The Evolution of Antisocial Punishment in Optional Public Goods Games

Citation	Rand, David G. and Martin A. Nowak. 2011. The evolution of antisocial punishment in optional public goods game. Nature Communications 2(8): 434.
Published Version	doi:10.1038/ncomms1442
Accessed	February 19, 2015 9:25:26 AM EST
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:8207502
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

The evolution of anti-social punishment in optional public goods games

David G. Rand[1,2] & Martin A. Nowak[1,3,4]

[1]Program for Evolutionary Dynamics, [2]Department of Psychology, [3]Department of Mathematics, and [4]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA.

Author contributions: DR and MN designed and executed the research and wrote the paper.

The authors have no competing interests to declare.

Editor's summary: *Anti-social punishment, where non-cooperators punish cooperators, is a puzzling empirical phenomenon excluded from most theoretical models. Here anti-social punishment is added to an optional public goods game, revealing that evolution favors anti-social punishment and punishment does not promote cooperation.*

Abstract

Cooperation, where one individual incurs a cost to help another, is a fundamental building block of the natural world and of human society. It has been suggested that costly punishment can promote the evolution of cooperation, with the threat of punishment deterring free-riders. Recent experiments, however, have revealed the existence of 'anti-social' punishment, where non-cooperators punish cooperators. While various theoretical models find that punishment can promote the evolution of cooperation, these models *a priori* exclude the possibility of anti-social punishment. Here we extend the standard theory of optional public goods games to include the full set of punishment strategies. We find that punishment no longer increases cooperation, and that selection favors substantial levels of anti-social punishment for a wide range of parameters. Furthermore, we conduct behavioral experiments, which lead to results that are consistent with our model predictions. As opposed to an altruistic act that promotes cooperation, punishment is mostly a self-interested tool for protecting oneself against potential competitors.

Introduction

Explaining the evolution of cooperation is a topic of importance to both biologists and social scientists, and significant progress has been made in this area[1-6]. Various mechanisms such as reciprocal altruism, spatial selection, kin selection and multi-level selection have been proposed to explain the evolution of cooperation. In addition to these mechanisms, the role of costly punishment in promoting cooperation has received much attention. Many behavioral experiments have demonstrated that people are willing to incur costs to punish others[7-14]. Complementing these empirical findings, several evolutionary models have been developed to explore the potential effect of punishment on promoting cooperation[15-20]. Many researchers have concluded that the propensity to punish can encourage cooperation, although this position has not gone unquestioned[12,21-24].

The positive role of punishment has been challenged by recent experimental work that shows the existence of a more sinister form of punishment: sometimes non-cooperators punish cooperators[14,24-31]. In western countries, this 'anti-social punishment' is generally rare, except when it comes in the form of retaliation for punishment received in repeated games[14,24-26,28]. A series of cross-cultural experiments, however, finds substantial levels of anti-social punishment which cannot be explained by explicit retaliation[27,30,31] (see Supplementary notes for further analysis). It is this phenomenon of punishment targeted at cooperators, rather than explicit retaliation, which is the focus of our paper.

Anti-social punishment is puzzling, as it is inconsistent with both rational self-interest and the hypothesis that punishment facilitates cooperation. Social preference models of economic decision-making also predict that it should not occur[32-35]. Due to its seemingly illogical nature, anti-social punishment has been excluded *a priori* from most previous theoretical models for the evolution of cooperation, which only allow cooperators to punish defectors (exceptions include refs 22,23,36,37). Yet empirically, sometimes cooperators are punished, raising interesting evolutionary questions. What are the effects of anti-social punishment on the co-evolution of punishment and cooperation? And can the punishment of cooperators be explained in an evolutionary framework?

In this paper, we extend the standard theory of optional Public Goods Games[17-19,38] to explore anti-social punishment of cooperators. We study a finite population of N individuals. In each round of the game, groups of size n are randomly drawn from the population to play a one-shot optional public goods game followed by punishment. Each player chooses whether to participate in the public goods game as a cooperator (C) or defector (D), or to abstain from the public goods game and operate as a loner (L). Each cooperator pays a cost c to contribute to the public good, which is multiplied by a factor $r > 1$, and split evenly among all participating players in the group. Loners pay no cost and receive no share of the public good, but instead receive a fixed payoff σ. This loner's payoff is less than the $(r-1)c$ payoff earned in a group of all cooperators, but greater than the 0 payoff earned in a group of all defectors. If only one group member chooses to participate, then all group members receive the loner's payoff σ. Following the public goods game, each player has the opportunity to punish any or all of the $n-1$ other members of the group. A given player pays a cost γ for each other player he chooses to punish, and incurs a cost β for each punishment that he receives ($\gamma < \beta$).

Each of the N players has a strategy, which specifies her action in the public goods game (C, D or L). Each player also has a decision rule for the punishment round that specifies whether she punishes those members of her group who cooperated in the public goods game, those who defected, or those who opted out. For example, a C-NPN strategist cooperates in the public goods game, does not punish cooperators, punishes defectors, and does not punish loners; and an L-PNN strategist opts out of the public goods game, punishes cooperators, and does not punish defectors or loners. In total there are 24 strategies. We contrast this full strategy set with the limited strategy set that has been considered before[17-19]. The limited set has only four strategies: cooperators that never punish, defectors that never punish, loners that never punish, and cooperators that punish defectors.

We study the transmission of strategies through an evolutionary process, which can be interpreted either as genetic evolution or as social learning. In both cases, strategies which earn higher payoffs are more likely to spread in the population, while lower payoff strategies tend to die out. Novel strategies are introduced by mutation in the case of genetic evolution, or innovation and experimentation in the case of social learning. We use a frequency dependent Moran process[39] with an exponential payoff function[40]. We perform exact numerical calculations in the limit of low mutation[41,42], which characterizes genetic evolution and the long-term evolution of societal norms, as well as agent based simulations for higher mutation rates which may be more appropriate for short-term learning and exploration dynamics[19,43] (see Supplementary Methods for details).

In summary, we find that although punishment dramatically increases cooperation when only cooperators can punish defectors, this positive effect of punishment disappears almost entirely when the full set of punishment strategies is allowed. Just as punishment protects cooperators from invasion by defectors, it also protects defectors from invasion by loners, and loners from invasion by cooperators. Thus punishment is not 'altruistic' or particularly linked to cooperation. Instead natural selection favors substantial amounts of punishment targeted at all three public goods game actions, including cooperation. Furthermore, we find that the parameter sets which lead to high levels of cooperation (and little anti-social punishment) are those with efficient public goods (large r) and very weak punishment (small β). Finally, we generate testable predictions using our evolutionary model, and present preliminary experimental evidence which is consistent with those predictions.

Results

Effect of allowing the full set of punishment strategies

In the absence of punishment, defectors invade cooperators, loners invade defectors and cooperators invade loners[44], as in a rock-paper-scissors cycle (Figure 1a). The system spends a similar amount of time in each of the three behavioral states, and cooperation is not the dominant outcome (although there is more cooperation than in the game without loners). If cooperators are allowed to punish defectors, however, the cooperator-defector-loner cycle is broken when the system reaches punishing cooperators (Figure 1b). For this limited strategy set, the population spends the vast majority of its time in a cooperative state[17].

But what happens when all punishment strategies are available? Now the cooperator-defector-loner cycle can be broken as easily in the loner or defector states as in a cooperative state (Figure 1c,d). Without punishment, loners are invaded by cooperators; but loners that punish cooperators are protected from such an invasion. Similarly, defectors are invaded by loners; but defectors who punish loners are protected. Thus, when all punishment strategies are available, the dynamics effectively revert back to the original cooperator-defector-loner cycle. The salient difference is that now the most successful strategies use punishment against threatening invaders. We see that adding punishment does not provide much benefit to cooperators once the option to punish is available to all individuals, instead of being artificially restricted to cooperators punishing defectors. Furthermore, in this light non-cooperative strategies that pay to punish cooperators seem less surprising, and we see why natural selection can lead to the evolution of anti-social punishment.

Robustness to parameter variation

These results are not particular to the parameter values used in Figure 1. We have examined the steady state frequency of each strategy averaged over 100,000 randomly sampled parameter sets (see Supplementary notes). The outcome is remarkably similar to what is observed in Figure 1. The average level of cooperation is 34% in the absence of punishment, jumps to 87% with restricted punishment, and falls back to 34% with the full punishment strategy set. Although restricted punishment makes cooperation the dominant outcome for the vast majority of parameter sets, the full punishment strategy set does not (Figure 2). For more than 98% of the randomly chosen parameter sets, the frequency of cooperation is below 0.5.

We also find that the evolutionary success of anti-social punishment in the full strategy set is robust to variation in the payoff values. The frequencies of all three forms of punishment averaged over the 100,000 parameter sets are quite similar (punish cooperators, 40%; punish defectors, 41%, punish loners, 37%), and the frequency of each punishment type varies relatively little across parameter sets.

Furthermore, these results are not unique to the low mutation limit. Agent based simulations for higher mutation rates show that (i) punishment does not increase the average frequency of cooperation in the full strategy set; and (ii) all three forms of punishment (anti-C, anti-D and anti-L) have similar average frequencies (see Supplementary notes for details).

Discussion

Here we have shown have evolution can lead to punishment targeted at cooperators. We find that in our framework, loners are largely responsible for this anti-social punishment, and that it protects them against invasion by cooperators. The concept of loners punishing cooperators may seem strange given that loners could be envisioned as trying to avoid interactions with others. However, we find strong selection pressure in favor of such behavior: loners who avoid others in the context of the public goods game, but subsequently seek out and punish cooperators, outcompete fully solitary loners.

Our findings raise serious questions about the commonly held view that punishment promotes the evolution of cooperation. There is no reason to assume *a priori* that only cooperators punish others. In fact, there is substantial empirical evidence to the contrary[24-31]. As we have shown, using the full strategy set dramatically changes the evolutionary outcome, and punishment no longer increases cooperation. These results highlight the importance of not restricting analysis to a subset of strategies, and emphasize the need to reexamine other models of the co-evolution of punishment and cooperation while including all possible punishment strategies[23,36].

But if our framework suggests that the taste for punishment did not evolve to promote cooperation, then why do humans display the desire to punish? And why do many behavioral experiments find that punishment discourages free-riding in the lab[7-12,14]? As opposed to being particularly suited to protecting cooperators from free-riders, our model suggests that costly punishment is an effective tool for subduing potential invaders of any kind. This finding is reminiscent of early work on the ability of punishment to stabilize disadvantageous norms[20], while adding the critical step of the emergence of punishing behavior. Our results are also suggestive of a type of in-group bias[45-50], as our most successful strategists punish those who are different from themselves, while not punishing those who are the same.

Therefore we would expect the level of anti-social punishment to vary depending on the makeup of the population. In populations with a high frequency of cooperators, such as those societies in which most previous behavioral experiments have been conducted, we anticipate punishment to be largely directed towards defectors. In populations where cooperators are less common, however, we expect higher levels of punishment targeted at cooperators. Consistent with this intuition, we find a clear inverse relationship between steady state cooperation frequency and anti-social punishment across randomly sampled parameter sets (Figure 3a).

Examining specific parameter sets we find that a larger cooperation multiplier, r, increases cooperation and decreases anti-social punishment (Figure 3b). The importance of r is further demonstrated by a sensitivity analysis calculating the marginal effect of each parameter (see Supplementary notes). In addition to having large r, we find that punishment must be largely ineffective in order to achieve a high level of cooperation. Among the parameter sets in Figure 3a with steady state cooperation over 65%, the average effect of punishment is $\beta = 0.11$ (compared to an overall average of $\beta = 7.45$). Systematic parameter variation gives further evidence of an interaction between r and β (see Supplementary notes). When punishment is weak, a wide range of outcomes is possible, including high levels of cooperation if r is large. When punishment is strong, however, all strategies can effectively protect against invasion. Thus neutral drift between punishing and non-punishing strategies dominates the dynamics, and the range of outcomes is tightly constrained. These relationships between cooperation, anti-social punishment and the payoff parameters r and β are consistent with cross-cultural sociological evidence (see Supplementary notes). Exploring the connection between model parameters, sociological variation and play in experimental games is an important direction for future research across societies.

Taken together with data from cross-cultural experiments[27,30,31], our evolutionary model generates testable predictions about behavior in the laboratory. Most previous experiments on public goods have explored compulsory games, which do not offer the choice to abstain in favor of a fixed loner's payoff (an exception is ref [38] where many people do take the loner's payoff when offered). In the compulsory framework, cross-cultural experiments find evidence of low contributors who punish high contributors. Our model, which is based on an optional public goods game, finds that most punishment of cooperators comes from loners rather than defectors. Therefore, our evolutionary model makes two testable predictions about behavior in the lab: (i) that many low contributors in compulsory games will opt for the loner's payoff if given the chance; and (ii) that players who take the loner's payoff in an optional game will engage in more anti-social punishment in a compulsory game.

To begin evaluating these predictions, we use the internet to recruit participants [51,52] for two incentivized behavioral experiments[46,47] (see Methods for experimental details, and Supplementary notes for statistical analysis). In the first experiment, subjects engage in both an optional and a compulsory one-shot public goods game. Consistent with our first theoretical prediction, we find that subjects who choose the loner's payoff in the optional game contribute significantly less in the compulsory game (Figure 4a). Thus many subjects who appear to be defectors in the compulsory game prefer to be loners, and may be bringing intuitions evolved as loners to bear in the experiment. To test our second theoretical prediction, the second experiment has subjects participate in two one-shot public goods games, an optional game followed by a compulsory game with costly punishment. The results are again in agreement with the model's prediction: subjects who opt out of the optional game engage in significantly more punishment of high contributors (Figure 4b). Thus these experiments provide preliminary empirical evidence in support of our theoretical framework, although intuitions evolved in optional games are not the only possible explanation for the data. Further experimental work exploring cooperation and punishment in optional games, as well as the relationship between play in optional and compulsory games, is an important direction for future research.

We have also performed an evolutionary analysis of the compulsory game, where opting out is not possible (see Supplementary notes). Here we still find that anti-social punishment is favored by selection; that cooperation and anti-social punishment are inversely correlated; and that non-negligible amounts of pro- and anti-social punishment co-occur in many parameter sets. In the compulsory game, however, cooperation is never favored over defection using the full strategy set, and anti-social defectors are always the most common strategy. Thus, modeling the game exactly as it is performed in the experiments cannot explain the behaviors that are often observed in such experiments. The preferences displayed by subjects in these experiments must have evolved under circumstances that are somewhat different from those encountered in the experiments (see Supplementary notes). Adding the possibility to abstain leads to a model which *can* describe the range of experimental behaviors.

We have shown that although punishment does not increase cooperation or aggregate payoffs in our model, there is nonetheless an incentive to punish. Once punishment becomes available, it is essential for each strategy type to adopt it so as to protect against dominance by similarly armed others. As opposed to shifting the balance of strategies towards cooperation, punishment works to maintain the status quo. This maintenance, however, comes at a high price. Punishment is destructive for all parties and thus reduces the average payoff, without creating the benefit of increased cooperation. If all parties could agree to abandon punishment, everyone would benefit; but in a world without punishment, a strategy that switched to punishing potential invaders would dominate. Therefore choosing to punish is not altruistic in our framework, but rather self-interested.

Punishing leads to a tragedy of the commons where all individuals are forced to adopt punishment strategies. Abstaining from punishment becomes an act of cooperation, while using punishment is a form of second-order defection. The cooperative imperative is not the promotion of punishment, which is costly yet ineffective in our model, but instead the maintenance of cooperation through non-destructive means[12,53,54].

Methods

Experimental overview

Together with previous experiments on compulsory games[27], our model makes testable predictions about the behavior of experimental subjects: loners are predicted to be lower contributors in compulsory games, and to be most likely to punish high contributors. To evaluate these predictions, we conducted two incentivized behavioral experiments. Experiment 1 investigates the contribution behavior of loners in a compulsory game, while Experiment 2 considers the degree of anti-social punishment exhibited by loners in a compulsory game.

Recruitment using Amazon Mechanical Turk

Both experiments were conducted via the internet, using the online labor market Amazon Mechanical Turk (AMT) to recruit subjects. AMT is an online labor market in which employers contract workers to perform short tasks (typically less than 5 minutes) in exchange for small payments (typically less than $1). The amount paid can be conditioned on the outcome of the task, allowing for performance-dependent payments and incentive-compatible designs. AMT therefore offers an unprecedented tool for quickly and inexpensively recruiting subjects for economic game experiments. Although potential concerns exist regarding conducting experiments over the internet, numerous recent papers have demonstrated the reliability of data gathered using AMT across a range of domains[51,52,55-58]. Most relevant for our experiments are two studies using economic games. The first shows quantitative agreement in contribution behavior in a repeated public goods game between experiments conducted in the physical lab and those conducted using AMT with approximately 10-fold lower stakes[58]. The second replication again found quantitative agreement between the lab and AMT, this time in cooperation in a one-shot Prisoner's Dilemma[51]. It has also been shown that AMT subjects display a level of test-retest reliability similar to what is seen in the traditional physical laboratory on measures of political beliefs, self-esteem, Social Dominance Orientation, and Big-Five personality traits[56], as well as demographic variables such as belief in God, age, gender, education level and income[52,57]; that AMT subjects do not differ significantly from college undergraduates in terms of attentiveness or basic numeracy skills, and demonstrate similar effect sizes as undergraduates in tasks examining framing effects, the conjunction fallacy and outcome bias[55]; and are significantly more representative of the American population than undergraduates[56]. Thus there is ample reason to believe in the validity of experiments conducted using subjects recruited from AMT.

Experiment 1 design

In October 2010, 124 subjects were recruited through AMT to participate in Experiment 1, and assigned to either a treatment or control condition. Subjects received a $0.20 showup fee for participating, and earned on average an additional $0.63 based on decisions made in the experiment. First, subjects read a set of instructions for a one-shot public goods game, in which groups of 4 interact with a cooperation multiplier of $r = 2$, each choosing how much of a $0.40 endowment to contribute (in increments of 2 cents to avoid fractional cent amounts). Subjects then answered two comprehension questions to ensure they understood the payoff structure, and only those who answered correctly were allowed to participate.

The 73 subjects randomly assigned to the treatment group were then informed that the game was optional, and they could choose either to participate or to abstain in favor of a fixed payoff of $0.50. Those who chose to participate then indicated their level of contribution. Next, subjects were informed that they would play a second game with 3 new partners, which was a compulsory version of the first game. They were further informed that one of the two games (optional or compulsory) would be randomly selected to determine their payoff. This was done to keep the payoff range the same as in the control experiment described below, where subjects played only one game. In order to prevent between-game learning, subjects were not informed about the outcome of the optional game before making their decision in the compulsory game. To test if behavior in the compulsory game was affected by the preceding optional game, the remaining 51 subjects participated in a control condition in which they participated only in a compulsory public goods game. Payoffs were determined exactly as described (no deception was used).

Experiment 2 design

In November 2010, 196 subjects were recruited through AMT to participate in Experiment 2. Subjects received a \$0.40 showup fee for participating, and earned on average an additional \$0.92 based on decisions made in the experiment. As in Experiment 1, subjects began by reading a set of instructions for a one-shot public goods game with groups of 4, a cooperation multiplier of $r = 2$, and a \$0.40 endowment, and then answered two comprehension questions in order to participate. Subjects were then informed that the game was optional, and decided whether to participate or opt out for a fixed \$0.50 payment. Subjects who chose to participate were given 5 contribution levels to choose from: 0, 10, 20, 30 or 40. Thus the first game of Experiment 2 is identical to that of Experiment 1's treatment condition, except for a more limited set of contribution options to facilitate punishment decisions as described below.

Subjects were then informed that they would play a second game with 3 new partners, which differed from the first game in two ways: it was compulsory, and it would be followed by a Stage 2 in which participants could interact directly with each other group member. In Stage 2, subjects had three direct actions to pick from: choosing option A had no effect on either player; choosing option B caused them to lose 4 cents while the other player lost 12 cents; and choosing option C caused them to lose 8 cents while the other player lost 24 cents. Thus we offered mild (B) and severe (C) punishment options with a 1:3 punishment technology. Subjects were allowed to condition their Stage 2 choice on the other player's contribution in the compulsory public goods game. To do so, we employ the strategy method: subjects indicate which action (A, B, C) they would take towards group members choosing each possible contribution level (0, 10, 20, 30, 40). It has been shown that using the strategy method to elicit punishment decisions has a quantitative (although not qualitative) effect on the level of punishment targeted at defectors, but has little effect on punishment targeted at cooperators[59]. As anti-social punishment is our main focus, we therefore feel confident in our use of the strategy method. In order to prevent between-game learning, subjects were not informed about the outcome of the optional game before making their decision in the compulsory game. Payoffs were determined exactly as described (no deception was used).

See Supplementary notes for further details of the experimental setup, sample instructions, and detailed analysis of the results.

1 Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314**, 1560-1563 (2006).
2 Sigmund, K. *The calculus of selfishness*. (Princeton Univ Press, 2010).
3 Ostrom, E. *Governing the commons: The evolution of institutions for collective action*. (Cambridge Univ Pr, 1990).
4 Cressman, R. *The stability concept of evolutionary game theory: a dynamic approach*. (Springer-Verlag, 1992).

5 Helbing, D. & Yu, W. The outbreak of cooperation among success-driven individuals under noisy conditions. *Proceedings of the National Academy of Sciences* **106**, 3680-3685, doi:10.1073/pnas.0811503106 (2009).
6 Worden, L. & Levin, S. A. Evolutionary escape from the prisoner's dilemma. *Journal of theoretical biology* **245**, 411-422, doi:DOI: 10.1016/j.jtbi.2006.10.011 (2007).
7 Yamagishi, T. The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology* **51**, 110-116 (1986).
8 Ostrom, E., Walker, J. & Gardner, R. Covenants With and Without a Sword: Self-Governance is Possible. *The American Political Science Review* **86**, 404-417 (1992).
9 Fehr, E. & Gächter, S. Cooperation and punishment in public goods experiments. *American Economic Review* **90**, 980-994 (2000).
10 Gurerk, O., Irlenbusch, B. & Rockenbach, B. The competitive advantage of sanctioning institutions. *Science* **312**, 108-111 (2006).
11 Rockenbach, B. & Milinski, M. The efficient interaction of indirect reciprocity and costly punishment. *Nature* **444**, 718-723 (2006).
12 Rand, D. G., Dreber, A., Ellingsen, T., Fudenberg, D. & Nowak, M. A. Positive Interactions Promote Public Cooperation. *Science* **325**, 1272-1275 (2009).
13 Ule, A., Schram, A., Riedl, A. & Cason, T. N. Indirect Punishment and Generosity Toward Strangers. *Science* **326**, 1701-1704, doi:10.1126/science.1178883 (2009).
14 Janssen, M. A., Holahan, R., Lee, A. & Ostrom, E. Lab Experiments for the Study of Social-Ecological Systems. *Science* **328**, 613-617 (2010).
15 Boyd, R., Gintis, H., Bowles, S. & Richerson, P. J. The evolution of altruistic punishment. *Proc Natl Acad Sci USA* **100**, 3531-3535 (2003).
16 Nakamaru, M. & Iwasa, Y. The coevolution of altruism and punishment: Role of the selfish punisher. *Journal of theoretical biology* **240**, 475-488 (2006).
17 Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A. & Sigmund, K. Via Freedom to Coercion: The Emergence of Costly Punishment. *Science* **316**, 1905-1907 (2007).
18 Fowler, J. H. Altruistic punishment and the origin of cooperation. *Proc Natl Acad Sci U S A* **102**, 7047-7049 (2005).
19 Traulsen, A., Hauert, C., De Silva, H., Nowak, M. A. & Sigmund, K. Exploration dynamics in evolutionary games. *Proceedings of the National Academy of Sciences* **106**, 709-712 (2009).
20 Boyd, R. & Richerson, P. Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups. *Ethology and Sociobiology* **13**, 171-195 (1992).
21 Burnham, T. & Johnson, D. The evolutionary and biological logic of human cooperation. *Analyse & Kritik* **27**, 113-135 (2005).
22 Janssen, M. A. & Bushman, C. Evolution of cooperation and altruistic punishment when retaliation is possible. *Journal of theoretical biology* **254**, 541-545 (2008).
23 Rand, D. G., Armao IV, J. J., Nakamaru, M. & Ohtsuki, H. Anti-social punishment can prevent the co-evolution of punishment and cooperation. *Journal of theoretical biology* **265**, 624-632 (2010).
24 Dreber, A., Rand, D. G., Fudenberg, D. & Nowak, M. A. Winners don't punish. *Nature* **452**, 348-351 (2008).
25 Cinyabuguma, M., Page, T. & Putterman, L. Can second-order punishment deter perverse punishment? *Experimental Economics* **9**, 265-279 (2006).

26 Denant-Boemont, L., Masclet, D. & Noussair, C. Punishment, counterpunishment and sanction enforcement in a social dilemma experiment. *Economic Theory* **33**, 145-167 (2007).
27 Herrmann, B., Thoni, C. & Gächter, S. Antisocial punishment across societies. *Science* **319**, 1362-1367 (2008).
28 Nikiforakis, N. Punishment and Counter-punishment in Public Goods Games: Can we still govern ourselves? *Journal of Public Economics* **92**, 91-112 (2008).
29 Wu, J.-J. *et al.* Costly punishment does not always increase cooperation. *Proceedings of the National Academy of Sciences* **106**, 17448-17451, doi:10.1073/pnas.0905918106 (2009).
30 Gächter, S. & Herrmann, B. Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society B: Biological Sciences* **364**, 791-806 (2009).
31 Gächter, S. & Herrmann, B. The Limits of Self-Governance when Cooperators Get Punished: Experimental Evidence from Urban and Rural Russia. *Eur. Econ. Rev.*, doi:10.1016/j.euroecorev.2010.04.003 (In press).
32 Rabin, M. Incorporating Fairness into Game Theory and Economics. *The American Economic Review* **83**, 1281-1302 (1993).
33 Fehr, E. & Schmidt, K. A theory of fairness, competition and cooperation. *Quarterly Journal of Economics* **114**, 817-868 (1999).
34 Bolton, G. E. & Ockenfels, A. ERC: A Theory of Equity, Reciprocity, and Competition. *The American Economic Review* **90**, 166-193 (2000).
35 Dufwenberg, M. & Kirchsteiger, G. A theory of sequential reciprocity. *Games and Economic Behavior* **47**, 268-298 (2004).
36 Rand, D. G., Ohtsuki, H. & Nowak, M. A. Direct reciprocity with costly punishment: Generous tit-for-tat prevails *J Theor Biol* **256**, 45-57 (2009).
37 Sethi, R. Evolutionary stability and social norms. *Journal of economic behavior & organization* **29**, 113-140, doi:10.1016/0167-2681(95)00053-4 (1996).
38 Semmann, D., Krambeck, H.-J. & Milinski, M. Volunteering leads to rock-paper-scissors dynamics in a public goods game. *Nature* **425**, 390-393 (2003).
39 Nowak, M. A., Sasaki, A., Taylor, C. & Fudenberg, D. Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646-650 (2004).
40 Traulsen, A., Shoresh, N. & Nowak, M. Analytical Results for Individual and Group Selection of Any Intensity. *Bulletin of Mathematical Biology* **70**, 1410-1424 (2008).
41 Imhof, L. A., Fudenberg, D. & Nowak, M. A. Evolutionary cycles of cooperation and defection. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 10797-10800 (2005).
42 Fudenberg, D. & Imhof, L. A. Imitation processes with small mutations. *Journal of Economic Theory* **131**, 251-262 (2006).
43 Traulsen, A., Semmann, D., Sommerfeld, R. D., Krambeck, H.-J. & Milinski, M. Human strategy updating in evolutionary games. *Proceedings of the National Academy of Sciences* **107**, 2962-2966 (2010).
44 Hauert, C., De Monte, S., Hofbauer, J. & Sigmund, K. Volunteering as Red Queen Mechanism for Cooperation in Public Goods Games. *Science* **296**, 1129-1132 (2002).
45 Tajfel, H., Billig, R. P. & Flament, C. Social categorization and intergroup behavior. *European Journal of Social Psychology* **1**, 149-178 (1971).

46 Rand, D. G. *et al.* Dynamic Remodeling of In-group Bias During the 2008 Presidential Election. *Proc Natl Acad Sci U S A* **106**, 6187-6191 (2009).
47 Brewer, M. B. In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological bulletin* **86**, 307 (1979).
48 Yamagishi, T., Jin, N. & Kiyonari, T. Bounded Generalized Reciprocity: Ingroup Boasting and Ingroup Favoritism. *Advances in Group Processes* **16**, 161-197 (1999).
49 Fowler, J. H. & Kam, C. D. Beyond the Self: Social Identity, Altruism, and Political Participation. *Journal of Politics* **69**, 813-827 (2007).
50 Fershtman, C. & Gneezy, U. Discrimination in a Segmented Society: An Experimental Approach. *The Quarterly Journal of Economics* **116**, 351-377 (2001).
51 Horton, J. J., Rand, D. G. & Zeckhauser, R. J. The Online Laboratory: Conducting Experiments in a Real Labor Market. *Experimental Economics*, doi:10.1007/s10683-10011-19273-10689, doi:10.1007/s10683-011-9273-9 (2011).
52 Rand, D. G. The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments. *Journal of theoretical biology* **DOI: 10.1016/j.jtbi.2011.03.004** (2011).
53 Milinski, M., Semmann, D. & Krambeck, H. J. Reputation helps solve the 'tragedy of the commons'. *Nature* **415**, 424-426 (2002).
54 Ellingsen, T. & Johannesson, M. Anticipated verbal feedback induces altruistic behavior. *Evolution and Human Behavior* **29**, 100-105 (2008).
55 Paolacci, G., Chandler, J. & Ipeirotis, P. G. Running Experiments on Amazon Mechanical Turk. *Judgment and Decision Making* **5**, 411-419 (2010).
56 Buhrmester, M. D., Kwang, T. & Gosling, S. D. Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science* **6**, 3-5 (2011).
57 Mason, W. & Suri, S. Conducting Behavioral Research on Amazon's Mechanical Turk. *Available at SSRN: http://ssrn.com/abstract=1691163* (2010).
58 Suri, S. & Watts, D. J. Cooperation and Contagion in Web-Based, Networked Public Goods Experiments. *PLoS ONE* **6**, e16836 (2011).
59 Falk, A., Fehr, E. & Fischbacher, U. Driving Forces Behind Informal Sanctions. *Econometrica* **73**, 2017-2030 (2005).

Acknowledgements

We thank Anna Dreber, Simon Gachter, Christoph Hauert, Benedikt Herrmann, Karl Sigmund, and Arne Traulsen for helpful comments and suggestions, Martijn Egas, Benedikt Herrmann Simon Gachter for sharing their experimental data, and Nicos Nikiforakis for the use of his publicly available data. Support from the John Templeton Foundation, the NSF/NIH joint program in mathematical biology (NIH grant R01GM078986), the Bill and Melinda Gates Foundation (Grand Challenges grant 37874), and J. Epstein is gratefully acknowledged.

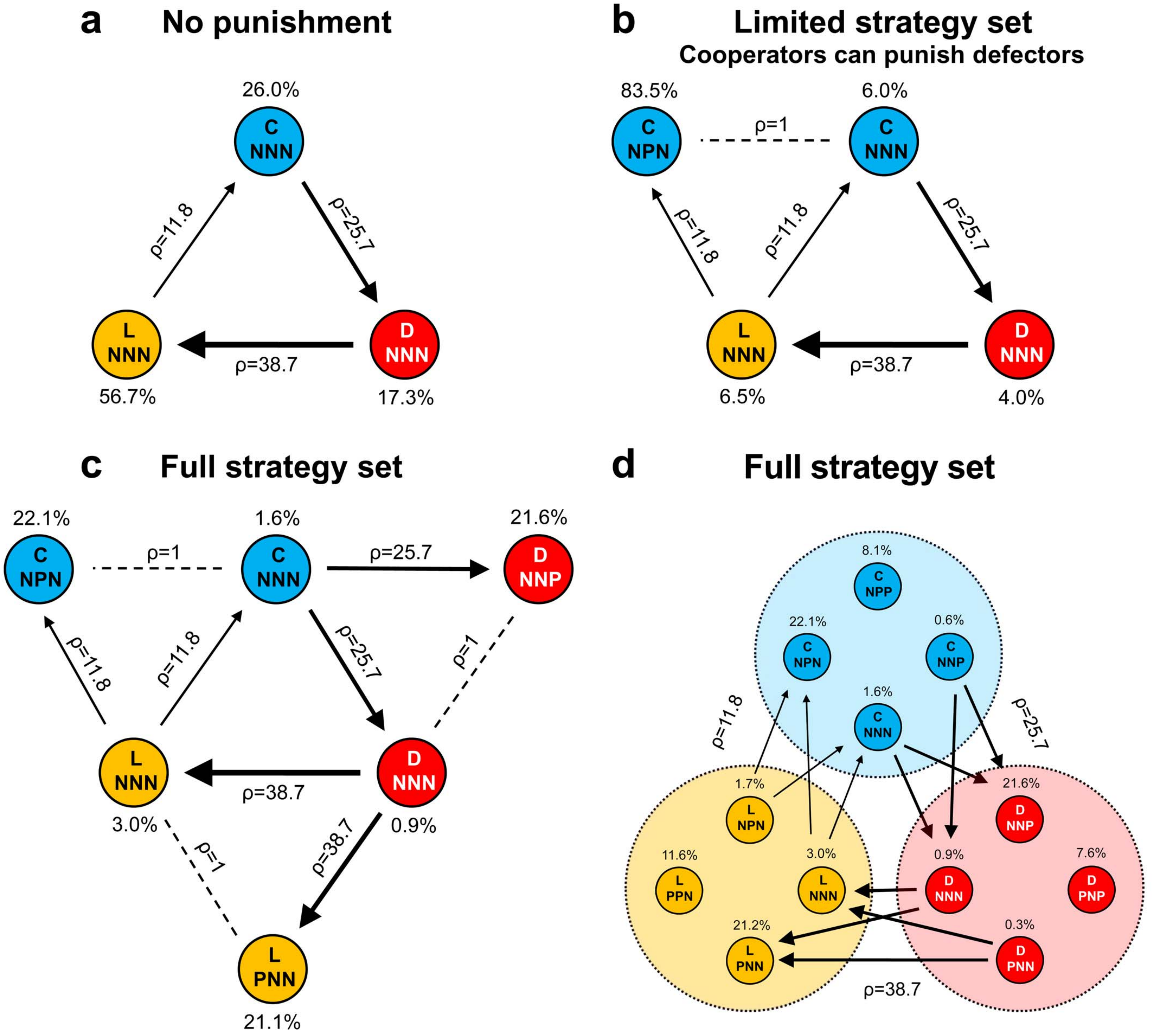
Figure and Table legends

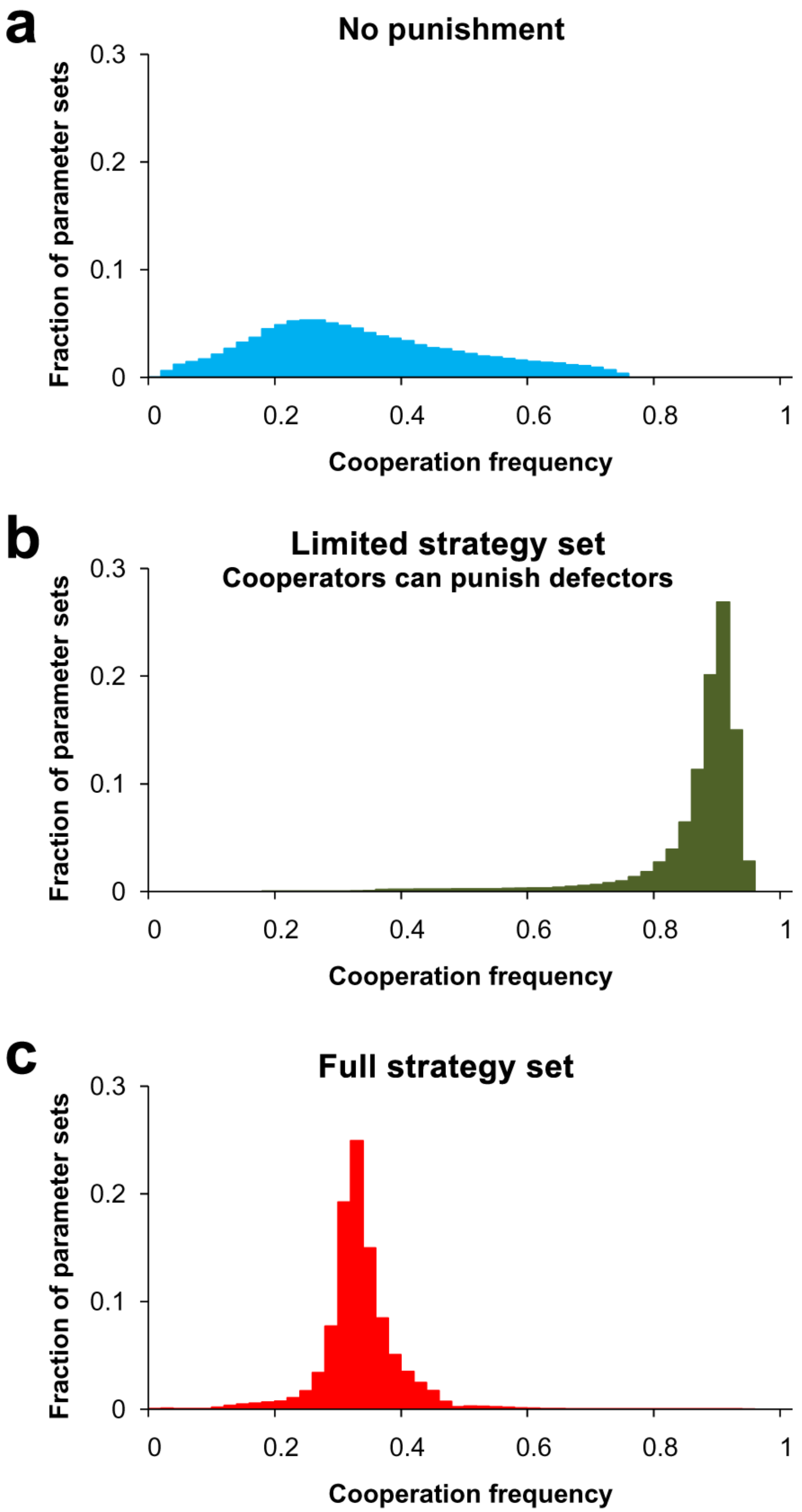
Figure 1 – Using the full punishment strategy set, anti-social punishment is common and punishment does not promote cooperation. Time-averaged frequencies of each strategy and transition rates between homogeneous populations, (**a**) without punishment, (**b**) when cooperators can punish defectors, and (**c,d**) with the full set of punishment strategies. A strategy X-$Y_1Y_2Y_3$ is defined by a public goods game action (X: C=cooperator, D=defector, L=loner), a punishment decision taken towards cooperators (Y_1: N = no action, P = punish), defectors (Y_2: N or P) and loners (Y_3: N or P). Transition rates ρ are the probability that a new mutant goes to fixation multiplied by the population size. We indicate neutral drift (ρ = 1, dotted lines), slow transitions (ρ = 11.8, thin lines), intermediate transitions (ρ = 25.7, medium lines) and fast transitions (ρ = 38.7, thick lines). Transitions with rates less than 0.1 are not shown. Parameter values are N = 100, n = 5, r = 3, c = 1, γ = 0.3, β = 1 and σ = 1. In panels **c** and **d**, strategies which punish others taking the same public goods action are included in the analysis, but not pictured because they are strongly disfavored by selection and virtually non-existent in the steady state distribution. For clarity, transitions with $\rho < 10$ are not shown in panel **d**.

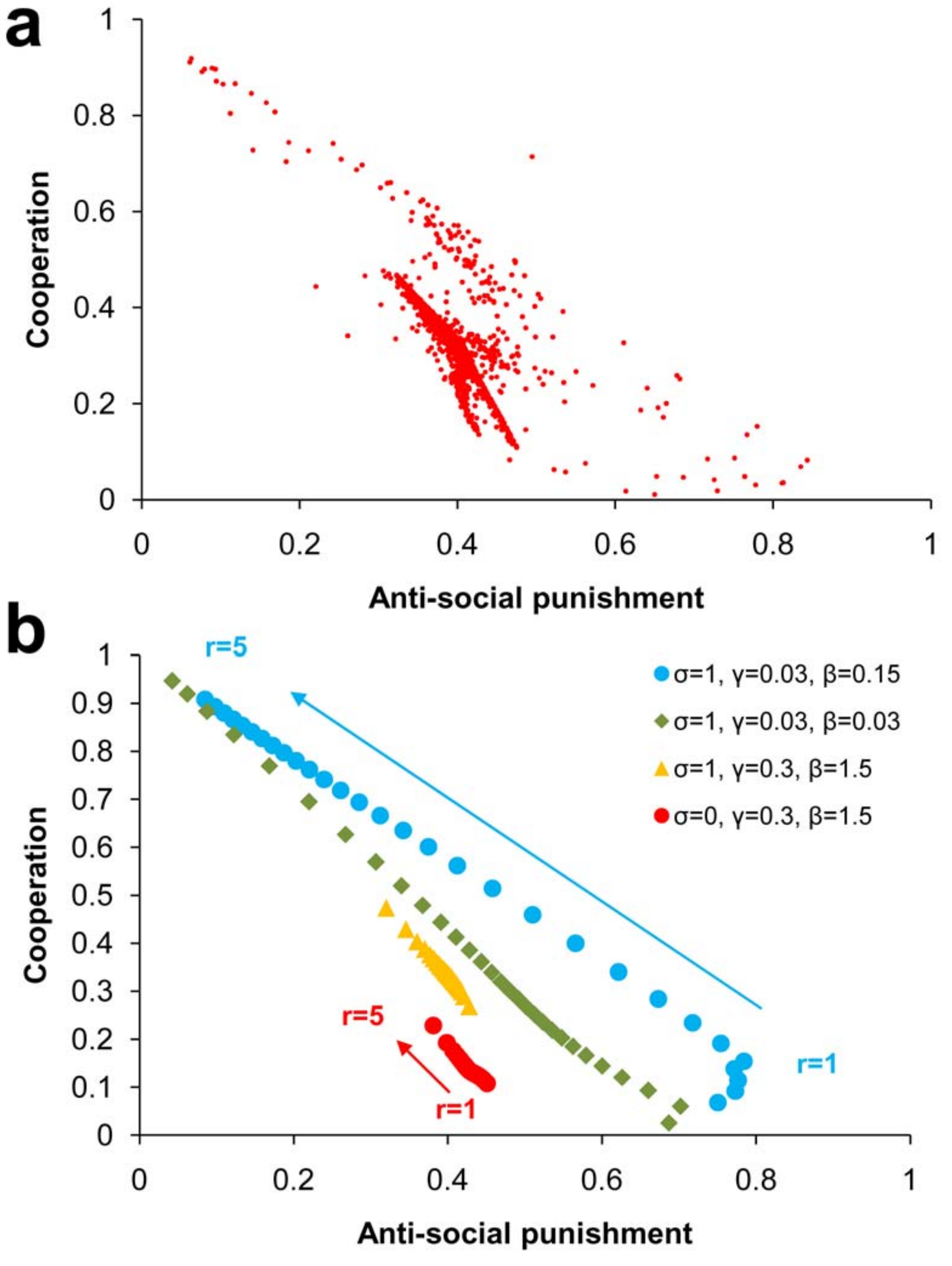
Figure 2 – Robustness of results across parameter sets. Shown are the results of 100,000 numerical calculations using N = 100, n = 5, c = 1 and randomly sampling from uniform distributions on the intervals $1 < r < 5$, $0 < \sigma < (r\text{-}1)c$, $0 < \gamma < 5$, and $\gamma < \beta < 5\gamma$. Results are shown for (**a**) the no-punishment strategy set, (**b**) the restricted punishment strategy set where only cooperators can punish loners, and (**c**) the full punishment strategy set.

Figure 3 – Inverse relationship between cooperation and anti-social punishment across parameters. (**a**) The steady state frequency of cooperation and anti-social punishment from 5,000 random parameter sets is shown. An inverse relationship is clearly visible: when anti-social punishment is rare, cooperation (and pro-social punishment) are common. (**b**) To explore this relationship, we vary r from (σ-c) to 5 for various values of σ, γ, and β, fixing N=100, n=5, and c=1. We see that increasing r always increases cooperation while decreasing anti-social punishment. We also see that when β is small, the range of cooperation and anti-social punishment values is large, whereas values are tightly constrained when β is larger. Achieving high levels of cooperation and low levels of anti-social punishment requires both large expected returns on public investment (large r) and symbolic punishments (small β).

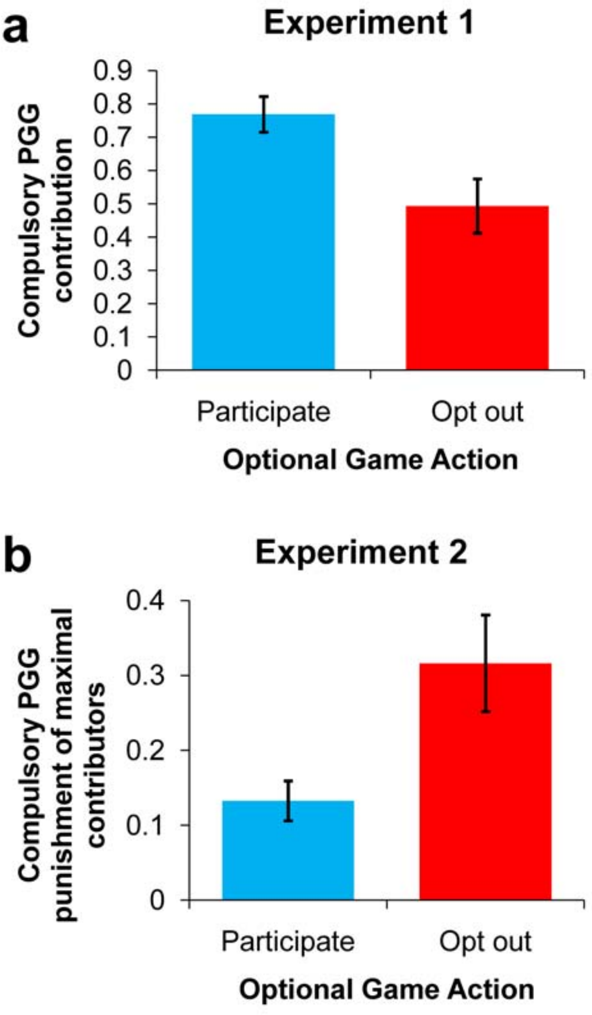
Figure 4 – Two behavioral experiments are consistent with model predictions. (**a**) In Experiment 1, subjects play an optional public goods game followed by a compulsory public goods game. The average fraction contributed in the compulsory game is significantly lower among subjects who opt out of the optional game (Rank-sum, N=73, p=0.006). Thus as predicted, loners contribute less in compulsory games. (**b**) In Experiment 2, subjects play an optional game followed by a compulsory game with costly punishment. Subjects indicate how much (0, 1 or 2) they would punish each possible contribution level in the compulsory game. Subjects who opt out of the optional game invest significantly more in punishing those that contribute the maximal amount (Rank-sum, N=196, p=0.003). Thus as predicted, loners engage in more anti-social punishment. All games are one-shot interactions among 4 players, with contributions to the public good multiplied by 2. Punishment technology is 3:1. Results are robust to various controls and alternate methods of analysis; see Supplementary notes for details.


a
No punishment
26.0%
C
NNN
ρ=25.7
D
NNN
17.3%
ρ=38.7
L
NNN
56.7%
ρ=11.8
b
Limited strategy set
Cooperators can punish defectors
83.5%
C
NPN
ρ=1
6.0%
C
NNN
ρ=25.7
D
NNN
4.0%
ρ=38.7
L
NNN
6.5%
ρ=11.8
ρ=11.8
c
Full strategy set
22.1%
C
NPN
ρ=1
1.6%
C
NNN
ρ=25.7
21.6%
D
NNP
ρ=11.8
ρ=11.8
ρ=25.7
ρ=1
L
NNN
3.0%
ρ=38.7
D
NNN
0.9%
ρ=1
ρ=38.7
L
PNN
21.1%
d
Full strategy set
8.1%
C
NPP
22.1%
C
NPN
0.6%
C
NNP
1.6%
C
NNN
ρ=11.8
ρ=25.7
1.7%
L
NPN
21.6%
D
NNP
11.6%
L
PPN
3.0%
L
NNN
0.9%
D
NNN
7.6%
D
PNP
21.2%
L
PNN
0.3%
D
PNN
ρ=38.7


a
No punishment
b
Limited strategy set
Cooperators can punish defectors
c
Full strategy set
Fraction of parameter sets
Cooperation frequency
0
0.1
0.2
0.3
0.4
0.6
0.8
1



a
Cooperation
Anti-social punishment
b
r=5
σ=1, γ=0.03, β=0.15
σ=1, γ=0.03, β=0.03
σ=1, γ=0.3, β=1.5
σ=0, γ=0.3, β=1.5
r=5
r=1
r=1
Cooperation
Anti-social punishment


a
Experiment 1
Compulsory PGG contribution
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Participate
Opt out
Optional Game Action
b
Experiment 2
Compulsory PGG punishment of maximal contributors
0.4
0.3
0.2
0.1
0
Participate
Opt out
Optional Game Action

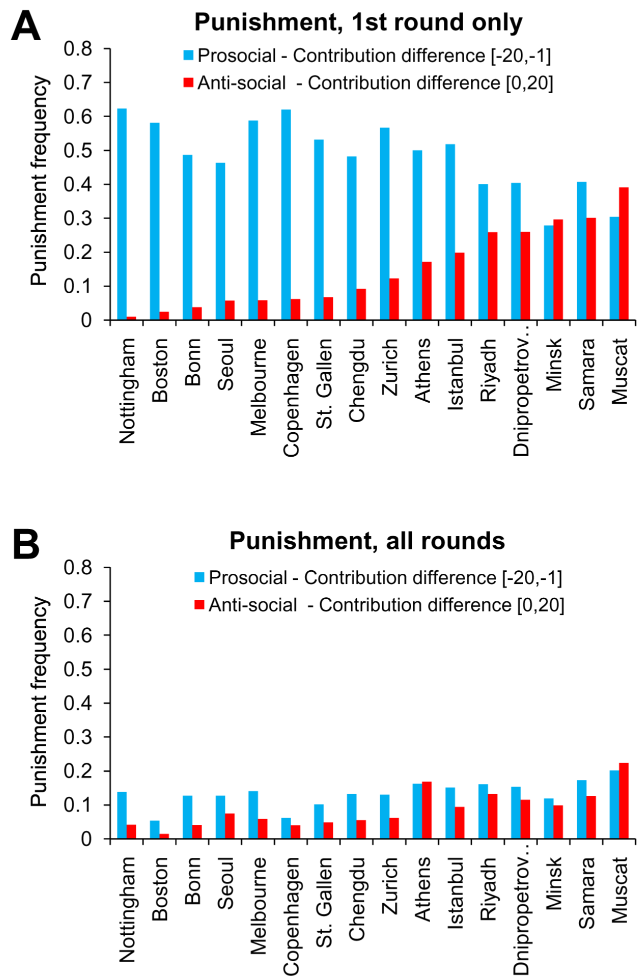
Supplementary information

The evolution of anti-social punishment in optional public goods games

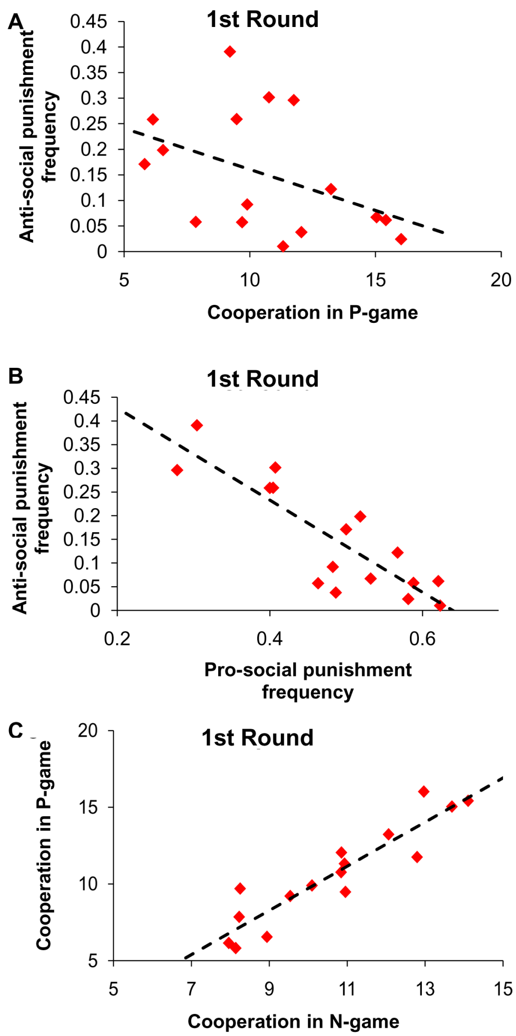
David G. Rand[1,2] & Martin A. Nowak[1,3,4]

[1]Program for Evolutionary Dynamics, Harvard University, Cambridge MA 02138, USA
[2]Department of Psychology, Harvard University, Cambridge, MA 02138, USA
[3]Department of Mathematics, Harvard University, Cambridge, MA 02138, USA
[4]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

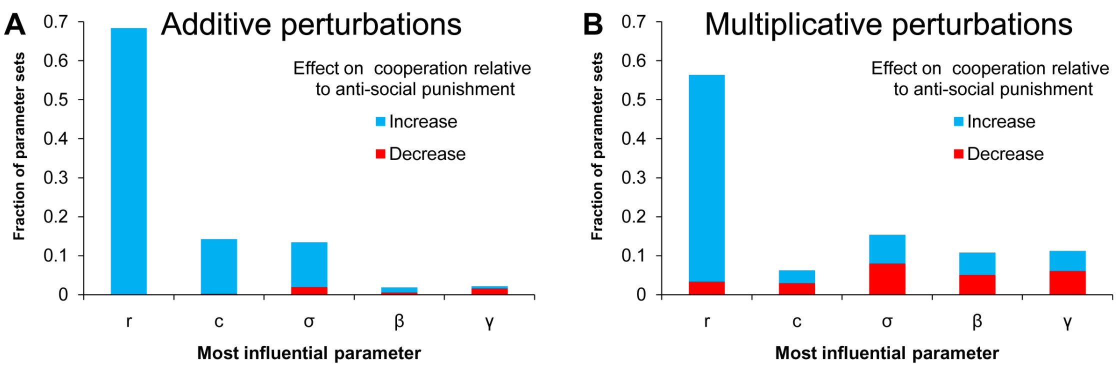
Supplementary Figures



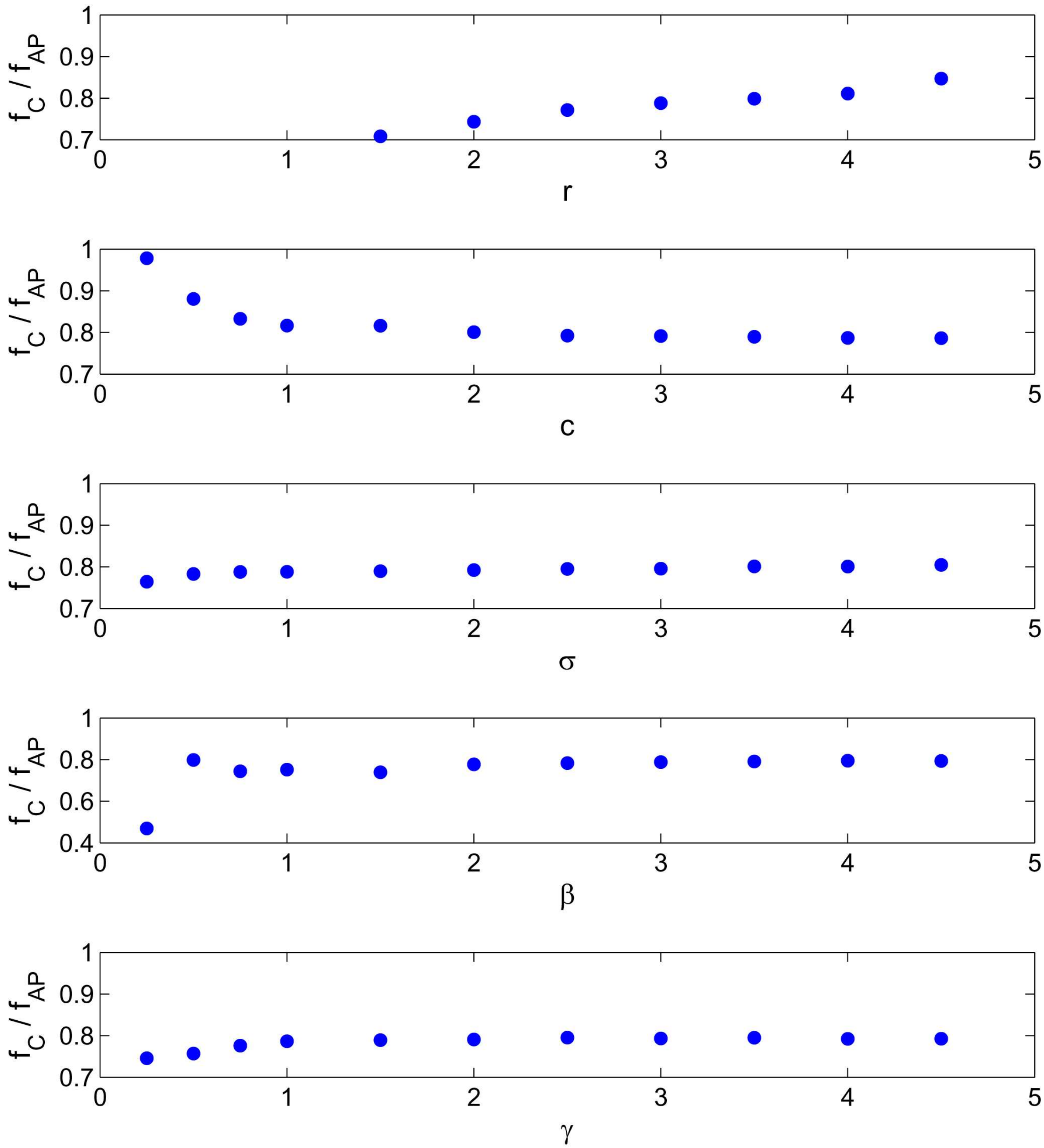
Supplementary Figure S1. Probability of punishing those who contributed less than you (pro-social, contribution difference between -20 and -1; blue) versus those who contributed as much or common than you (anti-social, contribution differnce between 0 and 20; red) in the first round (A) and overall (B). Data from Herrmann et al. 2008. Both panels are sorted by probability of punishing anti-socially in the first round.



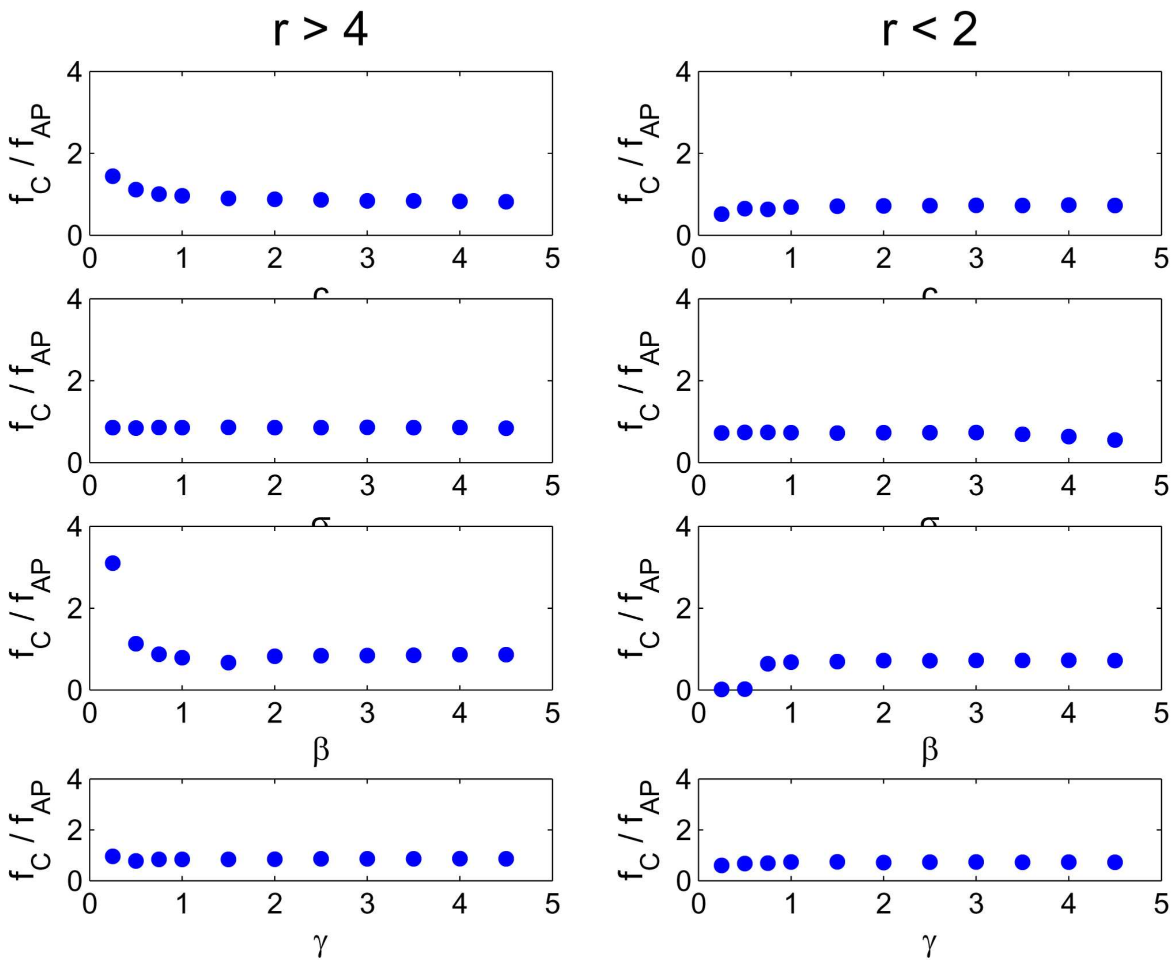
Supplementary Figure S2. (A) Probability of punishing those who contributed as much or common than you (anti-social) versus average contribution in the game with punishment. (B) Probability of punishing those who contributed less than you (pro-social) versus those who contributed as much or common than you (anti-social). (C) Average contribution in the game without punishment (N-game) versus with punishment (P-game). All data from Herrmann et al. 2008, considering only first round decisions.



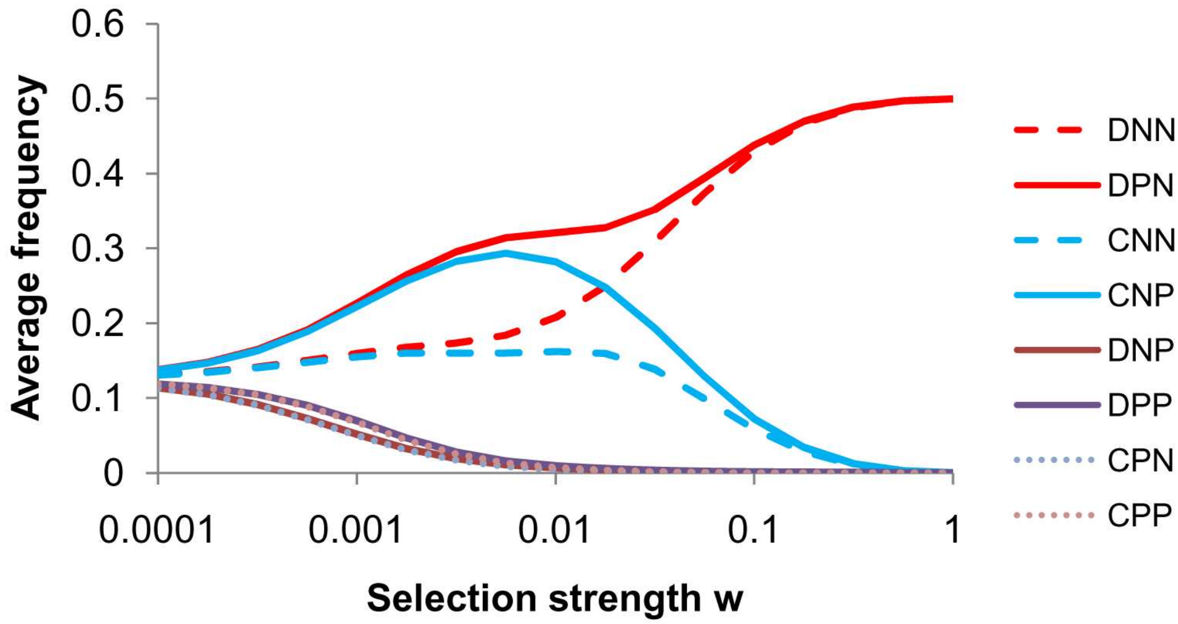
Supplementary Figure S3. Sensitivity analysis results. Across 10,000 random parameter sets sampled from the intervals $1 < r < 5$, $c = 1$, $0 < \sigma < (r-1)c$, $0 < \gamma < 5$, *and* $\gamma < \beta < 5\gamma$, *we calculate the marginal effect of increasing each parameter by 0.01 (A) or multiplying each parameter be 1.01 (B) on the ratio of cooperation to anti-social punishment. Though* c, σ, γ, *and* β *have some effect,* r *is the most influential parameter in a large majority of parameter sets.*



Supplementary Figure S4. Ratio of cooperation to anti-social punishment as a function of each parameter. For each data point the indicated parameter is fixed as shown, and the median of 500 parameter sets with the other payoff parameters randomly sampled is shown.



Supplementary Figure S5. Ratio of cooperation to anti-social punishment as a function of each parameter. For each data point, the indicated parameter is fixed as shown, 500 parameter sets with the other payoff parameters randomly sampled are generated, and the median of the subset obeying $r > 4$ (first column) or $r < 2$ (second column) is shown.



Supplementary Figure S6. Average frequency of each strategy in the low mutation limit, in a compulsory public goods game without loners. Strategy frequencies are averaged over 10,000 randomly sampled parameter sets.

Supplementary Tables

Country	Coeff	P-value
Nottingham	-5.09	< 0.001
Boston	-4.03	< 0.001
Bonn	-3.18	< 0.001
Seoul	-2.65	< 0.001
Melbourne	-3.14	< 0.001
Copenhagen	-3.22	< 0.001
St. Gallen	-2.76	< 0.001
Chengdu	-2.22	< 0.001
Zurich	-2.24	< 0.001
Athens	-1.58	0.004
Istanbul	-1.47	0.001
Riyadh	-0.65	0.080
Dnipropetrovsk	-0.66	0.064
Minsk	0.09	0.835
Samara	-0.46	0.082
Muscat	0.38	0.353

Supplementary Table S1
Logistic regression results assessing the difference in probability of punishing those who contribute less than you (pro-social punishment) versus those who contribute as much or more than you, using the data of Herrmann et al. 2008. Values of $p > 0.05$ *indicate no significant difference in probability to punish pro-socially versus anti-socially.*

Public goods action	Punish C?	Punish D?	Punish L?	No punishment	Limited strategy set	Full strategy set
C	N	N	N	33.7%	9.2%	2.6%
C	N	N	P			1.1%
C	N	P	N		77.6%	21.3%
C	N	P	P			8.6%
C	P	N	N			0.0%
C	P	N	P			0.0%
C	P	P	N			0.0%
C	P	P	P			0.0%
D	N	N	N	18.4%	5.6%	1.5%
D	N	N	P			20.0%
D	N	P	N			0.0%
D	N	P	P			0.0%
D	P	N	N			0.6%
D	P	N	P			7.5%
D	P	P	N			0.0%
D	P	P	P			0.0%
L	N	N	N	47.9%	7.6%	3.6%
L	N	N	P			0.0%
L	N	P	N			1.7%
L	N	P	P			0.0%
L	P	N	N			21.8%
L	P	N	P			0.0%
L	P	P	N			9.6%
L	P	P	P			0.0%

Supplementary Table S2
Steady state frequency of each strategy in the low mutation limit, averaged over 100,000 random parameter sets.

	No punishment	Limited strategy set	Full strategy set
Cooperators	33.7%	86.8%	33.6%
Defectors	18.4%	5.6%	29.6%
Loners	47.9%	7.6%	36.8%
Punishment of Cooperators			39.6%
Punishment of Defectors		77.6%	41.3%
Punishment of Loners			37.2%

Supplementary Table S3
Frequencies of cooperation, defection and opting out, as well as each type of punishment, in the low mutation limit averaged over 100,000 randomly sampled parameter sets.

.

Public goods action	Punish C?	Punish D?	Punish L?	$u \to 0$	$u = 0.001$	$u = 0.01$
C	N	N	N	2.6%	2.9%	10.7%
C	N	N	P	1.1%	1.1%	1.9%
C	N	P	N	21.3%	23.0%	19.4%
C	N	P	P	8.6%	6.3%	1.6%
C	P	N	N	0.0%	0.0%	0.1%
C	P	N	P	0.0%	0.0%	0.1%
C	P	P	N	0.0%	0.0%	0.1%
C	P	P	P	0.0%	0.0%	0.1%
D	N	N	N	1.5%	1.8%	6.8%
D	N	N	P	20.0%	21.9%	17.9%
D	N	P	N	0.0%	0.0%	0.1%
D	N	P	P	0.0%	0.0%	0.1%
D	P	N	N	0.6%	0.7%	1.3%
D	P	N	P	7.5%	5.2%	1.1%
D	P	P	N	0.0%	0.0%	0.1%
D	P	P	P	0.0%	0.0%	0.1%
L	N	N	N	3.6%	4.3%	14.1%
L	N	N	P	0.0%	0.0%	0.1%
L	N	P	N	1.7%	1.7%	2.1%
L	N	P	P	0.0%	0.0%	0.1%
L	P	N	N	21.8%	24.1%	20.7%
L	P	N	P	0.0%	0.0%	0.1%
L	P	P	N	9.6%	6.8%	1.8%
L	P	P	P	0.0%	0.0%	0.1%

Supplementary Table S4
Average steady state frequencies of each strategy in the full punishment set, from low mutation limit calculation (100,000 random parameter sets) and agent based simulations using $u = 0.001$ and $u = 0.01$ (1000 random parameter sets each).

	$u \to 0$	$u = 0.001$	$u = 0.01$
Cooperators	33.6%	33.4%	33.8%
Defectors	29.6%	29.6%	27.4%
Loners	36.8%	37.0%	38.9%
Punishment of Cooperators	39.6%	37.0%	25.3%
Punishment of Defectors	41.3%	38.0%	25.3%
Punishment of Loners	37.2%	34.6%	22.9%

Supplementary Table S5
Average frequencies of cooperation, defection and opting out, as well as average frequencies of each type of punishment. Determined using low mutation limit calculation (100,000 random parameter sets) and agent based simulations with $u = 0.001$ *and* $u = 0.01$ *(1000 random parameter sets each).*

.

	(1) OLS	(2) OLS	(3) OLS	(4) Tobit	(5) Tobit	(6) Tobit
Opt out	-11.02***	-10.88***	-8.571**	-34.76***	-34.70***	-25.57*
	(3.884)	(3.849)	(4.230)	(12.77)	(12.44)	(12.97)
Age		0.183	0.203*		0.838*	0.926*
		(0.124)	(0.116)		(0.498)	(0.492)
Female		3.204	4.685		11.28	19.58
		(3.769)	(3.726)		(13.33)	(13.72)
Risk-taking			1.228*			5.182*
			(0.726)			(2.718)
U. S. Resident		4.379	5.669		13.80	19.07
		(3.929)	(4.159)		(13.39)	(13.90)
Constant	30.76***	21.20***	11.73	57.55***	18.54	-21.64
	(2.149)	(5.677)	(7.484)	(10.02)	(19.06)	(26.88)
Observations	73	73	73	73	73	73
R-squared	0.103	0.157	0.189			

Supplementary Table S6
*Compulsory game contribution as a function of choosing to opt out of the optional game, for the treatment condition. Robust standard errors shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

	(1) OLS	(2) OLS	(3) OLS	(4) Tobit	(5) Tobit	(6) Tobit
Treatment	1.680	1.597	2.208	5.109	5.439	8.522
	(3.044)	(2.995)	(2.799)	(11.62)	(11.25)	(10.45)
Age		0.119	0.128		0.534	0.592
		(0.131)	(0.114)		(0.531)	(0.463)
Female		0.662	2.475		0.737	9.429
		(3.072)	(2.930)		(11.53)	(11.03)
Risk-taking			2.285***			8.947***
			(0.518)			(2.394)
U. S. Resident		6.278*	8.735***		23.61*	31.88**
		(3.196)	(3.187)		(12.22)	(12.30)
Constant	25.61***	18.43***	2.891	43.37***	13.81	-48.25*
	(2.394)	(5.066)	(5.681)	(9.697)	(19.71)	(24.43)
Observations	124	124	124	124	124	124
R-squared	0.003	0.055	0.166			

Supplementary Table S7
*Compulsory game contribution in treatment versus control conditions. Robust standard errors shown in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.*

	Treatment	Control
Observations	73	51
Age	33	33
Female	41%	43%
U.S. Resident	48%	47%
Risk-taking	5.7	5.8
Fraction opting out	32%	N/A
Optional PGG contribution	30.2	N/A
Compulsory PGG contribution	27.3	25.6

Supplementary Table S8
Summary statistics for Experiment 1.

	Experiment 2
Observations	196
Age	33
Female	48%
U.S. Resident	57%
Risk-taking	5.2
Fraction opting out	25%
Optional PGG contribution	30.0
Compulsory PGG contribution	19.9
Punishment of:	
0 contributors	0.62
10 contributors	0.65
20 contributors	0.51
30 contributors	0.43
40 contributors	0.36

Supplementary Table S9
Summary statistics for Experiment 2. Punishment amounts indicate average amount spent on punishing (A=0, B=1, C=2).

Punishment of those that contribute:	(1) 0	(2) 0	(3) 0	(4) 40	(5) 40	(6) 40
Compulsory contribution	0.0181***	0.0187***	0.0179***	-0.0135**	-0.0137**	-0.0134**
	(0.00521)	(0.00534)	(0.00540)	(0.00611)	(0.00613)	(0.00617)
Age		-0.0123	-0.0109		-0.000906	-0.00132
		(0.00977)	(0.00987)		(0.0110)	(0.0110)
Female		-0.0793	-0.0411		-0.0208	-0.0404
		(0.182)	(0.185)		(0.222)	(0.225)
U. S. Resident		-0.240	-0.177		-0.0781	-0.104
		(0.186)	(0.187)		(0.228)	(0.230)
Risk-taking			0.0748**			-0.0296
			(0.0379)			(0.0409)
Constant cut 1	0.684***	0.141	0.618	0.581***	0.486	0.306
	(0.138)	(0.328)	(0.404)	(0.148)	(0.382)	(0.455)
Constant cut 2	1.081***	0.533*	1.016**	0.782***	0.688*	0.508
	(0.141)	(0.319)	(0.398)	(0.149)	(0.381)	(0.449)
Observations	196	194	194	196	194	194

Supplementary Table S10
*Punishment of those that contribute 0 (regressions 1-3) and those that contribute 40 (regressions 4-6) versus compulsory game contribution. Ordinal probit regression. Robust standard errors shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

	(1)	(2)	(3)	(4)
Opt out	0.633***	0.661***	0.647***	0.579**
	(0.220)	(0.223)	(0.235)	(0.246)
Age		-0.00545	-0.00552	-0.00530
		(0.0106)	(0.0107)	(0.0109)
Female		0.0963	0.0859	0.0739
		(0.224)	(0.229)	(0.232)
U. S. Resident		-0.0915	-0.103	-0.0998
		(0.227)	(0.226)	(0.230)
Risk-taking			-0.0122	-0.00849
			(0.0434)	(0.0430)
Compulsory contribution				-0.0111*
				(0.00640)
Constant cut 1	1.014***	0.835**	0.755	0.558
	(0.125)	(0.343)	(0.462)	(0.480)
Constant cut 2	1.219***	1.041***	0.961**	0.767
	(0.131)	(0.344)	(0.459)	(0.475)
Observations	196	194	194	194

Supplementary Table S11
*Punishment of those that contribute the maximal amount (40) versus opting out of the optional game. Ordinal probit regression. Robust standard errors shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

	(1) OLS	(2) OLS	(3) OLS	(4) Tobit	(5) Tobit	(6) Tobit
Opt out	0.286**	0.282**	0.239*	0.971***	0.943***	0.762**
	(0.115)	(0.121)	(0.130)	(0.335)	(0.353)	(0.380)
Age		-0.00266	-0.00246		-0.00884	-0.00787
		(0.00467)	(0.00467)		(0.0159)	(0.0161)
Female		-0.0201	-0.0227		0.0276	0.00127
		(0.101)	(0.102)		(0.357)	(0.360)
U. S. Resident		-0.0806	-0.0789		-0.434	-0.437
		(0.102)	(0.102)		(0.353)	(0.357)
Risk-taking		-0.00445	-0.00233		-0.0388	-0.0256
		(0.0193)	(0.0190)		(0.0647)	(0.0634)
Compulsory contribution			-0.00536*			-0.0286***
			(0.00281)			(0.00922)
Constant	0.303***	0.473**	0.573***	-0.985***	-0.244	0.238
	(0.0480)	(0.204)	(0.212)	(0.250)	(0.662)	(0.670)
Observations	196	194	194	196	194	194
R-squared	0.039	0.045	0.065			

Supplementary Table S12
*Average punishment of those that contribute as much or more the punisher versus opting out of the optional game. Robust standard errors shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Punishment of those contributing:	(1) 0	(2) 10	(3) 20	(4) 30	(5) 40
Opt out	0.107	0.304	0.310	0.591***	0.647***
	(0.223)	(0.190)	(0.197)	(0.204)	(0.235)
Age	-0.0111	-0.0120	-0.00280	-0.00784	-0.00552
	(0.00992)	(0.00856)	(0.00787)	(0.00817)	(0.0107)
Female	-0.0238	-0.342*	-0.241	-0.250	0.0859
	(0.189)	(0.180)	(0.194)	(0.208)	(0.229)
U. S. Resident	-0.147	-0.126	0.0444	-0.0726	-0.103
	(0.188)	(0.178)	(0.193)	(0.202)	(0.226)
Risk-taking	0.0884**	0.0640	0.0290	-0.000489	-0.0122
	(0.0379)	(0.0392)	(0.0385)	(0.0402)	(0.0434)
Constant cut 1	0.362	-0.180	0.395	0.267	0.755
	(0.402)	(0.392)	(0.369)	(0.392)	(0.462)
Constant cut 2	0.743*	0.772**	1.106***	0.921**	0.961**
	(0.398)	(0.393)	(0.379)	(0.413)	(0.459)
Observations	194	194	194	194	194

Supplementary Table S13
*Punishment of those that contribute each possible contribution level. Ordinal probit regression. Robust standard errors shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

Public goods action	Punish C?	Punish D?	Punish L?	Method 1	Method 2	Method 3
C	N	N	N	2.6%	2.6%	2.0%
C	N	N	P	1.1%	1.1%	0.8%
C	N	P	N	21.3%	21.3%	20.0%
C	N	P	P	8.6%	8.3%	7.7%
C	P	N	N	0.0%	0.0%	0.0%
C	P	N	P	0.0%	0.0%	0.0%
C	P	P	N	0.0%	0.0%	0.0%
C	P	P	P	0.0%	0.0%	0.0%
D	N	N	N	1.5%	1.5%	2.1%
D	N	N	P	20.0%	20.5%	19.9%
D	N	P	N	0.0%	0.0%	0.0%
D	N	P	P	0.0%	0.0%	0.0%
D	P	N	N	0.6%	0.6%	1.1%
D	P	N	P	7.5%	7.6%	7.8%
D	P	P	N	0.0%	0.0%	0.0%
D	P	P	P	0.0%	0.0%	0.0%
L	N	N	N	3.6%	3.7%	4.5%
L	N	N	P	0.0%	0.0%	0.0%
L	N	P	N	1.7%	1.6%	2.2%
L	N	P	P	0.0%	0.0%	0.0%
L	P	N	N	21.8%	22.1%	21.9%
L	P	N	P	0.0%	0.0%	0.0%
L	P	P	N	9.6%	9.2%	10.0%
L	P	P	P	0.0%	0.0%	0.0%

Supplementary Table S14
Average steady state frequencies of each strategy in the low mutation limit, using three different parameter sampling methods (100,000 random parameter sets each).

Supplementary Methods

Payoff calculation

To calculate the payoff of each strategy in a given state of the population X, we extend the formulation of Traulsen et al. (2009) to include all 24 possible strategies[19]. Note that we are not increasing the complexity of available strategies, or making structural changes to the game such as adding a retaliation stage[22]. Instead we simply consider the complete strategy set. There are 24 possible strategies, specified by a 4 digit strategy name [a_1-$a_2a_3a_4$] where $a_1 \in [C, D, L]$ indicates a decision in the public goods game, $a_2 \in [P, N]$ indicates whether to punish cooperators, $a_3 \in [P, N]$ indicates whether to punish defectors and $a_4 \in [P, N]$ indicates whether to punish loners.

First we consider the payoff of a given strategy, s, in a single public goods game with n other players. Although there are 24 total strategies, it is useful to aggregate them along each strategy component: in the group of n players, let i_C be the number of players that cooperate (8 different strategies, C-***), i_D be the number of players who defect (8 different strategies, D-***), i_L be the number of players that opt out (8 different strategies, L-***), i_{pC} be the number of players that punish cooperators (12 strategies, *-P**), i_{pD} be the number of players that punish defectors (12 strategies, *-*P*) and i_{pL} be the number of players that punish loners (12 strategies, *-**P).

We also define 6 variables which describe the strategy of player s. Let $s_C = 1$ if strategy s is a cooperator and $s_C = 0$ otherwise; let $s_D = 1$ if strategy s is a defector and $s_D = 0$ otherwise; let $s_L = 1$ if strategy s is a loner and $s_L = 0$ otherwise; let $s_pC = 1$ if strategy s punishes cooperators and $s_pC = 0$ otherwise; let $s_pD = 1$ if strategy s punishes defectors and $s_pD = 0$ otherwise; and let $s_pL = 1$ if strategy s punishes loners and $s_pL = 0$ otherwise.

The payoff of strategy s in this group of n is the sum of (i) the payoff from the public goods game δ_1, which depends on i_C and i_D if $s_L = 0$, (ii) the payoff from being punished δ_2, which depends on i_{pC} if $s_C = 1$, i_{pD} if $s_D = 1$, or i_{pL} if $s_L = 1$, and (iii) the payoff from choosing to punish δ_3, which depends on i_C if $s_{pC} = 1$, i_D if $s_{pD} = 1$, and i_L if $s_{pL} = 1$. Note that we assume players never punish themselves.

The payoff of strategy s in the public goods game δ_1 is determined as follows. If $i_C + i_D < 1$ then there are not enough non-loners for the public goods game to occur, and so $\delta_1 = \sigma$ regardless of the strategy of player s. Otherwise,

$$\delta_1 = rc\frac{i_C}{i_C + i_D} - cs_C \tag{S1}$$

The payoff of strategy s from being punished δ_2 is given by

$$\delta_2 = -\beta\Bigg(s_C(i_{pC} - s_{pC}) + s_D(i_{pD} - s_{pD}) + s_L(i_{pL} - s_{pL})\Bigg) \tag{S2}$$

The payoff of strategy s from paying to punish others δ_3 is given by

$$\delta_3 = -\gamma\Bigg(s_{pC}(i_C - s_C) + s_{pD}(i_D - s_D) + s_{pL}(i_L - s_L)\Bigg) \tag{S3}$$

The total payoff of strategy s is then given by

$$P_s = \delta_1 + \delta_2 + \delta_3 \tag{S4}$$

Thus far we have calculated the payoff of a particular strategy playing the public goods game with punishment in a group with a particular set of $n-1$ other players. To calculate the expected (average) payoff of a strategy s in a population of size N, we must now consider the average group composition. Let X_i be the total number of players in the population of size N using strategy i, $i \in [1, 24]$. A randomly sampled group of size n has a specific composition given by the multivariate hypergeometric distribution,

$$H(\mathbf{I}, \mathbf{X}) = \frac{\binom{X_1}{i_1}\binom{X_2}{i_2}\cdots\binom{X_24}{i_{24}}}{\binom{N}{n}} \tag{S5}$$

where $\mathbf{I} = (i_1, i_2, \ldots, i_{24})$ and $\mathbf{X} = (X_1, X_2, \ldots, X_{24})$. The average payoff π_s for strategy s is then $\Sigma_{i_1}\Sigma_{i_2}\ldots\Sigma_{i_{24}}\mathbf{H}(\ldots)P_s$.

Substituting for each of the 24 strategies and simplifying gives

$$\pi_{D-NNN} = B(\mathbf{X}) - X_{pD}\beta D(N,n) \tag{S6}$$
$$\pi_{D-NNP} = B(\mathbf{X}) - (X_{pD}\beta + X_L\gamma)D(N,n) \tag{S7}$$
$$\pi_{D-NPN} = B(\mathbf{X}) - ((X_D - 1)\gamma + (X_{pD} - 1)\beta)D(N,n) \tag{S8}$$
$$\pi_{D-NPP} = B(\mathbf{X}) - ((X_D - 1 + X_L)\gamma + (X_{pD} - 1)\beta)D(N,n) \tag{S9}$$
$$\pi_{D-PNN} = B(\mathbf{X}) - (X_{pD}\beta + X_C\gamma)D(N,n) \tag{S10}$$
$$\pi_{D-PNP} = B(\mathbf{X}) - ((X_C + X_L)\gamma + X_{pD}\beta)D(N,n) \tag{S11}$$
$$\pi_{D-PPN} = B(\mathbf{X}) - ((X_C + X_D - 1)\gamma + (X_{pD} - 1)\beta)D(N,n) \tag{S12}$$
$$\pi_{D-PPP} = B(\mathbf{X}) - ((X_C + X_D - 1 + X_L)\gamma + (X_{pD} - 1)\beta)D(N,n) \tag{S13}$$
$$\pi_{C-NNN} = B(\mathbf{X}) - F(X_L)c - X_{pC}\beta D(N,n) \tag{S14}$$
$$\pi_{C-NNP} = B(\mathbf{X}) - F(X_L)c - (X_{pC}\beta + X_L\gamma)D(N,n) \tag{S15}$$
$$\pi_{C-NPN} = B(\mathbf{X}) - F(X_L)c - (X_{pC}\beta + X_D\gamma)D(N,n) \tag{S16}$$
$$\pi_{C-NPP} = B(\mathbf{X}) - F(X_L)c - ((X_D + X_L)\gamma + X_{pC}\beta)D(N,n) \tag{S17}$$
$$\pi_{C-PNN} = B(\mathbf{X}) - F(X_L)c - ((X_C - 1)\gamma + (X_{pC} - 1)\beta)D(N,n) \tag{S18}$$
$$\pi_{C-PNP} = B(\mathbf{X}) - F(X_L)c - ((X_C - 1 + X_L)\gamma + (X_{pC} - 1)\beta)D(N,n) \tag{S19}$$
$$\pi_{C-PPN} = B(\mathbf{X}) - F(X_L)c - ((X_C - 1 + X_D)\gamma + (X_{pC} - 1)\beta)D(N,n) \tag{S20}$$
$$\pi_{C-PPP} = B(\mathbf{X}) - F(X_L)c - ((X_C - 1 + X_D + X_L)\gamma + (X_{pC} - 1)\beta)D(N,n) \tag{S21}$$
$$\pi_{L-NNN} = \sigma - X_{pL}\beta D(N,n) \tag{S22}$$
$$\pi_{L-NNP} = \sigma - ((X_L - 1)\gamma + (X_{pL} - 1)\beta)D(N,n) \tag{S23}$$
$$\pi_{L-NPN} = \sigma - (X_{pL}\beta + X_D\gamma)D(N,n) \tag{S24}$$
$$\pi_{L-NPP} = \sigma - ((X_D + X_L - 1)\gamma + (X_{pL} - 1)\beta)D(N,n) \tag{S25}$$
$$\pi_{L-PNN} = \sigma - (X_{pL}\beta + X_C\gamma)D(N,n) \tag{S26}$$
$$\pi_{L-PNP} = \sigma - ((X_C + X_L - 1)\gamma + (X_{pL} - 1)\beta)D(N,n) \tag{S27}$$
$$\pi_{L-PPN} = \sigma - ((X_C + X_D)\gamma + X_{pL}\beta)D(N,n) \tag{S28}$$
$$\pi_{L-PPP} = \sigma - ((X_C + X_D + X_L - 1)\gamma + (X_{pL} - 1)\beta)D(N,n) \tag{S29}$$

where X_C is the total number of players in the population that cooperate, X_D is the total number of players that defect, X_L is the total number of players that are loners, X_{pC} is the total number of players that punish cooperators, X_{pD} is the total number of players that punish defectors, X_{pL} is the total number of players that punish loners, $B(\mathbf{X})$ is the payoff from the public goods game, $F(X_L)$ is the effective cost of contributing to the public good, and $D(N,n) = \frac{n-1}{N-1}$.

If $X_L \geq (N-1)$, then there is never more than 1 loner in a group, the public goods game is never played, and $B(\mathbf{X}) = \sigma$ and $F(X_L) = 0$.

Otherwise,

$$B(\mathbf{X}) = \frac{rcX_C}{N - X_L - 1}\left(1 - \frac{N}{n(N - X_L)}\right) + \frac{\binom{X_L}{n-1}}{\binom{N-1}{n-1}}\left(\sigma + \frac{rcX_C X_L - n + 1}{n(N - X_L - 1)(N - X_L)}\right) \quad \text{(S30)}$$

and

$$F(X_L) = 1 - \frac{r}{n}\frac{N - n}{N - X_L - 1} + \frac{\binom{X_L}{n-1}}{\binom{N-1}{n-1}}\left(\frac{r}{n}\frac{X_L + 1}{N - X_L - 1} + r\frac{N - X_L - 2}{N - X_L - 1} - 1\right) \quad \text{(S31)}$$

Description of the evolutionary dynamic

We study the transmission of strategies through an evolutionary process, which can be interpreted either as genetic evolution or as social learning. In both cases, strategies which earn higher payoffs are more likely to spread in the population, while lower payoff strategies tend to die out. Novel strategies are introduced by mutation in the case of genetic evolution, or innovation and experimentation in the case of social learning.

We use a frequency dependent Moran process[39] with an exponential payoff function[40]. In each round, agents interact at random. One agent is then randomly selected to change strategy. With probability u, a mutation occurs and the agent picks a new strategy at random. With probability $1 - u$, the agent adopts the strategy of another agent j, who is selected from the population with probability proportional to e^{π_j} where π_j is the payoff of agent j.

We begin by studying low mutation rates, which characterize both genetic evolution and the long-term evolution of societal norms. Later, we explore higher mutation rates, which may be more appropriate for short-term learning and exploration dynamics[19,43]. For low mutation a mutant either goes to fixation or dies out before another mutant appears[41,42]. Thus the population makes transitions between homogeneous states, where all agents use the same strategy. Here the success of a given strategy depends on its ability to invade other strategies, and to resist invasion by other strategies. We use an exact numerical calculation to determine the average frequency of each strategy in the stationary distribution (see below).

Low mutation limit calculation method

If viable mutants are very rare, the population spends almost all of its time in a homogeneous state. When a mutant arises, it either goes to fixation or dies out before the next mutant arises, returning the system to a homogeneous population. Let s_i be the frequency of strategy i, with a total of M strategies. We can then assemble a transition matrix between homogeneous states of the system. The transition probability from state i to state j is the product of the probability of a mutant of type j arising ($\frac{1}{M-1}$) and the fixation probability of a single mutant j in a population of i players, $\rho_{i,j}$. The probability of staying in state i is thus $1 - \frac{1}{M-1}\Sigma_j \rho_{j,i}$, where $\rho_{i,i} = 0$. This transition matrix can then be used to calculate the steady state frequency distribution s^* of strategies:

$$
\begin{pmatrix} s_1^* \\ s_2^* \\ \vdots \\ s_M^* \end{pmatrix} = \begin{pmatrix} 1-\Sigma_j \frac{\rho_{j,1}}{M-1} & \frac{\rho_{1,2}}{M-1} & \cdots & \frac{\rho_{1,M}}{M-1} \\ \frac{\rho_{2,1}}{M-1} & 1-\Sigma_j \frac{\rho_{j,2}}{M-1} & \cdots & \frac{\rho_{2,M}}{M-1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\rho_{M,1}}{M-1} & \frac{\rho_{M,2}}{M-1} & \cdots & 1-\Sigma_j \frac{\rho_{j,M}}{M-1} \end{pmatrix} \begin{pmatrix} s_1^* \\ s_2^* \\ \vdots \\ s_M^* \end{pmatrix} \tag{S32}
$$

The eigenvector corresponding to the largest eigenvalue (1) will give the steady state distribution of the stochastic process.

Using the Moran process, the fixation probability $\rho_{B,A}$ (the probability that a single A mutant introduced into a population of B-players will take over) can be calculated as follows. When the full set of punishment strategies are available, it is common for strategies to receive negative payoffs. Therefore, we use an exponential fitness function as opposed to a linear fitness function[17,19]. This allows us to maintain a strong intensity of selection using the full punishment set. In a population of i A-players and $N - i$ B-players, the average payoffs π_A and π_B can be calculated using Eq. (S38) - (S29). The fitness of an A-player f_i and B-player g_i are then defined as

$$f_i = e^{\pi_A} \tag{S33}$$
$$g_i = e^{\pi_B} \tag{S34}$$

.

Note that the results in the main text Figure 1A and 1B do not exactly match those of Hauert et al. 2007 because they use the linear fitness function $f_i = 1 - w + w\pi_A$ rather than the exponential payoff function described here.

The fixation probability of a single A-player in a population of B-players can then be calculated[17,19,40] as follows:

$$\rho_{B,A} = 1/\left(1 + \Sigma_{k=1}^{N-1}\Pi_{i=1}^{k}\frac{g_i}{f_i}\right) \tag{S35}$$

The calculations presented in the main text numerically evaluate Eq. (S35) for each strategy pair, and then numerically solve Eq. (S32) to determine the steady state frequency of each strategy.

Supplementary Notes

Review of experimental evidence for anti-social punishment

The experimental literature shows that punishment is sometimes targeted at cooperators across a range of different experimental designs. On one extreme are repeated games that allow for retaliation (punishment in response to being punished) because identities are kept constant over repeated interactions and subjects know who has punished them[26,28,29,30]. On the other extreme are one-shot games where explicit retaliation is impossible, yet nonetheless considerable anti-social punishment is observed in some cultures[30,31]. The landmark study of Herrmann et al 2008 is somewhere in between these two extremes[27]. The game is repeated, but identities are shuffled between rounds and subjects are only informed of the total amount of punishment they received, not which particular players have punished them. Thus explicit retaliation is impossible, although it is still possible for low contributors to assume that punishment came from high contributors, and to respond accordingly. However, the presence of anti-social punishment even in one-shot games shows that anti-social punishment is not entirely motivated by explicit retaliation.

To provide further evidence for anti-social punishment that is not linked to retaliation, we have reanalyzed the data from Herrmann et al 2008, focusing on punishment occurring in the very first punishment round. Here subjects have played one round of public goods game, and are punishing based on what has occurred in that first cooperation round. Retaliation is impossible, as no previous punishment has occurred. Supplementary Figure S1A shows the frequency of punishment given that the other person contributed less than you (pro-social punishment, shown in blue) and given that the other person contributed as much or more than you (anti-social punishment, shown in red). Supplementary Figure S1B shows the same analysis for all 10 rounds of play. In both panels, we see a substantial level of anti-social punishment in many societies.

We now ask for each subject pool whether there is a significant difference in the frequency of pro-social versus anti-social punishment in the first round. We do so using logistic regression with robust standard errors, clustered on subject and group to account for the non-independence of observations from a given subject and given group. The regression takes punishment decision in the first round (0=no deduction points alloted, 1=1 or more deduction points alloted) as the dependent variable, and other's relative contribution in the first round (0=less than you, 1=as much or more than you) as the independent variable. Regression results are shown in Supplementary Table S1. We see that in 5 out of 16 countries, anti-social punishment is as common as pro-social punishment in the first round ($p > 0.05$).

In addition to demonstrating the existence of anti-social punishment, Herrmann et al. 2008 show that anti-social punishment varies systematically across subject pools. Particularly, in Supplementary Figure 2A of their paper they show a significant negative correlation across subject pools between contribution level average over all periods and the mean level of anti-social punishment across all periods. We find the same effect when considering only contributions in the first period, and frequency of anti-social punishment in the first period (linear regression with robust standard errors, considering country-level averages; coeff = -11.65, p = 0.048). Additional analysis of the Hermmann et al. 2008 data also shows a significant negative correlation across subject pools between the frequency of pro-social and anti-social punishment in the first period (coeff = -0.97, $p < 0.001$); and a significant positive correlation between the level of first-round contribution in the no-punishment and punishment games (coeff = 1.43, $p < 0.001$). All three of these relationships are visualized in Supplementary Fig. S2.

We also note that the negative relationship between contribution level and anti-social punishment holds at the level of the individual as well as the level of the subject pool. Analyzing data from all subjects, a logistic regression with robust standard errors clustered on subject and group finds a significant negative correlation between contribution and probability to punish if the other contributed as much or more than you (1st round only: coeff = -0.097, $p < 0.001$; all rounds: coeff = -.0107, $p < 0.001$). Similarly, a linear regression with robust standard errors clustered on subject and group finds a significant negative relationship between contribution and amount spent on punishing if the other contributed as much or more than you (1st round only: coeff = -0.031, $p < 0.001$; all rounds: coeff = -.032, $p < 0.001$). These results are qualitatively unchanged by the addition of subject pool dummy variables.

It is important to keep in mind that the particular subject pools studied by Herrmann et al. 2008 do not represent a random sampling of peoples and cultures. Therefore the distribution of pro-social and anti-social punishment observed across these groups should not be considered representative. Instead, these experiments provide evidence that a sizeable degree of anti-social punishment exists (even absent explicit retaliatory motives), that the level of anti-social punishment varies across subject pools in systematic ways, and that in some subject pools anti-social punishment is as common, or even more common, than pro-social punishment.

In summary, there is a large body of experimental evidence for punishment targeted at cooperators. This basic phenomenon of anti-social punishment exists separately from retaliation in repeated games, and cannot be neglected in models of the co-evolution of cooperation and punishment. Moreover, anti-social punishment is an interesting behavior that is in need of an evolutionary explanation of its own. Addressing that question is the focus of this paper.

Strategy frequencies averaged over 100,000 parameter sets

To demonstrate that our results are not unique to the particular parameter values used in the main text Figure 1, we calculate the steady state frequency of each strategy averaged over 100,000 parameter sets using $N = 100, n = 5, c = 1$ and randomly sampling from uniform distributions on the intervals $1 < r < 5, 0 < \sigma < (r-1)c, 0 < \gamma < 5$, and $\gamma < \beta < 5\gamma$. In Supplementary Table S2, the average frequency of each strategy is shown for each strategy set (No punishment, limited strategy set, full strategy set). A strategy is defined by public goods game action (C=cooperator, D=defector, L=loner), and punishment decision taken towards cooperators (N = no action, P = punish), defectors (N or P) and loners (N or P). Supplementary Table S3 shows the aggregated frequency of cooperators, defectors and loners, as well as of strategies which punish cooperators, defectors and loners. Note that a single strategy can punish more than 1 public goods game move, and thus punishment of cooperators defectors and loners does not sum to 100%. These averages taken over 100,000 parameter sets are remarkably similar to the results in main text Figure 1, demonstrating the robustness of our findings.

Higher mutation rate analysis using agent based simulation

The calculations presented in the main text assume a vanishingly small mutation rate u. We now use agent based simulations to explore higher mutation rates. We simulate a Moran process with an exponential fitness function[40]. Payoffs for each agent are calculated as described in the Supplementary Methods, based on the frequency of each strategy in the population. Then in each generation, (i) one agent i is randomly chosen to change strategy, and (ii) with probability $1-u$, agent i picks another agent j to imitate, proportional to $exp[\pi_j]$; or with probability u, a 'mutation' occurs instead and agent i picks a new strategy at random.

For $u = 0.001$ and $u = 0.01$, we fix $c = 1$, $N = 100$ and $n = 5$, and randomly sample 1000 parameter sets from the region $1 < r < 5$, $0 < \sigma < (r - 1)c$, $0 < \gamma < 5$ and $\gamma < \beta < 5\gamma$. For each parameter set, we simulate for 10^8 generations, and calculate the time averaged frequency of each strategy over the second half of the simulation (Supplementary Table S4). The results using $u = 0.001$ are very similar to the low mutation limit calculation. At $u = 0.01$, the results are qualitatively similar, except that the strategies which do not punish are somewhat more common while the punishing strategies are correspondingly less common.

Supplementary Table S5 shows that the average frequency of cooperators, defectors and loners are strikingly consistent across the three different mutation rates. At $u = 0.001$, the frequency of cooperation is less than 0.5 in 95% of parameter sets. At $u = 0.01$, the frequency of cooperation is less than 0.5 in 93% of parameter sets. We also see that although punishment is less common when $u = 0.01$, it remains true that for a given mutation rate, the frequencies of all three types of punishment are roughly equal. Thus the results presented in the main text are not unique to the low mutation limit. Also here we find that punishment does not effectively promote cooperation when using the full punishment strategy set, and anti-social punishment is approximately as common as pro-social punishment.

Inverse relationship between cooperation and anti-social punishment: Sensitivity analysis

In the main text, Figure 3A shows a strong negative relationship between cooperation and anti-social punishment across parameter sets. Here we explore that relationship in more detail. We define the frequency of cooperation f_C to be the steady state frequency summed over all strategies which cooperate in the public goods game (C-***); and the frequency of anti-social punishment f_{AP} to be the steady state frequency summed over all strategies which punish cooperators (*-P**).

We begin by noting that almost all points in main text Figure 3A lie below the line $f_C + f_{AP} = 1$. In the full strategy set, there are several strategies which both cooperate and engage in anti-social punishment: C-PNN, C-PNP, C-PPN and C-PPP. Thus points in Figure 3A can in principle be distributed anywhere in the unit square, and there is no *a priori* $f_C + f_{AP} \leq 1$ constraint. The fact that very few points lie above $f_C + f_{AP} = 1$ is a result of natural selection: cooperators who also punish cooperators usually fare poorly and are strongly disfavored.

Moreover, the relationship between cooperation and anti-social punishment involves more than an implicit $f_C + f_{AP} \leq 1$ constraint. Randomly sampling points which satisfy this constraint and regressing f_C against f_{AP} results in $f_C = 0.5 - 0.5 f_{AP}$. Performing the same regression on the data in Figure 3A, however, gives $f_C = 0.9 - 1.43 f_{AP}$.

To explore the nature of this dependence, we perform a sensitivity analysis. We fix $N = 100$, $n = 5$ and $c = 1$, and randomly sample 10,000 parameter sets from the intervals $1 < r < 5$, $0 < \sigma < (r-1)c$, $0 < \gamma < 5$, and $\gamma < \beta < 5\gamma$. For each parameter set, we calculate the steady state frequency of each strategy in the low mutation limit and determine the ratio of cooperation to anti-social punishment f_C/f_{AP}. We then determine the marginal effect of each payoff parameter by incrementing each parameter individually by 0.01 and recalculating f_C/f_{AP}. The most influential parameter for a given parameter set is thus the parameter that when incremented results in the largest change in f_C/f_{AP} relative to the baseline. As shown in Supplementary Figure S3A, we find that r is the most influential parameter in a large majority of parameter sets. Multiplying each parameter by 1.01 instead of adding 0.01 gives similar results, with r again being most influential in a large majority of parameter sets (Supplementary Figure S3B).

To further visualize these findings, we systematically vary each parameter. We use $N = 100$ and $n = 5$, and fix a given parameter to a particular value, and then sample the other parameters from the intervals $1 < r < 5$, $0 < c < 5$, $0 < \sigma < (r-1)c$, $0 < \gamma < 5$, and $\gamma < \beta < 5\gamma$. For each parameter set, we calculate the steady state frequency of each strategy in the low mutation limit, and from this determine the ratio of cooperation to anti-social punishment f_C/f_{AP}. The median f_C/f_{AP} for each parameter value is shown in Supplementary Figure S4. We see that f_C/f_{AP} steadily and consistently increases with r over the full parameter range (for the game to be a social dilemma using $n = 5$, r is necessarily restricted to $1 < r < 5$). Conversely, f_C/f_{AP} only varies with c if $c < 1$, only varies with β when β is very small, and does not vary substantially with σ or γ. Supplementary Figure S4 therefore suggests that overall, the cooperation multiplier r is the most important parameter for determining the ratio of cooperation to anti-social punishment.

We also find important interactions between r and the other parameters. To illustrate this, we replicate the previous analysis in Supplementary Figure S4, restricting to $r > 4$ and $r < 2$ (Supplementary Figure S5). Most strikingly we see that decreasing β dramatically increases f_C/f_{AP} when r is large, and dramatically reduces f_C/f_{AP} when r is small. A qualitatively similar pattern, of markedly smaller magnitude, is seen for c. To summarize, we see that cooperation is most successful when the returns on cooperation are large and punishment is weak.

In addition to these results on the relative frequency of cooperation and anti-social punishment, our model also shows two other statistical regularities. We find a significant negative relationship between the frequency of pro-social punishment f_{PP} and anti-social punishment f_{AP} across parameter sets ($f_{PP} = -0.31 f_{AP} + 0.54$, p ¡ 0.001), and a significant positive relationship between the frequency of cooperation with punishment f_C and the frequency of cooperation without punishment f_{C-NP} ($f_C = 0.31 f_{C-NP} + 0.23, p < 0.001$).

Connections between model results and sociological variables

As shown in Figure 3 of the main text, as well as the previous section of the SI, we find that across parametert sets, cooperation and anti-social punishment levels are inversely related; pro-social and anti-social punishment frequencies are inversely related; and cooperation in the punishment and no-punishment conditions are positively related. We also find that higher levels of cooperation are associated with more efficient public goods (larger values of cooperation multiplier r) and with less severe costly punishments (smaller values of the punishment effectiveness β).

Each of these statistical regularities is present in the cross-cultural data of Herrmann et al. 2008. While non-cooperators bear the brunt of the punishment in countries which achieve high levels of cooperation, the opposite is true in less cooperative countries, where free-riders are more often willing to pay to punish those who contribute more to than them to the public good. This relationship is clearly demonstrated in Figure 2B and Table 1 of Herrmann et al. 2008, and in our Supplementary Discussion above. Our analysis also shows a significant negative correlation between pro-social and anti-social punishment in the data of Herrmann et al. 2008, as well as a positive correlation between cooperation in the games with and without punishment opportunities (see Supplementary Figure S2).

We now consider how our predictions about the role of the parameters r and β relate to sociological variables that predict cooperation and anti-social punishment across societies. It has been shown that that subjects from more affluent countries with stronger rule of law and higher GDP tend to engage in less anti-social punishment than those from poorer counties with weaker legal institutions[27]; and that societies with more market integration engage in more cooperation[62]. These results are consistent with our finding that the cooperation multiplier r, a measure of the efficiency of cooperation and public goods, is a key factor in determining the level of cooperation versus anti-social punishment. Weak rule of law and the associated corruption have been shown to reduce the efficiency of public goods by inflating the cost of public projects[63]; and market integration facilitates the specialization and division of labor[64] that allow groups to achieve more than each individual would alone. Thus if people evolved their intuitions according to the process described in our model, we would expect experimental subjects who developed their intuitions in high r settings to be more cooperative and less likely to punish anti-socially. This is true even if the game they play in the experiment is not identical to the setting in which their intuitions evolved, as discussed at greater length below.

We also find empirical support for our prediction that more cooperative, productive societies engage in less severe peer punishment. Legal scholars have shown that the majority of violent crimes are moralistic in nature, 'expression[s] of a grievance by unilateral aggression'[65] that map directly onto the type of vigilante justice represented by high-effectiveness (large β) peer punishment. Yet cooperative, trusting societies which are more economically successful have lower rates of violent crime[66], instead relying on less destructive (more symbolic) forms of norm enforcement. This is consistent with our result that successful cooperation requires relatively ineffectual punishments (small β). Thus once again, differences in evolved intuitions caused by the societal variation described here may lead subjects to behave differently across subject pools in economic game experiments, even when participating in games with objectively identical payoff structures. Exploring these connections between sociological variables and model parameters is an important area for future research.

Evolved intuitions versus in-game learning

The model we present makes predictions about the strategies, intuitions and heuristics that are favored by natural selection. When comparing model predictions with laboratory experiments, it is important to remember that human psychology did not evolve in the lab; instead, the lab gives us a window into our psychology which has evolved outside of the lab. No experiment is a perfect match to this reality: just as mathematical models are simplifications of complex scenarios, so too are laboratory experiments. Therefore it takes subjects some time to adjust away from their real world intuitions and acclimate to the details of the particular laboratory game they are playing. This adjustment process is clearly demonstrated by the learning evident in virtually all experiments with multiple trials. If we want to gain insight into evolved intuitions, rather than in-game learning and adaptation to the exogenously imposed experimental setting, it is therefore necessary to focus on behavior in the first trial or round (while keeping in mind that these intuitions involved in settings which likely differ from the experimental setting at hand).

As shown earlier in this Supplementary Discussion, large cross-cultural differences exist in first-round behavior, despite identical game setups. This variation in behavior reflects the variation in evolved intuitions subjects bring into the lab. Conversely, relatively little variation in first-round behavior is often observed between different experimental designs applied to the same subject pool. For example, consider two studies which vary the punishment technology β/γ in a compulsory public goods game[67,68]. Both studies find that larger punishment technologies lead to higher contribution levels, but only after several periods of learning. In the first period of both experiments, there is no significant difference in contribution across treatments (Linear regression with robust standard errors clustered on subject, taking the no-punishment treatment as the baseline and adding dummy variables for each punishment treatment; Egas & Riedl 2008: 1:1, coeff = 0.40, p = 0.53; 1:3, coeff = 0.73, p = 0.26; 3:1, coeff = -0.28, p = 0.65; 3:3, coeff = 0.65, p = 0.32; Nikiforakis & Normann 2008: 1:1, coeff = -2.12, p = 0.288; 1:2, coeff = 0.50, p = 0.80; 1:3, coeff = 0.08, p = 0.97; 1:4, coeff = 1.96,p = 0.33). Our model seeks to explain variation in first-round play resulting from intuitions developed in daily life, rather than this in-game learning over the course of the session. Therefore these results do not necessarily contradict our model. If instead we consider cross-cultural variation in first round play[27], we find sociological evidence supporting our model prediction. As described in the previous section, high cooperation is achieved by subjects from affluent societies who developed their intuitions with little violent peer norm enforcement[66], and the opposite is true of subjects from less affluent countries with more violent (truly costly) enforcement. Exploring these connections more deeply is an important area for future research.

The observed cross cultural differences in behavior[27,30,31] have important implications for all models of the evolution of cooperation and punishment. Successful models must explain the punishment of both non-cooperators and cooperators. The evolution of anti-social punishment is particularly challenging for models based on multi-level selection[15]. Selection at the level of the group would always favor cooperation and punishment of non-cooperators, and always disfavor anti-social punishment. Therefore such models cannot explain the tendency to punish cooperators exhibited in many societies. Our framework, however, supports the heterogeneity that has been empirically observed.

While providing a potential ultimate evolutionary explanation for anti-social punishment, our model does not address the proximate psychological mechanisms behind such behavior. The extent to which punishment of cooperators is motivated by dislike of do-gooders, desire for revenge against anticipated pro-social punishment, concern with absolute versus relative payoffs, or other related thought processes merits further study. It is also important to note that although our model considers discrete options (cooperation, defect or opt out, punish or do not punish), many experiments present subjects with a continuous choice of how much to contribute, and how much to punish.

Behavior of loners in compulsory games: evidence from two novel experiments

Almost all experiments studying punishment have used compulsory rather than optional public goods games. That does not mean, however, that only compulsory games are important for the evolution of human cooperation. Many important real-world situations involving cooperation are optional, such as entrepreneurs choosing to start a business together, individuals choosing to move into an apartment together or scientists choosing to collaborate with each other. Thus our instincts and understandings of appropriate behavior are likely shaped by optional as well as compulsory games. Consistent with this observation, many subjects choose to opt out in favor of a fixed loners payoff when given the chance[38], and this stabilizes cooperation over a series of games in the laboratory through the cooperator-defector-loner cycle predicted theoretically[17,18,19,44].

In order to fully understand behavior in compulsory public goods games, therefore, it is necessary to know how loners will play when not given the chance to abstain for a fixed payoff. Together with previous experiments on compulsory games[27], our model makes clear predictions about how loners will behave: based on our theoretical results, loners are predicted to be lower contributors in compulsory games, and to be most likely to punish high contributors. To evaluate these predictions, we conducted two incentivized behavioral experiments. Experiment 1 investigates the contribution behavior of loners in a compulsory game, while Experiment 2 considers the degree of anti-social punishment exhibited by loners in a compulsory game. In both experiments, subjects were recruited using the online labor market Amazon Mechanical Turk (AMT)[51]. On AMT, people from around the world can be hired to perform short tasks (generally 5 minutes or less) for small amounts of money (generally under $1). The amount paid can be conditioned on the outcome of the task, allowing for performance-dependent payments and incentive-compatible designs. AMT therefore offers an unprecedented tool to quickly and inexpensively recruit subjects for economic game experiments. A variety of standard results have been successfully replicated on AMT by scholars across the social science disciplines[51,55,56]. Most relevant to the current experiment is a demonstration of quantitative agreement between the traditional laboratory public goods games of Fehr & Gachter 2000[9] and low-stakes public goods games run on AMT with endowments of $0.10 or $0.05[58]; and between a one-shot Prisoner's Dilemma run in the physical laboratory and one run on AMT using identical instructions but 10-fold lower payoffs[51]. This relative insensitivity to stake size in social dilemmas is consistent with previous research in the traditional laboratory (see[69] for a thorough review). We will now describe each experiment in detail.

Experiment 1

In October 2010, 124 subjects were recruited through AMT to participate in Experiment 1. Subjects received a $0.20 showup fee for participating, and earned on average an additional $0.63 based on decisions made in the experiment. First, subjects read a set of instructions for a one-shot public goods game, in which groups of 4 interact with a cooperation multiplier of r = 2, each choosing how much of a $0.40 endowment to contribute (in increments of 2 cents to avoid fraction cent amounts). Subjects then answered two comprehension questions to ensure they understood the payoff structure, and only those who answered correctly were allowed to participate.

Treatment experiment

The 73 subjects randomly assigned to the treatment group were then informed that the game was optional, and they could choose either to participate or to abstain in favor of a fixed payoff of $0.50. Those who chose to participate then indicated their level of contribution. Next, subjects were informed that they would play a second game with 3 new partners, which was a compulsory version of the first game. They were further informed that one of the two games (optional or compulsory) would be randomly selected to determine their payoff. This was done to keep the payoff range the same as in the control experiment described below, where subjects played only one game. In order to prevent between-game learning, subjects were not informed about the outcome of the optional game before making their decision in the compulsory game. The instruction used are included at the end of this section, and payoffs were determined exactly as described (no deception was used). Throughout the analysis of all experiments, we present the results of linear regressions with robust standard errors when considering non-binary dependent variables. If tobit regression is instead used when considering contribution amount, to account for minimum and maximum values, qualitatively equivalent results are found unless otherwise noted.

Using this experimental setup, we can (i) classify subjects as loners or non-loners based on their decision in the first (optional) game, and then (ii) ask how loners versus non-loners behave in the second (compulsory) game. Comparing average contribution in compulsory game, we find that those who opt out in the optional game contribute significantly less in the compulsory game (Rank-sum, N=73, p=0.006; main text Figure 4a). This remains true when controlling for additional sources of variation in a linear regression with robust standard errors (see Supplementary Table S6). We regress contribution in the compulsory game as the dependent variable (0-40) against decision in the optional game (0 = participate (non-loner), 1 = opt out (loner)), including controls for age, gender (0 = male, 1 = female) and US residence (0 = non-US, 1 = US). We find a significant negative relationship between contribution in the compulsory game and opting out (coeff = -10.880, p = 0.006). Thus loners contribute significantly less in the compulsory game.

To explore the role that risk preferences play in our experiment, we assessed subjects' appetite for risk-taking with the question 'Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?'. Subjects responded using an 11-point Likert scale ranging from 'Unwilling to take risks' to 'Fully prepared to take risks'. This self-report risk question was used by Dohmen et al. (forthcoming)[7]6. In a representative sample of 450 individuals, they find that this measure provides a good proxy of actual risk taking behaviors, such as traffic offenses, portfolio choice, smoking, occupational choice, participation in sports and migration. We find that the relationship between opting out and contribution in the compulsory game remains significant (coeff = -8.570, p = 0.047) when controlling for risk-taking (although the relationship crosses the 5% significance threshold when using Tobit regression and controlling for risk, p = 0.053).

While our model considers a binary cooperation decision (cooperate or defect), our experiments follow the typical experimental design of a continuous decision (in our case, contributing between 0 and 40 units). The maximum benefit is generated when individuals contribute their entire endowment. Thus any deviation below the maximum amount constitutes some level of free-riding. Our data show that loners engage in substantially more free-riding than non-loners, as predicted by our model. The relationship between compulsory contribution and opting out has a substantial effect size as well as statistical significance, with non-loners contributing a 44% larger amount compared to non-loners (non-loners: 29.4, loners: 20.4).

Control experiment

To demonstrate that behavior in the compulsory game discussed above was not affected by the preceding optional game, the remaining 51 subjects participated in a control condition in which they played only a compulsory game. We find no significant difference in compulsory contribution levels between the treatment and the control (Rank-sum, $p = 0.620$). This result is robust to controls for age, gender, risk-taking and US residence in a linear regression with robust standard errors (Supplementary Table S7: coeff = 2.21, $p = 0.432$). Thus it does not seem that playing the optional game first in the treatment condition changed how subjects subsequently played in the compulsory game.

Summary statistics for both conditions of Experiment 1 are shown in Supplementary Table S8.

Experiment 2

In Experiment 1, we investigated contribution behavior of loners in a compulsory game without punishment. We now turn our attention to how loners play in a compulsory game with punishment. In November 2010, 196 subjects were recruited through AMT to participate in Experiment 2. Subjects received a $0.40 showup fee for participating, and earned on average an additional $0.92

based on decisions made in the experiment. As in Experiment 1, subjects began by reading a set of instructions for a one-shot public goods game with groups of 4, a cooperation multiplier of r = 2, and a $0.40 endowment, and then answered two comprehension questions in order to participate. Subjects were then informed that the game was optional, and decided whether to participate or opt out for a fixed $0.50 payment. Subjects who chose to participate were given 5 contribution levels to choose from: 0, 10, 20, 30 or 40. Thus the first game of Experiment 2 is identical to that of Experiment 1's treatment condition, except for a more limited set of contribution options to facilitate punishment decisions as described below.

Subjects were then informed that they would play a second game with 3 new partners, which differed from the first game in two ways: it was compulsory, and it would be followed by a Stage 2 in which participants could interact directly with each other group member. In Stage 2, subjects had three direct actions to pick from: choosing option A had no effect on either player; choosing option B caused them to lose 4 cents while the other player lost 12 cents; and choosing option C caused them to lose 8 cents while the other player lost 24 cents. Thus we offered mild (B) and severe (C) punishment options with a 1:3 punishment technology. Subjects were allowed to condition their Stage 2 choice on the other player's contribution in the compulsory public goods game. To do so, we employ the strategy method[70]: subjects indicate which action (A, B, C) they would take towards group members choosing each possible contribution level (0, 10, 20, 30, 40). It has been shown that using the strategy method to elicit punishment decisions has a quantitative (although not qualitative) effect on the level of punishment targeted at defectors, but has little effect on the punishment targeted at cooperators[71]. As anti-social punishment is our main focus, we therefore feel confident in our use of the strategy method. In order to prevent between-game learning, subjects were not informed about the outcome of the optional game before making their decision in the compulsory game. The instructions used are included at the end of this section, and payoffs were determined exactly as described (no deception was used).

Summary statistics are given in Supplementary Table S9.

Validation

We begin by showing that the compulsory game data from Experiment 2 displays expected statistical regularities based on past research. Consistent with our analysis of Herrmann et al. 2008 above, our AMT data show a negative relationship between contribution in the compulsory game and anti-social punishment. Note that in the current analysis it is not necessary to cluster regressions on either group or subject because of our use of the strategy method: subjects received no information about the actual contribution decisions of others before making their punishment decisions, and made only a single set of decisions.

An ordinal probit regression with robust standard errors (Supplementary Table S10) finds a significant positive correlation between your contribution in the compulsory game and the amount of punishment (A=0, B=1, C=2) you target at those who contribute nothing (coeff = 0.018, p = 0.001); and a significant negative correlation between your contribution in the compulsory game and the amount spent punishing those who contribute the maximum amount of 40 (coeff = -0.014, p = 0.027). Furthermore, consistent with the large body of experimental evidence on punishment in public goods game, we see significantly more punishment targeted at those who contribute nothing compared to those who contribute the maximum amount (Sign-rank, p=0.002). Thus pro-social punishment is more common than anti-social punishment. These results give us confidence that our experimental design is sufficiently similar to that of previous experiments.

Next we compare behavior in the optional game of Experiment 2 to that of Experiment 1's treatment condition. We find no significant difference in the fraction of subjects who opt out (Chi^2 test, p = 0.680), or in contribution among those who participate (Rank-sum, p = 0.339). Experiment 2 also replicates the result from Experiment 1 that loners contribute less in the compulsory game (linear regression with robust standard errors: coeff = -7.188, p = 0.005). A regression including data from both experiments continues to find a significant negative relationship between compulsory contribution and opting out (coeff = -7.785, $p < 0.001$), and finds no significant interaction between experiment and opting out (coeff = 1.76, p = 0.680). Thus the relationship between being a loner and a low contributor in the compulsory game is the same in both studies. These results demonstrate consistency across studies with similar designs on AMT.

Loners and anti-social punishment We now turn to our primary question of interest: are loners more likely to punish high contributors? Because we use the strategy method, we can conducted fine-grained analysis of anti-social punishment behavior. We begin by considering the most clear-cut case, punishment that is directed at those who contribute the maximum amount (40 units). This behavior is indisputably anti-social. We find that loners spend significantly more punishing maximal contributors (Rank-sum, p = 0.003, main text Figure 4b; Supplementary Table S11), and that this result is robust to controlling for age, gender, risk-taking and US residence in an ordinal probit regression with robust standard errors (coeff = 0.647, p = 0.006). It might that loners are more likely to punish high contributors merely because they themselves are lower contributors on average. However, the relationship between opting out and punishing maximal contributors is robust (coeff = 0.579, p = 0.019) to including an additional control for the punisher's compulsory contribution, demonstrating that this is not the case.

We also find that loners spend significantly more punishing those who contributed as much or more than themselves (amount spent on punishment averaged over all contribution levels equal to or greater than the subject's own contribution in the compulsory game; linear regression with robust standard errors, Supplementary Table S12; coeff = 0.286, p = 0.013).

Furthermore, we find the same positive correlation between opting out and punishing those that contribute 30, while conversely, no such correlations exist for punishment of lower contribution amounts (Supplementary Table S13). Taken together, these results show that subjects who opt out of optional games engage in more anti-social punishment than those who choose to participate.

Discussion

In the main text, we show that natural selection can favor punishment targeted at cooperators, and that this punishment is dealt out largely by loners. We then argue that this model may help understand cross-cultural experiments which show low contributors punishing high contributors in compulsory games.

Central to this argument is the idea that in compulsory games, many low contributors may in fact be loners who would opt out if given the chance. In Experiment 1, we provide empirical evidence consistent this prediction. We find that subjects that opt out of optional games contribute significantly less in compulsory games. These results are consistent with early experimental work showing that in a 2 player Prisoner's Dilemma games, giving the choice of not participating increases cooperation in the games which occur because defectors are more likely to opt out[72]. Our model also predicts that in compulsory games with punishment, most of the punishment targeted at cooperators comes from these defectors that would opt out if given the chance. The results of Experiment 2 are consistent with this second prediction. We find that subjects who opt out of optional games engage in significantly more punishment of high contributors, even when controlling for contribution level in the compulsory game and self-reported risk preferences. Interestingly, we find no difference in punishment of low contributors between those who do and do not choose to participate in the optional game: while non-loners preferentially punish low contributors, loners punish both low and high contributors. Extending our simple framework to include effects such as repetition and reputation is an important area for future work, which may help explain these features of the behavioral data. In sum, the results of these preliminary experiments highlight the importance of studying optional games experimentally. Extending previous cross-cultural experiments on anti-social punishment [27,30,31] to consider optional games is an important next step for future research.

Experimental instructions

Experiment 1 Treatment instructions

Page 1:

You have been randomly assigned to interact with 3 other people. All of you receive this same set of instructions. You cannot participate in this study more than once.

Each person in your group is given 40 cents for this interaction (in addition to the 20 cents you received already for participating). You each decide how much of your 40 cents to keep for yourself, and how much (if any) to contribute to the groups common project (in increments of 2 units: 0, 2, 4, 6 etc).

All money contributed to the common project is doubled, and then split evenly among the 4 group members. Thus, for every 2 cents contributed to the common project, each group member receives 1 cent.

If everyone contributes all of their 40 cents, everyones money will double: each of you will earn 80 cents.

But if everyone else contributes their 40 cents, while you keep your 40 cents, you will earn 100 cents, while the others will earn only 60 cents. That is because for every 2 cents you contribute, you get only 1 cent back. Thus you personally lose money on contributing.

The other people are REAL and will really make a decision there is no deception in this study.

Page 2:

You MUST answer these two questions correctly to have your work accepted and to receive a bonus!

1. What level of contribution is best for the group as a whole?

2. What level of contribution earns the highest payoff for you personally?

Page 3:

In this task, your bonus depends on both your decision and the decision of the others in your group.

If all group members (including you) contributed the maximum amount, all group members earn an $0.80 bonus. If all group members (including you) contributed nothing, all group members earn a $0.40 bonus. You can now choose whether to (i) actually engage in this task, or (ii) opt out, and instead receive a fixed bonus of $0.50 (regardless of what anyone else choses in the task). Which do you choose?

Page 4 [Only appears for subjects who did not opt out]:

Please choose the amount of money you wish to contribute.

Page 5:

Now you will interact with 3 new people in a 2nd task. This task is exactly the same, except that you do not have the option to opt out - you (and everyone else) must participate in the task. There is a 50% chance that your bonus will be determined by your decision in the first task, and a 50% chance that your bonus will be determined by the decision you make now. Please choose the amount of money you wish to contribute.

Experiment 1 Control instructions

Page 1:

You have been randomly assigned to interact with 3 other people. All of you receive this same set of instructions. You cannot participate in this study more than once.

Each person in your group is given 40 cents for this interaction (in addition to the 20 cents you received already for participating). You each decide how much of your 40 cents to keep for yourself, and how much (if any) to contribute to the groups common project (in increments of 2 units: 0, 2, 4, 6 etc).

All money contributed to the common project is doubled, and then split evenly among the 4 group members. Thus, for every 2 cents contributed to the common project, each group member receives 1 cent.

If everyone contributes all of their 40 cents, everyones money will double: each of you will earn 80 cents.

But if everyone else contributes their 40 cents, while you keep your 40 cents, you will earn 100 cents, while the others will earn only 60 cents. That is because for every 2 cents you contribute, you get only 1 cent back. Thus you personally lose money on contributing.

The other people are REAL and will really make a decision there is no deception in this study.

Page 2:

You MUST answer these two questions correctly to have your work accepted and to receive a bonus!

1. What level of contribution is best for the group as a whole?

2. What level of contribution earns the highest payoff for you personally?

Page 3:

Please choose the amount of money you wish to contribute.

Experiment 2 Treatment instructions

Page 1:

You have been randomly assigned to interact with 3 other people. All of you receive this same set of instructions. You cannot participate in this study more than once.

Each person in your group is given 40 cents for this interaction (in addition to the 40 cents you received already for participating).

You each decide how much of your 40 cents to keep for yourself, and how much (if any) to contribute to the groups common project.

All money contributed to the common project is doubled, and then split evenly among the 4 group members. Thus, for every 2 cents contributed to the common project, each group member receives 1 cent.

If everyone contributes all of their 40 cents, everyones money will double: each of you will earn 80 cents.

But if everyone else contributes their 40 cents, while you keep your 40 cents, you will earn 100 cents, while the others will earn only 60 cents. That is because for every 2 cents you contribute, you get only 1 cent back. Thus you personally lose money on contributing.

The other people are REAL and will really make a decision there is no deception in this study.

Page 2:

You MUST answer these two questions correctly to have your work accepted and to receive a bonus!

1. What level of contribution is best for the group as a whole?

2. What level of contribution earns the highest payoff for you personally?

Page 3:

In this task, your bonus depends on both your decision and the decision of the others in your group. If all group members (including you) contributed the maximum amount, all group members earn an $0.80 bonus. If all group members (including you) contributed nothing, all group members earn a $0.40 bonus. You can now choose whether to (i) actually engage in this task, or (ii) opt out, and instead receive a fixed bonus of $0.50 (regardless of what anyone else chooses in the task).

Page 4 [Only appears for subjects who did not opt out]:

Please choose the amount of money you wish to contribute.

[0 10 20 30 40]

Page 5:

Now you will interact with 3 new people in a 2nd task. This task is the same as the interaction you just participated in, with two important differences:

1) This interaction is compulsory. There is no choice to opt out, and instead everyone must participate in the interaction.

2) The decision about how much to contribute to the common project (Stage I) is followed by Stage II. In Stage II, all group members receive an extra 50 cents, and then you interact with each of the three other group members individually. You must decide between one of three possible actions, A B or C, toward each of the three other players. Each other player will also choose an action towards you.

If you choose A then you lose 0 cents, and the other player loses 0 cents.

If you choose B then you lose 4 cents, and the other player loses 12 cents.

If you choose C then you lose 8 cents, and the other player loses 24 cents.

You can base your Stage II decision on the amount the other group member contributed in Stage I.

————————-

Please choose the amount of money you wish to contribute in Stage I:

[0 10 20 30 40]

Page 6:

Now you must choose your Stage II actions.You must decide between one of three possible actions, A B or C, toward each of the three other players. Each other player will also choose an action towards you.

If you choose A then you lose 0 cents, and the other player loses 0 cents.

If you choose B then you lose 4 cents, and the other player loses 12 cents.

If you choose C then you lose 8 cents, and the other player loses 24 cents.

Each other player will also choose an action towards you.

You (and the other players) can base your decision on each other group member's contribution in Stage I:

What action will you take towards group members who contribute 0? [A B C]

What action will you take towards group members who contribute 10? [A B C]

What action will you take towards group members who contribute 20? [A B C]

What action will you take towards group members who contribute 30? [A B C]

What action will you take towards group members who contribute 40? [A B C]

Compulsory public goods games without loners

To explore the evolution of anti-social punishment in compulsory public goods games, we replicate our analysis in a game without loners. Strategy definitions and payoffs are the same as in the game with loners (Supplementary Methods), with two exceptions: (i) only two options are available for the public goods game, cooperation or defection; and (ii) the 4th strategy parameter (whether to punish loners) is irrelevant as there are no loners, and so we do not include it in our strategy definitions. Thus we have 8 strategies total: D-NN, D-NP, D-PN, D-PP, C-NN, C-NP, C-PN and C-PP, with payoffs as follows.

Simplifying Eq (S30) for $X_L = 0$ gives

$$B(\mathbf{X}) = \frac{rcX_C}{n}\frac{n-1}{N-1} \tag{S36}$$

Simplifying Eq (S31) for $X_L = 0$ gives

$$F = 1 - r\frac{N-n}{n(N-1)} \tag{S37}$$

The resulting payoffs are then given by

$$\pi_{D-NN} = B(\mathbf{X}) - X_{pD}\beta\frac{n-1}{N-1} \tag{S38}$$

$$\pi_{D-NP} = B(\mathbf{X}) - ((X_D-1)\gamma + (X_{pD}-1)\beta)\frac{n-1}{N-1} \tag{S39}$$

$$\pi_{D-PN} = B(\mathbf{X}) - (X_{pD}\beta + X_C\gamma)\frac{n-1}{N-1} \tag{S40}$$

$$\pi_{D-PP} = B(\mathbf{X}) - ((X_C+X_D-1)\gamma + (X_{pD}-1)\beta)\frac{n-1}{N-1} \tag{S41}$$

$$\pi_{C-NN} = B(\mathbf{X}) - Fc - X_{pC}\beta\frac{n-1}{N-1} \tag{S42}$$

$$\pi_{C-NP} = B(\mathbf{X}) - Fc - (X_{pC}\beta + X_D\gamma)\frac{n-1}{N-1} \tag{S43}$$

$$\pi_{C-PN} = B(\mathbf{X}) - Fc - ((X_C-1)\gamma + (X_{pC}-1)\beta)\frac{n-1}{N-1} \tag{S44}$$

$$\pi_{C-PP} = B(\mathbf{X}) - Fc - ((X_C-1+X_D)\gamma + (X_{pC}-1)\beta)\frac{n-1}{N-1} \tag{S45}$$

We now study the evolutionary dynamics of this game in the low mutation limit. We begin by fixing $c = 1$, $N = 100$ and $n = 5$, and randomly sampling 10,000 parameter sets from the intervals $1 < r < 5$, $0 < \sigma < (r-1)c$, $0 < \gamma < 5$ and $\gamma < \beta < 5\gamma$. For each parameter set we calculate the steady state frequency of each of the 8 strategies. Averaging over the 10,000 parameter sets gives D-NN=50.0% and D-PN=50.0%, and all other strategies with frequency $< 0.1\%$. Thus when selection is relatively strong, defection completely dominates cooperation, and anti-social defectors are as common as defectors that do not punish.

For further insight into the game without loners, we next explore the effect of reducing the intensity of selection. We modify the evolutionary dynamic such that player i's fitness is given by $f_i = e^{w\pi_i}$ (rather than $f_i = e^{\pi_i}$), where w controls the intensity of selection. Setting $w = 0$ gives neutral selection. Setting $w \to \inf$ gives deterministic game dynamics. Our simulations up until this point implicitly use a value of $w = 1$. Supplementary Fig S6 shows the steady state frequency of each strategy averaged over 10,000 parameter sets for w between 10^{-4} and 1. We see that at intermediate intensities of selection, the two most common strategies are defectors that punish cooperators D-PN, and cooperators that punish defectors C-NP, although D-PN is always more common than C-NP. Examining parameter sets individually shows that defection is always more common than cooperation, and that anti-social punishment is always more common than cooperation, regardless of w or the payoff parameters. Furthermore, as in the game without loners, we find a strong negative correlation when regressing cooperation level against anti-social punishment level across parameter sets ($w = 1$, coeff=-0.41; $w = 0.1$, coeff=-0.45; $w = 0.01$, coeff=-0.20).

Thus we find that in the low mutation limit, loners are not needed for evolution to favor anti-social punishment. Loners are, however essential for the success of cooperation in our framework, as has been shown previously[44]. Other mechanisms for the evolution of cooperation[1], such as spatial structure[73] or reputation[74], may also produce interesting results in conjunction with anti-social punishment, and are important topics for future research.

Parameter sampling method

Throughout our analysis, we present results aggregated over large sets of randomly sampled parameters. To do so, we randomly choose values of $r \in [1, 5]$, $c \in [0, 5]$, $\sigma \in [0, 5]$, $\gamma \in [0, 5]$ and punishment technology $\beta/\gamma \in [1, 5]$. If we find that the constraint $\sigma < (r - 1)c$ is not satisfied, we resample until the constraint is met. Here we show that alternate sampling procedures yield equivalent results. In Supplementary Table S14 we show the steady state frequencies of each strategy in the low mutation limit averaged over 100,000 parameter sets sampled using the following methods: (i) the procedure used in the main text (Method 1); (ii) sampling $r \in [1, 5]$, $c \in [0, 5]$, $\sigma \in [0, 5]$,

$\gamma \in [0, 5]$ and $\beta \in [0, 25]$, and then resampling if $\sigma > (r - 1)c$ or $\beta > 5\gamma$ (Method 2); and (iii) sampling $r \in [1, 5]$, $c \in [0, 5]$, $\sigma \in [0, (r-1)c]$, $\gamma \in [0, 5]$ and punishment technology $\beta/\gamma \in [1, 5]$ (Method 3). As can be seen, the results are extremely similar across the three sampling methods.

Second order punishment

The issue of second-order punishment has received some attention in the literature[17,75]. Here punishers pay to punish both defectors and cooperators that do not punish (second order free-riders). In our framework we do not consider second order punishment. Using our larger strategy set, it is unclear which strategies should be the targets of second order punishment. We also question the appropriateness of the second order free-rider concept in the context of our model (see main text discussion).

Supplementary references

60. Bergin, J. & Lipman, B. L. Evolution with state-dependent mutations. *Econometrica* **64**, 943-956, (1996).

61. Antal, T., Traulsen, A., Ohtsuki, H., Tarnita, C. E. & Nowak, M. A. Mutation-selection equilibrium in games with multiple strategies. *Journal of theoretical biology* **258**, 614-622 (2009).

62. Henrich, J. *et al*. Markets, Religion, Community Size, and the Evolution of Fairness and Punishment. *Science* **327**, 1480-1484 (2010).

63. Tanzi, V. & Davoodi, H. in *The Welfare State, Public Investment, and Growth: Selected Papers from the 53rd Congress of the International Institute of Public Finance*, eds Hirofumi Shibata & Toshihiro Ihori) Ch. 4, (Springer, 1997).

64. Becker, G. S. & Murphy, K. M. The Division of Labor, Coordination Costs, and Knowledge. *The Quarterly Journal of Economics* **107**, 1137-1160 (1992).

65. Black, D. Crime as Social Control. *Am. Sociol. Rev.* **48**, 34-45 (1983).

66. Lederman, D., Loayza, N. & Menendez, A. M. Violent Crime: Does Social Capital Matter? *Economic Development and Cultural Change* **50**, 509-539 (2002).

67. Egas, M. & Riedl, A. The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of the Royal Society B: Biological Sciences* **275**, 871-878 (2008).

68. Nikiforakis, N. & Normann, H. T. A comparative statics analysis of punishment in public-good

experiments. *Experimental Economics* **11**, 358-369 (2008).

69. Camerer, C. F. & Hogarth, R. M. The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework. *Journal of Risk and Uncertainty* **19**, 7-42 (1999).

70. Selten, R. Die strategiemethode zur erforschung des eingeschrankt rationalen verhaltens im rahmen eimes oligopolexperiments. In: Sauerman, H. (Ed.), Beitrage zur experimentellen Wirtschafts-forschung. J.C.B. Mohr (Paul Siebeck), Tubingen, 136168 (1967).

71. Falk, A., Fehr, E., Fischbacher, U. Driving Forces behind Informal Sanctions. *Econometrica* **73**, 2017-2030 (2005).

72. Orbell, J. M. & Dawes, R. M. Social Welfare, Cooperators' Advantage, and the Option of Not Playing the Game. *American Sociological Review* **58**, 787-800 (1993).

73. Nowak, M., Tarnita, C. & Antal, T. Evolutionary dynamics in structured populations. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 19-30 (2010).

74. Nowak, M. A. & Sigmund, K. Evolution of indirect reciprocity. *Nature* **437**, 1291-1298 (2005).

75. Henrich, J. & Boyd, R. Why People Punish Defectors: Weak Conformist Transmission can Stabilize Costly Enforcement of Norms in Cooperative Dilemmas. *Journal of theoretical biology* **208**, 79-89 (2001).

76. Dohmen, T. J., Falk, A., Huffman, D., Sunde, U., Schupp, J. & Wagner, G. G. Individual risk attitudes: measurement, determinants, and behavioral consequences. *Journal of the European Economic Association* **437**, 522-550 (2011).