# Reciprocity and the Tragedies of Maintaining and Providing the Commons

Simon Gächter[1,4,5,*], Felix Kölle[2] & Simone Quercia[3]

**Affiliations:**

[1] School of Economics, University of Nottingham, UK.

[2] Faculty of Management, Economics and Social Sciences, University of Cologne, Germany.

[3] Institute for Applied Microeconomics, University of Bonn, Germany.

[4] CESifo, Munich, Germany.

[5] IZA, Bonn, Germany.

*Correspondence to: simon.gaechter@nottingham.ac.uk.

27 July 2017

**Social cooperation often requires collectively beneficial but individually costly restraint to maintain a public good[1-4], or it needs costly generosity to create one[1,5]. Status quo effects[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'[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[7-10], 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.**

30    Humans are an exceptionally cooperative species able to collaborate for the creation

31    of common benefit[9,11-13]. Collective actions such as voting, participating in political

32    movements, the provision of the welfare state, charity, volunteering and teamwork are

33    examples of public goods that come into existence by the generosity of many people

34    that puts the greater good before self-interest[1,5]. Cooperation is, however, not always

35    about providing collectively valuable resources, but often about maintaining existing

36    ones[1-4]. Limiting $CO_2$ emissions, sustaining natural resources, or maintaining common

37    pastures and biodiversity are important examples of cooperation problems that require

38    restraint in exploiting existing socially beneficial public goods.

39       In this paper, we show experimentally and with simulations that cooperation for

40    maintaining an initially existing public good is substantially and systematically

41    weaker than cooperation for creating a new public good even if they are otherwise

42    identical social dilemmas. This is unexpected, given that many people are biased

43    towards the status quo and defaults[6], which should ease cooperation when the public

44    good already exists compared to when it needs to be provided.

45       We show that the reason for lower cooperation in the maintenance dilemma is

46    that reciprocity, a fundamental force behind the evolution of cooperation and human

47    sociality[7-10], is substantially diminished in maintaining compared to providing a

48    public good. Simulations show that, despite some variability, lower cooperation in

49    Maintenance than Provision is a systematic effect to be expected with a likelihood of

50    70%. The simulation results also provide an explanation for the mixed findings in

51    some related literature.[14-23]

52       In our experiments, we focus sharply on the behavioural differences between

53    initially existing and inexistent public goods (Fig. 1) and abstract from technological

54    complexities, loss aversion, time discounting and institutional details relevant in real

55  world social dilemmas[1,24-29]. In 'Maintenance', a group of four people possesses a

56  common pool of 80 tokens and each member can withdraw up to 20 tokens.

57  Upholding the status quo by withdrawing nothing earns each group member 32

58  money units (MU); if all withdraw maximally, everyone earns 20 MU. In 'Provision',

59  the common pool is initially empty and 80 tokens are distributed equally among group

60  members who decide simultaneously how many tokens (up to 20) to contribute to the

61  pool. In the status quo all earn 20 MU, and all contributing maximally earns each

62  member 32 MU.

63      Using the setup described in Fig. 1, we run three series of experiments with 704

64  participants who interact anonymously in three generic settings of social interaction

65  (see Methods).  All experiments involve a between-subjects comparison of

66  cooperation in Maintenance and Provision. We also elicit beliefs about group

67  members' contributions to (or withdrawals from) the public good. Participants need to

68  successfully complete a comprehension test before the experiment starts.

69      In the first experiment, called One-shot, participants ($n = 288$) take a single

70  decision only. This experiment is a basic measure of people's cooperativeness in the

71  absence of strategic incentives to cooperate. In a second experiment, called Strangers,

72  participants ($n = 256$) play the games of Fig. 1 for 27 iterations with randomly

73  changing group composition in each round. This experiment is a sequence of one-shot

74  interactions that permit learning about cooperativeness in the population[30-32]. The

75  third experiment ($n = 160$), called Partners, keeps group composition constant across

76  the 27 iterations, which creates strategic incentives for cooperation[32,33].

77      The effective size of the public good (after withdrawals or contributions) is

78  smaller in Maintenance than Provision in all experiments (Fig. 2, Supplementary

79  Table 1). In One-shot, the public good in Maintenance is on average 27% smaller than

80    in Provision (Fig. 2, Panel 1; 23.8 vs. 32.6; two-sided t-test, $t = -2.51$, $P = 0.014$).

81    Low levels of the public good (less than 10% of the optimal size of 80), are more

82    likely in Maintenance than Provision (23% vs. 0%; $\chi^2(1) = 9.51$, $P = 0.002$).

83        In Strangers, the public good starts out 23% lower in Maintenance than

84    Provision (22.7 vs. 29.5; two-sided t-test, $t = -1.92$, $P = 0.059$) and decays on average

85    to about 5% of the socially efficient level in both problems (Fig. 2, Panel 2). Thus, the

86    tragedy of the commons is almost maximal in both Maintenance and Provision.

87        In Partners, the public good starts 33% smaller in Maintenance than Provision

88    (27.7 vs. 41.3; two-sided t-test, $t = -2.96$, $P = 0.005$) and drops over time (Fig. 2,

89    Panel 3). On average the public good is 37.3% smaller in Maintenance than Provision

90    (10.6 vs.16.9; linear mixed effects model, $P = 0.035$).

91        Comparing Partners and Strangers reveals the extent to which strategic

92    incentives help the provision of the public good. We find that in Maintenance the

93    average size of the public good is only 3.6 units higher in Partners than Strangers

94    (10.6 vs. 7.0; linear mixed-effects model, $P = 0.346$, Supplementary Table 2), while

95    in Provision the public good is on average twice as large in Partners than Strangers

96    (16.9 vs. 8.2; linear mixed effects model, $P = 0.004$). Thus, strategic incentives to

97    increase cooperation are substantially weaker in Maintenance than Provision.

98        Taken together, these results show that high levels of the public good are harder

99    to achieve in Maintenance than Provision in One-shot and in the first period of

100    Partners and Strangers. This is surprising given that in Maintenance the public good

101    enjoys a head start because it is already provided at the outset. Furthermore, while in

102    Strangers the size of the public good converges to similar long-run equilibrium levels,

103    in Partners the initial differences are persistent and lead to different long-run

104    outcomes between Provision and Maintenance. The aim of our further analysis is to

105    understand the differences in cooperation outcomes by investigating whether initial

106    resource allocation affects reciprocity in response to restraint and generosity,

107    respectively.

108        Studying reciprocity is particularly interesting due to its fundamental role for

109    human sociality[7-10]. In our settings, reciprocity takes the form of conditional

110    cooperation: the willingness to cooperate provided others do the same[30,32,34,35]. Here,

111    we distinguish between two forms of conditional cooperation, which are inspired by

112    the concepts of weak and strong reciprocity[9,10,36]. Weak reciprocity can occur in stable

113    relationships and means behaving conditionally cooperative for self-regarding

114    strategic reasons. By contrast, strong reciprocity entails non-selfish conditional

115    cooperation not only in repeated interactions but also in one-shot games. Strong

116    reciprocity is a preference for conditional cooperation, whereas weak reciprocity is a

117    behavioural strategy deployed for self-regarding reasons.

118        Studying reciprocity as a preference requires looking beyond cooperation

119    outcomes and to measure attitudes to cooperation separately from outcomes. The

120    reason why this is important is that people who differ in their *ex ante* attitudes can *ex*

121    *post* make the same cooperation decision. To see why, consider that a conditional

122    cooperator's ex ante attitude is to cooperate only if they believe their group members

123    do so too. But there may also be 'free riders', who never want to contribute to the

124    public good irrespective of their beliefs how much others contribute. A conditional

125    cooperator who believes that others do not contribute and a person with a free rider

126    attitude both contribute nothing: their ex post behavior is observationally equivalent

127    despite different ex ante attitudes. Thus, if cooperation is a function of attitudes and

128    beliefs, the challenge is to separate them empirically. Our approach, which we call the

129    'ABC of cooperation', achieves this separation. This also allows us to compare strong

5

130    reciprocity as measured by the ABC approach with reciprocity estimated from

131    observed behaviour.

132         The ABC approach measures individual attitudes ($a_i$), beliefs ($b_i$), and effective

133    contributions ($c_i$) separately and explains cooperation as $a_i(b_i) \rightarrow c_i$. It is inspired

134    by[30] and implemented as follows. All three experiments start with an incentive-

135    compatible elicitation of attitudes without feedback in a one-shot version of either the

136    Maintenance or the Provision dilemma. The elicited attitudes are our main measure of

137    strong reciprocity. Eliciting attitudes involves specifying a vector $a_i$ of contributions

138    or withdrawals as a function of all possible average contributions or withdrawals of

139    other group members. We classify participants as conditional cooperators (that is,

140    strong reciprocators) if the entries in the vector $a_i$ are increasing in others'

141    contributions or withdrawals, or as a free rider if a participant's $a_i$ consists of only

142    zero contributions or maximal withdrawals. We refer to the remaining participants as

143    'others'. After attitude elicitation, the three experiments proceed as described above.

144    In all experiments, we elicit incentivized beliefs ($b_i$) about other group members'

145    average withdrawal or contribution and we observe effective contributions ($c_i$) to the

146    public good (see Methods).

147         In the repeated direct interactions of Strangers and Partners we measure

148    conditional cooperation in linear mixed-effects models by regressing individual

149    contributions or withdrawals on the average contribution or withdrawals of other

150    group members in the previous period (Supplementary Information). The relation

151    between these two variables, the coefficient $\beta_1$, is our measure of conditional

152    cooperation. We will call $\beta_1$ 'estimated reciprocity'.

153         In Strangers, $\beta_1$ is an estimate of strong reciprocity because there are no

154    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; multi-level 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,

180    panels 2 and 3; multilevel mixed-effects models, $P < 0.001$). Conditional cooperators

181    also display significantly higher $\beta_1$ in Partners than Strangers (multilevel mixed-

182    effects models, $P < 0.001$). As predicted, free riders in Strangers display low

183    estimated reciprocity but show increased $\beta_1$ in Partners compared to Strangers.

184    Participants classified as 'others' do display a substantial $\beta_1$ but do not react to

185    strategic incentives (Fig. 3b, panel 4; multilevel mixed-effects models, $P > 0.166$).

186        Our next step is to investigate whether the differences in reciprocity across

187    Maintenance and Provision can explain the observed differences in cooperation

188    outcomes (Fig. 2). We do this by applying our ABC framework that uses attitudes and

189    beliefs to explain effective contributions. We calculate predicted effective

190    contributions $[a_i(b_i) \rightarrow c_i^*]$ and compare them with actual effective contributions $c_i$

191    from One-shot as well as with the effective first-period contributions in the repeated

192    experiments (Methods). Predicted and actual effective contributions are highly

193    significantly positively correlated in One-shot as well as in all repeated games (all

194    Spearman's $\rho > 0.59$; $P < 0.001$).

195        We also calculate individual-level deviations from the predicted effective

196    contribution, $c_i^* - c_i$. In One-shot, this measure lies within $\pm 2$ tokens in 63% and

197    62% of the cases in Maintenance and Provision, respectively, with no differences

198    between treatments ($\chi^2(1) = 0.01$, $P = 0.903$). We obtain similar results for first-period

199    effective contributions in Strangers (66% and 63%; $\chi^2(1) = 0.43$, $P = 0.514$) and

200    Partners (74% and 64%; $\chi^2(1) = 1.86$, $P = 0.172$). Finally, effective contributions

201    differ significantly between attitude types: free riders contribute significantly less than

202    conditional cooperators and 'others' in all conditions (Supplementary Fig. 1).

203        The fact that the ABC approach predicts equally well in Maintenance and

204    Provision allows us to use the elicited attitudes and beliefs as a 'population pool' from

205     which we can sample at random to run 'simulated experiments' (Methods and

206     Supplementary Information).  The advantage of simulations is that we are not

207     restricted to a specific laboratory sample we happen to draw at a given instance (with

208     hitherto unobservable attitudes and beliefs); we can cost effectively perform a large

209     number of identical experiments and therefore elicit a distribution of likely

210     cooperation ratios of Maintenance relative to Provision. This also allows us to check

211     how systematic the results are that we observe.

212        The results of 1000 simulated experiments (Fig. 4) show that effective

213     cooperation levels in Maintenance are lower than in Provision in 70% of all simulated

214     experiments. This result shows that our findings that cooperation in Maintenance is

215     lower than in Provision are systematic.

216        Given that our results reveal important asymmetries in positive reciprocity

217     between Maintenance and Provision, it is interesting to study whether initial resource

218     allocation also affects negative reciprocity, which in our setting takes the form of

219     punishment[9,36]. Furthermore, punishment is an expression of moral disapproval and

220     social norms[37] that are important in many real world public goods[38]. If the differences

221     in positive reciprocity in Maintenance and Provision also translate into negative

222     reciprocity, we should observe less punishment in Maintenance than Provision and,

223     therefore, also a reduced effectiveness of punishment to stabilize cooperation in

224     Maintenance compared to Provision.

225        We study punishment in a fourth experiment ('Partners with Punishment';

226     $n = 172$), which is identical to Partners except for an added punishment stage in each

227     period after group members have made their withdrawal or contribution decisions[39].

228     In the punishment stage, each group member can assign up to 5 punishment points to

229    each other member, where each punishment point costs one MU and reduces the

230    earnings of the punished group member by three MU (see Methods).

231        The attitudes elicited prior to the experiment replicate the results from Fig. 3a

232    (Supplementary Fig. 2). Contrary to expectations, negative reciprocity, estimated as

233    assigned punishment in reaction to negative deviations of others from own effective

234    contribution, is substantially and significantly higher in Maintenance than in

235    Provision. This effect is present both overall and for each attitude type (Fig. 5a, panels

236    1-4), and it is not driven by different frequencies of punishers (Supplementary Figure

237    3). There are no treatment differences for positive deviations (Fig. 5a, panel 1;

238    Supplementary Table 7). Interestingly, in contrast to estimated positive reciprocity,

239    estimated negative reciprocity does not differ between conditional cooperators and

240    free riders (Fig. 5a, panels 2-4; Supplementary Table 8).

241        As expected[40], punishment increases the public goods to substantially higher

242    levels compared to Partners (Fig. 5b; linear mixed effect models; Maintenance: 43.1

243    vs. 10.6, $P < 0.001$; Provision: 44.1 vs. 16.9, $P < 0.001$; Supplementary Tables 9-10).

244    Remarkably, the sizes of public goods are now very similar in Maintenance and

245    Provision (linear mixed effect models; Maintenance: 43.1, Provision: 44.1, $P =$

246    0.904). Besides stronger negative reciprocity, a further reason for this result is that

247    reactions to received punishment (in terms of change in effective contributions) are

248    also stronger in Maintenance than Provision (Supplementary Table 11).

249        One way to reconcile the results on positive and negative reciprocity in Partners

250    and Partners with Punishment, respectively, is to argue that also in Partners people

251    engage in punishment by reducing their contributions in the current period as a

252    reaction to previous negative deviations of others from own effective contributions. If

253    such 'implicit' punishment is stronger in Maintenance than Provision, it could explain

254 why the decay in effective contributions is stronger in Maintenance than Provision.

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

256    We find that participants in Partners significantly increase their contributions in

257 round $t$ in response to positive deviations of others from own contributions in round

258 $t-1$; the reverse holds for negative deviations. However, we find both of these

259 reactions to be significantly more pronounced in Provision than in Maintenance

260 (linear mixed effect models; both $P < 0.018$; Supplementary Table 12). This confirms

261 once again stronger conditional cooperation in Provision compared to Maintenance in

262 Partners. It also suggests another interpretation of the results of Partners with

263 Punishment: because voluntary conditional cooperation is weaker in Maintenance

264 than Provision, stronger extrinsic incentives are needed, here in the form of

265 punishment, to stabilize cooperation in Maintenance at similar levels than in

266 Provision.

267    Our analysis has revealed that the important principles of human cooperation of

268 strong and weak reciprocity[7-10] are substantially diminished when cooperation

269 requires restraint in exploiting a public good as opposed to when cooperation calls for

270 generosity to provide a public good. Our findings are consistent with the observation

271 that failing to contribute to a public good is judged more morally blameworthy than

272 exploiting an existing public good.[41]

273    Our results can also be explained by a model of revealed altruism[42,43], according

274 to which initial resource allocation affects perceptions of generosity of actions and

275 hence subsequent reciprocity.  Because in Provision cooperation is the result of an act

276 of commission (contributing), while in Maintenance cooperation is achieved by

277 omission (not withdrawing), cooperation in Provision is perceived as more generous

278 than in Maintenance and thus Provision triggers stronger positive reciprocity than

279 Maintenance. By contrast, our results suggest that negative reciprocity, as expressed

280 by people's costly punishment, does not follow this logic because punishment is more

281 severe in Maintenance than Provision, likely to compensate for weaker voluntary

282 cooperation in Maintenance.

283       Our findings from the experiments without punishment and the simulations also

284 help explaining the mixed evidence from previous related literature, which, with a few

285 exceptions[35,44,45], only compared cooperation outcomes, that is, the effective size of

286 the public good after contribution or withdrawal decisions. Some of these studies find

287 higher cooperation in so-called 'give-some' vs. 'take-some' games[14-17], some find the

288 reverse[18] and some find no significant differences[19-23]. The simulations based on our

289 ABC approach can explain these mixed results (Fig. 4) but they also show that on

290 average cooperation in Maintenance is generally expected to be lower than in

291 Provision. The finding that Maintenance and Provision are systematically different

292 also suggests that future research should choose the game (Maintenance or Provision)

293 that comes closest to the social dilemma of interest.

294       Our results also have potential policy relevance.[46] Recent policy proposals to

295 foster cooperation build on the power of reciprocity in combination with economic

296 incentives[47,48]. Policy makers who reckon with reciprocity should therefore consider

297 that the extent of reciprocity that can be evoked is dilemma-specific. Moreover, a

298 problem of incentives is that they might 'crowd out' strong reciprocity because

299 incentives typically strengthen self-regarding motives to cooperate[49]. Our finding of

300 higher reciprocity in Provision than Maintenance suggests that crowding out may be

301 more problematic in provision problems than in maintenance problems, because in

302 Maintenance more people display non-reciprocal attitudes. Future research will need

303 to address these issues, including how reciprocity and incentives interact in non-linear

304  settings with thresholds, resource rivalry, discounting, and hybrid social dilemmas

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

306  **Methods**

307  **Isomorphism of Maintenance and Provision under monetary incentives.** In *Maintenance*,

308  each group of 4 members is initially endowed with 80 tokens placed in a "group project";

309  individual members have no endowment. Material incentives are described by equation (1):

310  $$\pi_i = w_i + 0.4\left(80 - \sum_{j=1}^{4} w_j\right) \tag{1}$$

311  where $0 \leq w_i \leq 20$ indicates the withdrawal of individual $i$ from the project.

312  In *Provision*, the "group project" is initially empty and each group member has an

313  endowment of 20 tokens instead. The material incentives for each individual $i$ are described by

314  equation (2):

315  $$\pi_i = 20 - c_i + 0.4 \sum_{j=1}^{4} c_j \tag{2}$$

316  where $0 \leq c_i \leq 20$ denotes the contribution of individual $i$ to the project.

317  Hence, under rationality and money maximization, Maintenance and Provision are

318  isomorphic social dilemmas. Using $c_j = 20 - w_j$ for $j = 1, ... , 4$ and substituting into eq. (2) we

319  obtain (1). Analogously, using $w_j = 20 - c_j$ for $j = 1, ... , 4$ and substituting into (1) yields (2).

320
321  **Experimental design details.** The experiments were approved by the Research Ethics

322  Committee in the School of Economics at the University of Nottingham. We conducted four

323  series of experiments using the two decision situations described above and in Fig. 1. Each

324  experiment was composed of three parts that allow to elicit the three components of our ABC

325  framework: an individual $i$'s attitude ($a_i$) towards cooperation ($i$'s 'type'), $i$'s beliefs ($b_i$) about

326  others' contribution, and $i$'s contribution decision ($c_i$). Participants knew that the experiment

327  consisted of several parts but only received information about the relevant part upon

328  progression of the experiment. To avoid spillover effects between different parts, information

329  about decisions and payoffs were given only at the very end of the experiment. Experimental

330  instructions are in the Supplementary Information.

331  In Part 1, participants were introduced to either the Maintenance or Provision problem.

332  Before continuing, participants answered a set of computerized control questions.

333  In Part 2, we elicited cooperation attitudes $a_i$ using a variant of the strategy method[50],

334  which allows eliciting an individual's willingness to cooperate as a function of the other group

335  members' cooperation decisions. Participants were asked to make an unconditional and a

336  conditional cooperation decision. In the unconditional decision, participants were simply asked

337  how much they want to withdraw from (contribute to) the common pool. In the conditional

338  contribution participants had to fill a withdrawal (contribution) table in which they had to state

339      their withdrawal (contribution) decision for each possible (rounded) average withdrawal

340      (contribution) of the other three group members. This gives us the vector $a_i$, our measure of

341      strong reciprocity. To achieve incentive compatibility, in each group a random mechanism

342      selected three members for which the unconditional decision was payoff-relevant and one

343      member for whom the conditional decision for the (rounded) average unconditional withdrawal

344      (contribution) of the three other group members was payoff-relevant.

345      Part 3 comprised a direct-response interaction that differed in its exact design protocol

346      across the four experiments as described in the main text (One-shot, Strangers, Partners, and

347      Partners with Punishment). This elicits component $c_i$ of the ABC framework.

348      In all repeated experiments (Strangers, Partners, and Partners with Punishment),

349      participants were matched in groups of four and interacted for 27 consecutive rounds under

350      payment rules (1) or (2). Participants were not told how many rounds the experiment would

351      last.[51] This avoids endgame effects and also seems realistic for many common resource

352      problems, which do not have a known endpoint. In Strangers, participants were re-matched

353      randomly in 16-participants matching groups after every round, while in Partners and Partners

354      with Punishment group composition remained constant across all 27 rounds. At the end of each

355      round, participants received aggregate feedback on choices and outcomes.

356      In all rounds of the direct-response interactions, we also elicited beliefs about average

357      effective contributions of the other three group members. Participants were paid for the

358      accuracy of their beliefs. They earned 3 points if their belief was exactly correct, and 2 (1)

359      points when their belief deviated by 1 (2) point(s) from the true average effective contribution.

360      If their estimation was off by more than two points, they received no additional money. This

361      elicits component $b_i$ of our framework.

362

363      **Data collection and subject-pool socio-demographics**. A total of $n = 876$ students

364      participated in our experiments (Maintenance: $n_M = 432$, Provision: $n_P = 444$). Participants were

365      recruited with the help of ORSEE[52] from the volunteer student subject pool at the University of

366      Nottingham. Participation was upon informed consent. The average age was 20.1 years (s.d.

367      2.25 years); 57% were females. 59% were British, 22% Asian, 12% from other European

368      countries and the rest from other countries. 20% were economics or business students, 18%

369      other social sciences, 20% humanities, 14% sciences, 12% engineering, and 12% medical

370      science, and 4% law. We conducted all experiments in the CeDEx lab at the University of

371      Nottingham using z-Tree.[53] The experiments lasted between 70 to 210 minutes depending on

372      the experimental condition. Participants earned on average £20.60.

373

374      **Predicting effective contributions**. In One-shot, Strangers, and Partners, the ABC approach

375      allows us to predict contributions using elicited cooperation attitudes $a_i$ and beliefs

376    $b_i: a_i(b_i) \rightarrow c_i^*$. By matching beliefs with the corresponding decision in the contribution

377    (withdrawal) table, we predict a contribution (withdrawal) decision $c_i^*$ for each subject and

378    compare $c_i^*$ with the actual contribution $c_i$ that we observe in the direct-response

379    experiment.[30,54]

380

381    **Classification of attitudes.** We analyse cooperation in the strategy-method experiment treating

382    each participant's effective contribution schedule (the vector $a_i$) as an independent observation.

383    We classify cooperation attitudes into three main behavioural types[30]: a participant is a

384    *conditional cooperator* if either his/her effective contribution schedule exhibits a (weakly)

385    monotonically increasing pattern, or if the Spearman correlation coefficient between his/her

386    schedule and the others' average contribution is positive and significant at the 1% level; a *free*

387    *rider* if he/she never contributes anything (always withdraws everything) irrespective of how

388    much others contribute (withdraw); (iii) *other* if neither (i) nor (ii) applies. Attitudes are a proxy

389    for cooperation preferences because they reflect a willingness to pay for cooperation as a

390    function of other group members' cooperation.

391

392    **Simulations.** For each simulated experiment, we randomly sample (with replacement) from the

393    participant pool of Maintenance experiments attitudes and beliefs ($n = 60$, the median sample

394    size in related studies also using linear public goods[14-23]) and calculate simulated effective

395    contributions [$a_i(b_j) \rightarrow \tilde{c}$]. We do the same for $n = 60$ Provision attitudes and beliefs. This

396    resembles an experiment where a researcher invites 60 participants per treatment and then

397    observes their effective contribution. As a participation pool, we use all $n = 876$ attitudes from

398    our four experiments (Maintenance: $n_M = 432$, Provision: $n_P = 444$), and $n = 544$ beliefs

399    (Maintenance: $n_M = 268$, Provision: $n_P = 276$) from One-shot as well as the first period of

400    Strangers. Details are in the Supplementary Information.

401

402    **Statistical analysis.** In the One-shot direct interaction, we treat ($b_i, c_i$) as independent

403    observations. In the repeated interactions, we treat beliefs and effective contributions at the

404    matching group level as an independent observation. Matching groups are composed by 16

405    participants in Strangers and by 4 participants in Partners and Partners with Punishment,

406    respectively. For the repeated experiments, all estimations are performed using linear mixed

407    models with random intercepts at the matching group and the individual level (see Statistical

408    Analysis in SI for details on model specifications).

409

410    **Data availability**. The data for the statistical analyses are stored in Dryad Data package title:

411    Reciprocity in Maintaining and Providing Public Goods;

412    http://dx.doi.org/10.5061/dryad.8d9t2

413

414    **Code availability**. We used STATA 14.2 for data analysis. The codes are stored in Dryad

415    Data package title: Reciprocity in Maintaining and Providing Public Goods;

416    http://dx.doi.org/10.5061/dryad.8d9t2

417

418


# References

420    1.    Van Lange, P. A. M., Balliet, D., Parks, C. D. & Van Vugt, M., *Social*
421          *dilemmas. The psychology of human cooperation*. (Oxford University Press,
422          Oxford, 2014).
423    2.    Hardin, G. The tragedy of the commons. *Science* **162**, 1243 (1968).
424    3.    Ostrom, E., *Governing the commons. The evolution of institutions for*
425          *collective action*. (Cambridge University Press, Cambridge, 1990).
426    4.    Collier, P., *The plundered planet*. (Penguin Books, London, 2011).
427    5.    Gintis, H., Bowles, S., Boyd, R. & Fehr, E. eds., *Moral sentiments and*
428          *material interests. The foundations of cooperation in economic life*. (MIT
429          Press, Cambridge, 2005).
430    6.    Samuelson, W. & Zeckhauser, R. Status quo bias in decision making. *J. Risk*
431          *Uncertain.* **1**, 7 (1988).
432    7.    Trivers, R. L. Evolution of reciprocal altruism. *Quarterly Review of Biology*
433          **46**, 35 (1971).
434    8.    Axelrod, R. & Hamilton, W. The evolution of cooperation. *Science* **211**, 1390
435          (1981).
436    9.    Bowles, S. & Gintis, H., *A cooperative species: Human reciprocity and its*
437          *evolution*. (Princeton University Press, Princeton, 2011).
438    10.   Rand, D. G. & Nowak, M. A. Human cooperation. *Trends in Cognitive*
439          *Sciences* **17**, 413 (2013).
440    11.   Nowak, M. A. & Highfield, R., *Supercooperators: Altruism, evolution, and*
441          *why we need each other to succeed*. (Free Press, New York, 2011).
442    12.   Turchin, P., *Ultrasociety: How 10,000 years of war made humans the greatest*
443          *cooperators on earth*. (Beresta Books, LLC, Chaplin, CT, 2016).
444    13.   Tomasello, M., *A natural history of human morality*. (Harvard University
445          Press, Cambridge, MA, 2016).
446    14.   Andreoni, J. Warm glow versus cold prickle - the effects of positive and
447          negative framing on cooperation in experiments. *Quarterly Journal of*
448          *Economics* **110**, 1 (1995).
449    15.   Park, E.-S. Warm-glow versus cold-prickle: A further experimental study of
450          framing effects on free-riding. *Journal of Economic Behavior and*
451          *Organization* **43**, 405 (2000).
452    16.   Khadjavi, M. & Lange, A. Doing good or doing harm: Experimental evidence
453          on giving and taking in public good games. *Experimental Economics* **18**, 432
454          (2015).

455    17.    Cox, C. A. Decomposing the effects of negative framing in linear public goods
456          games. *Economics Letters* **126**, 63 (2015).

457    18.    Sell, J. & Son, Y. Comparing public goods and common pool resources: Three
458          experiments. *Social Psychology Quarterly* **60**, 118 (1997).

459    19.    Van Dijk, E. & Wilke, H. Is it mine or is it ours? Framing property rights and
460          decision making in social dilemmas. *Organizational Behavior and Human*
461          *Decision Processes* **71**, 195 (1997).

462    20.    Cubitt, R., Drouvelis, M. & Gächter, S. Framing and free riding: Emotional
463          responses and punishment in social dilemma games. *Experimental Economics*
464          **14**, 254 (2011).

465    21.    Dufwenberg, M., Gächter, S. & Hennig-Schmidt, H. The framing of games
466          and the psychology of play. *Games and Economic Behavior* **73**, 459 (2011).

467    22.    Messer, K. D., Zarghamee, H., Kaiser, H. M. & Schulze, W. D. New hope for
468          the voluntary contributions mechanism: The effects of context. *Journal of*
469          *Public Economics* **91**, 1783 (2007).

470    23.    Cox, J. C., Ostrom, E., Sadiraj, V. & Walker, J. M. Provision versus
471          appropriation in symmetric and asymmetric social dilemmas. *Southern*
472          *Economic Journal* **79**, 496 (2013).

473    24.    Apesteguía, J. & Maier-Rigaud, F. P. The role of rivalry. Public goods versus
474          common-pool resources. *Journal of Conflict Resolution* **50**, 646 (2006).

475    25.    Levin, S. A. Public goods in relation to competition, cooperation, and spite.
476          *Proceedings of the National Academy of Sciences* **111**, 10838 (2014).

477    26.    Brewer, M. B. & Kramer, R. M. Choice behavior in social dilemmas: Effects
478          of social identity, group size, and decision framing. *Journal of Personality and*
479          *Social Psychology* **50**, 543 (1986).

480    27.    De Dreu, C. K. & McCusker, C. Gain–loss frames and cooperation in two-
481          person social dilemmas: A transformational analysis. *Journal of Personality*
482          *and Social Psychology* **72**, 1093 (1997).

483    28.    Van Dijk, E., Wilke, H., Wilke, M. & Metman, L. What information do we
484          use in social dilemmas? Environmental uncertainty and the employment of
485          coordination rules. *Journal of Experimental Social Psychology* **35**, 109 (1999).

486    29.    McCusker, C. & Carnevale, P. J. Framing in resource dilemmas: Loss aversion
487          and the moderating effects of sanctions. *Organizational Behavior and Human*
488          *Decision Processes* **61**, 190 (1995).

489    30.    Fischbacher, U. & Gächter, S. Social preferences, beliefs, and the dynamics of
490          free riding in public good experiments. *American Economic Review.* **100**, 541
491          (2010).

492    31.    Janssen, M. A., Holahan, R., Lee, A. & Ostrom, E. Lab experiments for the
493          study of social-ecological systems. *Science* **328**, 613 (2010).

494    32.    Keser, C. & van Winden, F. Conditional cooperation and voluntary
495          contributions to public goods. *Scandinavian Journal of Economics* **102**, 23
496          (2000).

497    33.    Andreoni, J. Why free ride - strategies and learning in public-goods
498          experiments. *Journal of Public Economics* **37**, 291 (1988).

499    34.    Rustagi, D., Engel, S. & Kosfeld, M. Conditional cooperation and costly
500          monitoring explain success in forest commons management. *Science* **330**, 961
501          (2010).

502    35.    Fosgaard, T. R., Hansen, L. G. & Wengström, E. Understanding the nature of
503          cooperation variability. *Journal of Public Economics* **120**, 134 (2014).

504 36. Gintis, H. Strong reciprocity and human sociality. *Journal of Theoretical*
505 *Biology* **206**, 169 (2000).
506 37. Fehr, E. & Fischbacher, U. Social norms and human cooperation. *Trends in*
507 *Cognitive Sciences* **8**, 185 (2004).
508 38. Nyborg, K. et al. Social norms as solutions. *Science* **354**, 42 (2016).
509 39. Fehr, E. & Gächter, S. Altruistic punishment in humans. *Nature* **415**, 137
510 (2002).
511 40. Sigmund, K. Punish or perish? Retaliation and collaboration among humans.
512 *Trends in Ecology and Evolution* **22**, 593 (2007).
513 41. Cubitt, R., Drouvelis, M., Gächter, S. & Kabalin, R. Moral judgments in social
514 dilemmas: How bad is free riding? *Journal of Public Economics* **95**, 253
515 (2011).
516 42. Cox, J. C., Friedman, D. & Sadiraj, V. Revealed altruism. *Econometrica* **76**,
517 31 (2008).
518 43. Cox, J. C., Servátka, M. & Vadovič, R. Status quo effects in fairness games:
519 Reciprocal responses to acts of commission versus acts of omission.
520 *Experimental Economics* **20**, 1 (2017).
521 44. Fleishman, J. A. The effects of decision framing and others' behavior on
522 cooperation in a social dilemma. *The Journal of Conflict Resolution* **32**, 162
523 (1988).
524 45. Frackenpohl, G., Hillenbrand, A. & Kube, S. Leadership effectiveness and
525 institutional frames. *Experimental Economics* **19**, 842 (2016).
526 46. Fehr-Duda, H. & Fehr, E. Game human nature. *Nature* **530**, 413 (2016).
527 47. MacKay, D. J. C., Cramton, P., Ockenfels, A. & Stoft, S. Price carbon — I
528 will if you will. *Nature* **326**, 315 (2015).
529 48. Rand, D. G., Yoeli, E. & Hoffman, M. Harnessing reciprocity to promote
530 cooperation and the provisioning of public goods. *Policy Insights from the*
531 *Behavioral and Brain Sciences* **1**, 263 (2014).
532 49. Bowles, S. Policies designed for self-interested citizens may undermine "the
533 moral sentiments": Evidence from economic experiments. *Science* **320**, 1605
534 (2008).
535 50. Fischbacher, U., Gächter, S. & Fehr, E. Are people conditionally cooperative?
536 Evidence from a public goods experiment. *Economics Letters* **71**, 397 (2001).
537 51. Rand, D. G. et al. Positive interactions promote public cooperation. *Science*
538 **325**, 1272 (2009).
539 52. Greiner, B. Subject pool recruitment procedures: Organizing experiments with
540 ORSEE. *J Econ Sci Assoc* **1**, 114 (2015).
541 53. Fischbacher, U. Z-tree: Zurich toolbox for ready-made economic experiments.
542 *Experimental Economics* **10**, 171 (2007).
543 54. Fischbacher, U., Gächter, S. & Quercia, S. The behavioral validity of the
544 strategy method in public good experiments. *Journal of Economic Psychology*
545 **33**, 897 (2012).
546 55. Bardsley, N. & Moffatt, P. G. The experimetrics of public goods: Inferring
547 motivations from contributions. *Theory and Decision* **62**, 161 (2007).
548
549

550 **Acknowledgements**

## Author Contributions

559     SG, FK, and SQ developed the research ideas and designed the study; FK and SQ conducted

560     the experiments, and analysed data. SG, FK, and SQ wrote the manuscript.

## Additional information

562     **Supplementary information** is available for this paper.

563     **Reprints and permissions information** is available at www.nature.com/reprints.

564     **Correspondence** should be addressed to S.G.

## Competing interest

566     The authors declare no competing interests.

**FIGURE LEGENDS**

568 **Figure 1 | The isomorphic social dilemmas of maintaining and providing a public**

569 **good.** The figure illustrates initial resource allocation in Maintenance and Provision

570 prior to decision-making. **a**, Maintenance**:** Initially, group members have 0 tokens and

571 80 tokens are provided in the public good. Group members can simultaneously

572 withdraw up to 20 tokens. **b,** Provision: Initially, each group member has 20 tokens

573 and the public good is empty. Group members can simultaneously contribute up to 20

574 tokens. Each token withdrawn or not contributed is worth 1 MU to the respective

575 group member alone. Each token in the common resource is worth 0.4 MU for each

576 group member. Material incentives therefore are to withdraw 20 tokens in

577 Maintenance and to contribute 0 tokens in Provision, yielding 20 MU for each group

578 member. The socially beneficial decisions of withdrawing nothing and contributing

579 everything earn each group member 32 MU. Further details are in Methods.

580
581 **Figure 2 | Public good levels**. Shown are the effective sizes of the public good per

582 round after contribution or withdrawal decisions ($\pm$ 1 s.e.m.). **a**, One-shot game, $n_M =$

583 140, $n_P = 148$. **b**, **c**, Effective public goods over the 27 rounds of interactions in

584 randomly changing groups (Strangers, $n_M = 128$, $n_P = 128$) and fixed groups

585 (Partners, $n_M = 80$, $n_P = 80$), in Maintenance and Provision, respectively.

586 Supplementary Table 1 reports further summary statistics, including on beliefs about

587 other group members' average effective contributions (which mirror the effective

588 contributions).

589

590 **Figure 3 | Reciprocity in Maintenance and Provision. a**, Strong reciprocity as

591 measured by cooperation attitudes; type classification as in[30], $n_M = 348$, $n_P = 356$. $\chi^2$-

592 tests, *** $P < 0.01$. Results are robust to alternative classification methods

593     (Supplementary Information, Section 1.2). **b**, Estimated reciprocity in repeated

594     interactions (± 1 s.e.m.); Strangers (S), $n = 256$; Partners (P), $n = 160$ by treatment

595     and attitude category (conditional cooperators, free riders, others). Positive reciprocity

596     is estimated as the coefficient of lagged average contributions of the other group

597     members ($\bar{C}_{-ij,t-1}$) from multilevel mixed-effects linear regressions (Supplementary

598     Information, Section 1.1; Supplementary Table 3). An alternative estimation approach

599     using finite mixture models[55] confirms these results (Supplementary Table 14).

600

601     **Figure 4 | Simulated effective contribution ratios**. Distribution of 1000 simulated

602     effective contribution ratios between Maintenance and Provision ($\bar{c}^M / \bar{c}^P$) using a

603     sample of $n = 60$ per treatment and simulated experiment. The sample size reflects the

604     median sample size in related literature.[14-23] The mean is 0.91, median is 0.89, and

605     IQR = 0.76 to 1.03. Further details are in Supplementary Information, Section 1.3. As

606     a robustness check, we ran a simulation with $n = 100$, which returns a mean of 0.90, a

607     median of 0.89, and an IQR of 0.79 to 1.00 (see also Supplementary Figure 4).

608

609     **Figure 5 | Partners with Punishment. a**, Estimated negative reciprocity (± 1 s.e.m.);

610     by treatment and attitude category (conditional cooperators, free riders, others). We

611     estimate negative reciprocity in multilevel mixed-effects linear regression as the

612     number of punishment points assigned to effective contributions that deviate negatively

613     from own contribution (Supplementary Table 7; Supplementary Information). **b**,

614     Shown are the effective levels of the public goods (± 1 s.e.m.) over the 27 rounds of

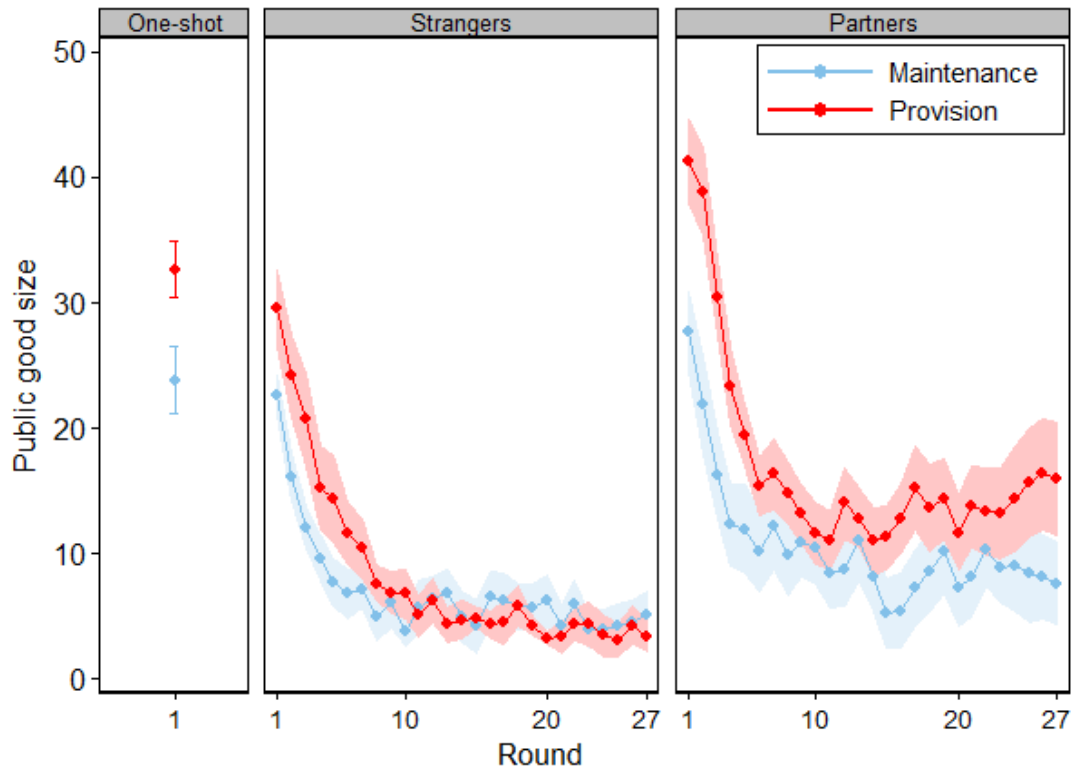615     interactions ($n_M = 84$, $n_P = 88$).

616

617

**FIGURE 1**

**a** Maintenance       **b** Provision
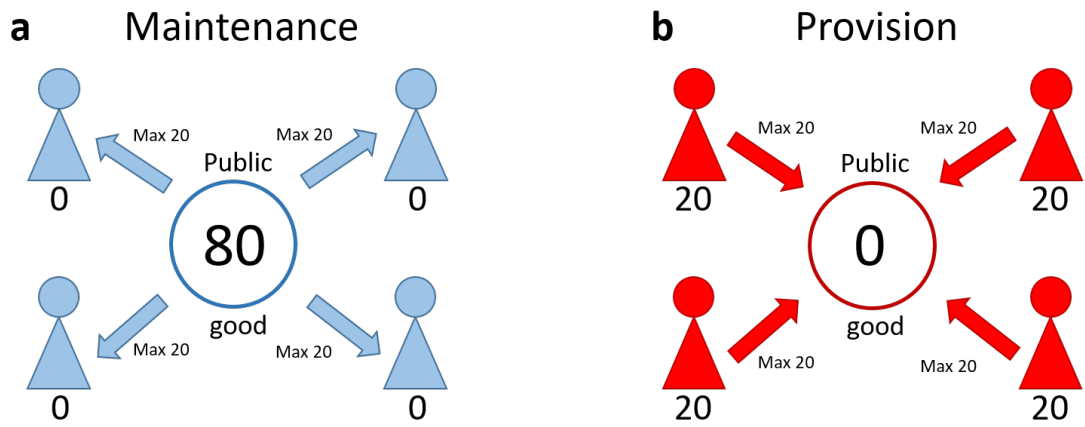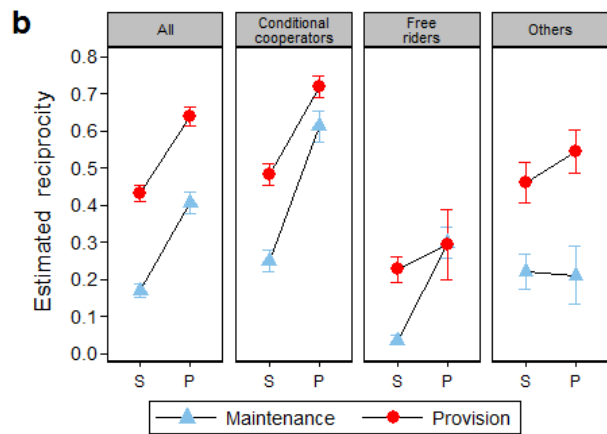
620
621
622
623
624
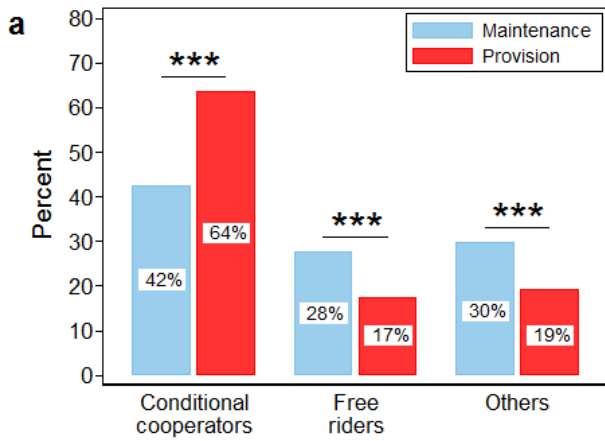625
626
627

628 **FIGURE 2**



629

630
631
632 **FIGURE 3**



633
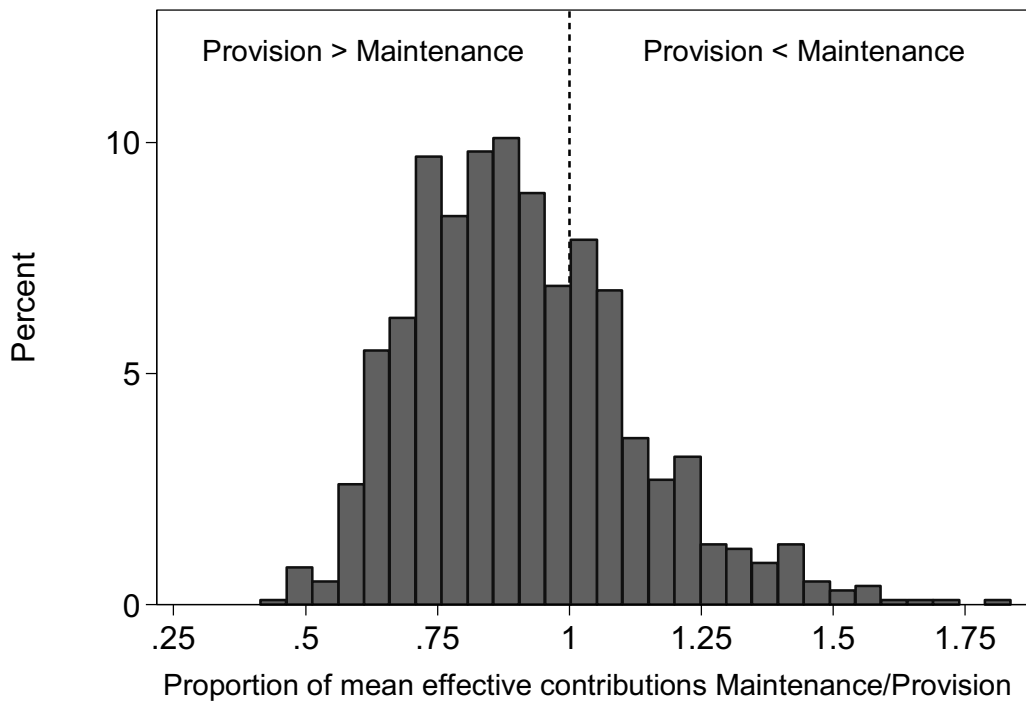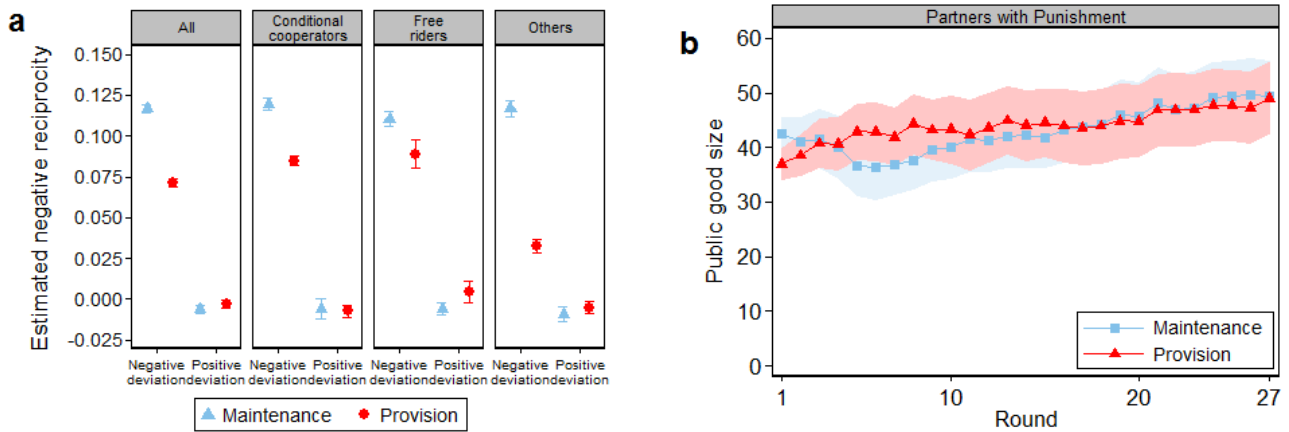
634
635
636
637

638 **FIGURE 4**



639
640
641
642
643
644 **FIGURE 5**



645

646

# 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

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:

$$PG_{jt} = \beta_0 + u_{0m} + \beta_1 Provision + \varepsilon_{jt} \tag{1}$$

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:

$$PG_{jt} = \beta_0 + u_{0m} + \beta_1 Partners + \varepsilon_{jt} \tag{2}$$

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:

$$C_{ijt} = \beta_0 + u_{0m} + u_{0i} + \beta_1 \bar{C}_{-ij,t-1} + \beta_4 Round + \varepsilon_{ijt} \tag{3}$$

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

3

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:

$$
\begin{aligned}
C_{ijt} = \ & \beta_0 + u_{0m} + u_{0i} + \ \beta_1 \bar{C}_{-ij,t-1} + \beta_2 Provision + \beta_3 \bar{C}_{-ij,t-1} \times Provision + \beta_4 Round \\
& + \beta_5 Round \times Provision + \varepsilon_{ijt}
\end{aligned}
$$

(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:

$$C_{ijt} = \beta_0 + u_{0m} + u_{0i} + \beta_1 \bar{C}_{-ij,t-1} + \beta_2 Partners + \beta_3 \bar{C}_{-ij,t-1} \times Partners + \beta_4 Round$$
$$+ \beta_5 Round \times Partners + \varepsilon_{ijt} \tag{5}$$

where *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 **Supplementary Table 4**.

Finally, we <u>compare reciprocity across attitude types in each treatment and experiment</u> **(Figure 3b, Panels 2-4)** by estimating the following model:

$$C_{ijt} = \beta_0 + u_{0m} + u_{0i} + \beta_1 \bar{C}_{-ij,t-1} + \beta_2 CC + \beta_3 \bar{C}_{-ij,t-1} \times CC + \beta_4 OT + \beta_5 \bar{C}_{-ij,t-1} \times OT$$
$$+ \beta_6 Round + \beta_7 Round \times CC + \beta_8 Round \times OT + \varepsilon_{ijt} \tag{6}$$

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 report results of these estimations in **Supplementary Table 5**.

*Comparing effective contributions across types (Supplementary Figure 1)*

To <u>compare effective contributions across attitude types in One-shot</u> **(Supplementary Figure 1a and 1b, Panel 1)** we run the following ordinary least squares (OLS) model:

$$C_i = \beta_0 + \beta_1 CC + \beta_2 OT + \varepsilon_i \tag{7}$$

where $C_i$ is the effective contribution of participant $i$ and $CC$ 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.

To <u>compare effective contributions across attitude types in Strangers and Partners</u> **(Supplementary Figure 1a and 1b, Panels 2 and 3)** we run the following linear mixed-effects model:

$$C_{ijt} = \beta_0 + u_{0m} + u_{0i} + \beta_1 CC + \beta_2 OT + \varepsilon_{ijt} \qquad (8)$$

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 <u>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</u> **(Figure 5a, all panels)**, we estimate the following linear mixed-effects model:

$$
\begin{aligned}
Y_{ijkt} = \ & \beta_0 + u_{0m} + u_{0i} + \beta_1 \max(c_{ijt} - c_{kjt}, 0) + \beta_2 \max(c_{kjt} - c_{ijt}, 0) + \beta_3 Provision \\
& + \beta_4 \max(c_{ijt} - c_{kjt}, 0) \times Provision + \beta_5 \max(c_{kjt} - c_{ijt}, 0) \times Provision \\
& + \beta_6 Round + \beta_7 Round \times Provision + \beta_8 \bar{C}_{-i-kj,t} + \beta_9 \bar{C}_{-i-kj,t} \times Provision \\
& + \varepsilon_{ijt}
\end{aligned}
$$
$$(9)$$

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:

$$
\begin{aligned}
Y_{ijkt} = \; & \beta_0 + u_{0m} + u_{0i} + \beta_1 \max(c_{ijt} - c_{kjt}, 0) + \beta_2 \max(c_{kjt} - c_{ijt}, 0) + \beta_3 CC + \beta_4 OT \\
& + \beta_5 \max(c_{ijt} - c_{kjt}, 0) \times CC + \beta_6 \max(c_{kjt} - c_{ijt}, 0) \times CC \\
& + \beta_7 \max(c_{ijt} - c_{kjt}, 0) \times OT + \beta_8 \max(c_{kjt} - c_{ijt}, 0) \times OT + \beta_9 Round \\
& + \beta_{10} Round \times CC + \beta_{11} Round \times OT + \beta_{12} \bar{C}_{-i-kjt} + \beta_{13} \bar{C}_{-i-kjt} \times CC \\
& + \beta_{14} \bar{C}_{-i-kjt} \times OT + \varepsilon_{ijt}
\end{aligned}
\tag{10}
$$

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:

$$
PG_{jt} = \beta_0 + u_{0m} + \beta_1 Provision + \varepsilon_{jt}
\tag{11}
$$

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.

To compare the public good size between Partners with Punishment and Partners, we estimate the following specification:

$$PG_{jt} = \beta_0 + u_{0m} + \beta_1 Partners\ with\ Punishment + \varepsilon_{jt} \qquad (12)$$

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. $Partners\ with\ Punishment$ 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:

$$
\begin{aligned}
C_{ij,t} - C_{ij,t-1} = {} & \beta_0 + u_{0m} + u_{0i} + \beta_1 PunReceived_{ij,t-1} + \beta_2 Provision \\
& + \beta_3 PunReceived_{ij,t-1} \times Provision + \beta_4 Round + \beta_5 Round \times Provision \\
& + \varepsilon_{ijt}
\end{aligned}
$$

$$(13)$$

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 $Provision$ and we interact the dummy with the variable $PunReceived_{ij,t-1}$. We additionally control for the time trend including the variable $Round$ as well as the interaction terms between this variable and $Provision$. 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:

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

$$(14)$$

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 *Provision* and we interact the dummy with the positive and negative deviation variables. We additionally control for the time trend including the variable *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 *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:

$$C_{ik} = \beta_0 + u_{0i} + \beta_1 \bar{C}_{-ik} + \beta_2 Provision + \beta_3 Provision \times \bar{C}_{-ik} + \varepsilon_{ik} \qquad (15)$$

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 [5]. 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 [5]:

$$d_{ij} = \sum_{k=1}^{21} |x_{ik} - x_{jk}|$$

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[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 $m = 1, \dots, M$ indexes the number of clusters and

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

where $i = 1, \dots, N$ 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.

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:     $C_{it}^* = \beta_0 + \beta_1 \bar{C}_{-i,t-1} + \beta_2 Round + \varepsilon_{it}$

STR:    $C_{it}^* = \gamma_0 + \gamma_1 Round + \nu_{it}$

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:

$$C_{it} = \begin{cases} 0 & if & C_{it}^* \leq 0 \\ C_{it}^* & if & 0 < C_{it}^* < 20 \\ 20 & if & C_{it}^* \geq 0 \end{cases}$$

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

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

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

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

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

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

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

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

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

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

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

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

For subject *i,* the likelihood function is:

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

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

$$p_{FR} \prod_{t=1}^{T} P(C_{it} = 0|FR)^{I_{C_{it}=0}} f(C_{it}|FR)^{I_{0<C_{it}<20}} P(C_{it} = 20|FR)^{I_{C_{it}=20}}$$

where $I(.)$ 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:
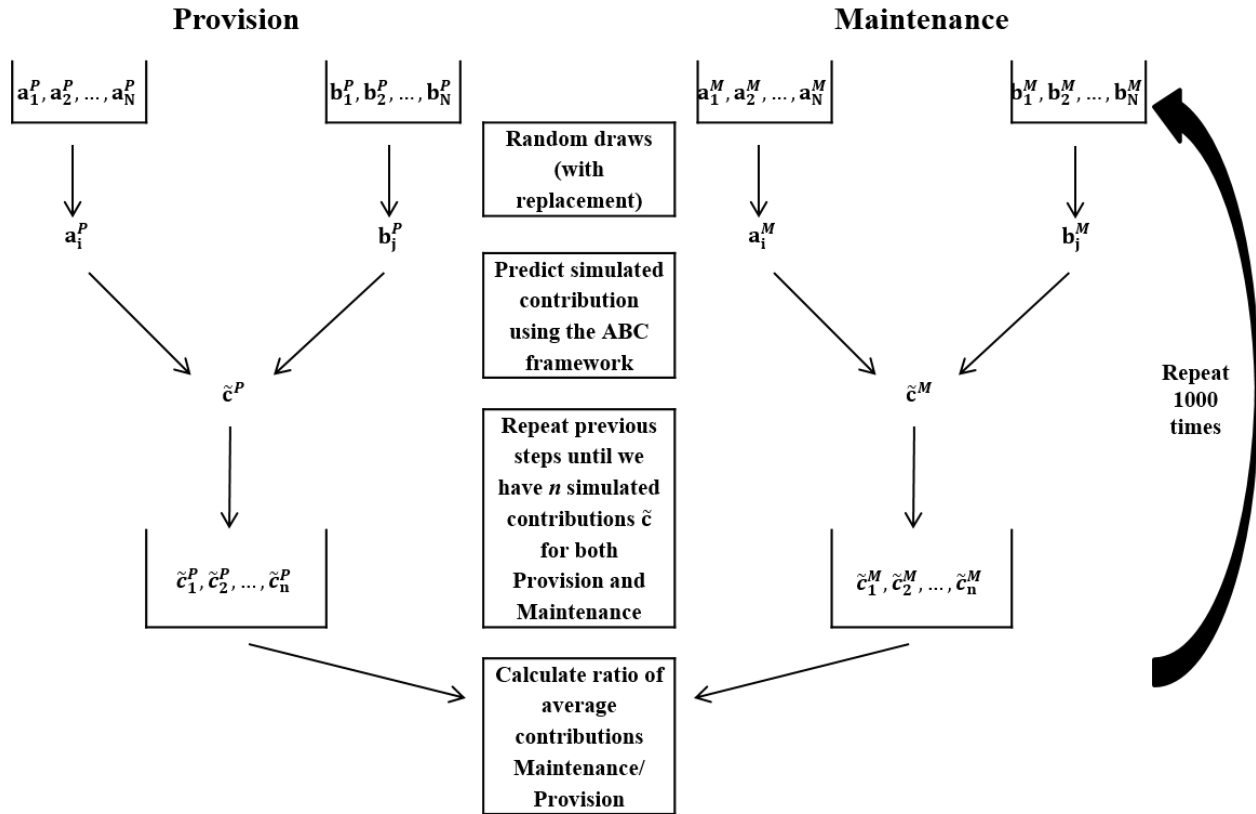
$$\tilde{c}^F = a_i^F\left(b_j^F\right) \tag{16}$$

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 $a_i^F$ and $b_j^F$, 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**.

**Provision**                                                                                       **Maintenance**

$a_1^P, a_2^P, \ldots, a_N^P$          $b_1^P, b_2^P, \ldots, b_N^P$                                $a_1^M, a_2^M, \ldots, a_N^M$          $b_1^M, b_2^M, \ldots, b_N^M$

Random draws
(with
replacement)

$a_i^P$                            $b_j^P$                                                    $a_i^M$                            $b_j^M$

Predict simulated
contribution
using the ABC
framework

Repeat
1000
times

$\tilde{c}^P$                                                                                        $\tilde{c}^M$

Repeat previous
steps until we
have $n$ simulated
contributions $\tilde{c}$
for both
Provision and
Maintenance

$\tilde{c}_1^P, \tilde{c}_2^P, \ldots, \tilde{c}_n^P$                                                  $\tilde{c}_1^M, \tilde{c}_2^M, \ldots, \tilde{c}_n^M$

Calculate ratio of
average
contributions
Maintenance/
Provision

## 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:

> ***Income from the project*** = 0.4 × (sum of contributions) [0.4 × (80 - sum of all tokens withdrawn from the project)]

If, for example, the sum of all contributions to the project [*tokens withdrawn from the project*] by you and your other group members is 60 [*20*] tokens, then you and each other member of your group would earn 60 [*80-20*] × 0.4 = 24 points out of the project. If the four members of the group contribute [*withdraw*] a total of 10 [*70*] tokens to [*from*] the project, you and the other members of your group would each earn 10 [*80-70*] × 0.4 = 4 points.

## Total income

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

> ***Your Total Income*** = *Income from your private account + Income from the project*
>
> = *20 – your contribution to the project + 0.4 × sum of all contributions to the project*
>
> [= *Tokens withdrawn from the project by you + 0.4 × (80-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. [*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. [*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. [*There are 80 tokens in the project. The other three members withdraw 30 tokens from the project.*]

a) What will your total income (in points) be, if - in addition to the 30 tokens contributed by others - you contribute 0 tokens to the project? [*What will your total income (in points) be, if - in addition to the 30 tokens withdrawn by others - you withdraw 20 tokens from the project?]*

b) What will your total income (in points) be, if - in addition to the 30 tokens contributed by others - you contribute 8 tokens to the project? [*What will your total income (in points) be, if - in addition to the 30 tokens withdrawn by others - you withdraw 12 tokens from the project?*]

c) What will your total income (in points) be, if - in addition to the 30 tokens contributed by others - you contribute 15 tokens to the project? [*What will your total income (in points) be, if - in addition to the 30 tokens withdrawn by others - you withdraw 5 tokens from the project?]*


4. Each group member has 20 tokens. Assume you invest 8 tokens to the project. [*There are 80 tokens in the project. Assume you withdraw 12 tokens from the project.*]

a) What will your total income (in points) be, if the other group members - in addition to your 8 tokens - contribute another 7 tokens to the project? [*What will your total income (in points)? be, if the other group members - in addition to your 12 tokens - withdraw another 53 tokens from the project.*]

b) What will your total income (in points) be, if the other group members - in addition to your 8 tokens - contribute another 12 tokens to the project? [*What will your total income (in points) be, if the other group members - in addition to your 12 tokens - withdraw another 48 tokens from the project?*]

c) What will your total income (in points) be, if the other group members - in addition to your 8 tokens - contribute another 22 tokens to the project? [*What will your total income (in points) be, if the other group members - in addition to your 12 tokens - withdraw another 38 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*):

Period
1   of   1

Your unconditional contribution to the project [____]

OK

Help
Please enter your unconditional contribution to the project. Press "Ok" when you are done.

After you have determined your unconditional contribution [*withdrawal*], please click "OK".

- Your second task is to fill in a "contribution [*withdrawal*] table" where you indicate how many tokens **you want to contribute [*withdraw*]** to [*from*] the project **for each possible average contribution [*withdrawal*] of the other group members** (rounded to the next integer). Here, you can condition your contribution [*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*):

Period
1   of   1

Your conditional contribution to the project (Contribution schedule)

| 0 | [____] | 7 | [____] | 14 | [____] |
| 1 | [____] | 8 | [____] | 15 | [____] |
| 2 | [____] | 9 | [____] | 16 | [____] |
| 3 | [____] | 10 | [____] | 17 | [____] |
| 4 | [____] | 11 | [____] | 18 | [____] |
| 5 | [____] | 12 | [____] | 19 | [____] |
| 6 | [____] | 13 | [____] | 20 | [____] |

OK

Help
Enter the amount which you want to contribute to the project if the others make the average contribution which stands to the left of the entry field. When you have completed your entries, press "OK".

The numbers to the left of the blue cells are the possible (rounded) average contributions [*withdrawals*] of the **other** group members to the project. You have to insert how many tokens you want to contribute to [*withdraw from*] the project into each input box – conditional on the indicated average contribution [*withdrawal*] by the other members of your group. **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 [*withdraw from*] the project if the others contribute [*withdraw*] 0 tokens on average to [*from*] the project; how much you contribute [*withdraw*] if the others contribute [*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 [*withdrawal*] and have filled in their contribution [*withdrawal*] table, a random mechanism will select one member from every group. For **this** group member, it is his **contribution [*withdrawal*] table** that will determine his actual contribution [*withdrawal*]; whereas, for the **other three** group members, it is their **unconditional contributions [*withdrawals*]** that will determine their actual contributions [*withdrawals*]. You will not know whom the random mechanism will select when you make your unconditional contribution [*withdrawal*] and fill in your contribution [*withdrawal*] table. You must therefore think carefully about both decisions because either could determine your actual contribution [*withdrawal*]. Two examples should make this clear.

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

**EXAMPLE 2:** Suppose **that the random mechanism does not select you;** and that your unconditional [*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 [*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 [*withdrawal*] table that she will contribute [*withdraw*] 1 [*19*] token[s] if the other three group members contribute [*withdraw*] on average 18 [*2*] tokens, then the total contribution of the group to the project is given by 16+18+20+1=55 [*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 0.4×55=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 [*withdrawal*] table that she contributes [*withdraws*] 19 [*1*] if the others contribute [*withdraw*] on average 18 [*2*] tokens, then the total contribution of the group to the project is 16+18+20+19=73 [*the total number of tokens left in the project is 80-(4+2+0+1)=73*] tokens. Each group member would therefore earn 0.4×73=29.2 points from the project plus their respective income from their own private account.

**The random selection of the group member whose contribution [*withdrawal*] table will determine his actual contribution [*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*].

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:

> ***Your Total Income*** *= Income from your private account + Income from the project*
>
> *=20 – your contribution to the project + 0.4 × sum of all contributions to the project*
>
> *[=Tokens withdrawn from the project by you +0.4 × (80-sum of all tokens withdrawn from the project)]*

The decision screen looks like this (*screenshot taken from the Provision treatment only*):

Your endowment 20

Your contribution to the project [ ]

What is your estimate of the average contribution from the OTHER group members in this period (rounded to an integer)? [ ]

OK

Help
Press "OK" when you made your entries.

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.

2) *Strangers*

<div align="center">**Instructions**</div>

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:

<div align="center">**1 point=£0.2**</div>

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:

---

*Your Total Income* = Income from your private account + Income from the project

= 20 – your contribution to the project + 0.4 × sum of all contributions to the project

[=Tokens withdrawn from the project by you +0.4 × (80-sum of all tokens withdrawn from the project)]

---

The decision screen looks like this (*screenshot taken from the Provision treatment only*):

Your endowment        20

Your contribution to the project    [   ]

What is your estimate of the average contribution from the OTHER group members in this period (rounded to an integer)?    [   ]

OK

Help
Press "OK" when you made your entries.

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.

27

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:

> ***Your Total Income*** *= Income from your private account + Income from the project*
>
> *=20 – your contribution to the project + 0.4 × sum of all contributions to the project*
>
> *[=Tokens withdrawn from the project by you +0.4 × (80-sum of all tokens withdrawn from the project)]*

The decision screen looks like this (*screenshot taken from the Provision treatment only*):

Your endowment 20

Your contribution to the project

What is your estimate of the average contribution from the OTHER group members in this period (rounded to an integer)?

OK

Help
Press "OK" when you made your entries.

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:**

*Your Total Income = Income from your private account + Income from the project*

*= 20 – your contribution to the project + 0.4 × sum of all contributions to the project*

*[=Tokens withdrawn from the project by you +0.4 × (80-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:

**Costs of received deduction points = 3 × 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:

---

EITHER

Your income from the first stage is greater than or equal to the cost of received deduction points:

**Total income** = Income from the first stage – 3 × (sum of received deduction points) –
– sum of deduction points you have assigned

OR

Your income from the first stage is less than the cost of received deduction points:

**Total income = 0** – sum of deduction points you have assigned

---

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

1.  Fischbacher, U., Gächter, S. & Fehr, E. Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters* **71**, 397 (2001).
2.  Fischbacher, U. & Gächter, S. Social preferences, beliefs, and the dynamics of free riding in public good experiments. *American Economic Review.* **100**, 541 (2010).
3.  Fischbacher, U., Gächter, S. & Quercia, S. The behavioral validity of the strategy method in public good experiments. *Journal of Economic Psychology* **33**, 897 (2012).
4.  Smith, A. Estimating the causal effect of beliefs on contributions in repeated public good games. *Experimental Economics* **16**, 414 (2013).
5.  Everitt, B., Landau, S. & Leese, M. Cluster analysis. 2001. *Arnold, London* (2001).
6.  Bardsley, N. & Moffatt, P. G. The experimetrics of public goods: Inferring motivations from contributions. *Theory and Decision* **62**, 161 (2007).

# 3. Supplementary Tables

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

**Panel A:** Public Good Size

| | One-shot | | | Strangers | | | Partners | | |
|---|---|---|---|---|---|---|---|---|---|
| | M | P | | M | P | | M | P | |
| Round 1 | 23.83 (16.00) | 32.65 (13.81) | P = 0.014 | 22.69 (13.48) | 29.53 (14.93) | P = 0.059 | 27.70 (14.07) | 41.30 (14.94) | P = 0.005 |
| Round 1-9 | | | | 10.39 (9.61) | 15.63 (10.73) | P = 0.053 | 14.83 (14.85) | 23.69 (12.58) | P = 0.008 |
| Round 10-18 | | | | 5.65 (9.06) | 5.21 (6.49) | P = 0.831 | 8.18 (13.17) | 12.63 (12.38) | P = 0.177 |
| Round 19-27 | | | | 4.88 (8.31) | 3.77 (6.02) | P = 0.541 | 8.69 (14.57) | 14.31 (16.79) | P = 0.177 |
| All rounds | 23.83 (16.00) | 32.65 (13.81) | P = 0.014 | 6.97 (8.99) | 8.20 (7.75) | P = 0.503 | 10.56 (14.20) | 16.87 (13.91) | P = 0.035 |

**Panel B:** Beliefs

| | One-shot | | | Strangers | | | Partners | | |
|---|---|---|---|---|---|---|---|---|---|
| | M | P | | M | P | | M | P | |
| Round 1 | 7.76 (5.82) | 9.62 (5.32) | P = 0.005 | 6.70 (5.92) | 8.71 (5.77) | P = 0.006 | 7.03 (6.28) | 11.06 (5.91) | P < 0.001 |
| Round 1-9 | | | | 3.17 (3.13) | 5.66 (4.07) | P = 0.002 | 4.13 (4.51) | 7.49 (4.32) | P < 0.001 |
| Round 10-18 | | | | 1.04 (2.05) | 1.72 (2.02) | P = 0.232 | 1.98 (3.36) | 3.61 (3.65) | P = 0.059 |
| Round 19-27 | | | | 0.82 (1.83) | 1.02 (1.82) | P = 0.632 | 1.94 (3.46) | 3.81 (4.56) | P = 0.079 |
| All rounds | 7.76 (5.82) | 9.62 (5.32) | P = 0.005 | 1.68 (2.34) | 2.80 (2.64) | P = 0.030 | 2.68 (3.78) | 4.97 (4.18) | P = 0.004 |

**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)**

| | Maintenance vs. Provision | | Strangers vs. Partners | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Strangers | Partners | Maintenance | Provision |
| Provision | 1.230 | 6.309** | | |
| *1 if Provision, 0 otherwise* | (1.837) | (2.987) | | |
| | | | | |
| Partners | | | 3.591 | 8.670*** |
| *1 if Partners, 0 otherwise* | | | (3.781) | (2.999) |
| | | | | |
| Constant | 6.973*** | 10.56*** | 6.973** | 8.204*** |
| | (1.299) | (2.112) | (3.188) | (2.528) |
| N | 432 | 1080 | 756 | 756 |

**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)

**Panel A:** Strangers

|  | (1)<br>All | (2)<br>CC | (3)<br>FR | (4)<br>OT |
|---|---|---|---|---|
| Provision<br>1 if Provision, 0 otherwise | -0.019<br>(0.383) | 0.302<br>(0.539) | -0.706**<br>(0.283) | -0.197<br>(1.079) |
| $\bar{C}_{-ij,t-1}$ | 0.172***<br>(0.020) | 0.254***<br>(0.033) | 0.038*<br>(0.021) | 0.222***<br>(0.046) |
| Provision × $\bar{C}_{-ij,t-1}$ | 0.257***<br>(0.028) | 0.228***<br>(0.043) | 0.190***<br>(0.033) | 0.238***<br>(0.073) |
| Round | -0.044***<br>(0.007) | -0.069***<br>(0.011) | -0.003<br>(0.007) | -0.056***<br>(0.016) |
| Provision × Round | -0.022**<br>(0.010) | -0.017<br>(0.015) | 0.045***<br>(0.012) | -0.093***<br>(0.026) |
| Constant | 1.920***<br>(0.265) | 2.191***<br>(0.398) | 0.138<br>(0.178) | 3.589***<br>(0.676) |
| *N* | 6656 | 3172 | 1898 | 1586 |

**Panel B:** Partners

|  | (1)<br>All | (2)<br>CC | (3)<br>FR | (4)<br>OT |
|---|---|---|---|---|
| Provision<br>*1 if Provision, 0 otherwise* | -0.407<br>(0.549) | 0.171<br>(0.636) | -0.623<br>(1.212) | -1.576<br>(1.498) |
| $\bar{C}_{-ij,t-1}$ | 0.420***<br>(0.030) | 0.621***<br>(0.043) | 0.292***<br>(0.048) | 0.202***<br>(0.075) |
| Provision × $\bar{C}_{-ij,t-1}$ | 0.217***<br>(0.039) | 0.096*<br>(0.052) | -0.002<br>(0.090) | 0.339***<br>(0.097) |
| Round | -0.027**<br>(0.012) | -0.001<br>(0.016) | -0.024<br>(0.020) | -0.103***<br>(0.028) |
| Provision × Round | 0.024<br>(0.017) | -0.023<br>(0.021) | 0.069*<br>(0.040) | 0.122***<br>(0.040) |
| Constant | 1.741***<br>(0.381) | 1.188**<br>(0.479) | 0.919<br>(0.608) | 3.804***<br>(1.052) |
| *N* | 4160 | 2366 | 884 | 910 |

**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)**

| | (1) All Maintenance | (2) All Provision | (3) CC Maintenance | (4) CC Provision | (5) FR Maintenance | (6) FR Provision | (7) OT Maintenance | (8) OT Provision |
|---|---|---|---|---|---|---|---|---|
| Partners | -0.131 | -0.557 | -0.938 | -1.120$^{**}$ | 0.780 | 0.931 | 0.221 | -1.165 |
| 1 if Partners, 0 otherwise | (0.514) | (0.427) | (0.700) | (0.487) | (0.487) | (1.214) | (1.326) | (1.192) |
| | | | | | | | | |
| $\bar{C}_{-ij,t-1}$ | 0.170$^{***}$ | 0.430$^{***}$ | 0.252$^{***}$ | 0.484$^{***}$ | 0.040 | 0.235$^{***}$ | 0.222$^{***}$ | 0.460$^{***}$ |
| | (0.022) | (0.025) | (0.035) | (0.032) | (0.026) | (0.044) | (0.052) | (0.060) |
| | | | | | | | | |
| Partners × $\bar{C}_{-ij,t-1}$ | 0.237$^{***}$ | 0.206$^{***}$ | 0.348$^{***}$ | 0.233$^{***}$ | 0.254$^{***}$ | 0.062 | -0.022 | 0.081 |
| | (0.033) | (0.033) | (0.051) | (0.041) | (0.039) | (0.077) | (0.086) | (0.079) |
| | | | | | | | | |
| Round | -0.044$^{***}$ | -0.066$^{***}$ | -0.069$^{***}$ | -0.086$^{***}$ | -0.003 | 0.043$^{***}$ | -0.056$^{***}$ | -0.149$^{***}$ |
| | (0.007) | (0.009) | (0.012) | (0.012) | (0.009) | (0.016) | (0.018) | (0.022) |
| | | | | | | | | |
| Partners × Round | 0.016 | 0.063$^{***}$ | 0.066$^{***}$ | 0.061$^{***}$ | -0.020 | 0.004 | -0.047 | 0.168$^{***}$ |
| | (0.012) | (0.014) | (0.018) | (0.017) | (0.015) | (0.033) | (0.031) | (0.032) |
| | | | | | | | | |
| Constant | 1.925$^{***}$ | 1.896$^{***}$ | 2.195$^{***}$ | 2.483$^{***}$ | 0.126 | -0.725 | 3.588$^{***}$ | 3.391$^{***}$ |
| | (0.338) | (0.266) | (0.479) | (0.328) | (0.339) | (0.742) | (0.751) | (0.774) |
| 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. $\bar{C}_{-ij,t-1}$ is the average effective contribution of the other three group members from the previous round. Random intercepts are included for matching groups and individuals. 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)**

| | (1) Strangers Maintenance | (2) Strangers Provision | (3) Partners Maintenance | (4) Partners Provision |
|---|---|---|---|---|
| CC | 2.076*** | 3.003*** | 0.231 | 1.077 |
| *1 if CC, 0 otherwise* | (0.679) | (0.466) | (0.744) | (1.392) |
| | | | | |
| OT | 3.450*** | 3.924*** | 2.955*** | 1.969 |
| *1 if OT, 0 otherwise* | (0.739) | (0.592) | (0.897) | (1.568) |
| | | | | |
| $\bar{C}_{-ij,t-1}$ | 0.040 | 0.228*** | 0.285*** | 0.300*** |
| | (0.031) | (0.044) | (0.050) | (0.084) |
| | | | | |
| $CC \times \bar{C}_{-ij,t-1}$ | 0.208*** | 0.258*** | 0.320*** | 0.419*** |
| | (0.042) | (0.052) | (0.066) | (0.089) |
| | | | | |
| $OT \times \bar{C}_{-ij,t-1}$ | 0.181*** | 0.232*** | -0.087 | 0.239** |
| | (0.046) | (0.065) | (0.082) | (0.100) |
| | | | | |
| Round | -0.003 | 0.042*** | -0.025 | 0.048 |
| | (0.011) | (0.016) | (0.021) | (0.038) |
| | | | | |
| CC × Round | -0.066*** | -0.127*** | 0.022 | -0.072* |
| | (0.014) | (0.019) | (0.026) | (0.041) |
| | | | | |
| OT × Round | -0.053*** | -0.191*** | -0.079** | -0.029 |
| | (0.016) | (0.024) | (0.032) | (0.046) |
| | | | | |
| Constant | 0.139 | -0.533 | 0.997 | 0.274 |
| | (0.508) | (0.386) | (0.606) | (1.298) |
| *N* | 5512 | 5616 | 2080 | 2080 |

**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. $\bar{C}_{-ij,t-1}$ 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)**

|  | (1)<br>One-shot<br>Maintenance | (2)<br>One-shot<br>Provision | (3)<br>Strangers<br>Maintenance | (4)<br>Strangers<br>Provision | (5)<br>Partners<br>Maintenance | (6)<br>Partners<br>Provision |
|---|---|---|---|---|---|---|
| CC | 7.521*** | 8.318*** | 1.611** | 1.872*** | 1.503*** | 1.923* |
| *1 if CC, 0 otherwise* | (1.396) | (1.437) | (0.634) | (0.345) | (0.584) | (1.088) |
| | | | | | | |
| OT | 6.464*** | 6.105*** | 3.159*** | 1.817*** | 1.846*** | 2.630** |
| *1 if OT, 0 otherwise* | (1.435) | (1.785) | (0.691) | (0.427) | (0.702) | (1.218) |
| | | | | | | |
| Constant | 0.379 | 1.455 | 0.226 | 0.643* | 1.535** | 2.353** |
| | (1.146) | (1.302) | (0.509) | (0.346) | (0.686) | (1.052) |
| *N* | 140 | 148 | 3456 | 3456 | 2160 | 2160 |

**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)**

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

**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. $\bar{C}_{-i-k,jt}$ 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. *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$.

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

| | (1)<br>Maintenance | (2)<br>Provision |
|---|---|---|
| CC | -0.056 | -0.382*** |
| *1 if CC, 0 otherwise* | (0.106) | (0.096) |
| OT | -0.166 | -0.237** |
| *1 if OT, 0 otherwise* | (0.109) | (0.108) |
| Positive deviation from $c_i$ | -0.006 | 0.005 |
| *max($c_k$ - $c_i$, 0)* | (0.004) | (0.005) |
| Positive deviation from $c_i$ × CC | 0.000 | -0.012* |
| | (0.008) | (0.007) |
| Positive deviation from $c_i$ × OT | -0.002 | -0.009 |
| | (0.006) | (0.007) |
| Negative deviation from $c_i$ | 0.111*** | 0.089*** |
| *max($c_i$ – $c_k$, 0)* | (0.005) | (0.006) |
| Negative deviation from $c_i$ × CC | 0.008 | -0.004 |
| | (0.006) | (0.007) |
| Negative deviation from $c_i$ × OT | 0.006 | -0.056*** |
| | (0.006) | (0.008) |
| Round | -0.009*** | -0.008*** |
| | (0.002) | (0.002) |
| Round × CC | 0.005** | 0.007*** |
| | (0.003) | (0.002) |
| Round × OT | 0.006** | -0.001 |
| | (0.003) | (0.003) |
| $\bar{C}_{-i-kjt}$ | 0.006 | -0.016*** |
| | (0.004) | (0.004) |
| $\bar{C}_{-i-kjt}$ × CC | -0.009 | 0.017*** |
| | (0.006) | (0.005) |
| $\bar{C}_{-i-kjt}$ × OT | 0.014** | 0.026*** |
| | (0.006) | (0.006) |
| Constant | 0.171** | 0.458*** |
| | (0.077) | (0.087) |
| N | 6804 | 7128 |

**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. *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$.

**Supplementary Table 9 (support for Figure 5b)**

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

**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)**

| | Maintenance vs. Provision | Partners vs. Partners with Punishment | |
|---|---|---|---|
| | (1) | (2) | (3) |
| | | Maintenance | Provision |
| Provision | 0.902 | | |
| *1 if Provision, 0 otherwise* | (7.501) | | |
| | | | |
| Partners with Punishment | | 32.59*** | 27.18*** |
| *1 if Partners with Punishment, 0 otherwise* | | (5.826) | (5.898) |
| | | | |
| Constant | 43.15*** | 10.56** | 16.87*** |
| | (5.365) | (4.169) | (4.269) |
| *N* | 1161 | 1107 | 1134 |

**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

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

**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

|  | (1) |
|---|---|
| Positive deviation from $c_i$ in *t-1* <br> *$max(\bar{c}_{-i,t-1} - c_{i,t-1}, 0)$* | $0.172^{***}$ <br> (0.04) |
| Negative deviation from $c_i$ in *t-1* <br> *$max(c_{i,t-1} - \bar{c}_{-i,t-1}, 0)$* | $-0.554^{***}$ <br> (0.05) |
| Positive deviation from $c_i$ in *t-1* $\times$ Provision | $0.126^{**}$ <br> (0.05) |
| Negative deviation from $c_i$ in *t-1* $\times$ Provision | $-0.154^{**}$ <br> (0.06) |
| Lagged own contribution $c_{i,t-1}$ | $0.814^{***}$ <br> (0.04) |
| Round | -0.005 <br> (0.01) |
| Lagged own contribution $c_{i,t-1}$ $\times$ Provision | 0.048 <br> (0.05) |
| Round $\times$ Provision | 0.010 <br> (0.02) |
| Provision <br> *1 if Provision, 0 otherwise* | 0.116 <br> (0.43) |
| Constant | $0.913^{***}$ <br> (0.29) |
| N | 4160 |

**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. *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$.

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

|  | (1) |
|---|---|
| Avg. contribution others | 0.333*** |
|  | (0.008) |
| Provision | -1.796*** |
| *1 if Provision, 0 otherwise* | (0.403) |
| Avg. contribution others × Provision | 0.176*** |
|  | (0.011) |
| Constant | 3.057*** |
|  | (0.286) |
| N | 14784 |

**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)**

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

**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. *** $P < 0.01$, ** $P < 0.05$, * $P < 0.1$.
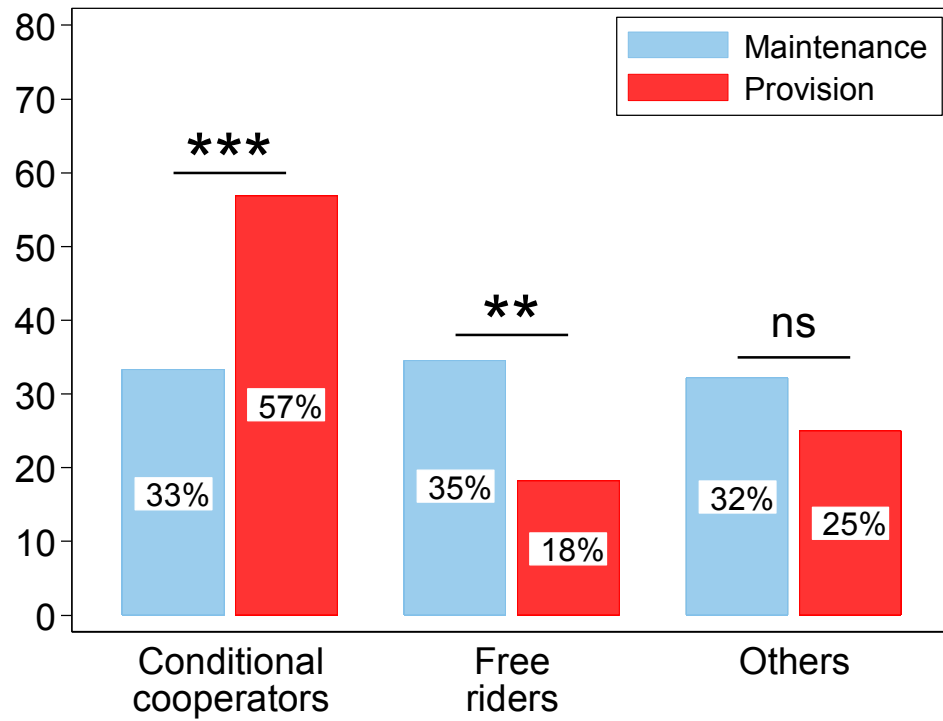
# 4. Supplementary Figures

**Supplementary Figure 1**
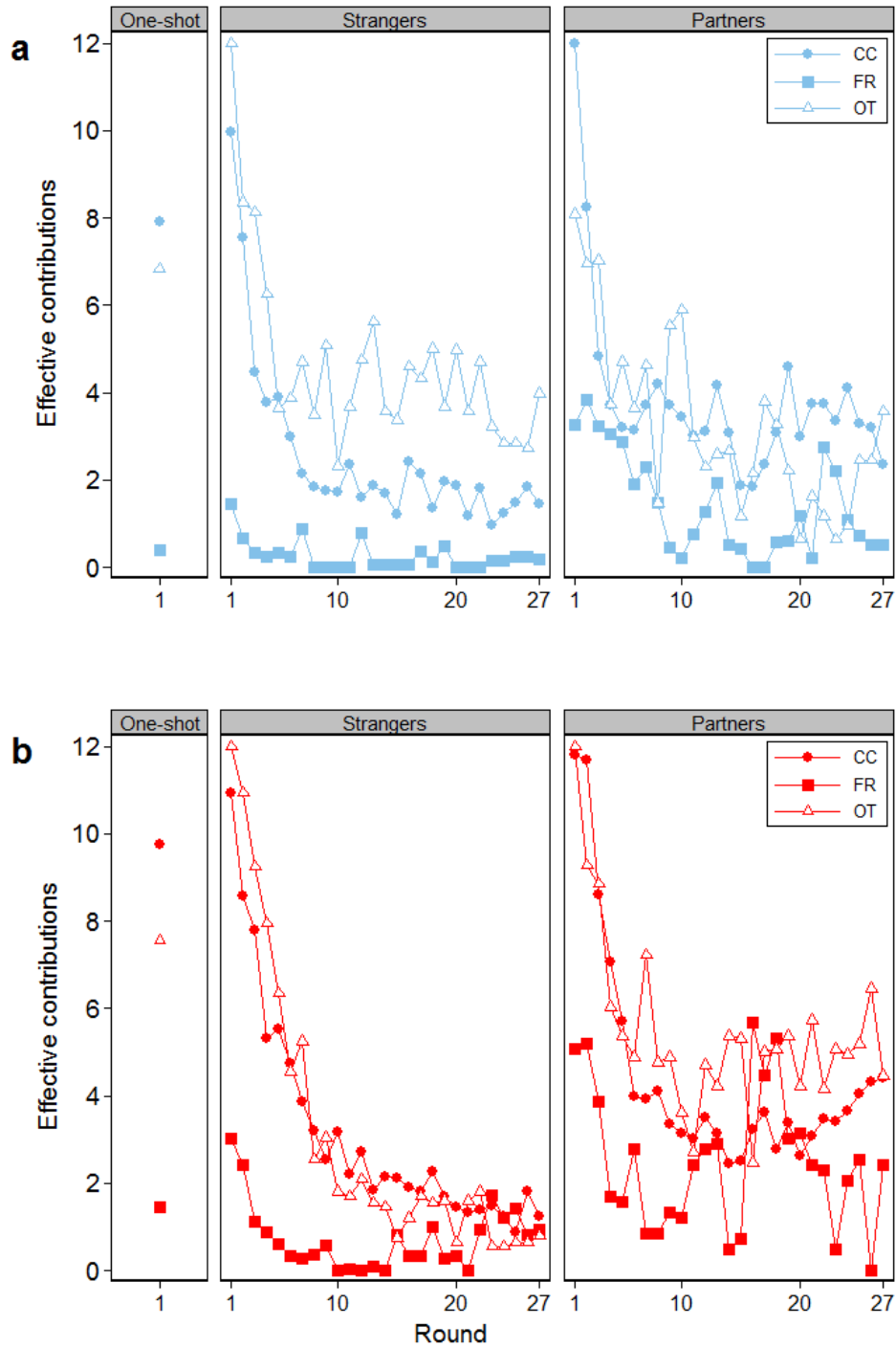


**Supplementary Figure 1 | Effective average contributions by attitude type and type of strategic interaction. A,**

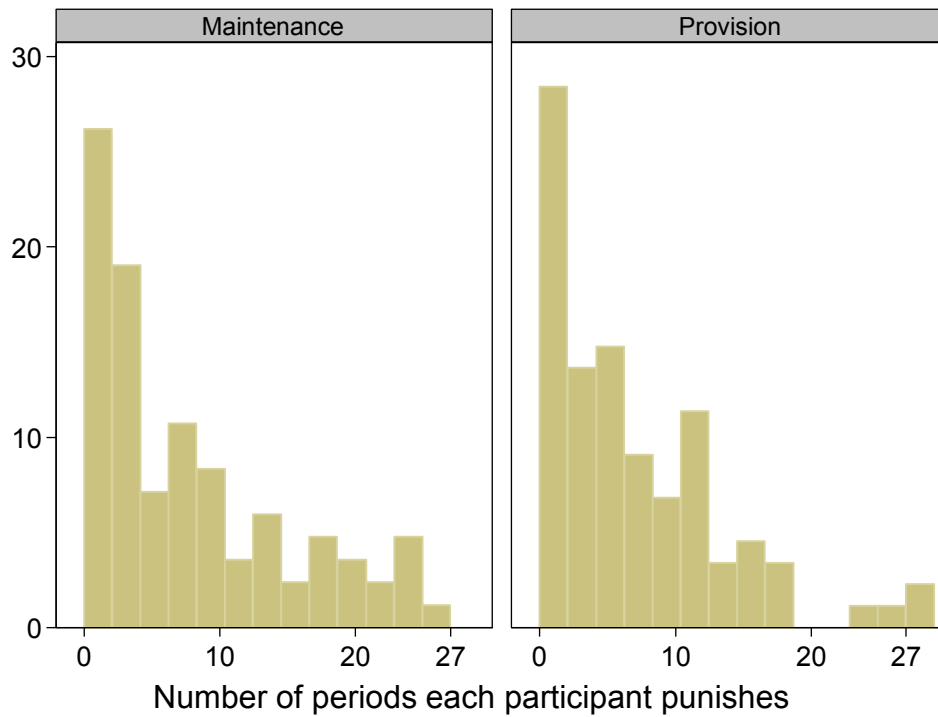Maintenance. **B,** Provision. CC: Conditional Cooperators, FR: Free Riders, OT: Others.

**Supplementary Figure 2 | 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) $P > 0.10$.

**Supplementary Figure 3**



**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**
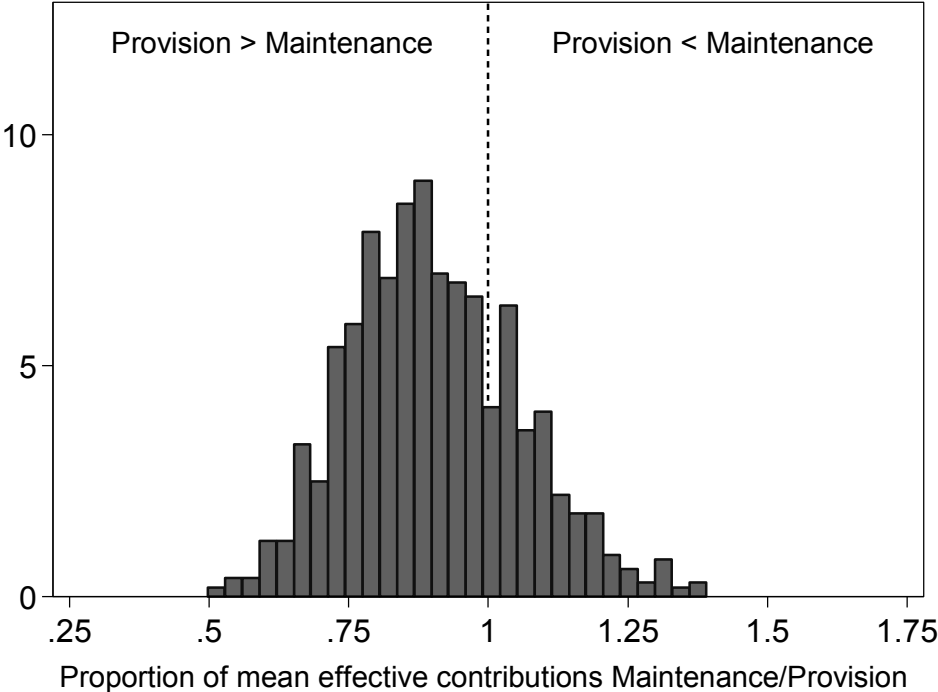


**Supplementary Figure 4 | Simulated effective contribution ratios**. Distribution of 1000 simulated effective contribution ratios between Maintenance and Provision ($\bar{c}^M / \bar{c}^P$) using a sample of $n = 100$ per treatment and simulated experiment.