SCIENTIFIC REPORTS

OPEN

Conditional punishment is a double-edged sword in promoting cooperation

Feng Huang [1,2], Xiaojie Chen[2] & Long Wang[1]

Punishment is widely recognized as an effective approach for averting from exploitation by free-riders in human society. However, punishment is costly, and thus rational individuals are unwilling to take the punishing action, resulting in the second-order free-rider problem. Recent experimental study evidences that individuals prefer conditional punishment, and their punishing decision depends on other members' punishing decisions. In this work, we thus propose a theoretical model for conditional punishment and investigate how such conditional punishment influences cooperation in the public goods game. Considering conditional punishers only take the punishing action when the number of unconditional punishers exceeds a threshold number, we demonstrate that such conditional punishment induces the effect of a double-edged sword on the evolution of cooperation both in well-mixed and structured populations. Specifically, when it is relatively easy for conditional punishers to engage in the punishment activity corresponding to a low threshold value, cooperation can be promoted in comparison with the case without conditional punishment. Whereas when it is relatively difficult for conditional punishers to engage in the punishment activity corresponding to a high threshold value, cooperation is inhibited in comparison with the case without conditional punishment. Moreover, we verify that such double-edged sword effect exists in a wide range of model parameters and can be still observed in other different punishment regimes.

The solutions to many challenges in human societies, such as the management of public resources[1–3] and the global warming[4–6], all boil down to resort to a broad collective action of cooperation. However, the dilemma of helping others at a cost to ourselves or refraining from doing so but still profiting from the efforts provided by others[7], always leads to the collapse of cooperation. As a solution to the dilemma of cooperation, costly punishment has attracted much attention both from the theoretical[8–11] and experimental[12–17] perspectives. But its side effect that the enforcement can lower the income of punishers is also highlighted[9,18]. Hence, whether contributing to the punishment pool or not becomes a similar dilemma as whether contributing to the public good or not[19–21].

The puzzle about the emergence of costly punishment can be solved by considering some additional factors, such as reputation[22–24], group selection[25–28], social exclusion[29,30], and consideration of sanctioning the second-order free-riders[10,31]. In addition, by including a loner strategy, voluntary participation also paves the way to solve the dilemma of costly punishment[8,9,32,33]. Based on the assumption that punishment is considered to be unconditional and uncoordinated individual action automatically triggered by defectors[34], however, it seems that the loner strategy does not effectively address the inherent dilemma for the initial emergence of costly punishment, since rare punishers must undertake enough punishment when defection are prevalent[18]. On the contrary, the coordinated effort among punishers is well documented both in ethnographic evidence and behavioral experiments with communication or with the option of coordinating behavior[35,36]. It seems that such coordinated strategy can provide another method to overcome the problem of costly punishment because punishers do not bear the cost of punishment permanently. Motivated by the ethnographic evidence and behavioral experiments, a theoretically seminal work on coordinated punishment shows that cooperation can be sustained and such punishment can proliferate when rare[34]. Moreover, some other variants of punishment, such as conditional punishment[37] defined by imposing a fine with a strength proportional to the number of punishers in their own groups and probabilistic sharing of punishment responsibility[38], also play an important role in solving the problem of second-order free-riders.

[1]Center for Systems and Control, College of Engineering, Peking University, Beijing, 100871, China. [2]School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, 611731, China. Correspondence and requests for materials should be addressed to X.C. (email: xiaojiechen@uestc.edu.cn)

It is worth mentioning that a recent behavior experiment found that individual's punishing decision is on average significantly positively proportional to other members' punishing decisions[39]. Actually, such sheep-flock effect of punishing behavior or the threshold effect of collective action[40], is very ubiquitous in human societies and in animals. For example, when robbers implement a robbery in a public place, policemen may behave righteously and bring the thief to justice at once. While general civilians may hesitate to engage in sanction and their punishing decisions to robbers should significantly depend on the number of individuals who perform the punishment. And this novel behavior among punishers is completely distinct from the coordinated punishment investigated in some aforementioned works[34,37,38]. Hence, it still remains unclear how such conditional punishment, under which whether to sanction free-riders or not depends on the number of unconditional punishers in the group, influences the evolution of cooperation.

In this work, we then propose a theoretical model for conditional punishment in the context of public goods games, and consider that conditional punishers will participate in the punishment activity with other unconditional punishers only when the number of unconditional punishers in the group is not less than a threshold number, otherwise they will just cooperate. In addition to the consideration of well-mixed populations, we also investigate the conditional punishment in structured populations out of the interest for dynamics in some real social systems[41,42]. Considering that very little work has addressed questions about the relative efficacy of different types of punishment pointed out in ref.[17], we further take into account different punishment forms[10,29,38]. As we will show in what follows, the introduction of conditional punishment induces the effect of a double-edged sword on the evolution of cooperation. That is, if the threshold for the number of unconditional punishers is low, more conditional punishers will jump on the bandwagon and punish free-riders, which sustains cooperation. Otherwise, a high threshold exacerbates the second-order free-rider problem of punishment. And we verify that such effect is robust against population structures and punishment regimes.

Model

We consider that individuals in a population play the public goods game in which G individuals are chosen randomly to form a group for playing the game. Each player is set as a pure cooperator (C), a pure defector (D), an unconditional punisher (P), or a conditional punisher (M). Except for defectors who contribute nothing to the common pool but exploit others' efforts, all three other strategists contribute a fixed amount c to the common pool. The sum of all contributions in the group will be multiplied by a synergy factor r, and then allotted equally among all group members irrespective of their contributions.

Subsequently, the punishment mechanism will work as long as there exists at least one defector and one punisher in the group. Each unconditional punisher will impose a fine α on each defector in the group. While all pure cooperators only contribute to the public good but refrain from punishing defectors, who are the second-order free-riders[18,43]. Conditional punishers are principally cooperators who contribute to the common pool, but meanwhile permanently observe the choices of other players in the group at an additional cost of γ. Such observation will assist conditional punishers to discern the number of unconditional punishers in the group. When the number of unconditional punishers is not less than the threshold H, which should be satisfied $0 < H < G$, the punishing action from conditional punishers will be triggered. Each conditional punisher will impose the same fine α on each defector as a reaction. Otherwise, they do nothing but cooperate. Thus, in our model the punishment decision of conditional punishers to a defector significantly relies on the number of unconditional punishers. And when H is low, it means that it is relatively easy for conditional punishers to participate in the punishment activity. While when H is high, it means that the environment for conditional punishers to participate in the punishment activity is more harsh. In addition, each defector penalized for free-riding will bring a cost β to the community of punishers. And the associated costs are equally shared among individuals who participate in the punishment activity following a previous work[38].

Accordingly, we designate the number of pure cooperators, pure defectors, unconditional punishers, and conditional punishers as N_C, N_D, N_P, and N_M among the other $G - 1$ players in the group, respectively. And hence the payoffs of cooperators (Π_C), defectors (Π_D), unconditional punishers (Π_P), and conditional punishers (Π_M) from the group are given by, respectively,

$$
\begin{aligned}
\Pi_C &= \frac{r(G - N_D)c}{G} - c, \\
\Pi_D &= \frac{r(G - N_D - 1)c}{G} - [N_P + \delta(N_P - H)N_M]\alpha, \\
\Pi_P &= \Pi_C - \frac{N_D\beta}{N_P + \delta(N_P + 1 - H)N_M + 1}, \\
\Pi_M &= \Pi_C - \delta(N_P - H)\frac{N_D\beta}{N_P + N_M + 1} - \gamma,
\end{aligned}
\tag{1}
$$

where $\delta(u)$ is the Heaviside step function: $\delta(u) = 1$ if $u \geq 0$, otherwise $\delta(u) = 0$. For the sake of comparison with the case of a structured population, we assume that the group size is $G = 5$ in this paper. Furthermore, without loss of generality, the contribution to the public good is considered to be $c = 1$. And to adhere to the existence of social dilemma[11,44,45], the interval of r values is constrained as $1 < r < G$.

As we have already defined, it is a key point that conditional punishers employ a more sophisticated strategy with following the trend, which characterizes the sheep-flock effect of punishing behavior. More specifically, such a player only behaves as a pure cooperator and refuses to engage in punishment if the number of unconditional punishers is less than a critical threshold H. Otherwise, they will undertake the obligation of punishing defectors, who play the role of unconditional punishers. Such propensities of following the trend for conditional punishers
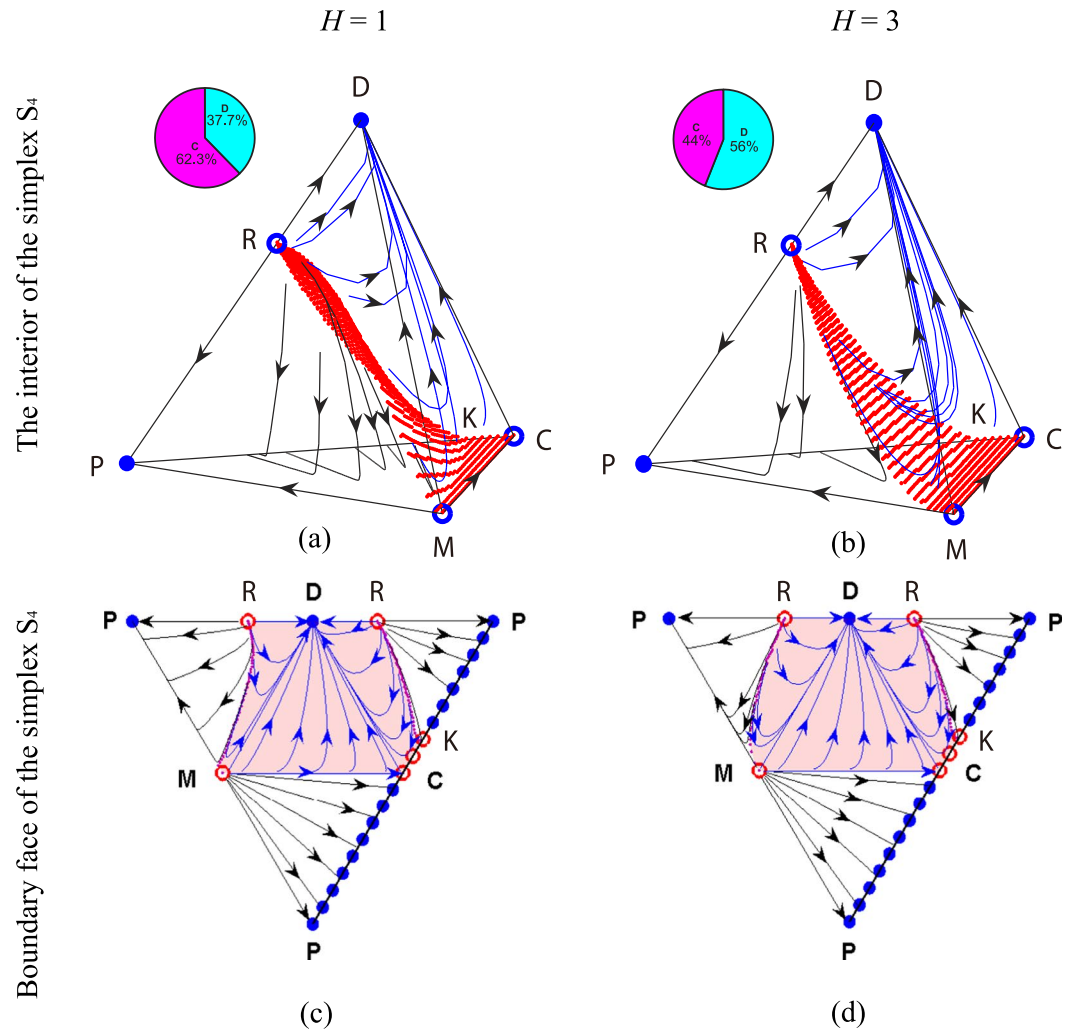
Figure 1. Flow diagrams in the interior of the simplex S_4 and on its boundary faces based on replicator dynamics. Top row (bottom row) depicts evolutionary dynamics in the interior (on the boundary faces) of the simplex S_4 for $H = 1$ (panels (**a**) and (**c**)) and $H = 3$ (panels (**b**) and (**d**)), respectively. Stable fixed points are depicted with solid blue circles, while unstable fixed points are depicted with open blue or red circles. Arrows indicate the direction of evolution. In the interior of the simplex S_4 and on its boundary faces, the system will evolve to the stable equilibrium—either all defectors (vertex D) or the coexistence of cooperators and unconditional punishers (segment PK), in dependence on the initial conditions. The red surface separates the basins of attraction for vertex D and segment PK. In each pie chart, the likelihood that the equilibrium segment PK evolves is indicated by red, and green for the likelihood that the equilibrium D evolves. Additionally, the light red domain in the second row represents the attractive basin of defection on the boundary faces of the simplex S_4. Parameters in panels (**a**) and (**c**): $r = 3$, $c = 1$, $G = 5$, $\alpha = 1.0$, $\beta = 0.7$, $\gamma = 0.05$, and $H = 1$. Parameters in panels (**b**) and (**d**): $r = 3$, $c = 1$, $G = 5$, $\alpha = 1.0$, $\beta = 0.7$, $\gamma = 0.05$, and $H = 3$.

are characterized via the δ function. In general, the value of H can characterize the level of willingness or difficulty for conditional punishers engaging in punishment. Thus, the threshold H is a key parameter in our model. In what follows, we will present the evolutionary dynamics both in well-mixed and structured populations for low and high values of H. In particular, we will show the effects of conditional punishment on the evolution of cooperation by comparing with the case in which conditional punishment is not introduced.

Results

Infinite well-mixed populations. Based on replicator equations, we first present the evolutionary dynamics in infinite well-mixed populations. In Fig. 1, the flow diagrams are shown in the interior of the simplex S_4 and on its boundary faces for two different threshold values, respectively. We find that when conditional punishment is considered, the system will evolve to either the state of all defectors (vertex D) or the coexistence state of cooperators and unconditional punishers (segment PK), no matter whether the threshold value is low or high (Fig. 1(a) and (b)). And such evolutionary outcomes are not changed in comparison with the case in which conditional punishment is not introduced (see the triangle PDC in Fig. 1(c) and (d)). In the simplex S_4, accordingly there exists a surface which divides the whole strategy state space into two basins of attraction. In particular,
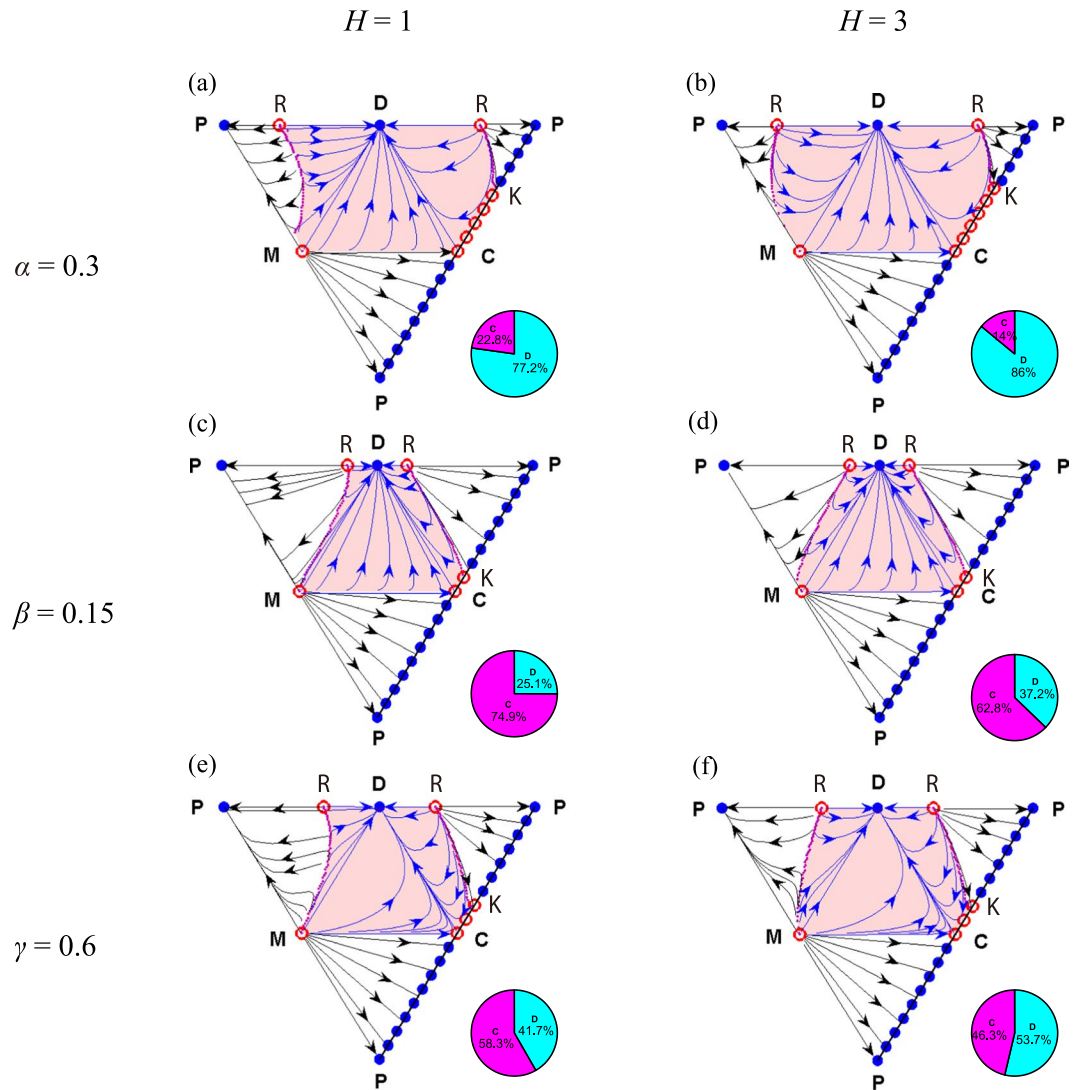
Figure 2. Evolution on the boundary faces of the simplex S_4 for different model parameters. Unstable equilibria are indicated by open red circles, and stable equilibria are indicated by solid blue circles. The defection basin of attraction is depicted by the light red shadow, and the remaining domain is the 'cooperative' basin of attraction. In each pie chart, the red domain represents the likelihood that the equilibrium segment PK evolves, and the green domain for the likelihood that the equilibrium D evolves in the interior of the simplex S_4. For the cases without the strategy of conditional punisher, the cooperative basin of attraction occupies 16.4% ((**a**) and (**b**)), 69.1% ((**c**) and (**d**)), and 48.4% ((**e**) and (**f**)) of the whole triangle DCP, respectively. Parameters: $r = 3$, $c = 1$, $G = 5$, $\alpha = 0.3$, $\beta = 0.7$, and $\gamma = 0.05$ in (**a**) and (**b**); $r = 3$, $c = 1$, $G = 5$, $\alpha = 1.0$, $\beta = 0.15$, and $\gamma = 0.05$ in (**c**) and (**d**); $r = 3$, $c = 1$, $G = 5$, $\alpha = 1.0$, $\beta = 0.7$, and $\gamma = 0.6$ in (**e**) and (**f**).

the unstable interior equilibrium R on the edge DP can be determined by the real root $z^* \in (0, 1)$ of the function $g(z) = \beta\{(1)/(z)[(1-z)^G - 1] + (\alpha(G-1)z)/(\beta) + (rc)/(G\beta) - (c)/(\beta) + 1\}$ (Methods for infinite populations).

Furthermore, we analyze the basin of attraction in the simplex S_4 by numerical calculations, as shown in the pie chart of Fig. 1. We find that for low $H = 1$, the cooperative basin of attraction occupies 62.3% of the whole strategy state space in the simplex S_4. While for high $H = 3$, it only occupies 44.0% of the whole strategy state space. It indicates that the cooperative basin of attraction decreases with increasing the threshold value H. On the other hand, we note that the cooperative basin of attraction occupies 51.5% of the whole triangle PDC shown in Fig. 1(c) and (d), which corresponds to the case without conditional punishment. Thus, we can conclude that the introduction of conditional punishment induces a double-edge sword effect on cooperation. That is, when it is easy for conditional punishers to participate in the punishment activity (low threshold value), cooperation is promoted in comparison with the case in which conditional punishment is not introduced. While when it is difficult for conditional punishers to participate in the punishment activity (high threshold value), cooperation is inhibited in comparison with the same case where conditional punishment is not introduced.

By respectively changing the model parameters (α, β, and γ), we show the evolution of strategies on the boundary faces of the simplex S_4 again (Fig. 2). It is found that the stability of the system does not change when
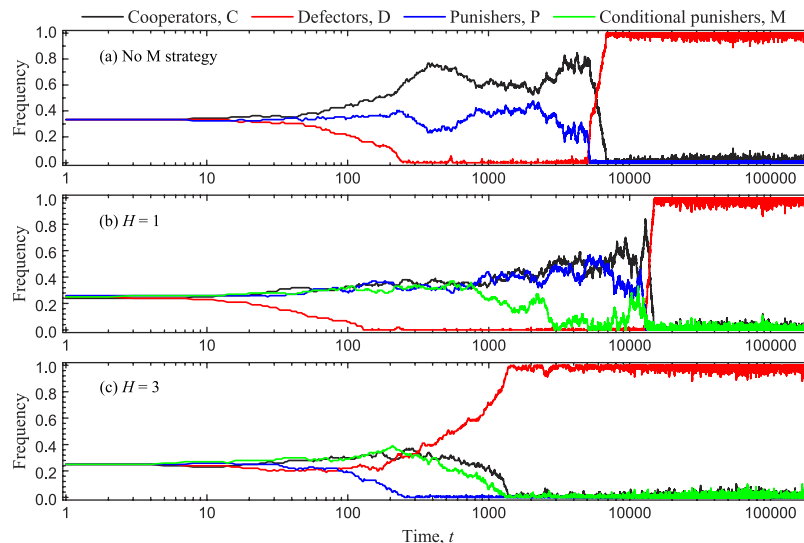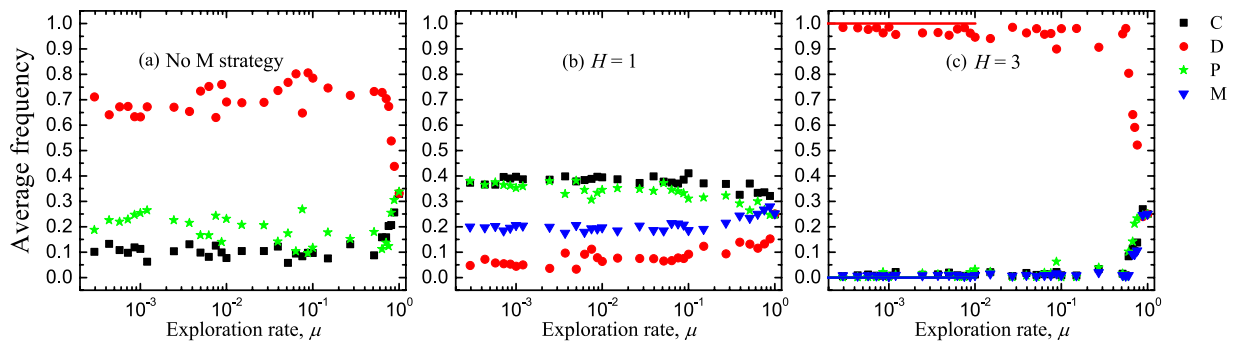
Figure 3. Time evolution of strategies for three different situations. Panel (**a**) shows the time evolution of all three strategies when conditional punishment is not considered. In the presence of conditional punisher strategy, panel (**b**) shows the time evolution of all four strategies for $H = 1$, but panel (**c**) for $H = 3$. Individual-based simulations run over 10^9 time steps, and here we only present the outcomes for 2×10^5 time steps. Parameters: $r = 3$, $c = 1$, $G = 5$, $N = 100$, $\alpha = 1.0$, $\beta = 0.7$, $\gamma = 0.05$, $s = 2.0$, and $\mu = 0.001$.

the parameter values are properly altered. And in comparison with Fig. 1, we further find that the cooperative basin of attraction decreases with decreasing the α value, or increasing the β value, no matter whether the threshold value H is low or high. Moreover, increasing the γ value for low H also decreases the cooperative basin of attraction. But this effect reverses for high H. Importantly, we still see that the double-edge sword effect exists even if these parameter values are changed significantly, which indicates that this finding of the double-edge sword effect remains valid in a broad range of model parameters.

Finite well-mixed populations.   We continue to study the effects of conditional punishment on the evolution of cooperation in finite well-mixed populations. Based on the social learning dynamics[10] with an arbitrary exploration rate μ (Methods for finite populations), we first show the time evolution of strategies for three different situations by individual-based simulations, as shown in Fig. 3. It is noted that for a relatively small threshold the population can temporarily evolve into a quasi-stable state[46] where defectors are suppressed, and only cooperators and unconditional punishers coexist due to neutral drift. Although such quasi-equilibrium is not the evolutionarily stable state no matter whether the conditional punishment is introduced or not (Fig. 3(a) and (b)), the time duration of such quasi-equilibrium can be changed significantly with the change of threshold values once the conditional punishment is introduced. In comparison with the time duration of the quasi-stable state for no conditional punishment as shown in Fig. 3(a), the time duration of the quasi-stable state for low $H = 1$ is longer (Fig. 3(b)). While for high $H = 3$, the quasi-stable state almost does not emerge and the system rapidly evolves into the globally stable equilibrium where the whole population is taken over completely by defectors (Fig. 3(c)). This indicates that the introduction of conditional punishment can still induce a double-edged sword effect in finite well-mixed populations.

In order to illustrate the robustness of the double-edged sword effect in finite populations, we further present the average frequencies of strategies as a function of mutation rate, as shown in Fig. 4. The simulation results are indicated by data points, and the analytical approximations for very small values of μ are indicated by solid lines (Methods for finite populations). We find that for a sufficiently large μ (close to 1), random exploration dominates and results in roughly equal average frequencies for all available strategies. But for a small or moderate μ, the results can be significantly influenced by the threshold H in comparison with the results without the strategy of conditional punisher. Specifically, for low $H = 1$ the average frequencies of cooperators and (unconditional and conditional) punishers are higher than the frequency of defectors (Fig. 4(b)). And importantly, these frequencies are much higher than those in the case where the conditional punisher strategy is not considered, as shown in Fig. 4(a). This shows that cooperation is promoted for low threshold values when conditional punisher strategy is introduced. However, for high $H = 3$ the average frequencies of cooperators and (unconditional and conditional) punishers are much lower than the frequency of defectors (Fig. 4(c)). Correspondingly, these frequencies are also lower than the cooperators' and punishers' frequencies in the case without conditional punisher strategy. This indicates that cooperation is inhibited for high threshold values when conditional punisher strategy is introduced. Therefore, the double-edged sword effect still exists in a broad range of μ in finite well-mixed populations.

Structured populations.   In contrast to the well-mixed case, the fact that the interactions among players are not typically random but rather that each player merely interacts with a set of fixed neighbors in the population[38,47–50], is taken into account in structured populations. Usually, it can lead to some novel and counterintuitive results which are absent in a well-mixed population[51].
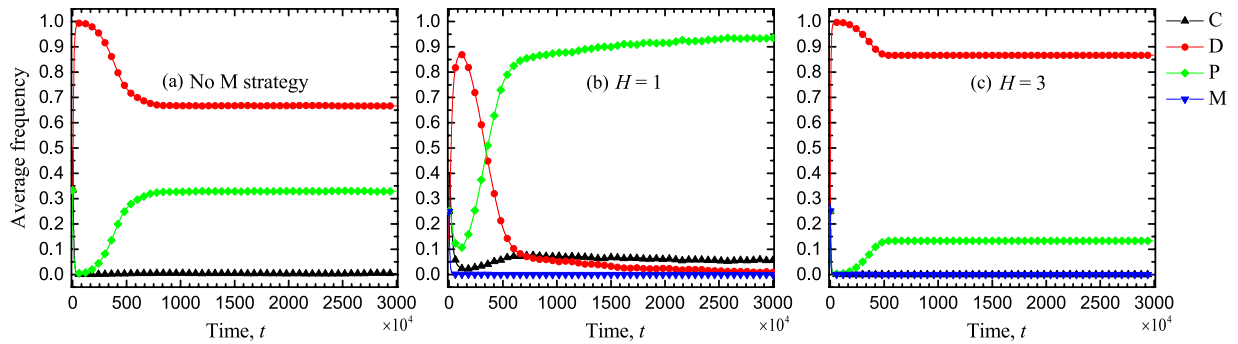
Figure 4. Average frequencies of strategies as a function of mutation rate in three different situations. Panel (**a**) shows the average frequencies of three strategies when conditional punishment is not considered. In the presence of conditional punisher strategy, panel (**b**) shows the average frequencies of all four strategies for $H = 1$, but panel (**c**) for $H = 3$. Symbols indicate results obtained from individual-based simulations (by averaging over 10^4–10^5 imitation steps for each player and by doing 50 independent runs) and solid lines indicate analytical approximations for very small values of μ. Parameters: $r = 3$, $c = 1$, $G = 5$, $N = 100$, $\alpha = 1.0$, $\beta = 0.7$, $\gamma = 0.05$, $s = 2.0$, and $\mu = 0.001$.
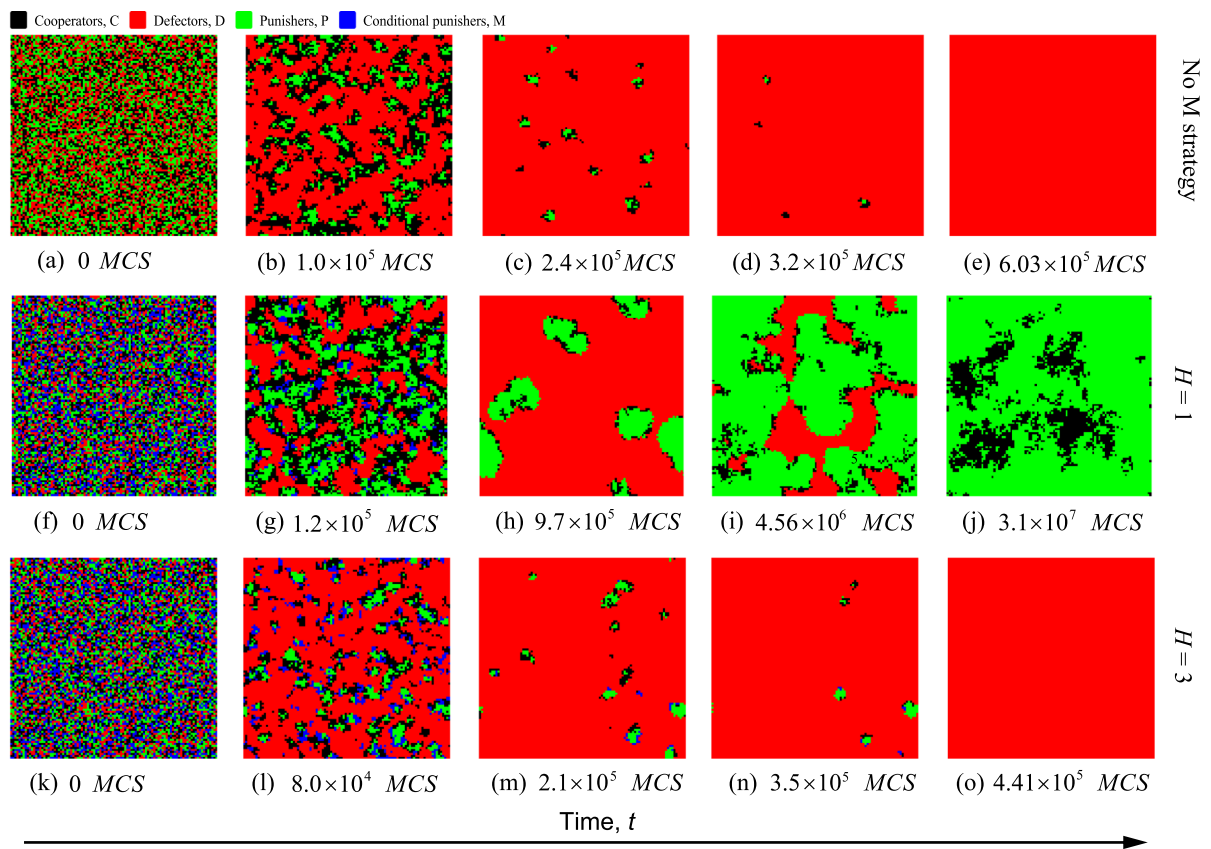
Figure 5. Spatial patterns of strategies over time for three different situations. Cooperators (C) are denoted by black, defectors (D) by red, unconditional punishers (P) by green, and conditional punishers (M) by blue. Top row depicts the typical snapshots over time in the case where the strategy of conditional punisher is not considered. Middle row depicts the typical snapshots over time in the case where the strategy of conditional punisher is introduced and the threshold value is $H = 1$, but $H = 3$ for bottom row. The times of strategy update by Monte Carlo simulation are also shown for each snapshot. Parameters: $G = 5$, $r = 2.0$, $\alpha = 1.0$, $\beta = 0.8$, $c = 1$, $\gamma = 0.05$, $s = 2$, and $L = 100$.

To explore the effects of conditional punishment on cooperation in structured populations, here we show a series of snapshots on a square lattice with the von Neumann neighborhood to depict the spatial formation of all strategies over time (Methods for structured populations), as shown in Fig. 5. First, some characteristic snapshots of the spatial formation are presented in the case where conditional punishment is not considered (top row of

Figure 6. Time evolution of average frequencies of strategies by averaging 30 independent runs in structured populations. Panel (**a**) depicts the average frequencies of the three strategies over time in the case where the strategy of conditional punisher is not considered. Panel (**b**) depicts the average frequencies of all four strategies over time in the case where the strategy of conditional punisher is introduced and the threshold value is $H = 1$, but $H = 3$ for panel (**c**). Parameters: $G = 5$, $r = 2.0$, $\alpha = 1.0$, $\beta = 0.8$, $c = 1$, $\gamma = 0.05$, $s = 2$, and $L = 100$.

Fig. 5). We find that defectors can quickly fight against other two strategists, which results in that cooperators or unconditional punishers can only form small tiny clusters. With the invasion of defectors, these small clusters formed by cooperative individuals finally disappear completely, which leaves defectors to take over the population. Nevertheless, when conditional punishment is considered and the threshold value is low (middle row of Fig. 5), defectors only have some competitive advantages over the other three strategists during the initial period of the evolution, and they can then utilize these advantages to rapidly invade the whole population. Ultimately, it results in the extinction of conditional punishers as well as the decrease of cooperators and unconditional punishers. But when cooperators and unconditional punishers form the compact clusters, they can reverse the invasion of defectors and expand across the whole population. Such results are indicated by the spatial formation that the isolated islands of defectors (depicted by red) are in the sea of cooperators (depicted by black) and punishers (depicted by green), and finally disappear completely. On the contrary, when the threshold H is high (bottom row of Fig. 5), the negative effect of conditional punishers on cooperation is highlighted. Because conditional punishers have less chances to participate in the punishment activity with unconditional punishers, this immediately provides a chance for defectors to obtain a relatively higher fitness. Consequently, defectors can permanently remain successful in the structured public goods game. Hence, in structured populations, the effect of conditional punishment on the evolution of cooperation is still a double-edged sword. When conditional punishment is not considered in the public goods game, the population evolves towards a homogenous state of full defectors (D-only phase), finally. However, if the strategy of conditional punisher is considered, a low threshold value will drive the population to evolve into the coexistence of cooperators and unconditional punishers (a mixed C + P phase), which shows the positive effect of conditional punishment on cooperation. But if a high threshold value is applied, the system will transform from the mixed C + P phase to a D-only phase, similar to the previous finding in spatial public goods game with four strategies[38,51,52]. And it takes a shorter time to reach the D-only phase in comparison with the case without conditional punishment, which shows the negative effect of conditional punishment on cooperation.

In order to eliminate the influence of randomness on the evolutionary outcomes, we further calculate the average frequencies of strategies over time by doing 30 independent runs for those three different situations studied in Fig. 5, as illustrated in Fig. 6. It is our goal to verify the existence of the double-edged sword effect in structured populations by evaluating the average frequencies of strategies in equilibrium. We find that when there is no conditional punishers engaging in structured public goods game, the average frequency of cooperators in equilibrium is zero. And the average frequency of punishers in equilibrium is about 0.33, which is much smaller than that of defectors whose frequency is about 0.67 (Fig. 6(a)). Whereas when conditional punishment is considered and the threshold value is low, we see that cooperators and unconditional punishers coexist in equilibrium and the average frequencies of the other two strategies converge to zero (Fig. 6(b)). It implies that the average level of cooperation is significantly increased in comparison with the case without conditional punishment. On the contrary, when the threshold value is high, we observe the similar results to the case without conditional punishment. But the average frequency of unconditional punishers in equilibrium is about 0.15 (Fig. 6(c)), which is much lower than that in the case without conditional punishment. This shows that cooperation is obviously inhibited when it compares with the results in the case without conditional punishment. We thus conclude that the double-edged sword effect is also embodied in structured populations. In addition, we have checked that such effect still exists even if we properly change the initial conditions for these three situations.

Discussion

Since a decision-making may cause damage to our own interests, the reaction made by depending on others' decisions is usually a dominant strategy most of the time. The effect of sheep-flock in our life is a typical example. Unlike the definition given by the previous study[37], our model regarding conditional punishment characterizes the sheep-flock effect of the punishing behavior which has been documented in the experimental research[39]. By means of theoretical analysis and computer simulations, we have explored the effects of conditional punishment on the evolution of cooperation. Conceptually similar to the conditional cooperation or conditional participation

in joint efforts[7,51,53,54], conditional punishers can utilize the advantages of both unconditional cooperators and unconditional punishers. But simultaneously, it also induces the effect of a double-edged sword on cooperation. When it is relatively easy for conditional punishers to participate in the punishment activity, the invasion of defectors can be controlled by punishment, and the cost caused by punishment can be also shared by more individuals. Thus, in comparison with the case without conditional punishment, cooperation can be further promoted. Whereas when it is relatively difficult for conditional punishers to participate in punishment activity, they are more willing to perform prosocial cooperation rather than spiteful punishment. In this way, the threat of sanctioning free-riders is so weak that the environment for cooperation to thrive becomes more harsh. Thus, cooperation is inhibited when it compares with the outcome in the case without conditional punishment.

Moreover, it is necessary to point out that in our model the ability to recognize the level of unconditional punishers is not self-serving but costly for conditional punishers, because the self-serving function does not seem to be the feature of punishment in real life[29,51]. Accordingly, the extra cost makes conditional punishers do not have the competitive advantages over other two cooperative strategists, no matter what the threshold value is. And meanwhile it also differentiates the strategy of conditional punisher from other cooperative strategies, essentially. Our study thus reveals the significant role of additional cost in the evolution of cooperation, and shows that the introduction of conditional punishment can alleviate or exacerbate the second-order free-rider problem[9,29,55], which strongly depends on the the threshold value. Consequently, conditional punishment induces the effect of a double-edged sword on the evolution of cooperation.

Although different punishment modes can result in different outcomes according to previous studies[10,17], we verify that the double-edged sword effect found in our model is also valid for other punishment regimes, such as peer punishment and a variant of its (see Supplementary Information). Undeniably, in the framework of our model, sanction is merely targeted at free-riders, and the possibility of anti-social punishment[24,56,57] that non-cooperators attack cooperators is excluded a priori. To make up this deficiency, we additionally consider a model variant that includes the possibility of anti-social punishment (see Supplementary Information). When defectors suffer from the sanction of punishers, it will trigger defectors to revenge all members in the group. Surprisingly, cooperation is still sustained in the population. In addition, as an important direction to develop our model, considering the heterogeneity of the threshold[52,58,59] for matching the reality well is worth the effort in the future.

Methods

Evolutionary dynamics in infinite well-mixed populations.

We study the evolution of strategies in infinite well-mixed populations based on replicator dynamics[60,61]. First, we define that the fraction of cooperators (C), defectors (D), unconditional punishers (P), and conditional punishers (M) can be denoted by x, y, z, and w, respectively. Thus we have $x + y + z + w = 1$. Accordingly, the replicator equations are given by

$$
\begin{aligned}
\dot{x} &= x(P_C - \overline{P}), \\
\dot{y} &= y(P_D - \overline{P}), \\
\dot{z} &= z(P_P - \overline{P}),
\end{aligned}
\tag{2}
$$

where dots denote the derivatives with respect to time t and P_i designates the expected payoff for each strategy i (i = C, D, P, or M), which is given by

$$
P_i = \sum_{0 \leq N_s \leq G-1} \frac{(G-1)!}{N_C! N_D! N_P! N_M!} x^{N_C} y^{N_D} z^{N_P} w^{N_M} \Pi_i,
\tag{3}
$$

where N_s is the number of players choosing strategy s (s = C, D, P, M) in a group, hence $\sum_s N_s = G - 1$. Π_i represents the payoff of strategy i, which is defined by Eq. (1). $\overline{P}$ describes the average payoff of the entire population, which is given by $\overline{P} = x P_C + y P_D + z P_P + w P_M$.

For discussing the evolution of these four strategies, we first consider there are no any punishers in the population. In this way, defectors can exploit the effort of cooperators permanently. Therefore, natural selection will always favor defectors to take over the population, irrespective of the initial conditions.

However, the introduction of punishers can effectively reverse the negative situation. Thus we consider that only defectors and unconditional punishers are presented in the population, namely $y + z = 1$. Then the replicator equation degenerates to $\dot{z} = z(1 - z)(P_P - P_D)$. In this situation, the average payoff of punishers P_P is given by

$$
\begin{aligned}
P_P &= \sum_{k=0}^{G-1} \binom{G-1}{k} z^k (1-z)^{G-k-1} \left[\frac{r(k+1)c}{G} - c - \frac{(G-k-1)\beta}{k+1} \right] \\
&= \frac{\beta}{z} [(1-z)^G - 1] + \frac{rc}{G} [(G-1)z + 1] + \beta - c.
\end{aligned}
\tag{4}
$$

Similarly, the average payoff of defectors P_D is given by

$$
\begin{aligned}
P_D &= \sum_{k=0}^{G-1} \binom{G-1}{k} z^k (1-z)^{G-k-1} \left[\frac{rkc}{G} - k\alpha \right] \\
&= \left(\frac{rc}{G} - \alpha \right)(G-1)z.
\end{aligned}
\tag{5}
$$

With these expressions, the replicator equation has two boundary equilibria, namely $z = 0$ and $z = 1$. On the other hand, the interior equilibria can be determined by the roots of the function $g(z) := P_P - P_D$, thus obtaining

$$g(z) = \frac{\beta}{z}[(1-z)^G - 1] + \alpha(G-1)z + \frac{rc}{G} + \beta - c. \tag{6}$$

It follows that $g(0) = \lim_{z \to 0^+} g(z) = \frac{rc}{G} - c + \beta(1-G) < 0$ with $1 < r < G$ and $c = 1$. Note that the function $g(z)$ can be approximated by $g(z) \approx (\alpha + G\beta/2)(G-1)z + \beta(1-G) + (rc)/(G) - c$. Thus the function $g(z)$ is strictly increasing since $g'(z) \approx (\alpha + G\beta/2)(G-1) > 0$. Accordingly, the interior equilibrium is determined by $g(1) = \alpha(G-1) + (rc)/(G) - c$, from which we have the following two conclusions:

1. When $\alpha > ((G-r)c)/(G(G-1))$, the replicator equation has only one interior equilibrium $z^* \in (0, 1)$, but it is unstable since $g'(z^*) > 0$. The two boundary equilibria $z = 0$ and $z = 1$ are both stable.
2. When $\alpha \leq ((G-r)c)/(G(G-1))$, the replicator equation has no interior equilibria in $(0, 1)$. $z = 0$ is a stable equilibrium, while $z = 1$ is an unstable equilibrium.

Moreover, if there are no defectors in the population, the average payoff of cooperators is equal to that the unconditional punishers obtain from the public goods game. And it is higher than the average payoff of conditional punishers, because the latter have to pay the observation cost. Thus natural selection will support the system to evolve into the coexistence state of cooperators and punishers because of neutral drift.

Evolutionary dynamics in finite well-mixed populations.

We denote that the population of size N contains X cooperators, Y defectors, Z unconditional punishers, and W conditional punishers. Thus we have $X + Y + Z + W = N$, and the average payoffs of cooperators (C), defectors (D), unconditional punishers (P), and conditional punishers (M) can be given by, respectively,

$$P_C = \sum_{k=0}^{G-1} \frac{\binom{N-1-Y}{k}\binom{Y}{G-k-1}}{\binom{N-1}{G-1}}\left[\frac{r(k+1)c}{G} - c\right]$$
$$= \frac{rc}{G}\left[\frac{(G-1)(N-Y-1)}{N-1} + 1\right] - c, \tag{7}$$

$$P_D = \sum_{k=0}^{G-1} \frac{\binom{N-Y}{k}\binom{Y-1}{N-k-1}}{\binom{N-1}{G-1}}\frac{rkc}{G} - \sum_{0 \leq l,m \leq G-1} \frac{\binom{Z}{l}\binom{W}{m}\binom{N-Z-W-1}{G-l-m-1}}{\binom{N-1}{G-1}}[l + \delta(l-H)m]\alpha, \tag{8}$$

$$P_P = P_C - \sum_{0 \leq l,m,n \leq G-1} \frac{\binom{Z-1}{l}\binom{W}{m}\binom{Y}{n}\binom{N-Z-W-Y}{G-l-m-n-1}}{\binom{N-1}{G-1}}\frac{n\beta}{l+1+\delta(l+1-H)m}, \tag{9}$$

and

$$P_M = P_C - \gamma - \sum_{0 \leq l,m,n \leq G-1} \frac{\binom{Z}{l}\binom{W-1}{m}\binom{Y}{n}\binom{N-Z-W-Y}{G-l-m-n-1}}{\binom{N-1}{G-1}}\delta(l-H)\frac{n\beta}{l+m+1}, \tag{10}$$

where k, l, m, and n represent the number of contributors, unconditional punishers, conditional punishers, and defectors among $G - 1$ players in a group, respectively.

Next, we employ a so-called social learning process[10] to describe the evolution of all strategies in finite well-mixed populations. Let us denote that P_u and P_v are the average payoffs of two randomly chosen players u and v, respectively. Under pairwise comparison rule[45,62,63], player u adopts the strategy of player v with a probability given by the Fermi function[64]

$$q = \frac{1}{1 + exp[-s(P_v - P_u)]}, \tag{11}$$

where the imitation strength $s \geq 0$ measures the intensity of selection that determines the level of uncertainty in the strategy imitation process[4,11]. Without loss of generality, we use a representative value $s = 2$[11,19,51] in finite well-mixed and structured populations, which implies that the better performing players are readily imitated, but it is not impossible to adopt the strategy of a player performing worse.

Then we denote that N_i is the number of players choosing strategy i. Hence the probability that one chosen as a focal player out of N_i players with strategy i imitates another player of the $N_j = N - N_i$ players with strategy j ($j \neq i$ and $j = C$, D, P, or M) is given by

$$\tau_{i \to j}(N_i) = \frac{N_i}{N} \frac{N - N_i}{N} \frac{1}{1 + exp[-s(P_j - P_i)]}.$$

(12)

As a result, the fixation probability that characterizes the fixation of the dissident strategy j caused by imitation in the population can be computed by

$$\rho_{ij} = \frac{1}{1 + \sum_{q=1}^{N-1} \prod_{N_j=1}^{q} \frac{\tau_{j \to i}(N_j)}{\tau_{i \to j}(N_j)}}.$$

(13)

It is noted that the equation $N_j = N - N_i$ is always met, so the fixation probability ρ_{ij} can be simplified to

$$\rho_{ij} = \frac{1}{1 + \sum_{q=1}^{N-1} exp[s \sum_{N_j=1}^{q} (P_i - P_j)]}.$$

(14)

Furthermore, let us denote that the homogeneous population with $N_i = N$ is All_i and the random exploration rate is μ. In the case of four strategies (C, D, P, and M), with probability $\mu/3$ a single individual randomly switches from strategy i to the strategy j ($j \neq i$). Thus the transition probability p_{ij} from All_i to All_j is $\mu\rho_{ij}/3$. In this way, the transition matrix of the complete Markov chain can be written as $Pr = [p_{ij}]_{4 \times 4}$. Accordingly, the stationary distribution which describes the percentage of time spent by the state of the population in the vicinity of the homogeneous state[10], is given by the normalized left eigenvector to the eigenvalue 1. In addition, it is shown that the stationary distribution of the full system converges to the stationary distribution of this 'embedded' Markov chain on the homogeneous states[65,66] for $\mu \to 0$, of which transition probabilities from All_i to $All_j(j \neq i)$ are given by $\rho_{ij}/3$[10]. Thus for four competitive strategies, the transition matrix can be concisely written by

$$Pr = \begin{array}{c} \\ C \\ D \\ P \\ M \end{array} \begin{array}{cccc} C & D & P & M \\ \begin{pmatrix} 1 - \frac{1}{3}\sum_j \rho_{Cj} & \frac{1}{3}\rho_{CD} & \frac{1}{3}\rho_{CP} & \frac{1}{3}\rho_{CM} \\ \frac{1}{3}\rho_{DC} & 1 - \frac{1}{3}\sum_j \rho_{Dj} & \frac{1}{3}\rho_{DP} & \frac{1}{3}\rho_{DM} \\ \frac{1}{3}\rho_{PC} & \frac{1}{3}\rho_{PD} & 1 - \frac{1}{3}\sum_j \rho_{Pj} & \frac{1}{3}\rho_{PM} \\ \frac{1}{3}\rho_{MC} & \frac{1}{3}\rho_{MD} & \frac{1}{3}\rho_{MP} & 1 - \frac{1}{3}\sum_j \rho_{Mj} \end{pmatrix} \end{array},$$

(15)

where j is subject to three other strategies in the group except the imitator itself.

In particular, in the limiting case of strong imitation ($s \to +\infty$), the transition matrix can be significantly simplified by

$$Pr = \begin{pmatrix} \frac{2N-1}{3N} & \frac{1}{3} & \frac{1}{3N} & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{3N} & 0 & \frac{3N-1}{3N} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix}.$$

(16)

And the stationary distribution (the left eigenvector to the eigenvalue 1) is easily given by (0, 1, 0, 0), which implies the population becomes a stable regime of defectors, leading to the tragedy of the commons.

Individual-based simulations for finite well-mixed populations. We consider a finite well-mixed population with a constant size N. Each individual achieves an expected payoff defined by Eqs (7)–(10) based on the random sampling of the interaction groups. Strategies evolve in dependence on a mutation-selection process defined in discrete time[7]. At each time step, a player u is randomly selected to update. With probability μ, the player u undergoes a mutation and randomly adopts one strategy from the space of available strategies. With probability $1 - \mu$, another individual v is randomly selected to act as a role model for player u. Then player u adopts the strategy of player v with a probability q defined by Eq. (11). Otherwise, player u sticks to its strategy with the probability $1 - q$.

Individual-based simulations for structured populations. Here, we consider a structured population where the public goods game is staged on a $L \times L$ square lattice with periodic boundary conditions. L^2 players are arranged into overlapping groups of size $G = 5$ such that everyone is connected to its $G - 1$ nearest neighbors, which implies that each individual is involved in G different groups. Hence the overall payoffs for each player are the sum of all the profits acquired from G groups. Initially, the player on every site is designated either as a cooperator, defector, unconditional punisher, or conditional punisher with equal probability. At every time step, a player u is randomly selected to play the public goods game with its four neighbors as a member of all five groups and obtains its overall payoffs P_u. Similarly, another player v, one of the four nearest neighbors, is chosen randomly and acquires its total payoffs P_v in the same way. If their strategies are different, the imitation is executed with the probability defined by Eq. (11). In each full round of the game, every player has one chance to imitate from one of their neighbors on average[19,37,51].

References

1. Ostrom, E. *Governing the Commons: The Evolution of Institutions for Collective Action*. (Cambridge University Press, Cambridge, UK, 1990).
2. Poteete, A. R., Janssen, M. A. & Ostrom, E. *Working together: collective action, the commons, and multiple methods in practice* (Princeton University Press, Princeton, NJ, 2010).
3. Sober, E. & Wilson, D. S. *Unto others: The evolution and psychology of unselfish behavior*. (Harvard University Press, Cambridge, MA, 1999).
4. Santos, F. C. & Pacheco, J. M. Risk of collective failure provides an escape from the tragedy of the commons. *Proc. Natl. Acad. Sci. USA* **108**, 10421–10425 (2011).
5. Schneider, S. H. What is 'dangerous' climate change? *Nature* **411**, 17–19 (2001).
6. Tavoni, A., Dannenberg, A., Kallis, G. & Löschel, A. Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proc. Natl. Acad. Sci. USA* **108**, 11825–11829 (2011).
7. Van Segbroeck, S., Pacheco, J. M., Lenaerts, T. & Santos, F. C. Emergence of fairness in repeated group interactions. *Phys. Rev. Lett.* **108**, 158104 (2012).
8. 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).
9. 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).
10. Sigmund, K., De Silva, H., Traulsen, A. & Hauert, C. Social learning promotes institutions for governing the commons. *Nature* **466**, 861–863 (2010).
11. Chen, X., Sasaki, T., Brännström, Å. & Dieckmann, U. First carrot, then stick: how the adaptive hybridization of incentives promotes cooperation. *J. R. Soc. Interface* **12**, 20140935 (2015).
12. Fehr, E. & Gächter, S. Cooperation and punishment in public goods experiments. *Am. Econ. Rev.* **90**, 980–994 (2000).
13. Rockenbach, B. & Milinski, M. The efficient interaction of indirect reciprocity and costly punishment. *Nature* **444**, 718–723 (2006).
14. Henrich, J. *et al.* Costly punishment across human societies. *Science* **312**, 1767–1770 (2006).
15. Gächter, S., Renner, E. & Sefton, M. The long-run benefits of punishment. *Science* **322**, 1510–1510 (2008).
16. Dreber, A., Rand, D. G., Fudenberg, D. & Nowak, M. A. Winners don't punish. *Nature* **452**, 348–351 (2008).
17. Raihani, N. J., Thornton, A. & Bshary, R. Punishment and cooperation in nature. *Trends Ecol. Evol.* **27**, 288–295 (2012).
18. Panchanathan, K. & Boyd, R. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature* **432**, 499–502 (2004).
19. Perc, M. & Szolnoki, A. Self-organization of punishment in structured populations. *New J. Phys.* **14**, 043013 (2012).
20. Colman, A. M. The puzzle of cooperation. *Nature* **440**, 744–745 (2006).
21. Perc, M. *et al*. Statistical physics of human cooperation. *Phys. Rep.* **687**, 1–51 (2017).
22. Nowak, M. A. & Sigmund, K. Evolution of indirect reciprocity by image scoring. *Nature* **393**, 573–577 (1998).
23. Fu, F., Hauert, C., Nowak, M. A. & Wang, L. Reputation-based partner choice promotes cooperation in social networks. *Phys. Rev. E* **78**, 026117 (2008).
24. Hilbe, C. & Traulsen, A. Emergence of responsible sanctions without second order free riders, antisocial punishment or spite. *Sci. Rep.* **2**, 458 (2012).
25. Boyd, R. & Richerson, P. J. Group selection among alternative evolutionarily stable strategies. *J. Theor. Biol.* **145**, 331–342 (1990).
26. Traulsen, A. & Nowak, M. A. Evolution of cooperation by multilevel selection. *Proc. Natl. Acad. Sci. USA* **103**, 10952–10955 (2006).
27. Perc, M., Gómez-Gardeñes, J., Szolnoki, A., Flora, L. M. & Moreno, Y. Evolutionary dynamics of group interactions on structured populations: a review. *J. R. Soc. Interface* **10**, 20120997 (2013).
28. Perc, M. Phase transitions in models of human cooperation. *Phys. Lett. A* **380**, 2803–2808 (2016).
29. Sasaki, T. & Uchida, S. The evolution of cooperation by social exclusion. *Proc. R. Soc. B* **280**, 20122498 (2013).
30. Li, K., Cong, R., Wu, T. & Wang, L. Social exclusion in finite populations. *Phys. Rev. E* **91**, 042810 (2015).
31. Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A. & Sigmund, K. Public goods with punishment and abstaining in finite and infinite populations. *Biol. Theor.* **3**, 114–122 (2008).
32. Semmann, D., Krambeck, H.-J. & Milinski, M. Volunteering leads to rock–paper–scissors dynamics in a public goods game. *Nature* **425**, 390–393 (2003).
33. Sigmund, K. *The calculus of selfishness* (Princeton University Press, Princeton, NJ, 2010).
34. Boyd, R., Gintis, H. & Bowles, S. Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science* **328**, 617–620 (2010).
35. Wiessner, P. Norm enforcement among the ju/'hoansi bushmen. *Hum. Nat.* **16**, 115–145 (2005).
36. Ertan, A., Page, T. & Putterman, L. Who to punish? individual decisions and majority rule in mitigating the free rider problem. *Eur. Econ. Rev.* **53**, 495–511 (2009).
37. Szolnoki, A. & Perc, M. Effectiveness of conditional punishment for the evolution of public cooperation. *J. Theor. Biol.* **325**, 34–41 (2013).
38. Chen, X., Szolnoki, A. & Perc, M. Probabilistic sharing solves the problem of costly punishment. *New J. Phys.* **16**, 083016 (2014).
39. Kamei, K. Conditional punishment. *Econ. Lett.* **124**, 199–202 (2014).
40. Pacheco, J. M., Santos, F. C., Souza, M. O. & Skyrms, B. Evolutionary dynamics of collective action in n-person stag hunt dilemmas. *Proc. R. Soc. B* **276**, 315–321 (2009).
41. Kandori, M., Mailath, G. J. & Rob, R. Learning, mutation, and long run equilibria in games. *Econometrica* **61**, 29–56 (1993).
42. Szabó, G. & Hauert, C. Phase transitions and volunteering in spatial public goods games. *Phys. Rev. Lett.* **89**, 118101 (2002).
43. Fowler, J. H. Human cooperation: second-order free-riding problem solved? *Nature* **437**, E8–E8 (2005).
44. Hauert, C., De Monte, S., Hofbauer, J. & Sigmund, K. Replicator dynamics for optional public good games. *J. Theor. Biol.* **218**, 187–194 (2002).
45. Szabó, G. & Fath, G. Evolutionary games on graphs. *Phys. Rep.* **446**, 97–216 (2007).
46. Holme, P., Trusina, A., Kim, B. J. & Minnhagen, P. Prisoners' dilemma in real-world acquaintance networks: Spikes and quasiequilibria induced by the interplay between structure and dynamics. *Phys. Rev. E* **68**, 030901 (2003).
47. Chen, X. & Szolnoki, A. Individual wealth-based selection supports cooperation in spatial public goods games. *Sci. Rep.* **6**, 32802 (2016).
48. Su, Q., Li, A., Zhou, L. & Wang, L. Interactive diversity promotes the evolution of cooperation in structured populations. *New J. Phys.* **18**, 103007 (2016).
49. Pei, Z., Wang, B. & Du, J. Effects of income redistribution on the evolution of cooperation in spatial public goods games. *New J. Phys.* **19**, 013037 (2017).
50. Allen, B. *et al*. Evolutionary dynamics on any population structure. *Nature* **544**, 227–230 (2017).
51. Szolnoki, A. & Chen, X. Benefits of tolerance in public goods games. *Phys. Rev. E* **92**, 042813 (2015).
52. Szolnoki, A. & Perc, M. Competition of tolerant strategies in the spatial public goods game. *New J. Phys.* **18**, 083021 (2016).
53. Szolnoki, A. & Perc, M. Conditional strategies and the evolution of cooperation in spatial public goods games. *Phys. Rev. E* **85**, 026104 (2012).
54. Sui, X., Wu, B. & Wang, L. Multiple tolerances dilute the second order cooperative dilemma. *Phys. Lett. A* **381**, 3785–3797 (2017).

55. Brandt, H., Hauert, C. & Sigmund, K. Punishing and abstaining for public goods. *Proc. Natl. Acad. Sci. USA* **103**, 495–497 (2006).
56. Rand, D. G. & Nowak, M. A. The evolution of anti-social punishment in optional public goods games. *Nat. Commun.* **2**, 434 (2011).
57. Hauser, O. P., Nowak, M. A. & Rand, D. G. Punishment does not promote cooperation under exploration dynamics when anti-social punishment is possible. *J. Theor. Biol.* **360**, 163–171 (2014).
58. Hauser, O. P., Traulsen, A. & Nowak, M. A. Heterogeneity in background fitness acts as a suppressor of selection. *J. Theor. Biol.* **343**, 178–185 (2014).
59. Kaveh, K., McAvoy, A. & Nowak, M. A. The effect of spatial fitness heterogeneity on fixation probability. *Rreprint arXiv* **1709**, 03031 (2017).
60. Taylor, P. D. & Jonker, L. B. Evolutionary stable strategies and game dynamics. *Math. Biosci.* **40**, 145–156 (1978).
61. Hofbauer, J. & Sigmund, K. *Evolutionary games and population dynamics* (Cambridge University Press, Cambridge, UK, 1998).
62. Traulsen, A., Claussen, J. C. & Hauert, C. Coevolutionary dynamics: from finite to infinite populations. *Phys. Rev. Lett.* **95**, 238701 (2005).
63. Traulsen, A., Pacheco, J. M. & Nowak, M. A. Pairwise comparison and selection temperature in evolutionary game dynamics. *J. Theor. Biol.* **246**, 522–529 (2007).
64. Szabó, G. & Tőke, C. Evolutionary prisoner's dilemma game on a square lattice. *Phys. Rev. E* **58**, 69 (1998).
65. Fudenberg, D. & Imhof, L. A. Imitation processes with small mutations. *J. Econ. Theory* **131**, 251–262 (2006).
66. Antal, T. & Scheuring, I. Fixation of strategies for an evolutionary game in finite populations. *Bull. Math. Biol.* **68**, 1923–1944 (2006).

Acknowledgements

Author Contributions

F.H. and X.C. conceived and performed the research as well as wrote the paper, L.W. conducted and analysed the results. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-017-18727-7.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.