International Institute for
Applied Systems Analysis
Schlossplatz 1
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342
Fax: +43 2236 71313
E-mail: publications@iiasa.ac.at
Web: www.iiasa.ac.at

Interim Report **IR-13-077**

Mixed strategy under generalized public goods games

Yanling Zhang
Te Wu
Xiaojie Chen (chenx@iiasa.ac.at)
Guangming Xie
Long Wang

Approved by

Ulf Dieckmann
Director, Evolution and Ecology Program

June 2015

Mixed strategy under nonlinear public goods games

Yanling Zhang[a], Te Wu[a], Xiaojie Chen[b], Guangming Xie[a], Long Wang[a]

[a] Center for Systems and Control, State Key Laboratory for Turbulence and Complex Systems, College of Engineering, Peking University, Beijing 100871, China

[b] Evolution and Ecology Program, International Institute for Applied Systems Analysis, Laxenburg, Austria

Classifycation: BIOLOGICAL SCIENCES – Population Biology

Manuscript information: 26 pages (including figure legend and appendix); 4 Figures;

Corresponding authors:
Yanling Zhang, Guangming Xie
Center for Systems and Control, State Key Laboratory for Turbulence and Complex Systems, College of Engineering, Peking University
No. 5 Yiheyuan Road Haidian District
Beijing 100871, China
Email: yanlzhang@pku.edu.cn, guangming xie@pku.edu.cn

Abstract: The relationship between public goods benefiting all group members irrespective of their contribution and the proportion of cooperators in the group, modeled as the production function, often exhibits nonlinearity. Previous nonlinear game theoretical models assume two discrete strategies or continuous investments. Here, we investigate the effect of nonlinear production functions on the evolution of the mixed strategy in finite populations. When the group size and population size become comparable, cooperation is doomed irrespective of the production function. Otherwise, nonlinear production functions may lead to a convergent evolutionary stable strategy (CESS) or a repeller, but can't yield the evolutionary branching in contrast with the evolution of continuous investments. In particular, we consider three representative families of production functions, in which most previous studied production functions are recovered as special cases. Full defection, full cooperation, a unique CESS or a unique repeller may occur for two families of production functions including concave and convex curves even if the group size is two. However, the parameter region of each evolutionary situation exhibits a great difference since the value of public goods produced is fixed for one class and is variable for the other when all group members cooperate. A third class encompassing symmetrically sigmoidal and inverse sigmoidal curves may lead to the coexistence of a CESS and a repeller impossible for the other two classes only when group size exceeds two. Intriguingly, the hysteresis effect is found in all these three classes. However, two saddle-node bifurcations appear for the third class but not for the other two classes.

Key words: public goods game, nonlinear production function, mixed strategy, adaptive dynamics

I. INTRODUCTION

Public goods dilemma is abundant ranging from bacteria to high advanced animal to human being [1–5], which describes a situation in which several players can cooperate to achieve a public good. A public good is a resource from which all may benefit, regardless of whether they have helped provide the good, such as grassland, environmental resources and peace. Public goods games have been widely used to model and elucidate the evolution of cooperation which is an enduring evolutionary intractable problem [6–20]. All group members are provided an identical endowment (*b*) that can be kept or invested in a pool and simultaneously decide to cooperate (invest endowment) or not (keep endowment). The key characteristic of public goods games is the relationship between the level of resources contributed toward the production of the public good and the level of the public good that is provided. This relationship is known as the production function [21, 22]. If all members have cooperated, the group's payoff is maximized, yet to defect is always better for a given player irrespective of the group's composition since the public good is nonexclusive and the contribution is costly. The social dilemmas result as the best strategy for players and that for a group do not coincide.

As in the linear public goods game, the per capita benefit is assumed to increase in a linear fashion with the number of cooperators, i.e., each unit of resource contributed produces the same return [6–13]. Some previous studies have relaxed this assumption and investigated the threshold production function in which each player has a benefit if and only if the number of cooperators is above a threshold, otherwise he gets nothing [14–16]. That is to say that little or no amount of the public good is produced until a certain level of contributions is reached, at which point a small increase in the level of contributions returns a large and discontinuous amount of the public good. In various natural situations, the threshold production function, let alone the linear production function, sometimes fails to capture the nature of the production of the public good. Particularly, both synergy allowing per capita public good produced to increase and discounting characterizing decreasing pubic good produced with increasing proportion of cooperators are prevalent from pheromone trails to human architecture [23–30]. The synergy can be represented by the convex function, the first part of the sigmoidal function and the second part of the inverse sigmoidal function. Yet the concave function, the second part of the sigmoidal function and the first part of the inverse sigmoidal function perform the discounting. In economics, the standard production function is assumed as an S-shaped curve [31], which accelerates at the beginning of production and then decelerates. The initial acceleration reflects the increasing marginal returns (i.e., synergy) from each additional initial contribution, while the ensuing deceleration reflects the decreasing marginal returns (i.e., discounting) when output limits are approached.

Besides linear and threshold production functions, concave, convex and sigmoidal production functions have been investigated in evolutionary game theory [17, 18]. However all these studies assume a priori that the amount of the investment takes two discrete values (i.e., all or nothing) or varies continuously within a certain range. In real life, whether to contribute to the production of public goods may not be either of the above two cases but rather probabilistically change according to circumstances and risk orientation. Consider voting in which cooperative participants vote for the proposal yielding the public good benefiting all group members and bear some costs, yet defectors, on the contrary, vote against such proposal and bear no costs. However, in practice, few people vote for such proposal every time they make a vote, partly due to the maxima 'don't put eggs in a basket'. In addition, most of previous studies upon nonlinearity just proceed in infinitely large populations [14, 17–20]. Obviously, the realistic systems exhibit finiteness of populations. Although there have been already attempts to study the dynamics of public goods games in finite populations which assume two discrete strategies or proceed in two person games [15, 16, 35], the adaptive dynamics of public goods games in finite populations deserve further and intensive investigations.

In this paper, we propose a general framework how nonlinear production functions affect the adaptive dynamics of the mixed strategy in finite populations where the mixed strategy is the probability that players invest all and otherwise invest nothing. The adaptive dynamics are widely used to study the long-term evolution of continuous strategy which say that the evolution on average takes the population up the gradient of 'invasion fitness'. 'Invasion fitness' has frequently been assumed to be the payoff of a single mutant, suggesting the population size is infinitely large [19, 20, 32, 33]. Such assumption upon "invasion fitness" has been theoretically confirmed for a frequency independent process [34]. In finite populations, researchers have argued that it is the fixation probability rather than the payoff that carries the important information for evolution [35–37]. In particular, we consider two classes of production functions consisting of concave and convex curves, one with the same value of public goods produced and the other with various values of public goods produced when all group members cooperate. In addition, a class of production functions including sigmoidal and inverse sigmoidal curves are also considered, which produce an identical public good when all group members cooperate. Note that some threshold production functions appear in them as special cases.

II. MODEL

Consider a well-mixed population of size N, in which each individual is endowed with 1 and adopts a mixed strategy x $(1 \leq x \leq 1)$. The strategy x represents the probability that a player cooperates, i.e., invests all his endowment 1 and the initial endowment added to his payoff is 0, or else he defects, i.e., invests nothing and and the initial endowment added to his payoff is c. Sampling of individuals playing games follows a hypergeometric distribution and the average payoff of individuals of strategy z in a population consisting of X individuals of strategy w and $N - X$ individuals of strategy z, $f(z, w, N - X)$ (see the payoff in Appendix), is

$$\begin{aligned} f(z, w, N - X) \;=\; & \sum_{i=0}^{n-1} \frac{\binom{N-X-1}{i}\binom{X}{n-1-i}}{\binom{N-1}{n-1}} \sum_{k=0}^{i} \sum_{l=0}^{n-1-i} \binom{i}{k}\binom{n-1-i}{l} \\ & z^k(1-z)^{i-k} w^l (1-w)^{n-1-i-l} \Big[zg(\tfrac{k+l+1}{n}) + (1-z)[g(\tfrac{k+l}{n}) + c] \Big] \end{aligned}$$

where n is the group size and $g(\theta)$ is the production function describing the relationship between the proportion of cooperators in a group and the public good produced. Clearly, $g(\frac{k}{n})$ is the benefit to each member if k players apt to cooperate and $n - k$ defect. The production function is $g(\theta) = r\theta$ in [6–13], $rf(\theta - \lambda)$ in [14, 16] and $r\theta f(\theta - \lambda)$ in [15] where r is the enhancement factor, $f(x)$ satisfies $f(x < 0) = 0$ and $f(x \geq 0) = 1$ and λ is the threshold point. Other relative studies assume $g(\theta) = \frac{r}{n}(1 + w + w^2 + \cdots + w^{n\theta-1})$ in [17] and $g(\theta) = r\frac{e^{\frac{n}{2}k}+1}{e^{\frac{n}{2}k}-1}\frac{1}{e^{-kn(\theta-0.5)}+1} - \frac{1}{e^{\frac{n}{2}k}-1}$ in [18], where w and k characterize the shape of production functions. All the above production functions, without exception, are increasing and satisfy $g(0) = 0$, which are also followed in our model.

III. RESULTS

Following the method of the derivation in [34] (see derivation in Appendix), the first-order deterministic approximation for the mean path of x is, $\frac{\mathrm{d}}{\mathrm{d}t}x(t) = C\frac{\partial}{\partial y}\Big|_{y=x}\rho(x; y)$, where $\rho(x; y) = [\sum_{k=0}^{N-1}\prod_{a=1}^{k}\frac{f(x,y,N-a)}{f(y,x,a)}]^{-1}$ is the fixation probability of a mutant with strategy y in a resident population with strategy x. Note throughout this paper $\frac{\partial}{\partial y}\Big|_{y=x}\rho(x; y)$ means that we first calculate the partial derivative of $\rho(x; y)$ with respect to y and then replace y with x. Since C is a constant for the evolutionary process and just scales time, the adaptive dynamics can be reduced to

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = \frac{\partial}{\partial y}\Big|_{y=x}\rho(x; y) \tag{1}$$

Substituting $\rho(x;y)$ into Eq.1 determines the evolution of the mixed strategy over evolutionary time (see calculation in Appendix):

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = G(x) = \frac{[(N-n)C(x)-(N-1)c]}{2Nf(x)}, \tag{2}$$

where $C(x) = \sum_{j=0}^{n-1}\binom{n-1}{j}x^j(1-x)^{n-1-j}\left[g(\frac{j+1}{n}) - g(\frac{j}{n})\right]$ and $f(x)$ is the payoff of x in a monomorphic population with x and is always positive in our model. As $N \to +\infty$, $G(x) = \frac{(C(x)-c)}{2f(x)} = \frac{1}{2f(x)}\frac{\partial}{\partial y}\Big|_{y=x} f(y,x,1)$, so our conclusion qualitatively recovers the dynamics used in [19, 20, 32, 33].

Clearly the evolution of the mixed strategy x is related with the global information of the production function $g(\theta)$. If $(N-n)C(x) > (N-1)c$, selection favors mutants with $y > x$ replacing x (i.e., $\rho(x;y) > 1/N$ with $y > x$), whereas if $(N-n)C(x) < (N-1)c$, this is only possible for mutants with $y < x$ (i.e., $\rho(x;y) > 1/N$ with $y < x$) [38]. In particular, whenever the group size equals the population size, $n = N$, the adaptive dynamics lead to the demise of cooperation irrespective of the shape of the production functions, which contrasts with the results in larger populations addressed later. A strategy satisfying $(N-n)C(x) = (N-1)c$ (i.e., $\rho(y;x) = 1/N$) is termed a singular strategy whose evolutionary direction is uncertain and further close investigations upon $G'(x)$ and $\frac{\partial\rho^2(x;y)}{\partial y^2}\Big|_{y=x}$ are needed [39], where $G'(x)$ is the derivative of $G(x)$ with respect to x.

At singular points, we have (see calculation in Appendix)

$$G'(x) = \frac{(n-1)D(x)}{Nf(x)}, \qquad \frac{\partial\rho^2(x;y)}{\partial y^2}\Big|_{y=x} = \frac{(n-1)D(x)}{3Nf(x)}, \tag{3}$$

where $D(x) = \sum_{j=0}^{n-2}\binom{n-2}{j}x^j(1-x)^{n-2-j}\left[g(\frac{j+2}{n}) - 2g(\frac{j+1}{n}) + g(\frac{j}{n})\right]$. In principle, the evolution of the mixed strategy opens up the possibility of the evolutionary branching requiring $G'(x) < 0$ and $\frac{\partial\rho^2(x;y)}{\partial y^2}\Big|_{y=x} > 0$, which clearly never appears in such adaptive dynamics. The singular strategy satisfying $D(x) < 0$ is a convergent evolutionary stable strategy (CESS or called as h-stable in [33]) since $G'(x) < 0$ and $\frac{\partial\rho^2(x;y)}{\partial y^2}\Big|_{y=x} < 0$, where once it has become established in a population, no further evolutionary change is possible for small mutations. If $D(x) > 0$ holds at the singular strategy, the singular strategy is a repeller (called as x-stable in [33]) since $G'(x) > 0$, leading bistable dynamics. In the case of $G'(x) = 0$, the first-order deterministic approximation fails to tell the evolutionary evolution of such point which can be determined by the third-order deterministic approximation. In this case, it's just like the adaptive dynamics yield no singular points (see calculation in Appendix).

A. Convex and concave production functions

If $g(\theta)$ is a convex function which means $D(x) > 0$, $G(x) = 0$ has at most one root x^* which satisfies $D(x^*) > 0$, thus there exists at most one singular strategy which, if it exists, is a repeller. However concave $g(\theta)$ leads to $D(x) < 0$ and the monotonicity of $G(x)$, hence there exists at most one singular strategy which, if it exists, is a CESS. In particular, we will consider two classes of production functions consisting of concave and convex curves.

Firstly, a family of production functions $\phi_m(\theta) = r\theta^m$ with $m > 0$ controlling the shape of production functions (Fig. 1a) are investigated. When $m = 1$, $\phi_m(\theta)$ is a linear production function which is probed as a benchmark. When $m > 1$, $\phi_m(\theta)$ is a convex curve where the amount of the public good produced by each additional cooperator is higher than the previous one. When $0 < m < 1$, $\phi_m(\theta)$ is a concave curve in which each additional cooperator produces decreasing public goods in contrast with the case of $m > 1$. Particularly when $m \to 0$ or $m \to +\infty$, $\phi_m(\theta)$ is a threshold function at $\theta = 0$ or $\theta = 1$.

Assume $m_1^* = \frac{\ln \frac{c(N-1)}{r(N-n)}}{\ln \frac{1}{n}}$ and $m_2^* = \frac{\ln(1-\frac{c(N-1)}{r(N-n)})}{\ln \frac{n-1}{n}}$ and a complete classification of the adaptive dynamics for $\phi_m(\theta)$ is provided as follows (see analysis in Appendix):

- The cooperative probability monotonically decreases to full defection (figure 2a) if $m \in [m_1^*, m_2^*]$ and $\frac{c}{r} > \frac{N-n}{n(N-1)}$. Moderate m and large ratio of the initial remaining endowment induced by defection and the enhancement factor, $\frac{c}{r}$, commonly lead full defection to evolve in the initially uniform populations. The parameter region of m varies with $\frac{c}{r}$, the population size N and the group size n. Increasing $\frac{c}{r}$ (figure 1b) or increasing n (figure 1d) expands the region, whereas increasing N (figure 1c) shrinks the region.

- The cooperative probability monotonically increases to full cooperation (figure 2b) if $m \in [m_2^*, m_1^*]$ and $\frac{c}{r} < \frac{N-n}{n(N-1)}$. Moderate m and small $\frac{c}{r}$ together induce full cooperation to evolve in the initially uniform populations. In contrast with the region of m favoring full defection, the parameter region of m favoring full cooperation wanes with growing $\frac{c}{r}$ (figure 1b) or growing n (figure 1d), yet waxes with growing N (figure 1c).

- Under the condition of $m > m_1^*$ and $\frac{c}{r} \leq \frac{N-n}{n(N-1)}$ or the condition of $m > m_2^*$ and $\frac{c}{r} \geq \frac{N-n}{n(N-1)}$, a unique repeller occurs in the adaptive dynamics (figure 2c). A repeller leads to bistable dynamics, where the initially uniform population finally converges to full cooperation or full defection depending on the initial condition. The existence of the repeller requires the sufficiently large m bounded by $m_1^* \geq 1$ for $\frac{c}{r} \leq \frac{N-n}{n(N-1)}$ and $m_2^* \geq 1$ for $\frac{c}{r} \geq \frac{N-n}{n(N-1)}$. The parameter region of m is first magnified and then shortened with increasing $\frac{c}{r}$ (figure 1b), increasing N (figure 1c) or increasing n (figure 1d), which is maximized to all convex production functions of $\phi_m(\theta)$ at $\frac{c}{r} = \frac{N-n}{n(N-1)}$.

- Under the condition of $m \in (0, m_2^*)$ and $\frac{c}{r} \leq \frac{N-n}{n(N-1)}$ or the condition of $m \in (0, m_1^*)$ and $\frac{c}{r} \geq \frac{N-n}{n(N-1)}$, a unique CESS appears in the adaptive dynamics (figure 2d). Irrespective of the initial condition, the initially uniform population converges and resides in a moderate cooperative probability which statistically means the coexistence of cooperators and defectors. In contrast with the requirement of the occurrence of a repeller, the generation of a CESS demands the sufficiently small m, which is bounded by $m_2^* \leq 1$ for $\frac{c}{r} \leq \frac{N-n}{n(N-1)}$ and $m_1^* \leq 1$ for $\frac{c}{r} \geq \frac{N-n}{n(N-1)}$. However, similar to the parameter region yielding a repeller, growing $\frac{c}{r}$ (figure 1b), growing N (figure 1c) or growing n (figure 1d) first magnifies and then shortens the region of m yielding a CESS, which is maximized to all concave production functions of $\phi_m(\theta)$ at $\frac{c}{r} = \frac{N-n}{n(N-1)}$.

- Evolution un-changes the state of the initially uniform population when the linear production function is adopted and $\frac{c}{r}$ arrives at $\frac{N-n}{n(N-1)}$. Such result is less meaning and is no longer considered in our paper.

Note when $\frac{c}{r} > \frac{N-n}{n(N-1)}$ satisfies, a hysteresis effect occurs in which if the population with a repeller initially is in the state of full cooperation and $m_1^* \leq m \leq m_2^*$ is reached, the population will settle in the state of full defection and any cooperative state will not be established even if $m > m_2^*$ again holds (figures 1b–d).

Secondly, another family of production functions characterized by $\psi_w(\theta) = \frac{r}{n}(1 + w + w^2 + \cdots + w^{n\theta-1})$ (figure 3a) are investigated, which have been adopted to consider the evolution of two discrete strategies in infinitely large populations [17]. The shape of such production functions is controlled by an exponent w and clearly they show that the first cooperator contributes $\frac{r}{n}$ to the public goods, the ensuing cooperator adds $\frac{r}{n}w$ to the public goods, and so on, to the last cooperator of $n\theta$ producing $\frac{r}{n}w^{n\theta-1}$ of the public goods. Easily verified, $\psi_{w=1}(\theta) = \phi_{m=1}(\theta) = \frac{r}{n}\theta$, $\psi_{w<1}$ similar to $\phi_{m<1}$ is concave, and $\psi_{w>1}$ similar to $\phi_{m>1}$ is convex. In particular, $\psi_{w\to0}$ and $\psi_{m\to+\infty}$ are threshold production functions at $\theta = 1/n$. The difference between ψ_w and ϕ_m is that $\psi_{w<1}$ and $\phi_{m<1}$ completely locate on either side of the curve $\psi_{w=1}$ which can also occurs for both $\psi_{w>1}$ and $\phi_{m>1}$ when only the values of $\psi_m(\theta)$ at $\theta = 0, 1/n, \cdots, 1$ are considered. Additionally, the value of the public goods provided when all group members cooperate (i.e., $\theta = 1$) is variable for all $\psi_w(\theta)$ but is fixed for all $\phi_m(\theta)$.

Substituting $\psi_w(\theta)$ into Eq.2 determines the adaptive dynamics,

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = G(x) = \frac{[\frac{r}{n}(1+(w-1)x)^{n-1} - c\frac{N-1}{N-n}]}{2Nf(x)}.$$

The above five evolutionary scenarios for ϕ_m also occur for ψ_w. However, the parameter ranges corresponding to each evolutionary scenario exhibit great difference in these two families of production functions. Sufficiently large w ($w \geq (\frac{cn(N-1)}{r(N-n)})^{\frac{1}{n-1}}$) in ψ_w navigates the direction of evolution to full cooperation when $\frac{c}{r} < \frac{N-n}{n(N-1)}$ holds, yet which demands the moderate m in ϕ_m. On contrary, when $\frac{c}{r} > \frac{N-n}{n(N-1)}$ holds, sufficiently small w ($w \leq (\frac{cn(N-1)}{r(N-n)})^{\frac{1}{n-1}}$) in ψ_w induces the direction of evolution to full defection yet which requires the moderate m in ϕ_m. Moreover if ϕ_m is adopted, a unique CESS may occur in the case of $\frac{c}{r} > \frac{N-n}{n(N-1)}$ and a unique repeller in the case of $\frac{c}{r} < \frac{N-n}{n(N-1)}$, both of which are impossible for ψ_w. In deference to the adaptive dynamics for ϕ_m, the parameter range favoring full cooperation in ψ_w shrinks with growing $\frac{c}{r}$ (figure 3b), decreasing N (figure 3c) or growing n (figure 3d) which is the opposite for that favoring full defection. Different from the adaptive dynamics for ϕ_m, the parameter range yielding a unique repeller or a unique CESS in ψ_w only exists until some thresholds are reached and then shrinks or expands with growing $\frac{c}{r}$ (figure 3b), decreasing N (figure 3c) or growing n (figure 3d). A hysteresis effect completely similar to ϕ_m also occurs in ψ_w.

B. Sigmoidal or inverse sigmoidal production functions

We have investigated the adaptive dynamics of the mixed strategy under convex and concave production functions which are segments of the standard S-shaped production functions. However, the adaptive dynamics under the convex and concave production functions can't predict that of sigmoidal production functions since the adaptive dynamics are related with the whole shape of the production function. Following we will consider sigmoidal as well as inverse sigmoid production functions for generality which may lead to more than one singular strategy since $G'(0)$ and $G'(1)$ have different signs. The generalized sigmoidal or inverse sigmoidal production functions are intractable analytically, so we illustrate the adaptive dynamics induced by special types, i.e., symmetrically sigmoidal and symmetrically inverse sigmoidal production functions. Specifically, symmetrically sigmoidal $g(\theta)$ (or symmetrically inverse sigmoidal $g(\theta)$) is convex (or concave) in $[0, 1/2)$ and concave (or convex) in $[1/2, 1]$ satisfying, $\forall \theta_1, \theta_2 \leq 1/2$, $g(\theta_1) - g(\theta_2) = -g(1-\theta_1) + g(1-\theta_2)$. Here, the adaptive dynamics yield at most two singular strategies $0 < x_1 < 1/2 < x_2 < 1$, where x_1 is a repeller (or a CESS) and x_2 is a CESS (or a repeller) (see analysis in Appendix). In principle, sigmoidal or inverse sigmoidal production functions can generate between zero and two singular strategies. Moreover, there are up to six different dynamical scenarios since the stability of adjacent singular strategies must alternate (see analysis in Appendix). Due to the symmetry of such production functions, the unique singular strategy, if it appears, is $x = \frac{1}{2}$ and satisfies $G'(x) = 0$. Here, the adaptive dynamics don't feel the existence of such singular strategy and perform just as the system yields no singular strategies.

Consider a third series of production functions $\varphi_s(\theta)$ whose shape is characterized by the parameter $s > -0.5$ in figure 4a

$$\varphi_s(\theta) = \begin{cases} r\frac{1-(1-2\theta)^{\frac{1}{2s+1}}}{2} \ , \ 0 \leq \theta \leq \frac{1}{2}, \\ r\frac{1+(2\theta-1)^{\frac{1}{2s+1}}}{2} \ , \ \frac{1}{2} \leq \theta \leq 1. \end{cases} \tag{4}$$

Easily verified, $\varphi_{s=0}(\theta) = \phi_{m=1}(\theta) = \psi_{w=1}(\theta) = r\theta$. In the case of $s > 0$, $\varphi_s(\theta)$ is a symmetrically sigmoidal function and is similar to $\zeta_{k>0}(\theta) = \zeta_{k<0}(\theta) = r\frac{e^{\frac{n}{2}k}+1}{e^{\frac{n}{2}k}-1}\frac{1}{e^{-kn(\theta-0.5)}+1} - \frac{1}{e^{\frac{n}{2}k}-1}$ which is also symmetrically sigmoid and is adopted in [18]. In the case of $-\frac{1}{2} < s < 0$, $\varphi_s(\theta)$ is a symmetrically inverse sigmoidal function. Particularly in the case of $s \to -0.5$ or $s \to +\infty$, $\varphi_s(\theta)$ is a threshold function at $\theta = 0, 1$ or at $\theta = 0.5$. Note that when all group members cooperate, φ_s or ϕ_m produces the same public goods with varying s or m and ψ_w produces different public goods with varying w.

Two evolutionary scenarios occurring in ϕ_m, cooperators eventually vanish and cooperators dominate the whole population, also appear in the adaptive dynamics for φ_s. Two reasons underlying such evolutionary results are the absence of the interior singular strategy and the existence of only a unique singular strategy satisfying $G'(x) = 0$. The latter case never happens in the adaptive dynamics for ϕ_m. Similar to the adaptive dynamics for ϕ_m, moderate s ($s_2^* \leq s \leq s_1^*$) along with sufficiently small $\frac{c}{r}$ ($\frac{c}{r} < \frac{N-n}{n(N-1)}$) guarantees unidirectional evolution to full cooperation, and moderate s ($s_1^* \leq s \leq s_2^*$) together with sufficiently large $\frac{c}{r}$ ($\frac{c}{r} > \frac{N-n}{n(N-1)}$) navigates unidirectional evolution to full defection, where s_1^* and s_2^* are the solutions of $G(0) = 0$ and $G(\frac{1}{2}) = 0$ with respect to s (see analysis in Appendix). Moreover, increasing $\frac{c}{r}$, n or decreasing N shrinks the parameter region of s favoring full cooperation and expands the parameter region of s favoring full defection (figure 4b–d).

In contrast with the adaptive dynamics for ϕ_m, two evolutionary situations yielding a unique

CESS and a unique repeller no longer appear in the adaptive dynamics for φ_s and two new evolutionary situations occur for φ_s as follows (see analysis in Appendix).

- If $s > s_1^*$ as well as $\frac{c}{r} \leq \frac{N-n}{n(N-1)}$ holds or $s > s_2^*$ as well as $\frac{c}{r} \geq \frac{N-n}{n(N-1)}$ holds (figure 2e), a smaller interior repeller coexists with a larger interior CESS as compared with $\frac{1}{2}$. Full defection and a larger cooperative probability may be the final state of the initially uniform population up to the start-up condition. Such dynamics require the sufficiently large s bounded by $s_1^* \geq 0$ for $\frac{c}{r} \leq \frac{N-n}{n(N-1)}$ and $s_2^* \geq 0$ for $\frac{c}{r} \geq \frac{N-n}{n(N-1)}$. The parameter region of s is first magnified and then shortened with increasing $\frac{c}{r}$, N or n, and maximized to all sigmoid production functions of $\varphi_s(\theta)$ at $\frac{c}{r} = \frac{N-n}{n(N-1)}$ (figure 4b–d).
- If $s < s_2^*$ as well as $\frac{c}{r} \leq \frac{N-n}{n(N-1)}$ holds or $s < s_1^*$ as well as $\frac{c}{r} \geq \frac{N-n}{n(N-1)}$ holds (figure 2f), a smaller interior CESS coexists with a larger interior repeller as compared with $\frac{1}{2}$. Full cooperation and a smaller cooperative probability may be the final state of the initially uniform population depending on the initial condition. Such dynamics require the sufficiently small s bounded by $s_2^* \leq 0$ for $\frac{c}{r} \leq \frac{N-n}{n(N-1)}$ and $s_1^* \leq 0$ for $\frac{c}{r} \geq \frac{N-n}{n(N-1)}$. The parameter region of s is first magnified and then shortened with increasing $\frac{c}{r}$, N or n, and maximized to all inverse sigmoid production functions of $\varphi_s(\theta)$ at $\frac{c}{r} = \frac{N-n}{n(N-1)}$ (Fig.4b–d).

Unlike the adaptive dynamics for ϕ_m, we have two saddle-node bifurcations at s_1^* and s_2^* (figure 4b–d). Here, as s initially below s_1^* (or s_2^*) increases when $\frac{c}{r} > \frac{N-n}{n(N-1)}$ (or $\frac{c}{r} < \frac{N-n}{n(N-1)}$) holds, a smaller stable singular strategy and a larger unstable one first disappear and no singular strategies exist and then a larger stable singular strategy and a smaller unstable one are created. Moreover, the hysteresis effect can also be found in φ_s for $\frac{c}{r} > \frac{N-n}{n(N-1)}$ (figure 4b–d). When $s > s_2^*$, the population initially residing in the CESS will settle in full defection if s decreases below s_2^* and above s_1^*. However even if s rises above s_2^* again, any cooperative state will not be established again. Similarly, if the population with $s < s_1^*$ initially locates in full cooperation and s increases between s_1^* and s_2^*, the population will reside in full defection and only the CESS less cooperative than full cooperation is reached even if s returns below s_1^*.

IV. DISCUSSION

Besides the interest in relaxing the assumption of infinitely large populations and two alternative pure strategies in the theory of public goods games, clearly it is natural to go beyond such premises and investigate the evolution of mixed strategy in finite populations. Our extensions to public goods games with arbitrary production functions allow a natural relation between the benefit and the number of cooperators. A linear public goods game in a group of size n is equivalent to $n - 1$ pairwise prisoner's dilemma under some transformation [40]. However when nonlinear production functions are adopted, this equivalence fails and thus the public goods games are no longer the simple addition of several two-person games. Although the adaptive dynamics of the mixed strategy in two-player games have been investigated [35], the evolution in multiple players games still deserves deep focus and turns out to exhibit complex adaptive dynamics. In principle, the evolution of the mixed strategy just yields two types of singular strategies, the CESS and the repeller, which contrasts with the evolution of the continuously varying investment leading to the evolutionary branching as well as the CESS and the repeller in [19, 20]. The adaptive dynamics produce a repeller if a singular strategy exists when a convex production function is adopted, yet a CESS occurs if a singular strategy exists when a concave production is taken. Such complex dynamics are not only possible for multi-player public goods games but also possible for even two player games, unlike the results exhibited in the evolution of two alternative discrete strategies [14]. The symmetrically sigmoidal production function brings about a smaller repeller and a larger CESS as compared with $\frac{1}{2}$, yet the symmetrically inverse sigmoidal production function results into a smaller CESS and a larger repeller as compared with $\frac{1}{2}$. Such complex dynamics only appear in multi-player games not possible for two-player games similar to [14]. In particular, cooperation is doomed when the group size is equal to the population size irrespective of production functions, similar to the evolution of two pure strategies in finite populations [15, 16].

The adaptive dynamics of the mixed strategy are related with the whole shape of the production functions. We quantitatively analyze how nonlinear production functions exert the effect on the evolutionary dynamics on the basis that the shape of a production function can be controlled by a parameter. Two classes of production functions considered ϕ_m and ψ_w are separately characterized by the parameters m and w, encompassing convex ($m, w > 1$), linear ($m, w = 1$), and concave curves ($0 < m, w < 1$). Then we take into account another family of production functions φ_s represented by the parameter s, including sigmoidal ($s > 0$), linear ($s = 0$), and inverse sigmoidal curves ($0.5 < s < 0$). When all group members cooperate, ϕ_m or ψ_w produces the same public good for varying m or w, yet φ_s produces different public goods for varying s. We find the various evolutionary scenarios are interconnected through variations of the continuous parameters the shape of production functions (m, s, w), the ratio of the remaining initial endowment induced by defection and the enhancement factor (c/r), the group size (n) and the population size (N), which seamlessly relates seemingly disparate biological situations.

Full cooperation evolves for moderate m, moderate s, or sufficiently large w when $\frac{c}{r} < \frac{N-n}{n(N-1)}$ holds, yet for moderate m, moderate s, or sufficiently small w when $\frac{c}{r} > \frac{N-n}{n(N-1)}$ holds, full defection evolves. The unique CESS or unique repeller only occurs in ϕ_m and ψ_w. The occurrence of a unique CESS requires sufficiently small m for all $\frac{c}{r}$ but sufficiently small w only for $\frac{c}{r} < \frac{N-n}{n(N-1)}$, whereas a unique repeller occurs when m is sufficiently large for all $\frac{c}{r}$ but when w is sufficiently large only for $\frac{c}{r} < \frac{N-n}{n(N-1)}$. In contrast with the adaptive dynamics exhibiting in ϕ_m and ψ_w, the adaptive dynamics emerging in φ_s reduce two evolutionary situations leading to a unique CESS or a unique repeller and increase two new evolutionary dynamics. A smaller repeller coexists with a larger CESS as compared with $\frac{1}{2}$ with m above some positive value, yet with m below some negative value, a smaller CESS co-occurs with a larger repeller as compared with $\frac{1}{2}$.

We found the hysteresis effect in all these three classes of production functions, where the population in a cooperative state will experience a sudden transition and settle in defection if some parameters are changed. However, a subsequent complete recovery of parameter values will not recover the initial cooperative state. Unlike the adaptive dynamics in ϕ_m and ψ_w, we have two saddle-node bifurcations at s_1^* and s_2^* in φ_s where as s initially below s_1^* (or s_2^*) increases and $\frac{c}{r} > \frac{N-n}{n(N-1)}$ (or $\frac{c}{r} < \frac{N-n}{n(N-1)}$) holds, two singular strategies, with the smaller stable and the larger unstable, first disappear and no singular strategies exist and then two new singular strategies are created, with the larger stable and the smaller unstable. In all three classes of production functions, the parameter region favoring full cooperation shrinks and that favoring full defection expands with growing $\frac{c}{r}$, decreasing N or growing n. The parameter range leading to a unique CESS or a unique repeller in ϕ_m first expands and then shrinks with growing $\frac{c}{r}$, N or n, similar to the parameter range bringing about the coexistence of a CESS and a repeller in φ_s. However in ψ_w, the parameter range yielding a unique CESS expands with growing $\frac{c}{r}$, decreasing N or growing n, and suddenly diminish when $\frac{c}{r} = \frac{N-n}{n(N-1)}$ is reached, which is the opposite for the parameter range producing a unique repeller.

Acknowledgements

We are grateful for support from the National Natural Science Foundation of China (NSFC, No. 60774089, 10972003, 10926195).

---

[1] Fortunato, A., Strassmann, J. E., Santorelli, L. & Queller, D. C. 2003 Co-occurrence in nature of different clones of the social amoeba, *Dictyostelium discoideum*. *Mol. Ecol.* **12,** 1031–1038.

[2] Stander, P. E. 1992 Cooperative hunting in lions: the role of the individual. *Behav. Ecol. Sociobiol.* **29,** 445–454.

[3] Gazda, S. K., Connor, R. C., Edgar, R. K. & Cox, F. 2005 A division of labour with role specialization in group-hunting bottlenose dolphins (*Tursiops truncatus*) off Cedar Key, Florida. *Proc. R. Soc. B* **272,** 135–140.

[4] Heinsohn, R. & Packer, C. 1995 Complex cooperative strategies in group-territorial African lions. *Science* **269,** 1260–1262.

[5] Milinski, M., Sommerfeld, R. D., Krambeck, H. J., Reed, F. A. & Marotzke, J. 2008 The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proc. Natl Acad. Sci. USA* **105,** 2291–2294.

[6] Semmann, D., Krambeck, H. J. & Milinski, M. 2003 Volunteering leads to rock-paper-scissors dynamics in a public goods game. *Nature* **425,** 390–393.

[7] Chen, X. J., Liu, Y. K., Zhou, Y. H., Wang, L. & Perc, M. 2012 Adaptive and bounded investment returns promote cooperation in spatial public goods games. *PLoS ONE* **7,** e36895.

[8] 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.

[9] Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A. & Sigmund, K. 2007 Via freedom to coercion: the emergence of costly punishment. *Science* **316,** 1905–1907.

[10] Rand, D. G. & Nowak, M. A. 2011 The evolution of antisocial punishment in optional public goods games. *Nat. Commun.* **2,** 434.

[11] Perc, M. 2012 Sustainable institutionalized punishment requires elimination of second-order free-riders. *Sci. Rep.* **2,** 344.

[12] Eldakar, O. T. & Wilson, D. S. 2008 Selfishness as second-order altruism. *Proc. Natl Acad. Sci. USA*

105, 6982–6986.

[13] Hauert, C., Holmes, M. & Doebeli, M. 2006 Evolutionary games and population dynamics: maintenance of cooperation in public goods games. *Proc. R. Soc. B* **273,** 2565–2571.

[14] Bach, L. A., Helvik, T. & Christiansen, F. B. 2006 The evolution of n-player cooperation-threshold games and ESS bifurcations. *J. Theor. Biol.* **238,** 426–434.

[15] Pacheco, J. M., Santos, F. C., Souza, M. O. & Skyrms, B. 2009 Evolutionary dynamics of collective action in N- person stage hunt dilemmas. *Proc. R. Soc. B* **276,** 315–321.

[16] Souza, M. O., Pacheco, J. A. & Santos, F. C. 2009 Evolution of cooperation under N-person snowdrift games. *J. Theor. Biol.* **260,** 581–588.

[17] Hauert, C., Michor, F., Nowak, M. A. & Doebeli, M. 2006 Synergy and discounting of cooperation in social dilemmas. *J. Theor. Biol.* **239,** 195–202.

[18] Archetti, M. & Scheuring, I. 2011 Coexistence of cooperation and defection in public goods games. *Evolution* **65,** 1140–1148.

[19] Doebeli, M., Hauert, C. & Killingback, T. 2004 The evolutionary origin of cooperators and defectors. *Science* **306,** 859–862.

[20] Cornforth, D. M., Sumpter, J. T., Brown, S. P. & Brännström Å. 2012 Synergy and group size in microbial cooperation. *The American Naturalist* **180,** 296–305.

[21] Marwel, G. & Oliver, P. 1993 *The critical mass in collective action: a micro-social theory*. Cambridge University Press.

[22] Heckathor, D. D. 1996 The dynamics and dilemmas of collective action. *Am. Sociol. Rev.* **61,** 250–277.

[23] Clutton-Brock, T. 2002 Breedint together: kin selection and mutualism in cooperative vertebrates. *Science* **296,** 69–72.

[24] Sumpter, D. J. T. 2010 *Collective animal behavior*. Princeron University Press.

[25] Chuang, J. S., Rivoire, O. & Leibler, S. 2010 Cooperation and Hamilton's rule in a simple synthetic microbial system. *Molecular Systems Biology* **6,** 398.

[26] Smith, J., Dyken, J. D. V. & Zee, P. C. 2010 A generalization of Hamilton's rule for the evolution of microbial cooperation. *Science* **328,** 1700–1703.

[27] Packer, C. & Ruttan, L. 1988 The evolution of cooperative hunting. *Am. Nat.* **132,** 159–198.

[28] Kim, K. W., Krafft, B. & Choe, J. C. 2005 Cooperative prey capture by young subsocial spider I. Functional value. *Behav. Ecol. Sociobiol.* **59,** 92–100.

[29] Boesch, C. 1994 Cooperative hunting in wild chimpanzees. *Anim. Behav.* **48,** 653–667.

[30] MacNulty, D. R., Smith, D. W., Mech, L. D., Vucetich, J. A. & Packer, C. 2011 Nonlinear effects of group size on the success of wolves hunting elk. *Behav. Ecol.* **23,** 75-82.

[31] Heckathorn, D. D. 1996 The dynamics and dilemmas of collective action. *American Sociological Review* **61,** 250–277.

[32] Nowak, M. & Sigmund, K. 1990 The evolution of stochastic strategies in the prisoner's dilemma. *Acta Appl. Math.* **20,** 247–265.

[33] Hofbauer, J. & Sigmund, K. 1990 Adaptive dynamics and evolutionary stability. *Appl. Math. Lett.* **3,** 75–79.

[34] Dieckmann, U. & Law, R. 1996 The dynamical theory of coevolution: a derivation from stochastic ecological processes. *J. Math. Biol.* **34,** 579–612.

[35] Imhof, L. A. & Nowak, M. A. 2010 Stochastic evolutionary dynamics of direct reciprocity. *Proc. R. Soc. B* **277,** 463–468.

[36] Proulx, S. R. & Day, T. 2001 What can invasion analyses tell us about evolution under stochasticity in finite populations? *Selection* **2,** 1–15.

[37] Wild, G. & Taylor, P. D. 2004 Fitness and evolutionary stability in game theoretic models of finite populations. *Proc. R. Soc. B* **271,** 2345–2349.

[38] Nowak, M. A., Sasaki, A., Taylor, C. & Fudenberg, D. 2004 Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428,** 646–650.

[39] Geritz, S. A. H., Metz, J. A. J., Kisdi, É. & Meszéna, G. 1997 Dynamics of adaptation and evolutionary branching. *Phys. Rev. L.* **78,** 2024–2027.

[40] Hauert, C. & Szabó , G. 2003 Prisoner's dilemma and public goods games in different geometries: compulsory versus voluntary interactions. *Complexity* **8**, 31–38.

[41] Champagnat, N., Ferrière, R. & Arous, G. B. 2001 The canonical equation of adaptive dynamics: a mathematical view. *Selection* **2,** 73–83.

Figure legends:

Figure 1 Normalized production functions $\phi_m(\theta)/r$ with $m > 0$ and phase diagrams illustrating different dynamical regimes induced by them. (a) The shape of $\phi_m(\theta)$ is controlled by m. Convex $\phi_m(\theta)$ corresponds to $m > 1$ (dashed), linear $\phi_m(\theta)$ corresponds to $m = 1$ (solid), and concave $\phi_m(\theta)$ corresponds to $0 < m < 1$ (dotted). (b) For $N = 100$ and $n = 50$, the dynamics are determined by c/r and m. (c) Dynamics as determined by N and m for $n = 50$ and $c/r = 0.01$. (d) For $N = 100$ and $c/r = 0.02$, the dynamics are determined by n and m. The parameter regions in $(b) - -(d)$ are separated into four subregions for full cooperation, full defection, a unique repeller and a unique CESS by $m = m_1^*$ (Dashed) and $m = m_2^*$ (solid).

Figure 2 Samples of evolutionary dynamics where color shades indicate higher frequencies of a strategy. (a and b) Unidirectional evolutionary dynamics in the absence of singular strategies in which full defection evolves (a) and full cooperation evolves (b). (c) A unique repeller leads the population to evolve to full defection or full cooperation up to the initial condition. (d) A unique CESS in which the population finally converges and resides. (e) The co-existence of a smaller CESS and a large repeller. (f) The co-occurrence of a smaller repeller and a larger CESS. Results were obtained from numerical simulations in Appendix. Parameters: population size $N = 100$, group size $n = 4$, mutation rate $u = 0.001$, mutation variance $\sigma = 0.0001$, and the following production functions: (a) $g(\theta) = 3.5$ for $0 \leq \theta \leq 1$, (b) $g(\theta) = 5$ for $0 \leq \theta \leq 1$, (c) $g(\theta) = 0$ for $0 \leq \theta < 1$, $g(\theta) = 3.5$ for $\theta = 1$, (d) $g(\theta) = 0$ for $\theta = 0$, $g(\theta) = 3.5$ for $0 < \theta \leq 1$, (e) $g(\theta) = 0$ for $\theta = 0$, $g(\theta) = 1.75$ for $0 < \theta < 0.5$ and $g(\theta) = 3.5$ for $\theta = 1$, (f) $g(\theta) = 0$ for $0 \leq \theta < 0.5$, $g(\theta) = 1.75$ for $\theta = 0.5$ and $g(\theta) = 3.5$ for $0.5 < \theta \leq 1$.

Figure 3 Normalized production functions $\psi_w(\theta)/r$ with $w > 0$ and phase diagrams illustrating different dynamical regimes induced by them. (a) The shape of $\psi_w(\theta)$ is controlled by m. Convex $\psi_w(\theta)$ corresponds to $w > 1$ (dashed), linear $\psi_w(\theta)$ corresponds to $w = 1$ (solid) and concave $\psi_w(\theta)$ corresponds to $0 < w < 1$ (dotted). (b) For $N = 100$ and $n = 50$, the dynamics are determined by c/r and w. (c) Dynamics as determined by N and w for $n = 50$ and $c/r = 0.01$. (d) For $N = 100$ and $c/r = 0.02$, the dynamics are determined by n and w. The parameter regions in $(b) - -(d)$ are separated into four subregions for full cooperation, full defection, a unique repeller and a unique CESS by $w = (\frac{cn(N-1)}{r(N-n)})^{\frac{1}{n-1}}$ (solid) and $\frac{c}{r} = \frac{N-n}{n(N-1)}$ (dashed).

Figure 4 Normalized production functions $\varphi_s(\theta)/r$ with $s > -0.5$ and phase diagrams illustrating different dynamical regimes induced by them. (a) The shape of $\varphi_s(\theta)$ is controlled by s. Symmetrically inverse sigmoidal $\varphi_s(\theta)$ corresponds to $-0.5 < s < 0$ (dashed), linear $\varphi_s(\theta)$ corresponds to $s = 0$ (solid) and symmetrically sigmoidal $\varphi_s(\theta)$ corresponds to $s > 0$ (dotted). (b) For $N = 100$ and $n = 50$, the dynamics are determined by c/r and s. (c) Dynamics as determined by N and s for $n = 50$ and $c/r = 0.01$. (d) For $N = 100$ and $c/r = 0.02$, the dynamics are determined by n and s. The parameter regions in (b)–(d) are separated into four subregions for full cooperation, full defection, the coexistence of a smaller repeller and a large CESS and the co-occurrence of a smaller CESS and a larger repeller by $s = s_1^*$ (solid) and $s = s_2^*$ (dashed).

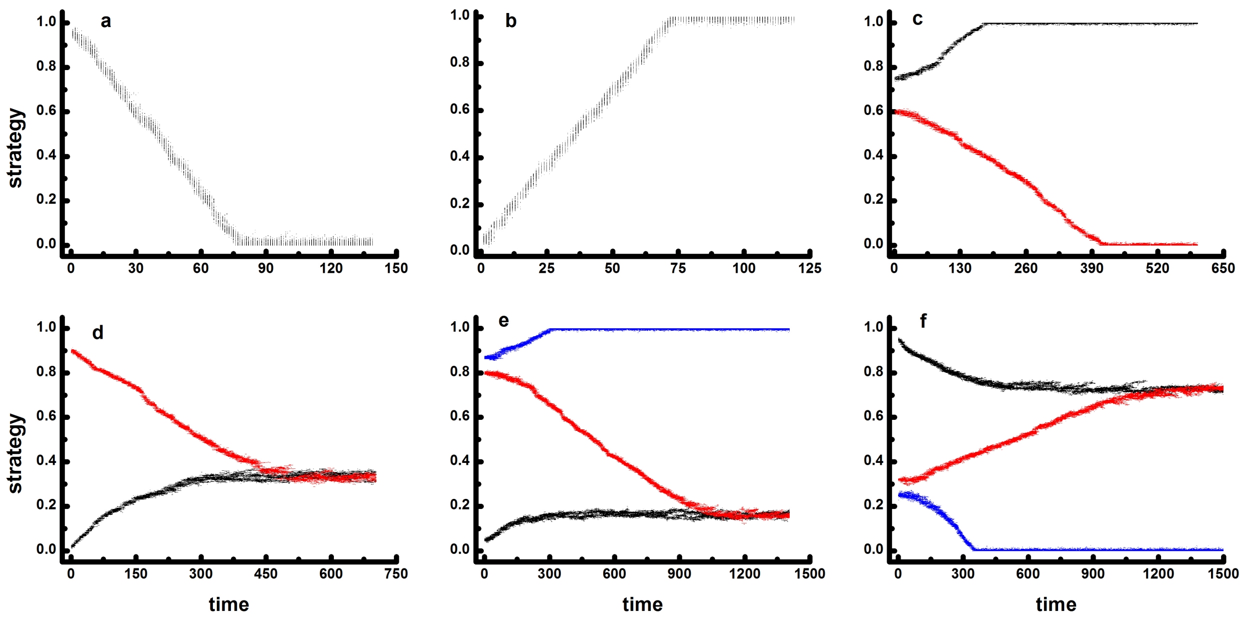

FIG. 1: Inserted here

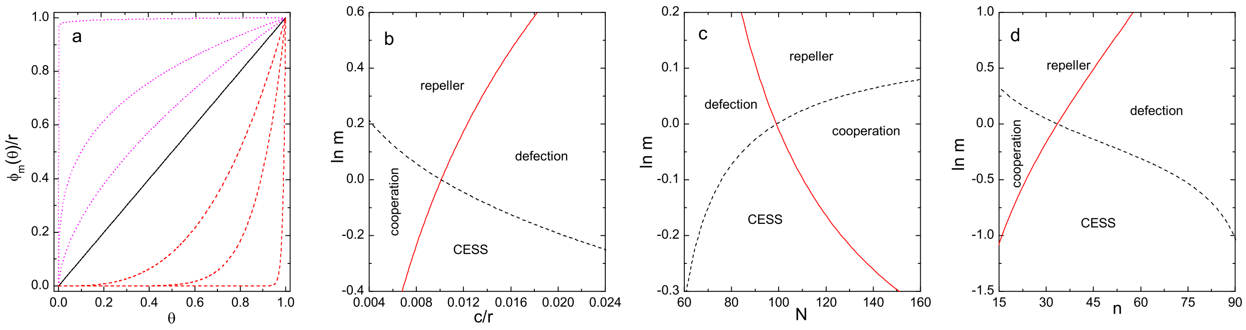

FIG. 2: Inserted here

APPENDIX A: THE PAYOFF

If the population consists of X individuals of strategy w and $N - X$ individuals of strategy z, a z strategist engages a group consisting of i individuals of strategy z and $n-1-i$ individuals of strategy w in its $n-1$ interactive partners with the probability $H(i, n-1, N-X-1, N-1) \overset{\text{def}}{=} \frac{\binom{N-X-1}{i}\binom{X}{n-1-i}}{\binom{N-1}{n-1}}$. Here a z strategist encounters k cooperators and $i-k$ defectors from strategy z, and l cooperators and $n-1-i-l$ defectors from strategy w with the probability $\binom{i}{k}\binom{n-1-i}{l} z^k (1-z)^{i-k} w^l (1-w)^{n-1-i-l}$. It follows that the average payoff to strategy z, $P(z, w, i+1)$, in a group having $i+1$ individuals of strategy z and $n-1-i$ individuals of strategy w is $P(z, w, i+1) = \sum_{k=0}^{i} \sum_{l=0}^{n-1-i} \binom{i}{k}\binom{n-1-i}{l} z^k (1-z)^{i-k} w^l (1-w)^{n-1-i-l} \Big[z g(\frac{k+l+1}{n}) + (1-z)[g(\frac{k+l}{n}) + c] \Big]$. Here $g(x)$ is the production function, describing the relationship between the proportion of cooperators in a group

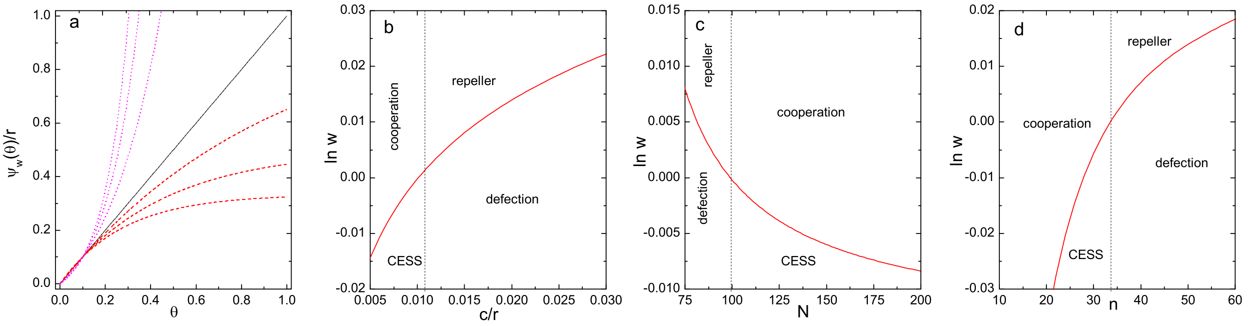

FIG. 3: Inserted here

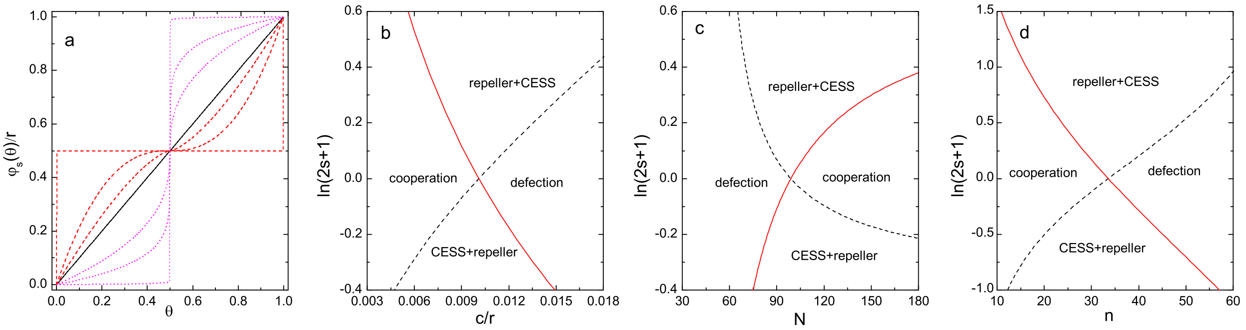

FIG. 4: Inserted here

and the public good produced, which is an increasing function and satisfies $g(0) = 0$. Taking the weighted average over all possible numbers of strategy z's individuals in the interactive partners with the weights given by the probability $H(i, n-1, N-X-1, N-1)$, yields the average payoff of individuals of strategy z in a population consisting of X individuals of strategy w and $N - X$ individuals of strategy z, $f(z, w, N-X)$, $f(z, w, N-X) = \sum_{i=0}^{n-1} H(i, n-1, N-X-1, N-1) P(z, w, i+1)$.

APPENDIX B: THE DERIVATION OF THE ADAPTIVE DYNAMICS IN FINITE POPULATIONS

We adopt the frequency dependent Moran process to update the mixed strategy. In each time step, all individuals initially compete to reproduce an offspring. The probability that each individual reproduces is proportional to his payoff. With probability u, the offspring inherits the strategy of his parent. With complementary probability $1 - u$, a mutant y emerges whose strategy obeys a symmetric probability distribution with the parent strategy as mean and with the variance denoted by σ. Subsequently, a randomly chosen individual is replaced by the offspring. Therefore, the population size N remains constant during the evolutionary process.

The principle of mutual exclusion says that two adaptive strategies cannot coexist indefinitely in the population when not renewed by mutations. We assume that the mutation rate u is sufficiently small so that the mutant strategy or the resident strategy reaches fixation before a new mutant occurs. In the long term, there is only a single strategy prevailing in the population at almost any point in time and the evolutionary process can be envisaged a sequence of strategy substitutions. Let $p(x,t)$ be the probability that the strategy in the population is x at the time t. By virtue of the Markov property,

$$\frac{\mathrm{d}}{\mathrm{d}t}p(x,t) = \int [w(x,y)p(y,t) - w(y,x)p(x,t)]dy, \tag{B1}$$

where $w(y,x)$ is the transition probability per unit time for the strategy substitution $x \to y$. Since mutation and selection are uncorrelated, $w(y,x) = \mu N M(y)\rho(x;y)$ where $\rho(x;y)$ is the fixation probability of a mutant y in a resident population x.

The mean path of the strategy substitutions is denoted by $\langle x\rangle(t)$ and defined as

$$\langle x\rangle(t) = \int xp(x,t)\mathrm{d}x.$$

Neglecting the order of integration and differential, we can obtain the dynamics of the mean path from Eq.(B1)

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle x\rangle(t) = \int x\frac{\mathrm{d}}{\mathrm{d}t}p(x,t)\mathrm{d}x = \langle a_1(x)\rangle(t),$$

where $a_1(x) = \int (y-x)w(y,x)\mathrm{d}y$. On the condition that the derivations of the stochastic realizations from the mean path are relatively small which means that the variance of the mutation process is sufficient small, the above equation can be approximated as

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle x\rangle(t) = a_1(\langle x\rangle)(t) = \int (y-\langle x\rangle)w(y,\langle x\rangle)\mathrm{d}y = \mu N\int (y-\langle x\rangle)M(y)\rho_N(\langle x\rangle;y)\mathrm{d}y.$$

Using the first-order approximation of the fixation probability and the symmetry of the mutation, we obtain the dynamics of the mean path

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle x\rangle(t) = C_0\frac{\partial}{\partial y}\Big|_{y=\langle x\rangle}\rho_N(\langle x\rangle;y), \tag{B2}$$

where C_0 is a constant which reflects the populations size and the mutation process. Using the third-order approximation of the fixation probability, we obtain the dynamics of the mean path

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle x\rangle(t) = C_1\frac{\partial}{\partial y}\Big|_{y=\langle x\rangle}\rho(\langle x\rangle;y) + C_2\frac{\partial^3}{\partial y^3}\Big|_{y=\langle x\rangle}\rho_N(\langle x\rangle;y), \tag{B3}$$

where C_1 and C_2 are constants which reflect the populations size and the mutation process. Note the bracket denoting the mean will ceased using for simplicity in main texts.

APPENDIX C: THE CALCULATION OF THE ADAPTIVE DYNAMICS IN OUR MODEL

The payoff of y $f(y,x,a)$ and the payoff of x $f(x,y,N-a)$ in a population with a individuals of y and $N-a$ individuals of x are

$$\begin{aligned} f(y,x,a) \;&=\; \textstyle\sum_{i=0}^{n-1} H(i,n-1,a-1,N-1)\sum_{k=0}^{i}\sum_{l=0}^{n-1-i}\binom{i}{k}\binom{n-1-i}{l} \\ &\qquad y^k(1-y)^{i-k}x^l(1-x)^{n-1-i-l}[yg(\tfrac{k+l+1}{n})+(1-y)(g(\tfrac{k+l}{n})+c)], \\ f(x,y,N-a) \;&=\; \textstyle\sum_{i=0}^{n-1} H(i,n-1,N-a-1,N-1)\sum_{k=0}^{i}\sum_{l=0}^{n-1-i}\binom{i}{k}\binom{n-1-i}{l} \\ &\qquad x^k(1-x)^{i-k}y^l(1-y)^{n-1-i-l}[xg(\tfrac{k+l+1}{n})+(1-x)(g(\tfrac{k+l}{n})+c)]. \end{aligned}$$

The calculations of the first, second or third order partial derivatives of $\rho(x;y)$ can be transformed into those of $f(y,x,a)$ and $f(x,y,N-a)$,

$$\begin{aligned} \frac{\partial}{\partial y}\Big|_{y=x} f(y,x,a) &= \frac{(a-1)(n-1)+N-1}{N-1}C(x)-c, \\ \frac{\partial^2}{\partial y^2}\Big|_{y=x} f(y,x,a) &= \frac{(a-1)(n-1)[(a-2)(n-2)+2(N-2)]}{(N-1)(N-2)}D(x), \\ \frac{\partial^3}{\partial y^3}\Big|_{y=x} f(y,x,a) &= \frac{(a-1)(a-2)(n-1)(n-2)[(a-3)(n-3)+3(N-3)]}{(N-1)(N-2)(N-3)}A(x), \\ \frac{\partial}{\partial y}\Big|_{y=x} f(x,y,N-a) &= \frac{a(n-1)}{N-1}C(x), \\ \frac{\partial^2}{\partial y^2}\Big|_{y=x} f(x,y,N-a) &= \frac{a(a-1)(n-1)(n-2)}{(N-1)(N-2)}D(x), \\ \frac{\partial^3}{\partial y^3}\Big|_{y=x} f(x,y,N-a) &= \frac{a(a-1)(a-2)(n-1)(n-2)(n-3)}{(N-1)(N-2)(N-3)}A(x). \end{aligned}$$

where

$$\begin{aligned} C(x) &= \sum_{j=0}^{n-1}\binom{n-1}{j}x^j(1-x)^{n-1-j}\left[g(\frac{j+1}{n})-g(\frac{j}{n})\right], \\ D(x) &= \sum_{j=0}^{n-2}\binom{n-2}{j}x^j(1-x)^{n-2-j}\left[g(\frac{j+2}{n})-2g(\frac{j+1}{n})+g(\frac{j}{n})\right], \\ A(x) &= \sum_{j=0}^{n-3}\binom{n-3}{j}x^j(1-x)^{n-3-j}\left[g(\frac{j+3}{n})-3g(\frac{j+2}{n})+3g(\frac{j+1}{n})-g(\frac{j}{n})\right]. \end{aligned}$$

Adopting the first-order deterministic approximation for $\rho(x;y)$, the adaptive dynamics are,

$$\dot{x} = G(x) = \frac{\partial}{\partial y}\Big|_{y=x}\rho(x;y) = \frac{(N-n)C(x)-c(N-1)}{2Nf(x)}.$$

At singular points, we have

$$\begin{aligned} G^{'}(x) &= \frac{(n-1)(N-n)D(x)}{2Nf(x)}, \\ \frac{\partial^2}{\partial y^2}\Big|_{y=x}\rho(x;y) &= \frac{(n-1)(N-n)D(x)}{3Nf(x)}. \end{aligned}$$

For singular points satisfying $D(x)$ whose evolutionary direction the first-order deterministic approximation fails to tell, employing the third-order deterministic approximation for $\rho(x;y)$, we have

$$x'(t) = \frac{\partial^3}{\partial y^3}\bigg|_{y=x} \rho(x;y) = \frac{(n-1)(n-2)(N-n)A(x)}{4Nf(x)}$$

Here, the adaptive dynamics at x^* is obviously directional and it's just like the adaptive dynamics yield no singular points since $A(x)$ is the derivative of $D(x)$.

APPENDIX D: THE ANALYSIS OF $\phi_m(\theta)$

Consider a family of production functions $\phi_m(\theta) = r\theta^m$ where $m > 0$ is an exponent controlling the shape of production functions. When $m = 1$, $\phi_m(\theta)$ is a linear production function where for all $x \in [0, 1]$, $G(x) > 0$ if $\frac{c}{r} < \frac{N-n}{n(N-1)}$, $G(x) = 0$ if $\frac{c}{r} = \frac{N-n}{n(N-1)}$ and $G(x) < 0$ if $\frac{c}{r} > \frac{N-n}{n(N-1)}$. $\phi_{m>1}(\theta)$ leads to $D(x) > 0$, thus $G(x)|_{m>1}$ has a minimum at $x = 0$ and a maximum at $x = 1$, conversely $\phi_{0<m<1}(\theta)$ leads to $D(x) < 0$, thus $G(x)|_{0<m<1}$ has a minimum at $x = 1$ and a maximum at $x = 0$. Let m_1^* and m_2^* be the solutions of $G(0) = 0$ and $G(1) = 0$ with respect to m, which are given by $m_1^* = \frac{\ln \frac{c(N-1)}{r(N-n)}}{\ln \frac{1}{n}}$ and $m_2^* = \frac{\ln(1-\frac{c(N-1)}{r(N-n)})}{\ln \frac{n-1}{n}}$,

- For $\frac{c}{r} > \frac{N-1}{n(N-n)}$, we have $G(x)\big|_{m=1} < 0$ for $x \in [0, 1]$ and $m_1^* < 1 < m_2^*$:
  Since $G(0)\big|_{m>1} < G(0)\big|_{m=1} < 0$ and $G(1)\big|_{0<m<1} < G(1)\big|_{m=1} < 0$, both $G(x)\big|_{m>1}$ and $G(x)\big|_{0<m<1}$ have negative minima. According to the sign of $G(1)\big|_{m>1}$ or the sign of $G(0)\big|_{0<m<1}$, $G(x)\big|_{m>1} = 0$ or $G(x)\big|_{0<m<1} = 0$ has one or no root.

- For $\frac{c}{r} < \frac{N-1}{n(N-n)}$, we have $G(x)\big|_{m=1} > 0$ for $x \in [0, 1]$ and $m_2^* < 1 < m_1^*$.:
  Since $G(1)\big|_{m>1} > G(1)\big|_{m=1} > 0$ and $G(0)\big|_{0<m<1} > G(0)\big|_{m=1} > 0$, both $G(x)\big|_{m>1}$ and $G(x)\big|_{0<m<1}$ have positive maxima. According to the sign of $G(0)\big|_{m>1}$ or $G(1)\big|_{0<m<1}$, $G(x)\big|_{m>1} = 0$ or $G(x)\big|_{0<m<1} = 0$ has one or no root.

- For $\frac{c}{r} = \frac{N-1}{n(N-n)}$, we have $G(x)\big|_{m=1} = 0$ for $x \in [0, 1]$ and $m_1^* = 1 = m_2^*$:
  $G(x)\big|_{m>1} = 0$ has one interior root since $G(0)\big|_{m>1} < G(0)\big|_{m=1} = 0$ and $G(1)\big|_{m>1} > G(1)\big|_{m=1} = 0$. And $G(x)\big|_{0<m<1} = 0$ has one interior root since $G(0)\big|_{0<m<1} > G(0)\big|_{m=1} = 0$ and $G(1)\big|_{0<m<1} < G(1)\big|_{m=1} = 0$.

APPENDIX E: THE STABILITY OF TWO ADJACENT SINGULAR POINTS ALTERNATES

We assume two adjacent singular points denoted by x_1^* and x_2^* hold for $G'(x) > 0$.

- If $G'(x) \geq 0$ holds for all $x \in (x_1^*, x_2^*)$, $G(x_2^*) > G(x_1^*) = 0$ contradicts with the singularity of x_2^*.
- If there exists $x \in (x_1^*, x_2^*)$ satisfying $G'(x) < 0$, then one of the following three possible cases occurs, $G(x_2^*) > G(x_1^*) = 0$, $G(x_2^*) < G(x_1^*) = 0$ or $G(x_2^*) = G(x_1^*) = 0$. In the first two cases, it is obvious that contradiction results, the third case leads to the emergence of another singular point and contradiction results again.

Similarly, contradiction will result if $G^{'}(x_1^*) < 0$ and $G^{'}(x_2^*) < 0$. Therefore, we can clam that $G^{'}(x_1^*)$ and $G^{'}(x_2^*)$ have different signs and the stability of two adjacent singular points alternates.

APPENDIX F: THE ANALYSIS OF SYMMETRICALLY SIGMOIDAL OR INVERSE SIGMOIDAL $g(s)$

The expression of $D(x)$ can be transformed into $D(x) = \sum_{j=0}^{k-1} \binom{n-2}{j} x^j (1-x)^j [(1-x)^{n-2-2j} - x^{n-2-2j}][g(\frac{j+2}{n}) - 2g(\frac{j+1}{n}) + g(\frac{j}{n})]$ where $k = \frac{n-2}{2}$ holds for even n or $k = \frac{n-1}{2}$ holds for odd n.

- Symmetrically sigmoidal $g(\theta)$ is convex in $[0, 1/2)$ and concave in $[1/2, 1]$ satisfying, $\forall \theta_1, \theta_2 \leq 1/2$, $g(\theta_1) - g(\theta_2) = -g(1-\theta_1) + g(1-\theta_2)$. It is easy to see that $D(1/2) = 0$, $D(x) > 0$ for $x \in [0, 1/2)$ and $D(x) < 0$ for $x \in (1/2, 1]$. There exist at almost two singular strategies $0 \leq x_1 < 1/2 < x_2 \leq 1$, where x_1 is a repeller and x_2 is a CESS.
- Symmetrically inverse sigmoidal $g(\theta)$ is concave in $[0, 1/2)$ and convex in $[1/2, 1]$ satisfying, $\forall \theta_1, \theta_2 \leq 1/2$, $g(\theta_1) - g(\theta_2) = -g(1-\theta_1) + g(1-\theta_2)$. It is easy to see that $D(1/2) = 0$, $D(x) < 0$ for $x \in [0, 1/2)$ and $D(x) > 0$ for $x \in (1/2, 1]$. There exist almost two singular strategies $0 \leq x_1 < 1/2 < x_2 \leq 1$, where x_1 is a CESS and x_2 is a repeller.

APPENDIX G: THE ANALYSIS OF $\varphi_s(\theta)$

Consider a series of production functions $\varphi_s(\theta)$ whose shape is characterized by the parameter $s > -0.5$

$$\varphi_s(\theta) = \left(\begin{array}{ll} r\frac{1-(1-2\theta)^{\frac{1}{2s+1}}}{2} & , \ 0 \leq \theta \leq \frac{1}{2}, \\ r\frac{1+(2\theta-1)^{\frac{1}{2s+1}}}{2} & , \ \frac{1}{2} \leq \theta \leq 1. \end{array} \right. \tag{G1}$$

In the case of $s = 0$, the production function $\varphi_s(\theta)$ is a linear function. Just as the aforementioned, $G(x) > 0$ holds for $x \in [0, 1]$ if $\frac{c}{r} < \frac{N-n}{n(N-1)}$, $G(x) = 0$ holds for $x \in [0, 1]$ if $\frac{c}{r} = \frac{N-n}{n(N-1)}$ and $G(x) < 0$ holds for $x \in [0, 1]$ if $\frac{c}{r} > \frac{N-n}{n(N-1)}$. The expression of $D(x)$ can be transformed into $D(x) = \sum_{j=0}^{k-1} \binom{n-2}{j} x^j (1-x)^j [(1-x)^{n-2-2j} -$

$x^{n-2-2j}][g(\frac{j+2}{n}) - 2g(\frac{j+1}{n}) + g(\frac{j}{n})]$ where $k = \frac{n-2}{2}$ holds for even n or $k = \frac{n-1}{2}$ holds for odd n. In the case of $s > 0$, $\varphi_s(\theta)$ is a symmetrically sigmoidal function and results into $G'(x) > 0$ for $x \in [0, 1/2)$ and $G'(x) < 0$ for $x \in (1/2, 1]$, thus $G(x)$ has a maximum at $x = 1/2$ and a minimum at $x = 0$ along with $x = 1$. In the case of $-0.5 < s < 0$, $\varphi_s(\theta)$ is a symmetrically inverse sigmoid function and results in $G'(x) < 0$ for $x \in [0, 1/2)$ and $G'(x) > 0$ for $x \in (1/2, 1]$, thus $G(x)$ has a minimum at $x = 1/2$ and a maximum at $x = 0$ along with $x = 1$. The solutions of $G(0) = 0$ and $G(\frac{1}{2}) = 0$ with respect to s are denoted by s_1^* and s_2^*, where the explicit expression of s_1^* is $s_1^* = \frac{\ln(1-\frac{2}{n})}{2\ln(1-2\frac{c}{r}\frac{N-1}{N-n})} - \frac{1}{2}$. The explicit expression of s_2^* is difficult to obtain for general n, whereas we have $s_2^* = \frac{\ln\frac{1}{2}}{2\ln 4\frac{c}{r}\frac{N-1}{N-4}-\frac{1}{2}} - \frac{1}{2}$ for $n = 4$. Note that $\varphi_s(\theta) - \varphi_0(\theta) = \varphi_0(1-\theta) - \varphi_s(1-\theta)$ holds for any $\theta < 1/2$, so we have $2^{n-1}[G(\frac{1}{2})\big|_{s=s_1} - G(\frac{1}{2})\big|_{s=0}] = 2\sum_{j=0}^{k-1}\left[\binom{n-1}{j} - \binom{n-1}{j+1}\right]\left[\varphi_{s_1}(\frac{j+1}{n}) - \varphi_0(\frac{j+1}{n})\right]$ where $k = \frac{n-2}{2}$ holds for even n or $k = \frac{n-1}{2}$ holds for odd n. Easily verified, we have $G(1/2)\big|_{0.5<s<0} < G(1/2)\big|_{s=0}$ and $G(1/2)\big|_{s>0} > G(1/2)\big|_{s=0}$.

- If $\frac{c}{r} > \frac{N-n}{n(N-1)}$ holds, we get $G(x)\big|_{s=0} < 0$ for $x \in [0, 1]$ and $s_1^* < 0 < s_2^*$:
  Due to $G(0)\big|_{s>0} < G(0)\big|_{s=0} < 0$ and $G(1/2)\big|_{-0.5<s<0} < G(1/2)\big|_{s=0} < 0$, both the minimum of $G(x)\big|_{s>0}$ and the minimum of $G(x)\big|_{-0.5<s<0}$ are negative. $G(x)\big|_{s>0} = 0$ or $G(x)\big|_{-0.5<s<0} = 0$ has two, one or no roots by the signs of $G(1/2)\big|_{s>0}$ and $G(0)\big|_{-0.5<s<0}$.

- If $\frac{c}{r} < \frac{N-n}{n(N-1)}$ holds, we get $G(x)\big|_{s=0} > 0$ for $x \in [0, 1]$ and $s_2^* < 0 < s_1^*$:
  Due to $G(1/2)\big|_{s>0} > G(1/2)\big|_{s=0} > 0$ and $G(0)\big|_{-0.5<s<0} > G(0)\big|_{s=0} > 0$, both the maximum of $G(x)\big|_{s>0}$ and the maximum of $G(x)\big|_{-0.5<s<0}$ are positive. $G(x)\big|_{s>0} = 0$ or $G(x)\big|_{-0.5<s<0} = 0$ has two, one or no roots by the signs of $G(0)\big|_{s>0}$ and $G(1/2)\big|_{-0.5<s<0}$.

- If $\frac{c}{r} = \frac{N-n}{n(N-1)}$ holds, we get $G(x)\big|_{s=0} = 0$ for $x \in [0, 1]$ and $s_2^* = 0 = s_1^*$:
  Due to $G(0)\big|_{s>0} < G(0)\big|_{s=0} = 0$ and $G(1/2)\big|_{s>0} > G(1/2)\big|_{s=0} = 0$, $G(x)\big|_{s>0} = 0$ has two roots. Due to $G(0)\big|_{-0.5<s<0} > G(0)\big|_{s=0} = 0$ and $G(1/2)\big|_{-0.5<s<0} < G(1/2)\big|_{s=0} = 0$, $G(x)\big|_{-0.5<s<0} = 0$ has two roots.

APPENDIX H: SIMULATION.

The population of size N is initialized with a homogeneous state. We asynchronously update the population by replacing a random focal individual α with an offspring as follows. The payoff of the focal individual α, F_α, is determined by randomly choosing an n-person interaction including the focal individual α. Then a reference individual β differing from a is randomly chosen, whose payoff, F_β, is obtained by another random n-person interaction. The focal individual α takes on the parent role with the probability

$w = \frac{F_\alpha - F_\beta}{c}$ (c guarantees $w \leq 1$), otherwise the reference individual β does so. The offspring stays the same strategy with the parent without mutation. If a mutant emerges (with the probability μ), the strategy of the offspring follows a Gaussian distribution with the parent strategy as mean and with a very small standard variance σ. We trace realistic trajectories of all individuals in a population. It has been rigorously proved that this evolutionary process converges in law to the solution of the adaptive dynamics [41], as the distribution variance of mutation steps goes to zero. Therefore, with a sufficiently small mutation variance, a single realization can represent the average evolution of the population.