

Avoiding or Restricting Defectors in Public Goods Games?

The Anh Han ^{$\alpha, \beta, \gamma, *$} , Luís Moniz Pereira ^{λ} and Tom Lenaerts ^{$\alpha, \beta, *$}

November 29, 2014

^{α} AI lab, Computer Science Department, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

^{β} MLG, Département d'Informatique, Université Libre de Bruxelles, Boulevard du Triomphe CP212, 1050 Brussels, Belgium

^{γ} School of Computing, Teesside University, Borough Road, Middlesbrough, UK TS1 3BA (present address)

^{λ} Centro de Inteligência Artificial (CENTRIA), Departamento de Informática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

*corresponding authors: T.Han@tees.ac.uk, Tom.Lenaerts@ulb.ac.be

Abstract

When creating a public good, strategies or mechanisms are required to handle defectors. We first show mathematically and numerically that prior agreements with posterior compensations provide a strategic solution that leads to substantial levels of cooperation in the context of Public Goods games, results that are corroborated by available experimental data. Notwithstanding this success, one cannot, as with other approaches, fully exclude the presence of defectors, raising the question of how they can be dealt with to avoid the demise of the common good. We show that both avoiding creation of the common good, whenever full agreement is not reached, and limiting the benefit that disagreeing defectors can acquire, using costly restriction mechanisms, are relevant choices. Nonetheless, restriction mechanisms are found the more favorable, especially in larger group interactions. Given decreasing restriction costs, introducing restraining measures to cope with public goods free-riding issues is the ultimate advantageous solution for all participants, rather than avoiding its creation.

Keywords: evolutionary games, cooperation, commitment, public goods.

1 Introduction

Arranging a prior commitment or agreement is an essential ingredient to encourage cooperative behavior in a wide range of relationships, ranging from personal to political and religious ones¹⁻⁵. Prior agreements clarify the intentions and preferences of other players. Hence, refusing to establish an agreement may be considered as intending or preferring not to cooperate (non-committers)⁵⁻⁷. Prior agreements may be highly rewarding in group situations, as in the case of Public Goods Games (PGG)⁸, as it forces the other participants to signal their willingness to achieve a common goal. Especially for increasing group sizes, such prior agreements could be ultimately rewarding, as it becomes more and more difficult to assess the aspirations of all participants.

In a PGG, where players meet in groups of size N ^{9,10}, all players can choose whether to cooperate and contribute an amount, c , to the public good or to defect and take advantage of the public good without contributing to it. The total contribution is multiplied by a constant public goods producing factor, $r > 1$, and the result is afterwards distributed equally among all players. With r smaller than the group size ($r < N$), non-contributing free-riders always gain more than contributors. Evolutionary game dynamics has shown that under those conditions cooperation disappears, which is famously known as the ‘tragedy of commons’^{10,11}. Various mechanisms, such as direct and indirect reciprocity, kin and group selections, and costly punishment, have been proposed and evaluated both theoretically and experimentally, which explain the evolution of cooperation nevertheless^{10,12,13}, ranging from microbial systems to animals and humans societies^{12,14-18}.

Here, we examine a strategic solution based on prior agreements to address the problem of the evolution of cooperation in PGG. Prior to the PGG, commitment proposing players ask their co-players to commit to contribute to the PGG, paying a personal proposer’s cost to establish

25 that agreement. If all the requested co-players accept the commitment, the proposers assume
26 that everyone will contribute to the public good. Those individuals that commit yet later do
27 not contribute receive a penalty and are forced to compensate the proposers at a cost^{6,19,20}. As
28 such, our model explicitly and novelly addresses the relevance of the commitment proposing
29 behavior regarding posterior compensations in group interactions, which has been suggested to
30 be a major pathway to the emergence of cooperation^{1,5}.

31 As commitment proposers may encounter also non-committers, they require strategies that
32 can deal with this kind of individuals^{1,6,21}. When dealing with non-committers, the simplest
33 strategy is to not participate in the creation of the common good or, when the interaction is
34 mandatory, to simply not contribute, i.e. defect^{9,22}, when not everyone commits. Yet, this
35 *avoidance* strategy also removes the benefits for those that wished to establish the public good,
36 hindering any advancements they could harness from this novel resource. Alternatively, one
37 can try to establish boundaries on the common good so that only those that commit to make
38 it work have access or that the benefit non-contributors can acquire from the common good is
39 reduced, as is the case for food sharing, aid in social health and defence against predators^{23–25}.
40 An extreme case of exercising restriction is ostracism, which can be enforced through financial
41 or social means^{26,27}. Experimental studies with PGG have shown that the threat of excluding
42 or ostracizing non-cooperative members from the PGG can significantly increase contribution
43 and cooperation^{28,29}. As public goods are by definition non-excludable^{8,30}, ostracism may not
44 be possible and non-committers may only be excluded to a certain degree. Moreover, a cost
45 may be associated with this *exclusion* strategy, where the capacity to ostracize may be too
46 costly. Evidence regarding restriction abounds in biological and social contexts: Animals fence
47 and defend territory and resources²⁵. Trade restrictions against non-participating countries are
48 widely implemented in international and environmental treaties^{8,31}, yet may be circumvented.

49 While showing the relevance of our prior conclusions on commitment obtained for the pair-

50 wise prisoner's dilemma (PD)²⁰ within the context of the more complex PGG, we focus here
51 on showing mathematically and numerically how best to deal with individuals that do not wish
52 to make prior agreements and do not contribute to the common good. This issue is not only
53 essential in the general discussion of the PGG, it is also fundamental in case of the strategic
54 commitment behavior since we observed in the PD that the number of non-committers, to-
55 gether with those that free-ride on the investment of committers, increase markedly with the
56 increase of the cost of setting up the commitment²⁰. We will examine under which conditions
57 avoidance, which is a generalization of the PD commitment behavior towards the PGG, and re-
58 stricting strategies are beneficial in the PGG, determining at the same time the condition when
59 the latter strategy is preferred over the former. The effect of the different parameters implicit
60 to the strategies on their viability is carefully analyzed. Interestingly, we will show that group
61 size is an important factor in determining the conditions for which restriction may be better than
62 simply avoiding non-committers.

63 **2 Results**

64 **2.1 Commitment strategies in PGG**

65 Commitment strategies can propose a commitment deal to all members of the group before
66 playing the PGG. The proposer(s) share the cost ϵ_P , while those who do not (but still can
67 join the commitment) pay nothing. If all the requested co-players agree to the commitment,
68 they are assumed to contribute to the public good. Those that commit though later do not
69 contribute have to compensate the commitment proposers at their personal cost δ , and their
70 compensation is shared among all proposers. Additionally, there may be group members that
71 refuse the agreement, wishing to play the game without any prior commitment. As refusal
72 may be conceived as a future defection (no contribution), proposers may wish to either avoid

73 interacting with them or set up some mechanism that restricts their access to the public good.
 74 We define those strategies as follows:

- 75 • **AVOID**: refuses to play the game when there are non-committers in the group (hence, the
 76 PGG does not take place and each player receives 0 payoff).
- 77 • **RESTRICT**: sets up, at an extra cost ϵ_R , a mechanism to restrict the access of the non-
 78 committers to the public good. This restriction is modeled through a factor, $\psi < 1$,
 79 representing the fraction of the common resource the non-committers receive compared
 80 to the committed players (i.e. the smaller ψ the greater the effect).

81 Next to the traditional unconditional contributors (C, who always commit when being proposed
 82 a commitment deal, contribute whenever the PGG is played, but do not propose commitment)
 83 and unconditional non-contributors (D, who do not accept commitment, defect when the PGG is
 84 played, and do not propose commitment), we consider two commitment free-riding strategies,
 85 which we have shown to become dominant under certain conditions in the pair-wise PD situa-
 86 tion²⁰: (i) fake committers (FAKE), who accept a commitment proposal yet do not contribute
 87 whenever the PGG is played. These players assume that they can exploit the commitment
 88 proposing players without suffering the consequences; (ii) commitment free-riders (FREE),
 89 who defect unless being proposed a commitment, which they then accept and cooperate subse-
 90 quently in the PGG. In other words, these players are willing to contribute when a commitment
 91 is proposed but are not prepared to pay the cost of setting it up.

92 We consider a well-mixed, finite population of a constant size Z , potentially composed
 93 of those five strategies, i.e. AVOID (or RESTRICT), C, FREE, D, and FAKE. To simplify the
 94 notations, the strategies are numerated 1, 2, 3, 4, and 5, respectively (where 1 can either indicate
 95 the AVOID or RESTRICT strategy). In each interaction, N individuals are randomly selected
 96 from the population for playing the PGG. Among N randomly selected players, the number

of players in the group of size N using strategy i is denoted by N_i , $i = 1, \dots, 5$, such that
 $N = N_1 + N_2 + N_3 + N_4 + N_5$.

We compute the payoffs of either the AVOID or RESTRICT strategy in the population in relation to the other strategies (see Methods and Supporting Information (SI)). In case of the AVOID strategy, if there is a D player in the group, i.e. $N_4 \geq 1$, the game is not played and every player in the group obtains 0 (the results will be unchanged if the game is not optional, as in that case, AVOID players contribute nothing to the public good). In case of the RESTRICT strategy, the game is always played and RESTRICT players share an additional cost, $\epsilon_R > 0$, to restrict the benefits D players can obtain from the public good. The committing player (RESTRICT, C, FREE and FAKE) and non-committing player (D) in the PGG gain respectively $\frac{r(N_1+N_2+N_3)}{N_1+N_2+N_3+\psi N_4+N_5}c$ and $\frac{r(N_1+N_2+N_3)\psi}{N_1+N_2+N_3+\psi N_4+N_5}c$. The gain for a RESTRICT player is reduced by $c + \frac{1}{N_1}\epsilon_P - \frac{N_5}{N_1}\delta$, if $N_4 = 0$, and by $c + \frac{1}{N_1}(\epsilon_P + \epsilon_R) - \frac{N_5}{N_1}\delta$, otherwise. The payoff for a C and a FREE player is reduced by c . Finally, the payoff for a FAKE player is reduced by δ . The detailed calculation of the payoff for each strategy as well as the payoff matrix is provided in SI.

2.2 Constraints for viability of AVOID and RESTRICT

We derive the conditions for which commitment strategies, AVOID and RESTRICT, are evolutionary viable in a PGG, showing when they are risk-dominant (see Methods) against all defectors and free-riders (i.e. FAKE, FREE and D players). Yet more importantly, we determine when RESTRICT becomes more advantageous than AVOID; that is, when it is worthwhile to pay extra cost to invest in restriction technologies and infrastructure that limit the benefits of non-committers (D).

Equation (16) (see Methods) allows one to determine when AVOID and RESTRICT are

120 risk-dominant against FREE. This occurs when

$$\sum_{k=1}^N (rc - c - \epsilon_P/k) \geq (N-1)(rc - c), \quad (1)$$

121 which can be simplified to

$$\epsilon_P \leq c(r-1)/F_N, \quad \text{with } F_N = \sum_{k=1}^N 1/k \quad (2)$$

122 AVOID and RESTRICT are risk-dominant against FAKE if

$$\sum_{k=1}^N \left(\left(\frac{rk}{N} - 1 \right) c + \frac{N\delta - \epsilon_P}{k} - \delta \right) \geq \sum_{k=1}^{N-1} \left(\frac{rk}{N} c - \delta \right). \quad (3)$$

123 Which can also be simplified to

$$\delta \geq \frac{N-r}{NF_{N-1}} c + \frac{F_N}{NF_{N-1}} \epsilon_P. \quad (4)$$

124 Now, as we aim to examine when restriction works better than avoiding non-committers, we

125 first examine independently when AVOID or RESTRICT are risk-dominant against D players.

126 In case of the AVOID strategy this occurs when:

$$\epsilon_P \leq N(r-1)c. \quad (5)$$

127 As Equation (2) is more restrictive than Equation (5), the two conditions in Equations (2) and

128 (4) define when AVOID is risk-dominant against all types of defectors and free-riders (see SI for

129 simplifications of these formulas using inequalities for F_N). Both conditions can be understood

130 intuitively. For a successful commitment, the cost of arranging the commitment needs to be

131 justified with respect to the benefit of (mutual) cooperation (i.e. $r(c-1)$), and a compensation

132 needs to be arranged (see Equation 4) that is proportional to the player's contribution and the
 133 investment cost she paid for setting up the commitment.

134 This observation becomes clearer when looking at the transition probabilities and stationary
 135 distribution in a population of AVOID players with the other four strategies, as shown in Figure
 136 1a. Note the cycles from C to defection strategies (FREE, D and FAKE) and back to AVOID
 137 strategists, showing that defection strategies cannot completely be avoided in the PGG context
 138 (see also Figure S1): When the cost of arranging commitment, ϵ_P , is sufficiently small, the
 139 population spends most of the time in the homogeneous state with AVOID players, regardless of
 140 the initial composition of the population (Figure 1b). For low ϵ_P , nearly homogeneous AVOID
 141 populations are almost always reached for sufficiently large δ . More interestingly, this high
 142 frequency is not affected by changes in the compensation δ , once a certain threshold is reached.
 143 Accordingly, as for the PD²⁰, the arrangement cost is the essential parameter for the emergence
 144 and survival of AVOID and mutual cooperation. Additionally, we observe that for a variety
 145 of group sizes N , the region of ϵ_P wherein AVOID is a viable strategy increases (see Figure
 146 S3, considering also the fact that right hand side of Equation (2) is an increasing function of
 147 N): AVOID can handle the commitment free-riding strategies for a wider range of arrangement
 148 costs. Yet as the groups size increases, the frequency of AVOID, for similar small values of ϵ_P ,
 149 decreases, revealing that other strategies may be necessary to cope with the increasing number
 150 of defectors in the groups and the population (Figure S3). These results provide a novel insight,
 151 when moving from the PD to the PGG, which is that avoiding defectors by refusing to play the
 152 game when someone does not agree to commit might lead to cooperation at higher arrangement
 153 costs, yet may in turn be detrimental for the overall level of cooperation in the game.

154 In turn, RESTRICT is risk-dominant against D when

$$\sum_{k=1}^{N-1} \frac{rkc(1-\psi)}{k+\psi(N-k)} \geq (N-r)c + F_N(\epsilon_P + \epsilon_R). \quad (6)$$

155 Because the left hand side of Equation (6) is a strictly decreasing function of ψ (see SI), the
 156 necessary condition for RESTRICT to be risk-dominant against D is (i.e. when $\psi = 0$)

$$(N - 1)rc \geq (N - r)c + F_N(\epsilon_P + \epsilon_R), \quad (7)$$

157 which is equivalent to

$$\epsilon_P + \epsilon_R \leq \frac{N(r - 1)}{F_N}c. \quad (8)$$

158 As the left hand side of Equation (6) is a continuous function of ψ , the satisfaction of Equation
 159 (8) guarantees that, for any given ϵ_P and ϵ_R , there exists a threshold ψ^D such that RESTRICT is
 160 risk-dominant against D for any ψ below it. This restriction threshold could be interpreted as the
 161 organizational or technological advancement required to guarantee success against individuals
 162 that exploit the non-exclusive public good. Moreover, it specifies what the limit is on the cost,
 163 combining the restriction and the proposing costs, for this to work.

164 These observations are supported by Figure 2a, where we show the transition probabilities
 165 and stationary distribution in a population of RESTRICT players with other non-commitment
 166 proposing strategies. The main difference with Figure 1a, is the increase in the fixation prob-
 167 ability from D to RESTRICT. This effect depends on the value of restriction factor ψ , as is
 168 shown in Figure 2b. In general, the better the effect of restriction on non-committers (i.e. the
 169 smaller ψ), the higher the frequency of RESTRICT and cooperation in the long run, regardless
 170 of the initial composition of the population. These observations are robust against changes in
 171 the restriction ϵ_R , as seen in Figure 2c, where we show the frequency of RESTRICT varying
 172 both ψ and ϵ_R . One can observe in both figures that ψ is the decisive parameter on the frequency
 173 of RESTRICT: To achieve significantly high frequency of RESTRICT, and as a consequence
 174 cooperation, a stringent restriction of non-committers must be possible. Even when costless
 175 restriction (i.e. $\epsilon_R = 0$) is available, the frequency of RESTRICT decreases quickly when ψ

176 approaches 1. Notice that when $\psi = 1$, RESTRICT is never risk-dominant against D players,
 177 as can be seen from Equation (6).

178 This notable success of the RESTRICT strategy in dealing with non-committers becomes
 179 even more significant when the group size increases (see Figure 2d). When the cost of restriction
 180 is extreme (e.g. $\epsilon_R = 2$), D players dominate when the group size is small. But when the group
 181 size is sufficiently large, thereby reducing the individual cost of implementing the restriction,
 182 which is shared by the proposers, RESTRICT becomes dominant. The frequency of RESTRICT
 183 is even higher when ϵ_R is small (see already Figure 3b). It is also interesting to note that the
 184 necessary condition for RESTRICT to be risk-dominant against D, as specified in Equation (8),
 185 is simpler for larger N , since the right hand side of the equation is an increasing function of N
 186 (see SI).

187 **2.3 What to do with the non-committers?**

188 Since there is no difference between AVOID and RESTRICT, except when playing D, one only
 189 needs to compare the stationary distribution of AVOID in a population with only D players
 190 against the stationary distribution of RESTRICT in a similar population. One can show that the
 191 frequency of RESTRICT is greater than AVOID if and only if the ratio of transition probabilities
 192 from D to RESTRICT and vice versa is greater than the ratio of transition probabilities from
 193 D to AVOID and vice versa, which can further be simplified, in the large population limit, to³²
 194 (see SI)

$$\sum_{k=1}^{N-1} \frac{rkc(1-\psi)}{k+\psi(N-k)} - (N-r)c - F_N(\epsilon_P + \epsilon_R) \geq rc - c - \epsilon_P/N, \quad (9)$$

195 which is equivalent to

$$\sum_{k=1}^{N-1} \frac{k(1-\psi)}{\psi N + k(1-\psi)} rc - F_N \epsilon_R - F_{N-1} \epsilon_P - (N-1)c \geq 0. \quad (10)$$

196 The left hand side of this equation is a decreasing function of ψ , ϵ_R and ϵ_P , but increasing in r
 197 (see proofs in SI). Note that the cost ϵ_P still persists because RESTRICT players need to pay
 198 this cost when there are D players in the group, while AVOID players do not have to pay it as
 199 they refuse to play with D players (thereby not arranging any commitment deal).

200 An observation that can be derived immediately from Equation (10) is that, when $\psi = 1$,
 201 the equation is not satisfied. That is, unless restriction is possible, at least to some degree, it is
 202 better to refuse to play with those who explicitly do not agree to commit. Furthermore, with
 203 $\psi = 0$ we obtain the necessary condition for RESTRICT to be favored over AVOID:

$$(r - 1)c \geq \frac{F_N}{N - 1}\epsilon_R + \frac{F_{N-1}}{N - 1}\epsilon_P. \quad (11)$$

204 It means the cost of arranging the agreement and restricting the benefit of non-contributors
 205 needs to be justified with respect to the benefit of the PGG. Only in that case can restriction
 206 become justified over avoidance or non-participation. If this equation is not satisfied, AVOID
 207 is the better commitment strategy, however good the restriction of non-committers that can be
 208 brought about, including full exclusion or ostracism.

209 In addition, as the left hand side of Equation (10) is a continuous function of ψ , this equa-
 210 tion guarantees that, for any given ϵ_P and ϵ_R , there exists a threshold ψ^{AVOID} which, for any
 211 $\psi < \psi^{AVOID}$, RESTRICT performs better than AVOID. This threshold moreover increases with
 212 r and can approach infinitely close to 1 when r tends to infinity, because, as mentioned above,
 213 the left hand side of this equation is a decreasing function of ψ , but increasing in r (see Figures
 214 S4 and S5 in SI). In addition, the equation provides the criteria for whether it is worthwhile
 215 to develop the restriction infrastructure or mechanism. As technological evolution may signifi-
 216 cantly (and unlimitedly) reduce the cost-to-impact ratio, for instance with respect to restriction
 217 mechanisms that require computing and communication power^{33,34}, one can postulate that un-

218 der this condition RESTRICT will eventually be available and deployable, even when currently
 219 AVOID seems to be the best choice.

220 Figure 3a shows the region of parameters where RESTRICT is better than AVOID, gener-
 221 ated from the analytical condition in Equation (10). There is a large range of costs for restricting
 222 the access of non-committers, ϵ_R , and the effect of restriction, ψ , where RESTRICT is better
 223 than AVOID. Yet, if ϵ_R is too large, even a full exclusion (ostracism) does not lead to a better
 224 commitment strategy (as also can be seen from Equation (11)). These analytical results are com-
 225 patible with the numerical simulation results shown in Figure 2c (see SI for more comparisons,
 226 Figure S6).

227 In addition, because the right hand side of Equation (11) is a decreasing function of N ,
 228 converging to 0 when N approaches infinity (see proofs in SI), AVOID is less preferred for
 229 increasing N : As group size increases (see Figure 3b), the frequency of the strategy decreases.
 230 This indicates that avoiding non-committers in larger groups is less successful than avoiding
 231 them in smaller ones. Interestingly, RESTRICT seems to cope better with changes in group
 232 size: As group size increases, RESTRICT becomes more frequent than AVOID for a larger
 233 value of the restriction cost. When the cost of restriction is small (e.g. $\epsilon_R = 0.25$), RESTRICT
 234 ensures a higher level of committers, and as a consequence contributors, in comparison to the
 235 AVOID strategy. Even when the restriction cost is high, there is a certain group size for which
 236 RESTRICT is more abundant than AVOID, see Figure 3b. Furthermore, as can be seen in
 237 Figures S4 and S5, when N increases, RESTRICT is the better strategy for strictly larger ranges
 238 of ϵ_P , ϵ_R and ψ , which becomes even more apparent for increasing values of the public good's
 239 multiplication factor r .

240 **3 Discussion**

241 We have shown that arranging prior commitment can lead to the evolution of cooperation in
242 a PGG when the cost of arranging commitment is justified with respect to the benefit of co-
243 operation. As such, this result generalizes the conclusions previously obtained for the PD²⁰,
244 underlining again the evolutionary advantage of this capacity to make prior agreements in com-
245 bination with the capacity to send signals and act accordingly^{1,35}. Moreover, we show that even
246 though the commitment strategies become viable for a wider range of the arrangement cost
247 when moving from the PD to the PGG, and more generally, when the group size of the PGG
248 increases, it might be detrimental for the overall level of cooperation when the cost of arrange-
249 ment is low. Nonetheless, prior agreements remain more efficient in achieving cooperation
250 when being compared to simple peer punishment^{9,36} (see Figure S7) in the PGG, an important
251 result we observed also in case of the PD scenario²⁰.

252 Notwithstanding this efficiency with respect to punishment, individuals that do not accept
253 an agreement (D) or individuals that free-ride on the investment that commitment proposers
254 make (FREE) may increase in frequency within both the PD and PGG contexts as the cost of
255 setting up the commitment increases. We examined here in detail how to deal with the former:
256 Either the commitment proposers can decide to avoid interacting with those non-committers
257 (AVOID) or implement an infrastructure that allows them to restrict their access (RESTRICT),
258 which corresponds to a reduction in the benefit they can obtain from the public good. For both
259 strategies it is assumed that an institution is present in order to enforce the compensation when
260 someone in the commitment does not honor it. As such, both strategies are closely related to
261 pool punishment models in terms of the presence of a third party required for the execution of
262 the process^{9,22}. Yet AVOID and RESTRICT are also different from pool punishment as in the
263 latter system no prior agreement on the posterior compensations is made. Additionally, the main

264 difference between AVOID and RESTRICT, in terms of execution, is that, AVOID does not pay
 265 to set up the agreement when someone does not accept the agreement, whereas RESTRICT
 266 always initiates the agreement to play and pays an additional fee to the institution to reduce the
 267 access of the (explicit) non-committers.

268 We show here that for both strategies one can identify intuitive conditions, defined by costs
 269 and compensations, that lead to increased likelihood in receiving contributions for the common
 270 good. Furthermore, we have compared the AVOID and RESTRICT strategies and provided
 271 analytical conditions for when one mechanism is better than the other. Our results show that
 272 RESTRICT is better than AVOID if non-committers can be restricted to a certain degree, with
 273 a small enough cost. Otherwise, it is better to rely on AVOID, i.e. simply refuse to interact if
 274 there are non-committers in the group (or, if the interaction is mandatory, to not contribute when
 275 playing the PGG). The restriction effect ψ , which is defined as the reduction of the benefit of
 276 the non-committer, was shown to be a decisive factor. Interestingly, its threshold ψ^{AVOID} (i.e.
 277 for all $\psi < \psi^{AVOID}$, RESTRICT is better than AVOID) increases with r and can be infinitely
 278 close to 1. This indicates that for given costs and compensations of commitment, any restriction
 279 mechanism can be more advantageous than AVOID when the PGG is sufficiently beneficial.

280 Moving from pairwise to multi-player interactions requires an analysis of the role of the
 281 group size in relation to the viability of the commitment strategies. Indeed, as group size in-
 282 creases, RESTRICT becomes more viable than AVOID even when it is costlier to implement
 283 the restriction measure. Furthermore, it is so for a strictly larger range of the commitment costs
 284 and restriction factors, especially when the public good's producing factor (r) is high. As such,
 285 these results indicate that larger public goods with higher levels of contribution can be estab-
 286 lished once restriction becomes possible. The amount of restriction depends on the gain each
 287 participant gets from the public good: the lower the gain, the tougher the restriction needs to
 288 be.

289 These results differ from other observations related to the impact of group size on the level
290 of cooperation in PGG^{37,38}: Lehmann et al.³⁷ showed that if the group size can be expanded
291 stochastically, for instance, as a result of an increase in fecundity and/or a decrease in mortality,
292 the kin-competition pressure induced by the limited dispersal in their networked model, can
293 be significantly reduced, thereby favoring the evolution of cooperative behavior. In a similar
294 manner, Alizon and Taylor³⁸ showed that if the group size and compositions can be adapted
295 over time, in a way that reduces the competition among relatives in a structured population (by
296 allowing groups or patches with high fecundity rates to grow faster), the cooperation level is
297 increased.

298 In contrast, the strategic mechanisms we examine here do not consider any forms of relat-
299 edness between group members or structured populations. The essential message is that for
300 smaller groups one is better off to avoid individuals that do not wish to accept an agreement,
301 prior to the game, when the cost of restriction is too large. Yet as group size increases, re-
302 striction mechanisms are more efficient in achieving cooperation, depending on the association
303 with the restriction. As such these results provide a completely new perspective on the role of
304 group size on the level of cooperation. We envisage that AVOID and RESTRICT may also be
305 more efficient in a structured population because the free-riders can be avoided and excluded
306 permanently by removing links with them. Furthermore, an introduction of relatedness among
307 individuals may reduce the need for arranging commitments as it provides additional incentive
308 to not free-ride. Both issues may be explored in subsequent papers.

309 The results presented here are in accordance with the outcomes of different behavioral com-
310 mitment experiments^{6,19,28,29}. High levels of cooperation were observed in a PGG experiment
311 where a binding agreement, which was enabled through a prior communication stage among the
312 members of the group, could be arranged before the PGG interaction occurs⁶. The experiment
313 showed that whenever a commitment deal is not binding or not enforced, corresponding to a low

314 compensation cost δ in our commitment models, defectors are widespread and the contribution
315 level is low. Commitment can also take the form of a deposit-refund scheme¹⁹, where those
316 who agree to commit have to deposit an amount which will be refunded only if they fulfill the
317 commitment and contribute to the public good. Similarly to our results for the AVOID strategy,
318 the most successful commitment strategy in that work was shown to be the one that refuses to
319 set up the public good whenever there is a non-committer in the group (RESTRICT was not con-
320 sidered). The outcome of this deposit-refund experiment showed that when the deposit amount,
321 corresponding to δ in our model, is sufficiently high, the contribution level is significant¹⁹. Note,
322 however, that in both these experiments^{6,19}, the cost of setting up the commitment is always set
323 to 0, thereby leading to effortless and effective commitment strategies. But as we have shown,
324 this cost is the decisive factor for the viability of commitments strategies. This said, despite the
325 fact that commitment has been shown experimentally to be a successful strategy for promoting
326 cooperation in the PGG, our results further the understanding of the mechanism by identify-
327 ing under which region of the parameters' values the mechanism works (as in the experiments)
328 and when it does not. As a result, the outcome of our analysis suggests the need to study how
329 varying the parameters would affect the outcomes of those commitment experiments.

330 Furthermore, PGG experiments, where exclusion of disapproved members (for example,
331 through voting) is allowed, exhibit a high level of contribution and commitment^{28,29}. But therein
332 exclusion is carried out after the PGG takes place, towards the observed non-contributors, as
333 in the model of Sasaki and Uchida³⁹, which is different from our model where restriction oc-
334 curs before the game takes place. This suggests that social exclusion or ostracism, even when
335 it requires an additional cost and/or has a reduced effect in terms of the restriction, is an im-
336 portant mechanism for promoting group cooperation^{26,29}. However, we envisage that exclusion
337 imposed through arranging prior commitment as in this work may be more suitable in the case
338 where there is rivalry in the game, as is the case for Common-pool Resource games⁸, as in that

339 case posterior restriction would not hinder the participating players in collecting their benefit.

340 Various extensions to the current model can be addressed. First, one can consider to move
341 beyond the symmetric commitments, where the cost for arranging and managing the agreement
342 is equally shared among the proposers. Asymmetric commitments, where the contribution to
343 manage the agreement may depend on the wealth and the potential benefits of each member
344 as in inequality models^{40,41}, may further increase the realism of the conclusions one can draw
345 from these models. Moreover, when extending to the repeated interaction scenario⁴², it is also
346 natural to consider that those who benefited more from the previous interactions should con-
347 tribute more to the management of the commitments. We envisage that these seemingly fairer
348 ways of sharing the benefit and cost of commitment can elevate the willingness to commit and
349 contribute. In this repeated interaction context, commitments can also be made incrementally,
350 conditional on behaviors of others in the previous round of interaction; this option has been
351 shown to promote a higher level of contribution in a repeated PGG experiment⁴³.

352 In summary, our results have demonstrated that arranging prior commitments provides an
353 important pathway for the emerge of cooperation in the one-shot Public Good Games, suggest-
354 ing that good agreements make good friends²⁰ also in group interactions. Furthermore, always
355 avoiding to play with those unwilling to commit is detrimental for the overall level of contri-
356 bution, especially when interacting in large groups, and restriction towards those players might
357 provide a better path to enhance the contribution level.

358 **4 Methods**

359 **4.1 Population setup and evolutionary dynamics**

360 Both the analytical and numerical results obtained here use Evolutionary Game Theory meth-
361 ods for finite populations^{10,44,45}. In such a setting, individuals' payoff represents their *fitness*

362 or social *success*, and evolutionary dynamics is shaped by social learning^{10,46,47}, whereby the
 363 most successful individuals will tend to be imitated more often by the others. In the current
 364 work, social learning is modeled using the so-called pairwise comparison rule⁴⁸, assuming that
 365 an individual A with fitness f_A adopts the strategy of another individual B with fitness f_B with
 366 probability given by the Fermi function, $(1 + e^{-\beta(f_B - f_A)})^{-1}$. The parameter β represents the
 367 ‘imitation strength’ or ‘intensity of selection’, i.e., how strongly the individuals base their de-
 368 cision to imitate on fitness comparison. For $\beta = 0$, we obtain the limit of neutral drift – the
 369 imitation decision is random. For large β , imitation becomes increasingly deterministic.

370 In the absence of mutations or exploration, the end states of evolution are inevitably monomor-
 371 phic: once such a state is reached, it cannot be escaped through imitation. We thus further
 372 assume that, with a certain mutation probability, an individual switches randomly to a different
 373 strategy without imitating another individual. In the limit of small mutation rates, the behav-
 374 ioral dynamics can be conveniently described by a Markov Chain, where each state represents
 375 a monomorphic population, whereas the transition probabilities are given by the fixation proba-
 376 bility of a single mutant^{9,45,49}. The resulting Markov Chain has a stationary distribution, which
 377 characterizes the average time the population spends in each of these monomorphic end states.

In finite populations, the groups engaging in PGG are given by multivariate hypergeometric sampling. For transition between two pure states (small mutation rate), this reduces to sampling (without replacement) from a hypergeometric distribution⁹. Denote

$$H(k, N, m, Z) = \frac{\binom{m}{k} \binom{Z-m}{N-k}}{\binom{Z}{N}}$$

378 Let $\Pi_{ij}(k)$ and $\Pi_{ji}(k)$ denote the payoff of a strategists of type i and j , respectively, when
 379 the random sampling consists of k players of type i and $N - k$ players of type j (as given in

380 the payoff matrix in Equation (1) in SI). Hence, in a population of x i -strategists and $(Z - x)$
 381 j -strategists, the average payoffs to i - and j - strategists are^{9,10}:

$$\begin{aligned}
 P_{ij}(x) &= \sum_{k=0}^{N-1} H(k, N-1, x-1, Z-1) \Pi_{ij}(k+1) \\
 &= \sum_{k=0}^{N-1} \frac{\binom{x-1}{k} \binom{Z-x}{N-1-k}}{\binom{Z-1}{N-1}} \Pi_{ij}(k+1)
 \end{aligned} \tag{12}$$

$$\begin{aligned}
 P_{ji}(x) &= \sum_{k=0}^{N-1} H(k, N-1, x, Z-1) \Pi_{ji}(k) \\
 &= \sum_{k=0}^{N-1} \frac{\binom{x}{k} \binom{Z-1-x}{N-1-k}}{\binom{Z-1}{N-1}} \Pi_{ji}(k)
 \end{aligned} \tag{13}$$

382 Note that several $P_{ij}(x)$ can be further simplified (see Supporting Information). Now, the prob-
 383 ability to change the number k of individuals using strategy i by ± 1 in each time step can be
 384 written as⁴⁸

$$T^{\pm}(k) = \frac{Z-k}{Z} \frac{k}{Z} [1 + e^{\mp\beta[P_{ij}(k) - P_{ji}(k)]}]^{-1}. \tag{14}$$

385 The fixation probability of a single mutant with a strategy i in a population of $(Z-1)$ individuals
 386 using j is given by^{45,48-50}

$$\rho_{j,i} = \left(1 + \sum_{i=1}^{N-1} \prod_{j=1}^i \frac{T^-(j)}{T^+(j)} \right)^{-1}. \tag{15}$$

387 In the limit of neutral selection (i.e. $\beta = 0$), $\rho_{B,A}$ equals the inverse of population size, $1/Z$.

388 Considering a set $\{1, \dots, q\}$ of different strategies, these fixation probabilities determine a
 389 transition matrix $M = \{T_{ij}\}_{i,j=1}^q$, with $T_{ij,j \neq i} = \rho_{ji}/(q-1)$ and $T_{ii} = 1 - \sum_{j=1, j \neq i}^q T_{ij}$, of a

390 Markov Chain. The normalized eigenvector associated with the eigenvalue 1 of the transposed
 391 of M provides the stationary distribution described above^{45,49,50}, describing the relative time the
 392 population spends adopting each of the strategies.

393

394 4.2 Risk-dominance condition

395 An important analytical criteria to determine the viability of a given strategy is whether it is risk-
 396 dominant with respect to other strategies^{13,32}. Namely, one considers which selection direction
 397 is more probable: an i mutant fixating in a homogeneous population of individuals playing j or
 398 a j mutant fixating in a homogeneous population of individuals playing i . When the first is more
 399 likely than the latter, i is said to be *risk-dominant* against j ³², which holds for any intensity of
 400 selection and in the limit of large Z when

$$\sum_{k=1}^N \Pi_{ij}(k) \geq \sum_{k=0}^{N-1} \Pi_{ji}(k). \quad (16)$$

401 Acknowledgments

402 TAH and TL acknowledges the support from the FWO Vlaanderen. TL also acknowledges the
 403 support of the *Fonds de la Recherche Scientifique* - FNRS.

404 References

- 405 1. Nesse RM. Evolution and the capacity for commitment. Russell Sage Foundation series
 406 on trust. Russell Sage; 2001.

- 407 2. Frank RH. Cooperation through Emotional Commitment. In: Nesse RM, editor. *Evolution and the capacity for commitment*. New York: Russell Sage; 2001. p. 55–76.
408
- 409 3. Sosis R. Religion and intra-group cooperation: preliminary results of a comparative
410 analysis of utopian communities. *Cross-Cultural Research*. 2000;34:70–87.
- 411 4. Irons W. Religion as a hard-to-fake sign of commitment. In: Nesse RM, editor. *Evolution
412 and the capacity for commitment*. New York: Russell Sage; 2001. p. 292–309.
- 413 5. Sterelny K. *The evolved apprentice*. MIT Press; 2012.
- 414 6. Chen XP, Komorita SS. The effects of communication and commitment in a public
415 goods social dilemma. *Organizational Behavior and Human Decision Processes*.
416 1994;60(3):367–386.
- 417 7. Han TA, Pereira LM, Santos FC. Intention Recognition Promotes The Emergence of
418 Cooperation. *Adaptive Behavior*. 2011;19(3):264–279.
- 419 8. Ostrom E. *Governing the commons: The evolution of institutions for collective action*.
420 Cambridge university press; 1990.
- 421 9. Hauert C, Traulsen A, Brandt H, Nowak MA, Sigmund K. Via freedom to coercion: The
422 emergence of costly punishment. *Science*. 2007;316:1905–1907.
- 423 10. Sigmund K. *The Calculus of Selfishness*. Princeton University Press; 2010.
- 424 11. Hardin G. The tragedy of the commons. *Science*. 1968;162:1243–1248.
- 425 12. West SA, Griffin AA, Gardner A. Evolutionary Explanations for Cooperation. *Current
426 Biology*. 2007;17:R661–R672.

- 427 13. Nowak MA. Five Rules for the Evolution of Cooperation. Science.
428 2006;314(5805):1560. DOI: 10.1126/science.1133755.
- 429 14. Hamilton WD. The genetical evolution of social behaviour. I. Journal of Theoretical
430 Biology. 1964;7(1):1 – 16. Available from: [http://www.sciencedirect.com/
431 science/article/pii/0022519364900384](http://www.sciencedirect.com/science/article/pii/0022519364900384).
- 432 15. Axelrod R, Hamilton WD. The evolution of cooperation. Science. 1981;211:1390–1396.
- 433 16. Chuang JS, Rivoire O, Leibler S. Simpson’s paradox in a synthetic microbial system.
434 Science. 2009;323(5911):272–275.
- 435 17. Xavier JB, Kim W, Foster KR. A molecular mechanism that stabilizes cooperative se-
436 cretions in *Pseudomonas aeruginosa*. Molecular microbiology. 2011;79(1):166–179.
- 437 18. de Vargas Roditi L, Boyle KE, Xavier JB. Multilevel selection analysis of a microbial
438 social trait. Molecular systems biology. 2013;9(1).
- 439 19. Cherry TL, McEvoy DM. Enforcing compliance with environmental agreements in the
440 absence of strong institutions: An experimental analysis. Environmental and Resource
441 Economics. 2013;54(1):63–77.
- 442 20. Han TA, Pereira LM, Santos FC, Lenaerts T. Good Agreements Make Good Friends.
443 Scientific reports. 2013;3(2695).
- 444 21. Han TA. Intention Recognition, Commitments and Their Roles in the Evolution of Coop-
445 eration: From Artificial Intelligence Techniques to Evolutionary Game Theory Models.
446 vol. 9. Springer SAPERE series; 2013.
- 447 22. Sigmund K, Silva HD, Traulsen A, Hauert C. Social learning promotes institutions for
448 governing the commons. Nature. 2010;466:7308.

- 449 23. Woodburn J. Egalitarian societies. *Man*. 1982;p. 431–451.
- 450 24. Sugiyama LS. Illness, injury, and disability among Shiwiar forager-horticulturalists:
451 Implications of health-risk buffering for the evolution of human life history. *American*
452 *Journal of Physical Anthropology*. 2004;123(4):371–389.
- 453 25. Hofmann HA, Schildberger K. Assessment of strength and willingness to fight during
454 aggressive encounters in crickets. *Animal Behaviour*. 2001;62(2):337–348.
- 455 26. Williams KD. *Ostracism: The Power of Silence. Emotions and Social Behavior*. Guilford
456 Publication; 2002.
- 457 27. Robertson TE, Delton AW, Klein SB, Cosmides L, Tooby J. Keeping the benefits of
458 group cooperation: domain-specific responses to distinct causes of social exclusion. *Evo-*
459 *lution and Human Behavior*. 2014;35(6):472–480.
- 460 28. Cinyabuguma M, Page T, Putterman L. Cooperation under the threat of expulsion in a
461 public goods experiment. *Journal of Public Economics*. 2005;89(8):1421–1435.
- 462 29. Maier-Rigaud FP, Martinsson P, Staffiero G. Ostracism and the provision of a pub-
463 lic good: experimental evidence. *Journal of Economic Behavior & Organization*.
464 2010;73(3):387–395.
- 465 30. Hummel JR, Lavoie D. National Defense and the Public-Goods Problem. *Anarchy and*
466 *the law: the political economy of choice*. 2007;p. 127–148.
- 467 31. Barrett S. *Environment and Statecraft: The Strategy of Environmental Treaty-Making:*
468 *The Strategy of Environmental Treaty-Making*. Oxford University Press; 2003.

- 469 32. Gokhale CS, Traulsen A. Evolutionary games in the multiverse. Proceedings of the Na-
470 tional Academy of Sciences of the United States of America. 2010 Mar;107(12):5500–
471 5504.
- 472 33. Moore GE. Cramming more components onto integrated circuits. Electronics.
473 1965;38(8).
- 474 34. Hilbert M, López P. The world's technological capacity to store, communicate, and
475 compute information. Science. 2011;332(6025):60–65.
- 476 35. Skyrms B. Signals: Evolution, Learning, and Information. Oxford University Press;
477 2010.
- 478 36. Sigmund K, Hauert C, Nowak M. Reward and punishment. P Natl Acad Sci USA.
479 2001;98(19):10757–10762.
- 480 37. Lehmann L, Perrin N, Rousset F. Population demography and the evolution of helping
481 behaviors. Evolution. 2006;60(6):1137–1151.
- 482 38. Alizon S, Taylor P. Empty sites can promote altruistic behavior. Evolution.
483 2008;62(6):1335–1344.
- 484 39. Sasaki T, Uchida S. The evolution of cooperation by social exclusion. Proceedings of
485 the Royal Society B: Biological Sciences. 2013;280(1752):20122498.
- 486 40. Tavoni A, Dannenberg A, Kallis G, Löschel A. Inequality, communication and the avoid-
487 ance of disastrous climate change in a public goods game. Proceedings of the National
488 Academy of Sciences USA. 2011;108:11825–11829.
- 489 41. Vasconcelos VV, Santos FC, Pacheco JM, Levin SA. Climate policies under wealth
490 inequality. Proceedings of the National Academy of Sciences. 2014;111(6):2212–2216.

- 491 42. Van Segbroeck S, Pacheco JM, Lenaerts T, Santos FC. Emergence of fairness in repeated
492 group interactions. *Physical review letters*. 2012;108(15):158104.
- 493 43. Kurzban R, McCabe K, Smith VL, Wilson BJ. Incremental commitment and reci-
494 procity in a real-time public goods game. *Personality and Social Psychology Bulletin*.
495 2001;27(12):1662–1673.
- 496 44. Nowak MA, Sasaki A, Taylor C, Fudenberg D. Emergence of cooperation and evolu-
497 tionary stability in finite populations. *Nature*. 2004;428:646–650.
- 498 45. Imhof LA, Fudenberg D, Nowak MA. Evolutionary cycles of cooperation and defec-
499 tion. *Proceedings of the National Academy of Sciences of the United States of America*.
500 2005;102:10797–10800.
- 501 46. Hofbauer J, Sigmund K. *Evolutionary Games and Population Dynamics*. Cambridge
502 University Press; 1998.
- 503 47. Rendell L, Boyd R, Cownden D, Enquist M, Eriksson K, Feldman MW, et al.
504 Why Copy Others? Insights from the Social Learning Strategies Tournament. *Sci-*
505 *ence*. 2010;328(5975):208–213. Available from: [http://www.sciencemag.org/
506 content/328/5975/208.abstract](http://www.sciencemag.org/content/328/5975/208.abstract).
- 507 48. Traulsen A, Nowak MA, Pacheco JM. Stochastic Dynamics of Invasion and Fixation.
508 *Phys Rev E*. 2006;74:11909.
- 509 49. Fudenberg D, Imhof LA. Imitation processes with small mutations. *Journal of Economic*
510 *Theory*. 2005;131:251–262.
- 511 50. Karlin S, Taylor HE. *A First Course in Stochastic Processes*. Academic Press, New
512 York; 1975.

513 **Figure Legends**

Figure 1. (a) Stationary distribution and fixation probabilities. The population spends most of the time in the homogenous state of AVOID. The black arrows identify the advantageous transitions, where $\rho_N = 1/Z$ denotes the neutral fixation probability. The dashed lines denote neutral transitions. Note the cyclic pattern from cooperation to defection to commitment strategies and back. **(b) Contour plot of the frequency of AVOID as a function of ϵ_P and δ .** For a small enough cost of arranging the commitment, AVOID is abundant whenever a sufficient compensation is associated with the commitment deal. Parameters: $N = 5, Z = 100, r = 3; \beta = 0.1$; In panel a, $\epsilon_P = 0.25, \delta = 2$.

Figure 2. (a) Transition probabilities and stationary distributions in case of RESTRICT. For an efficient restriction ($\epsilon_R = 0.5$ and $\psi = 0.25$), the population spends most of the time in the homogenous state of RESTRICT. Notations are the same as in Figure 1a. **(b) Frequencies of each strategy for varying ψ , in case of RESTRICT.** For a given cost of restriction ($\epsilon_R = 0.5$), in general the better the effect of restriction on non-committers (i.e. the smaller ψ), the greater the frequency of RESTRICT. **(c) Frequency of RESTRICT as a function of ϵ_R and ψ , in a population with C, D, FREE and FAKE strategies.** For a large range of cost for restricting the access of non-committers, ϵ_R , and the restriction, ψ , RESTRICT is highly frequent, having a higher frequency than AVOID. The double-stroke line corresponds to the part having the same frequency as AVOID (i.e. 0.64, with the same parameter values), and the area below this line identifies the area in which RESTRICT is more frequent than AVOID. In general, the larger ϵ_R , the smaller ψ is required for RESTRICT to be advantageous to AVOID. **(d) Frequencies of each strategy as a function of the group size, N .** RESTRICT becomes more frequent when the group size increases, even for a rather high cost of restriction ($\epsilon_R = 2.0$). Parameters: In panels a, b, c: $N = 5$; in all cases, $Z = 100, r = 3; \epsilon_P = 0.25, \delta = 2; \beta = 0.1$.

Figure 3. (a) Range of parameters ψ, ϵ_R and ϵ_P , generated from the analytical formula in Equation (10), in which RESTRICT is better than AVOID. For a large range of cost for restricting the access of non-committers, ϵ_R , and the effect of restriction, ψ , RESTRICT is better than AVOID. In general, the larger ϵ_R , the smaller ψ is required for RESTRICT to be advantageous to AVOID. **(b) Group size is an important factor for making RESTRICT more viable than AVOID.** We compute, as a function of the group size, N , the frequencies of RESTRICT for different values of restriction cost ϵ_R (the curves without markers), in comparison to the frequency of AVOID (the red curve with circled markers). In general, the lower the cost of restriction, the higher the frequency of RESTRICT. Also, the threshold of N above which RESTRICT is more frequent than AVOID is smaller. Parameters: in panel b, $Z = 100, \epsilon_P = 0.25, \psi = 0.25, \beta = 0.1$; In both panels, $N = 5, r = 3$.