Science AAAS

www.sciencemag.org/cgi/content/full/316/5833/1905/DC1

Supporting Online Material for

Via Freedom to Coercion: The Emergence of Costly Punishment

Christoph Hauert, Arne Traulsen, Hannelore Brandt, Martin A. Nowak, Karl Sigmund*

*To whom correspondence should be addressed. E-mail: karl.sigmund@univie.ac.at

Published 29 June, *Science* **316**, 1905 (2007)
DOI: 10.1126/science.xxxxxxxxxxx

This PDF file includes:

SOM Text
Figs. S1 to S5
References

Supporting online material: Via freedom to coercion: the emergence of costly punishment

Christoph Hauert [1], Arne Traulsen [1], Hannelore Brandt [2], Martin A. Nowak [1] & Karl Sigmund [3,4,*]

[1] Program for Evolutionary Dynamics, Department of Organismic and Evolutionary Biology, Department of Mathematics, Harvard University, One Brattle Square, Cambridge, MA 02138, USA

[2] Vienna University of Economics and Business Administration, A-1090 Vienna, Austria

[3] Faculty of Mathematics, University of Vienna, A-1090 Vienna, Austria

[4] International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria

April 20, 2007

Analysis

In order to obtain analytic expressions, we chose, among the plethora of possible updating rules governing the transmission of strategies, a specific stochastic process. There are many candidates. For instance, with 'imitate the better', two players are chosen at random and the one with the lower payoff adopts the strategy of the player with the higher payoff. With the 'proportional imitation' rule, the player with the lower payoff adopts the strategy of the more successful player, with a probability proportional to the payoff difference. For our analysis, we adopt what is known as the frequency dependent Moran process (*1*, *2*): each individual updates by imitating a player who is selected with a probability proportional to its 'fitness'. We define each players' fitness as $1 - s + sP$, the convex combination of the 'baseline fitness', which is normalised to 1 for all players, and the payoff P from the optional public goods game with punishment. The relative importance of the two components is determined by s, which is usually called selection strength. We shall assume that occasionally, each player can change strategy by imitating a player chosen with a probability proportional to that player's fitness. This mimics a learning process similar to the Moran process describing natural selection: more successful players are more likely to be copied. In addition, we shall assume that with a small probability μ, a player can switch to another strategy irrespective of its payoff (this 'mutation term' corresponds to blindly experimenting with anything different).

The analysis of the corresponding stochastic dynamics is greatly simplified in the limiting case $\mu \to 0$. The population consists almost always of one or two types at most. This holds because for $\mu = 0$ the four monomorphic states are absorbing, and for sufficiently small μ the fate of a mutant (i.e. its elimination or fixation) is settled before the next mutant appears. Thus the transitions between the four pure states - cooperators, defectors, non-participants and punishers - occur when a mutant appears and spreads to fixation (*3*).

In finite populations, the groups engaging in a public goods game are given by multivariate hypergeometric sampling. For transitions between two pure states, this reduces to sampling (without replacement) from a hypergeometric distribution. In a population of size M with m_i individuals of type i and $m_j = M - m_i$ of type j, the probability to select k individuals of type i and $N - k$ individuals of type j in N trials is

$$H(k, N, m_i, M) = \frac{\binom{m_i}{k}\binom{M-m_i}{N-k}}{\binom{M}{N}}. \tag{1}$$

Thus, in a population of x cooperators and $y = M - x$ defectors, the average payoffs to cooperators P_{xy} and defectors P_{yx} are

$$P_{xy} = \sum_{k=0}^{N-1} H(k, N-1, x-1, M-1)\left(\frac{k+1}{N}r - 1\right) = \frac{r}{N}\left(1 + (x-1)\frac{N-1}{M-1}\right) - 1$$

$$P_{yx} = \sum_{k=0}^{N-1} H(k, N-1, x, M-1)\left(\frac{k}{N}r\right) = \frac{r(N-1)}{N(M-1)}x.$$

Similarly, the payoffs P_{ij} of strategy type i competing against type j for the other possible pairings are

$$
\begin{aligned}
P_{zx} &= P_{zy} = \sigma \\
P_{xz} &= P_{wz} = r - 1 - \frac{\binom{z}{N-1}}{\binom{M-1}{N-1}}(r - 1 - \sigma) \\
P_{yz} &= \frac{\binom{z}{N-1}}{\binom{M-1}{N-1}}\sigma \\
P_{xw} &= P_{wx} = r - 1 \\
P_{yw} &= \frac{(N-1)(r - N\beta)}{N(M-1)}w \\
P_{wy} &= \frac{r}{N} - 1 - \gamma(N-1) + \frac{(N-1)(r + N\gamma)}{N(M-1)}(w-1) \\
P_{zw} &= \sigma - \frac{N-1}{M-1}\delta\beta w \\
P_{wz} &= r - 1 - \frac{\binom{z}{N-1}}{\binom{M-1}{N-1}}(r - 1 - \sigma) - \frac{N-1}{M-1}\delta\gamma z
\end{aligned}
$$

The fitness of an individual of type i in a mixed population of types i and j is then given by $1 - s + sP_{ij}$. Since the fitness has to be positive, there is an upper limit on the intensity of selection s given by $s_{\max} = 1/(1 - \min P_{ij})$ for all strategic types i, j under consideration. The above payoffs together with s determine the probability to change the number of individuals m_i of type i by ± 1, $T_{ij}^{\pm}$:

$$
T_{ij}^{+} = \frac{m_i(1 - s + sP_{ij})}{M(1-s) + s(m_i P_{ij} + (M - m_i)P_{ji})}\frac{M - m_i}{M} \tag{2a}
$$

$$
T_{ij}^{-} = \frac{(M - m_i)(1 - s + sP_{ji})}{M(1-s) + s(m_i P_{ij} + (M - m_i)P_{ji})}\frac{m_i}{M} \tag{2b}
$$

From these transition probabilities, the fixation probability ρ_{ij} of a single mutant strategy of type i in a resident population of type j can be derived (*2*, *4*):

$$
\rho_{ij} = \frac{1}{\sum_{k=0}^{M-1}\prod_{m_i=1}^{k}\frac{T_{ij}^{-}}{T_{ij}^{+}}} = \frac{1}{\sum_{k=0}^{M-1}\prod_{m_i=1}^{k}\frac{1 - s + sP_{ji}}{1 - s + sP_{ij}}} \tag{3}
$$

Finally, the fixation probabilities ρ_{ij} define the transition probabilities of a Markov process between the four different homogeneous states of the population. The transition matrix $\mathbf{A}$ is given by:

$$\mathbf{A} = \begin{pmatrix} 1-\rho_{yx}-\rho_{zx}-\rho_{wx} & \rho_{xy} & \rho_{xz} & \rho_{xw} \\ \rho_{yx} & 1-\rho_{xy}-\rho_{zy}-\rho_{wy} & \rho_{yz} & \rho_{yw} \\ \rho_{zx} & \rho_{zy} & 1-\rho_{xz}-\rho_{yz}-\rho_{wz} & \rho_{zw} \\ \rho_{wx} & \rho_{wy} & \rho_{wz} & 1-\rho_{xw}-\rho_{yw}-\rho_{zw} \end{pmatrix} \tag{4}$$

The normalized right eigenvector to the largest eigenvalue (which is 1 for the matrix $\mathbf{A}$) determines the stationary distribution, i.e. indicates the probability to find the system in one of the four homogeneous states. It is given by

$$\phi = \frac{1}{\mathcal{N}} \left[\begin{matrix} \begin{pmatrix} \rho_{wy}\rho_{wz}\rho_{xw} + \rho_{wz}\rho_{xy}\rho_{xw} + \rho_{wy}\rho_{xz}\rho_{xw} + \rho_{xy}\rho_{xz}\rho_{xw} + \rho_{wy}\rho_{yz}\rho_{xw} + \rho_{xy}\rho_{yz}\rho_{xw} \\ +\rho_{wz}\rho_{zy}\rho_{xw} + \rho_{xz}\rho_{zy}\rho_{xw} + \rho_{wz}\rho_{xy}\rho_{yw} + \rho_{xy}\rho_{xz}\rho_{yw} + \rho_{xy}\rho_{yw}\rho_{yz} \\ +\rho_{wy}\rho_{xz}\rho_{zw} + \rho_{xy}\rho_{xz}\rho_{zw} + \rho_{xy}\rho_{yz}\rho_{zw} + \rho_{xz}\rho_{yw}\rho_{zy} + \rho_{xz}\rho_{zw}\rho_{zy} \end{pmatrix} \\ \begin{pmatrix} \rho_{wx}\rho_{wz}\rho_{yw} + \rho_{wx}\rho_{xz}\rho_{yw} + \rho_{wz}\rho_{yx}\rho_{yw} + \rho_{xz}\rho_{yx}\rho_{yw} + \rho_{wx}\rho_{yz}\rho_{yw} + \rho_{yx}\rho_{yz}\rho_{yw} \\ +\rho_{wz}\rho_{zx}\rho_{yw} + \rho_{yz}\rho_{zx}\rho_{yw} + \rho_{wz}\rho_{xw}\rho_{yx} + \rho_{xw}\rho_{xz}\rho_{yx} + \rho_{xw}\rho_{yx}\rho_{yz} \\ +\rho_{xz}\rho_{yx}\rho_{zw} + \rho_{wx}\rho_{yz}\rho_{zw} + \rho_{yx}\rho_{yz}\rho_{zw} + \rho_{xw}\rho_{yz}\rho_{zx} + \rho_{yz}\rho_{zw}\rho_{zx} \end{pmatrix} \\ \begin{pmatrix} \rho_{wx}\rho_{wy}\rho_{zw} + \rho_{wx}\rho_{xy}\rho_{zw} + \rho_{wy}\rho_{yx}\rho_{zw} + \rho_{wy}\rho_{zx}\rho_{zw} + \rho_{xy}\rho_{zx}\rho_{zw} + \rho_{wx}\rho_{zy}\rho_{zw} \\ +\rho_{yx}\rho_{zy}\rho_{zw} + \rho_{zx}\rho_{zy}\rho_{zw} + \rho_{wy}\rho_{xw}\rho_{zx} + \rho_{xw}\rho_{xy}\rho_{zx} + \rho_{xy}\rho_{yw}\rho_{zx} \\ +\rho_{wx}\rho_{yw}\rho_{zy} + \rho_{xw}\rho_{yx}\rho_{zy} + \rho_{yw}\rho_{yx}\rho_{zy} + \rho_{xw}\rho_{zx}\rho_{zy} + \rho_{yw}\rho_{zx}\rho_{zy} \end{pmatrix} \\ \begin{pmatrix} \rho_{wx}\rho_{wy}\rho_{wz} + \rho_{wx}\rho_{xy}\rho_{wz} + \rho_{wy}\rho_{yx}\rho_{wz} + \rho_{wy}\rho_{zx}\rho_{wz} + \rho_{xy}\rho_{zx}\rho_{wz} + \rho_{wx}\rho_{zy}\rho_{wz} \\ +\rho_{yx}\rho_{zy}\rho_{wz} + \rho_{zx}\rho_{zy}\rho_{wz} + \rho_{wx}\rho_{wy}\rho_{xz} + \rho_{wx}\rho_{xy}\rho_{xz} + \rho_{wy}\rho_{xz}\rho_{yx} \\ +\rho_{wx}\rho_{wy}\rho_{yz} + \rho_{wx}\rho_{xy}\rho_{yz} + \rho_{wy}\rho_{yx}\rho_{yz} + \rho_{wy}\rho_{yz}\rho_{zx} + \rho_{wx}\rho_{xz}\rho_{zy} \end{pmatrix} \end{matrix} \right] \tag{5}$$

where the normalisation factor $\mathcal{N}$ has to be chosen such that the elements of ϕ sum up to one. S 1 and S 2 display the stationary distribution as a function of the selection strength s.

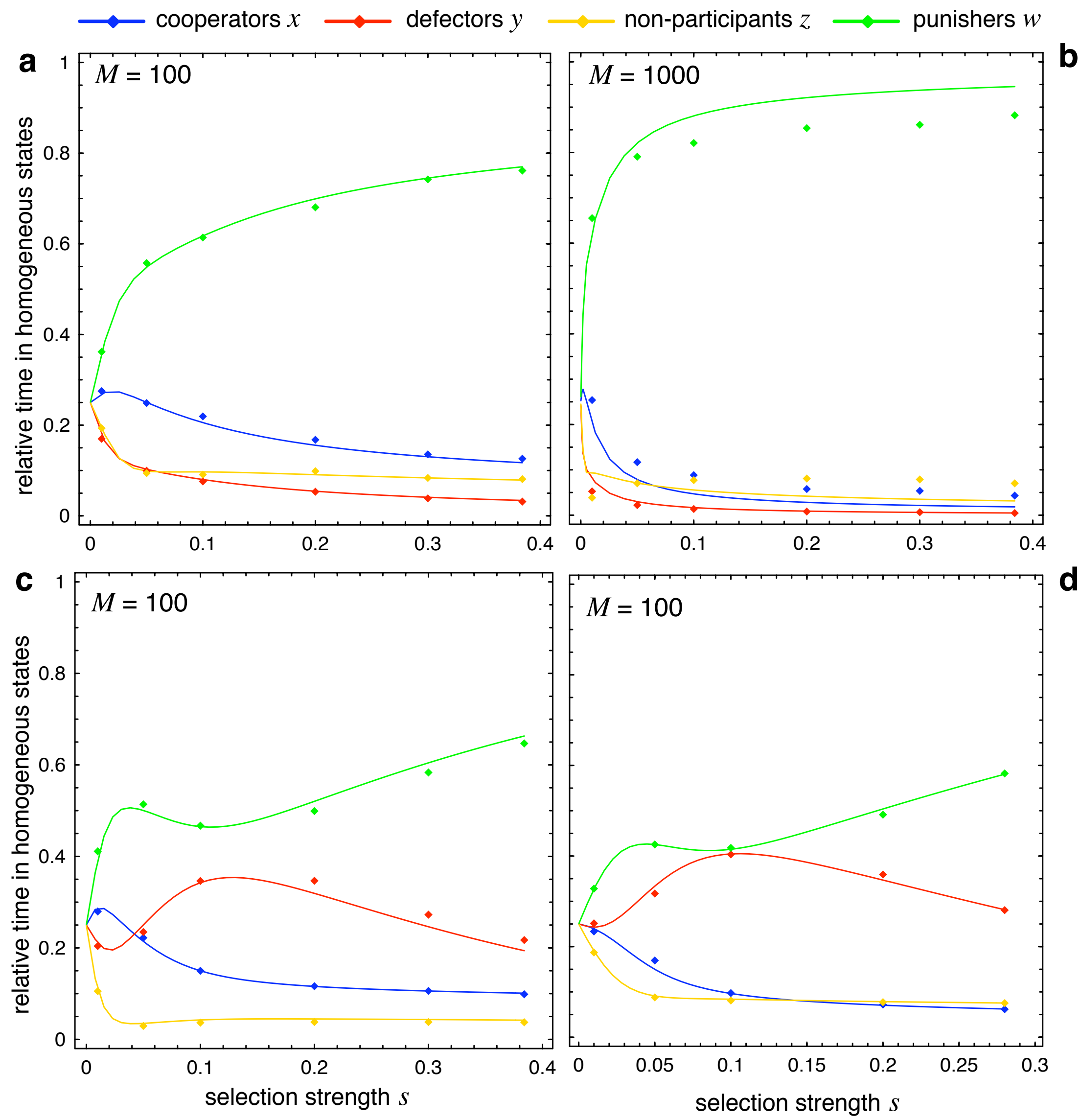

Figure 1: Punishment and abstaining in public goods games. For rare mutations, the dynamics is restricted to transitions between the four homogeneous states consisting entirely of cooperators (blue), defectors (red), non-participants (yellow) and punishers (green). All panels depict the probabilities of each state as a function of the selection strength $s < s_{\max} = 0.384$ for $N = 5, r = 3, \sigma = 1, \gamma = 0.3, \beta = 1$. In the limit of neutral evolution ($s = 0$), all states become equally likely. **a** Punishers are clearly dominant in voluntary joint enterprises with punishment. **b** The success of punishers is even more pronounced for larger population sizes. *(continued)*

Figure 1 *(continued)*: Individual-based simulation data confirms the analytical results for small mutation rates (colored dots, $\mu = 0.001$ **a** $M = 100$, sampling time $T = 10^7$, **b** $M = 1000, T = 10^6$). Whenever $> 90\%$ of the population opt for one strategy it is counted as being in the respective homogeneous state. The payoff determination, the mutation rate and the 90% threshold are responsible for the systematic deviations but also illustrate the robustness of the model. **c** Lowering the payoff of non-participants to $\sigma = 0.1$ (one tenth of σ in **a**,**b**) obviously reduces the risk of the public goods game, i.e. encourages participation and hence supports defectors. Nevertheless, punishers reign unchallenged. **d** Additionally, lowering the return of the joint effort to $r = 1.8$ impedes both cooperators and punishers but punishers win (somewhat surprisingly since in infinite populations, non-participants win and joint enterprises are abandoned for $r < 2$).

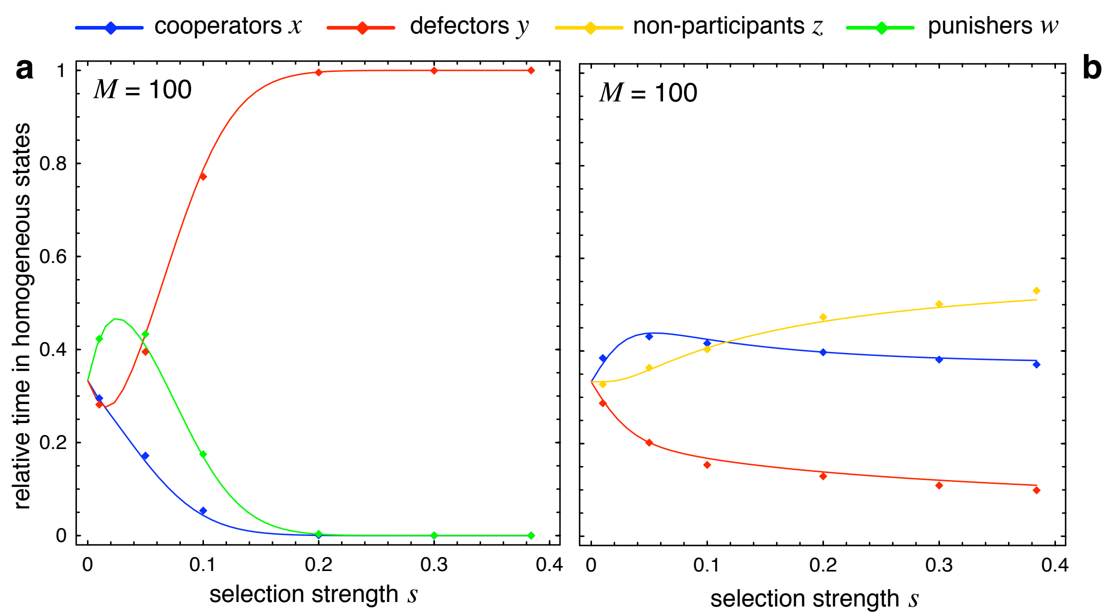

Figure 2: The role of non-participants and punishers in public goods games (all parameters are identical to S 1a). **a** Despite punishment, defectors reign in compulsory games except for very weak selection. **b** In the absence of punishers, no strategy clearly dominates due to the systems' tendency to cycle between cooperators, defectors and non-participants. However, the system spends significantly more time in the states with all cooperators or non-participants than in the defector state.

Fixation time of punishers

In the limit of rare mutations, the average time to reach the punisher state for the first time can be derived analytically. This fixation time τ_i when starting in a pure state of i (which can be x, y or z for cooperators, defectors, and non-participants, respectively) satisfies

$$\tau_i = 1 + \sum_{j=x,y,z} \tau_j \cdot R_{ji}, \tag{6}$$

where the time τ_i is measured in updating periods (which consist of M individual update steps) and $R_{ji} = \delta_{ji} + \nu_j (A_{ji} - \delta_{ji})$ where A_{ji} denotes the transition probability from pure state i to pure state j (see Eq. 4), δ_{ji} denotes the Kronecker symbol and ν_j denotes the rate at which mutants of type j are produced. If all mutants are equally likely, this rate is simply $\nu_j = \mu M/3$ (on average there are μM mutations per generation). Solving for τ_j leads to

$$\begin{pmatrix} \tau_x \\ \tau_y \\ \tau_z \end{pmatrix} = \frac{3}{\mu M} \begin{pmatrix} \rho_{yx} + \rho_{zx} + \rho_{wx} & -\rho_{yx} & -\rho_{zx} \\ -\rho_{xy} & \rho_{xy} + \rho_{zy} + \rho_{wy} & -\rho_{zy} \\ -\rho_{xz} & -\rho_{yz} & \rho_{xz} + \rho_{yz} + \rho_{wz} \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{7}$$

Thus, in the limit of rare mutations, the average waiting time to reach the punishment state scales inversely with the mutation rate μ. According to numerical simulations, this relation still holds for larger μ as shown in S 3. However, for μ of the order $1/M$ or larger, pure states and their close vicinity may no longer be accessible.

Note, that the fixation times follow an exponential distribution. Hence the average fixation time equals its standard deviation and thus has limited predictive power for specific realizations. In the limit of neutral selection ($s = 0$) the fixation time of punishers reduces to $\tau_x = \tau_y = \tau_z = 3/\mu$.

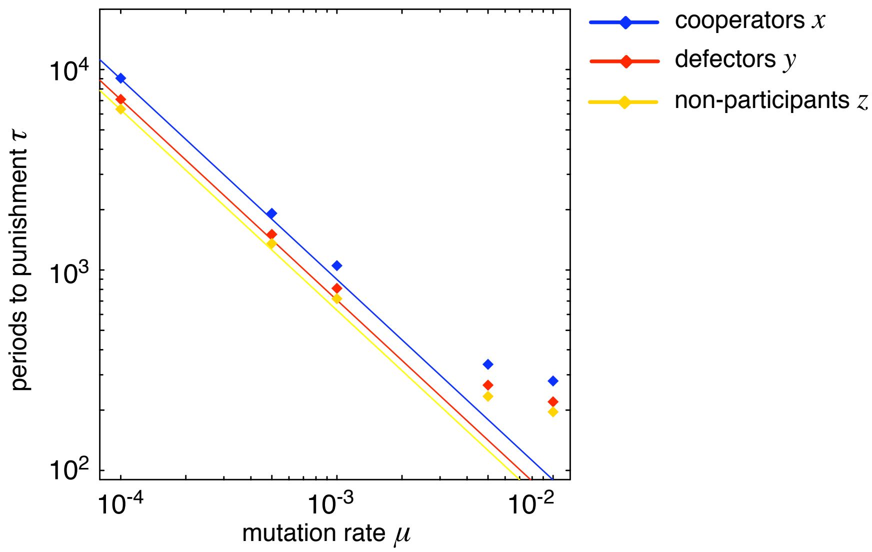

Figure 3: Average waiting time τ to reach the punisher state when starting with all cooperators (blue), defectors (red) or non-participants (yellow) as a function of the mutation rate μ. The lines depict the analytical solution (see Eq. 10) for maximal selection strength. The symbols indicate simulation results for different mutation rates μ. In order to determine whether the punisher state has been reached, a threshold of 90% was used. Parameters: $N = 5, r = 3, \sigma = 1, \gamma = 0.3, \beta = 1, \alpha = 0, M = 100, s = 0.384$.

Enforcing cooperation

Punishers can attempt to enforce cooperation through different means. In order to avoid second order free-riding, punishers can punish cooperators who have failed to punish defectors ($\alpha > 0$) or they can enforce participation in the joint effort game by punishing the non-participants ($\delta > 0$). The very small effects of second order punishment ($\alpha > 0$) are illustrated in S 4. Note that α does not enter the analytical approximation for the limit of vanishing μ. For larger values of μ, it barely affects the simulation results.

For $\delta > 0$ a new unstable equilibrium point appears along the non-participant–punisher edge. This results in bi-stability between the two states, just as between the defector and the punisher state. However, for small δ and/or weak selection, punishers can still relatively easily invade non-participant populations and reach the critical threshold through random drift. In this case the general conclusions of the main text remain unaffected (S 5).

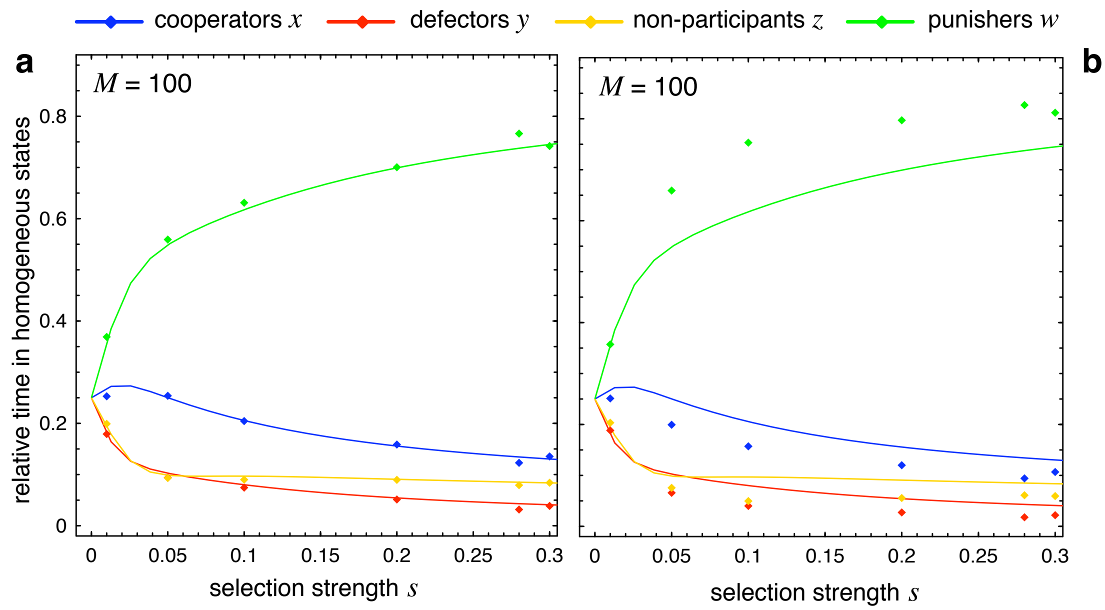

Figure 4: Effects of second order free-riding ($\alpha > 0$) in simulations as compared to the analytical results that are independent of α. (**a**) For mild punishment of non-punishing cooperators, $\alpha = 0.1$, the result is essentially indistinguishable from the case $\alpha = 0$ (c.f. S 1a). (**b**) Equal punishment of cooperators and defectors, $\alpha = 1$, strengthens the position of punishers and the system spends even more time in the punisher state. However, punishers can invade less easily. Parameters: $N = 5, r = 3, \sigma = 1, \gamma = 0.3, \beta = 1, \delta = 0, M = 100, \mu = 0.01, s_{\text{max}} = 0.384$; **a** $\alpha = 0.1$; **b** $\alpha = 1$.

For small δ and strong selection, the systems generally spends $> 60\%$ in the punisher state and the defector state is suppressed to levels $< 10\%$ (S 5a). Similarly, punishers dominate for large δ and weak selection (S 5b). However, for large δ and strong selection the bi-stability essentially prevents transitions from the non-participant state to the punisher state or vice versa. Thus, punishment can only be established by invading cooperators through random drift. As a consequence, the system spends more time in the cooperator state and less in the punisher state. Quite intriguingly, punishment of non-participants does not diminish their success but, in fact, actually increases the time in the non-participant state. Nevertheless, even for large δ and strong selection punishers dominate with 40% and the systems spends roughly equal times cooperator and non-participant state with around 25% each, leaving merely 10% for defectors. Interestingly, such strict social coercion maintains close to 60% cooperation in total (cooperators

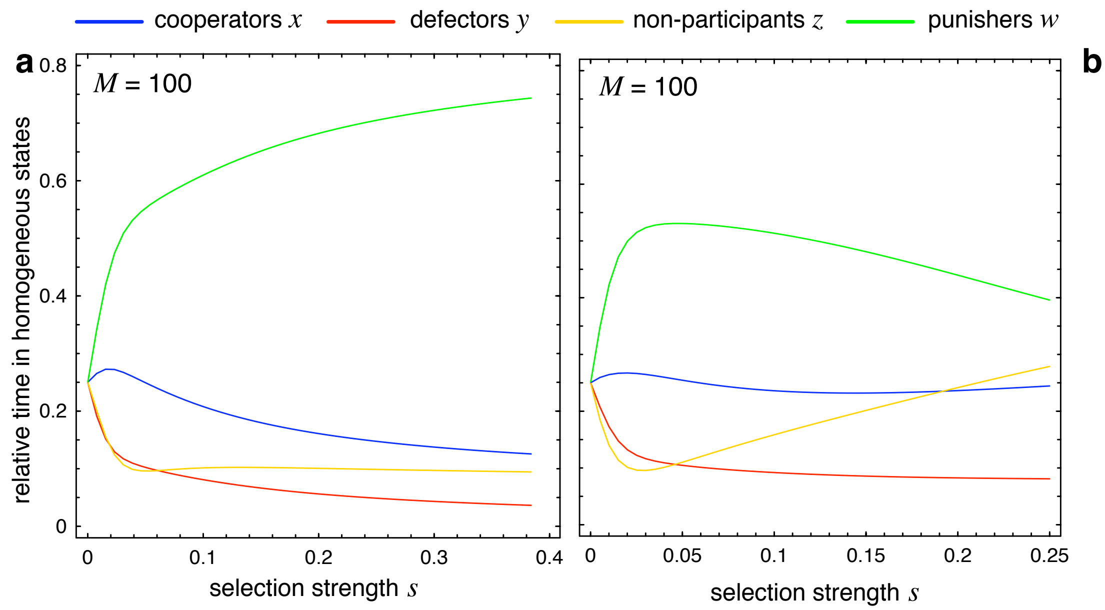

Figure 5: Effects of social coercion by punishing non-participants. (**a**) For mild measures against non-participants, the outcome is barely affected: punishers dominate (c.f. S 1a). (**b**) In contrast, heavy measures against non-participants tend to undermine the success of punishment and actually result in an increase of cooperators and non-participants while leaving the frequency of defectors largely unaffected. Parameters: $N = 5, r = 3, \sigma = 1, \gamma = 1, \beta = 2, \alpha = 0, M = 100$; **a** $\delta = 0.1, s_{\text{max}} = 0.151$; **b** $\delta = 1, s_{\text{max}} = 0.125$.

plus punishers), whereas the scenario with purely voluntary participation ($\delta = 0$) results in more than 80% cooperation (c.f. S 1a). Hence the system responds to coercion by actually reducing the readiness to cooperate.

References

[1] P. A. P. Moran, *The Statistical Processes of Evolutionary Theory* (Clarendon, Oxford, UK, 1962).

[2] M. A. Nowak, A. Sasaki, C. Taylor, D. Fudenberg, *Nature* **428**, 646 (2004).

[3] L. A. Imhof, D. Fudenberg, M. A. Nowak, *Proc. Natl. Acad. Sci. USA* **102**, 10797 (2005).

[4] S. Karlin, H. M. Taylor, *First Course in Stochastic Processes* (Academic Press, London, 1975), second edn.