



The Singaporean model in public goods dilemmas with benevolent leaders and bribery

Yinhai Fang^a, Matjaž Perc^{b,c,d,*}, Haiyan Xu^{a,*}

^a College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 211100, China

^b Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, 2000 Maribor, Slovenia

^c Department of Medical Research, China Medical University Hospital, China Medical University, Taichung, Taiwan

^d Complexity Science Hub Vienna, Josefstädterstraße 39, 1080 Vienna, Austria

ARTICLE INFO

Article history:

Received 18 February 2020

Revised 12 May 2020

Accepted 19 May 2020

Available online 22 May 2020

Keywords:

Evolutionary game theory

Cooperation

Punishment

Public goods

Bribery

ABSTRACT

Public goods dilemmas are at the heart of some of the greatest challenges of our time, including climate inaction, growing inequality, and the overuse of natural resources. The public goods game in which cooperators contribute to a common pool that is then shared equally with defectors who contribute nothing captures the gist of the problem. Cooperators therefore cannot prevail, which ultimately leads to the tragedy of the commons. Actions such as punishment, rewards, and exclusion have been shown to help, but they are costly, therefore rendering cooperators second-order free-riders due to their lack of participation in these actions. In the search for a remedy, we study the public goods game with benevolent leaders who, at a personal cost, have the ability to exclude defectors from using common pool resources. We also consider bribers who can pay the leaders to relax their exclusion efforts. In a traditional setting, this setup yields the standard second-order free-rider problem, where, ironically, the leaders are overcome by cooperators, who then themselves succumb to defectors. We show, however, that the Singaporean model – where a leader's payoff is determined not only by the regular sharing income from the firm production but also by the success of gross firm production as an incentive – can resolve the second-order free-rider problem. We also show that the detrimental effect of bribery can always be, no matter how high the bribe, held in check as long as the number of individuals engaged in this activity is low compared to the number of benevolent leaders. Otherwise, an abrupt transition to a cooperator-less state becomes unavoidable. We discuss the implications of our research for designing successful cooperation and anti-corruption strategies in public goods dilemmas.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

Altruistic behavior – the act of voluntarily sacrificing personal benefits to improve the welfare of others – is a remarkable aspect of our biology, and it can also be observed in bacteria, ants, bees, birds, many other higher mammals (Axelrod, 1984; Henrich et al., 2001; Fehr and Fischbacher, 2004; Rand and Nowak, 2013). Why such behavior exists, how it evolved and why it prevails, however, are some of the greatest questions of evolution. It is namely, at odds with the fundamental principles of evolutionary success, which suppose a relentless drive towards the maximization of individual fitness. Clearly, there should thus be no place for coop-

eration in an evolutionary landscape lest the organism is bound to die out.

Nevertheless, cooperation is indeed widespread in nature, especially among humans, and it is one of the central pillars of our evolutionary success Nowak and Highfield (2011). Nonetheless, cooperation is also fragile since it is constantly contested by defection. In public goods dilemmas (Santos et al., 2008), which include vaccination (Fu et al., 2011; Wang et al., 2016), climate inaction (Vasconcelos et al., 2013; Pacheco et al., 2014), and antibiotic overuse (Chen and Fu, 2018), defectors enjoy all the same benefits as cooperators without contributing to the common pool. Population dynamics is considered to be one of the causes of this oscillatory coexistence of cooperators and defectors (Hauert et al., 2008). Moreover, not surprisingly then, individuals are often prepared to cooperate beyond just contributing to the common pool by also contributing to actions such as punishing or excluding defectors (Jiao et al., 2020; Yamagishi, 1986; Fehr and Gächter, 2000; Gächter et al., 2008; Sekiguchi and Nakamaru, 2009; Jordan

* Corresponding authors at: Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, 2000 Maribor, Slovenia (Matjaž Perc).

E-mail addresses: matjaz.perc@gmail.com (M. Perc), xu haiyan@nuaa.edu.cn (H. Xu).

et al., 2016; Liu et al., 2019), or rewarding cooperators (Andreoni et al., 2003; Szolnoki and Perc, 2010; Sasaki and Unemi, 2011; Chen et al., 2014; Sasaki and Uchida, 2014; Okada et al., 2015; Sasaki et al., 2015), in an attempt to bolster altruistic behavior. However, since these actions are costly, those cooperators that do not participate in them become second-order free-riders (Fehr, 2004; Milinski and Rockenbach, 2008), and they become impediments to the success of the population and themselves (Panchanathan and Boyd, 2004; Helbing et al., 2010; Chen et al., 2014; dos Santos, 2015; Szolnoki and Perc, 2017).

The second-order free-rider problem is thus an important challenge to the effectiveness of costly actions that support cooperation. Here, we consider the Singaporean model as a means to resolve the second-order free-rider problem in public goods dilemmas. In particular, we consider exclusion as a means of punishment in the public goods game, which can be enacted by benevolent leaders. The latter are cooperators, but they also carry additional costs associated with excluding defectors from the use of public goods. We also consider bribers who can bribe the leader into relaxing their exclusion criteria as the fourth competing strategy. In the absence of restricted interactions (Szolnoki and Perc, 2015; Szolnoki and Perc, 2017), this setup yield the classical second-order free-rider problem, wherein cooperators outperform the leaders, which ultimately leads to the collapse of cooperation and to the tragedy of the commons.

The Singaporean model (Yew, 2000), however, provides a small but significant twist in the considered public goods game. Namely, the payoffs of the leaders among the population are determined not only by the regular share of the public goods but also by the success of the public goods game in their group. This is akin to the policy in Singapore (Muthukrishna et al., 2017) where a leader's payoff is determined not only by a regular salary, as is the same with most of the individuals where she works but also, as an incentive, by the success of the whole firm or organization. As we will show, this has important consequences for the evolution of cooperation in public goods dilemmas since it can often overcome the second-order free-rider problem, thus paving the way towards sustained cooperation under adverse conditions. Our model also provides important insight into the detrimental effects of bribery. Namely, even if bribes are high, we show that as long as the bribers are rare compared to the benevolent leaders, the downfall of benevolent leaders can be avoided. It is when bribers are common, even if they do not bribe highly, that pressure on the benevolent leaders becomes too strong to enact the efficient exclusion of defectors. This phenomenon can then lead to sudden transitions to cooperator-less states.

The remainder of this paper is organized as follows. In Section 2, we first introduce the public goods game with benevolent leaders besides cooperators and defectors but without bribery. Secondly, we present the results of our model, which include the equilibria and stability analysis and the numerical simulations. The arrangement of Section 3 is the same as Section 2, but the public goods game considers cooperators, defectors, and benevolent leaders, as well as bribers. In Section 4, we summarize the main results and discuss their implications for designing successful cooperation and anti-corruption strategies in social dilemma situations, where what is best for an individual is at odds with what is best for society as a whole.

2. Exclusion without bribery

2.1. Typical methods

We consider a well-mixed, infinitely large population consisting of cooperators, defectors, and benevolent leaders, in which N

($N > 2$) individuals are randomly selected to form a group to play a public goods game (Liu et al., 2019). In the public goods game, each cooperator (C) contributes a constant amount c to the public pool and the defector (D) donates nothing. The benevolent leaders (BLs) play two roles. The first role is the normal collaborator donating c to the public pool. The second is a punisher who contributes μ to the sanction pool to exclude defectors. The total contribution to the public pool will be multiplied by the synergy factor r ($1 \leq r \leq N$) as the final collective payoff and then assigned to the participants in light of a specific method given below. To solve the phenomenon of second-order free-riding, we introduce the Singaporean model into the public goods game. According to this rule, the leader's income consists of the basic payoff and performance salary. The former part is the same as the pure cooperators' payoff and the performance salary accounts for α of the final collective payoff, which serves as a reward for their exclusion behavior. It is determined and supported by the sanction pool. In the absence of corruption, we assume that benevolent leaders are strict and enact perfect exclusion, which means that defectors can get nothing as long as benevolent leaders exist in the group, and the final collective payoff is allocated among cooperators and benevolent leaders equally. Otherwise, the final collective payoff will be allocated evenly among all the cooperators and defectors in the group if benevolent leaders are absent.

To calculate the expected payoff of each kind of individual, we denote the frequencies of the cooperators, defectors and benevolent leaders as x_C, x_D and x_{BL} , respectively. Here, $x_C, x_D, x_{BL} \geq 0$ and $x_C + x_D + x_{BL} = 1$. Thus, the expected payoff of benevolent leaders can be given as follows:

$$P_{BL} = (1 + \alpha N)rc \sum_{N_{BL}=1}^N C_{N-1}^{N_{BL}-1} x_{BL}^{N_{BL}-1} x_C^{N-N_{BL}} + [1 + \alpha(N - N_D)]rc \sum_{N_D=1}^{N-1} \sum_{N_{BL}=1}^{N-N_D} C_{N-1}^{N_{BL}-1} C_{N-N_{BL}}^{N_D} x_{BL}^{N_{BL}-1} x_D^{N_D} x_C^{N-N_{BL}-N_D} - \mu - c \quad (1)$$

where N_{BL} denotes the number of benevolent leaders and N_D denotes the number of defectors. For the equation above, the first part is the sum of the basic payoff and the performance payoff in a group consisting of only benevolent leaders or both benevolent leaders and cooperators. The second part is the sum of the basic payoff and the performance payoff in a group consisting of both benevolent leaders and defectors or all the three kinds of players (for further theoretical details of this part, see Section 1.1 in the [Supplementary Information](#)).

Cooperators can get the payoff in four circumstances which include only cooperators, both cooperators and defectors, both cooperators and benevolent leaders, and all of the three kinds of players in the group. In detail, the expected payoff of cooperators can be given below:

$$P_C = \sum_{N_C=1}^N C_{N-1}^{N_C-1} x_C^{N_C-1} x_D^{N-N_C} \frac{rcN_C}{N} + \sum_{N_{BL}=1}^{N-1} \sum_{N_C=1}^{N-N_{BL}} C_{N-1}^{N_C-1} C_{N-N_{BL}}^{N_{BL}} x_C^{N_C-1} x_{BL}^{N_{BL}} x_D^{N-N_C-N_{BL}} rc - c \quad (2)$$

where N_C denotes the number of cooperators.

Defectors can get the payoff only in the scenario without benevolent leaders. The expected payoff of defectors can be given as below:

$$P_D = \sum_{N_C=0}^{N-1} C_{N-1}^{N_C} x_C^{N_C} x_D^{N-N_C-1} \frac{rcN_C}{N} \quad (3)$$

After some calculations, the expected payoff of benevolent leaders, cooperators and defectors can be simplified into the following forms:

$$P_{BL} = rc + \alpha[(x_C + x_{BL})(N-1) + 1]rc - \mu - c \quad (4)$$

$$P_C = rc - \frac{(x_C + x_D)^{N-2} x_D (N-1)}{N} rc - c \quad (5)$$

$$P_D = \frac{(x_C + x_D)^{N-2} x_C (N-1) rc}{N} \quad (6)$$

Consequently, the evolutionary dynamics can be analyzed by using the replicator equations (for further theoretical details of our model, see Section 1.2 in the [Supplementary Information](#)).

2.2. Results

2.2.1. Equilibria and stability analysis

In addition to three vertex fixed points, we explore the dynamics of the simplex S_3 and find that it may include some other equilibrium points in the system (for further theoretical details, see Section 1.3 in the [Supplementary Information](#)). On the edge of D-BL, there is one equilibrium point $(0, 1 - (\mu + c - rc - \alpha rc)/[\alpha(N-1)rc], (\mu + c - rc - \alpha rc)/[\alpha(N-1)rc])$ if $0 < \mu + c - rc - \alpha rc < \alpha(N-1)rc$. On the edge of C-BL, all the points can be equilibrium points if $\mu = \alpha Nrc$. We can get an interior equilibrium point

$$(x_C^o, x_D^o, x_{BL}^o) \text{ if } 0 < \frac{\alpha Nrc - \mu}{\alpha(N-1)rc - \left[\frac{r(N-1)}{N(r-1)}\right]^{\frac{1}{N-1}}(r-1)c} < \left[\frac{N(r-1)}{r(N-1)}\right]^{\frac{1}{N-1}} < 1. \text{ The values}$$

of x_{BL}^o, x_D^o and x_C^o are calculated below:

$$x_{BL}^o = 1 - \left[\frac{N(r-1)}{r(N-1)}\right]^{\frac{1}{N-1}} \quad (7)$$

$$x_D^o = \frac{\alpha Nrc - \mu}{\alpha(N-1)rc - \left[\frac{r(N-1)}{N(r-1)}\right]^{\frac{1}{N-1}}(r-1)c} \quad (8)$$

$$x_C^o = \left[\frac{N(r-1)}{r(N-1)}\right]^{\frac{1}{N-1}} - \frac{\alpha Nrc - \mu}{\alpha(N-1)rc - \left[\frac{r(N-1)}{N(r-1)}\right]^{\frac{1}{N-1}}(r-1)c} \quad (9)$$

We next investigate the evolutionary direction on the edge of the simplex S_3 . On the edge of C-D, the evolutionary direction of the dynamics goes from C to D. On the edge of C-BL, the evolutionary direction goes from C to BL if $\mu < \alpha Nrc$. In contrast, the direction goes from BL to C if $\mu > \alpha Nrc$. Specifically, all the points on C-BL stay still if $\mu = \alpha Nrc$. On the edge of D-BL, the direction goes from D to BL when $(\mu + c - rc - \alpha rc)/[\alpha(N-1)rc] \leq 0$ and from BL to D if $(\mu + c - rc - \alpha rc)/[\alpha(N-1)rc] \geq 1$. Otherwise, in the condition of $0 < (\mu + c - rc - \alpha rc)/[\alpha(N-1)rc] < 1$, the direction goes from BL to D for $x_{BL} \in (0, (\mu + c - rc - \alpha rc)/[\alpha(N-1)rc])$, and it goes from D to BL when $x_{BL} \in ((\mu + c - rc - \alpha rc)/[\alpha(N-1)rc], 1)$.

The equilibrium points are stable under certain conditions (for further theoretical details, see Section 1.5 in the [Supplementary Information](#)). For the fixed point $(0, 1, 0)$, the condition for this point to be stable is $rc + \alpha rc - \mu - c < 0$. For the fixed point $(0, 0, 1)$, the condition for this point to be stable is $\mu < \alpha Nrc$. All the points on the edge of C-BL are stable if $\mu = \alpha Nrc$. The stability of the interior equilibrium point (x_C^o, x_D^o, x_{BL}^o) depends on the specific settings of the various system parameters, and so it is hard to theoretically determine its stability. To explore the details of the evolutionary dynamics, we do some numerical simulations for the representative sets of system parameters in the next section.

2.2.2. Numerical simulations

The characteristics of the evolutionary dynamics in a public goods game with the benevolent leaders under the influence of the Singapore model are depicted in [Fig. 1](#). It is easy to see that,

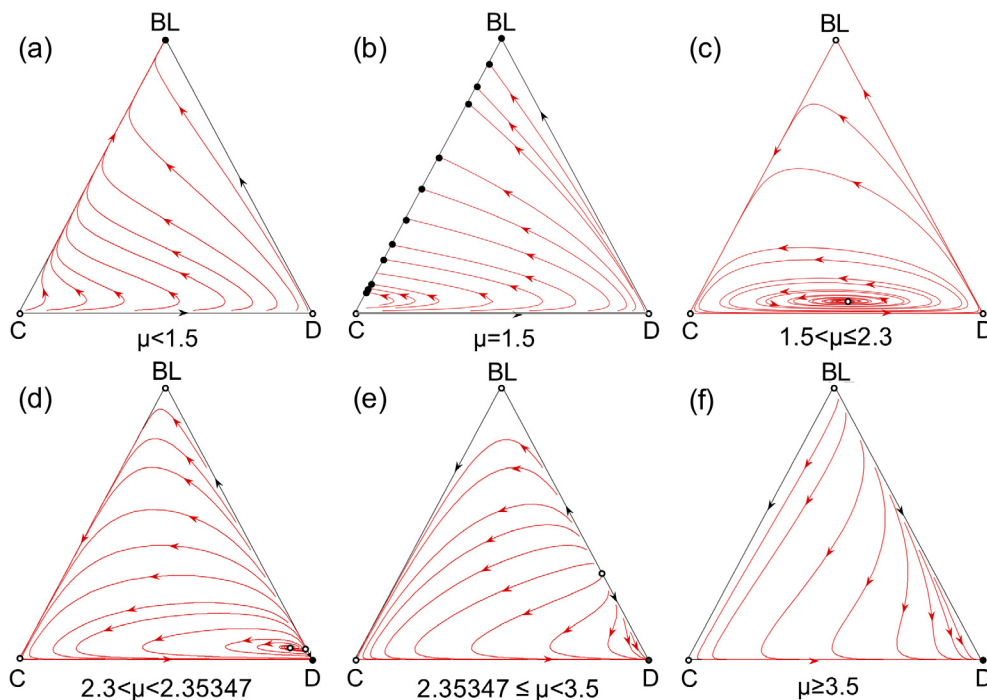


Fig. 1. Effects of benevolent leaders on cooperation for different exclusion costs μ with the low performance pay level $\alpha = 0.1$. The triangles represent the state space $S_3 = \{(x_C, x_D, x_{BL}) : x_C, x_D, x_{BL} \geq 0, \text{ and } x_C + x_D + x_{BL} = 1\}$, where x_C, x_D and x_{BL} are the probabilities of C, D and BL, respectively. Filled circles represent the stable states whereas open circles represent unstable states. The exclusion costs are (a) $\mu < 1.5$, (b) $\mu = 1.5$, (c) $1.5 < \mu \leq 2.3$, (d) $2.3 < \mu < 2.35347$, (e) $2.35347 \leq \mu < 3.5$ and (f) $\mu \geq 3.5$. The other parameters are the following: $N = 5$, $r = 3$ and $c = 1$.

for different exclusion costs μ , the patterns of the evolutionary trajectories are absolutely dissimilar. In general, as the exclusion cost increases, individuals imitate defectors instead of benevolent leaders. In detail, all-BL is the only absorbing state at the end of the evolution when $\mu < 1.5$. It is worth mentioning that the trajectories inside of S_3 evolve towards the edge of C-BL and then keep still after they reach that edge when $\mu = 1.5$. It means that both cooperators and benevolent leaders will coexist in the system (see Fig. 1(b)). The heteroclinic cycles ($\cdots \rightarrow C \rightarrow D \rightarrow BL \rightarrow C \rightarrow D \rightarrow BL \cdots$) are formed in the case of $1.5 < \mu \leq 2.3$ (see Fig. 1(c)). The reason is that defectors invade cooperators, benevolent leaders exclude defectors and cooperators dominate benevolent leaders (see Fig. 2(c)). The system tends to be all-D when $\mu > 2.3$ with the trajectories shown in Figs. 1(d)–(f). However, some details vary on the edge of D-BL. The equilibrium point on D-BL moves from D to BL as μ increases (see Fig. 1)) and disappears when $\mu \geq 3.5$ (see Fig. 1(f)). Note that there is an equilibrium point inside of S_3 when $1.5 < \mu < 2.35347$ (see Fig. 1(c) and (d)).

To clarify the particulars behind the dynamic characteristics depicted above, we continue to present the evolution of the frequencies of C, D and BL for various exclusion costs μ , which match six different dynamic characteristics, as presented in Fig. 1. It denotes that both C and BL invade D at first and then C is invaded by BL after D dies out in Fig. 2(a), which corresponds to the evolutionary direction on C-BL from C to BL, as presented in Fig. 1(a). In fact, cooperators can-not always invade defectors and it only exists at a certain stage of the evolutionary process. According to Fig. 1(a), the evolutionary path of the whole system is that D invades C and BL invades C and D at the same time when the frequency of BL is small, which results in an increase of BL and D and a decrease of C. However, with the increase of BL, the invasion of BL to D dominates the evolutionary process, the number of Ds decreases sharply, and then C invading D appearing in the system.

With the increase of μ , the ability of BL to invade C gradually declines and they can coexist when $\mu = 1.5$ as shown in Fig. 2(b).

The invasion of C to BL dominates the evolutionary process after $\mu > 1.5$ (see Figs. 2(c)–(f)). Note that C dominates BL, D dominates C and BL dominates D, which leads to a periodic oscillation when $1.5 < \mu \leq 2.3$, as shown in Fig. 2(c). As μ continues to increase, D can then invade BL, destroying the periodic oscillation, and the system eventually evolves into all-D when $\mu > 2.3$ (see Figs. 2(d)–(f)). From the viewpoint of the evolutionary processes, for the population with the same initial state, the peak frequency of C decreases and the valley frequency of D increases during the process as μ increases gradually, which means that D's ability to occupy the entire population is constantly improving and the whole evolutionary process gets worse. In general, both the final evolutionary steady state and the evolutionary process, before the steady state appears, deteriorate as the exclusion cost increases.

To further analyze the evolutionary dynamics in a high-performance payoff circumstance and test the robustness of the conclusions obtained above, Figs. 3 and 4 present the simulation results of the evolutionary dynamics when $\alpha = 0.18$. Compared with the low-performance payoff case, the condition that the evolutionary steady state is absolutely all-BL relaxes as shown for $\mu \leq 2.54$ in Fig. 3(a) instead of $\mu < 1.5$ in Fig. 1(a). The stable state of the coexistence of C and BL is raised from the original $\mu = 1.5$ to $\mu = 2.7$. The absolute all-D stable state appears when $\mu > 2.7$ instead of $\mu > 2.3$ (see Fig. 1(a) and Fig. 1(d)–(f)). There are multiple stable equilibria, and thus the initial conditions will determine which one will be reached as shown in Figs. 3(b)–(d). These results show that the improvement of the performance payoff α is conducive to promoting the individuals to greatly punish the defectors in the game. For details of the evolutionary process, the peak value of the frequency of C is gradually increased to 90.5% (see Fig. 4(a)–(d)) instead of 81.3% (see Fig. 2(a)–(b)) except for the special periodic oscillation (during this case, the peak frequency of C can reach 100%). The number of time steps until C vanishes is significantly extended when $\alpha = 0.18$ (e.g., the time until C vanishes > 100 time steps in Fig. 4(c) instead of < 30 time steps in Fig. 2(e)). Overall, the

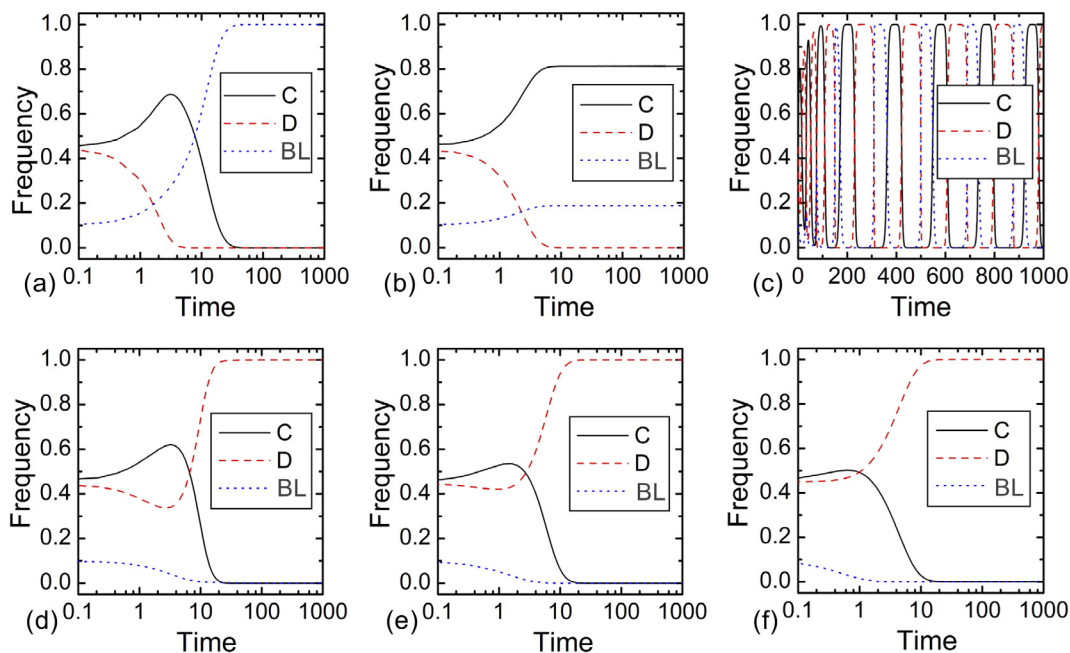


Fig. 2. Time evolution of the frequencies of C (black solid line), D (red dashed line) and BL (blue dotted line) as the exclusion cost μ increased. To explain the evolutionary characteristics of the different patterns in Fig. 1, here, (a)–(f) represent the effects of $\mu = 1.3, 1.5, 2.1, 2.32, 2.8$ and 3.9 on the evolutionary process, respectively, which match six different intervals of μ exactly in Fig. 1. The initial state of the population is $(x_C, x_D, x_{BL}) = (0.45, 0.45, 0.1)$. All the results are obtained for $N = 5, r = 3, c = 1$ and $\alpha = 0.1$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

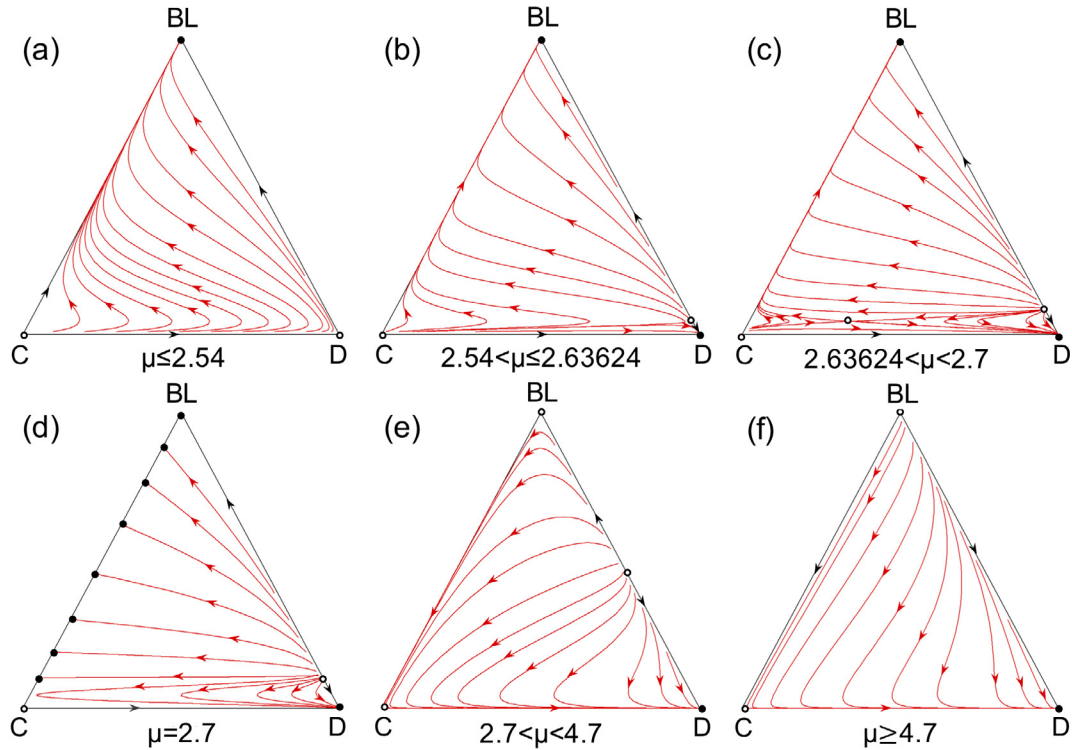


Fig. 3. Effects of benevolent leaders on cooperation for different exclusion costs μ with the high performance pay level $\alpha = 0.18$. The triangles represent the state space $S_3 = \{(x_C, x_D, x_{BL}) : x_C, x_D, x_{BL} \geq 0, \text{ and } x_C + x_D + x_{BL} = 1\}$, where x_C, x_D and x_{BL} are the probabilities of C, D and BL, respectively. Filled circles represent stable states whereas open circles represent unstable states. The exclusion costs are (a) $\mu \leq 2.6$, (b) $2 < \mu \leq 2.70694$, (c) $2.70694 < \mu < 3$, (d) $\mu = 3$, (e) $3 < \mu < 5$ and (f) $\mu \geq 5$. The other parameters are as follows: $N = 5$, $r = 3$ and $c = 1$.

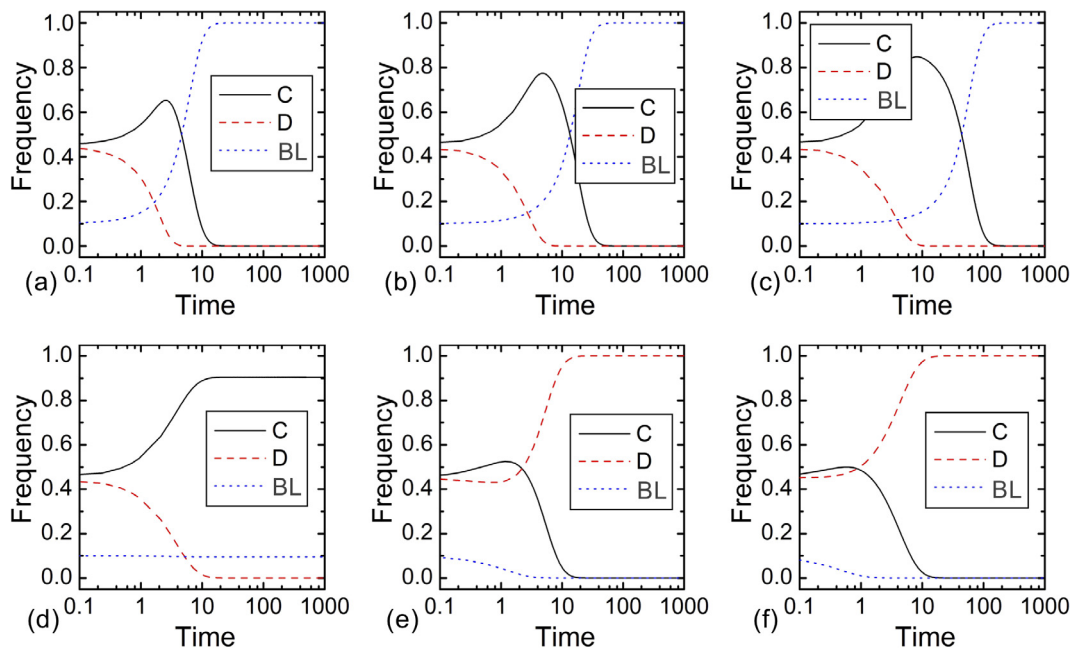


Fig. 4. Time evolution of the frequencies of C (black solid line), D (red dashed line) and BL (blue dotted line) as the exclusion cost μ increased. The same as in Fig. 2, here, (a)–(f) represent the effects of $\mu = 2.2, 2.6, 2.64, 2.7, 3.1$ and 4.9 on the evolutionary process, respectively, which match six different intervals of μ exactly in Fig. 3. The initial state of the population is $(x_C, x_D, x_{BL}) = (0.45, 0.45, 0.1)$. All the results are obtained for $N = 5$, $r = 3$, $c = 1$ and $\alpha = 0.18$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

increase of α is not only conducive to optimizing the final stable state, but it also has a significant positive effect during the evolutionary process.

3. Exclusion with bribery

3.1. Typical methods

Considering that the behavior of leaders in real society is usually affected by corruption, here we further introduce a new role, namely, the briber (B), into the model. As previously, the cooperator (C) and defector (D) adopt the same behavior as described in Section 2. However, as a new kind of participant, bribers contribute nothing to the public goods pool but they each donate δ to the corruption pool, which is used to commit bribery. Note that large differences exist between the leaders (L) in this game and the benevolent leaders in the first game under the effect of bribery. First, the compositions of their payoffs are different because the total amount of corruption will be distributed equally to all the leaders participating in this game. Second, the sanctioning behavior is also different. Here, leaders take probabilistic exclusion, depending on the degree of corruption, instead of perfect exclusion conducted by benevolent leaders, which is more in line with the actual situation. The frequencies of C, D, L and B are denoted by x_C, x_D, x_L and x_B , respectively, which satisfy $x_C + x_D + x_L + x_B = 1$ and $x_C, x_D, x_L, x_B \geq 0$. The probabilistic exclusion here is for both defectors and bribers by considering that the bribers donate to the sanction pool instead of bribing specific leaders directly. We define the probability that defectors are successfully excluded by leaders as

$$\rho = \frac{1}{1 + e^{\alpha_B N \delta - x_L N \mu}} \quad (10)$$

where N is the number of individuals in the public goods game and μ is the exclusion cost of leaders. It is easy to see that defectors will have a higher probability of being successfully excluded as the degree of corruption decreases and the exclusion cost increases. In particular, defectors will have the same chance of being successfully excluded or remaining ($\rho = 0.5$) if the corruption pool equals the sanction pool.

Then we can get the replicator equations of the system with cooperators, defectors, leaders and bribers based on the expected payoff of each kind of participant (for further theoretical details, see Section 2.2 in the [Supplementary Information](#)) as below:

$$\begin{cases} \frac{dx_C}{dt} = x_C [P_C - (x_C P_C + x_D P_D + x_L P_L + x_B P_B)] \\ \frac{dx_D}{dt} = x_D [P_D - (x_C P_C + x_D P_D + x_L P_L + x_B P_B)] \\ \frac{dx_L}{dt} = x_L [P_L - (x_C P_C + x_D P_D + x_L P_L + x_B P_B)] \\ \frac{dx_B}{dt} = x_B [P_B - (x_C P_C + x_D P_D + x_L P_L + x_B P_B)] \end{cases} \quad (11)$$

3.2. Results

3.2.1. Equilibria and stability analysis

We know that $P_B < P_D$ because the bribers pay a cost $\delta > 0$ to the corruption pool. However, the possibility of being excluded is the same. Therefore, there is no interior equilibrium in the simplex S_4 . The stability of the four vertex equilibria and the dynamics on the boundary faces of the simplex S_4 have been investigated (for further theoretical details, see Section 2.3 in the [Supplementary Information](#)). For vertex equilibria $(0, 1, 0, 0)$, the condition for this point to be stable is $2\alpha Nrc + Nrc - 2Nc - 2N\mu + rc < 0$. For vertex equilibria $(0, 0, 1, 0)$, the condition for this point to be stable is $\mu < \alpha Nrc$ and $\frac{(N-1)rc}{N} \left(1 - \frac{1}{1+e^{-N\mu}}\right) - \alpha Nrc - rc + c + \mu < 0$.

On the edge of C-D, the direction of the evolution goes from C to D. On the edge of C-L, the direction of the evolution goes from C to L if $\mu < \alpha Nrc$ and it goes from L to C if $\mu > \alpha Nrc$. Specifically, all the points on C-L stay still if $\mu = \alpha Nrc$. On the edge of C-B, the direction of the evolution goes from C to B if $\delta < c - \frac{rc}{N}$ and the direction goes from B to C if $\delta > c - \frac{rc}{N}$. Specifically, all the points on C-B stay still if $\delta = c - \frac{rc}{N}$. On the edge of D-B, the direction of the evolution goes from B to D because of $P_D > P_B$. On the edge of D-L and B-L, the direction of evolution is difficult to theoretically determine since it depends on the specific settings of various system parameters (for further theoretical details in this part, see Section 2.4 in the [Supplementary Information](#)). We will conduct some numerical simulations for the representative sets of system parameters to explore the details of the evolutionary dynamics in the next section.

3.2.2. Numerical simulations

The phase portrait of the model with corruption described in Section 3.2 has been shown in [Fig. 5](#). Specifically, [Fig. 5\(a\)](#) presents the evolutionary trajectories in simplex S_4 and [Fig. 5\(b\)](#) depicts the replicator dynamics on the four different boundary faces of S_4 . Four strategies, namely cooperators, defectors, leaders and bribers, correspond to the four fixed points C, D, L and B, respectively. It is easy

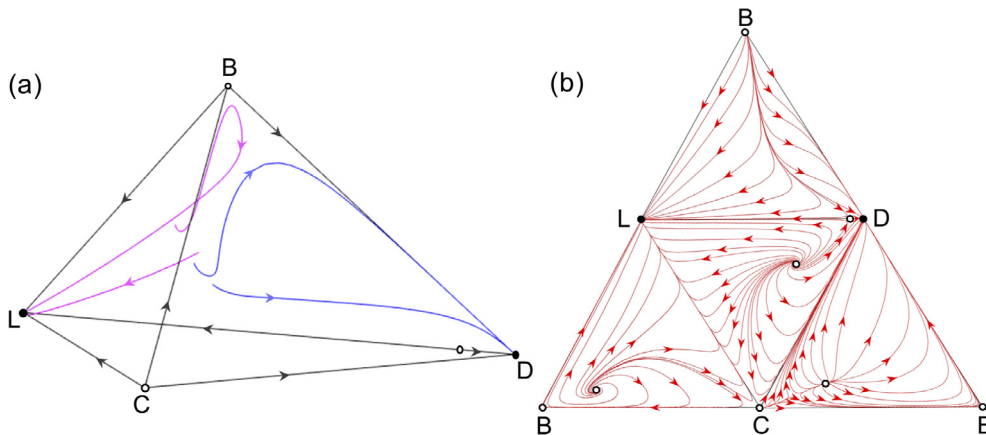


Fig. 5. Effects of corruption committed by bribers on cooperation. (a) Evolutionary trajectories in simplex S_4 , which contains cooperators (C), defectors (D), leaders (L) and bribers (B). Filled circles represent stable states and open circles are the opposite. (b) Replicator dynamics of the four different boundary faces of S_4 . All the results are obtained with the following parameters: $N = 5, r = 3, c = 1, \alpha = 0.1, \mu = 1.3$ and $\delta = 0.2$.

to see that L and D are both stable among these fixed points. The system will evolve to one or the other depending on the initial conditions. The game environments of Figs. 5 and 1(a) are similar except for the probabilistic exclusion by the introduction of bribers in Fig. 5. The final stable state in Fig. 1(a) is BL while the final stable state in Fig. 5 becomes L or D. This fully demonstrates that the existence of corruption has a significant degrading effect on the evolution of the system. The emergence of bribers in the population makes it possible for the system to evolve into all-D. Furthermore, it is interesting to note that as the frequency of bribers decreases, the system is more likely to evolve into an all-L state if there is no cooperator in the population because the separatrix on the boundary face of L-D-B shifts towards the direction of D from the up to bottom (it means that the frequency of B decreases) (see Fig. 5 (b)). However, bribers have no influence on the final stable point in the absence of defectors and L is the only stable state as shown in both Figs. 5(a) and (b).

To further analyze the effects of the initial states on the dynamics evolution, Fig. 6 presents the distribution areas of the final stable states D and L for different initial states of the population. Specifically, we use the ratio of B to L to indicate the level of corruption in the game, for which the ratio of B to L is increasing from (a) to (d) in Fig. 6. Overall, it is easy to see that the area covered by D is getting increasingly larger, which means that an increase in the corruption in the community will eventually lead to a higher possibility of an all-D appearing in the evolutionary stable state. Furthermore, there is a common feature in (a)–(b) of Fig. 6, which

is that L is the stable state in the lower-left area while D is the stable state near the hypotenuse. This result means that the fitness of D is significantly better than that of C when the total number of Cs and Ds in the initial population is overwhelming, and the system evolves into the All-D state. Otherwise, the stable state of the system is replaced by all-L instead of all-D if the frequency of B and L is relatively high in the system (lower-left area). Note that the unipolarity of the population structure (only either C or D in the initial population is a significant majority) is beneficial to suppressing the occurrence of an all-D stable state in a lowly corrupt society, as shown in Figs. 6(a)–(b) (the shape of the red area that is invaded by blue on the curved edge). There is a special area at the base of each figure in Fig. 6, at where the stable state of the evolution is always L regardless of the ratio of B to L. It shows that for the population with few defectors, corruption can lead to the emergence of a large number of leaders instead of defectors.

In addition to the ratio of bribers to leaders, the amount that each briber invests in the corruption pool (δ) can also reflect the extent of corruption. To understand the impact of δ on the population's evolution, we present the distribution of the stable state for $\delta = 0.05$ and $\delta = 0.5$ in Fig. 6, and that for $\delta = 0.2$ in Fig. 6(c). Different from the effect of the ratio of bribers to leaders on the evolutionary process, with the increase of δ , the evolutionary process evolves in a direction favoring an all-L state. Especially, this kind of impact is most obvious for the population with plenty of defectors, as presented by the red area rising rapidly near the vertical axis. Therefore, the input cost of bribers mainly affects the income

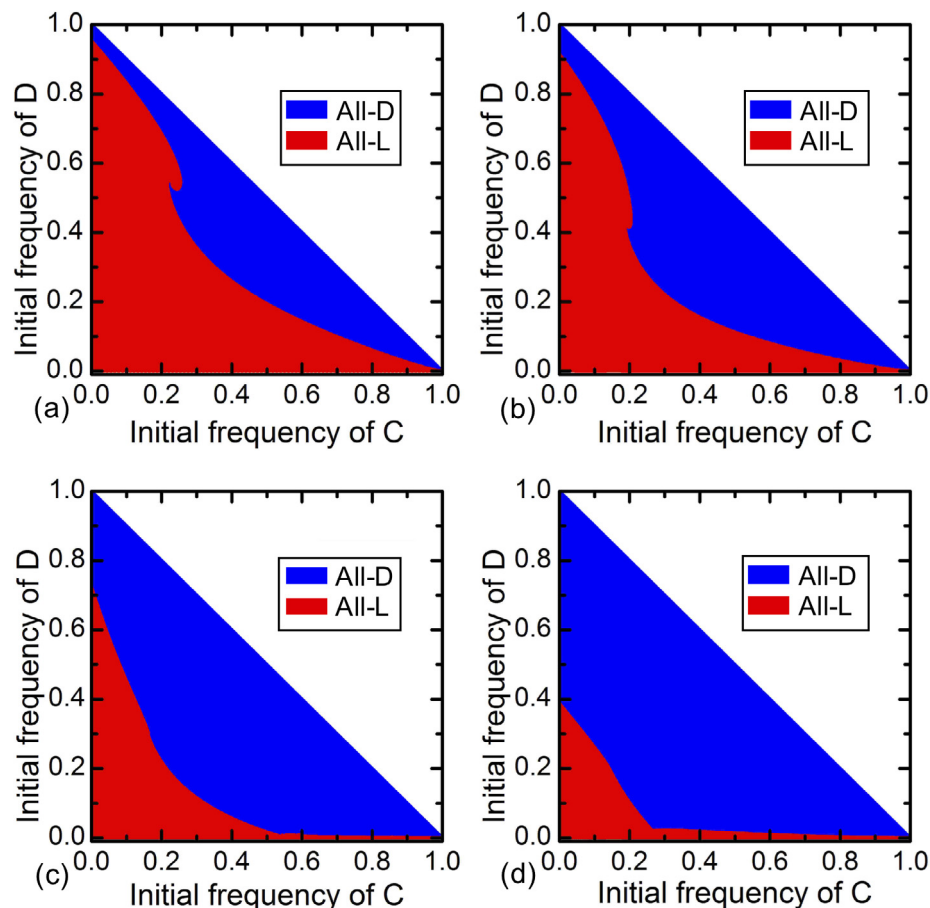


Fig. 6. The effect of the initial population structure on cooperation. The initial ratio of B to L in the population is 0.1, 0.5, 1 and 2 from (a) to (d), respectively. The blue color indicates that the final stable state of evolution is the defector and the red color indicates the final stable state is Leader. All the results are obtained with the following parameters: $N = 5$, $r = 3$, $c = 1$, $\alpha = 0.1$, $\mu = 1.3$ and $\delta = 0.2$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

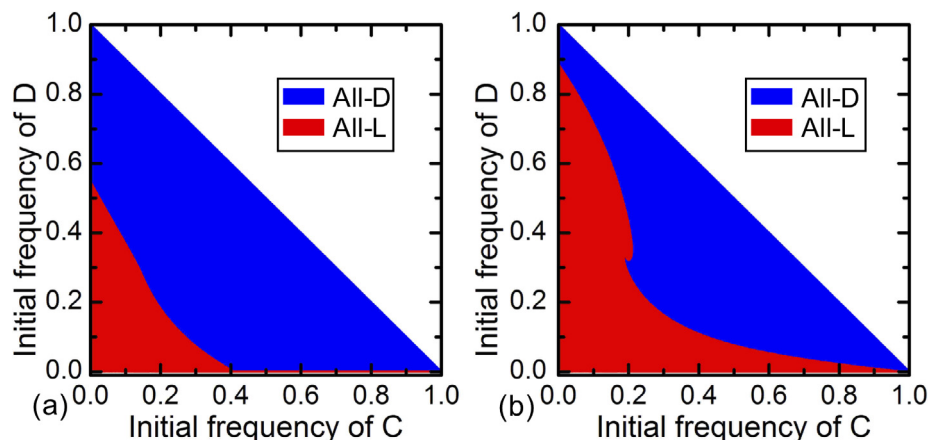


Fig. 7. The effect of the bribery cost on cooperation. For comparison, both low corruption ($\delta = 0.05$) and high corruption ($\delta = 0.5$) are considered in (a) and (b), respectively. The frequencies of leaders and bribers are the same as in the initial state. The blue color indicates that the final stable state of the evolution is defector and the red color indicates that the final stable state is Leader. All the results are obtained with the following parameters: $N = 5$, $r = 3$, $c = 1$, $\alpha = 0.1$ and $\mu = 1.3$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of leaders. In particular, the greater the briber invests, the greater the income of the leader, and the defectors have a lower advantage in the population's evolutionary process (see Fig. 7).

4. Discussion

We have studied the impact of the Singaporean model on cooperation in the public goods game with benevolent leaders and bribery. According to the model (Yew, 2000; Muthukrishna et al., 2017), the payoffs of leaders were determined not only by the success of the public goods game in their group, which is the traditional setup, but also by the average payoff of all the players in the population. This mimics the policy in Singapore, taking into account the success of the firm and the gross domestic product of the state to determine salary.

First, we considered cooperators, defectors, and benevolent leaders as the three competing strategies, thus not considering bribery. Therein, benevolent leaders were willing to pay a cost to exclude defectors from using the public goods. Second, we also introduced bribers as the fourth competing strategy, which led to the exclusion by benevolent leaders going from perfect to probabilistic. The more common or the stronger the bribes were, the lesser the probability of defector exclusion by benevolent leaders. We have studied the evolutionary dynamics in both the 3-strategy and the 4-strategy public goods game by means of using the replicator equation to determine the equilibria and their stability, as well as performing numerical simulations to determine all possible evolutionary outcomes.

For the 3-strategy game, we have observed that, as the cost of exclusion increases, the benevolent leaders became weaker, which in turn strengthened the position of defectors. Specifically, the evolutionary stable state has undergone a transition from an all-BL \rightarrow BL + C \rightarrow BL + C + D \rightarrow all-D for low performance incentives related to the expected collective payoff of the group or from an all-BL \rightarrow BL + D \rightarrow BL + C + D \rightarrow all-D for high performance incentives related to the expected collective payoff of the group. For example, the threshold of the exclusion cost to observe an all-BL state increased from 1.5 to 2 as the incentives related to the collective payoff of the population increased from 0.1 to 0.15. More importantly, we have observed that for sufficiently strong incentives related to the collective payoff of the population, the second-order free-riders problem can be resolved. For the 4-strategy game, which included bribers that render the exclusion of benevolent leaders probabilis-

tic rather than perfect, we have observed that increasing the ratio of bribers to leaders can lead to an all-D steady state. However, the increase of bribery costs is conducive to the evolution of an all-L steady state. Therefore, based on these two rules, we can obtain a strategy chain to restrain defectors, which is high bribery costs \rightarrow high ratio of L \rightarrow low ratio of bribers to leaders \rightarrow inhibit defector. No matter whether or not corruption exists in the game, all-C (those who donate nothing to the sanction pool) can never be the stable state, thus confirming that the Singaporean model is a viable option to avoid the second-order free-rider problem.

Our results support the narrative that healthy levels of cooperation can be achieved through positive incentives rather than punishment. Although punishment has received substantially more attention in the past, it is worth noting that recent research related to antisocial punishment (Rand and Nowak, 2011) and rewarding (Rand et al., 2009; Szolnoki and Perc, 2012) is questioning the aptness of sanctioning for elevating collaborative efforts and raising social welfare. Indeed, while the majority of previous theoretical studies addressing the stick-versus-carrot dilemma concluded that punishment may be more effective than reward in promoting public cooperation (Perc et al., 2017), evidence suggesting that rewards may be as effective as punishment and lead to higher total earnings are accumulating (Dreber et al., 2008; Szolnoki and Perc, 2015; Hu et al., 2020). Linking the Singaporean model with the public goods game amounts to a positive incentive, in that the success of the firm and the gross domestic product of the state both determine the payoff. However, if the state or the firm start failing, the decrease in payoff can also become a negative incentive, i.e., a punishment of sorts, such that in fact the Singaporean model incorporates both in a simple yet effective manner. Our approach thus invites the consideration of doubly strategies that flip incentives based on performance, which in turn suggests that coevolutionary rules (Perc and Szolnoki, 2010) to rewarding and punishment might be a viable alternative that is worth exploring more in the future.

We conclude by noting that successful cooperation and anti-corruption strategies in social dilemma situations are crucial for mitigating some of the most pressing challenges of our time, with examples ranging from improved vaccination policies (Fu et al., 2011; Wang et al., 2016) to climate inaction (Vasconcelos et al., 2013; Pacheco et al., 2014). Our research shows that the aspect of overall welfare for individual performance is crucial for a closed loop of responsibility, which can ultimately lead to healthier levels of public cooperation. In this regard, Singapore is a good role model

due to its policies in which the gross domestic product is a determinant of individual salaries. Future research should address the impact of such policies in structured populations, and in the presence of other policies aimed at improving social welfare.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Yinhai Fang: Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Writing - original draft, Writing - review & editing. **Matjaž Perc:** Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Writing - original draft, Writing - review & editing. **Haiyan Xu:** Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Writing - original draft, Writing - review & editing.

Acknowledgments

Yinhai Fang was supported by China Association for Science and Technology (2018CASTQNJL23) and Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX18_0237). Matjaž Perc was supported by the Slovenian Research Agency (Grant Nos. J4-9302, J1-9112, and P1-0403). Haiyan Xu was supported by National Natural Science Foundation of China (71971115).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.jtbi.2020.110345>.

References

- Andreoni, J., Harbaugh, W., Vesterlund, L., 2003. The carrot or the stick: rewards, punishments, and cooperation. *Am. Econ. Rev.* 93, 893–902.
- Axelrod, R., 1984. *The Evolution of Cooperation*. Basic Books, New York.
- Chen, X., Fu, F., 2018. Social learning of prescribing behavior can promote population optimum of antibiotic use. *Front. Phys.* 6, 193.
- Chen, X., Sasaki, T., Brännström, Å., Dieckmann, U., 2014. First carrot, then stick: how the adaptive hybridization of incentives promotes cooperation. *J. R. Soc. Interface* 12, 20140935.
- Chen, X., Szolnoki, A., Perc, M., 2014. Probabilistic sharing solves the problem of costly punishment. *New J. Phys.* 16, 083016.
- dos Santos, M., 2015. The evolution of anti-social rewarding and its countermeasures in public goods games. *Proc. R. Soc. B* 282, 20141994.
- Dreber, A., Rand, D.G., Fudenberg, D., Nowak, M.A., 2008. Winners don't punish. *Nature* 452, 348–351.
- Fehr, E., 2004. Don't lose your reputation. *Nature* 432, 449–450.
- Fehr, E., Fischbacher, U., 2004. Social norms and human cooperation. *Trends Cogn. Sci.* 8, 784–790.
- Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods experiments. *Am. Econ. Rev.* 90, 980–994.
- Fu, F., Rosenbloom, D.I., Wang, L., Nowak, M.A., 2011. Imitation dynamics of vaccination behaviour on social networks. *Proc. R. Soc. B* 278, 42–49.
- Gächter, S., Renner, E., Sefton, M., 2008. The long-run benefits of punishment. *Science* 322, 1510.
- Hauert, C., Wakano, J.Y., Doebeli, M., 2008. Ecological public goods games: cooperation and bifurcation. *Theor. Popul. Biol.* 73, 257–263.
- Helbing, D., Szolnoki, A., Perc, M., Szabó, G., 2010. Evolutionary establishment of moral and double moral standards through spatial interactions. *PLoS Comput. Biol.* 6, e1000758.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., 2001. In search of homo economicus: behavioral experiments in 15 small-scale societies. *Am. Econ. Rev.* 91, 73–78.
- Hu, L., He, N., Weng, Q., Chen, X., Perc, M., 2020. Rewarding endowments lead to a win-win in the evolution of public cooperation and the accumulation of common resources. *Chaos Solitons Fract.* 134, 109694.
- Jiao, Y., Chen, T., Chen, Q., 2020. Probabilistic punishment and reward under rule of trust-based decision-making in continuous public goods game. *J. Theor. Biol.* 486, 110103.
- Jordan, J.J., Hoffman, M., Bloom, P., Rand, D.G., 2016. Third-party punishment as a costly signal of trustworthiness. *Nature* 530, 473–476.
- Liu, L., Chen, X., Perc, M., 2019. Evolutionary dynamics of cooperation in the public goods game with pool exclusion strategies. *Nonlinear Dyn.* 97, 749–766.
- Milinski, M., Rockenbach, B., 2008. Punisher pays. *Nature* 452, 297–298.
- Muthukrishna, M., Francois, P., Pourahmadi, S., Henrich, J., 2017. Corrupting cooperation and how anti-corruption strategies may backfire. *Nat. Human Behav.* 1, 0138.
- Nowak, M.A., Highfield, R., 2011. *Supercooperators: Altruism, Evolution, and Why We Need Each Other to Succeed*. Free Press, New York.
- Okada, I., Yamamoto, H., Toriumi, F., Sasaki, T., 2015. The effect of incentives and meta-incentives on the evolution of cooperation. *PLoS Comput. Biol.* 11, e1004232.
- Pacheco, J.M., Vasconcelos, V.V., Santos, F.C., 2014. Climate change governance, cooperation and self-organization. *Phys. Life Rev.* 11, 573–586.
- Panchanathan, K., Boyd, R., 2004. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature* 432, 499–502.
- Perc, M., Jordan, J.J., Rand, D.G., Wang, Z., Boccaletti, S., Szolnoki, A., 2017. Statistical physics of human cooperation. *Phys. Rep.* 687, 1–51.
- Perc, M., Szolnoki, A., 2010. Coevolutionary games – a mini review. *BioSystems* 99, 109–125.
- Rand, D.G., Dreber, A., Ellingsen, T., Fudenberg, D., Nowak, M.A., 2009. Positive interactions promote public cooperation. *Science* 325, 1272–1275.
- Rand, D.G., Nowak, M.A., 2011. The evolution of antisocial punishment in optional public goods games. *Nat. Commun.* 2, 434.
- Rand, D.G., Nowak, M.A., 2013. Human cooperation. *Trends Cogn. Sci.* 17, 413–425.
- Santos, F.C., Santos, M.D., Pacheco, J.M., 2008. Social diversity promotes the emergence of cooperation in public goods games. *Nature* 454, 213–216.
- Sasaki, T., Uchida, S., 2014. Rewards and the evolution of cooperation in public good games. *Biol. Lett.* 10, 20130903.
- Sasaki, T., Uchida, S., Chen, X., 2015. Voluntary rewards mediate the evolution of pool punishment for maintaining public goods in large populations. *Sci. Rep.* 5, 8917.
- Sasaki, T., Unemi, T., 2011. Replicator dynamics in public goods games with reward funds. *J. Theor. Biol.* 287, 109–114.
- Sekiguchi, T., Nakamaru, M., 2009. Effect of the presence of empty sites on the evolution of cooperation by costly punishment in spatial games. *J. Theor. Biol.* 256, 297–304.
- Szolnoki, A., Perc, M., 2010. Reward and cooperation in the spatial public goods game. *EPL* 92, 38003.
- Szolnoki, A., Perc, M., 2012. Evolutionary advantages of adaptive rewarding. *New J. Phys.* 14, 093016.
- Szolnoki, A., Perc, M., 2015. Antisocial pool rewarding does not deter public cooperation. *Proc. R. Soc. B* 282, 20151975.
- Szolnoki, A., Perc, M., 2017. Second-order free-riding on antisocial punishment restores the effectiveness of prosocial punishment. *Phys. Rev. X* 7, 041027.
- Vasconcelos, V.V., Santos, F.C., Pacheco, J.M., 2013. A bottom-up institutional approach to cooperative governance of risky commons. *Nat. Clim. Chang.* 3, 797–801.
- Wang, Z., Bauch, C.T., Bhattacharyya, S., d'Onofrio, A., Manfredi, P., Perc, M., Perra, N., Salathé, M., Zhao, D., 2016. Statistical physics of vaccination. *Phys. Rep.* 664, 1–113.
- Yamagishi, T., 1986. The provision of a sanctioning system as a public good. *J. Pers. Soc. Psychol.* 51, 110–116.
- Yew, L.K., 2000. *From Third World to First: The Singapore Story: 1965–2000*. Harper, New York.