the ABC framework.

In all repeated experiments (Strangers, Partners, and Partners with Punishment), participants were matched in groups of four and interacted for 27 consecutive rounds under payment rules (1) or (2). Participants were not told how many rounds the experiment would last. This avoids endgame effects and also seems realistic for many common resource problems, which do not have a known endpoint. In Strangers, participants were re-matched randomly in 16-participants matching groups after every round, while in Partners and Partners with Punishment group composition remained constant across all 27 rounds. At the end of each round, participants received aggregate feedback on choices and outcomes.

In all rounds of the direct-response interactions, we also elicited beliefs about average effective contributions of the other three group members. Participants were paid for the accuracy of their beliefs. They earned 3 points if their belief was exactly correct, and 2 (1) points when their belief deviated by 1 (2) point(s) from the true average effective contribution. If their estimation was off by more than two points, they received no additional money. This elicits component $b_i$ of our framework.

**Data collection and subject-pool socio-demographics.** A total of $n = 876$ students participated in our experiments (Maintenance: $n_M = 432$, Provision: $n_P = 444$). Participants were recruited with the help of ORSEE from the volunteer student subject pool at the University of Nottingham. Participation was upon informed consent. The average age was 20.1 years (s.d. 2.25 years); 57% were females. 59% were British, 22% Asian, 12% from other European countries and the rest from other countries. 20% were economics or business students, 18% other social sciences, 20% humanities, 14% sciences, 12% engineering, and 12% medical science, and 4% law. We conducted all experiments in the CeDEx lab at the University of Nottingham using z-Tree. The experiments lasted between 70 to 210 minutes depending on the experimental condition. Participants earned on average £20.60.

**Predicting effective contributions.** In One-shot, Strangers, and Partners, the ABC approach allows us to predict contributions using elicited cooperation attitudes $a_i$ and beliefs...
\[ b_t: a_t(b_t) \rightarrow c_t^*. \] By matching beliefs with the corresponding decision in the contribution (withdrawal) table, we predict a contribution (withdrawal) decision \( c_t^* \) for each subject and compare \( c_t^* \) with the actual contribution \( c_t \) that we observe in the direct-response experiment.\textsuperscript{30,54}

**Classification of attitudes.** We analyse cooperation in the strategy-method experiment treating each participant’s effective contribution schedule (the vector \( a_t \)) as an independent observation. We classify cooperation attitudes into three main behavioural types\textsuperscript{30}: a participant is a *conditional cooperator* if either his/her effective contribution schedule exhibits a (weakly) monotonically increasing pattern, or if the Spearman correlation coefficient between his/her schedule and the others’ average contribution is positive and significant at the 1\% level; a *free rider* if he/she never contributes anything (always withdraws everything) irrespective of how much others contribute (withdraw); (iii) *other* if neither (i) nor (ii) applies. Attitudes are a proxy for cooperation preferences because they reflect a willingness to pay for cooperation as a function of other group members’ cooperation.

**Simulations.** For each simulated experiment, we randomly sample (with replacement) from the participant pool of Maintenance experiments attitudes and beliefs \((n = 60, \text{the median sample size in related studies also using linear public goods}^{14-23})\) and calculate simulated effective contributions \([a_t(b_t) \rightarrow \hat{c}]\). We do the same for \(n = 60\) Provision attitudes and beliefs. This resembles an experiment where a researcher invites 60 participants per treatment and then observes their effective contribution. As a participation pool, we use all \(n = 876\) attitudes from our four experiments (Maintenance: \(n_M = 432\), Provision: \(n_P = 444\)), and \(n = 544\) beliefs (Maintenance: \(n_M = 268\), Provision: \(n_P = 276\)) from One-shot as well as the first period of Strangers. Details are in the Supplementary Information.

**Statistical analysis.** In the One-shot direct interaction, we treat \((b_t, c_t)\) as independent observations. In the repeated interactions, we treat beliefs and effective contributions at the matching group level as an independent observation. Matching groups are composed by 16 participants in Strangers and by 4 participants in Partners and Partners with Punishment, respectively. For the repeated experiments, all estimations are performed using linear mixed models with random intercepts at the matching group and the individual level (see Statistical Analysis in SI for details on model specifications).
Data availability. The data for the statistical analyses are stored in Dryad Data package title: Reciprocity in Maintaining and Providing Public Goods; http://dx.doi.org/10.5061/dryad.8d9t2

Code availability. We used STATA 14.2 for data analysis. The codes are stored in Dryad Data package title: Reciprocity in Maintaining and Providing Public Goods; http://dx.doi.org/10.5061/dryad.8d9t2

References


12. Turchin, P., Ultrasociety: How 10,000 years of war made humans the greatest cooperators on earth. (Beresta Books, LLC, Chaplin, CT, 2016).


Acknowledgements

Weisel, and various seminar audiences for helpful comments. B. Beranek provided valuable 
research assistance. This work was supported by the ERC-AdG 295707 COOPERATION and 
the ESRC Network for Integrated Behavioural Sciences (NIBS, ES/K002201/1). The funders 
had no role in study design, data collection and analysis, decision to publish, or preparation of 
the manuscript.

**Author Contributions**

SG, FK, and SQ developed the research ideas and designed the study; FK and SQ conducted 
the experiments, and analysed data. SG, FK, and SQ wrote the manuscript.

**Additional information**

Supplementary information is available for this paper.

Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints).

Correspondence should be addressed to S.G.

**Competing interest**

The authors declare no competing interests.
FIGURE LEGENDS

Figure 1 | The isomorphic social dilemmas of maintaining and providing a public good. The figure illustrates initial resource allocation in Maintenance and Provision prior to decision-making. **a**, Maintenance: Initially, group members have 0 tokens and 80 tokens are provided in the public good. Group members can simultaneously withdraw up to 20 tokens. **b**, Provision: Initially, each group member has 20 tokens and the public good is empty. Group members can simultaneously contribute up to 20 tokens. Each token withdrawn or not contributed is worth 1 MU to the respective group member alone. Each token in the common resource is worth 0.4 MU for each group member. Material incentives therefore are to withdraw 20 tokens in Maintenance and to contribute 0 tokens in Provision, yielding 20 MU for each group member. The socially beneficial decisions of withdrawing nothing and contributing everything earn each group member 32 MU. Further details are in Methods.

Figure 2 | Public good levels. Shown are the effective sizes of the public good per round after contribution or withdrawal decisions (± 1 s.e.m.). **a**, One-shot game, \(n_M = 140, n_P = 148\). **b**, **c**, Effective public goods over the 27 rounds of interactions in randomly changing groups (Strangers, \(n_M = 128, n_P = 128\)) and fixed groups (Partners, \(n_M = 80, n_P = 80\)), in Maintenance and Provision, respectively. Supplementary Table 1 reports further summary statistics, including on beliefs about other group members’ average effective contributions (which mirror the effective contributions).

Figure 3 | Reciprocity in Maintenance and Provision. **a**, Strong reciprocity as measured by cooperation attitudes; type classification as in\(^{30}\), \(n_M = 348, n_P = 356\). \(\chi^2\)-tests, *** \(P < 0.01\). Results are robust to alternative classification methods.
Estimated reciprocity in repeated interactions (± 1 s.e.m.); Strangers (S), n = 256; Partners (P), n = 160 by treatment and attitude category (conditional cooperators, free riders, others). Positive reciprocity is estimated as the coefficient of lagged average contributions of the other group members \( C_{ij,t-1} \) from multilevel mixed-effects linear regressions (Supplementary Information, Section 1.1; Supplementary Table 3). An alternative estimation approach using finite mixture models\(^{55}\) confirms these results (Supplementary Table 14).

**Figure 4** | **Simulated effective contribution ratios.** Distribution of 1000 simulated effective contribution ratios between Maintenance and Provision \( c^M / c^p \) using a sample of \( n = 60 \) per treatment and simulated experiment. The sample size reflects the median sample size in related literature.\(^{14-23}\) The mean is 0.91, median is 0.89, and IQR = 0.76 to 1.03. Further details are in Supplementary Information, Section 1.3. As a robustness check, we ran a simulation with \( n = 100 \), which returns a mean of 0.90, a median of 0.89, and an IQR of 0.79 to 1.00 (see also Supplementary Figure 4).

**Figure 5** | **Partners with Punishment.** a, Estimated negative reciprocity (± 1 s.e.m.); by treatment and attitude category (conditional cooperators, free riders, others). We estimate negative reciprocity in multilevel mixed-effects linear regression as the number of punishment points assigned to effective contributions that deviate negatively from own contribution (Supplementary Table 7; Supplementary Information). b, Shown are the effective levels of the public goods (± 1 s.e.m.) over the 27 rounds of interactions \( n_M = 84, n_P = 88 \).
FIGURE 1

(a) Maintenance

(b) Provision

80

20

20

20

20
FIGURE 2

![Graph showing public good size over rounds for one-shot, strangers, and partners with maintenance and provision lines.]

FIGURE 3

**a**

![Bar chart showing percent of conditional cooperators, free riders, and others with maintenance and provision lines.]

**b**

![Graphs showing estimated reciprocity for all, conditional cooperators, free riders, and others with maintenance and provision lines.]

[Further textual content related to the figures and data may follow.]
FIGURE 4

Proportion of mean effective contributions Maintenance/Provision

FIGURE 5

Estimated negative reciprocity

Partners with Punishment

Public good size
Supplementary Information for
Reciprocity and the Tragedies of Maintaining and Providing the Commons
Simon Gächter, Felix Kölle & Simone Quercia
correspondence to: simon.gaechter@nottingham.ac.uk

27 July 2017

1. Supplementary Methods
   1.1 Supporting Statistical Analysis
   1.2 Robustness Checks
   1.3 Simulation Analysis
   1.4 Experimental Instructions

2. Supplementary References

3. Supplementary Tables

4. Supplementary Figures
1. Supplementary Methods

1.1 Supporting Statistical Analysis

We analyse cooperation in the strategy method experiment treating each participant’s effective contribution schedule (the vector $a_i$) as an independent observation. In the One-shot direct interaction, we treat each belief ($b_i$) and effective contribution ($c_i$) as an independent observation. Finally, in the repeated interactions we treat beliefs and effective contributions at the matching group level as independent observations. Matching groups are composed of 16 participants in Strangers (see Methods in main text) and of 4 participants in Partners and Partners with Punishment, respectively.

*Figure 2 Panels 2-3 – Comparisons across treatments and experiments*

To compare the size of the public good across Maintenance and Provision (*Figure 2, Panels 2-3*), we run linear mixed-effects models with a random intercept at the matching group level. For each condition (Strangers and Partners) we estimate the following specification:

\[
PG_{jt} = \beta_0 + u_{0m} + \beta_1 \text{Provision} + \epsilon_{jt}
\]

where $PG_{jt}$ is the size of the public good in group $j$ at round $t$; $\beta_0$ is a constant; $u_{0m}$ is a random intercept at the matching group level. *Provision* is a treatment dummy that takes value one for Provision and zero for Maintenance. We estimate this model separately for Strangers and Partners.

To compare the size of the public good across Strangers and Partners (*Figure 2, Panels 2-3*) we estimate the following specification:

\[
PG_{jt} = \beta_0 + u_{0m} + \beta_1 \text{Partners} + \epsilon_{jt}
\]
where \( PG_{jt} \) is the size of the public good in group \( j \) at round \( t \); \( \beta_0 \) is a constant; \( u_{0m} \) is a random intercept at the matching group level. \( Partners \) is a treatment dummy that takes value one for Partners and zero for Strangers. We estimate this model separately for Maintenance and Provision. The results of the estimates from models 1 and 2 are reported in Supplementary Table 2.

**Figure 3a - Cooperation attitudes**

Following previous literature\(^1\)\(^-\)\(^3\), we classify cooperation attitudes into three main behavioral types: *conditional cooperators*, *free riders*, and *others*. Specifically, we classify a participant as a

(i) *conditional cooperator* if either his/her effective contribution schedule (the vector \( a_i \)) exhibits a (weakly) monotonically increasing pattern, or if the Spearman correlation coefficient between his/her schedule and the others’ average contribution is positive and significant at the 1\% level;

(ii) *free rider* if he/she never contributes anything (always withdraws everything) irrespective of how much the others contribute (withdraw); (iii) *other* if neither (i) nor (ii) applies (see Section 2.1 for robustness checks on the classification procedure).

**Figure 3b – Linear mixed-effects models**

To obtain a measure of estimated reciprocity, we estimate the following linear mixed-effects model:

\[
C_{ijt} = \beta_0 + u_{0m} + u_{0i} + \beta_1 \bar{C}_{-i,t-1} + \beta_4 Round + \varepsilon_{ijt}
\]

where \( C_{ijt} \) is the effective contribution of individual \( i \) in group \( j \) at round \( t \); \( \beta_0 \) is a constant; \( u_{0m} \) and \( u_{0i} \) are random intercepts at the matching group and individual level, respectively. \( \bar{C}_{-i,t-1} \) is the average contribution of the other three group members from the previous round. The variable \( Round \) indicates the round of the experiment and estimates a time trend. The coefficient \( \beta_1 \) is our measure of estimated reciprocity that we depict in Figure 3b in the main text.

We also estimated a model where contributions in period \( t \) are explained by beliefs about others’ contribution in period \( t \). We find significantly lower reciprocity (a positive contributions-belief correlation) in Maintenance than Provision in Strangers; in Partners reciprocity in Maintenance is also lower than in Provision, but not significantly so. The problem is, however, that beliefs in
period $t$ are strongly influenced by contributions of others in $t-1$ but the coefficients on contributions of others in $t-1$ are less than 1 in all treatments. This implies that beliefs in $t$ are revised downwards beyond the observation of others’ contributions in $t-1$. One possible reason is that beliefs in $t$ are to some extent a rationalization of planned own contributions in $t$, that is, beliefs are not fully exogenous but to some extent endogenous. We believe that this problem arises mainly in the repeated games, while in One-shot we interpret elicited beliefs as an exogenous component of the ABC framework.

A. Smith$^4$ proposed a solution to the endogeneity problem in repeated public goods games by using beliefs and effective others’ contributions in periods $t-2$ and $t-3$ as instruments for beliefs in period $t$. For the instruments to be valid, they need to be causal for beliefs but not for contributions. Following Smith$^4$, p. 422, we run Sargan and Basmann $\chi^2$ tests to determine the validity of the set of instruments. However, the null hypothesis of valid instruments is clearly rejected in our dataset (all $P < 0.001$), making this approach infeasible. We therefore use lagged effective contributions of the other group members, which are less likely to cause endogeneity problems and do not suffer from issues of reverse causality. As a consequence, our estimates might be seen as a combination of the differences in reciprocal responses to others’ previous contributions and beliefs about others’ contributions in the current period.

To compare reciprocity across Maintenance and Provision (Figure 3b, all panels), we estimate the following model:

$$C_{ijt} = \beta_0 + u_{0m} + u_{0t} + \beta_1 \bar{C}_{ij,t-1} + \beta_2 \text{Provision} + \beta_3 \bar{C}_{ij,t-1} \times \text{Provision} + \beta_4 \text{Round}$$

$$+ \beta_5 \text{Round} \times \text{Provision} + \varepsilon_{ijt}$$

(4)

where the dummy Provision indicates the treatment. We control for different time trends across Maintenance and Provision. We estimate the model above separately for Strangers and Partners both in the full sample and for each attitude type separately. The results of the estimations from model 4 are reported in Supplementary Table 3.

To compare reciprocity across Strangers and Partners (Figure 3b, all panels), we use the following model:
\[ C_{ijt} = \beta_0 + u_{0m} + u_{0i} + \beta_1 \bar{C}_{-ij,t-1} + \beta_2 \text{Partners} + \beta_3 \bar{C}_{-ij,t-1} \times \text{Partners} + \beta_4 \text{Round} \\
+ \beta_5 \text{Round} \times \text{Partners} + \epsilon_{ijt} \]  
(5)

where \text{Partners} is a dummy that takes value one in Partners and zero in Strangers. We estimate the model above separately for Maintenance and Provision both in the full sample and for each type separately. We report the results of these estimations in \textbf{Supplementary Table 4}.

Finally, we compare reciprocity across attitude types in each treatment and experiment (\textbf{Figure 3b, Panels 2-4}) by estimating the following model:

\[ C_{ijt} = \beta_0 + u_{0m} + u_{0i} + \beta_1 \bar{C}_{-ij,t-1} + \beta_2 C + \beta_3 \bar{C}_{-ij,t-1} \times C + \beta_4 OT + \beta_5 \bar{C}_{-ij,t-1} \times OT \\
+ \beta_6 \text{Round} + \beta_7 \text{Round} \times C + \beta_8 \text{Round} \times OT + \epsilon_{ijt} \]  
(6)

where \(C\) and \(OT\) are dummies that indicate whether the subject is classified in the attitude categories of \textit{conditional cooperators} or \textit{others}, respectively. We use \textit{free riders} as the omitted category. We report results of these estimations in \textbf{Supplementary Table 5}.

\textit{Comparing effective contributions across types (Supplementary Figure 1)}

To compare effective contributions across attitude types in One-shot (\textbf{Supplementary Figure 1a and 1b, Panel 1}) we run the following ordinary least squares (OLS) model:

\[ C_i = \beta_0 + \beta_1 C + \beta_2 OT + \epsilon_i \]  
(7)

where \(C_i\) is the effective contribution of participant \(i\) and \(C\) and \(OT\) are dummies for \textit{conditional cooperators} and \textit{others}, respectively. We use \textit{free riders} as the omitted category. We estimate this model separately for Maintenance and Provision.

To compare effective contributions across attitude types in Strangers and Partners (\textbf{Supplementary Figure 1a and 1b, Panels 2 and 3}) we run the following linear mixed-effects model:
\[ C_{ijt} = \beta_0 + u_{0m} + u_{0i} + \beta_1 CC + \beta_2 OT + \varepsilon_{ijt} \]  

(8)

where \( C_{ijt} \) is the effective contribution of participant \( i \) in group \( j \) at round \( t \). \( \beta_0 \) is a constant; \( u_{0m} \) and \( u_{0i} \) are random constants at the matching group and individual level, respectively. \( CC \) and \( OT \) are dummies for conditional cooperators and others, respectively. We use free riders as the omitted category. We estimate this model separately for Maintenance and Provision. The results from models 7 and 8 are reported in Supplementary Table 6.

**Figure 5a – Linear mixed-effects models in Partners with Punishment**

To investigate whether punishment behavior differs across Maintenance and Provision for given levels of positive/negative deviations between the contribution of the punisher and the punished person (Figure 5a, all panels), we estimate the following linear mixed-effects model:

\[
\begin{align*}
Y_{ijkt} &= \beta_0 + u_{0m} + u_{0i} + \beta_1 \max(c_{ijt} - c_{kjt}, 0) + \beta_2 \max(c_{kjt} - c_{ijt}, 0) + \beta_3 \text{Provision} \\
&\quad + \beta_4 \max(c_{ijt} - c_{kjt}, 0) \times \text{Provision} + \beta_5 \max(c_{kjt} - c_{ijt}, 0) \times \text{Provision} \\
&\quad + \beta_6 \text{Round} + \beta_7 \text{Round} \times \text{Provision} + \beta_8 \bar{C}_{i-kjt} + \beta_9 \bar{C}_{i-kjt} \times \text{Provision} \\
&\quad + \varepsilon_{ijt}
\end{align*}
\]  

(9)

where \( Y_{ijkt} \) is the number of punishment points assigned by individual \( i \) in group \( j \) to individual \( k \) at round \( t \); \( \beta_0 \) is a constant; \( u_{0m} \) and \( u_{0i} \) are random intercepts at the matching group and individual level, respectively. The coefficient \( \beta_1 \) estimates the effect of a negative deviation of individual \( k \)’s contribution compared to individual \( i \)’s contribution on the number of punishment points assigned from individual \( i \) to individual \( k \). The coefficient \( \beta_2 \) is analogous but for positive deviations. The coefficient \( \beta_1 \) is our estimate of negative reciprocity, that is pro-social punishment, while \( \beta_2 \) is an estimate for anti-social punishment. We depict estimates of \( \beta_1 \) and \( \beta_2 \) in Figure 5B. We also include the dummy Provision and we interact the dummy with the negative and positive deviation variables. We additionally control for the time trend including the variable Round and for the average contribution of the other two group members (\( \bar{C}_{i-kjt} \)) as well as the interaction terms.
between these variables and *Provision*. We estimate this model both for the entire sample and for each attitude type separately. We report the results from model 9 in **Supplementary Table 7**.

To compare negative reciprocity across attitude types (**Figure 5a, Panels 2-4**), we estimate a model similar to 9 including interaction terms for attitude types:

\[
Y_{ijkt} = \beta_0 + u_{0m} + u_{0i} + \beta_1 \max(c_{ijt} - c_{kjt}, 0) + \beta_2 \max(c_{kjt} - c_{ijt}, 0) + \beta_3 CC + \beta_4 OT \\
+ \beta_5 \max(c_{ijt} - c_{kjt}, 0) \times CC + \beta_6 \max(c_{kjt} - c_{ijt}, 0) \times CC \\
+ \beta_7 \max(c_{ijt} - c_{kjt}, 0) \times OT + \beta_8 \max(c_{kjt} - c_{ijt}, 0) \times OT + \beta_9 Round \\
+ \beta_{10} \text{Round} \times CC + \beta_{11} \text{Round} \times OT + \beta_{12} \bar{C}_{i-kjt} + \beta_{13} \bar{C}_{i-kjt} \times CC \\
+ \beta_{14} \bar{C}_{i-kjt} \times OT + \epsilon_{ijt}
\]

(10)

where *CC* and *OT* are dummies that indicate whether the subject is classified in the attitude categories of *conditional cooperators* or *others*, respectively. We use *free riders* as the omitted category. We estimate this model separately for Maintenance and Provision. We report estimates of this model in **Supplementary Table 8**.

**Figure 5b – Treatment comparisons in Partners with Punishment**

To compare the size of the public good between Maintenance and Provision (**Figure 5b**), we run linear mixed-effects models with random intercepts at the matching group level. Similar to Partners and Strangers, we estimate the following specification:

\[
PG_{jt} = \beta_0 + u_{0m} + \beta_1 \text{Provision} + \epsilon_{jt}
\]

(11)

where *PG* <span><i>jt</i></span> is the size of the public good in group *j* at round *t*; *\beta_0* is a constant; *u_{0m}* is a random intercept at the matching group level. *Provision* is a treatment dummy that takes value one for Provision and zero for Maintenance.

To compare the public good size between Partners with Punishment and Partners, we estimate the following specification:
\[ PG_{jt} = \beta_0 + u_{0m} + \beta_1 \text{Partners with Punishment} + \epsilon_{jt} \]  

(12)

where \( PG_{jt} \) is the size of the public good in group \( j \) at round \( t \); \( \beta_0 \) is a constant; \( u_{0m} \) is a random intercept at the matching group level. \( \text{Partners with Punishment} \) is a treatment dummy that takes value one for Partners with Punishment and zero for Partners. We estimate the model separately for Maintenance and Provision. The regression results from models 11 and 12 are reported in Supplementary Table 10.

Reactions to received punishment

As a final step in our analysis, we investigate the effectiveness of punishment by analyzing whether the change in contribution from round \( t - 1 \) to \( t \) is different in Maintenance and Provision given the same number of punishment points received in round \( t - 1 \). To compare Maintenance and Provision, we estimate the following model:

\[ C_{ij,t} - C_{ij,t-1} = \beta_0 + u_{0m} + u_{0i} + \beta_1 \text{PunReceived}_{ij,t-1} + \beta_2 \text{Provision} \\
+ \beta_3 \text{PunReceived}_{ij,t-1} \times \text{Provision} + \beta_4 \text{Round} + \beta_5 \text{Round} \times \text{Provision} \\
+ \epsilon_{ijt} \]  

(13)

where \( C_{ij,t} - C_{ij,t-1} \) is the change in contribution of individual \( i \) in group \( j \) from round \( t - 1 \) to round \( t \); \( \beta_0 \) is a constant; \( u_{0m} \) and \( u_{0i} \) are random intercepts at the matching group and individual level, respectively. The coefficient \( \beta_1 \) estimates the impact of the number of punishment points received at round \( t - 1 \) on the subsequent change in contribution at round \( t \). We also include the dummy \( \text{Provision} \) and we interact the dummy with the variable \( \text{PunReceived}_{ij,t-1} \). We additionally control for the time trend including the variable \( \text{Round} \) as well as the interaction terms between this variable and \( \text{Provision} \). To control for differential effects of pro-social and anti-social punishment, we run the above model separately for contributions that are below or above the average contribution of the group in a given round. The estimates from these models are reported in Supplementary Table 11.
Reactions to deviations of the average contribution of others from own contributions in Partners

To investigate whether contribution behavior differs across Maintenance and Provision in reaction to positive/negative deviations between the average contribution of others and own contributions in the previous period, similar to model 9 in Partners with Punishment we estimate the following linear mixed-effects model:

\[
C_{ijt} = \beta_0 + u_{0m} + u_{0i} + \beta_1 \max(\bar{C}_{-ij,t-1} - c_{ij,t-1}, 0) + \beta_2 \max(c_{ij,t-1} - \bar{C}_{-ij,t-1}, 0) \\
+ \beta_3 \text{Provision} + \beta_4 \max(\bar{C}_{-ij,t-1} - c_{ij,t-1}, 0) \times \text{Provision} \\
+ \beta_5 \max(c_{ij,t-1} - \bar{C}_{-ij,t-1}, 0) \times \text{Provision} + \beta_6 \text{Round} + \beta_7 \text{Round} \\
\times \text{Provision} + \beta_8 c_{ij,t-1} + \beta_9 c_{ij,t-1} \times \text{Provision} + \varepsilon_{ijt}
\]

(14)

where \(C_{ijt}\) is the effective contribution of individual \(i\) in group \(j\) at round \(t\); \(\beta_0\) is a constant; \(u_{0m}\) and \(u_{0i}\) are random intercepts at the matching group and individual level, respectively. The coefficient \(\beta_1\) estimates the effect of a positive deviation of average effective contribution of the other group members compared to individual \(i\)’s contribution in round \(t - 1\) on \(i\)’s contribution in round \(t\). The coefficient \(\beta_2\) is analogous but for negative deviations. We also include the dummy \text{Provision} and we interact the dummy with the positive and negative deviation variables. We additionally control for the time trend including the variable \text{Round} and for the average contribution of individual \(i\) in round \(t - 1\) (\(c_{ij,t-1}\)) as well as the interaction terms between these variables and \text{Provision}. We report the results from model 14 in the main text and in Supplementary Table 12.
1.2 Robustness Checks

Robustness checks for elicited attitudes (Figure 3a)

To verify that our results on cooperation attitudes are robust, we perform two checks. In the first one, we do not classify participants but simply compare the effective average schedule (the vector $a_i$) between Maintenance and Provision. Recall that for each participant we have a vector $a_i$ comprised of 21 effective contributions, one for each possible rounded average effective contribution of the other group members. We specify elements of the vector $a_i$ as $c_{ik}$, where $k = 1, ..., 21$. We estimate the following linear mixed-effects model:

$$c_{ik} = \beta_0 + u_{0i} + \beta_1 \bar{c}_{-ik} + \beta_2 \text{Provision} + \beta_3 \text{Provision} \times \bar{c}_{-ik} + \epsilon_{ik}$$

where $c_{ik}$ is the contribution of the individual $i$ in entry $k$ of the strategy method table. Our regressors are the average contribution of the others $\bar{c}_{-ik}$ in entry $k$, a treatment dummy Provision, and the interaction term between the average contribution of others and the treatment dummy. We also include a constant $\beta_0$ and a random intercept at the individual level $u_{0i}$.

We report the results of these estimates in Supplementary Table 13. We find a positive and highly significant coefficient $\hat{\beta}_2$, indicating that in Maintenance participants behave on average reciprocally, i.e., they cooperate more the higher the other group members’ effective contributions. We further find a positive and highly significant coefficient $\hat{\beta}_3$ indicating that the reaction to an increase in average contribution of other group members is stronger in Provision than in Maintenance. This confirms the result of higher reciprocity in Provision than in Maintenance.

In our second robustness check, we use hierarchical clustering to classify participants into attitude types. Hierarchical clustering allows to partition the data into subsets, so-called clusters, according to measures of proximity in behavior. The advantage of this method is that it groups data according to their similarity without making any ex-ante assumptions on how behavior looks like. As a measure of proximity between any two effective contribution schedules ($a_i$), we use the ‘city block distance’ measure:

$$d_{ij} = \sum_{k=1}^{21} |x_{ik} - x_{jk}|$$
where $d_{ij}$ is an index of proximity between any two effective contribution schedules, $k$ indexes each entry in the effective contribution schedule, and $x_{ik}$ indicates entry $k$ for individual $i$. Using a different proximity measure (Euclidean distance), does not affect our results.

We then used an agglomerative method to create clusters according to our proximity measure. Agglomerative methods are probably the most widely used type of hierarchical methods. These methods start from $n$ single-observation clusters and merge sequentially clusters until obtaining only one cluster with $n$ observations. In particular, we use Ward’s linkage\textsuperscript{5} method in which the merger of two clusters is based on the minimization of an error term equivalent to the total within-cluster sum of squares, that is:

$$\min E = \sum_{m=1}^{M} E_m$$

where $m = 1, ..., M$ indexes the number of clusters and

$$E_m = \sum_{i=1}^{N} \sum_{k=1}^{21} |x_{ikm} - \bar{x}_{km}|$$

where $i = 1, ..., N$ indexes the number of observations (in our data one observation means one effective contribution schedule, $a_i$); $k$ indexes each entry in the effective contribution schedule; $x_{ikm}$ indicates entry $k$ for individual $i$ in cluster $m$; and $\bar{x}_{km}$ indicates the average entry $k$ for cluster $m$. Clearly, at the start of the routine where we have $n$ single-observation clusters, the error $E$ is equal to zero and it increases as the routine starts merging observations to form clusters. The objective of the method is to merge observations to minimize the increase in $E$.

Finally, we used a formal method to assess the optimal number of clusters to partition our dataset. In particular, we use the Duda and Hart Je(2)/Je(1) and pseudo T-squared indexes, that indicate six clusters as optimal number in our dataset.

To label the six categories we plot the average effective contribution schedule ($a_i$) for each type classified according to the cluster analysis (the figure is available upon request). From visual observation of the average schedule, we label the six groups created in the cluster analysis as strong conditional cooperators, weak conditional cooperators, selfish, altruists, midrange, and triangle contributors. Strong conditional cooperators start out with a contribution of 0.3 when the effective
contribution of others is equal to zero and increase their contribution to 18.3 when other group members’ effective contribution is equal to twenty. Weak conditional cooperators start similarly with a contribution of 0.1 on average and increase to 12.7 when the other group members are fully cooperative. Subjects categorized as selfish contribute very low amounts for the entire average effective contribution schedule with a maximum of 0.25. Altruists are at the other end of the spectrum contributing very high amounts with a minimum of 18.7. Midrange exhibit a slightly decreasing contributing pattern with average effective contributions of 14.6 when the other group members contribute on average zero tokens and 10.5 when the others contribute on average twenty tokens. Finally, triangle contributors are hump-shaped with a contribution of 1.2 when the others contribute zero, a maximum at 7.6 when the others contribute ten tokens, and a contribution of 2.6 when the others are fully cooperative.

The distribution of attitude types classified in the cluster analysis is significantly different across Maintenance and Provision ($\chi^2(5) = 35.31, P < 0.001$). We find significantly less strong and weak conditional cooperators in Maintenance than Provision (32% vs. 39%; $\chi^2(1) = 3.93, P = 0.048$, and 7% vs. 20%; $\chi^2(1) = 18.36, P < 0.001$, respectively). We also find significantly more selfish (30% vs. 22%; $\chi^2(1) = 5.01, P = 0.025$) and midrange (13% vs. 7%; $\chi^2(1) = 11.23, P = 0.001$) in Maintenance than Provision. We find weak and no significant differences for altruists and triangle contributors, respectively (5% vs. 2%; $\chi^2(1) = 2.95, P = 0.086$; and 12% vs. 11%; $\chi^2(1) = 0.34, P = 0.560$). Overall, these results confirms weaker conditional cooperation in Maintenance compared to Provision.

Interestingly, if we compare the distribution of attitudes from our original classification with the one obtained from the cluster analysis, we find that 100% of participants classified as free riders according to the former criterion are classified as selfish in the latter. Furthermore, 92% of participants classified as conditional cooperators in the former are classified as either strong or weak conditional cooperators in the latter. Participants classified as others are mostly classified as triangle contributors or midrange (38% triangle contributors, 32% midrange, 14% selfish, 13% altruists, 2% strong conditional cooperators, and 1% weak conditional cooperators). Overall, this shows high consistency with the classification method used in the main text.
**Robustness checks for estimated reciprocity (Figure 3b)**

To check the robustness of the results of different reciprocity between Maintenance and Provision in Strangers and Partners, we estimate finite mixture models. Following\(^6\), we assume three types: *conditional cooperators* (CC) whose effective contribution depends on the average effective contribution of the other group members in the previous round, *strategic free riders* (STR) who contribute at the beginning but reduce their contributions over time no matter what the other group members do and *free riders* (FR) who contribute zero for all rounds.

We estimate two-limit Tobit models with limits at 0 and 20. The latent variable is the effective contribution of individual \(i\) in round \(t, C^*_it\). For each type, it depends linearly on a set of variables:

**CC:** \[
C^*_it = \beta_0 + \beta_1 \bar{C}_{-i,t-1} + \beta_2 \text{Round} + \varepsilon_{it}
\]

**STR:** \[
C^*_it = \gamma_0 + \gamma_1 \text{Round} + \nu_{it}
\]

Effective contributions of *conditional cooperators* depend positively on the average effective contribution of the other group members in the previous round and negatively on a time trend (we expect \(\beta_1 > 0\) and \(\beta_2 < 0\)). *Strategic free-riders* start with high effective contributions in the first rounds but then lower their effective contributions over time to exploit the other group members (\(\gamma_1 < 0\)). Hence, their behavior depends only on the time trend and not on others’ effective contributions.

The relationship between the latent variable and the observed effective contribution for CC and STR is as follows:

\[
C_{it} = \begin{cases} 
0 & \text{if } C^*_it \leq 0 \\
C^*_it & \text{if } 0 < C^*_it < 20 \\
20 & \text{if } C^*_it \geq 0
\end{cases}
\]

For free riders (FR):

\[
C_{it} = 0 \quad \forall t
\]
To take into account censoring, the maximum likelihood function is the combination of three estimation regimes, depending on the value of $C_{it}$:

**Regime 1 ($C_{it} = 0$):**

$$P(C_{it} = 0|i = CC) = \Phi \left( \frac{-\beta_0 + \beta_1 \bar{C}_{i,t-1} + \beta_2 \text{Round}}{\sigma_1} \right)$$

$$P(C_{it} = 0|i = STR) = \Phi \left( \frac{-\gamma_0 + \gamma_1 \text{Round}}{\sigma_2} \right)$$

$$P(C_{it} = 0|i = FR) = 1$$

**Regime 2 ($0 < C_{it} < 20$):**

$$f(C_{it}|i = CC) = \frac{1}{\sigma_1} \Phi \left( \frac{C_{it} - \beta_0 + \beta_1 \bar{C}_{i,t-1} + \beta_2 \text{Round}}{\sigma_1} \right)$$

$$f(C_{it}|i = STR) = \frac{1}{\sigma_1} \Phi \left( \frac{C_{it} - \gamma_0 + \gamma_1 \text{Round}}{\sigma_2} \right)$$

$$f(C_{it}|i = FR) = 0$$

**Regime 3 ($C_{it} = 20$):**

$$P(C_{it} = 20|i = CC) = 1 - \Phi \left( \frac{20 - \beta_0 + \beta_1 \bar{C}_{i,t-1} + \beta_2 \text{Round}}{\sigma_1} \right)$$

$$P(C_{it} = 20|i = STR) = 1 - \Phi \left( \frac{20 - \gamma_0 + \gamma_1 \text{Round}}{\sigma_2} \right)$$

$$P(C_{it} = 20|i = FR) = 0$$

For subject $i$, the likelihood function is:

$$L_i = p_{CC} \prod_{t=1}^{T} P(C_{it} = 0|CC)^{I_{C_{it}=0}} f(C_{it}|CC)^{I_{0 < C_{it} < 20}} P(C_{it} = 20|CC)^{I_{C_{it}=20}} +$$
\[ p_{STR} \prod_{t=1}^T P(C_{it} = 0 | STR)^{I_{C_{it} = 0}} f(C_{it} | STR)^{I_{0 < C_{it} < 20}} P(C_{it} = 20 | STR)^{I_{C_{it} = 20}} + p_{FR} \prod_{t=1}^T P(C_{it} = 0 | FR)^{I_{C_{it} = 0}} f(C_{it} | FR)^{I_{0 < C_{it} < 20}} P(C_{it} = 20 | FR)^{I_{C_{it} = 20}} \]

where \( I(\cdot) \) is an indicator function, taking value 1 if the subscript is true and 0 otherwise.

In **Supplementary Table 14** we report the maximum likelihood estimations and the resulting estimated mixing proportions of types separately for Strangers and Partners. In both cases, the distribution of types deduced from posterior probabilities is significantly different between Provision and Maintenance (\( \chi^2(2) = 22.05, P < 0.001 \) and \( \chi^2(2) = 16.78, P < 0.001 \) in Strangers and Partners, respectively). In particular, we find significantly more *conditional cooperators* (45% vs. 18%, \( \chi^2(1) = 21.84, P < 0.001 \) and 55% vs. 30%, \( \chi^2(1) = 10.23, P = 0.001 \) in Strangers and Partners, respectively) and significantly less *free riders* in Provision than in Maintenance (22% vs. 36%, \( \chi^2(1) = 6.16, P = 0.013 \) and 5% vs. 25%, \( \chi^2(1) = 12.55 \) and \( P < 0.001 \) in Strangers and Partners, respectively). In Strangers, we also find significantly less *strategic free-riders* in Provision than in Maintenance (33% vs. 46%, \( \chi^2(1) = 0.48, P = 0.029 \), while this is not the case in Partners (40% vs. 45%, \( \chi^2(1) = 0.41, P = 0.552 \)).
1.3 Simulation Analysis

Each simulated contribution is derived by matching one randomly drawn attitude and one randomly drawn belief from our sample. Each simulated contribution $\tilde{c}^F$ is therefore given by:

$$\tilde{c}^F = a_i^F(b_j^F)$$  \hfill (16)

where the superscript $F$ indicates the sample ($P$ for Provision and $M$ for Maintenance) from which each component is randomly drawn.

Our procedure comprises the following steps which are also summarized in the diagram below:

a. Fix a sample size $n$ (observations per game).

b. Set $F = P$ for $a_i^P$ and $b_j^P$, and randomly draw (with replacement) one $a_i^P$ and one $b_j^P$ from the Provision distribution of attitudes and beliefs, respectively. Use equation (16) to calculate $\tilde{c}^P$. Repeat this step until we have $n$ simulated contributions.

c. Redo step b. setting $F = M$ for all components, i.e., $a_i^M$ and $b_j^M$. Use equation (16) to calculate $\tilde{c}^M$. Repeat this step until we have $n$ simulated contributions.

d. Compare the two samples of size $n$ derived from b. (Provision) and c. (Maintenance) by calculating the ratio of average effective contributions between Maintenance and Provision.

e. Repeat steps b. - d. 1000 times.

Figure 4 in the main text reports simulation results of 1000 random samples of size $n = 60$, the median sample size in previous related literature (see refs. 14-23 in the main text). As a robustness check, we also ran a simulation with a sample size of $n = 100$. The results from this simulation are shown in **Supplementary Figure 4**.
1.4 Experimental Instructions

Here, we document the experimental instructions we used in the experiments. We document the exact texts used in Provision and show the changed texts used in Maintenance in [italics].

**Part 1 – Introduction to Provision [Maintenance]**

Instructions

You are participating in a study in which you will earn some money. The amount will depend on the outcome of a game you will play. The amount of money which you earned with your decisions will be paid to you in cash at the end of the experiment. We will not speak of Pounds during the experiment, but rather of points. At the end, the total number of points you have earned will be converted to Pounds at the following rate:

1 point = £0.2

These instructions are solely for your private information. You are not allowed to communicate during the experiment. If you have any questions, please raise your hand. A member of the experimental team will come to you and answer them in private.

All participants will be divided into groups of four members. Only the experimenters will know who is in which group.

**The decision situation**

We first introduce you to the basic decision situation. Then, you will complete a pre-study questionnaire on the screen in front of you, which is intended to help you understand the decision situation.

In each group, every member has to decide the allocation of 20 tokens. You can put these 20 tokens into your private account or you can put some or all of them into a project. [In each group, there are 80 tokens in a project. You can withdraw up to 20 tokens from the project and put them into your private account or you can leave them fully or partially in the project.] The other three members of your group have to make the same decision.

**Your income from the private account**

You will earn 1 point for each token you put into your private account. For example, if you put all 20 tokens into your private account, your income from your private account would be 20 points. If you put 6 tokens into your private account, your income from this account would be 6 points. No one except you earns anything from tokens you put in your private account.

**Your income from the project**

Each group member will profit equally from the amount you or any other group member put into [leave in] the project. The income for each group member from the project will be determined as follows:
If, for example, the sum of all contributions to the project \( \text{tokens withdrawn from the project} \) by you and your other group members is 60 \( \text{[20]} \text{tokens} \), then you and each other member of your group would earn 60 \( \text{[80-20]} \times 0.4 = 24 \) points out of the project. If the four members of the group contribute \( \text{withdraw} \) a total of 10 \( \text{[70]} \text{tokens} \) to \( \text{from} \) the project, you and the other members of your group would each earn 10 \( \text{[80-70]} \times 0.4 = 4 \) points.

**Total income**

Your total income is the sum of your income from your private account and from the project:

\[
\text{Your Total Income} = \text{Income from your private account} + \text{Income from the project}
\]

\[
geq 20 - \text{your contribution to the project} + 0.4 \times \text{sum of all contributions to the project}
\]

\[
\text{[= Tokens withdrawn from the project by you} + 0.4 \times (80-\text{sum of all tokens withdrawn from the project})]
\]

**Comprehension test**

Please answer all the following questions, to help you understand the determination of your income.

1. Each group member has 20 tokens. Assume that none of the four group members (including you) contributes anything to the project. \( \text{[There are 80 tokens in the project. Assume that everyone in your group withdraws 20 tokens from the project.]} \)
   a) What will your total income (in points) be?
   b) What will the total income (in points) of each of the other group members be?

2. Each group member has 20 tokens. You contribute 20 tokens in the project. Each of the other three members of the group also contributes 20 tokens to the project. \( \text{[There are 80 tokens in the project. You withdraw 0 tokens from the project. Each of the other three members of the group also withdraws 0 tokens from the project.]} \)
   a) What will your total income (in points) be?
   b) What will the total income (in points) of each of the other group members be?

3. Each group member has 20 tokens. The other three members contribute a total of 30 tokens to the project. \( \text{[There are 80 tokens in the project. The other three members withdraw 30 tokens from the project.]} \)
Part 2 – Strategy method experiment (elicitation of attitudes)

The Experiment

The experiment is based on the decision situation just described to you, conducted once. You will enter your decisions in the screen in front of you.

As you know, you will have 20 tokens at your disposal. You can put them into a private account or into a project. [As you know, there are 80 tokens in a project. You can withdraw tokens from the project which will be automatically placed into your private account or you can leave them in the project.] Each subject has to make two types of decisions in this experiment, which we will refer to below as the “unconditional contribution [withdrawal]” and the “contribution [withdrawal] table”.

- In the unconditional contribution [withdrawal] you simply decide how many of the 20 [80] tokens you want to put in [withdraw from] the project. Please indicate your contribution [withdrawal] in the following screen (screenshot taken from the Provision treatment only):
After you have determined your unconditional contribution \([\text{withdrawal}]\), please click “OK”.

- Your second task is to fill in a “contribution \([\text{withdrawal}]\) table” where you indicate how many tokens you want to contribute \([\text{withdraw}]\) to \([\text{from}]\) the project for each possible average contribution \([\text{withdrawal}]\) of the other group members (rounded to the next integer). Here, you can condition your contribution \([\text{withdrawal}]\) on that of the other group members. This will be immediately clear to you if you take a look at the following table.

This table will be presented to you in the experiment (screenshot taken from the Provision treatment only):
The numbers to the left of the blue cells are the possible (rounded) average contributions \([\text{withdrawals}]\) of the \textbf{other} group members to the project. You have to insert how many tokens you want to contribute to \([\text{withdraw from}]\) the project into each input box – conditional on the indicated average contribution \([\text{withdrawal}]\) by the other members of your group. \textbf{You must enter a number between 0 and 20 inclusive in each input box.} For example, you have to indicate how much you contribute to \([\text{withdraw from}]\) the project if the others contribute \([\text{withdraw}]\) 0 tokens on average to \([\text{from}]\) the project; how much you contribute \([\text{withdraw}]\) if the others contribute \([\text{withdraw}]\) 1, 2, or 3 tokens on average; etc. Once you have made an entry in each input box, click “OK”.

After all participants of the experiment have made an unconditional contribution \([\text{withdrawal}]\) and have filled in their contribution \([\text{withdrawal}]\) table, a random mechanism will select one member from every group. For \textbf{this} group member, it is his \textbf{contribution} \([\text{withdrawal}]\) table that will determine his actual contribution \([\text{withdrawal}]\); whereas, for the \textbf{other three} group members, it is their \textbf{unconditional contributions} \([\text{withdrawals}]\) that will determine their actual contributions \([\text{withdrawals}]\). You will not know whom the random mechanism will select when you make your unconditional contribution \([\text{withdrawal}]\) and fill in your contribution \([\text{withdrawal}]\) table. You must therefore think carefully about both decisions because either could determine your actual contribution \([\text{withdrawal}]\). Two examples should make this clear.

**EXAMPLE 1:** Suppose that \textbf{the random mechanism selects you}; and that the other three group members made unconditional contributions \([\text{withdrawals}]\) of 0, 2, and 4 \([20, 18, \text{and } 16]\) tokens, respectively. The average contribution \([\text{withdrawal}]\) of these three group members is, therefore, 2 \([18]\) tokens. If you indicated in your contribution \([\text{withdrawal}]\) table that you will contribute \([\text{withdraw}]\) 1 \([19]\) token[s] if the others contribute \([\text{withdraw}]\) 2 \([18]\) tokens on average, then the total contribution to the project is given by \(0+2+4+1=7\) \textit{[the total number of tokens left in the project is given by \(80-(20+18+16+19)=7\) tokens]. Each group member would, therefore, earn \(0.4\times7=2.8\) points from the project plus their respective income from their own private account. If, instead, you indicated in your contribution \([\text{withdrawal}]\) table that you would contribute \([\text{withdraw}]\) 19 tokens \([1\text{ token}]\) if the others contribute \([\text{withdraw}]\) 2 \([18]\) tokens on average, then the total contribution of the group to the project would be given by \(0+2+4+19=25\) \textit{[the total number of tokens left in the project would be given by \(80-(20+18+16+1)=25\) tokens}. Each group member would earn \(0.4\times25=10\) points from the project plus their respective income from their own private account.

**EXAMPLE 2:** Suppose that \textbf{the random mechanism does not select you}; and that your unconditional \([\text{withdrawal}]\) contribution is 16 \([4]\) tokens, while those of the other two group members not selected by the random mechanism are 18 \([2]\) and 20 \([0]\) tokens, respectively. Your average unconditional contribution \([\text{withdrawal}]\) and that of these two other group members is, therefore, 18 \([2]\) tokens. If the group member whom the random mechanism did select indicates in her contribution \([\text{withdrawal}]\) table that she will contribute \([\text{withdraw}]\) 1 \([19]\) token[s] if the other three group members contribute \([\text{withdraw}]\) on average 18 \([2]\) tokens, then the total contribution of the group to the project is given by \(16+18+20+1=55\) \textit{[the total number of tokens left in the project is given by \(80-(4+2+0+19)=55\) tokens]. Each group member will therefore earn \(0.4\times55=22\) points from the project plus their respective income from their own private account. If, instead, the randomly selected group member indicates in her contribution \([\text{withdrawal}]\) table that she contributes \([\text{withdraws}]\) 19 \([1]\) if the others contribute \([\text{withdraw}]\) on average 18 \([2]\) tokens, then the total contribution of the group to the project is \(16+18+20+19=73\) \textit{[the total number of tokens left in the project is \(80-(4+2+0+1)=73\) tokens]. Each group member would therefore earn \(0.4\times73=29.2\) points from the project plus their respective income from their own private account. The random selection of the group member whose contribution \([\text{withdrawal}]\) table will determine his actual contribution \([\text{withdrawal}]\) will be made as follows. Each group member is assigned a Group
Member ID between 1 and 4, which denote his/her number inside his group. Moreover, one participant was randomly selected at the very beginning of the experiment. This participant will draw a ball from an urn after all participants have made their unconditional contribution [withdrawal] and have filled out their contribution [withdrawal] table. Each ball in the urn has a different colour and each colour corresponds to a Group Member ID: orange=1, blue=2, yellow=3, green=4. The resulting number will be entered into the computer. If the randomly selected participant draws the Group Member ID that was assigned to you, then your contribution [withdrawal] table will determine your contribution [withdrawal] and their unconditional contributions [withdrawals] will determine the contribution [withdrawals] of the other group members. Otherwise, your unconditional [withdrawal] contribution determines your contribution [withdrawal].
Part 3 – Direct-response experiments

1) One-shot

Instructions

You are now taking part in a second experiment. The money you earn in this experiment will be added to what you earned in the first one. As before, we will not speak of Pounds during the experiment, but rather of points. At the end, the number of points you have earned will be converted to Pounds at the following rate:

1 point = £0.2

As in the previous experiment you are in a group composed by 4 people. However, the composition of the group is entirely new. None of the participants who were in your group in the second experiment will be in your group in this experiment.

The decision situation is the same as the one described on the first instruction sheet of the previous experiment. Each member of the group has to decide about the usage of the 20 tokens. [In each group there are 80 tokens in a project.] You can put these 20 tokens into your private account or you can put them fully or partially into a project. [You can withdraw up to 20 tokens from the project or you can leave them fully or partially in the project.] Each token you do not put into the project [withdraw from the project] is automatically placed into your private account. Your income will be determined in the same way as before.

Reminder:

\[
\text{Your Total Income} = \text{Income from your private account} + \text{Income from the project} \\
= 20 - \text{your contribution to the project} + 0.4 \times \text{sum of all contributions to the project} \\
\text{[=Tokens withdrawn from the project by you } + 0.4 \times (80-\text{sum of all tokens withdrawn from the project})]
\]

The decision screen looks like this (screenshot taken from the Provision treatment only):
1. First you have to decide on your contribution to withdrawal from the project, that is, you have to decide how many of the 20 tokens you want to contribute to the project, and how many tokens you want to put into your private account. Each other member of your group has to make the corresponding decision. This is the only contribution withdrawal decision that you or they make in this experiment. There is no contribution withdrawal table.

2. Afterwards you have to estimate the average contribution to withdrawal from the project (rounded to an integer) of the other three group members. You will be paid for the accuracy of your estimate:
   - If your estimate is exactly right (that is, if your estimate is exactly the same as the actual average contribution withdrawal of the other group members), you will get 3 points in addition to your other income from the experiment.
   - If your estimate deviates by one point from the correct result, you will get 2 additional points.
   - A deviation by 2 points still earns you 1 additional point.
   - If your estimate deviates by 3 or more points from the correct result, you will not get any additional points.
2) *Strangers*

**Instructions**

You are now taking part in a second experiment. The money you earn in this experiment will be added to what you earned in the first one. As before, we will not speak of Pounds during the experiment, but rather of points. At the end, the number of points you have earned will be converted to Pounds at the following rate:

1 point=£0.2

This experiment lasts **several rounds**, in which you and the other group members have to make decisions. You will not know how many rounds the experiment will last and will be told when the experiment is finished. As in the previous experiment, every group consists of **4 people**. The formation of the group changes at random after every round. **So your group will typically consist of different people every round.**

The decision situation is the same as the one described on the first instruction sheet of the previous experiment. Each member of the group has to decide about the usage of the 20 tokens. *[In each group there are 80 tokens in a project.]* You can put these 20 tokens into your private account or you can put them fully or partially into a project. *[You can withdraw up to 20 tokens from the project or you can leave them fully or partially in the project.]* Each token you do not put into the project [withdraw from the project] is automatically placed into your private account. Your income will be determined in the same way as before.

Reminder:

\[
\text{Your Total Income} = \text{Income from your private account} + \text{Income from the project} \\
= 20 - \text{your contribution to the project} + 0.4 \times \text{sum of all contributions to the project} \\
[= \text{Tokens withdrawn from the project by you} + 0.4 \times (80 - \text{sum of all tokens withdrawn from the project})]
\]

The decision screen looks like this *(screenshot taken from the Provision treatment only)*:
1. First you have to **decide on your contribution to [withdrawal from] the project**, that is, you have to decide how many of the 20 tokens you want to contribute to the project, and how many tokens you want to put into your private account. [you have to decide how many of the 80 tokens you want to withdraw from the project and put into your private account.] Each other member of your group has to make the corresponding decision. This is the only contribution [withdrawal] decision that you or they make in this experiment. There is no contribution [withdrawal] table.

2. Afterwards you have to estimate the average contribution to [withdrawal from] the project (rounded to an integer) of the other three group members. You will be paid for the accuracy of your estimate:
   - If your estimate is exactly right (that is, if your estimate is **exactly** the same as the actual average contribution [withdrawal] of the other group members), you will get **3 points** in addition to your other income from the experiment.
   - If your estimate deviates by one point from the correct result, you will get 2 additional points.
   - A deviation by 2 points still earns you 1 additional point.
   - If your estimate deviates by 3 or more points from the correct result, you will not get any additional points.

3. You will receive information about the outcome at the end of each round.
3) **Partners**

**Instructions**

You are now taking part in a second experiment. The money you earn in this experiment will be added to what you earned in the first one. As before, we will not speak of Pounds during the experiment, but rather of points. At the end, the number of points you have earned will be converted to Pounds at the following rate:

1 point = £0.2

This experiment lasts **several rounds**, in which you and the other group members have to make decisions. You will not know how many rounds the experiment will last and will be told when the experiment is finished.

As in the previous experiment, every group consists of **4 people**. However, the composition of the group is entirely new. None of the participants who were in your group in the first experiment will be in your group in this experiment. You and the other three group members will remain in this same group throughout the entire experiment. So your group will consist of the same people every round.

The decision situation is the same as the one described on the first instruction sheet of the previous experiment. Each member of the group has to decide about the usage of the 20 tokens. **[In each group there are 80 tokens in a project.]** You can put these 20 tokens into your private account or you can put them fully or partially into a project. **[You can withdraw up to 20 tokens from the project or you can leave them fully or partially in the project.]** Each token you do not put into the project (**withdraw from the project**) is automatically placed into your private account. Your income will be determined in the same way as before.

Reminder:

<table>
<thead>
<tr>
<th>Your Total Income</th>
<th>= Income from your private account + Income from the project</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>= 20 – your contribution to the project + 0.4 × sum of all contributions to the project</td>
</tr>
<tr>
<td></td>
<td>[=Tokens withdrawn from the project by you + 0.4 × (80-sum of all tokens withdrawn from the project)]</td>
</tr>
</tbody>
</table>

The decision screen looks like this (**screenshot taken from the Provision treatment only**):
1. First you have to **decide on your contribution to [withdrawal from] the project**, that is, you have to decide how many of the 20 tokens you want to contribute to the project, and how many tokens you want to put into your private account. [you have to decide how many of the 80 tokens you want to withdraw from the project and put into your private account.] Each other member of your group has to make the corresponding decision. This is the only contribution [withdrawal] decision that you or they make in this experiment. There is **no contribution [withdrawal] table**.

2. Afterwards you have to estimate the average contribution to [withdrawal from] the project (rounded to an integer) of the other three group members. You will be paid for the accuracy of your estimate:
   - If your estimate is exactly right (that is, if your estimate is exactly the same as the actual average contribution [withdrawal] of the other group members), you will get **3 points** in addition to your other income from the experiment.
   - If your estimate deviates by one point from the correct result, you will get 2 additional points.
   - A deviation by 2 points still earns you 1 additional point.
   - If your estimate deviates by 3 or more points from the correct result, you will not get any additional points.

3. You will receive information about the outcome at the end of each round.
4) **Partners with Punishment**

**Instructions**

You are now taking part in a second experiment. The money you earn in this experiment will be added to what you earned in the first one. As before, we will not speak of Pounds during the experiment, but rather of points. At the end, the number of points you have earned will be converted to Pounds at the following rate:

1 point = £0.02

This experiment lasts **several rounds**, in which you and the other group members have to make decisions. You will not know how many rounds the experiment will last and will be told when the experiment is finished.

As in the previous experiment, every group consists of **4 people**. You and the other three group members will **remain in this same group throughout the entire experiment. So your group will consist of the same people every round.**

The decision situation is the same as the one described on the first instruction sheet of the previous experiment. Each member of the group has to decide about the usage of the 20 tokens. **[In each group there are 80 tokens in a project.]** You can put these 20 tokens into your private account or you can put them fully or partially into a project. **[You can withdraw up to 20 tokens from the project or you can leave them fully or partially in the project.]** Each token you do not put into the project [withdraw from the project] is automatically placed into your private account. Your income will be determined in the same way as before.

**Reminder:**

\[
\text{Your Total Income} = \text{Income from your private account} + \text{Income from the project} \\
= 20 - \text{your contribution to the project} + 0.4 \times \text{sum of all contributions to the project} \\
[=\text{Tokens withdrawn from the project by you} + 0.4 \times (80 - \text{sum of all tokens withdrawn from the project})]
\]

**The Experiment**

Each round consists of **two stages**. In the first stage you will be endowed with tokens and have to decide how many tokens you would like to contribute to a project. **[In the first stage a project is endowed with tokens and you have to decide how many tokens you would like to withdraw from the project.]** In the second stage you will be informed about the contributions [withdrawals] of the other three group members. You will then decide whether or how much to reduce their earnings from the first stage by distributing points to them.

**STAGE 1**

The decision screen looks like this (**screenshot taken from the Provision treatment only**):
2. First you have to decide on your contribution to withdrawal from the project, that is, you have to decide how many of the 20 tokens you want to contribute to the project, and how many tokens you want to put into your private account. You have to decide how many of the 80 tokens you want to withdraw from the project and put into your private account. Each other member of your group has to make the same decision.

3. Afterwards you have to estimate the average contribution to withdrawal from the project (rounded to an integer) of the other three group members. You will be paid for the accuracy of your estimate:
   - If your estimate is exactly right (that is, if your estimate is exactly the same as the actual average contribution withdrawal of the other group members), you will get 3 points in addition to your other income from the experiment.
   - If your estimate deviates by one point from the correct result, you will get 2 additional points.
   - A deviation by 2 points still earns you 1 additional point.
   - If your estimate deviates by 3 or more points from the correct result, you will not get any additional points.

After that the first stage is over and the second stage begins.

**STAGE 2**

In the second stage you will learn your income from the first stage and you will see how much each group member contributed to withdrew from the project. Moreover, in this stage you can decide whether to decrease the income of each other group member by assigning deduction points. The other group members can also decrease your income if they wish to. This is apparent from the input screen of the second stage displayed below:
Your income and your contribution [withdrawal] from the first stage are displayed in the first two rows. The contributions [withdrawal] of the other group members are shown in the three columns below. Note that the order in which others’ contributions [withdrawals] are displayed will be determined at random in every round. The contribution [withdrawal] in the first column, for example, could represent a different group member in different rounds. The same holds true for the second and third column.

You will have to decide how many deduction points to assign to each of the other three group members. You must enter a number for each of them. If you do not wish to change the income of a specific group member then you must enter 0. You can assign up to 5 points to each group member.

You will incur costs from assigning deduction points. Every deduction point you assign costs you 1 point. For example, suppose you assign 2 deduction points to one member, this costs you 2 points; if, in addition, you assign 4 deduction points to another member this costs you an additional 4 points. Suppose further that you assign 0 deduction points to the third member. In total you will have assigned 6 points and your total costs therefore amount to 6 points.

If you assign 0 deduction points to a particular group member (i.e., enter “0”), you will not alter his or her income. However, if you assign one deduction point to a group member you will decrease the income of this group member by 3 points. If you assign a group member 2 deduction points you will decrease the group member’s income by 6 points, and so on. Each deduction point that you assign to another group member will reduce his or her income by 3 points. Similarly, each deduction point assigned to you by another group member will reduce your first stage income by three points:

$$\text{Costs of received deduction points} = 3 \times \text{Sum of received deduction points}$$
How much the income at the second stage is decreased depends on the sum of deduction points received. For instance, if somebody receives a total of 3 deduction points (from all other group members), his or her income would be decreased by 9 points. If somebody receives a total of 4 deduction points, his or her income is reduced by 12 points.

There is one exception to this rule. If the cost of received deduction points exceeds the group member’s first stage income, his or her first stage income will be reduced to zero. However, even in this case the group member must still incur the costs of any deduction points he or she assigned.

For each round, your total income from the two stages is therefore calculated as follows:

<table>
<thead>
<tr>
<th><strong>EITHER</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Your income from the first stage is greater than or equal to the cost of received deduction points:</td>
</tr>
<tr>
<td><strong>Total income =</strong> Income from the first stage $- 3 \times$ (sum of received deduction points) $- \text{sum of deduction points you have assigned}$</td>
</tr>
<tr>
<td>OR</td>
</tr>
<tr>
<td>Your income from the first stage is less than the cost of received deduction points:</td>
</tr>
<tr>
<td><strong>Total income =</strong> 0 $- \text{sum of deduction points you have assigned}$</td>
</tr>
</tbody>
</table>

Please note that your income in points at the end of the second stage can be negative if the costs of your assigned points exceed your income from the first stage minus the income reduction by the received deduction points. You can, however, avoid such losses with certainty through your own decisions!

After all participants have made their decision, the results from the round including your final income from that round will be displayed. After you have viewed the income screen the round is over and the next round begins.
2. Supplementary References


3. Supplementary Tables

Supplementary Table 1 (support for Figure 2, all panels)

<table>
<thead>
<tr>
<th>Panel A: Public Good Size</th>
<th>One-shot</th>
<th>Strangers</th>
<th>Partners</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>P</td>
<td>M</td>
</tr>
<tr>
<td>Round 1</td>
<td>23.83(16.00)</td>
<td>32.65(13.81)</td>
<td>22.69(13.48)</td>
</tr>
<tr>
<td></td>
<td>P = 0.014</td>
<td></td>
<td>P = 0.059</td>
</tr>
<tr>
<td>Round 1-9</td>
<td>10.39(9.61)</td>
<td>15.63(10.73)</td>
<td>14.83(14.85)</td>
</tr>
<tr>
<td></td>
<td>P = 0.053</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Round 10-18</td>
<td>5.65(9.06)</td>
<td>5.21(6.49)</td>
<td>8.18(13.17)</td>
</tr>
<tr>
<td></td>
<td>P = 0.831</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Round 19-27</td>
<td>4.88(8.31)</td>
<td>3.77(6.02)</td>
<td>6.97(14.57)</td>
</tr>
<tr>
<td></td>
<td>P = 0.541</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All rounds</td>
<td>23.83(16.00)</td>
<td>32.65(13.81)</td>
<td>6.97(8.99)</td>
</tr>
<tr>
<td></td>
<td>P = 0.014</td>
<td></td>
<td>P = 0.503</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Beliefs</th>
<th>One-shot</th>
<th>Strangers</th>
<th>Partners</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>P</td>
<td>M</td>
</tr>
<tr>
<td>Round 1</td>
<td>7.76(5.82)</td>
<td>9.62(5.32)</td>
<td>6.70(5.92)</td>
</tr>
<tr>
<td></td>
<td>P = 0.005</td>
<td></td>
<td>P = 0.006</td>
</tr>
<tr>
<td>Round 1-9</td>
<td>3.17(3.13)</td>
<td>5.66(4.07)</td>
<td>4.13(4.51)</td>
</tr>
<tr>
<td></td>
<td>P = 0.002</td>
<td></td>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td>Round 10-18</td>
<td>1.04(2.05)</td>
<td>1.72(2.02)</td>
<td>1.98(3.36)</td>
</tr>
<tr>
<td></td>
<td>P = 0.232</td>
<td></td>
<td>P = 0.059</td>
</tr>
<tr>
<td>Round 19-27</td>
<td>0.82(1.83)</td>
<td>1.02(1.82)</td>
<td>1.94(3.46)</td>
</tr>
<tr>
<td></td>
<td>P = 0.632</td>
<td></td>
<td>P = 0.079</td>
</tr>
<tr>
<td>All rounds</td>
<td>7.76(5.82)</td>
<td>9.62(5.32)</td>
<td>1.68(2.34)</td>
</tr>
<tr>
<td></td>
<td>P = 0.005</td>
<td></td>
<td>P = 0.030</td>
</tr>
</tbody>
</table>

Supplementary Table 1 | Descriptive statistics and tests on the size of the public good (Panel A) and on beliefs (Panel B). Panel A: shown is the average (std. dev.) size of the public good in all conditions. Standard deviations are calculated using differences across groups within a period, averaged across periods. Panel B: shown are the participants’ average (std. dev.) beliefs about the average effective contributions of the other group members. Standard deviations are calculated using differences across individuals within a period, averaged across periods. M = Maintenance, P = Provision. P values are based on two-sided t-tests for One-shot and the first round of Strangers and Partners. All the remaining P values are from linear mixed-effects models with random intercepts at the matching group level. Further details are in the Supplementary Analysis.
Supplementary Table 2 (support for Figure 2, Panels 2-3)

<table>
<thead>
<tr>
<th></th>
<th>Maintenance vs. Provision</th>
<th>Strangers vs. Partners</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>Strangers</td>
<td>Partners</td>
</tr>
<tr>
<td>Provision</td>
<td>1.230</td>
<td>6.309**</td>
</tr>
<tr>
<td>1 if Provision, 0 otherwise</td>
<td>(1.837)</td>
<td>(2.987)</td>
</tr>
<tr>
<td>Partners</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 if Partners, 0 otherwise</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>6.973***</td>
<td>10.56***</td>
</tr>
<tr>
<td></td>
<td>(1.299)</td>
<td>(2.112)</td>
</tr>
<tr>
<td>N</td>
<td>432</td>
<td>1080</td>
</tr>
</tbody>
</table>

Supplementary Table 2 | Comparing public good size between Maintenance and Provision, and between Strangers and Partners. Shown are estimated fixed effects from linear mixed-effects models (details on the estimation can be found in section 1.1 - Equations 1-2). Dependent variable: public good size. Random intercepts are included for matching groups. Models (1) and (2) compare the public good size between Maintenance and Provision separately for Strangers and Partners. Models (3) and (4) compare the public good size between Strangers and Partners separately for Maintenance and Provision. Standard errors in parentheses. *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$. 
Supplementary Table 3 (support for Figure 3b, all panels)

<table>
<thead>
<tr>
<th>Panel A: Strangers</th>
<th>(1) All</th>
<th>(2) CC</th>
<th>(3) FR</th>
<th>(4) OT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provision</td>
<td>-0.019</td>
<td>0.302</td>
<td>-0.706**</td>
<td>-0.197</td>
</tr>
<tr>
<td>1 if Provision, 0 otherwise</td>
<td>(0.383)</td>
<td>(0.539)</td>
<td>(0.283)</td>
<td>(1.079)</td>
</tr>
<tr>
<td>$\bar{C}_{ij,t-1}$</td>
<td>0.172***</td>
<td>0.254***</td>
<td>0.038*</td>
<td>0.222***</td>
</tr>
<tr>
<td>Provision $\times$ $\bar{C}_{ij,t-1}$</td>
<td>(0.020)</td>
<td>(0.033)</td>
<td>(0.021)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>Round</td>
<td>-0.044***</td>
<td>-0.069***</td>
<td>-0.003</td>
<td>-0.056***</td>
</tr>
<tr>
<td>Provision $\times$ Round</td>
<td>(0.007)</td>
<td>(0.011)</td>
<td>(0.007)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>Constant</td>
<td>1.920***</td>
<td>2.191***</td>
<td>0.138</td>
<td>3.589***</td>
</tr>
<tr>
<td>$N$</td>
<td>6656</td>
<td>3172</td>
<td>1898</td>
<td>1586</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Partners</th>
<th>(1) All</th>
<th>(2) CC</th>
<th>(3) FR</th>
<th>(4) OT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provision</td>
<td>-0.407</td>
<td>0.171</td>
<td>-0.623</td>
<td>-1.576</td>
</tr>
<tr>
<td>1 if Provision, 0 otherwise</td>
<td>(0.549)</td>
<td>(0.636)</td>
<td>(1.212)</td>
<td>(1.498)</td>
</tr>
<tr>
<td>$\bar{C}_{ij,t-1}$</td>
<td>0.420***</td>
<td>0.621***</td>
<td>0.292***</td>
<td>0.202***</td>
</tr>
<tr>
<td>Provision $\times$ $\bar{C}_{ij,t-1}$</td>
<td>(0.030)</td>
<td>(0.043)</td>
<td>(0.048)</td>
<td>(0.075)</td>
</tr>
<tr>
<td>Round</td>
<td>-0.027**</td>
<td>-0.001</td>
<td>-0.024</td>
<td>-0.103***</td>
</tr>
<tr>
<td>Provision $\times$ Round</td>
<td>(0.012)</td>
<td>(0.016)</td>
<td>(0.020)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Constant</td>
<td>1.741***</td>
<td>1.188**</td>
<td>0.919</td>
<td>3.804***</td>
</tr>
<tr>
<td>$N$</td>
<td>4160</td>
<td>2366</td>
<td>884</td>
<td>910</td>
</tr>
</tbody>
</table>

Supplementary Table 3 | Comparing estimated reciprocity between Maintenance and Provision. Shown are estimated fixed effects from linear mixed-effects models (details on the estimation can be found in section 1.1 - Equation 4). **Panel A** reports estimates for Strangers, **Panel B** for Partners. Dependent variable: effective contributions. $\bar{C}_{ij,t-1}$ is the average effective contribution of the other three group members from the previous round. Random intercepts are included for matching groups and individuals. Model (1) is estimated using the entire sample. Models (2-4) use only the subset of participants classified as **conditional cooperators** (CC), **free riders** (FR), and **others** (OT), respectively. Standard errors in parentheses. *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$. 
### Supplementary Table 4 (support for Figure 3b, all panels)

<table>
<thead>
<tr>
<th></th>
<th>(1) All Maintenance</th>
<th>(2) All Provision</th>
<th>(3) CC Maintenance</th>
<th>(4) CC Provision</th>
<th>(5) FR Maintenance</th>
<th>(6) FR Provision</th>
<th>(7) OT Maintenance</th>
<th>(8) OT Provision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partners</td>
<td>-0.131</td>
<td>-0.557</td>
<td>-0.938</td>
<td>-1.120*</td>
<td>0.780</td>
<td>0.931</td>
<td>0.221</td>
<td>-1.165</td>
</tr>
<tr>
<td>1 if Partners, 0 otherwise</td>
<td>(0.514)</td>
<td>(0.427)</td>
<td>(0.700)</td>
<td>(0.487)</td>
<td>(0.487)</td>
<td>(1.214)</td>
<td>(1.326)</td>
<td>(1.192)</td>
</tr>
<tr>
<td>( \tilde{C}_{i,j,t-1} )</td>
<td>0.170***</td>
<td>0.430***</td>
<td>0.252***</td>
<td>0.484***</td>
<td>0.040</td>
<td>0.235***</td>
<td>0.222***</td>
<td>0.460***</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.025)</td>
<td>(0.035)</td>
<td>(0.032)</td>
<td>(0.026)</td>
<td>(0.044)</td>
<td>(0.052)</td>
<td>(0.060)</td>
</tr>
<tr>
<td>Partners × ( \tilde{C}_{i,j,t-1} )</td>
<td>0.237***</td>
<td>0.206***</td>
<td>0.348***</td>
<td>0.233***</td>
<td>0.254***</td>
<td>0.062</td>
<td>-0.022</td>
<td>0.081</td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.033)</td>
<td>(0.051)</td>
<td>(0.041)</td>
<td>(0.039)</td>
<td>(0.077)</td>
<td>(0.086)</td>
<td>(0.079)</td>
</tr>
<tr>
<td>Round</td>
<td>-0.044***</td>
<td>-0.066***</td>
<td>-0.069***</td>
<td>-0.086***</td>
<td>-0.003</td>
<td>0.043***</td>
<td>-0.056***</td>
<td>-0.149***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.009)</td>
<td>(0.012)</td>
<td>(0.012)</td>
<td>(0.009)</td>
<td>(0.016)</td>
<td>(0.018)</td>
<td>(0.022)</td>
</tr>
<tr>
<td>Partners × Round</td>
<td>0.016</td>
<td>0.063***</td>
<td>0.066***</td>
<td>0.061***</td>
<td>-0.020</td>
<td>0.004</td>
<td>-0.047</td>
<td>0.168***</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.014)</td>
<td>(0.018)</td>
<td>(0.017)</td>
<td>(0.015)</td>
<td>(0.033)</td>
<td>(0.031)</td>
<td>(0.032)</td>
</tr>
<tr>
<td>Constant</td>
<td>1.925***</td>
<td>1.896***</td>
<td>2.195***</td>
<td>2.483***</td>
<td>0.126</td>
<td>-0.725</td>
<td>3.588***</td>
<td>3.391***</td>
</tr>
<tr>
<td></td>
<td>(0.338)</td>
<td>(0.266)</td>
<td>(0.479)</td>
<td>(0.328)</td>
<td>(0.339)</td>
<td>(0.742)</td>
<td>(0.751)</td>
<td>(0.774)</td>
</tr>
</tbody>
</table>

**N** 5408 5408 2288 3250 1742 1040 1378 1118

**Supplementary Table 4 | Comparing estimated reciprocity between Strangers and Partners.** Shown are estimated fixed effects from linear mixed-effects models (details on the estimation can be found in section 1.1 - Equation 5). Dependent variable: effective contributions. \( \tilde{C}_{i,j,t-1} \) is the average effective contribution of the other three group members from the previous round. Random intercepts are included for matching groups and individuals. Models (1-2) are estimated using the entire sample. Models (3-4), (5-6), and (7-8) use only the subset of participants classified as conditional cooperators (CC), free riders (FR), and others (OT), respectively. Standard errors in parentheses. *** \( P < 0.01 \), ** \( P < 0.05 \), * \( P < 0.1 \).
Supplementary Table 5 (support for Figure 3b, Panels 2-4)

<table>
<thead>
<tr>
<th></th>
<th>(1) Strangers Maintenance</th>
<th>(2) Strangers Provision</th>
<th>(3) Partners Maintenance</th>
<th>(4) Partners Provision</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC 1 if CC, 0 otherwise</td>
<td>2.076***</td>
<td>3.003***</td>
<td>0.231</td>
<td>1.077</td>
</tr>
<tr>
<td>OT 1 if OT, 0 otherwise</td>
<td>3.450***</td>
<td>3.924***</td>
<td>2.955***</td>
<td>1.969</td>
</tr>
<tr>
<td>$\hat{C}_{-ij,t-1}$</td>
<td>0.040</td>
<td>0.228***</td>
<td>0.285***</td>
<td>0.300***</td>
</tr>
<tr>
<td>CC $\times$ $\hat{C}_{-ij,t-1}$</td>
<td>0.208***</td>
<td>0.258***</td>
<td>0.320***</td>
<td>0.419***</td>
</tr>
<tr>
<td>OT $\times$ $\hat{C}_{-ij,t-1}$</td>
<td>0.181***</td>
<td>0.232***</td>
<td>-0.087</td>
<td>0.239**</td>
</tr>
<tr>
<td>Round</td>
<td>-0.003</td>
<td>0.042***</td>
<td>-0.025</td>
<td>0.048</td>
</tr>
<tr>
<td>CC $\times$ Round</td>
<td>-0.066***</td>
<td>-0.127***</td>
<td>0.022</td>
<td>-0.072*</td>
</tr>
<tr>
<td>OT $\times$ Round</td>
<td>-0.053***</td>
<td>-0.191***</td>
<td>-0.079**</td>
<td>-0.029</td>
</tr>
<tr>
<td>Constant</td>
<td>0.139</td>
<td>-0.533</td>
<td>0.997</td>
<td>0.274</td>
</tr>
<tr>
<td>N</td>
<td>5512</td>
<td>5616</td>
<td>2080</td>
<td>2080</td>
</tr>
</tbody>
</table>

Supplementary Table 5 | Comparing estimated reciprocity across attitude types. Shown are estimated fixed effects from linear mixed-effects models (details on the estimation can be found in section 1.1 - Equation 6). Dependent variable: effective contributions. $\hat{C}_{-ij,t-1}$ is the average effective contribution of the other three group members from the previous round. CC is a dummy for conditional cooperators and OT is a dummy for others. Free riders (FR) are the omitted category. Random intercepts are included for matching groups and individuals. Standard errors in parentheses. *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$. 
**Supplementary Table 6** (support for Supplementary Figure 1)

<table>
<thead>
<tr>
<th></th>
<th>(1) One-shot Maintenance</th>
<th>(2) One-shot Provision</th>
<th>(3) Strangers Maintenance</th>
<th>(4) Strangers Provision</th>
<th>(5) Partners Maintenance</th>
<th>(6) Partners Provision</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>7.521***</td>
<td>8.318***</td>
<td>1.611**</td>
<td>1.872***</td>
<td>1.503***</td>
<td>1.923†</td>
</tr>
<tr>
<td>1 if CC, 0 otherwise</td>
<td>(1.396)</td>
<td>(1.437)</td>
<td>(0.634)</td>
<td>(0.345)</td>
<td>(0.584)</td>
<td>(1.088)</td>
</tr>
<tr>
<td>OT</td>
<td>6.464***</td>
<td>6.105***</td>
<td>3.159***</td>
<td>1.817***</td>
<td>1.846***</td>
<td>2.630**</td>
</tr>
<tr>
<td>1 if OT, 0 otherwise</td>
<td>(1.435)</td>
<td>(1.785)</td>
<td>(0.691)</td>
<td>(0.427)</td>
<td>(0.702)</td>
<td>(1.218)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.379</td>
<td>1.455</td>
<td>0.226</td>
<td>0.643*</td>
<td>1.535**</td>
<td>2.353**</td>
</tr>
<tr>
<td></td>
<td>(1.146)</td>
<td>(1.302)</td>
<td>(0.509)</td>
<td>(0.346)</td>
<td>(0.686)</td>
<td>(1.052)</td>
</tr>
</tbody>
</table>

**Supplementary Table 6 | Comparing effective contributions across attitude types.** Models (1-2) report estimates from ordinary least squares (OLS) models. Models (3-6) report fixed effects estimates from linear mixed-effects models with random intercepts for matching groups and individuals (details on the estimation can be found in section 1.1 – Equations 7-8). Dependent variable: effective contributions. CC is a dummy for *conditional cooperators* and OT is a dummy for *others.* Free riders (FR) are the omitted category. Standard errors in parentheses. *** *P < 0.01, ** *P < 0.05, * *P < 0.1.*
Supplementary Table 7 (support for Figure 5a, all panels)

<table>
<thead>
<tr>
<th></th>
<th>(1) All</th>
<th>(2) CC</th>
<th>(3) FR</th>
<th>(4) OT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provision</td>
<td>0.076</td>
<td>-0.030</td>
<td>0.229**</td>
<td>0.214</td>
</tr>
<tr>
<td>1 if Provision, 0 otherwise</td>
<td>(0.066)</td>
<td>(0.078)</td>
<td>(0.106)</td>
<td>(0.144)</td>
</tr>
<tr>
<td>Positive deviation from $c_i$</td>
<td>-0.006**</td>
<td>-0.006</td>
<td>-0.006</td>
<td>-0.009**</td>
</tr>
<tr>
<td>$\max{c_k - c_i, 0}$</td>
<td>(0.002)</td>
<td>(0.005)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Provision $\times$ positive deviation from $c_i$</td>
<td>0.003</td>
<td>-0.001</td>
<td>0.009</td>
<td>0.004</td>
</tr>
<tr>
<td>(0.004)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td></td>
</tr>
<tr>
<td>Negative deviation from $c_i$</td>
<td>0.117***</td>
<td>0.119***</td>
<td>0.111***</td>
<td>0.117***</td>
</tr>
<tr>
<td>$\max{c_i - c_k, 0}$</td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.005)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Provision $\times$ negative deviation from $c_i$</td>
<td>-0.045***</td>
<td>-0.034***</td>
<td>-0.022**</td>
<td>-0.084***</td>
</tr>
<tr>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.009)</td>
<td>(0.007)</td>
<td></td>
</tr>
<tr>
<td>Round</td>
<td>-0.005***</td>
<td>-0.003**</td>
<td>-0.009***</td>
<td>-0.003</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>Provision $\times$ Round</td>
<td>0.001</td>
<td>0.002</td>
<td>0.001</td>
<td>-0.006**</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
</tr>
<tr>
<td>$\bar{C}_{i-k,j,t}$</td>
<td>0.007***</td>
<td>-0.003</td>
<td>0.006</td>
<td>0.022***</td>
</tr>
<tr>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td>Provision $\times$ $\bar{C}_{i-k,j,t}$</td>
<td>-0.006**</td>
<td>0.004</td>
<td>-0.018***</td>
<td>-0.011*</td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.006)</td>
<td>(0.007)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.109**</td>
<td>0.109*</td>
<td>0.177***</td>
<td>-0.013</td>
</tr>
<tr>
<td>(0.047)</td>
<td>(0.061)</td>
<td>(0.061)</td>
<td>(0.098)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>13932</td>
<td>6318</td>
<td>3645</td>
<td>3969</td>
</tr>
</tbody>
</table>

**Supplementary Table 7 | Comparing negative reciprocity across Maintenance and Provision.** Shown are estimated fixed effects from linear mixed-effects models (details on the estimation can be found in section 1.1 – Equation 9). Dependent variable: assigned punishment points. Model (1) compares assigned punishment points from individual $i$ to individual $k$ across Maintenance and Provision in reaction to a negative or positive deviation of $k$’s contribution from $i$’s contribution. $\bar{C}_{i-k,j,t}$ is the average contribution of the other two members of the group (excluding $i$ and $k$). Models (2-4) are the same as Model (1) but separately estimated for conditional cooperators (CC), free riders (FR), and others (OT), respectively. Random intercepts are included for matching groups and individuals. Standard errors in parentheses. *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$. 

41
Supplementary Table 8 (support for Figure 5a, Panel 2-4)

<table>
<thead>
<tr>
<th></th>
<th>(1) Maintenance</th>
<th>(2) Provision</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>-0.056</td>
<td>-0.382***</td>
</tr>
<tr>
<td>1 if CC, 0 otherwise</td>
<td>(0.106)</td>
<td>(0.096)</td>
</tr>
<tr>
<td>OT</td>
<td>-0.166</td>
<td>-0.237***</td>
</tr>
<tr>
<td>1 if OT, 0 otherwise</td>
<td>(0.109)</td>
<td>(0.108)</td>
</tr>
<tr>
<td>Positive deviation from $c_i$</td>
<td>0.006</td>
<td>0.005</td>
</tr>
<tr>
<td>$\max(c_i - c_k, 0)$</td>
<td>(0.004)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Positive deviation from $c_i \times CC$</td>
<td>0.000</td>
<td>-0.012*</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Positive deviation from $c_i \times OT$</td>
<td>-0.002</td>
<td>-0.009</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Negative deviation from $c_i$</td>
<td>0.111***</td>
<td>0.089***</td>
</tr>
<tr>
<td>$\max(c_i - c_k, 0)$</td>
<td>(0.005)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Negative deviation from $c_i \times CC$</td>
<td>0.008</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Negative deviation from $c_i \times OT$</td>
<td>0.006</td>
<td>-0.056***</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Round</td>
<td>-0.009***</td>
<td>-0.008***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Round $\times CC$</td>
<td>0.005**</td>
<td>0.007***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Round $\times OT$</td>
<td>0.006**</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>$\tilde{c}_{i-kjt}$</td>
<td>0.006</td>
<td>-0.016***</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>$\tilde{c}_{i-kjt} \times CC$</td>
<td>-0.009</td>
<td>0.017***</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>$\tilde{c}_{i-kjt} \times OT$</td>
<td>0.014**</td>
<td>0.026***</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.171**</td>
<td>0.458***</td>
</tr>
<tr>
<td></td>
<td>(0.077)</td>
<td>(0.087)</td>
</tr>
</tbody>
</table>

**Supplementary Table 8 | Comparing negative reciprocity across attitude types.** Shown are estimated fixed effects from linear mixed-effects models (details on the estimation can be found in section 1.1 – Equation 10). Dependent variable: assigned punishment points. CC is a dummy for conditional cooperators and OT is a dummy for others. FR is the omitted category. Model (1) compares assigned punishment points from individual $i$ to individual $k$ across attitudes types in Maintenance. Model (2) reports the same estimates for Provision. Random intercepts are included for matching groups and individuals. Standard errors in parentheses. *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$. 

N | 6804 | 7128
**Supplementary Table 9 (support for Figure 5b)**

<table>
<thead>
<tr>
<th>Partners with Punishment – Size of the public good</th>
<th>M</th>
<th>P</th>
<th>P =</th>
</tr>
</thead>
<tbody>
<tr>
<td>Round 1</td>
<td>42.43</td>
<td>37.05</td>
<td>0.204</td>
</tr>
<tr>
<td></td>
<td>(13.96)</td>
<td>(13.38)</td>
<td></td>
</tr>
<tr>
<td>Round 1-9</td>
<td>39.17</td>
<td>41.39</td>
<td>0.730</td>
</tr>
<tr>
<td></td>
<td>(23.41)</td>
<td>(21.99)</td>
<td></td>
</tr>
<tr>
<td>Round 10-18</td>
<td>42.28</td>
<td>43.81</td>
<td>0.854</td>
</tr>
<tr>
<td></td>
<td>(26.70)</td>
<td>(29.93)</td>
<td></td>
</tr>
<tr>
<td>Round 19-27</td>
<td>48.00</td>
<td>46.95</td>
<td>0.905</td>
</tr>
<tr>
<td></td>
<td>(29.22)</td>
<td>(30.73)</td>
<td></td>
</tr>
<tr>
<td>All rounds</td>
<td>43.15</td>
<td>44.05</td>
<td>0.904</td>
</tr>
<tr>
<td></td>
<td>(26.44)</td>
<td>(27.55)</td>
<td></td>
</tr>
</tbody>
</table>

**Supplementary Table 9 | Descriptive statistics on public good size in Partners with Punishment.** Average (Std. Dev.) public good size in Partners with Punishment. Standard deviations are calculated using differences across groups within a period, averaged across periods. M = Maintenance, P = Provision. P values are based on two-sided t-test in Round 1. All remaining P values are from linear mixed-effects models with random intercepts at the matching group level.
## Supplementary Table 10 (support for Figure 5b)

<table>
<thead>
<tr>
<th></th>
<th>Maintenance vs. Provision</th>
<th>Partners vs. Partners with Punishment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maintenance</td>
</tr>
<tr>
<td>Provision</td>
<td>0.902</td>
<td>(7.501)</td>
</tr>
<tr>
<td>$1$ if Provision, $0$ otherwise</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partners with Punishment</td>
<td>32.59***</td>
<td>27.18***</td>
</tr>
<tr>
<td>$1$ if Partners with Punishment, $0$ otherwise</td>
<td></td>
<td>(5.826)</td>
</tr>
<tr>
<td>Constant</td>
<td>43.15***</td>
<td>10.56**</td>
</tr>
<tr>
<td></td>
<td>(5.365)</td>
<td>(4.169)</td>
</tr>
<tr>
<td>$N$</td>
<td>1161</td>
<td>1107</td>
</tr>
</tbody>
</table>

**Supplementary Table 10 | Comparing the public good size in Partners with Punishment.** Shown are estimated fixed effects from linear mixed-effects models (details on the estimation can be found in section 1.1 – Equations 11-12). Dependent variable: size of the public good. Random intercepts are included for matching groups. Model (1) compares the public good size between Maintenance and Provision. Models (2-3) compare the public good size between Partners and Partners with Punishment separately for Maintenance and Provision. Standard errors in parentheses. *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$. 
### Supplementary Table 11

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Below average contribution</td>
<td>Above average contribution</td>
</tr>
<tr>
<td>Received # punishment points</td>
<td>1.235***</td>
<td>0.257***</td>
</tr>
<tr>
<td></td>
<td>(0.065)</td>
<td>(0.092)</td>
</tr>
<tr>
<td>Provision</td>
<td>1.454**</td>
<td>0.675***</td>
</tr>
<tr>
<td>1 if Provision, 0 otherwise</td>
<td>(0.617)</td>
<td>(0.261)</td>
</tr>
<tr>
<td>Provision × received # punishment points</td>
<td>-0.612***</td>
<td>-0.088</td>
</tr>
<tr>
<td></td>
<td>(0.101)</td>
<td>(0.136)</td>
</tr>
<tr>
<td>Round</td>
<td>0.019</td>
<td>0.077***</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Provision × Round</td>
<td>-0.011</td>
<td>-0.039***</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.418</td>
<td>-1.795***</td>
</tr>
<tr>
<td></td>
<td>(0.454)</td>
<td>(0.184)</td>
</tr>
<tr>
<td>N</td>
<td>1210</td>
<td>3262</td>
</tr>
</tbody>
</table>

**Supplementary Table 11 | Comparing reaction to received punishment across Maintenance and Provision.** Shown are estimated fixed effects from linear mixed-effects models (details on the estimation can be found in section 1.1 – Equation 13). Dependent variable: change in contributions from round $t-1$ to round $t$. Model (1) estimates the model for all cases where the contribution of individual $i$ in round $t-1$ is below the average contribution of the other three group members. Model (2) estimates the same model for the cases in which the contribution of individual $i$ in round $t-1$ is above the average contribution of the other three group members. Random intercepts are included for matching groups and individuals. Standard errors in parentheses. *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$. 
### Supplementary Table 12

<table>
<thead>
<tr>
<th>Term</th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive deviation from $c_i$ in $t-1$</td>
<td>0.172***</td>
<td>(0.04)</td>
</tr>
<tr>
<td>$\max(\bar{c}<em>{i,t-1} - c</em>{i,t-1}, 0)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative deviation from $c_i$ in $t-1$</td>
<td>-0.554***</td>
<td>(0.05)</td>
</tr>
<tr>
<td>$\max(c_{i,t-1} - \bar{c}_{i,t-1}, 0)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive deviation from $c_i$ in $t-1 \times$ Provision</td>
<td>0.126**</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Negative deviation from $c_i$ in $t-1 \times$ Provision</td>
<td>-0.154**</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Lagged own contribution $c_{i,t-1}$</td>
<td>0.814***</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Round</td>
<td>-0.005</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Lagged own contribution $c_{i,t-1} \times$ Provision</td>
<td>0.048</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Round $\times$ Provision</td>
<td>0.010</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Provision</td>
<td>0.116</td>
<td>(0.43)</td>
</tr>
<tr>
<td>$I$ if Provision, 0 otherwise</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.913***</td>
<td>(0.29)</td>
</tr>
</tbody>
</table>

### Supplementary Table 12 | Comparing reactions to positive and negative deviations from own contributions in Partners.

Shown are estimated fixed effects from linear mixed-effects models (details on the estimation can be found in section 1.1 – Equation 14). Dependent variable: effective contributions. Model (1) shows the reaction of contributions of individual $i$ to positive and negative deviations from others’ average contributions in the previous period, $\bar{c}_{i,t-1}$. Random intercepts are included for matching groups and individuals. Standard errors in parentheses. *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$. 

N 4160
**Supplementary Table 13 (Robustness check for Figure 3a)**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. contribution others</td>
<td>0.333***</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
</tr>
<tr>
<td>Provision</td>
<td>-1.796***</td>
</tr>
<tr>
<td>1 if Provision, 0 otherwise</td>
<td>(0.403)</td>
</tr>
<tr>
<td>Avg. contribution others × Provision</td>
<td>0.176***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
</tr>
<tr>
<td>Constant</td>
<td>3.057***</td>
</tr>
<tr>
<td></td>
<td>(0.286)</td>
</tr>
<tr>
<td>N</td>
<td>14784</td>
</tr>
</tbody>
</table>

**Supplementary Table 13 | Comparing cooperation attitudes between Maintenance and Provision.** Shown are estimated fixed effects from a linear mixed-effects model (details on the estimation can be found in section 1.2 - Equation 15). Dependent variable: effective contribution in the contribution schedule $C_{ik}$. Random intercepts are included for each individual. Standard errors in parentheses. *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$. 
Supplementary Table 14 (Robustness check for Figure 3b)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Strangers Maintenance</td>
<td>Strangers Provision</td>
<td>Partners Maintenance</td>
<td>Partners Provision</td>
</tr>
<tr>
<td><strong>CC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \bar{C}_{ij,t-1} )</td>
<td>0.334**</td>
<td>0.659***</td>
<td>0.897***</td>
<td>0.974***</td>
</tr>
<tr>
<td></td>
<td>(0.142)</td>
<td>(0.051)</td>
<td>(0.071)</td>
<td>(0.039)</td>
</tr>
<tr>
<td><strong>Round</strong></td>
<td>-0.051</td>
<td>-0.204***</td>
<td>-0.045</td>
<td>-0.052***</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.022)</td>
<td>(0.038)</td>
<td>(0.019)</td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>5.409***</td>
<td>3.093***</td>
<td>0.911</td>
<td>0.318</td>
</tr>
<tr>
<td></td>
<td>(1.099)</td>
<td>(0.427)</td>
<td>(0.815)</td>
<td>(0.397)</td>
</tr>
<tr>
<td><strong>STR</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Round</strong></td>
<td>-0.651***</td>
<td>-0.993***</td>
<td>-0.911***</td>
<td>-0.664***</td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.155)</td>
<td>(0.155)</td>
<td>(0.146)</td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>-1.027</td>
<td>-3.118*</td>
<td>-4.015**</td>
<td>-3.888*</td>
</tr>
<tr>
<td></td>
<td>(0.779)</td>
<td>(1.778)</td>
<td>(2.034)</td>
<td>(2.294)</td>
</tr>
<tr>
<td><strong>Estimated mixing proportions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CC</strong></td>
<td>0.184***</td>
<td>0.456***</td>
<td>0.300***</td>
<td>0.549***</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.047)</td>
<td>(0.057)</td>
<td>(0.057)</td>
</tr>
<tr>
<td><strong>STR</strong></td>
<td>0.460***</td>
<td>0.327***</td>
<td>0.451***</td>
<td>0.401***</td>
</tr>
<tr>
<td></td>
<td>(0.045)</td>
<td>(0.044)</td>
<td>(0.061)</td>
<td>(0.056)</td>
</tr>
<tr>
<td><strong>FR</strong></td>
<td>0.356***</td>
<td>0.217***</td>
<td>0.249***</td>
<td>0.050**</td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.037)</td>
<td>(0.048)</td>
<td>(0.024)</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>3328</td>
<td>3328</td>
<td>2080</td>
<td>2080</td>
</tr>
<tr>
<td><strong>Log L</strong></td>
<td>-3109.1</td>
<td>-4270.3</td>
<td>-2624.6</td>
<td>-3892.7</td>
</tr>
<tr>
<td><strong>AIC</strong></td>
<td>6236.2</td>
<td>8558.7</td>
<td>5267.1</td>
<td>7803.4</td>
</tr>
</tbody>
</table>

**Supplementary Table 14 | Maximum likelihood estimation of two-limit Tobit finite mixture models.** Details on the estimation procedure can be found in section 1.2. Dependent variable: effective contributions. \( \bar{C}_{ij,t-1} \) is the average contribution of the other three group members from the previous round. CC = *conditional cooperators*, STR = *strategic free riders*, and FR = *free riders*. The bottom part of the table shows the estimated mixing proportions. Estimates for FR are deduced from the sum of proportions of CC and STR, the standard error for FR is obtained from the covariance matrix of the estimates of CC and STR. AIC is the Akaike’s information criterion. Standard errors in parentheses. **\( P < 0.01 \), ***\( P < 0.05 \), *\( P < 0.1 \).
4. Supplementary Figures

Supplementary Figure 1

Supplementary Figure 1 | Effective average contributions by attitude type and type of strategic interaction. A, Maintenance. B, Provision. CC: Conditional Cooperators, FR: Free Riders, OT: Others.
Supplementary Figure 2

Supplementary Figure 2 | Replication of results on strong reciprocity in Partners with Punishment.

Strong reciprocity as measured by cooperation attitudes elicited prior to the repeated experiment; $n_M = 84$, $n_P = 88$. $\chi^2$-tests, *** $P < 0.01$, ** $P < 0.05$, ns (not significant) $P > 0.10$. 

![Bar chart showing conditional cooperators, free riders, and others with percentage values and significance levels.]
**Supplementary Figure 3**

![Bar chart showing distribution of the number of periods in which a participant punished a group member in Partners with Punishment.](image)

**Supplementary Figure 3 | Distribution of the number of periods in which a participant punished a group member in Partners with Punishment.** The figure is constructed counting for each individual in how many out of the 27 periods they punish some other group member. The mean (median) number of rounds where participants punish is 7.9 (6) and 7.1 (6) in Maintenance and Provision, respectively, with no differences across treatments (Kolmogorov Smirnov test, $P = 0.870$).
Supplementary Figure 4 | Simulated effective contribution ratios. Distribution of 1000 simulated effective contribution ratios between Maintenance and Provision ($\bar{e}^M / \bar{e}^P$) using a sample of $n = 100$ per treatment and simulated experiment.