International Institute for
Applied Systems Analysis
Schlossplatz 1
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342
Fax: +43 2236 71313
E-mail: publications@iiasa.ac.at
Web: www.iiasa.ac.at

Interim Report **IR-12-068**

The take-it-or-leave-it option allows small penalties to overcome social dilemmas

Tatsuya Sasaki
Åke Brännström
Ulf Dieckmann (dieckmann@iiasa.ac.at)
Karl Sigmund

Approved by

Pavel Kabat
Director General and Chief Executive Officer

February 2015

The take-it-or-leave-it option allows small penalties to overcome social dilemmas

Tatsuya Sasaki[a], Åke Brännström[b,a], Ulf Dieckmann[a] & Karl Sigmund[c,a,1]

Author affiliations:

[a]Evolution and Ecology Program, International Institute for Applied Systems Analysis (IIASA), 2361 Laxenburg, Austria
[b]Department of Mathematics and Mathematical Statistics, Umeå University, 90187 Umeå, Sweden
[c]Faculty of Mathematics, University of Vienna, 1090 Vienna, Austria

[1]To whom correspondence should be addressed

Tel: +43 1 4277 50612
Fax: +43 1 4277 9506
E-mail: karl.sigmund@univie.ac.at

20 Jan 2012

For the submission to the IIASA Interim Report series

Abstract

Self-interest frequently causes individuals engaged in joint enterprises to choose actions that are counterproductive. Free-riders can invade a society of cooperators, causing a tragedy of the commons. Such social dilemmas can be overcome by positive or negative incentives. Even though an incentive-providing institution may protect a cooperative society from invasion by free-riders, it cannot always convert a society of free-riders to cooperation. In the latter case, both norms, cooperation and defection, are stable: To avoid a collapse to full defection, cooperators must be sufficiently numerous initially. A society of free-riders is then caught in a social trap, and the institution is unable to provide an escape, except at a high, possibly prohibitive cost. Here, we analyze the interplay of (*a*) incentives provided by institutions and (*b*) the effects of voluntary participation. We show that this combination fundamentally improves the efficiency of incentives. In particular, optional participation allows institutions punishing free-riders to overcome the social dilemma at a much lower cost, and to promote a globally stable regime of cooperation. This removes the social trap and implies that whenever a society of cooperators cannot be invaded by free-riders, it will necessarily become established in the long run, through social learning, irrespective of the initial number of cooperators. We also demonstrate that punishing provides a 'lighter touch' than rewarding, guaranteeing full cooperation at considerably lower cost.

Key words: punishment; rewards; public goods; social contract; evolutionary games

Introduction

In many species, cooperation has evolved through natural selection. In human societies, it can additionally be promoted through institutions. Institutions may be viewed as 'tools that offer incentives to enable humans to overcome social dilemmas', to paraphrase Ostrom (1). The threat of punishment or the promise of reward can induce self-interested players to prefer actions that sustain the public good, and turn away from free-riding (2-13).

It is easy to understand the outcome of public good games in terms of the size of the incentive. If the incentive is too small, it has no effect and selfish players keep defecting by refraining from contributing to the public good (Fig. 1*a*). If, on the other hand, the incentive is sufficiently large, it compels all players to cooperate by contributing to the public good (Fig. 1*d*). It is the range of intermediate incentives that is of interest, and here, the effects of positive and negative incentives differ. Rewarding causes the stable coexistence of defectors and cooperators, with a larger proportion of cooperators when rewards are higher (Fig. 1*b*). Punishing, in contrast, leads to alternative stable states. As a result of the competition between cooperators and defectors, one or the other behavior will become established, but there can be no long-term coexistence (Fig. 1*c*). Whatever behavior prevails initially becomes fully established. Thus, each of the two behaviors may be viewed as a social norm: as long as the others stick to it, it does not pay to deviate. In particular, when cooperators are initially rare, the population will remain trapped in the asocial norm, with everyone defecting. Social learning cannot lead, in that case, to the more beneficial, pro-social norm of cooperating.

Here, we show that the option to abstain from the joint enterprise (14-17) offers an escape from the social trap. Indeed, when free-riding is the norm, players will turn away from unpromising joint ventures. This leads to the decline of exploiters and allows the re-emergence of cooperators. If the incentives are too low, this is followed by the comeback of defectors, in a rock-paper-scissors type of cycle (18, 19) (Fig. 2*a*). However, even a modest degree of punishment breaks the rock-paper-scissors cycle and allows the fixation of the cooperative norm (Fig. 2*e*-*g*). Thus, optional participation allows a permanent escape from the social trap. In contrast, we show that optional participation has little impact on rewarding systems (Fig. 2*b*-*d*).

Methods

Specifically, we apply evolutionary game theory (20) to cultural evolution, based on (*a*) social learning (i.e., the preferential imitation of more successful strategies) and (*b*) occasional exploratory steps (modeled as small and rare random perturbations). Because the diversity of public good interactions and sanctioning mechanisms is huge, we first present a fully analytical investigation of a prototypical case (Supporting Information, SI). We posit a large, well-mixed population of players. From time to time, a random sample of $n \geq 2$ players is faced with an opportunity to participate in a public good game, at a cost $g > 0$. We denote by m the number of players willing to participate ($0 \leq m \leq n$) and assume that $m \geq 2$ players

are required for the game to take place. If it does, each of the m players decides whether or not to contribute a fixed amount $c > 0$, knowing that it will be multiplied by r (with $1 < r < n$) and distributed equally among all $m - 1$ other members of the group. If all group members invest into the common pool, each obtains a payoff $(r - 1)c - g$, which we assume to be positive. The social dilemma arises because players can improve their payoffs by not contributing. If all do so, each obtains the negative payoff $-g$. Thus, they would have done better to refrain from participation.

We now introduce the incentive. It is convenient to write the total incentive stipulated by an authority ('the institution') in the form mI, where I is the per capita incentive. If rewards are used, the total incentive will be shared among those players who cooperated. Hence each cooperator obtains a reward mI/m_{C}, where m_{C} denotes the number of cooperators among the m players. If penalties are used, players who defect have their payoffs analogously reduced by mI/m_{D}, where m_{D} denotes the number of defectors among the m players. We will see that in the compulsory case, there exist two alternative stable norms for intermediate strength of punishment. In particular, a homogeneous population of defectors is unable to escape from the social trap (Fig. 1). In the optional case, cultural evolution leads to a stable homogenous population of cooperators (Fig. 2*e-g*), irrespective of the initial number of cooperators. Thus, voluntary participation overcomes the social trap plaguing the compulsory case. Remarkably, this is achieved at a fraction $1/n$ of the cost necessary in the compulsory case (Section S2 in the SI).

We base our analysis of the underlying evolutionary game on replicator dynamics (e.g., 20) for the three strategies C (cooperators), D (defectors), and N (non-participants), with frequencies x, y, and z. The state space Δ is the triangle of all (x, y, z) with $x, y, z \geq 0$ and $x + y + z = 1$. If $0 < g < (r - 1)c$, these three strategies form a rock-scissors-paper cycle in the absence of incentives, as shown in Fig. 2*a*: D beats C, N beats D, and C beats N. In the interior of the state space, all trajectories of the replicator dynamics originate from, and converge to, the state N of non-participation $(z = 1)$ (21). Hence, cooperation can only emerge in brief bursts, sparked by random perturbations. The long-term payoff is that of non-participants (i.e., 0).

Results

If the game is compulsory, i.e., if all n players are obliged to participate $(z = 0)$, the outcome changes with increasing per capita incentive I (Fig. 1). For small I, defection dominates. The replicator dynamics have two equilibria: one stable (a homogeneous population of D-players) and one unstable (a homogeneous population of C-players). In the case of rewarding, as I crosses the threshold $I_- = c/n$, the equilibrium D becomes unstable, spawning a stable equilibrium R at a mixture of C- and D-players. As I increases further, the fraction of cooperators becomes larger and larger. Finally, when I reaches the threshold $I_+ = c$, the stable mixture merges with the formerly unstable equilibrium C, which becomes stable. In the case of punishing, as I crosses the threshold I_-, it is the unstable equilibrium C that becomes

stable, spawning an unstable equilibrium R at a mixture of C- and D-players. R thus separates the regions of attraction of the equilibria C and D. With increasing I, the region of attraction of D becomes smaller and smaller, until I attains the value I_+. Here, the unstable equilibrium R merges with the formerly stable equilibrium D, which becomes unstable. For larger values of I, everyone cooperates. As shown in Section S2 in the SI, the values of I_+ and I_- are the same, irrespective of whether we consider rewarding or punishing.

We next investigate the interplay of (*a*) institutional incentives and (*b*) optional participation. Clearly, if the public good game is too expensive [i.e., if $g \geq (r-1)c + I$, in the case of rewarding or $g \geq (r-1)c$ in the case of punishing], players will opt for non-participation. We do not further consider this trivial case.

We first examine the case of punishing, for increasing per capita incentives I. For $I < I_-$, the effect of the incentive is negligible and all trajectories converge to N. As I crosses the threshold I_-, the equilibrium R appears on the CD-edge. At first, it is a saddle point. A trajectory leading from N to R separates the interior of Δ into two regions (Fig. 2*e*). One region is filled with trajectories issuing from N and converging to C, and the other is filled with trajectories issuing from and returning to N. If we assume that arbitrarily small random perturbations can, from time to time, affect the population (corresponding to occasional individual explorations of an alternative strategy), we see that the population will eventually end up at the stable equilibrium C. If I increases beyond a threshold K_-, an equilibrium Q enters Δ at R through a saddle-node bifurcation. With increasing I, the point Q moves along a straight line to N, while R keeps moving, along the CD-edge, to D (Fig. 2*f*). In the SI, we show that Q is the unique equilibrium in the interior of the state space Δ (i.e., with all three strategies present) and that it is a saddle point. If I increases still further and crosses a threshold K_+, the equilibrium Q exits Δ through N. The point R becomes a source and remains so until it merges with D (for $I = I_+$) (Fig. 2*g*). Almost all trajectories in Δ either converge directly to C or to N. However, N is not stable. If the population is in the vicinity of N, arbitrarily small and rare random perturbations will eventually send it into the region of attraction of C. Hence, the population ultimately settles at the stable equilibrium C whenever $I > I_-$. This means that as soon as a homogeneous population of cooperators is immune against invasion by rare defectors, it becomes established in the long run.

In the case of rewarding, for $I < I_-$, the incentive has a negligible effect and all trajectories converge to N. As I crosses the threshold I_-, the equilibrium R appears on the CD-edge. Again, it is a saddle, but a trajectory now leads away from R to N (Fig. 2*b*). It separates a region where all trajectories lead from D to N from a region filled with trajectories issuing from and returning to N. As I increases and crosses a threshold J_-, a saddle-node bifurcation occurs at R, spawning an equilibrium Q into Δ (Fig. 2*c*) Again, one can show that this interior equilibrium is unique, and is a saddle point (see the SI). If I crosses a threshold J_+, the equilibrium Q exits Δ through N. All trajectories in the interior of Δ converge to R (Fig. 2*d*). As I increases beyond I_+, the stable equilibrium R merges with C and all trajectories converge to C, just as in the case of punishment (Fig. 2*h*).

For enhancing a group's welfare, rewarding obviously works better than punishing (just as in the classical behaviourist analysis of reinforcements). However, the price of the rewarding has to be substantial. Punishing can achieve all-out cooperation (in the long run) for a much smaller price, namely, I_- (which is the smaller the larger the group). From the viewpoint of institutionalizing a sanctioning mechanism, punishing thus has an advantage over rewarding: it achieves a higher average payoff at lower costs.

So far, we have treated g (the price an individual is willing to pay to participate in a joint enterprise) and I (the per capita size of the total incentive) as independent parameters. However, if individuals can freely decide whether or not to participate in the game, it makes sense to assume that they pay for the institution providing the incentives. For instance, I could be some fraction of the entrance cost g, or (equivalently) the total entrance cost could be viewed as the sum $g + aI$ of a part g kept by the authority and a part aI used for the incentive, with $a > 0$ (it is natural to assume that this part is proportional to the per capita incentive I). A rewarding system, if $a = 1$, simply redistributes the payoff without increasing group welfare, whereas a punishing system decreases it even if no one has to be punished. (We have to pay for the costly apparatus of law enforcement even if no one defaults.)

In the case of rewarding, optional participation increases the group welfare only marginally to 0 (Fig. 3*b*), for the small range $I_- < I < J_-$, where compulsory participation leads to negative average payoffs. In that range, combining rewarding with optional participation even reduces the cooperator frequency to 0 (Fig. 3*a*). For punishing, the situation is very different. The group welfare is highest when I is just barely larger than the minimum $I_- = c/n$ required to obtain full cooperation (Fig. 3*d*). The learning process, in that case, will take some time, and the population can undergo violent oscillations between the N-, C-, and D-states; however, in the end, the C-norm will prevail (Fig. 3*c*).

In the SI, we test by extensive numerical investigations the robustness of our analytical results with respect to alternative model variants:

i) If we assume that part of the contribution to the public good returns to the contributing player, the dynamics becomes more complex but the evolutionary outcome remains unchanged (Section S3 and Figs. S1 and S2 in the SI).
ii) Requiring participants to pay a fee for the sanctioning system also has little effect on the predicted outcome, as long as this fee does not become unreasonably large (Fig. 3 and Section S5).
iii) Moreover, when unused fees are returned, small negative per capita incentives suffice to maximize social welfare (Section S5).

We can also model the sanctioning system in different ways. Results remain unchanged as long as reward, or punishment, decreases with the number of free-riders:

iv) This is the case, for instance, if only one defector is exemplarily punished, because the probability for being singled out decreases [in the old Navy, the slowest sailor was liable to get 'prompted' (i.e., beaten)] (Section S4).

v) It also holds whenever the institution needs to spend some resource (e.g., time) to punish a convicted free-rider. Indeed, this diminishes the resources to hunt for other free-riders. Such a 'handling time' [to borrow an expression from predator-prey models (22)] will reduce the average punishment expected per defector, which is proportional to $mI/(a + bm_{\mathrm{D}})$, with $a, b > 0$ (Section S4).

vi) Also, the capping of individual penalties leaves our qualitative findings unchanged (Section S4).

For these and related scenarios, optional participation leads to the establishment of full cooperation whenever the sanction is strong enough to deter free-riders from invading. Surprisingly, in all cases we have considered, the cost of the negative incentive required to establish a norm of full cooperation is a small fraction of the cost needed in the case of compulsory participation.

Discussion

In his famous *Leviathan*, published in 1651, Hobbes stressed the necessity of an authority to curb the selfish motivations of individuals. He attributed its existence to a social contract intended to promote the commonwealth. Here, we assume that such a Leviathan-like authority exists, and is able to provide sanctions in the form of penalties and rewards. Indeed, most of our joint enterprises are protected by an elaborate apparatus of regulations, controls, and contract-enforcement devices to provide the necessary coercion. The theory of the social contract is a major topic in political philosophy, and a rich field of applications for game theory (e.g., 13).

The large majority of economic experiments and theoretical studies dealing with sanctions use peer-punishment, and thus make do without Leviathan, at least at first sight. Players can decide, independent of each other, whether to punish co-players or not. This setting is of particular interest for investigating how pro-social coercion evolved, out of a world of anarchy (e.g., 1). Studies of peer-punishment attempt to address such a scenario (23-32). It seems clear, however, that in all economic experiments, Leviathan looms in the background. Players can pick their decisions, but usually only in a very narrow, regularized framework of alternatives. In modern human societies, anarchy is rare and players can almost always appeal to a higher authority.

There are many intermediate stages between pure peer-punishment and institutionalized punishment. Several authors have considered scenarios in which punishment is meted out only if two, or a majority, of players opt for it, or have allowed players to vote between treatments with or without peer-punishment (33-35). Thus, sanctions were supported by some social consensus, which can be mediated by communication ['cheap talk' (36)]. In other

studies, players could contribute, before engaging in the public good game, to a punishment pool. This is like paying the wages of a police force before knowing whether, or against whom, it will be deployed (4, 37). Both theory and experiments have shown that delegating punishment is an efficient way to promote cooperation (38-40). Often, however, players of a public good game can engage in second-order free-riding by not paying toward the sanctions, which, in turn, raises the issue of second-order punishment. In our model, whoever wants to join the game has to pay an entrance fee. Second-order free-riding is no option, nor is asocial punishment targeted against cooperators (30). Leviathan sees to it.

The interplay of punishing, on the one hand, and optional participation, on the other hand, has already been investigated in several papers (21, 41-43). However, these studies mainly examined the problem of second-order free-riding. In contrast to these papers, we consider institutional punishment enforced by a higher authority. To our knowledge, this is the first time that evolutionary game theory is applied to the implementation of an authority through social contract (by allowing individuals to voluntarily participate in a joint interaction). This establishes an interesting analogy with the suppression of competition occurring in several fields of evolutionary biology (e.g., ‘selfish genes’) (44).

Voluntary submission under a sanctioning institution occurs in many real-life instances of cooperation. Practically all joint commercial and industrial enterprises are protected by enforceable contracts. Adherence is voluntary but commits the parties to mutually beneficial contributions. Punitive clauses ensure that non-compliance will be sanctioned. This principle also works, although at a less regulated level, in small-scale societies (1, 5, 38) and permits the sustainable use of common grazing or fishing grounds, or the construction and maintenance of irrigation systems. Medieval guilds delegated authority to chosen agents, and settlers hired sheriffs to deter villains. In day-to-day life, we may think of janitors, umpires, referees, or wardens who uphold rules in housing blocks, team games, private clubs, or public parks. All these examples rely on formal or informal agreements that can be freely joined but are then backed up by a higher authority. Thus, the situation we have addressed in our model is both fundamental and widespread.

References

1. Ostrom E (2005) *Understanding Institutional Diversity* (Princeton Univ Press, Princeton).
2. Hardin G (1968) The tragedy of the commons. *Science* 162:1243-1248.
3. Olson E (1965) *The Logic of Collective Action: Public Goods and the Theory of Groups* (Harvard Univ Press, Cambridge, MA).
4. Yamagishi T (1986) The provision of a sanctioning system as a public good. *J Pers Soc Psychol* 51:110-116.
5. Ostrom E (1990) *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge Univ Press, New York).
6. Trivers RL (1971) The evolution of reciprocal altruism. *Q Rev Biol* 46:35-57.
7. Camerer C (2003) *Behavioral Game Theory: Experiments in Strategic Interaction* (Russell Sage Foundation, New York).
8. Dickinson DL (2001) The carrot vs. the stick in work team motivation. *Exp Econ* 4:107-124.
9. Henrich J, et al. (2006) Costly punishment across human societies. *Science* 312:1767-1770.
10. Sigmund K (2007) Punish or perish? Retaliation and collaboration among humans. *Trends Ecol Evol* 22:593-600.
11. Skyrms B (2004) *The Stag Hunt and the Evolution of Social Structure* (Cambridge Univ Press, Cambridge, UK).
12. Sugden R (1998) *The Economics of Rights, Cooperation and Welfare* (Blackwell, Oxford).
13. Binmore KG (1994) *Playing Fair: Game Theory and the Social Contract* (MIT Press Cambridge, MA).
14. Orbell JM, Dawes RM (1993) Social welfare, cooperators' advantage, and the option of not playing the game. *Am Sociol Rev* 58:787-800.
15. Batali J, Kitcher P (1995) Evolution of altruism in optional and compulsory games. *J Theor Biol* 175:161-171.
16. Semmann D, Krambeck HJ, Milinski M (2003) Volunteering leads to rock-paper-scissors dynamics in a public goods game. *Nature* 425:390-393.
17. Sasaki T, Okada I, Unemi T (2007) Probabilistic participation in public goods games. *Proc Biol Sci* 274:2639-2642.
18. Hauert C, De Monte S, Hofbauer J, Sigmund K (2002) Volunteering as Red Queen mechanism for cooperation in public goods games. *Science* 296:1129-1132.
19. Hauert C, De Monte S, Hofbauer J, Sigmund K (2002) Replicator dynamics for optional public good games. *J Theor Biol* 218:187-194.
20. Hofbauer J, Sigmund K (1998) *Evolutionary Games and Population Dynamics* (Cambridge Univ Press, Cambridge, UK).
21. De Silva H, Hauert C, Traulsen A, Sigmund K (2009) Freedom, enforcement, and the social dilemma of strong altruism. *J Evol Econ* 20:203-217.
22. Holling CS (1959) Some characteristics of simple types of predation and parasitism. *Can Entomol* 91:385-398.

23. Boyd R, Richerson P (1992) Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethol Sociobiol* 13:171-195.
24. Fehr E, Gächter S (2002) Altruistic punishment in humans. *Nature* 415:137-140.
25. Fehr E, Rockenbach B (2003) Detrimental effects of sanctions on human altruism. *Nature* 422:137-140.
26. Gardner A, West SA (2004) Cooperation and punishment, especially in humans. *Am Nat* 164:753-764.
27. Gürerk O, Irlenbush B, Rockenbach B (2006) The competitive advantage of sanctioning institutions. *Science* 312:108-111.
28. Egas M, Riedl A (2008) The economics of altruistic punishment and the maintenance of cooperation. *Proc Biol Sci* 275:871-878.
29. Dreber A, Rand DG, Fudenberg D, Nowak MA (2008) Winners don't punish. *Nature* 452:348-351.
30. Herrmann B, Thöni C, Gächter S (2008) Antisocial punishment across societies. *Science* 319:1362-1367.
31. Casari M (2005) On the design of peer punishment experiments. *Exp Econ* 8:107-115.
32. Nakamaru M, Dieckmann U (2009) Runaway selection for cooperation and strict-and-severe punishment. *J Theor Biol* 257:1-8.
33. Boyd R, Gintis H, Bowles S (2010) Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science* 328:617-620.
34. Ertan A, Page T, Putterman L (2009) Who to punish? Individual decisions and majority rule in mitigating the free rider problem. *Eur Econ Rev* 53:495-511.
35. Kosfeld M, Okada A, Riedl A (2009) Institution formation in public goods games. *Am Econ Rev* 99:1335-1355.
36. Bochet O, Page T, Putterman L (2006) Communication and punishment in voluntary contribution experiments, *J Econ Behav Organ* 60: 11-26.
37. Sigmund K, De Silva H, Traulsen A, Hauert C (2010) Social learning promotes institutions for governing the commons. *Nature* 466:861-863.
38. Poteete A, Janssen M, Ostrom E (2010) *Working Together: Collective Action, the Commons, and Multiple Methods in Practice* (Princeton Univ Press, Princeton).
39. O'Gorman R, Henrich J, Van Vugt M. (2009) Constraining free riding in public goods games: Designated solitary punishers can sustain human cooperation. *Proc Biol Sci* 276:323-329.
40. Baldassarri D, Grossman G (2011) Centralized sanctioning and legitimate authority promote cooperation in humans. *Proc Natl Acad Sci USA* 108:11023-11026.
41. Fowler JH (2005) Altruistic punishment and the origin of cooperation. *Proc Natl Acad Sci USA* 102:7047-7049.
42. Hauert C, Traulsen A, Brandt H, Nowak MA, Sigmund K (2007) Via freedom to coercion: The emergence of costly punishment. *Science* 316:1905-1907.
43. Mathew S, Boyd R (2009) When does optional participation allow the evolution of cooperation. *Proc Biol Sci* 276:1167-1174.
44. Frank SA (1995) Mutual policing and repression of competition in the evolution of cooperative groups. *Nature* 377:520-522.

Acknowledgements: This study was enabled by financial support by the Austrian Science Fund to U.D. (TECT I-106 G11), through a grant for the research project *The Adaptive Evolution of Mutualistic Interactions* as part of the multinational collaborative research project *Mutualisms, Contracts, Space, and Dispersal* (BIOCONTRACT) selected by the European Science Foundation as part of the EUROCORES Programme *The Evolution of Cooperation and Trading* (TECT). U.D. gratefully acknowledges additional support by the European Commission, the European Science Foundation, the Austrian Ministry of Science and Research, and the Vienna Science and Technology Fund**.** K.S. thanks TECT I-104 G1.

Author Contributions: All authors participated in model design, in model analysis, and in the writing of the paper. T.S. carried out the analytical and numerical investigations.

Figure Legends

Figure 1 | Effects of institutional rewarding and punishing on the compulsory public good game for different per capita incentives I. For rewarding and punishing, full cooperation requires large incentives, even though the transition from full defection to full cooperation differs for the two types of incentive (b and c). (a) If I is smaller than $I_- = c/n$, the incentives have no effect on the outcome of the public good game and defection prevails. (d) If I is larger than $I_+ = c$, the incentives reverse the outcome and cooperation prevails. (b and c) For intermediate incentive I, rewarding leads to the stable coexistence of cooperation and defection, whereas punishing leads to alternative stable states. C and D correspond to the two homogenous states in which the population consists exclusively of cooperators and defectors, respectively. With increasing incentive I, the equilibrium R moves toward C in the case of rewarding and toward D in the case of punishing.

Figure 2 | Effects of institutional rewarding and punishing on the optional public good game for different per capita incentives I. Combining punishing with optional participation enables full cooperation for a small fraction of the cost needed in the compulsory case. The triangles represent the state space $\Delta = \{(x, y, z): x, y, z \geq 0, x + y + z = 1\}$, where x, y, and z are the frequencies of cooperators, defectors, and non-participants, respectively. The three vertices C, D, and N correspond to the three homogeneous states in which the population consists exclusively of cooperators ($x = 1$), defectors ($y = 1$), or non-participants ($z = 1$). (a) If I is smaller than $I_- = c/n$, the incentives have no effect on the outcome of the public good game. The interior of Δ is filled with trajectories issuing from and converging to the vertex N of non-participation in the joint enterprise. In that state, arbitrarily small random perturbations lead to short bursts of cooperation, immediately subverted by defection and followed by a return to non-participation. (h) If I is larger than $I_+ = c$, the incentives alter the outcome and cooperation prevails. All trajectories converge to C, the state of full cooperation. For the range of incentives in between a and h, the impacts of rewards and penalties differ. **Rewarding:** (b) For $I_- < I < J_-$, the equilibrium R on the CD-edge is a saddle point. All trajectories in the interior of Δ lead to N. (c) For $J_- < I < J_+$, an interior saddle point Q moves, with increasing I, along the dashed line from the CD-edge to N. Trajectories either converge to R, now a sink, or else to N. From there, an arbitrarily small random perturbation will send the state into the region of attraction of R. (d) For $J_+ < I < I_+$, the interior equilibrium Q has exited through N, and all trajectories converge to R, implying stable coexistence of defectors and cooperators. **Punishing:** (e) For $I_- < I < K_-$, the equilibrium R on the CD-edge is a saddle point. A trajectory from N to R separates a region where all trajectories lead to C from a region where all trajectories lead to N. An arbitrarily small random perturbation of N can lead to the region of attraction of C, and hence to the fixation of full cooperation. (f) For $K_- < I < K_+$, an interior saddle point Q moves, with increasing I, along the dashed line from the CD-edge to N. R is now a source. (g) For $K_+ < I < I_+$, the interior equilibrium Q has exited through N. In f and g, trajectories converge to C, either directly, or after a small random perturbation away from N. In summary, combining punishing with optional participation causes full cooperation from any initial condition for per

capita incentives exceeding I_-, whereas combining rewarding with optional participation achieves this only for per capita incentives exceeding I_+. Parameters: $n = 5$, $r = 3$, $c = 1$, $g = 0.5$, and $I = 0$ (*a*); 0.25 (*b* and *e*); 0.35 (*c*); 0.55 (*f*); 0.7 (*d* and *g*); or (punishment) 1.2 (*h*).

Figure 3 | 'User-pays' variant. In this variant, players are obliged to pay an entrance fee $g + aI$. The panels show co-operator frequencies (*a* and *c*) and long-term average payoffs in the population (*b* and *d*), for rewarding (*a* and *b*) and punishing (*c* and *d*) and different per capita incentives I, Parameters: $n = 5$, $r = 3$, $c = 1$, $a = 1$, and $g = 0.5$.

Fig. 1

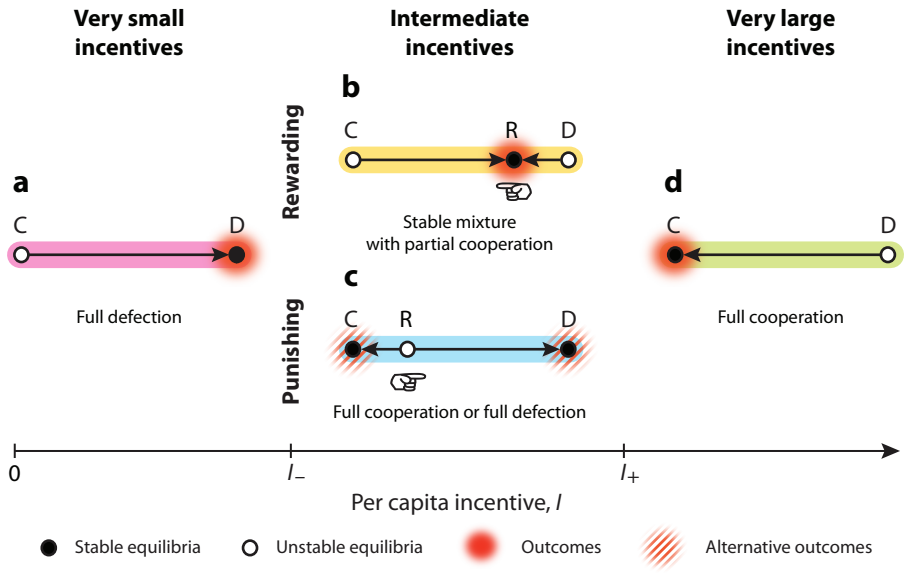

Very small incentives
Intermediate incentives
Very large incentives
a
C
D
Full defection
b
Rewarding
C
R
D
Stable mixture with partial cooperation
c
Punishing
C
R
D
Full cooperation or full defection
d
C
D
Full cooperation
0
I_-
I_+
Per capita incentive, I
Stable equilibria
Unstable equilibria
Outcomes
Alternative outcomes

Fig. 2

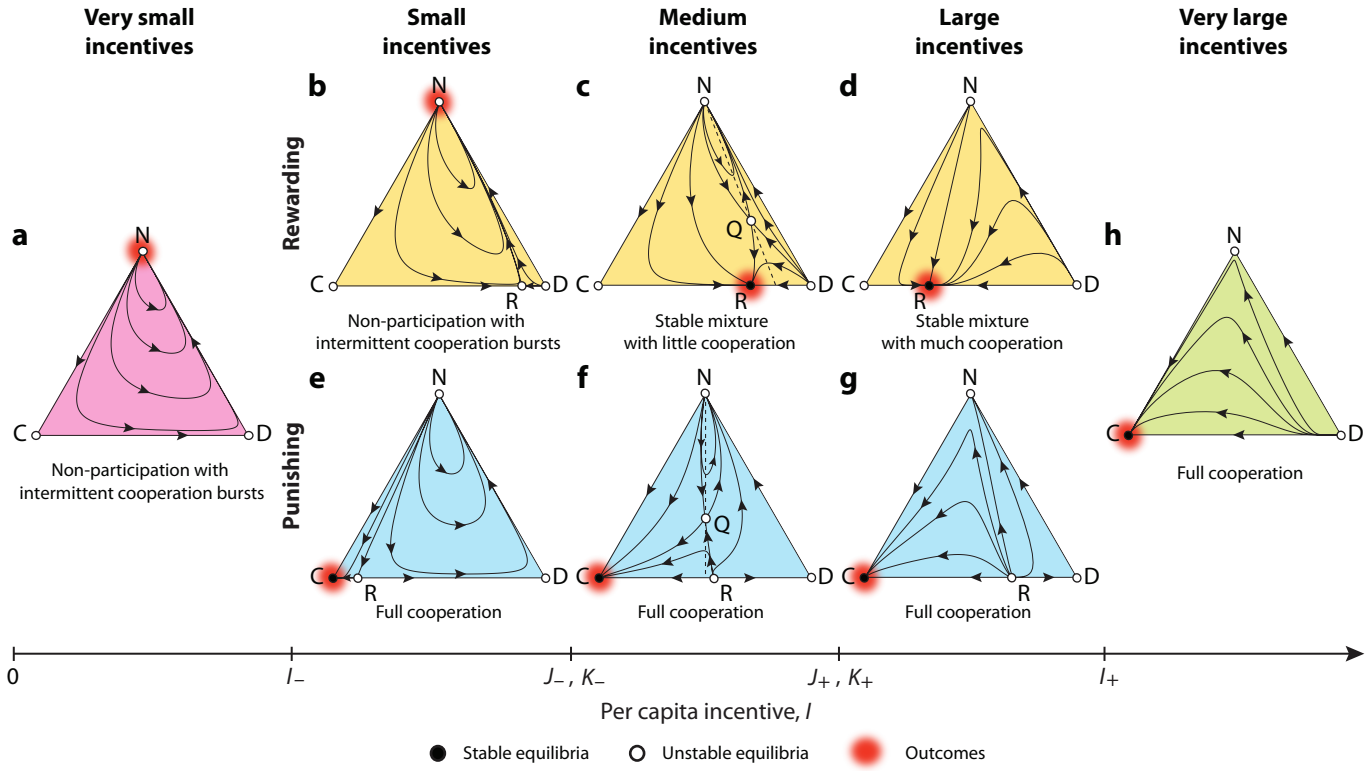
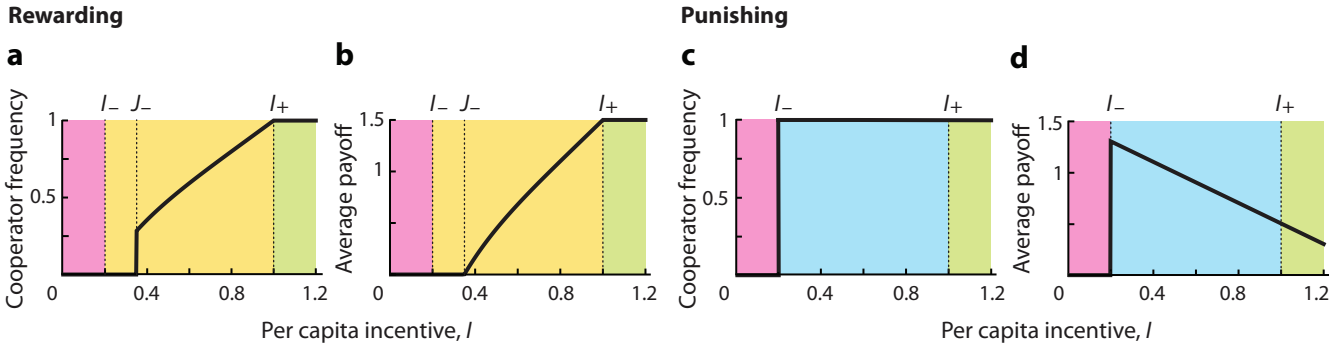

Very small incentives
Small incentives
Medium incentives
Large incentives
Very large incentives
Rewarding
Punishing
a
b
c
d
e
f
g
h
N
C
D
R
Q
Non-participation with intermittent cooperation bursts
Stable mixture with little cooperation
Stable mixture with much cooperation
Full cooperation
0
I_-
J_- , K_-
J_+ , K_+
I_+
Per capita incentive, I
Stable equilibria
Unstable equilibria
Outcomes

Fig. 3


Rewarding
Punishing
a
b
c
d
Cooperator frequency
Average payoff
Per capita incentive, l
l−
J−
l+
0
0.4
0.8
1.2
0.5
1
1.5

Supporting Information for

The take-it-or-leave-it option allows small penalties to overcome social dilemmas

Tatsuya Sasaki, Åke Brännström, Ulf Dieckmann & Karl Sigmund

We begin with the proofs of the results for our prototypical model. We first determine, in Section S1, the payoffs for optional public good games with rewarding and punishing and continue, in Section S2, with an analytical investigation of the resultant dynamics. We then numerically investigate variants, considering first, in Section S3, the 'self-returning' variant of public good games and then, in Section S4, variants of the incentive scheme. Finally, in Section S5, we consider a 'user-pays' variant, in which players themselves have to finance the total incentive.

S1 Payoffs

We calculate the average payoff $\bar{P}^s$ for the population and the expected payoff values P_{C}^s and P_{D}^s for cooperators and defectors, where $s =$ o, r, p is used to specify one of three sanctioning systems: 'without incentives', 'with rewarding', and 'with punishing'. We denote by x, y, and z the respective frequencies of cooperators, C; defectors, D; and non-participants, N. Because non-participants have a payoff of 0, the average payoff in the population is given by $\bar{P}^s = xP_{\mathrm{C}}^s + yP_{\mathrm{D}}^s$.

Without incentives, a defector in a group with $m-1$ co-players ($m = 2, \dots, n$) obtains from the common good a payoff of $rcx/(1-z)$ on average (21). Hence,

$$P_{\mathrm{D}}^{\mathrm{o}} = \left(rc\frac{x}{1-z} - g\right)(1-z^{n-1}). \tag{S1}$$

Clearly, z^{n-1} is the probability of finding no co-player, and thus to be reduced to non-participation. In addition, cooperators contribute c with a probability $1-z^{n-1}$. Hence,

$$P_{\mathrm{D}}^{\mathrm{o}} - P_{\mathrm{C}}^{\mathrm{o}} = c(1-z^{n-1}). \tag{S2}$$

The average payoff in the population is then

$$\bar{P}^{\mathrm{o}} = (1-z^{n-1})[(r-1)cx - (1-z)g]. \tag{S3}$$

We now turn to the two cases with positive or negative incentives. The total incentive is assumed to be proportional to the group size m, and hence of the form mI. The coefficient, where $I > 0$, is the (potential) per capita incentive. When rewards are used as incentives for cooperation, the total incentive is shared equally among cooperators. Hence, each cooperator obtains a reward mI/m_{C}, where m_{C} denotes the number of cooperators in the group of m players. When penalties are used as incentives for cooperation, each defector analogously have his or her individual payoff reduced by mI/m_{D}, where m_{D} denotes the number of defectors in the group of m players ($m = m_{\mathrm{C}} + m_{\mathrm{D}}$).

First, we consider rewarding. Because defectors never receive rewards, we have $P_{\mathrm{D}}^{\mathrm{r}} = P_{\mathrm{D}}^{\mathrm{o}}$. In a group in which the $m-1$ co-players include k cooperators, switching from C to D implies losing a reward $mI/(k+1)$. Hence,

$$\begin{aligned} P_{\mathrm{D}}^{\mathrm{r}} - P_{\mathrm{C}}^{\mathrm{r}} &= (P_{\mathrm{D}}^{\mathrm{o}} - P_{\mathrm{C}}^{\mathrm{o}}) \\ &\quad - \sum_{m=2}^{n} \binom{n-1}{m-1} (1-z)^{m-1} z^{n-m} \left[\sum_{k=0}^{m-1} \binom{m-1}{k} \left(\frac{x}{1-z}\right)^{k} \left(\frac{y}{1-z}\right)^{m-1-k} \frac{mI}{k+1} \right] \\ &= (P_{\mathrm{D}}^{\mathrm{o}} - P_{\mathrm{C}}^{\mathrm{o}}) - I\left[(1-z^{n-1}) + \frac{y}{x}(1-(1-x)^{n-1}) \right], \end{aligned} \tag{S4}$$

and thus,

$$\bar{P}^{\mathrm{r}} = \bar{P}^{\mathrm{o}} + I[x(1-z^{n-1}) + y(1-(1-x)^{n-1})]. \tag{S5}$$

Next, we consider punishing. It is now the cooperators who are unaffected, implying $P_{\mathrm{C}}^{\mathrm{p}} = P_{\mathrm{C}}^{\mathrm{o}}$. In a group in which the $m-1$ co-players include k cooperators (and thus, $m-1-k$ defectors), switching from C to D entails a penalty $mI/(m-k)$. Hence,

$$P_{\mathrm{D}}^{\mathrm{p}} - P_{\mathrm{C}}^{\mathrm{p}} = (P_{\mathrm{D}}^{\mathrm{o}} - P_{\mathrm{C}}^{\mathrm{o}}) - I\left[(1-z^{n-1}) + \frac{x}{y}(1-(1-y)^{n-1}) \right], \tag{S6}$$

and thus,

$$\bar{P}^{\mathrm{p}} = \bar{P}^{\mathrm{o}} - I[y(1-z^{n-1}) + x(1-(1-y)^{n-1})]. \tag{S7}$$

S2 Analytical Investigation of Game Dynamics

The replicator equations for the frequencies of three strategies are

$$\begin{aligned} \dot{x} &= x(P_{\mathrm{C}}^{s} - \bar{P}^{s}), \\ \dot{y} &= y(P_{\mathrm{D}}^{s} - \bar{P}^{s}), \\ \dot{z} &= z(P_{\mathrm{N}}^{s} - \bar{P}^{s}), \end{aligned} \tag{S8}$$

where the dots denote time derivatives. The frequencies x, y, and z can vary within the state space Δ, given by the combination of all (x,y,z) with $x,y,z \geq 0$ and $x+y+z=1$. As a first step, it is easy to understand the dynamics on the three edges of Δ. On the CD-edge, on which $z=0$, the dynamics correspond to compulsory participation; thus, the system of replicator equations reduces to $\dot{x} = -x(1-x)(P_{\mathrm{D}}^{s} - P_{\mathrm{C}}^{s})$. With rewarding, the difference in average payoff between a defector and a cooperator is

$$P_{\mathrm{D}}^{\mathrm{r}} - P_{\mathrm{C}}^{\mathrm{r}} = \frac{1}{1-y}[c(1-y) - I(1-y^{n})] = c - I\sum_{i=0}^{n-1} y^{i}, \tag{S9}$$

whereas, with punishing, the corresponding difference is

$$P_{\mathrm{D}}^{\mathrm{p}} - P_{\mathrm{C}}^{\mathrm{p}} = \frac{1}{1-x}[c(1-x) - I(1-x^{n})] = c - I\sum_{i=0}^{n-1} x^{i}. \tag{S10}$$

Because $I > 0$, the difference $P_{\mathrm{D}}^{\mathrm{r}} - P_{\mathrm{C}}^{\mathrm{r}}$ strictly increases, and $P_{\mathrm{D}}^{\mathrm{p}} - P_{\mathrm{C}}^{\mathrm{p}}$ strictly decreases, with $x = 1 - y$. The condition that there exists an interior equilibrium R on the CD-edge is

$$I_- < I < I_+ \quad \text{with} \quad I_- = c/n \quad \text{and} \quad I_+ = c\,. \tag{S11}$$

The dynamics on the two other edges are unidirectional: On the NC-edge, the dynamics always lead from N to C, and on the DN-edge, they always lead from D to N.

Having understood the dynamics on the three edges, we now consider the interior of Δ. We start by proving that if an interior equilibrium Q exists for the system of replicator equations (Eq. S8), it is unique. For this purpose, we introduce the coordinate system (f, z) in $\Delta \backslash \{z = 1\}$, with $f = x/(x + y)$. Using $P_{\mathrm{N}}^s = 0$, we can write the system of replicator equations (Eq. S8) as

$$\begin{aligned} \dot{f} &= -f(1-f)(P_{\mathrm{D}}^s - P_{\mathrm{C}}^s), \\ \dot{z} &= -z\bar{P}^s. \end{aligned} \tag{S12}$$

At an interior equilibrium $\mathrm{Q} = (\hat{f}, \hat{z})$, the three strategies must have equal payoffs, which means that they must all equal 0 in our model. The conditions $P_{\mathrm{C}}^{\mathrm{p}} = 0$ and $P_{\mathrm{D}}^{\mathrm{r}} = 0$ imply that $\hat{f}$ is independent of $\hat{z}$, and is given by

$$\hat{f} = \frac{c+g}{rc} = \hat{f}_{\mathrm{p}} \quad \text{for punishing and} \quad \hat{f} = \frac{g}{rc} = \hat{f}_{\mathrm{r}} \quad \text{for rewarding.} \tag{S13}$$

Thus, an interior equilibrium Q, if it exists, must be located on the line given by

$$\frac{x}{y} = \frac{\hat{f}}{1-\hat{f}}\,. \tag{S14}$$

We next show that $\hat{z}$ is uniquely determined. We first consider punishing. The equation $P_{\mathrm{D}}^{\mathrm{p}} - P_{\mathrm{C}}^{\mathrm{p}} = 0$ has, at most, one solution with respect to z. Indeed, using Eq. S6, this equation can be rewritten as

$$\begin{aligned} & c(1 - z^{n-1}) - I\left[(1 - z^{n-1}) + \frac{x}{y}(1 - (1-y)^{n-1})\right] = 0 \\ \Leftrightarrow\ & (c - I)(1 - z^{n-1}) - I\left[\frac{f}{1-f}(1 - (f + (1-f)z)^{n-1})\right] = 0 \\ \Leftrightarrow\ & \frac{(c-I)(1-f)}{If} = \frac{1 - [f + (1-f)z]^{n-1}}{1 - z^{n-1}}. \end{aligned} \tag{S15}$$

We denote the right-hand side of the last line by $G(f, z)$ and note that $G(f, 0) = 1 - f^{n-1}$ and $G(f, 1) = \lim_{z \to 1} G(f, z) = 1 - f$. It is sufficient to show that $G(f, z)$ is strictly monotonic with respect to $z \in (0,1)$. A straightforward computation yields

$$\frac{\partial}{\partial z} G(f, z) = \frac{(n-1)}{(1 - z^{n-1})^2}\left[z^{n-2} - (f + (1-f)z)^{n-2}((1-f) + fz^{n-2})\right]$$

$$= \frac{(n-1)z^{n-2}}{(1-z^{n-1})^2}\left[1-\left(\frac{f+(1-f)z}{z}\right)^{n-2}\left((1-f)+fz^{n-2}\right)\right]$$
$$= \frac{(n-1)z^{n-2}}{(1-z^{n-1})^2}\left[1-\left[\left(\frac{f+(1-f)z}{z}\right)\left((1-f)+fz\right)\right]^{n-2}\frac{(1-f)+fz^{n-2}}{\left((1-f)+fz\right)^{n-2}}\right]. \quad \text{(S16)}$$

We note that

$$\left(\frac{f+(1-f)z}{z}\right)\left((1-f)+fz\right) = 1+f(1-f)\left(z-2+\frac{1}{z}\right) = 1+f(1-f)\frac{(1-z)^2}{z} > 1, \quad \text{(S17)}$$

and

$$\frac{(1-f)+fz^{n-2}}{((1-f)+fz)^{n-2}} \geq 1. \quad \text{(S18)}$$

This inequality obviously holds for $n = 2$, and, by induction, for every larger n: If it holds for n, it must hold for $n+1$, because

$$\frac{(1-f)+fz^{n+1}}{((1-f)+fz)^{n+1}} - \frac{(1-f)+fz^{n}}{((1-f)+fz)^{n}} = \frac{1}{((1-f)+fz)^{n+1}}$$
$$\times\left[(1-f)+fz^{n+1}-((1-f)+fz)((1-f)+fz^{n})\right]$$
$$= \frac{1}{\left((1-f)+fz\right)^{n+1}} f(1-f)(1-z)(1-z^{n})$$
$$> 0. \quad \text{(S19)}$$

Consequently,

$$1-\left[\left(\frac{f+(1-f)z}{z}\right)\left((1-f)+fz\right)\right]^{n-2}\frac{(1-f)+fz^{n-2}}{\left((1-f)+fz\right)^{n-2}} < 0. \quad \text{(S20)}$$

Thus, $\partial G/\partial z(f,z) < 0$ for every $z \in (0,1)$, which implies strict monotonicity of G in z.

We now consider rewarding. In this case, using Eq. S4, we can rewrite $P_{\mathrm{D}}^{\mathrm{r}} - P_{\mathrm{C}}^{\mathrm{r}} = 0$ as

$$c(1-z^{n-1}) - I\left[(1-z^{n-1}) + \frac{y}{x}(1-(1-x)^{n-1})\right] = 0$$
$$\Leftrightarrow (c-I)(1-z^{n-1}) - I\left[\frac{\bar{f}}{1-\bar{f}}\left(1-\left(\bar{f}+\left(1-\bar{f}\right)z\right)^{n-1}\right)\right] = 0$$
$$\Leftrightarrow \frac{(c-I)(1-\bar{f})}{I\bar{f}} = \frac{1-\left[\bar{f}+\left(1-\bar{f}\right)z\right]^{n-1}}{1-z^{n-1}}, \quad \text{(S21)}$$

where $\bar{f} = y/(x+y) = 1-f$. Using the same argument as above, we see that $P_{\mathrm{D}}^{\mathrm{r}} - P_{\mathrm{C}}^{\mathrm{r}} = 0$ has, at most, one solution with respect to z. This concludes our proof of the uniqueness of Q.

We next prove that the interior equilibrium Q is a saddle point. For this purpose, we investigate the local dynamics around Q. We first consider punishing. Dividing the right-hand side of Eq. S12 by $f(1-z^{n-1})$, which is positive in the interior of Δ, corresponds to a change of velocity and does not affect the shape of trajectories in Δ. This yields

$$\dot{f} = f\left(c - I - \frac{c-I}{f} + IG(f,z)\right),$$
$$\dot{z} = z(1-z)\left(-(r-1)c - I + \frac{g+I}{f} + IG(f,z)\right). \tag{S22}$$

Because the large parentheses above vanish at Q, the Jacobian at Q is given by

$$J_{\mathrm{Q}} = \left.\begin{pmatrix} f\left(\frac{c-I}{f^2} + I\frac{\partial G(f,z)}{\partial f}\right) & fI\frac{\partial G(f,z)}{\partial z} \\ z(1-z)\left(-\frac{g+I}{f^2} + I\frac{\partial G(f,z)}{\partial f}\right) & z(1-z)I\frac{\partial G(f,z)}{\partial z} \end{pmatrix}\right|_{\mathrm{Q}}. \tag{S23}$$

Using $\partial G/\partial z(f,z) < 0$, this yields

$$\det J_{\mathrm{Q}} = (c+g)I\frac{z(1-z)}{f}\frac{\partial G(f,z)}{\partial z} < 0. \tag{S24}$$

Hence, J_{Q} has eigenvalues that are real and of opposite sign. Therefore, the unique interior equilibrium Q is a saddle point, and is thus unstable.

We now consider rewarding. An appropriate change of velocity results from dividing the right-hand side of Eq. S12 by $(1-f)(1-z^{n-1})$, which yields

$$\dot{f} = (1-f)\left(c - I - \frac{c-I}{1-f} + IG(1-f,z)\right),$$
$$\dot{z} = z(1-z)\left((r-1)c + I - \frac{(r-1)c - g + I}{1-f} - IG(1-f,z)\right). \tag{S25}$$

Because the large parentheses above vanish at Q, the Jacobian at Q is given by

$$J_{\mathrm{Q}} = \left.\begin{pmatrix} (1-f)\left(-\frac{c-I}{(1-f)^2} + I\frac{\partial G(1-f,z)}{\partial f}\right) & (1-f)I\frac{\partial G(1-f,z)}{\partial z} \\ -z(1-z)\left(\frac{(r-1)c-g+I}{(1-f)^2} + I\frac{\partial G(1-f,z)}{\partial f}\right) & -z(1-z)I\frac{\partial G(1-f,z)}{\partial z} \end{pmatrix}\right|_{\mathrm{Q}}. \tag{S26}$$

From our assumption that $(r-1)c > g$, it follows that

$$\det J_{\mathrm{Q}} = (rc-g)I\frac{z(1-z)}{1-f}\frac{\partial G(1-f,z)}{\partial z} < 0. \tag{S27}$$

Therefore, the unique interior equilibrium Q is again a saddle point.

We turn now to the investigation of the boundary equilibrium R and the interior equilibrium Q. We first consider punishing. On the CD-edge ($z = 0$), we obtain from Eq. S7

$$\bar{P}^{\mathrm{p}} = rc\left(x - \hat{f}_{\mathrm{p}}\right) + c(1 - x) - I(1 - x^n). \tag{S28}$$

As the per capita incentive I increases, the equilibrium R enters the edge at C ($x = 1$) and then moves to D ($x = 0$). It is a repellor on the CD-edge. From Eq. S10, we see that $\mathrm{R} = (x_{\mathrm{R}}, y_{\mathrm{R}}, 0)$, with $y_{\mathrm{R}} = 1 - x_{\mathrm{R}}$ given by the (unique) solution of $c(1 - x_{\mathrm{R}}) - I(1 - x_{\mathrm{R}}^n) = 0$. Hence, the average payoff at R is

$$\bar{P}^{\mathrm{p}} = rc\left(x_{\mathrm{R}} - \hat{f}_{\mathrm{p}}\right). \tag{S29}$$

Because $\dot{z} = -z\bar{P}^{\mathrm{p}}$, R is stable against invasion by non-participants (and R is thus a saddle point), if $\hat{f} < x_{\mathrm{R}} < 1$. If, conversely, $0 < x_{\mathrm{R}} < \hat{f}$, R can be invaded (and R is thus a source).

We now consider rewarding. On the CD-edge, Eq. S5 yields

$$\bar{P}^{\mathrm{r}} = rc\left(x - \hat{f}_{\mathrm{r}}\right) - c(1 - y) + I(1 - y^n). \tag{S30}$$

As I increases, the equilibrium R enters the CD-edge through D ($x = 0$) and then moves to C ($x = 1$). It is an attractor on the CD-edge. Using Eq. S9, we see that $c(1 - y_{\mathrm{R}}) - I(1 - y_{\mathrm{R}}^n) = 0$ holds at R. A similar argument as before then implies that the average payoff at R is

$$\bar{P}^{\mathrm{r}} = rc\left(x_{\mathrm{R}} - \hat{f}_{\mathrm{r}}\right). \tag{S31}$$

R can be invaded by non-participants (and R is thus a saddle point), if $0 < x_{\mathrm{R}} < \hat{f}$. If, conversely, $\hat{f} < x_{\mathrm{R}} < 1$, the equilibrium R is protected against invasion (and R is thus a sink).

The interior equilibrium $\mathrm{Q} = (\hat{x}, \hat{y}, \hat{z})$ splits off from R when the per capita incentive I crosses the threshold value corresponding to $x_{\mathrm{R}} = \hat{f}$. Indeed, the right-hand side of Eqs. S15 and S21 is decreasing with respect to z. Moreover, the left-hand side of these equations is decreasing with respect to I (for $I < c$). This implies that $\hat{z}$, the unique solution of Eqs. S15 and S21, increases with I.

For punishing, Eq. S15 implies that $G\left(\hat{f}, 0\right) = 1 - \hat{f}_{\mathrm{p}}^{\,n-1}$. Thus,

$$I = \frac{c}{1 + \hat{f}_{\mathrm{p}} + \cdots + \hat{f}_{\mathrm{p}}^{\,n-1}} =: K_{-}, \tag{S32}$$

which is larger than $I_{-} = c/n$. Similarly, $G\left(\hat{f}, 1\right) = 1 - \hat{f}_{\mathrm{p}}$, and thus

$$I = \frac{c}{1 + \hat{f}_{\mathrm{p}}} =: K_{+}, \tag{S33}$$

which is smaller than $I_{+} = c$. Analogously, for rewarding, Eq. S21 implies that $G\left(1 - \hat{f}, 0\right) = 1 - (1 - \hat{f}_{\mathrm{r}})^{n-1}$, and thus

$$I = \frac{c}{1 + \left(1 - \hat{f}_{\mathrm{r}}\right) + \cdots + \left(1 - \hat{f}_{\mathrm{r}}\right)^{n-1}} =: J_{-}, \tag{S34}$$

which is larger than $I_- = c/n$. For $\hat{z} = 1$, we obtain $G(1-\hat{f}, 1) = 1-(1-\hat{f}_r) = \hat{f}_r$, and thus

$$I = \frac{c}{1+(1-\hat{f}_r)} = \frac{c}{2-\hat{f}_r} =: J_+, \tag{S35}$$

which is smaller than $I_+ = c$.

We now summarize the results obtained so far, in terms of the thresholds given by Eqs. S11 and S32-S35. As I increases, first, the boundary equilibrium R enters the CD-edge at one end, for $= I_-$, and then moves toward the other end. Next, for $I = K_- > I_-$, the equilibrium Q enters the state space Δ through R, at $(\hat{f}, 1-\hat{f}, 0)$. It then moves towards N along the line given by $(1-\hat{f})x = \hat{f}y$. Eventually, for $I = K_+ < I_+$, the equilibrium Q collides with N. For still larger values of I, Δ contains no interior equilibrium. Finally, R attains the other end of the CD-edge for $I = I_+$.

We note that the dynamics around the non-hyperbolic equilibrium N can be fully analyzed by the blowing-up technique, using $x = f(1-z)$ and $y = (1-f)(1-z)$. This will be the subject of a separate analysis.

S3 Self-Returning Variant of Public Good Games

We next turn to a variant of public good games, called self-returning, in which the contribution of a player is multiplied by a factor $r > 1$ and then divided among all players (including the contributor, who therefore receives a fraction r/m in return). The social dilemma vanishes, in this case, if $r > m$. For the case without incentives, we can use known results (18, 19). A defector in a group with $m-1$ co-players ($m = 2, \cdots, n$) obtains from the common good a payoff of $rcx/(1-z)(1-m^{-1})$ on average. Hence,

$$P_D^o = -(1-z^{n-1})g + rc\frac{x}{1-z}\left(1-\frac{1-z^n}{n(1-z)}\right). \tag{S36}$$

Switching from C to D yields a difference in payoff of $c(1-r/m)$ in a group with $m-1$ co-players. This leads to

$$P_D^o - P_C^o = c + (r-1)cz^{n-1} - \frac{rc}{n}\frac{1-z^n}{1-z}. \tag{S37}$$

The average payoff in the population is then

$$\bar{P}^o = (1-z^{n-1})[(r-1)cx - (1-z)g], \tag{S38}$$

matching Eq. S3 for our main model (the 'others-only' variant). Also, the payoffs originating from the incentive mechanism are the same in both model variants.

Without incentives, the three strategies form a rock-scissors-paper cycle, as shown in Fig. S1*a*. For $2 < r < n$, the three strategies undergo periodic oscillations around an equilibrium, a center we denote by P. If $1 < r \leq 2$, just as in the others-only variant, all orbits issue from, and then again converge to, the state $z = 1$ of non-participation. In that case, cooperation can only emerge in brief bursts. In each case, the time average of all payoffs is 0.

It is our analytic result that with increasing I, an equilibrium R appears on the CD-edge, issuing from one end and moving to the other, just as in the 'others-only' case. The only difference is that the threshold values are now given by $I_- = c(1 - r/n)/n$ and $I_+ = c(1 - r/n)$, instead of by $I_- = c/n$ and $I_+ = c$.

According to numerical simulations, rewarding stabilizes the center P (Fig. S1*b*) as long as $2 < r < n$. For small I, P is a global attractor. The fraction of cooperators at Q is higher with than without rewarding, but the average payoff at Q remains equal to 0 in both cases. As I increases and exceeds I_-, the equilibrium R appears on the CD-edge. It is stable within that edge. However, as long as I is not too large, R can be invaded by non-participants, such that P remains the global attractor (Fig. S1*c*). When I reaches a critical value, P collides with R. For larger I, R becomes the global attractor (Fig. S1*d*). As I increases beyond I_+, the stable equilibrium R merges with C and all trajectories converge to C, just as in the case of punishment (Fig. S1*h*).

In contrast, punishing destabilizes the center P (Fig. S1*e*). For small I, all trajectories in the interior of the state space converge to the cycle on the boundary, staying in the vicinity of N for most of the time. As I increases and exceeds I_-, the equilibrium R appears on the CD-edge. It is a source, and C becomes a global attractor (Fig. S1*f*). This still holds after P has collided with R (Fig. S1*g*). For $I \leq I_-$, the time average of the frequency of cooperation, as well as the time average of the mean payoff in the population, remain 0. However, for $I > I_-$, these two averages increase to 1 and $(r - 1)c - g$, respectively.

For $1 < r \leq 2$, there is no equilibrium in the interior of the state space, as long as I is small. If I increases beyond a certain threshold, the equilibrium P enters the state space through N. It is an attractor in the case of rewarding and a repellor in the case of punishing. The further development, for increasing I, closely resembles that in the analysis above for $2 < r < n$.

So far, we have described Fig. S1. For a narrow range of parameter values, numerical investigations show that an additional twist can occur as a subplot of the self-returning variant (both with rewarding and with punishing) through the appearance of a second equilibrium Q in the interior of the state space, in addition to P (Fig. S2). As I increases, Q enters the state space through R (which thus turns into a sink with rewarding and into a source with punishing). As I increases further, P and Q approach each other and, when they collide, disappear in a saddle-node bifurcation. With punishing, the vertex C representing full cooperation remains a global attractor; thus, the long-term outcome is not affected. With rewarding, R resumes its role as a global attractor after the two interior equilibria have annihilated each other.

S4 Variants in the Incentive Scheme

We can investigate some variants in the incentive scheme. The underlying public good game, again, is the others-only variant, as in the main text.

First, we relax our assumption that the per capita penalty decreases proportionally with the number of defectors. For example, in many real-life situations, the size of the penalty is constant, and thus does not depend on how many players misbehave. Another special case is that of 'exemplary punishment': One defector has to pay the maximal penalty mI, whereas the other $m_{\mathrm{D}} -$

1 defectors have to pay no penalty. In this case, the expected penalty is still mI/m_{D}, just as analyzed in the main text and Sections S1 and S2. More generally, however, it makes sense to assume that if the sanctioning institution spends some resources on executing the punishment of a defector (e.g., by consuming time to process a ticket), it has less resources available for penalizing other defectors. In general, law-enforcers, on meeting defrauders, need some time to deal with them before resuming their chase for other abusers. This means that the chance for getting caught, and hence the expected penalty, is reduced if there are many defectors.

Borrowing the notion of 'handling time' used to study predatory behavior (22), we are led to model the size of the expected penalty as proportional to $mI/(a + bm_{\mathrm{D}})$, with two positive constants a and b. Depending on the ratio a/b, we can obtain a continuum of cases that include as limits a constant expected penalty ($b = 0$) and an expected penalty that is inversely proportional to the number of free-riders m_{D} ($a = 0$). For simplicity, we assume that $b = h$ and $a = 1 - h$ with $0 \leq h \leq 1$. If the handling time h decreases, the model smoothly transforms, from the inversely proportional case ($h = 1$) considered so far to the case of a constant punishment ($h = 0$). Investigating this generalization numerically, we find that the general outcome of our model remains unchanged, whereas the size of the interval (I_{-}, I_{+}) in which compulsory participation causes alternative stable states decreases with h. It is only in the limiting case $h = 0$ that this interval vanishes. Indeed, for $h = 0$, cooperation gets established if and only if $I > c/n$, no matter whether participation is optional or compulsory.

These conclusions also apply to rewarding. This means that our main result, that full cooperation is achieved at a much lower cost through negative incentives, is robust.

As a further robustness check, we can assume that there is a ceiling, $u > 0$, for the magnitude of the penalty or reward imposed on any one individual player. This results in a piecewise function for the per capita incentive. Once more, numerical investigations confirm that our results are qualitatively unaffected by this variation.

S5 User-Pays Variant

As a further variant, we can assume that in addition to the participation fee g, participants are obliged to pay a fee aI with $a > 0$ for the institution providing the incentives. We call this the user-pays variant: Players are obliged to come up with the total incentive. The expected payoff for a participant is thus reduced by $aI(1 - z^{n-1})$, with $1 - z^{n-1}$ being the probability that the public good game takes place. This leads to the following changes: With rewarding, the expected payoffs equal $P_{\mathrm{D}}^{\mathrm{r}} = P_{\mathrm{D}}^{\mathrm{o}} - aI(1 - z^{n-1})$, and

$$\bar{P}^{\mathrm{r}} = \bar{P}^{\mathrm{o}} + I[x(1 - z^{n-1}) + y(1 - (1 - x)^{n-1})] - aI(1 - z)(1 - z^{n-1}), \tag{S39}$$

whereas with punishing, they equal $P_{\mathrm{C}}^{\mathrm{p}} = P_{\mathrm{C}}^{\mathrm{o}} - aI(1 - z^{n-1})$, and

$$\bar{P}^{\mathrm{p}} = \bar{P}^{\mathrm{o}} - I[y(1 - z^{n-1}) + x(1 - (1 - y)^{n-1})] - aI(1 - z)(1 - z^{n-1}). \tag{S40}$$

The payoff difference between cooperators and defectors, $P_{\mathrm{D}}^{S} - P_{\mathrm{C}}^{S}$, obviously remains unaffected, as does the evolutionary dynamics on the CD-edge. Numerical results show the following.

With rewarding, optional participation increases the group welfare only marginally to 0, for such a small range of I that $I_- < I < J_-$ (Fig. 3*b*), in which compulsory participation causes the negative average payoffs. In the range, combining rewarding with optional participation even reduces the cooperator frequency to 0 (Fig. 3*a*). With punishing, the situation is very different. The group welfare is highest when I just barely exceeds the minimum $I_- = c/n$ required to obtain full cooperation (Fig. 3*d*). In this case, the learning process identifying the most efficient per capita incentive I will take some time; however, in the end, the cooperative norm will prevail (Fig. 3*c*).

As a further robustness check, we can examine a refund scheme for this user-pays sanctioning system. We consider an institution that punishes defectors; however, when there are none, that institution returns the fee aI to all participants. In this case, there are no 'lost deposits'. Clearly, this refinement renders the punitive protection of cooperators from free-riders less expensive. In particular, the value of the threshold I_- becomes smaller; thus, full cooperation is ensured with smaller per capita incentives I. Moreover, this refinement also avoids the reduction otherwise occurring in social welfare when the per capita incentive I is unnecessarily large, being not accurately matched to the optimal value I_- (Fig. 3*d*). In other words, this refinement guarantees maximal social welfare for any $I > I_-$ also in the user-pays variant.

Figure legends

Figure S1 | Effects of institutional rewarding and punishing on the 'self-returning' optional public good game for different per capita incentives I, when $2 < r < n$. (*a*) Without incentives, the interior equilibrium P is a center surrounded by closed trajectories. (*b*-*d*) With rewarding, the interior equilibrium P is stable. In *b* and *c*, it is a global attractor. In *c*, the CD-edge contains a saddle point R which can be invaded by non-participants. In *d*, P has reached the boundary and merged with R, turning it into a global attractor. (*e*-*g*) With punishing, P is unstable. In *f* and *g*, C is a global attractor. In *e*, trajectories stay in the vicinity of N for most of the time. In *f*, the CD-edge contains a saddle point R. In *g*, P has reached the boundary and merged with R, turning it into a source. (*h*) For very large incentives, full cooperation prevails. For very small or no incentives (*a*, *b*, and *e*), the average payoff equals 0 independent of the incentive used. Parameters: $n = 5$, $r = 3$, $c = 1$, $g = 0.5$, and $I = 0$ (*a*); 0.07 (*b* and *e*); 0.1 (*c* and *f*); 0.3 (*d* and *g*); or (punishing) 0.5 (*h*).

Figure S2 | Multiple interior equilibria. For a narrow range of parameter values, optional 'self-returning' public good games with incentives can exhibit two interior equilibria. (*a*) With rewarding, these equilibria are an attractor P and a saddle point Q. The boundary equilibrium R is a sink. The dynamics have alternative outcomes: Trajectories converge either to P or to R, depending on initial conditions. (*b*) With punishing, the two interior equilibria are a source P and a saddle point Q. C is an attractor, and the boundary equilibrium R is a source. Parameters: $n = 5$, $c = 1$, $r = 1.5$, $I = 0.2$, and $g = 0.2$ (*a*) or 0.3 (*b*).

Fig. S1

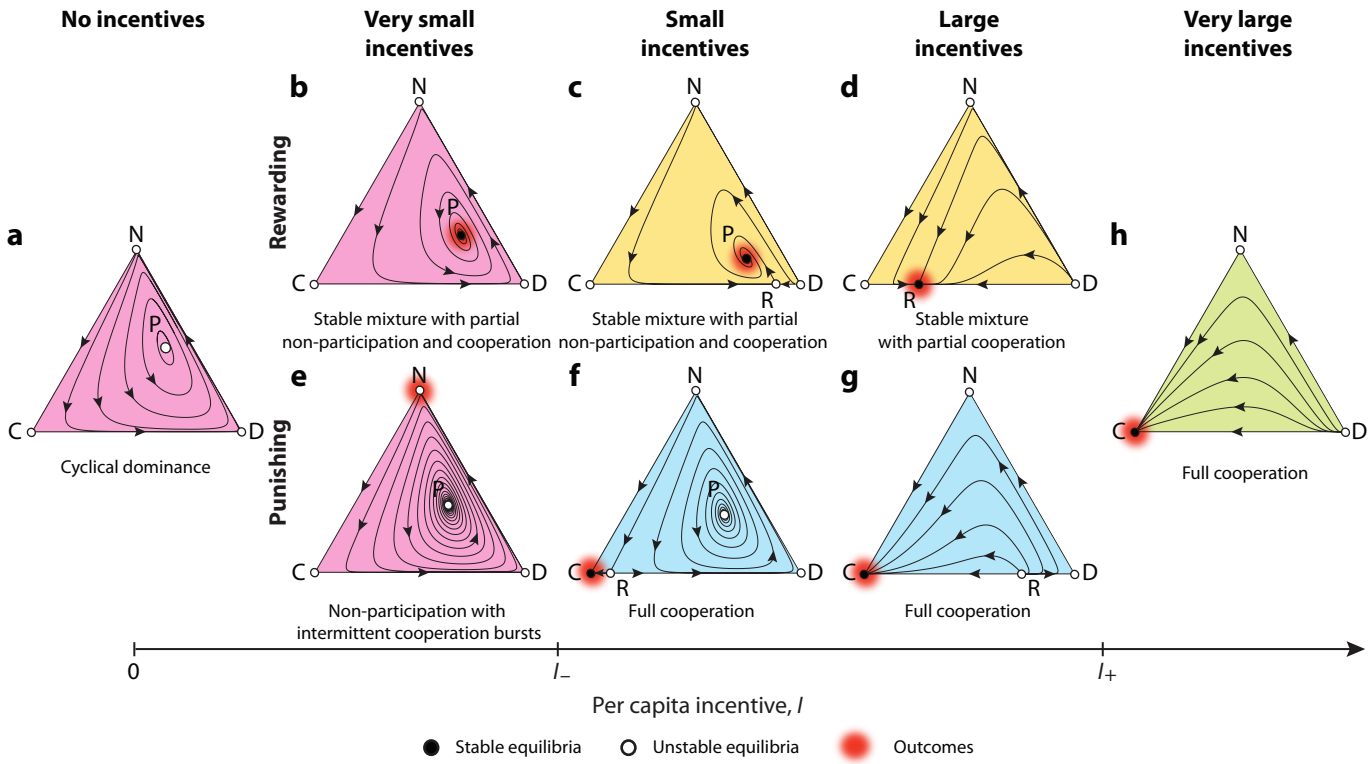
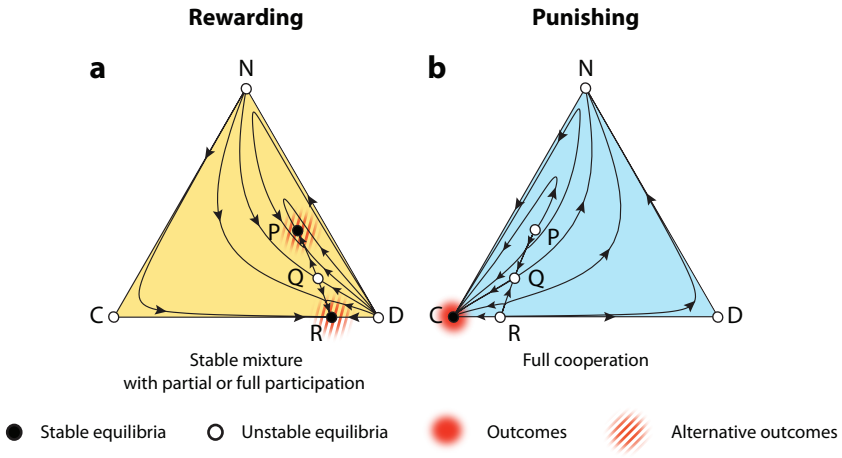

No incentives
Very small incentives
Small incentives
Large incentives
Very large incentives
a
b
c
d
h
e
f
g
Rewarding
Punishing
N
C
D
R
P
Cyclical dominance
Stable mixture with partial non-participation and cooperation
Stable mixture with partial non-participation and cooperation
Stable mixture with partial cooperation
Full cooperation
Non-participation with intermittent cooperation bursts
Full cooperation
Full cooperation
0
I_-
I_+
Per capita incentive, I
Stable equilibria
Unstable equilibria
Outcomes

Fig. S2


Rewarding
Punishing
a
b
N
C
D
P
Q
R
Stable mixture
with partial or full participation
Full cooperation
Stable equilibria
Unstable equilibria
Outcomes
Alternative outcomes