Journal of Theoretical Biology 321 (2013) 78–82

Contents lists available at SciVerse ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Does insurance against punishment undermine cooperation in the evolution of public goods games?

Jianlei Zhang [a,b], Tianguang Chu [a,*], Franz J. Weissing [b]

[a] State Key Laboratory for Turbulence and Complex Systems, College of Engineering, Peking University, Beijing 100871, China
[b] Theoretical Biology Group, Centre for Ecological and Evolutionary Studies, University of Groningen, The Netherlands

HIGHLIGHTS

- We consider a public goods game with punishment and insurance against punishment.
- We investigate whether the insurance destabilize the cooperation.
- In principle, cooperation can break down due to insurance.
- For realistic, assumptions on the cost of insurance, cooperation remain stable.

ARTICLE INFO

Article history:
Received 27 June 2012
Received in revised form
8 December 2012
Accepted 21 December 2012
Available online 2 January 2013

Keywords:
Game theory
Cooperation
Punishment
Speculation

ABSTRACT

In a public goods game, cooperation can be a stable outcome if defectors are facing efficient punishment. In some public goods systems, punishment is undermined by an insurance system where speculators buy a policy that sequentially covers all punishment costs. Here, we study a simple model to investigate the question whether stable cooperation can break down in the presence of such speculation. We do indeed find scenarios where speculation either leads to the reduction of the basin of attraction of the cooperative equilibrium or even the loss of stability of this equilibrium. This however only happens if the costs of the insurance are lower than the expected fines faced by a defector. We argue that an insurance of this type is not viable and conclude that under realistic assumptions speculation does not destabilize cooperation.

1. Introduction

Even if cooperation is beneficial for all members of a community, it is often not easy to achieve since individuals have an incentive to reap the benefits of cooperation by others without contributing to the costs of cooperation. The prototype example for this 'social dilemma' involved in cooperation is the public goods game (Brandt et al., 2006; Fehr and Gächter, 2002; Hardin, 1968; Heckathorn, 1996) where each player has two strategic options: cooperation and defection. Cooperators invest into the public good, and their investment is multiplied by a factor $r > 1$. Hence such investment increases the welfare of the group. The problem is that the revenue of the investment is distributed over all group members, irrespective of whether they contributed to the public good or not. Accordingly, an individual's return on its investment c is rc/N, which is smaller than the investment if $r < N$. As a consequence, it is individually rational not to contribute to the public good in this case, although the cooperative outcome would be much preferred by everyone than the non-cooperative outcome. Why then is cooperation that widespread in human society?

A variety of solutions for this dilemma has been discussed in the literature (for reviews see Nowak, 2006; West et al., 2007). Cooperation can, for example, be stabilized if interactions among players are not random and if cooperators tend to be associated with cooperators while defectors tend to be associated with defectors (Nowak and May, 1992; Nowak and Sigmund, 1992; Perc and Szolnoki, 2010; Szabó and Fáth, 2007). In biology, kin structure (Reece et al., 2003) or local dispersal and differential productivity related to cooperation (Xavier et al., 2009) can lead to such assortment. In humans, cooperation can be stabilized if individuals tend to stay in cooperative neighborhoods (Fu et al., 2008) or tend to leave non-cooperative neighborhoods (Pacheco et al., 2006; Zhang et al., 2012). But even in a well-mixed community cooperation can be a stable outcome if a social dilemma is embedded in a richer game theoretical structure (Maynard Smith, 1982; McNamara and Weissing, 2010). For

* Corresponding author. Tel.: +86 010 62752737.
E-mail addresses: jianleizhang@pku.edu.cn (J. Zhang), chutg@pku.edu.cn (T. Chu), f.j.weissing@rug.nl (F.J. Weissing).

0022-5193/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jtbi.2012.12.019

example, reciprocity in repeated interactions (Boyd and Lorberbaum, 1987; Pacheco et al., 2008) or cooperation based on the reputation on one's interaction partners (Fu et al., 2008; Rand et al., 2009; Szolnoki and Perc, 2010) can transform the short-term benefits of defection into long-term costs, thereby stabilizing cooperation.

Cooperation can also be stabilized by punishment. In theory and in experiments, punishment has turned out to be a simple but effective mechanism to prevent cheating even under full anonymity (Van den Berg and Weissing, 2012; Piazza and Bering, 2008). There is now a rich literature on whether and how various forms of punishment are effective in bringing about cooperation (Rand et al., 2009; Szolnoki and Perc, 2010), peer punishment (Boyd et al., 2003; Fowler, 2005; Hauert et al., 2007; Nakamaru and Iwasa, 2006), pool punishment (Perc, 2012; Sigmund et al., 2010; Szolnoki et al., 2011), and anti-social punishment (Rand and Nowak, 2011). Although punishment can be effective in stabilizing cooperation, the emergence of punishment is not self-evident. In fact, costly punishment may be viewed as a public good itself. As a consequence, punishment may be destabilized by a second-order free rider incentives (Van den Berg and Weissing, 2012; Perc, 2012). Here we are less interested in the establishment of an effective system of punishment, but in the possible destabilization of such a system once it is present. We consider a system where defectors are caught with a certain probability α and where a fine β is imposed on them when caught. Hence the expected costs of defection are $\alpha\beta$, which are imposed upon defectors by an external agency. We investigate the question whether insurance against such punishment, an additional option found in several systems, may lead to the collapse of punishment-stabilized cooperation.

A good example for what we have in mind is reported for the Paris Metro system. A group of Parisians has started an insurance funds to pay the fines of Paris Metro fare dodgers who are caught by ticket controllers who periodically patrol the metro cars. The members of such an insurance group pay some money per month for the insurance, which in turn will pay their fines when being caught without a metro ticket. The question is whether such an insurance system can threaten the persistence of a public good like a well-running Metro system. To address this question, we consider a public goods game with a external-agency punishment system indicated above. We extend this game by adding a third strategy, called speculation, which corresponds to buying an insurance policy covering the costs of punishment when caught defecting. In another words, by paying a fixed cost for their insurance policy, speculators can defect without paying any fine.

The outcome of this extended public goods game will of course depend on the costs of the insurance policy. We make the reasonable assumption that these costs decrease with the number of participants in the insurance system. Based on this assumption, we consider various scenarios for which we answer the question whether and to what extent speculation can lead to the break down of cooperation.

2. Model

Here, we consider the following model. Each player receives a certain benefit bx_c which is proportional to the relative frequency of cooperative individuals x_c. The costs associated with behavior differ between strategies. Cooperators have a fixed cost c corresponding to the investment in the public goods game. Defectors do not have this cost, but are confronted with punishment when caught. Their expected fine is $\alpha\beta$, which reflects the α of being detected and the fine β in cost of detective. Speculator also neither have to pay the cost investment nor a fine when caught, instead they have to pay an amount λ corresponding to the insurance policy. Notably, it is plausible that the insurance is a profit management, thus the value of λ is not a constant number. For example, when the insurance company has more premium, it can use the money for investing and the earns can make the company charge much less from its clients. Therefore, we can make a safe assumption that λ is a decreasing function of x_s. Here x_s is the fraction of speculators in a population.

These assumptions result in the payoffs for cooperation, defection and speculation respectively:

$$\begin{cases} P_c = bx_c - c \\ P_d = bx_c - \alpha\beta \\ P_s = bx_c - \lambda(x_s) \end{cases} \tag{1}$$

This payoff structure corresponds to the standard public goods game in a infinite population (Fehr and Gächter, 2002). In this game, each unit of investment is multiplied by a factor r and the product is distributed among all players irrespective of their strategies. Accordingly, here, b corresponds to rc.

However, in many situations, the number of players is finite. We consider a finite model as follows. From time to time, N players are chosen randomly from a large population of players. Within such a group, if N_c denotes the number of cooperators among the public goods players the net payoffs of the three strategies are respectively given by

$$\begin{cases} P_c = \dfrac{rN_c c}{N} - c \\ P_d = \dfrac{rN_c c}{N} - \alpha\beta \\ P_s = \dfrac{rN_c c}{N} - \lambda(x_s) \end{cases} \tag{2}$$

where r denotes the amplification effect on the common pool.

The chance that a given player occurs in a group of which m other individuals cooperate is

$$\binom{N-1}{m} x_c^m (1-x_c)^{m-1} \tag{3}$$

Here, x_c is the fraction of the cooperators in the whole population. Defectors (playing D) that withhold their share and, therefore, exploit other players, take the risk of an imposed fine of $-\beta$ ($\beta > 0$) with a probability of α ($0 < \alpha < 1$). Hence, the expected payoff for a defector in such a group is

$$\begin{aligned} P_d &= \sum_{m=0}^{N-1} \frac{rmc}{N} \binom{N-1}{m} x_c^m (1-x_c)^{m-1} \\ &= \frac{rcx_c(N-1)}{N} - \alpha\beta \end{aligned} \tag{4}$$

The payoff of a cooperator is as follows:

$$\begin{aligned} P_c &= \sum_{m=0}^{N-1} \left[\frac{rc(m+1)}{N} - 1\right] \binom{N-1}{m} x_c^m (1-x_c)^{m-1} \\ &= \frac{rcx_c(N-1)}{N} + \frac{rc}{N} - c \end{aligned} \tag{5}$$

The payoff of a speculator is as follows:

$$\begin{aligned} P_s &= \sum_{m=0}^{N-1} \left(\frac{rcm}{N} - \lambda\right) \binom{N-1}{m} x_c^m (1-x_c)^{m-1} \\ &= \frac{rcx_c(N-1)}{N} - \lambda(x_s) \end{aligned} \tag{6}$$

In the above model, $N > 1$, $1 < r < N$, $0 < \alpha < 1$ and $\beta > 0$, using x_c as the proportion of cooperators in the investigated population.

We can recover the payoffs obtained from Eq. (1) by defining $\tilde{b} = rc(N-1)/N$ and $\tilde{c} = c(1-r/N)$. We get the payoffs for cooperation, defection and speculation respectively:

$$\begin{cases} P_c = \tilde{b}x_c - \tilde{c} \\ P_d = \tilde{b}x_c - \alpha\beta \\ P_s = \tilde{b}x_c - \lambda(x_s) \end{cases} \tag{7}$$

For $1 < r < N$ (if all cooperate, they are better off than if all defect), the public goods game is a social dilemma and zero investment is the individually optimal strategy. The first term in the expression represents the benefit that the agent obtains from the public goods, while the second term denotes cost. For cooperators, the cost is $\tilde{c}$, instead of the investment to the public goods c, and for speculators, the cost is the payment to the insurance. Selfish individuals will therefore always avoid the cost of altruism, i.e. a collective of selfish players will never cooperate.

3. Results

We are interested in how the game dynamics are affected by the cost of the insurance policy λ and the depends of λ on the frequency of policy holders x_s. In total, we consider 12 insurance policies, which are discussed in scenario C_1 till C_6 in Fig. 1 and D_1 till D_6 in Fig. 2. Fig. 1 focuses on the situation $\alpha\beta > \tilde{c}$ implying that the fine for defectors is higher than the costs of cooperation. Hence, punishment is efficient in the sense that in the absence of speculation cooperation is the dominant strategy. The question is whether cooperation will break down in the presence of speculation. Fig. 2 considers the opposite case $\alpha\beta < \tilde{c}$, where defection is the dominating strategy. Here the question is whether speculation can get off the ground and whether it might even be favorable for cooperation.

As illustrated in Fig. 1, the game dynamics takes on six qualitatively different scenarios, which will be discussed one by one.

Scenario C_1 ($\lambda(0) > \lambda(1) > \alpha\beta$): In this case, for all x_s, speculation (S) and defection are both unstable equilibria of the game dynamics. The cooperation equilibrium (C) is stable and in fact a global attractor.

Scenario C_2 ($\lambda(0) > \alpha\beta > \lambda(1) > \tilde{c}$): In this case, $\lambda(x_s)$ will decline below $\alpha\beta$ beyond a certain frequency of speculators x_s. There is an additional border equilibrium consisting of speculation and defection. But this equilibrium is unstable. As before, the cooperative equilibrium is a global attractor.

Scenario C_3 ($\alpha\beta > \lambda(0) > \lambda(1) > \tilde{c}$): In comparison with scenario C_1, defection is now a repellor while speculation is now a saddle point. Still, cooperation is the only stable equilibrium and in fact a global attractor.

Scenario C_4 ($\alpha\beta > \lambda(0) > \tilde{c} > \lambda(1)$): Now $\lambda(x_s)$ will decline below $\tilde{c}$ beyond a certain value of x_s. This creates a border equilibrium consisting of speculation and cooperation. This equilibrium is unstable. It separates the two stable equilibria corresponding to pure cooperation (C) and pure speculation (S). Hence, the outcome of the game dynamics depends on initial conditions.

Scenario C_5 ($\lambda(0) > \alpha\beta > \tilde{c} > \lambda(1)$): The game dynamics is very much the same as in scenario C_4. However, there is now an additional unstable equilibrium consisting of speculation and defection.

Scenario C_6 ($\tilde{c} > \lambda(0) > \lambda(1)$): In this case, pure cooperation (C) is no longer stable neither globally nor locally. Pure speculation (S) is a global attractor.

Summarizing the six scenarios corresponding to $\alpha\beta > \tilde{c}$, we can conclude that pure cooperation is still a global attractor if $\lambda(1) > \tilde{c}$ (scenarios C_1, C_2, and C_3) while there is no longer the case for $\lambda(1) < \tilde{c}$. In the latter case, cooperation is still locally stable if $\lambda(0) > \tilde{c}$ (scenarios C_4 and C_5). Cooperation is fully destabilized if $\lambda(0) < \tilde{c}$. In this case, pure speculation is a global attractor (scenario C_6).

Similar to the above, we also obtain six scenarios for the case $\tilde{c} > \alpha\beta$ (illustrated in Fig. 2), where punishment is not sufficiently frequent or severe to deter defection.

Scenario D_1, ($\lambda(0) > \lambda(1) > \tilde{c}$): For all x_s, cooperation dominates speculation while defection dominates both cooperation and speculation. Pure defection (D) is a global attractor.

Scenario D_2 ($\lambda(0) > \tilde{c} > \lambda(1) > \alpha\beta$): Pure defection is still a global attractor, but now there is an additional unstable equilibrium consisting of cooperation and speculation.

Scenario D_3 ($\tilde{c} > \lambda(0) > \lambda(1) > \alpha\beta$): In comparison with scenario D_1, cooperation is now a repellor while speculation has turned into a saddle point. Still, defection remains stable and a global attractor.

Scenario D_4 ($\tilde{c} > \lambda(0) > \alpha\beta > \lambda(1)$): Now an unstable border equilibrium appears that consisting of speculation and defection. This equilibrium separates the two stable equilibria consisting of pure speculation (S) and pure defection (D). The outcome of the game dynamics depends on the initial conditions.

Scenario D_5 ($\lambda(0) > \tilde{c} > \alpha\beta > \lambda(1)$): Similar to scenario D_4, but with an additional unstable border equilibrium consisting of speculation and cooperation.

Scenario D_6 ($\alpha\beta > \lambda(0) > \lambda(1)$): Now pure defection (D) is no longer stable. Pure speculation (S) is a global attractor.

Summarizing the six scenarios corresponding to $\tilde{c} > \alpha\beta$, we can conclude that pure defection is still a global attractor if $\lambda(1) > \alpha\beta$ (scenarios D_1, D_2, and D_3). When $\lambda(1) < \alpha\beta$, defection is still locally

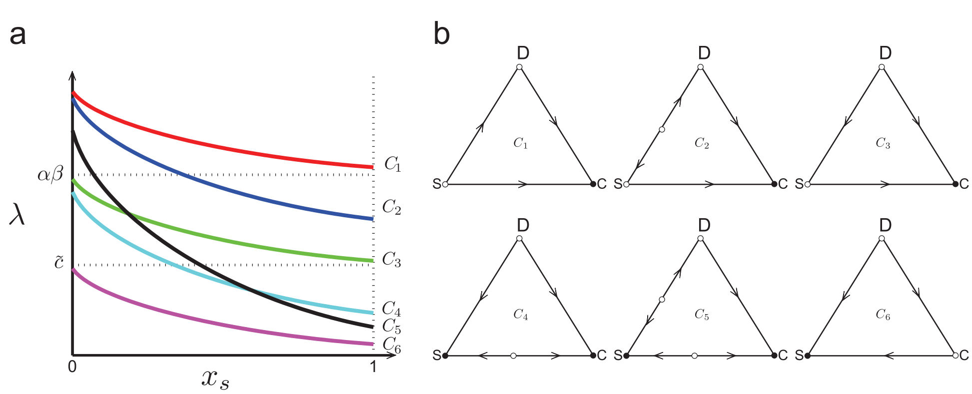

Fig. 1. Effects of insurance against punishment in case $\alpha\beta > \tilde{c}$, where in the absence of speculation, cooperation dominates defection. In the presence of speculation, there are six scenarios for the game dynamics that are characterized by the magnitude of the costs of the insurance policy λ and the steepness of the decline of λ with the frequency of speculation x_s. (a) Example functions $\lambda(x_s)$ corresponding to the six scenarios C_1–C_6 and (b) resulting game dynamics in the six scenarios. Due to the absence of an interior equilibrium point, the dynamics is fully characterized by the flow on the boundary of the strategy simplex. The corners C, D, and S are equilibrium points. Open dots are unstable equilibrium points and closed dots are stable equilibrium points.

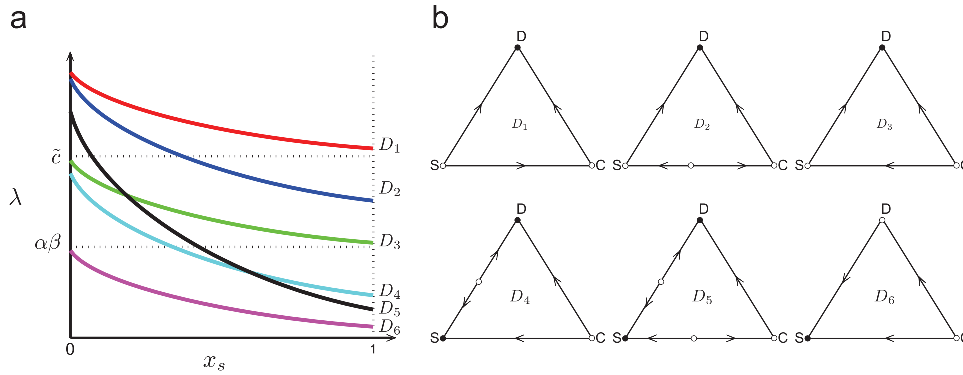

Fig. 2. Effects of insurance against punishment in case $\alpha\beta < \tilde{c}$, where in the absence of speculation, defection dominates cooperation. In the presence of speculation, there are six scenarios for the game dynamics that are characterized by the magnitude of the costs of the insurance policy λ and the steepness of the decline of λ with the frequency of speculation x_s. (a) Example functions $\lambda(x_s)$ corresponding to the six scenarios D_1–D_6 and (b) resulting game dynamics in the six scenarios. Open dots are unstable equilibrium points and closed dots are stable equilibrium points.

stable if $\lambda(0) > \alpha\beta$ (scenarios D_4 and D_5). In case of $\lambda(0) < \alpha\beta$, pure speculation is a global attractor (scenario D_6).

4. Conclusions

Although punishment can lead to the establishment of stable cooperation, it is not self-evident that a system of punishment can get established in the first place. Even if an effective system of punishment is in place, it can often be destabilized by additional strategic options. Here we considered one such situation. For simplicity we assumed that punishment of a given effectiveness ($\alpha\beta$) is externally imposed upon the defectors in a public goods game. We do not consider the question how the punishment system was established or who carries the costs of punishment. If all members of the communities have to pay the costs for the punishment institution (e.g. if this institution is paid by taxes), including theses costs will not affect our analysis in anyway; if only cooperators have to pay these costs (e.g. if the institution is paid from ticket revenues of the metro), they can be subsumed in the costs of cooperation. Irrespective of how the punishment system got established and who pays the costs, we asked the question whether, and to what extent, punishment-enforced cooperation can be undermined by insurance against punishment.

When punishment is effective in the sense that cooperation is a dominant strategy in the absence of speculation ($\tilde{c} < \alpha\beta$) our answer is twofold. In principle, speculation can destabilize co operation, by either turning the cooperative equilibrium from a global attractor into a locally stable equilibrium (scenarios C_4 and C_5) or by destabilizing it completely (scenario C_6). However, such destabilization only occurs if the costs of the insurance policy lambda(x_s) are smaller than $\alpha\beta$ for at least some values of x_s. Such low costs do not seem realistic. In fact, the expected fines to be paid by the insurance company for each policy holder amount to $\alpha\beta$. Taking into account additional overhead costs and a profit margin, one would expect that an insurance system is only viable if the costs of the policy are (substantially) larger than $\alpha\beta$. If this is the case, however, cooperation remains a global attractor.

A similar conclusion can be drawn for the case $\tilde{c} > \alpha\beta$ where a punishment system is in place, although it is not effective in the sense that it does not deter against defection. In all scenarios, the cooperative equilibrium remains unstable; hence speculation cannot stabilize cooperation. Defection can be destabilize by speculation (scenarios D_4, D_5, D_6), but only in the case that the costs of the insurance policy are below $\alpha\beta$ and, hence, unreasonably low.

Acknowledgments

This work was supported by the National Basic Research Program of China (973 Program, No. 2012CB821200), the National Natural Science Foundation of China (No. 60974064) and the Erasmus Mundus Action 2 (EMA2) Lotus project. We thank J. van Gestel for helpful comments and discussion.

References

Boyd, R., Gintis, H., Bowles, S., Richerson, P.J., 2003. The evolution of altruistic punishment. Proc. Natl. Acad. Sci. USA 100, 3531–3535.
Boyd, R., Lorberbaum, J.P., 1987. No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game. Nature 327, 58–59.
Brandt, H., Hauert, C., Sigmund, K., 2006. Punishing and abstaining for public goods. Proc. Natl. Acad. Sci. USA 103, 495–497.
Fehr, E., Gächter, S., 2002. Altruistic punishment in humans. Nature 415, 137–140.
Fowler, J.H., 2005. Altruistic punishment and the origin of cooperation. Proc. Natl. Acad. Sci. USA 102, 7047–7049.
Fu, F., Hauert, C., Nowak, M.A., Wang, L., 2008. Reputation-based partner choice promotes cooperation in social networks. Phys. Rev. E 78, 026117.
Hardin, G., 1968. The tragedy of the commons. Science 162, 1243–1248.
Hauert, C., Traulsen, A., Brandt, H., Nowak, M.A., Sigmund, K., 2007. Via freedom to coercion: the emergence of costly punishment. Science 316, 1905–1907.
Heckathorn, D.D., 1996. The dynamics and dilemmas of collective action. Am. Soc. Rev. 61, 250–277.
Maynard Smith, J., 1982. Evolution and the Theory of Games. Cambridge University Press, Cambridge, UK.
McNamara, J., Weissing, F., 2010. Evolutionary game theory. In: T. Székeley, A.J. Moore, J. Komdeur (Eds.), Social Behaviour. Genes, Ecology and Evolution. Cambridge, UK: Cambridge University Press, pp. 109–133.
Nakamaru, M., Iwasa, Y., 2006. The coevolution of altruism and punishment: role of the selfish punisher. Journal of Theoretical Biology 240, 475–488.
Nowak, M.A., 2006. Five rules for the evolution of cooperation. Science 314, 1560–1563.
Nowak, M.A., May, R.M., 1992. Evolutionary games and spatial chaos. Nature 359, 826–829.
Nowak, M.A., Sigmund, K., 1992. Tit for tat in heterogeneous population. Nature 355, 250–253.
Pacheco, J.M., Traulsen, A., Nowak, M.A., 2006. Coevolution of strategy and structure in complex networks with dynamical linking. Phys. Rev. Lett. 97, 258103.
Pacheco, J.M., Traulsen, A., Ohtsuki, H., Nowak, M.A., 2008. Repeated games and direct reciprocity under active linking. J. Theor. Biol. 250, 723–731.
Perc, M., 2012. Sustainable institutionalized punishment requires elimination of second-order free-riders. Sci. Rep. 2.
Perc, M., Szolnoki, A., 2010. Coevolutionary games: a mini review. BioSystems 99, 109–125.
Piazza, J., Bering, J., 2008. The effects of perceived anonymity on altruistic punishment. Evol. Psychol. 6, 487–501.
Rand, D., Dreber, A., Ellingsen, T., Fudenberg, D., Nowak, M., 2009. Positive interactions promote public cooperation. Science 325, 1272–1275.
Rand, D., Nowak, M., 2011. The evolution of antisocial punishment in optional public goods games. Nat. Commun. 2, 434.
Reece, S., Shuker, D., Pen, I., Duncan, A., Choudhary, A., Batchelor, C., West, S., 2003. Kin discrimination and sex ratios in a parasitoid wasp. J. Evol. Biol. 17, 208–216.

Sigmund, K., De Silva, H., Traulsen, A., Hauert, C., 2010. Social learning promotes institutions for governing the commons. Nature 466, 861–863.

Szabó, G., Fáth, G., 2007. Evolutionary games on graphs. Phys. Rep. 446, 97–216.

Szolnoki, A., Perc, M., 2010. Reward and cooperation in the spatial public goods game. Europhys. Lett. 92, 38003.

Szolnoki, A., Szabó, G., Perc, M., 2011. Phase diagrams for the spatial public goods game with pool punishment. Phys. Rev. E 83, 036101.

Van den Berg, P., Weissing, F., 2012. The social costs of punishment. Behav. Brain Sci. 35, 42–43.

West, S., Griffin, A., Gardner, A., 2007. Evolutionary explanations for cooperation. Curr. Biol. 17, R661–R672.

Xavier, J., Martinez-Garcia, E., Foster, K., 2009. Social evolution of spatial patterns in bacterial biofilms: when conflict drives disorder. Am. Nat. 174, 1–12.

Zhang, C., Zhang, J., Weissing, F., Perc, M., Xie, G., Wang, L., 2012. Different reactions to adverse neighborhoods in games of cooperation. PloS One 7, e35183.