Reward versus punishment: averting the tragedy of the commons in eco-evolutionary dynamics

Samrat Sohel Mondal*, Mayank Pathak and Sagar Chakraborty

Department of Physics, Indian Institute of Technology Kanpur, Uttar Pradesh 208016, India
* Author to whom any correspondence should be addressed.
E-mail: samrat@iitk.ac.in

Keywords: dynamic games, evolutionary games, noncooperative game, prisoner’s dilemma, natural resources

Abstract

We consider an unstructured population of individuals who are randomly matched in an underlying population game in which the payoffs depend on the evolving state of the common resource exploited by the population. There are many known mechanisms for averting the overexploitation (tragedy) of the (common) resource. Probably one of the most common mechanism is reinforcing cooperation through rewards and punishments. Additionally, the depleting resource can also provide feedback that reinforces cooperation. Thus, it is an interesting question that how reward and punishment comparatively fare in averting the tragedy of the common (TOC) in the game-resource feedback evolutionary dynamics. Our main finding is that, while averting the TOC completely, rewarding cooperators cannot get rid of all the defectors, unlike what happens when defectors are punished; and as a consequence, in the completely replete resource state, the outcome of the population game can be socially optimal in the presence of the punishment but not so in the presence of the reward.

1. Introduction

In human society, the fact that a man earns his just reward for his good deeds and just punishment for his crimes is the very foundation of the conception of both human and divine justice. Rational individuals—in the absence of any designed reinforcers like reward and punishment—selfishly defect from helping others with a view to maximizing their utilities while leading to the tragedy of the commons (TOC) [1–3], the overexploitation of the shared common resources. Uncontrolled population growth [3], water pollution and water crisis [4], pollution of the earth’s atmosphere [5], property rights, communal rights or state regulation [6], and wildlife crimes [7] are a few of the illustrating examples of the TOC. The TOC leads a system to a state of a depleted shared resource. For such systems, designing a proper reinforcement and feedback mechanism that regulates resource consumption is a burning issue.

Since the payoffs and strategies of the individuals in a population competing for a common resource should practically be dependent on the state of the resource, another interesting question presents itself: in the course of preventing the TOC, when can the outcome of the underlying population game be rendered socially optimal? Social optimality is the desired state of affairs for a social group. A strategic interaction is socially optimum if it maximizes the sum of the payoffs to all the individuals (players or agents) in the population [8]. Prevention of the TOC is obviously closely tied with a dominance of cooperators in the population. Reward and punishment are the most straightforward mechanisms for the emergence and sustenance of cooperation in both human and animal behavior. However, it is not apparent a priori whether reward and punishment act as two sides of the same coin or whether these are two kinds of reinforcers that play fundamentally distinct roles [9–13]. Thus, it is of interest to ask how effective they are in averting the TOC and which out of the two has more efficacy in furnishing a socially optimal outcome.

Reward and punishment were shown to exhibit a statistically equivalent positive effect on cooperation [14]. Sometimes [9, 15] their simultaneous action is desirable for averting the TOC. Both reward [16] and punishment [17] can initiate the first seed of cooperation for it to take over globally. Both are incapable of achieving
global cooperation where the population is divided into localized subgroups, and the payoffs are locally inefficient [18]. There are some subtle differences also, e.g., rewarding those who reward has been shown to be preferred to punishing those who do not punish [19, 20]. In an experimental setting of a variant of the public goods game, it was shown that punishment had a more substantial effect than reward on cooperation if considered by itself; however, reward had a more significant influence when attached with optional participation [21]. Of course, the effectiveness of reward and punishment in inducing cooperation is expected to be non-trivially dependent on their intensity [22] and the underlying game [23].

The TOC is omnipresent; it goes beyond the human behavior and action. It can also be observed in non-human biological evolutionary systems [24–37] where the players may have almost no rationality. Punishment is not unheard of in such systems [28, 38–41]. In such biological systems, the commons undergoing the tragedy have a rather more general interpretation: it either can be a pre-existing resource, or it can be a social good formed either by cooperation or by abstaining from conflict. Furthermore, the tragedy is classified as either collapsing or component depending on whether the corresponding resource is respectively completely or partially lost.

Obviously, the evolutionary game theory [42–45] is a relevant formalism to investigate the TOC and hence, the use of the paradigmatic replicator equation [45–50] in this context is only natural. For example, the replicator equation has been used in the public goods game to observe the success of altruistic punishment [51] and to show that reward may encourage cooperation but, unlike punishment, may fail to stabilize it [52]; in the latter complex dynamics and unpredictable oscillations are also observed. More importantly, in close relation to the central idea of the present paper, the game-resource feedback dynamics have been studied within the paradigm of the replicator equation to reveal how harsh punishments may not save the renewable resource [53].

2. Model

In this paper, we work with the simplest non-trivial setup of the TOC where an external resource is exploited by players—either cooperators or defectors—playing the prisoner’s dilemma game where defecting is the dominant strategy. We allow for the effect of punishment and reward by modifying the game payoff matrix. Furthermore, we allow for the feedback of the depleting resource in the game payoff matrix that is used to model the effective fitness of the cooperators whose frequency evolves in accordance with the replicator equation. The indirect [54] and direct [55] roles of environment in sustaining cooperation is well known. Subsequently, we do an insightful comparative study of the effects of the reward and the punishment. But before we discuss the results, let us first elaborate on the mathematical model that is central to our work.

2.1. Game-environment feedback: general formulation

Let us consider \( \mu \) different strategies that can be adopted by any individual member of a very large unstructured population of constant size. Let \( x_i(t) \) be the frequency with which \( i \)th strategy is used at any instant \( t \). The vector \( x(t) \equiv (x_1(t), x_2(t), \ldots, x_\mu(t)) \) defines the frequency distribution of the strategies over the consumer population. Let \( n(t) \equiv (n_1(t), n_2(t), \ldots, n_\nu(t)) \) denotes \( \nu \) different common shared resources in the environment. For convenience and without any loss of generality, we assume that for any \( j \) variable \( n_j(t) \) is normalized so as to lie in the unit interval, i.e., \( 0 \leq n_j(t) \leq 1 \). The state of the composite system consisting of the consumer population and the common shared resources can be represented by \( s(t) \equiv (x(t), n(t)) \). The resultant dynamical system is governed by \( \mu \) differential equations for \( x(t) \) (since \( \sum_{i=1}^{\mu} x_i = 1 \)) and \( \nu \) differential equations for \( n(t) \).

The state of the population evolves under the replication-selection process; the time-continuous replicator equation is a suitable model dynamic. The replicator equation can be obtained in several similar ways [45, 49, 56–58], either rigorously in the setting of one-locus-many-allele theory or phenomenologically for the interplay between the behavioural or strategic actions of the players. The fitness \( f_i \) of \( i \)th strategied individual, in full generality, should be a function of \( s(t) \equiv (x(t), n(t)) \). The per capita rate of increase of \( i \)th strategy, \( \frac{d}{dr}[\log x_i] \), is measure of evolutionary success; the basic tenet of Darwinism suggests that this evolutionary success can be expressed as the difference between fitness \( f(s(t)) \) of \( i \)th strategy and the average fitness \( \sum_{j=1}^{\mu} f_j(s(t)) \). Thus, we write

\[
\frac{dx_i}{dr} = x_i \left[ f_i(s(t)) - \sum_{j=1}^{\mu} f_j(s(t)) x_j \right],
\]

which is the replicator equation. To give an explicit functional form to the fitness function in the evolutionary game-theoretic set-up, we define a \( \mu \times \mu \) dimensional payoff matrix \( U(n) \), whose element \( U_{ij}(n) \) represents the
payoff obtained by an individual with ith strategy when it interacts with an opponent employing jth strategy. The payoff matrix is independent of $x(t)$ whose dependence comes into the fitness through following relation:

$$f(s) = \sum_{j=1}^{\mu} U_j(n)x_j.$$  \hspace{1cm} (2)

Complete description of the system requires specification of payoff matrix $U(n)$.

Finally, the model needs to be provided with a dynamics of state $n$ of the common shared resources. In general, these resources may have intrinsic dynamics in the absence of the exploitation by the consumer population. Let the intrinsic rate of change of $n_i(t)$ be $g^{(in)}_i(n(t))$ $(1 \leq i \leq \nu)$. Corresponding equations of motion can be written as

$$\frac{dn_i}{dt} = \lambda^{(in)}_i g^{(in)}_i(n(t)) + \lambda^{(ex)}_i g^{(ex)}_i(s(t)),$$  \hspace{1cm} (3)

where $g^{(ex)}_i(s(t))$ is the effect of consumption of consumer population on ith resource. The parameters $\lambda^{(in)}_i$ and $\lambda^{(ex)}_i$ respectively determine the speeds of the intrinsic and extrinsic dynamics of ith resource [60].

Equations (1) and (3) constitute the set of dynamical equations that describe the eco-evolutionary dynamics of the composite system’s state, $s(t) \equiv (x(t), n(t))$.

### 2.2. Social dilemmas and external reinforcers

We now adopt the simplest non-trivial specific setting that is conducive to studying the TOCs with added reinforcers, e.g., rewards and punishments. Often, the rise of the defectors who lead to the TOC is exemplified through the prisoner’s dilemma game ($T > R > P > S$) [61]—a one-shot two-player–two-strategy game ($\mu = 2$)—in which the players play non-Pareto-optimal [62] Nash equilibrium [63] strategy (defection) although mutual cooperation can fetch the players comparatively more payoff. Its normal bi-matrix form is given by

$$\begin{bmatrix} R, R & S, T \\ T, S & P, P \end{bmatrix},$$

where the first and the second comma-separated elements in each entry are the payoffs of two respective players. We denote by $x_1$ and $x_2$ as the fractions of ‘cooperators’ and ‘defectors’, respectively. Here, it is obvious that $\mu = 2$. Since $x_1(t)$ and $x_2(t) = 1 - x_1(t)$ are not independent variables, we use $x(t) \equiv x_1(t)$ as the only variable to specify the state of the consumer population. Furthermore, in order to engage in a simplified discussion of the eco-evolutionary model discussed above, we choose $\nu = 1$. Therefore we may write $s(t) \equiv (x(t), n(t))$ as the simplified determiner of the composite state in this case.

As we discussed earlier, our goal is to study the relative effects of the reinforcers like reward and punishment. To this end, we could assume that if a player defects against a cooperator, she is punished by reducing the payoff by a positive real number $p$. Similarly, we assume that if a player cooperates against a defector, she is rewarded with an additional payoff of $r$, a positive real number.

### 2.3. Reinforcement by shared resource’s feedback

In addition to the effect of the external reinforcers like reward and punishment, we must consider an inherent reinforcing mechanism introduced by the shared resource’s feedback. When the individuals in the population start to exploit the shared resource, the state of the shared resource depletes. Consequently, there is a potential change in the interactions pattern amongst the individuals. This concept is mathematized by the form of the total payoff matrix $U(n)$ as given below [60, 64–66]

$$U(n) = nU_1 + (1 - n)U_0.$$  \hspace{1cm} (4)

In the above equation $U_k$ $(k \in \{0, 1\})$ is the symbolic notation for $U(n = k)$, where $k = 0$ signifies the state of the completely depleted common resource and on the other hand $k = 1$ is for the fully replete state. Therefore, in the presence of reward or punishment, the mathematical form of $U_1$ is respectively

$$U_1 = \begin{bmatrix} R_1 & S_1 + r \\ T_1 & P_1 \end{bmatrix} \quad \text{or} \quad U_1 = \begin{bmatrix} R_1 & S_1 \\ T_1 - p & P_1 \end{bmatrix}. \hspace{1cm} (5)$$

Additionally, we note from equation (4), as the depletion starts the changed interaction pattern due to the feedback of the shared resource is mathematically reinforced through $U_0$ which for the consistency of the notations can be displayed in its most general form as

$$U_0 = \begin{bmatrix} R_0 & S_0 \\ T_0 & P_0 \end{bmatrix}. \hspace{1cm} (6)$$
In our set up, individuals in the fully replete state are involved in selfish behaviour modelled by the prisoner’s dilemma game matrix, i.e., \( T_1 > R_1 > P_1 > S_1 \). But as the degradation of the resource starts the cooperation can emerge through reinforcement. The form of \( U_0 \), thus, could deviate from the prisoner’s dilemma.

In general, depending on how the preferences of the players change as degradation of the shared resource takes place, the \( U_0 \) could be the payoff matrix for any of the four exhaustive and mutually exclusive classes of games \([67–71]\) classified based on the correspondence of the Nash equilibria with cooperation and defection:

(a) Harmony game: \( R_0 > T_0 \) and \( S_0 > P_0 \); the only symmetric Nash equilibrium is mutual cooperation.
(b) Anti-coordination game: \( R_0 < T_0 \) and \( S_0 > P_0 \); there exists a unique mixed symmetric Nash equilibrium in which the players play a mixed strategy randomized over the pure strategies.
(c) Prisoner’s dilemma: \( T_0 > R_0 > P_0 > S_0 \); the mutual defection is the only symmetric Nash equilibrium.
(d) Coordination game: \( R_0 > T_0 \geq P_0 > S_0 \); there are two symmetric pure Nash equilibria (mutual defection and mutual cooperation) and one mixed symmetric Nash equilibrium similar to the one in the anti-coordination game.

2.4. Dynamics of cooperators and common resource

With the choice of the form of equation (4) for the payoff matrix, the variations of the fraction of cooperators with time is governed by the following equation

\[
\frac{dx}{dt} = x \left[ \sum_{j=1}^{2} U_{ij}(n)x_j - \sum_{i=1}^{2} \sum_{j=1}^{2} U_{ij}(n)x_i x_j \right].
\]  

(7)

In our model, we ignore the intrinsic dynamics of the resource, for simplicity. We incorporate the fact the cooperator consumers can cause augmentation of the resource while there is always deterioration of the anti-cooperator consumers. If \( \theta > 0 \) is the ratio of the efficacy of augmentation rate to deterioration rate, then it has been shown \([64]\) that a simple model for the state of the shared resource \( n \) can be written as

\[
\frac{dn}{dt} = \lambda_1^{\text{ex}} g_1^{\text{ex}}(s(t)) \equiv \epsilon \{ \theta x - (1 - x) \} n(1 - n),
\]  

(8)

where parameter \( \epsilon = \lambda_1^{\text{ex}} \ll 1 \) because the dynamic of the environment is assumed to be slower than the dynamics of strategists’ frequencies.

Equations (7) and (8) constitute the set of dynamical equations that mathematically describes the eco-evolutionary dynamics of cooperators and defectors consuming a single shared resource in the additional presence of the external reinforcement through reward or punishment.

3. Results

For future convenience we first introduce four parameters \([60, 64–66]\) that can be intuitively interpreted as four distinct types of incentives for changes in the strategies:

\[
\Delta_{RT}^1 \equiv R_1 - T_1, \\
\Delta_{SP}^1 \equiv S_1 - P_1, \\
\Delta_{RT}^0 \equiv R_0 - T_0, \\
\Delta_{SP}^0 \equiv S_0 - P_0.
\]  

(9a), (9b), (9c), (9d)

In the literature \([72–75]\), \(- \Delta_{RT}^0 \) and \(- \Delta_{SP}^1 \) are the gamble-intending dilemma strengths, and \(- \Delta_{SP}^1 \) and \(- \Delta_{RT}^1 \) are the risk averting dilemma strengths for completely depleted or replete shared resource states respectively. It is easy to see that by construction, the four quadrants of \( \Delta_{RT}^1 - \Delta_{SP}^1 \) plane correspond to the aforementioned four classes of games: the first quadrant is for harmony game, the second quadrant is for coordination game, the third quadrant is for prisoner’s dilemma, and the fourth quadrant is for the anti-coordination game. (As is customary, the quadrant numbers are counted anticlockwise starting from the first quadrant in the top-right quadrant in figure 1.)

Furthermore, we also introduce a parameter, \( \delta_0 \equiv \Delta_{RT}^0 / \Delta_{SP}^0 = (\Delta_{RT}^1 / - \Delta_{SP}^1) \), that quantifies by what multiplicative factor a player has more affinity to defect against a cooperator than against a defector, when the shared resource state is in the worst possible state, i.e., \( n = 0 \). We draw two lines with \( \delta_0 = \Delta_{RT}^1 / \Delta_{SP}^1 \) and with \( \delta_0 = - \theta \), to divide \( \Delta_{RT}^1 - \Delta_{SP}^1 \) plane into seven distinct regions as seen in figure 1.
3.1. Equilibrium states

Now we perform the linear stability analysis of the eco-evolutionary dynamics, i.e., equations (7) and (8) along with equations (4)–(6).

For the case of either reward or punishment, we get a total of seven fixed points \((x^*, n^*)\) and four of them are the four corners of the phase space, i.e., \((0, 0), (0, 1), (1, 0), \) and \((1, 1)\) which are common for both the cases. These four fixed points always exist for all parameters values. Another common fixed point, \((x^*_b, 0)\), lies on the bottom edge \((n = 0)\) of the \(x-n\) phase space, where

\[
x^*_b = \frac{\Delta^0_{SP}}{\Delta^0_{SP} - \Delta^0_{RT}}.
\]

This fixed point exists if \(\delta_0 < 0\).

However, the fixed points on the top edge \((n = 1)\) and in the interior of the phase space are different depending on the reinforcers. For the case of punishment, the fixed point in the top edge is \((x^*_p, 1)\), where \(x^*_p\) is given by

\[
x^*_p = \frac{\Delta^1_{SP}}{-\Delta^1_{RT} + \Delta^1_{SP} - p}.
\]

and the interior fixed is \((x^*_p, n^*_p)\), where

\[
x^*_p = \frac{1}{1 + \theta},
\]

\[
n^*_p = \frac{\Delta^0_{RT} + \theta \Delta^0_{SP}}{(\Delta^0_{RT} + \theta \Delta^0_{SP}) - (\Delta^1_{RT} + \theta \Delta^1_{SP}) - p}.
\]

For the case of reward, the fixed point in the top edge is \((x^*_r, 1)\), where

\[
x^*_r = \frac{r + \Delta^1_{SP}}{r + \Delta^1_{SP} - \Delta^1_{RT}},
\]

and the interior fixed is \((x^*_r, n^*_r)\), where
Table 1. We tabulate the conditions that need to be satisfied for a fixed point to acquire its possible type in the case of reward. The $+/−$ sign refers to the sign that the corresponding parameter must take for the fixed point to be of the type mentioned in the last column. ‘na’ (not applicable) stands for the fact that the nature of the fixed point is independent of the corresponding parameter. Here, $ξ_i ≡ (θ + δ_i) − (θ + δ_i^*)$.

| Fixed point | $Δ_{RT}^1$ | $r + Δ_{RT}^1$ | $Δ_{SP}^0$ | $θ + δ_i$ | $θ + δ_i^*$ | $ξ_i$ | Type |
|-------------|------------|----------------|----------|-----------|-----------|-------|------|
| (0, 0)      | na         | na             | −        | na        | na        | Stable node |
|             | na         | na             | +        | na        | na        | Saddle |
| (0, 1)      | na         | +              | na       | na        | na        | Unstable node |
|             | na         | −              | na       | na        | na        | Saddle |
| (1, 0)      | −          | na             | −        | na        | −         | Unstable node |
|             | +          | na             | +        | na        | na        | Saddle |
| (1, 1)      | na         | na             | na       | −         | −         | Saddle |
| (x_i*, 0)   | −          | na             | −        | −         | −         | Stable node |
|             | +          | na             | +        | −         | +         | Saddle |
| (x_i*, 1)   | na         | +              | na       | −         | −         | Saddle |
|             | na         | −              | −        | −         | −         | Unstable node |

$x_{ip}^* = \frac{1}{1 + θ^*}$, \hspace{1cm} (14a)

$n_{ip}^* = \frac{θΔ_{SP}^0 + Δ_{RT}^0}{(Δ_{RT}^0 + θΔ_{SP}^0) − (Δ_{RT}^0 + θΔ_{SP}^0) − rθ^*}$, \hspace{1cm} (14b)

It is important to note that the locations of fixed points on the top edge and the interior of the phase space are dependent on the value of the reinforcement; for certain values of $r$ or $p$, these fixed points can leave the physically allowed phase space and become effectively non-existent. For the case of reward, $(x_{ip}^*, 1)$ exists if $δ_i ≡ Δ_{RT}^1/(r + Δ_{RT}^1) < 0$. Whereas for the case of punishment, the analogous fixed point, $(x_{ip}^*, 1)$ exists if $δ_i ≡ (p + Δ_{RT}^1)/Δ_{SP}^1 < 0$. Also, the interior fixed points $(x_{ip}^*, n_{ip}^*)$ and $(x_{ip}^*, n_{ip}^*)$ exist if the inequalities \((r + Δ_{SP}^1)(θ + δ_i)/[Δ_{SP}^0(θ + δ_i)] < 0\) and \(Δ_{SP}^1(θ + δ_i)/[Δ_{SP}^0(θ + δ_i)] < 0\) are respectively satisfied.

Our next goal is to find out the stability of the aforementioned fixed points, which corresponds physically to the composite system’s equilibrium states. We accomplish this by employing the linear stability analysis exhaustively for all the cases under consideration. Next, we calculate $2 \times 2$ Jacobian matrices at all the fixed points for the linearized dynamical system (equations (7) and (8)) and find the corresponding eigenvalues whose signs decide the type (hence, stability) of the fixed points (see tables 1 and 2).

In figure 1, we present the gist of the mathematical results pictorially. In the figure, we elaborately show the bifurcations in the phase portraits as the intensity of reward and punishment are increased. While reading the figure, one should translate any stable fixed point as an equilibrium state that can eventually be reached starting from some initial state in the corresponding basin of attraction; any unstable fixed point—unstable focus, unstable node, or saddle—is not realized in practice. A close inspection of the figure reveals some interesting features of the system that we discuss in what follows.

### 3.2. Bistability

The very first thing we note is the phenomenon of bistability, which is an omnipresent feature of many natural and man-made systems. In the case under study, the bistability partitions (almost all) initial conditions into two mutually exclusive classes depending on where the initial conditions are eventually attracted. The two classes can correspond to two of the three possible equilibrium states of the resource, viz, complete prevention of the TOC ($n = 1$), partial prevention of the TOC ($0 < n < 1$), and complete realization of the TOC ($n = 0$).
Table 2. We tabulate the conditions that need to be satisfied for a fixed point to acquire its possible type in the case of punishment. The +/− sign refers to the sign that the corresponding parameter must take for the fixed point to be of the type mentioned in the last column. ‘na’ (not applicable) stands for the fact that the nature of the fixed point is independent of the corresponding parameter.

| Fixed point | $\Delta_{1}$p | $p + \Delta_{1}$ | $\Delta_{p}$ | $\theta + \delta_{p}$ | $\theta + \delta_{1}$ | $\zeta$ | Type   |
|-------------|---------------|-----------------|--------------|------------------------|------------------------|--------|--------|
| (0, 0)      | na            | na              | na           | na                      | na                      | Stable node | Saddle |
| (0, 1)      | na            | na              | na           | na                      | na                      | Saddle       |        |
| (1, 0)      | −             | na              | na           | na                      | na                      | Unstable node | Saddle |
| (1, 1)      | na            | +               | na           | na                      | na                      | Stable node   | Saddle |
| $(x_{*}', 0)$ | −             | na              | na           | −                       | na                      | Stable node   |        |
| $(x_{*}', 1)$ | na            | +               | na           | −                       | na                      | Unstable node | Saddle |
| $(x_{*}', n_{1}^p)$ | na            | −               | −            | +                       | na                      | Saddle       |        |
|              | na            | −               | +            | −                       | na                      | Stable node   |        |
|              | na            | −               | −            | −                       | na                      | Stable node   |        |

The bistability is not realized in quadrant I (refer figure 1) in the presence of either reward or punishment. In passing, we remark that in this quadrant, oscillatory TOC (mathematically modelled by a heteroclinic cycle [64]) is known to be present in the absence of any reward or punishment, whose presence beyond a certain threshold, however, averts the oscillatory TOC. In the remaining three quadrants, both reward and punishment can affect bistability. However, the exact nature of the bistability is dependent on whether reward or punishment is being employed.

We note that while the emergence of bistability leads to prevention of the TOC that is inevitable in the absence of reward or punishment (refer quadrant II, quadrant III, and region IVA in figure 1), the prevention is not guaranteed; rather, it depends on the initial state of the system in addition to the extent of reward and punishment. In these cases, when the cooperator fraction is too low in an extremely poor state of the resource, any finite amount of reward and punishment cannot avert the TOC.

We also note (in region IA and region IVB in figure 1) that if partial TOC is present in the absence of reward or punishment, enough reward or punishment leads to complete prevention of the TOC. Unlike the case of the reward, the complete prevention of the TOC due to punishment passes through a stage of bistability (region IVB in figure 1(b)) between the states of partial TOC and completely averted TOC.

3.3. Socially optimal outcome

One notes that the most replete states of the resource effected by reward and punishment differ in an important respect: the fraction of defectors in the case of reward is not zero, unlike what happens in the case of the punishment. This observation has an interesting consequence that we discuss below.

We find that high enough punishment leads to a stable rich-environment state which furnishes socially-optimal outcome for the underlying population game. This is rather easy to see when we note that the stable rich-environment state is actually $n = 1$. Consider the payoff matrix, $U(n)$ when environment is replete, i.e., $n = 1$ (reference equation (4)). Linear stability analysis tells us that when the punishment is high enough---$p > -\Delta_{1}p$, to be precise—the stable fixed point $(1, 1)$ comes into existence, denoting that in a rich environment, the population is completely overwhelmed by the cooperators. In the context of the payoff elements of $U(1)$, this means that $R_1$ is the largest element because $R_1 > (T_1 - p)$ and $R_1 > P_1 > S_1$ (by construction). Since $R_1$ is the largest element, this is the maximum average payoff that can be achieved in the population. Obviously, the average payoff of the population would simply be $R_1$ if only cooperators are present; any other composition

---

*J. Phys. Complexity, 3 (2022) 025005 (11pp)*

SS Mondal
of the population would result in a smaller average payoff. Hence, the richest resource state, when achieved with only cooperators around, produces the socially optimal outcome for the population game in the presence of enforced punishment.

In contrast, high enough reward does not lead to a socially-optimal outcome for the population game in the rich-environment state. Incidentally, here also, the most replete state achievable corresponds to \( n = 1 \). The linear stability analysis concludes that high enough reward—specifically, \( r > -\Delta_{SP}^1 \)—is required for the fixed point \((x^*_r, 1)\) to become attracting. Furthermore, recall that the payoff matrix, \( U(n) \) when environment is replete, i.e., \( n = 1 \) (reference equation (4)). The average payoff of the whole population consisting of \( x \) fraction of cooperators and \((1-x)\) fraction of defectors is

\[
A(x) \equiv x^2 (R_1 - S_1 - r - T_1 + P_1) + x(S_1 + r + T_1 - 2P_1) + P_1,
\]

which is extremized (i.e., \( dA/dx = 0 \)) at

\[
x = x_r \equiv \frac{S_1 + T_1 - 2P_1 + r}{2(S_1 + T_1 - R_1 - P_1 + r)}.
\]

The second order derivative of \( A(x) \),

\[
\frac{d^2}{dx^2} A(x) \bigg|_{x=x_r} = 2[\Delta_{RT}^1 - (\Delta_{SP}^1 + r)],
\]

is a negative quantity implying that the extrema is a maximum. However, as already noted, the TOC-averted rich state due to reward is \((x^*_r, 1)\) and not \((x_r, 1)\). Hence, unlike punishment, reward induced TOC-averted resource state is incapable of furnishing a socially optimal outcome for the population game.

4. Discussion and conclusion

In this paper, we have extended the framework of deterministic eco-evolutionary dynamics to explore the effects of reward and punishment on the TOC. We have employed linear stability analysis to exhaustively study the nonlinear dynamical system. We have found that the phenomenon of bistability is brought forth through reward and punishment. However, our study reveals that these two kinds of reinforcers have fundamentally distinct roles in the prevention of the TOC.

Reward and punishment are not the same reinforcing factor acting oppositely. Intuitively, in a system with the replete resource in the presence of high cooperation level, the reward as a mechanism for sustaining cooperation is quite costly as all cooperators are supposed to be rewarded. However, if the system implements punishment, it is not relatively cheaper because the system has to punish only the defector fraction of the population; recall that in the replete state, defector and cooperators coexist in the case of punishment. This is the source of the asymmetric impact of reward and punishment on the prevention of the TOC. Our study shows that within the framework of our study, harsh punishment—unlike generous reward—results in a socially optimal outcome when the TOC is averted completely.

The eco-evolutionary dynamics we have considered, by construction, models the mean-field behaviour of a well-mixed population of individuals sharing a common resource. In reality, however, any population is usually structured. In such a case, one needs to go beyond the usage of differential equations for modelling the dynamics and adopt the relevant tools of non-equilibrium statistical mechanics. Of course, the most general set-up of a finite (arbitrarily) structured population is analytically intractable and can only be handled in a case-by-case manner using agent-based simulations. Nevertheless, remarkably, in the limit of weak selection, the frequencies of strategies in a structured large population can be modelled using the replicator equation with a transformed payoff matrix if the interaction structure is that of a regular graph with one player at each vertex. The transformed matrix depends on the local strategy update rules. Hence, the eco-evolutionary dynamics discussed in this paper can be straightforwardly adopted to at least such structured populations.

The folk theorem of evolutionary game theory (along with similar theorems for the evolutionarily stable strategy) establishes that many characteristics of the asymptotic dynamics of the replicator equation can be gauged merely by analyzing the corresponding payoff matrix for the Nash equilibrium and the evolutionarily stable strategy. These theorems have paved the way for (game-theoretic type) strategic reasoning in the field of population biology. Thus, when one notes that the outcome of the analyses presented in this paper can also be understood through the strategic reasoning using the effective (\( n \)-dependent) payoff matrix \( U(n) \), one should understand it as a paradigm analogous to that created by the aforementioned folk theorem; the only difference now is that the replicator equation is coupled with an equation for the common resource. Consequently, on the face of it, the mathematical framework presented herein might appear as an overkill.
However, we recall that mere existence of some game theoretic equilibria does not mean they are realized in practice as well because they may be dynamically unstable. Moreover, for games with many strategies and games with discrete-time dynamics, more complex dynamics [69, 71, 79–81] (like oscillations, invariant cycles, and chaos) beyond the simple convergent (fixed point type) solutions appear. Such complex solutions are known not to be related to any standard game-theoretic equilibrium concept (cf [82, 83]). Thus, the eco-evolutionary mathematical framework becomes indispensable in such situations.

We believe that the mathematical framework explored in this paper opens up the possibility of further exciting research themes in the game-resource feedback dynamics. For example, one could construct a model that takes the effect of changing population size [66] into account. Specifically, one could consider a renewable resource with intrinsic growth dynamics and study the effect of a finite growing population on the mechanism of reward and punishment. More elaborately, one could develop microscopic stochastic birth–death models for both the population and the resource with reward and punishment realized respectively by enhanced birth and death rates. Also, the impact of simultaneous action [84] of reward and punishment, and the time delayed effect [85] of players’ actions on a common resource are also worth investigating within the framework of this paper. Lastly, we hope that some experiments can be designed to put the conclusions of this paper to test.

Acknowledgments

The authors are thankful to Joy Das Bairagya and Debashish Chowdhury for helpful discussions. SC acknowledges support of SERB (DST, govt. of India) project MTR/2021/000119.

Data availability statement

No new data were created or analysed in this study.

ORCID iDs

Samrat Sohel Mondal  
https://orcid.org/0000-0003-2993-2234

Sagar Chakraborty  
https://orcid.org/0000-0001-7568-0598

References

[1] Malthus T and Gilbert G 1999 An Essay on the Principle of Population, Oxford World’s Classics (Oxford: Oxford University Press)
[2] Lloyd W 1833 Two Lectures on the Checks to Population (Oxford: J.H. Parker)
[3] Hardin G 1968 The tragedy of the commons Science 162 1243
[4] Shiklomanov I A 2000 Appraisal and assessment of world water resources Water Int. 25 11
[5] Jacobson M and Jacobson P 2002 Atmospheric Pollution: History, Science, and Regulation (Cambridge: Cambridge University Press)
[6] Ostrom E 1999 Coping with tragedies of the commons Annu. Rev. Polit. Sci. 2 493
[7] Moreno W D and Pires S F 2011 Preventing wildlife crimes: solutions that can overcome the tragedy of the commons Eur. J. Crim. Policy Res. 17 101
[8] Davis O A and Whinston A 1962 Externalities, welfare, and the theory of games J. Political Econ. 70 241
[9] Kareva I, Morin B and Karev G 2013 Preventing the tragedy of the commons through punishment of over-consumers and encouragement of under-consumers Bull. Math. Biol. 75 565
[10] Hintze A, Staudacher J, Gelhar K, Potthann A, Rasch J and Wildegger D 2020 Inclusive groups can avoid the tragedy of the commons Sci. Rep. 10 22392
[11] Steel A, Silson E H, Stagg C J and Baker C I 2016 The impact of reward and punishment on skill learning depends on task demands Sci. Rep. 6 36056
[12] Liu J, Li M, Alam M, Chen Y and Wu T 2019 A game theoretic reward and punishment unwanted traffic control mechanism Mobile Netw. Appl. 24 1279
[13] Van Lange P, Rockenbach B and Yamagishi T 2014 Reward and Punishment in Social Dilemmas (Series in Human Cooperation) (Oxford: Oxford University Press)
[14] Balliet D, Mulder L B and Van Lange P A M 2011 Reward, punishment, and cooperation: a meta-analysis Psychol. Bull. 137 594
[15] Fang Y, Benko T P, Perc M, Xu H and Tan Q 2019 Synergistic third-party rewarding and punishment in the public goods game Proc. R. Soc. A 475 20190349
[16] Göös A R, Santos F P, Pacheco J M and Santos F C 2019 Reward and punishment in climate change dilemmas Sci. Rep. 9 16193
[17] Johnson S 2015 Escaping the tragedy of the commons through targeted punishment R. Soc. Open Sci. 2 150223
[18] Ozono H, Kamijo Y and Shimizu K 2020 The role of peer reward and punishment for public goods problems in a localized society Sci. Rep. 10 8211
[19] Kiyonari T and Barclay P 2008 Cooperation in social dilemmas: free riding may be thwarted by second-order reward rather than by punishment J. Person. Soc. Psychol. 95 826
[20] Barclay P and Kiyonari T 2014 Why sanction? Functional causes of punishment and reward Reward and Punishment in Social Dilemmas (Oxford: Oxford University Press) pp 182–96
[21] Bravo G and Squazzoni F 2013 Exit, punishment and rewards in commons dilemmas: an experimental study PLoS One 8 e69871
Jiang L-L, Perc M and Szolnoki A 2013 If cooperation is likely punish mildly: insights from economic experiments based on the snowdrift game PLoS One 8 e64677

Stoop J, van Soest D and Vyrastekova J 2018 Rewards and cooperation in social dilemma games J. Environ. Econ. Manag. 88 300

Gersani M, Brown J S, O’Brien E E, Maina G M and Abramsky Z 2001 Tragedy of the commons as a result of root competition J. Ecol. 89 660

Zea-Cabrera E, Iwasa Y, Levin S and Rodriguez-Iturbe I 2006 Tragedy of the commons in plant water use Water Resour. Res. 42 W06102

O’Brien E E and Brown J S 2008 Games roots play: effects of soil volume and nutrients J. Ecol. 96 438

Baquero F and Campos J 2003 The tragedy of the commons in antimicrobial chemotherapy Rev. Esp. Quimioter. 16 11

Rankin D J, Bargum K and Kokko H 2007 The tragedy of the commons in evolutionary biology Trends Ecol. Evol. 22 643

MacLean R C 2008 The tragedy of the commons in microbial populations: insights from theoretical, comparative and experimental studies Heredity 100 471

Smith P and Schuster M 2019 Public goods and cheating in microbesCurr. Biol. 29 R442

Schuster M, Fonał, E, Finch D, Smith H and De Leenheer P 2017 Tragedy of the commons in the chemostat PLoS One 12 1

Rainey P B and Rainey K 2003 Evolution of cooperation and conflict in experimental bacterial populations Nature 425 72

West S A and Buckling A 2003 Cooperation, virulence and siderophore production in bacterial parasites Proc. R. Soc. B 270 37

Bourke A F G and Ratnieks F L W 1999 Kin conflict over caste determination in social hymenoptera Behav. Ecol. Sociobiol. 46 287

Moreno-Camacho C A, Montoya-Torres J R, Jaegler A and Gondran N 2019 Sustainability metrics for real case applications of the supply chain network design problem: a systematic literature review J. Cleaner Prod. 231 600

Gore J, Youk H and van Oudenaarden A 2009 Snowdrift game dynamics and facultative cheating in yeast Nature 459 253–6

Crespi B J 2001 The evolution of social behavior in microorganisms Trends Ecol. Evol. 16 178

Frank S A 1995 Mutual policing and repression of competition in the evolution of cooperative groups Nature 377 520

Clutton-Brock T H and Parker G A 1995 Punishment in animal societies Nature 373 209

Ratnieks F L W and Wenseleers T 2005 Policing insect societies Science 307 54

Kiers T, Rousseau R and Denison R 2006 Measured sanctions: legume hosts detect quantitative variation in rhizobium cooperation and punish accordingly Evol. Ecol. Res. 8 1077–86

Smith J M and Price G R 1973 The logic of animal conflict Nature 246 44

Smith J 1982 Evolution and the Theory of Games (Cambridge: Cambridge University Press)

Nowak M 2006 Evolutionary Dynamics: Exploring the Equations of Life (Cambridge, MA: Harvard University Press)

Hofbauer J and Sigmund K 1998 Evolutionary Games and Population Dynamics (Cambridge University Press)

Taylor P D and Jonker L B 1978 Evolutionary stable strategies and game dynamics Math. Biosci. 40 145

Schuster P and Sigmund K 1983 Replicator dynamics J. Theor. Biol. 100 533

Schuster P and Sigmund K 1985 Dynamics of evolutionary optimization Berichte der Bunsengesellschaft f¨ur physikalische Chemie 89 668

Page K M and Nowak M A 2002 Unifying evolutionary dynamics J. Theor. Biol. 219 93

Cressman R, Ansell C and Binmore K 2003 Evolutionary Dynamics and Extensive Form Games, Economic Learning and Social Evolution (Cambridge, MA: MIT Press)

Greenwood G W 2016 Altruistic punishment can help resolve the tragedy of the commons social dilemmas 2016 IEEE Conf. Computational Intelligence and Games (CIG) (Piscataway, NJ: IEEE)

Hauert C 2010 Replicator dynamics of reward & reputation in public goods games J. Theor. Biol. 267 22

Chen X and Szolnoki A 2018 Punishment and inspection for governing the commons in a feedback–evolving game PLoS Comput. Biol. 14 e1006347

Szolnoki A, Antonioni A, Tomassini M and Perc M 2014 Binary birth–death dynamics and the expansion of cooperation by means of self-organized growth Europhys. Lett. 105 48001

Szolnoki A and Chen X 2017 Environmental feedback drives cooperation in spatial social dilemmas Europhys. Lett. 120 58001

Rice S 1961 Evolutionary Theory: Mathematical and Conceptual Foundations 1st edn (Massachusetts: Sinauer Associates, Inc.)

Traulsen A, Claussen J C and Hauert C 2005 Coevolutionary dynamics: from finite to infinite populations Phys. Rev. Lett. 95 238701

Mukhopadhyay A and Chakraborty S 2021 Replicator equations induced by microscopic processes in nonoverlapping population playing bimatrix games Chaos 31 023123

Cressman R and Tao Y 2014 The replicator equation and other game dynamics Proc. Natl Acad. Sci. USA 111 10810

Tilman A R, Plotkin J B and Akc ¸ay E 2020 Evolutionary games with environmental feedbacks Nat. Commun. 11 913

Rapoport A and Chammah A 1965 Prisoner’s Dilemma (Ann Arbor, MI: University of Michigan Press)

Pareto V 1964 Cours d’´Economie Politique. Professoré a l’Université de Lausanne ed F Rouge (Lausanne: Librairie de l’Université) pp 1896–7

Nash J F 1950 Equilibrium points in n-person games Proc. Natl Acad. Sci. USA 36 48

Weitz J S, Eksin C, Paarporn K, Brown S P and Ratcliff W C 2016 An oscillating tragedy of the commons in replicator dynamics with game-environment feedback Proc. Natl Acad. Sci. USA 113 E7518

Lin Y-H and Weitz J S 2019 Spatial interactions and oscillatory tragedies of the commons Phys. Rev. Lett. 122 148102

Bairagya I D, Mondal S S, Chowdhury D and Chakraborty S 2021 Game-environment feedback dynamics in growing population: effect of finite carrying capacity Phys. Rev. E 104 044407

Rapoport A 1967 A note on the index of cooperation for prisoner’s dilemma J. Conflict Resol. 11 100

Hummerc S, Bohl K, Basanta D, Deutsch A, Werner S, Theißen G, Schroeter A and Schuster S 2014 Evolutionary game theory: cells as players Mol. BioSyst. 10 3014

Pandit V, Mukhopadhyay A and Chakraborty S 2018 Weight of fitness deviation governs strict physical chaos in replicator dynamics Chaos 28 033104

Hilbe C, Simsa S, Chatterjee K and Nowak M A 2018 Evolution of cooperation in stochastic games Nature 559 246

Mittal S, Mukhopadhyay A and Chakraborty S 2020 Evolutionary dynamics of the delayed replicator-mutator equation: limit cycle and cooperation Phys. Rev. E 101 024210

Wang Z, Kokubo S, Jusup M and Tanimoto J 2015 Universal scaling for the dilemma strength in evolutionary games Phys. Life Rev. 14 1

Tudge S J and Brede M 2015 A tale of two theorems Phys. Life Rev. 14 49

Areﬁn M R, Kabir K M A, Jusup M, Ito H and Tanimoto J 2020 Social efficiency deﬁcit deciphers social dilemmas Sci. Rep. 10 1
[75] Ito H and Tanimoto J 2018 Scaling the phase-planes of social dilemma strengths shows game-class changes in the five rules governing the evolution of cooperation R. Soc. Open Sci. 5 181085
[76] Perc M, Jordan J J, Rand D G, Wang Z, Boccaletti S and Szolnoki A 2017 Statistical physics of human cooperation Phys. Rep. 687 1
[77] Ohtsuki H and Nowak M A 2006 The replicator equation on graphs J. Theor. Biol. 243 86
[78] Hofbauer J and Sigmund K 2003 Evolutionary game dynamics Bull. Am. Math. Soc. 40 479
[79] Weissing F J 1991 Evolutionary stability and dynamic stability in a class of evolutionary normal form games Game Equilibrium Models I (Berlin: Springer) pp 29–97
[80] Vilone D, Robledo A and Sánchez A 2011 Chaos and unpredictability in evolutionary dynamics in discrete time Phys. Rev. Lett. 107 038101
[81] Mukhopadhyay A, Chakraborty S and Chakraborty S 2021 Chaos and coexisting attractors in replicator-mutator maps J. Phys. Complex. 2 035005
[82] Mukhopadhyay A and Chakraborty S 2020 Periodic orbit can be evolutionarily stable: case study of discrete replicator dynamics J. Theor. Biol. 497 110288
[83] Mukhopadhyay A and Chakraborty S 2020 Deciphering chaos in evolutionary games Chaos 30 121104
[84] Sun W, Liu L, Chen X, Szolnoki A and Vasconcelos V V 2021 Combination of institutional incentives for cooperative governance of risky commons iScience 24 102844
[85] Yan F, Chen X, Qiu Z and Szolnoki A 2021 Cooperator driven oscillation in a time-delayed feedback-evolving game New J. Phys. 23 053017