PROCEEDINGS A

rspa.royalsocietypublishing.org

Research

Article submitted to journal

Subject Areas:
91A22,91Cxx,60J10

Keywords:
institutional incentives, evolutionary game theory, evolution of cooperation

Author for correspondence:
M. H. Duong
e-mail: h.duong@bham.ac.uk

Cost efficiency of institutional incentives for promoting cooperation in finite populations

Manh Hong Duong[1], The Anh Han[2]

[1]School of Mathematics, University of Birmingham, Birmingham B15 2TT, UK.
[2]School of Computing, Engineering and Digital Technologies, Teesside University, TS1 3BX, UK.

Institutions can provide incentives to enhance cooperation in a population where this behaviour is infrequent. This process is costly, and it is thus important to optimize the overall spending. This problem can be mathematically formulated as a multi-objective optimization problem where one wishes to minimize the cost of providing incentives while ensuring a minimum level of cooperation, sustained over time. Prior works that consider this question usually omit the stochastic effects that drive population dynamics. In this paper, we provide a rigorous analysis of this optimization problem, in a finite population and stochastic setting, studying both pairwise and multi-player cooperation dilemmas. We prove the regularity of the cost functions for providing incentives over time, characterize their asymptotic limits (infinite population size, weak selection and large selection) and show exactly when reward or punishment is more cost efficient. We show that these cost functions exhibit a phase transition phenomena when the intensity of selection varies. By determining the critical threshold of this phase transition, we provide exact calculations for the optimal cost of incentive, for any given intensity of selection. Numerical simulations are also provided to demonstrate analytical observations. Overall, our analysis provides for the first time a selection-dependent calculation of the optimal cost of institutional incentives (for both reward and punishment) that guarantees a minimum level of cooperation over time. It is of crucial importance for real-world applications of institutional incentives since intensity of selection is often found to be non-extreme and specific for a given population.

THE ROYAL SOCIETY PUBLISHING

1. Introduction

The problem of promoting the evolution of cooperative behaviour within populations of self-regarding individuals has been intensively investigated across diverse fields of behavioural, social and computational sciences (Han, 2013; Nowak, 2006b; Perc et al., 2017; Sigmund, 2010; West et al., 2007). Various mechanisms responsible for promoting the emergence and stability of cooperative behaviours among such individuals have been proposed. They include kin and group selection (Hamilton, 1964; Traulsen and Nowak, 2006), direct and indirect reciprocities (Han et al., 2012; Krellner and Han, 2020; Nowak and Sigmund, 2005; Ohtsuki and Iwasa, 2006; Okada, 2020), spatial networks (Antonioni and Cardillo, 2017; Peña et al., 2016; Perc et al., 2013; Santos et al., 2006), reward and punishment (Boyd et al., 2003, 2010; Fehr and Gachter, 2000; Hauert et al., 2007a; Herrmann et al., 2008; Sigmund et al., 2001), and pre-commitments (Han et al., 2013, 2016; Martinez-Vaquero et al., 2017; Nesse, 2001; Sasaki et al., 2015). Institutional incentives, namely, rewards for cooperation and punishment of wrongdoing, are among the most important ones (Chen et al., 2015; García and Traulsen, 2019; Góis et al., 2019; Han and Tran-Thanh, 2018; Powers et al., 2018; Sigmund et al., 2001, 2010a; Vasconcelos et al., 2013; Wang et al., 2019; Wu et al., 2014). Differently from other mechanisms, in order to carry out institutional incentives, it is assumed that there exists an *external* decision maker (e.g. institutions such as the United Nations and the European Union) that has a budget to interfere in the population to achieve a desirable outcome. Institutional enforcement mechanisms are crucial for enabling large-scale cooperation. Most modern societies implemented certain forms of institutions for governing and promoting collective behaviors, including cooperation, coordination, and technology innovation (Bardhan, 2005; Bowles, 2009; Bowles and Gintis, 2002; Han et al., 2021; Ostrom, 1990; Scotchmer, 2004).

Providing incentives is costly and it is therefore important to minimize the cost while ensuring a sustained level of cooperation over time (Chen et al., 2015; Han and Tran-Thanh, 2018; Ostrom, 1990). Despite its paramount importance, so far there have been only few works exploring this question. In particular, Wang et al. (2019) use optimal control theory to provide an analytical solution for cost optimization of institutional incentives assuming deterministic evolution and infinite population sizes (modeled using replicator dynamics). This work therefore does not take into account various stochastic effects of evolutionary dynamics such as mutation and non-deterministic behavioral update (Hofbauer and Sigmund, 1998; Sigmund, 2010; Traulsen et al., 2006). In a deterministic system consisting of cooperators and defectors, once the latter disappear (for instance through strong institutional punishment), there is no further change to the system and thus no further interference in it is required. When mutation is present, this behaviour can however reoccur and become abundant over time, requiring institutions to spend more budget on providing further incentives. Moreover, a key factor of behavioral update, the intensity of selection (Sigmund, 2010)—which determines how strongly an individual bases her decision to copy another individual's strategy on their fitness difference—might strongly impact an institutional incentives strategy and its cost efficiency. Its value is usually found to be specific for a given population (Domingos et al., 2020; Rand et al., 2013; Traulsen et al., 2010; Zisis et al., 2015) and thus should be taken into account when designing suitable cost-efficient incentives. For instance, when selection is weak such that behavioral update is close to a random process (i.e. an imitation decision is independent of how large the fitness difference is), providing incentives would make little difference to cause behavioral change, however strong it is. When selection is strong, incentives that ensure a minimum fitness advantage to cooperators would already ensure a positive behavioral change.

In a stochastic, finite population context, so far this problem has been investigated primarily based on agent-based and numerical simulations (Chen et al., 2015; Cimpeanu et al., 2019, 2021; Han and Tran-Thanh, 2018; Han et al., 2018; Sasaki et al., 2012). Results demonstrate several interesting phenomena, such as the significant influence of the intensity of selection on incentive strategies and optimal costs. However, there is no satisfactory rigorous analysis available at present that allows one to determine the optimal way of providing incentives. This is a challenging problem because of the large but finite population size and the complexity of stochastic processes governing the population dynamics.

In this paper, we provide exactly such a rigorous analysis. We study cooperation dilemmas in both pairwise (the Donation game) and multi-player (the Public Goods game) settings (Sigmund, 2010). They are among the most well studied models for studying the evolution of cooperative behaviour where individually defection is always preferred over cooperation while mutual cooperation is the preferred collective outcome for the population as a whole. Adopting a popular stochastic evolutionary game approach for analysing well-mixed finite populations (Imhof et al., 2005; Nowak, 2006a; Nowak et al., 2004), we derive the total expected costs of providing institutional reward or punishment, characterize their asymptotic limits (namely, for infinite population, weak selection and strong selection) and show the existence of a phase transition phenomena in the optimization problem when the intensity of selection varies. We calculate the critical threshold of phase transitions and study the minimization problem when the selection is under and above the critical value. We furthermore provide numerical simulations to demonstrate the analytical results.

The rest of the paper is organized as follows. In Section 2 we introduce the models and methods, deriving mathematical optimization problems that will be studied. The main results of the paper are presented in Section 3. In Section 4 we discuss possible extensions for future work. Finally, detailed computations, technical lemmas and proofs the main results are provided in the attached Supporting Information (SI).

2. Models and methods

(a) Cooperation dilemmas

We consider a well-mixed, finite population of N self-regarding individuals or players, who interact with each other using one of the following one-shot (i.e. non-repeated) cooperation dilemmas, the Donation Game (DG) or its multi-player version, the Public Goods Game (PGG). In these games, a player can either choose to cooperate (i.e. a cooperator, or C player) or to defect (i.e. a defector, or D player).

Let $\Pi_C(i)$ and $\Pi_D(i)$ be the average payoffs of a C player and a D player in a population with i C players and $N-i$ D players, respectively (see also Section 2.3 for more details). We show below that the difference $\delta = \Pi_C(i) - \Pi_D(i)$ does not depend on i. For cooperation dilemmas, it is always the case that $\delta < 0$.

Donation Game (DG)

The payoff matrix of the DG (for row player) is given as follows

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix}, \end{array}$$

where c and b represent the cost and benefit of cooperation, where $b > c$. DG is a special version of the Prisoner's Dilemma game (PD).

Denoting $\pi_{X,Y}$ the payoff of a strategist X when playing with strategist Y from the payoff matrix above, we obtain

$$\Pi_C(i) = \frac{(i-1)\pi_{C,C} + (N-i)\pi_{C,D}}{N-1} = \frac{(i-1)(b-c) + (N-i)(-c)}{N-1},$$

$$\Pi_D(i) = \frac{i\pi_{D,C} + (N-i-1)\pi_{D,D}}{N-1} = \frac{ib}{N-1}.$$

Thus,

$$\delta = \Pi_C(i) - \Pi_D(i) = -(c + \frac{b}{N-1}).$$

Public Goods Game (PGG)

In a PGG, players interact in a group of size n, where they decide to cooperate, contributing an amount $c > 0$ to a common pool, or to defect and contributes nothing to the pool. The total contribution in a group will be multiplied by a factor r, where $1 < r < n$ (for the PGG to be a social dilemma), which is then shared equally among all members of the group, regardless of their strategy.

We obtain (Hauert et al., 2007b)

$$\Pi_C(i) = \sum_{j=0}^{n-1} \frac{\binom{i-1}{j}\binom{N-i}{n-1-j}}{\binom{N-1}{n-1}} \left(\frac{(j+1)rc}{n} - c\right) = \frac{rc}{n}\left(1 + (i-1)\frac{n-1}{N-1}\right) - c,$$

$$\Pi_D(i) = \sum_{j=0}^{n-1} \frac{\binom{i}{j}\binom{N-1-i}{n-1-j}}{\binom{N-1}{n-1}} \frac{jrc}{n} = \frac{rc(n-1)}{n(N-1)} i.$$

Thus,

$$\delta = \Pi_C(i) - \Pi_D(i) = -c\left(1 - \frac{r(N-n)}{n(N-1)}\right).$$

(b) Cost of institutional reward and punishment

To reward a cooperator (respectively, punish a defector), the institution has to pay an amount θ/a (resp., θ/b) so that the cooperator's (defector's) payoff increases (decreases) by θ, where $a, b > 0$ are constants representing the efficiency ratios of providing the corresponding incentive. As we study reward and punishment separately, without losing generality, we set $a = b = 1$ (Chen et al., 2015; Sigmund et al., 2001). Thus, the key question here is: *what is the optimal value of the individual incentive cost θ that ensures a sufficient desired level of cooperation in the population (in the long run) while minimizing the total cost spent by the institution?*

Deriving the expected cost of providing institutional incentives

We adopt here the finite population dynamics with the Fermi strategy update rule (Traulsen et al., 2006), stating that a player A with fitness f_A adopts the strategy of another player B with fitness f_B with a probability given by, $P_{A,B} = \left(1 + e^{-\beta(f_B - f_A)}\right)^{-1}$, where β represents the intensity of selection (see details in Section (c)). We compute the expected number of times the population contains i C players, $1 \le i \le N-1$. For that, we consider an absorbing Markov chain of $(N+1)$ states, $\{S_0, ..., S_N\}$, where S_i represents a population with i C players. S_0 and S_N are absorbing states. Let $U = \{u_{ij}\}_{i,j=1}^{N-1}$ denote the transition matrix between the $N-1$ transient states, $\{S_1, ..., S_{N-1}\}$. The transition probabilities can be defined as follows, for $1 \le i \le N-1$:

$$\begin{aligned} u_{i,i\pm j} &= 0 \qquad \text{for all } j \ge 2, \\ u_{i,i\pm 1} &= \frac{N-i}{N}\frac{i}{N}\left(1 + e^{\mp\beta[\Pi_C(i) - \Pi_D(i) + \theta]}\right)^{-1}, \\ u_{i,i} &= 1 - u_{i,i+1} - u_{i,i-1}. \end{aligned} \tag{2.1}$$

The entries n_{ij} of the so-called fundamental matrix $\mathcal{N} = (n_{ij})_{i,j=1}^{N-1} = (I - U)^{-1}$ of the absorbing Markov chain gives the expected number of times the population is in the state S_j if it is started in the transient state S_i (Kemeny and Snell, 1976). As a mutant can randomly occur either at S_0 or S_N, the expected number of visits at state S_i is thus, $\frac{1}{2}(n_{1i} + n_{N-1,i})$.

The total cost per generation is

$$\theta_i = \begin{cases} i \times \theta & \text{in the case of institutional reward,} \\ (N-i) \times \theta & \text{in the case of institutional punishment.} \end{cases}$$

Hence, the expected total cost of interference for institutional reward and institutional punishment are respectively

$$E_r(\theta) = \frac{\theta}{2} \sum_{i=1}^{N-1} (n_{1i} + n_{N-1,i}) i \quad \text{and} \quad E_p(\theta) = \frac{\theta}{2} \sum_{i=1}^{N-1} (n_{1i} + n_{N-1,i})(N-i). \tag{2.2}$$

Cooperation frequency

Since the population consists of only two strategies, the fixation probabilities of a C (D) player in a homogeneous population of D (C) players when the interference scheme is carried out are, respectively,

$$\rho_{D,C} = \left(1 + \sum_{i=1}^{N-1} \prod_{k=1}^{i} \frac{1 + e^{\beta(\Pi_C(k) - \Pi_D(k) + \theta)}}{1 + e^{-\beta(\Pi_C(k) - \Pi_D(k) + \theta)}}\right)^{-1},$$

$$\rho_{C,D} = \left(1 + \sum_{i=1}^{N-1} \prod_{k=1}^{i} \frac{1 + e^{\beta(\Pi_D(k) - \Pi_C(k) - \theta)}}{1 + e^{-\beta(\Pi_D(k) - \Pi_C(k) - \theta)}}\right)^{-1}.$$

Computing the stationary distribution using these fixation probabilities, we obtain the frequency of cooperation (See Section 2.3)

$$\frac{\rho_{D,C}}{\rho_{D,C} + \rho_{C,D}}.$$

Hence, this frequency of cooperation can be maximized by maximizing

$$\max_{\theta} \left(\rho_{D,C}/\rho_{C,D}\right). \tag{2.3}$$

The fraction in Equation (2.3) can be simplified as follows (Nowak, 2006a)

$$\begin{aligned} \frac{\rho_{D,C}}{\rho_{C,D}} &= \prod_{k=1}^{N-1} \frac{T^-(k)}{T^+(k)} = \prod_{k=1}^{N-1} \frac{1 + e^{\beta[\Pi_C(k) - \Pi_D(k) + \theta]}}{1 + e^{-\beta[\Pi_C(k) - \Pi_D(k) + \theta]}} \\ &= e^{\beta \sum_{k=1}^{N-1} (\Pi_C(k) - \Pi_D(k) + \theta)} \\ &= e^{\beta(N-1)(\delta + \theta)}. \end{aligned} \tag{2.4}$$

In the above transformation, $T^-(k)$ and $T^+(k)$ are the probabilities to increase or decrease the number of C players (i.e. k) by one in each time step, respectively.

We consider non-neutral selection, i.e. $\beta > 0$ (under neutral selection, there is no need to use incentives). Assuming that we desire to obtain at least an $\omega \in [0, 1]$ fraction of cooperation, i.e. $\frac{\rho_{D,C}}{\rho_{D,C} + \rho_{C,D}} \geq \omega$, it follows from Equation (2.4) that

$$\theta \geq \theta_0(\omega) = \frac{1}{(N-1)\beta} \log\left(\frac{\omega}{1-\omega}\right) - \delta. \tag{2.5}$$

Therefore it is guaranteed that if $\theta \geq \theta_0(\omega)$, at least an ω fraction of cooperation can be expected. From this condition it implies that the lower bound of θ monotonically depends on β. Namely, when $\omega \geq 0.5$, it increases with β while decreases for $\omega < 0.5$.

Optimization problems

Bringing all ingredients together, we obtain the following cost-optimization problems of institutional incentives in stochastic finite populations

$$\min_{\theta \geq \theta_0(\omega)} E(\theta), \tag{2.6}$$

where E is either E_r or E_p, defined in (2.2), which respectively corresponds to institutional reward and punishment. We show in Supporting Information (SI) that $\theta \mapsto E(\theta)$ is a smooth function on $\mathbb{R}$.

(c) Methods: Evolutionary Dynamics in Finite Populations

We adopt in our analysis the Evolutionary Game Theory (EGT) methods for finite populations (Imhof et al., 2005; Nowak, 2006a; Nowak et al., 2004). Herein, individuals' payoff represents their *fitness* or social *success*, and evolutionary dynamics is shaped by social learning (Hofbauer and Sigmund, 1998; Sigmund, 2010), whereby the most successful players will tend to be imitated more often by the other players. Here, social learning is modeled using the pairwise comparison rule (Traulsen et al., 2006), that is, a player A with fitness f_A adopts the strategy of another player B with fitness f_B with probability given by the Fermi function,

$$P_{A,B} = \left(1 + e^{-\beta(f_B - f_A)}\right)^{-1},$$

where β conveniently describes the selection intensity ($\beta = 0$ represents neutral drift while $\beta \to \infty$ represents increasingly deterministic selection).

In the absence of mutations or exploration, the end states of evolution are inevitably monomorphic: once such a state is reached, it cannot be escaped through social learning. We assume that, with a certain mutation probability, an individual switches randomly to a different strategy without imitating another individual. In addition, we assume here the small mutation limit (Fudenberg and Imhof, 2005; Imhof et al., 2005; Nowak et al., 2004). Thus, at most two strategies are present in the population at a time. The evolutionary dynamics can be described by a Markov Chain, where each state represents a homogeneous population and the transition probabilities between any two states are given by the fixation probability of a single mutant (Fudenberg and Imhof, 2005; Imhof et al., 2005; Nowak et al., 2004). The resulting Markov Chain has a stationary distribution, which describes the average time the population spends in an end state. The small mutation limit allows us to obtain an analytical form of the frequency of cooperation (see below). It is noteworthy that although we focus here on the small mutation limit, this approach has been shown to be widely applicable to scenarios which go well beyond the strict limit of very small mutation rates (Domingos et al., 2020; Rand et al., 2013; Sigmund et al., 2010b; Zisis et al., 2015).

The fixation probability that a single mutant A taking over a whole population with $(N-1)$ B players is as follows (see e.g. references for details (Karlin and Taylor, 1975; Nowak et al., 2004; Traulsen et al., 2006))

$$\rho_{B,A} = \left(1 + \sum_{i=1}^{N-1} \prod_{j=1}^{i} \frac{T^-(j)}{T^+(j)}\right)^{-1},$$

where $T^{\pm}(k) = \frac{N-k}{N}\frac{k}{N}\left[1 + e^{\mp\beta[\Pi_A(k) - \Pi_B(k)]}\right]^{-1}$ describes the probability to change the number of A players by $\pm$ one in a time step. Specifically, when $\beta = 0$, $\rho_{B,A} = 1/N$, representing the transition probability at neural limit.

Considering the set of two strategies C and D (see (Fudenberg and Imhof, 2005; Imhof et al., 2005) for the calculation for any number of strategies). Their stationary distribution is given by the normalised eigenvector associated with the eigenvalue 1 of the transposed of a matrix (Fudenberg and Imhof, 2005; Imhof et al., 2005)

$$M = \begin{pmatrix} 1-\rho_{C,D} & \rho_{C,D} \\ \rho_{D,C} & 1-\rho_{D,C} \end{pmatrix},$$

which is $\{\frac{\rho_{D,C}}{\rho_{D,C}+\rho_{C,D}}, \frac{\rho_{C,D}}{\rho_{D,C}+\rho_{C,D}}\}$. The first term is the frequency of cooperation and the second one is that of defection.

3. Main results

The present paper provides a rigorous analysis for the expected total cost of providing institutional incentive (2.2) and the associated optimization problem (2.6). In this section, we state our main analytical results, Theorems 3.1, 3.2 and 3.3, and provide numerical simulations to illustrate the analytical results. The proofs of these results, which require a delicate analysis of the cost functions, are presented in SI.

In the following theorems, E denotes the cost function either for institutional reward, E_r, or institutional punishment, E_p, as obtained in (2.2). Also, H_N denotes the well-known harmonic number

$$H_N := \sum_{j=1}^{N-1} \frac{1}{j}. \tag{3.1}$$

Our first main result provides qualitative properties and asymptotic limits of E.

Theorem 3.1 (Qualitative properties and asymptotic limits of total cost functions)**.**

(I) *(finite population estimates) The expected total cost of providing incentive satisfies the following estimates for all finite populations of size N*

$$\frac{N^2\theta}{2}\Big(H_N+\frac{1}{N-1}\Big) \le E(\theta) \le N(N-1)\theta\Big(H_N+1\Big). \tag{3.2}$$

(II) *(infinite population limit) The expected total cost of providing incentive satisfies the following asymptotic behaviour when the population size N tends to $+\infty$*

$$\lim_{N\to+\infty} \frac{E(\theta)}{\frac{N^2\theta}{2}(\ln N+\gamma)} = \begin{cases} 1+e^{-\beta|\theta-c|} & \textit{for DG}, \\ 1+e^{-\beta|\theta-c|}e^{\beta c\frac{r}{n}} & \textit{for PGG}, \end{cases} \tag{3.3}$$

where $\gamma = 0.5772\ldots$ is the Euler-Mascheroni constant.

(III) *(weak selection limit) The expected total cost of providing incentive satisfies the following asymptotic limit when the selection strength β tends to 0*

$$\lim_{\beta\to 0} E(\theta) = N^2\theta H_N. \tag{3.4}$$

(IV) *(strong selection limit) The expected total cost of providing incentive satisfies the following asymptotic limit when the selection strength β tends to $+\infty$*

$$\lim_{\beta\to+\infty} E_r(\theta) = \begin{cases} \frac{N^2}{2}\theta\Big(\frac{1}{N-1}+H_N\Big) & \textit{for} \quad \theta < -\delta, \\ N^2\theta H_N \quad \textit{for} \quad \theta = -\delta, \\ \frac{N^2}{2}\theta\Big(1+H_N\Big) \quad \textit{for} \quad \theta > -\delta. \end{cases} \tag{3.5}$$

$$\lim_{\beta\to+\infty} E_p(\theta) = \begin{cases} \frac{N^2\theta}{2}\Big(1+H_N\Big) & \textit{for} \quad \theta < -\delta, \\ N^2\theta H_N \quad \textit{for} \quad \theta = -\delta, \\ \frac{N^2\theta}{2}\Big(H_N+\frac{1}{N-1}\Big) \quad \textit{for} \quad \theta > -\delta. \end{cases} \tag{3.6}$$

The lower and upper bounds obtained in part (I) of the theorem suggest the total expected cost function E for both reward and punishment behaves asymptotically in order of $(N^2 H_N)\times\theta$

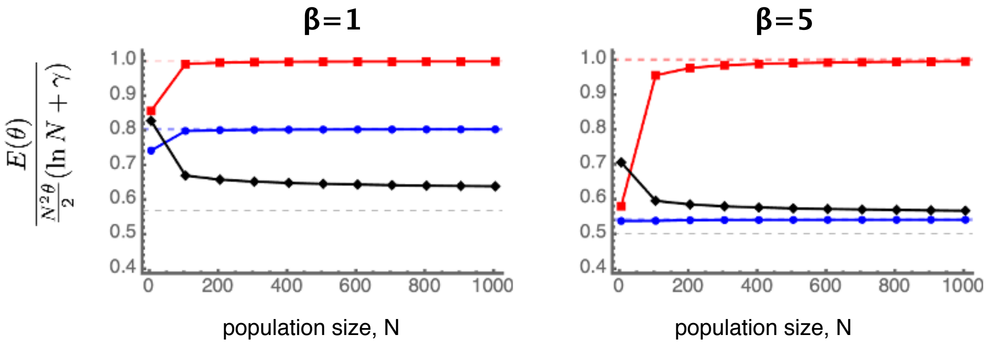

Figure 1. Large population size limit. We calculate numerically the expected total cost of incentive E for reward and punishment, varying population size N, for different values of θ and β. The dashed lines represent the corresponding theoretical limiting values obtained in Theorem 3.1 for the large population size limit, $N \to +\infty$. We observe that numerical results are in close accordance with those obtained theoretically. Results are obtained for DG with $b = 2$, $c = 1$.

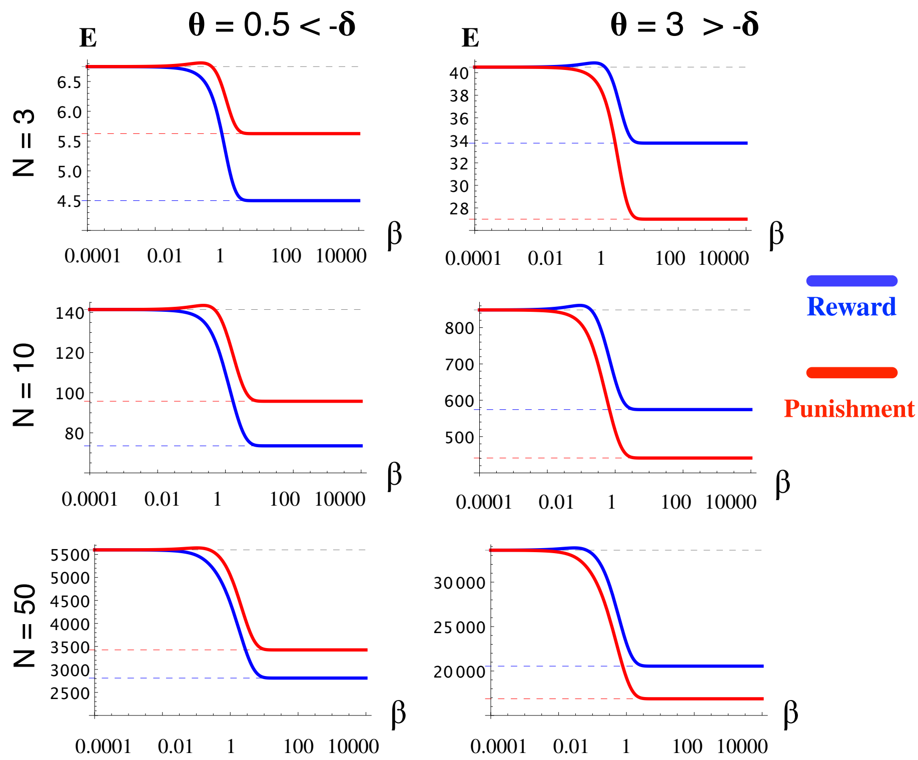

Figure 2. Weak and strong selection limits. We calculate numerically the total expected cost of incentive E for reward and punishment, for varying the intensity of selection, for different values of N and β. The dashed lines represent the corresponding theoretical limiting values obtained in Theorem 3.1 for weak and strong selection limits. We observe that numerical results are in close accordance with those obtained theoretically. Results are obtained for DG with $b = 2$, $c = 1$.

for sufficiently large N. It is confirmed in part (II), noting that $H_N \sim \ln N$. We also show that the leading asymptotic coefficient of E depends on the game (i.e., DG or PGG) and its parameters. Hence, it is important to adopt a precise optimal value of θ (e.g., obtained by solving the optimization problem (2.6)), as a small increase of this individual incentive cost can lead to significant increase in E, especially when the population size is large. Figure 1 numerically demonstrates this asymptotic limit.

Parts (III) and (IV) of the theorem provide theoretical estimations of E under the weak ($\beta \to 0$) and strong ($\beta \to +\infty$) selection limits. For the weak selection limit, the expected total costs are the same for reward and punishment, i.e. $E_r(\theta) = E_p(\theta)$. For the strong selection limit, E_r is smaller, equal or greater than E_p depending on whether θ is smaller, equal, or greater than $-\delta$. Figure 2 provides numerical validation of the theoretical weak and strong selection asymptotic behaviors of E, for different population sizes N. We can observe that, for a given individual incentive cost θ, the range of E increases significantly for larger N.

Our second main result concerns the optimization problem (2.6). We show that the cost function E exhibits a phase transition when the selection intensity β varies.

Theorem 3.2 (Optimization problems and phase transition phenomenon)**.**

(I) (phase transition phenomena and behaviour under the threshold) Define

$$F^* = \begin{cases} \min\{F(u) : P(u) > 0\} & \textit{in the reward case,} \\ \min\{\hat{F}(u) : \hat{P}(u) > 0\} & \textit{in the punishment case,} \end{cases}$$

where $P(u)$ and $F(u)$ as well as $\hat{P}$ and $\hat{F}$ are defined in the Supporting Information (See Section 1 and Section 2 there, respectively). There exists a threshold value β^* given by

$$\beta^* = -\frac{F^*}{\delta} > 0,$$

such that $\theta \mapsto E(\theta)$ is non-decreasing for all $\beta \le \beta^*$ and is non-monotonic when $\beta > \beta^*$. As a consequence, for $\beta \le \beta^*$

$$\min_{\theta \ge \theta_0} E(\theta) = E(\theta_0). \tag{3.7}$$

(II) (behaviour above the threshold value) For $\beta > \beta^*$, the number of changes of the sign of $E'(\theta)$ is at least two for all N and there exists an N_0 such that the number of changes is exactly two for $N \le N_0$. As a consequence, for $N \le N_0$, there exist $\theta_1 < \theta_2$ such that for $\beta > \beta^*$, $E(\theta)$ is increasing when $\theta < \theta_1$, decreasing when $\theta_1 < \theta < \theta_2$ and increasing when $\theta > \theta_2$. Thus, for $N \le N_0$,

$$\min_{\theta \ge \theta_0} E(\theta) = \min\{E(\theta_0), E(\theta_2)\}.$$

The proof of Theorems 3.1 and 3.2 for the case of reward and punishment are given in Section 1 and Section 2 in the SI, respectively. We also provide explicit computations for $N = 3$ and $N = 4$ to illustrate these theorems in Section 3 in the SI. Based on numerical simulations, we conjecture that the requirement that $N \le N_0$ could be removed and Theorem 3.2 is true for all finite N. In SI (Figure S2), using numerical calculation we have shown that $N_0 = 100$ satisfies the conjecture, ensuring the validity of the numerical examples below. Theorem 3.2 gives rise to the following algorithm to determine the optimal value θ^* for $N \le N_0$.

Algorithm 3.1 (Finding optimal cost of incentive $\theta^\star$).

***Inputs**: i) $N \le N_0$: population size, ii) β: intensity of selection, iii) game and parameters: PD (c and b) or PGG (c, r and n), iv) ω: minimum desired cooperation level.*

(1) Compute δ $\left\{\text{in PD: } \delta = -(c + \frac{b}{N-1}); \text{ in PGG: } \delta = -c\left(1 - \frac{r(N-n)}{n(N-1)}\right)\right\}$.

(2) Compute $\theta_0 = \frac{1}{(N-1)\beta} \log\left(\frac{\omega}{1-\omega}\right) - \delta$;

(3) Compute

$$F^* = \begin{cases} \min\{F(u) : P(u) > 0\} & \textit{in the reward case,} \\ \min\{\hat{F}(u) : \hat{P}(u) > 0\} & \textit{in the punishment case,} \end{cases}$$

where $P(u)$ and $F(u)$, as well as $\hat{P}$ and $\hat{F}$ are defined in the Supporting Information.

(4) Compute $\beta^* = -\frac{F^*}{\delta}$.

(5) If $\beta \le \beta^*$:

$$\theta^* = \theta_0, \quad \min E(\theta) = E(\theta_0).$$

(6) Otherwise (i.e. if $\beta > \beta^*$)

- *(a) Compute u_2 that is the largest root of the equation $F(u) + \beta\delta = 0$ for the reward case or that of $\hat{F}(u) + \beta\delta = 0$ for the punishment case.*
- *(b) Compute $\theta_2 = \frac{\log u_2}{\beta} - \delta$.*
  - *If $\theta_2 \le \theta_0$: $\theta^* = \theta_0, \quad \min E(\theta) = E(\theta_0)$;*
  - *Otherwise (if $\theta_2 > \theta_0$):*
    - *If $E(\theta_0) \le E(\theta_2)$: $\theta^* = \theta_0, \quad \min E(\theta) = E(\theta_0)$;*
    - *if $E(\theta_2) < E(\theta_0)$: $\theta^* = \theta_2, \quad \min E(\theta) = E(\theta_2)$.*

Output: *θ^* and $E(\theta^*)$.*

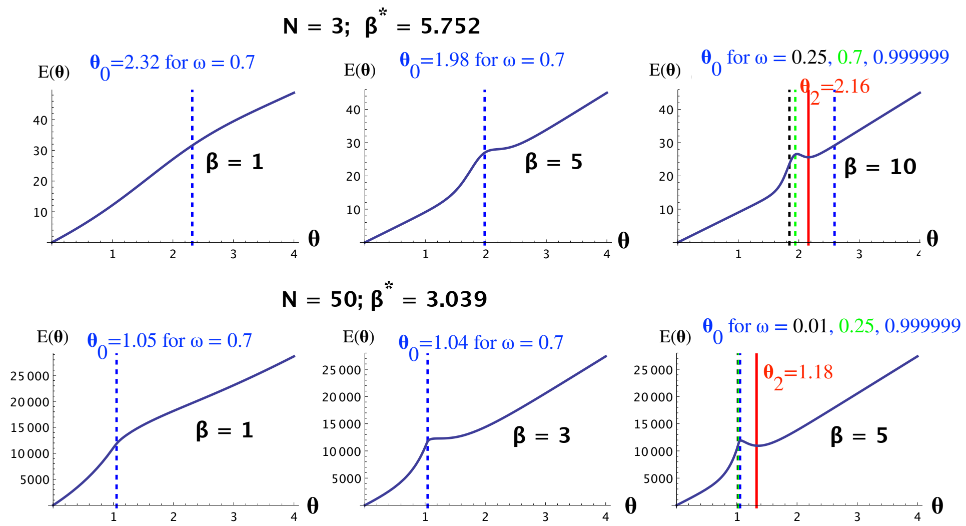

Figure 3. Using Algorithm 3.1 to find optimal θ that minimizes $E(\theta)$ (for institutional reward) while ensuring a minimum level of cooperation ω. We use as examples a small population size ($N = 3$, top row) and a larger one ($N = 50$, bottom row), for DG ($b = 1.8$, $c = 1$).

To illustrate Theorem 3.2 and Algorithm 3.1, we focus on the case of reward. Figure 3 shows the cost function E_r as a function of θ, for different values of N, β and ω for illustrating the phase transition when varying β, in a DG. We can see that in all cases, these numerical observations are in close accordance with theoretical results. For example, with $N = 3$ (see top row), we found $\beta^\star = f^\star/\delta = 10.9291/1.9 = 5.752$. For $\beta < \beta^\star$, $E(\theta)$ are increasing functions of θ. Thus, the optimal cost of incentive $\theta^\star = \theta_0$, for a given required minimum level of cooperation ω. For example, with $N = 3$, for $\beta = 1$ to ensure at least 70% of cooperation ($\omega = 0.7$), then $\theta^\star = \theta_0 = 2.32$. When $\beta \ge \beta^\star$ one needs to compare $E(\theta_0)$ and $E(\theta_2)$. For example, with $N = 3$, $\beta = 10$: for $\omega = 0.25$ (black dashed

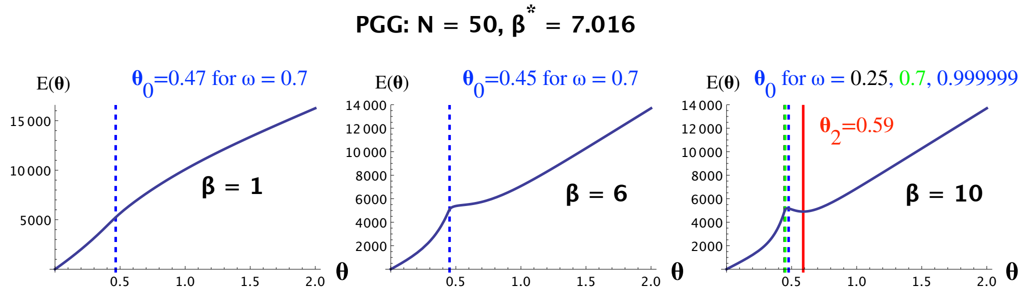

Figure 4. Using Algorithm 3.1 to find optimal θ that minimizes $E(\theta)$ while ensuring a minimum level of cooperation ω, for PGG ($r=3$, $n=5$, $c=1$) with $N=50$. Similar observations to those in DG, are obtained.

line), then $E(\theta_0)=23.602<25.6124=EC(\theta_2)$, so $\theta^\star=\theta_0=1.845$; for $\omega=0.7$ (green dashed line), then $E(\theta_0)=26.446>25.6124=EC(\theta_2)$, so $\theta^\star=\theta_2=2.16$ (red solid line); for $\omega=0.999999$ (blue dashed line), since $\theta_2<\theta_0$, $\theta^\star=\theta_0=2.59078$.

Similarly, with a larger population size ($N=50$, see Figure 1 in the SI, bottom row), we obtained $\beta^\star=3.15/1.03673=3.039$. In general, similar observations are obtained as in case of a small population size $N=3$. Except that when N is large, the values of θ_0 for different non-extreme values of minimum required cooperation ω (say, $\omega\in(0.01,0.99)$) is very small (given the log scale of $\omega/(1-\omega)$ in the formula of ω_0). This value is also smaller than θ_0, with a cost $E(\theta_0)>E(\theta_2)$, making θ_2 the optimal cost of incentive. Similar results are obtained for PGG (see Figure 4). When ω is extremely high (i.e. greater than $1-10^{-k}$, for a large k) (we don't look at extremely low value since we would like to ensure at least a sufficient level of cooperation), then we can also see other scenarios where the optimal cost is θ_0 (see Figure 1 in the SI, bottom row). We thus can observe that for $\omega\in(0.01,0.99)$, for sufficiently large population size N and large enough β ($\beta>\beta^\star+$ a bit more), then the optimal value of ω is always θ_2. Otherwise, θ_0 is the optimal cost.

Our last result provides a comparison of the expected total costs for providing institutional reward and punishment, for different individual incentive cost θ.

Theorem 3.3 (reward vs punishment costs)**.** *The difference between the expected total costs of reward and punishment is given by*

$$(E_r-E_p)(\theta)=\begin{cases}<0, & \textit{for} \quad \theta<-\delta,\\ =0, & \textit{for} \quad \theta=-\delta,\\ >0, & \textit{for} \quad \theta>-\delta.\end{cases} \tag{3.8}$$

As a consequence, when $\beta\le\min\{\beta_r^*,\beta_p^*\}$ *we have*

$$E_r^*=E_r(\theta_0),\quad E_p^*=E_p(\theta_0).$$

In this case,

$$(E_r^*-E_p^*)=E_r(\theta_0)-E_p(\theta_0)=\begin{cases}<0 & \textit{for} \quad \omega<0.5,\\ =0 & \textit{for} \quad \omega=0.5,\\ >0 & \textit{for} \quad \omega>0.5.\end{cases} \tag{3.9}$$

The proof of Theorem 3.4 is given in Section 3 in the SI. Numerical calculation in Figure 5 shows the expected total costs for reward and punishment (DG), for varying θ. We observe that reward is less costly than punishment ($E_r<E_p$) for $\theta<-\delta$ and vice versa when $\theta>-\delta$. It is exactly as shown analytically in Theorem 3.3. This analytical result is confirmed here for different population size N and intensity of selection β. Figure 6 also confirms the second part of

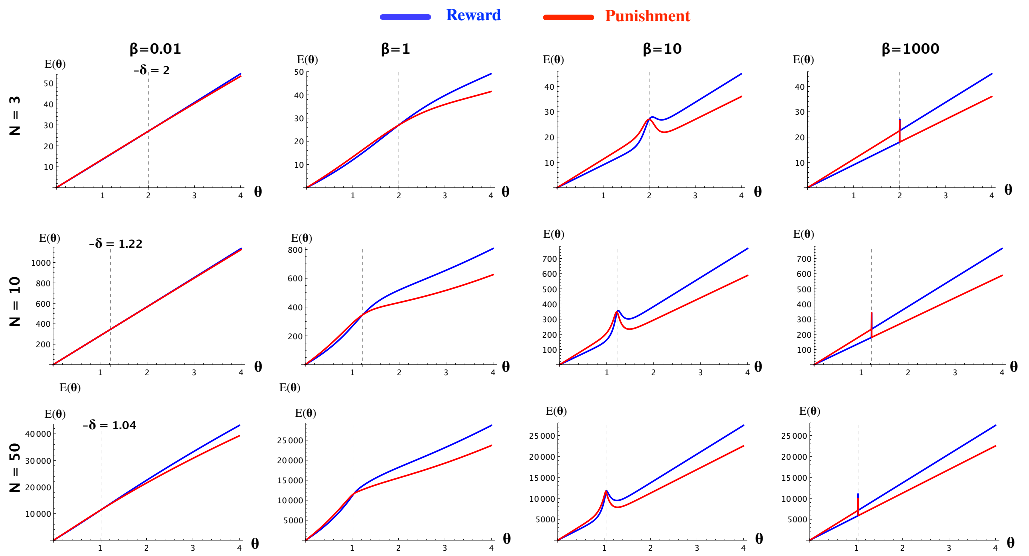

Figure 5. Compare the total costs E for reward and punishment as a function of θ, for different values of N and β. Reward is less costly than punishment ($E_r < E_p$) for small θ and vice versa otherwise. The threshold of θ for this change was obtained analytically (see Theorem 1), which is exactly equal to $-\delta$. Results are obtained for DG with $b = 2$, $c = 1$.

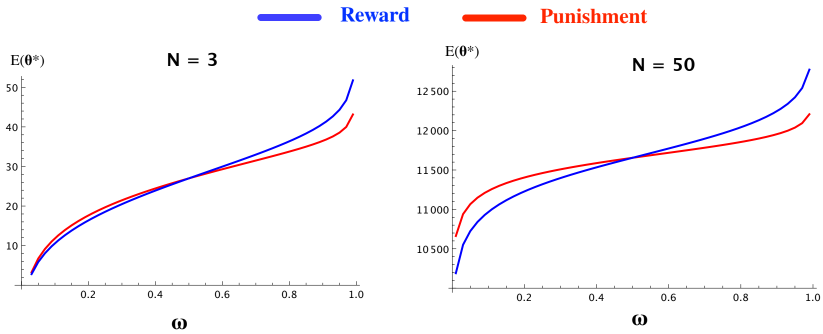

Figure 6. Compare the total costs E for reward and punishment at the optimal value $\theta^\star$ (obtained using Algorithm 3.1), for varying the minimum required level of cooperation, ω. Reward is more cost efficient for small ω, while punishment is more cost efficient when ω is larger. In both cases, the threshold is around $\omega = 0.5$. Other parameters: $\beta = 1$, DG with $b = 2$, $c = 1$.

the theorem, where for small β, if one can choose the type of incentive to use, either reward or punishment, then the former can provide a lower cost when requiring less than 50% cooperation at minimum and the later otherwise. This is in line with previous work showing that reward mechanisms work very well to promote cooperation in environments in which it is rare, while punishment mechanisms are better at maintaining high levels of cooperation (see e.g., (Chen et al., 2015; Sasaki et al., 2012; Wang et al., 2019)).

4. Discussion

Institutional incentives such as punishment and reward provide an effective tool for promoting the evolution of cooperation in social dilemmas. Both theoretical and experimental analysis has been made (Baldassarri and Grossman, 2011; Bardhan, 2005; Dong et al., 2019; García and Traulsen, 2019; Gürerk et al., 2006; Sasaki et al., 2012; Wu et al., 2014). However, past research usually ignores the question of how institutions' overall spending, i.e. the total cost of providing these incentives, can be minimized, while at the same time guaranteeing a minimum desired level of cooperation over time. Answering this question allows one to estimate exactly how incentives should be provided, that is how much to reward a cooperator and how severely to punish a wrongdoer. Existing works that consider this question usually omit the stochastic effects that drive population dynamics, namely, when the intensity of selection varies.

Resorting to a stochastic evolutionary game approach for finite, well-mixed populations, we have provided theoretical results for the optimal cost of incentives that ensure a desired level of cooperation while minimizing the total budget, for a given intensity of selection, β. We show that this cost strongly depends on the value of β, due to the existence of a phase transition in the cost functions when β varies. This behavior is missing in works that consider a deterministic evolutionary approach (Wang et al., 2019). The intensity of selection plays an important role in evolutionary processes. Its value differs depending on the payoff structure (i.e., scaling game payoff matrix by a factor is equivalent to dividing β by that factor) and is usually found to be specific for a given population, which can be estimated through behavioral experiments (Domingos et al., 2020; Rand et al., 2013; Traulsen et al., 2010; Zisis et al., 2015). Thus, our analysis provides a way to calculate the optimal incentive cost for a given population and game payoff matrix at hand.

As of theoretical importance, we characterized asymptotic behaviors of the total cost functions for both reward and punishment (namely, in the limits of large population, weak selection and strong selection) and compared these functions for the two types of incentive. We show that punishment is alway more costly for a small (individual) incentive cost (θ) but less so when this cost is above a certain threshold. We provided an exact formula for this threshold. This result provides insights into the choice of which type of incentives to use.

In the context of institutional incentives modelling, a crucial issue is the question of how to maintain the budget of incentives providing (Hilbe et al., 2014; Sigmund et al., 2010b). The problem of who pays or contributes to the budget is a social dilemma itself, and how to escape this dilemma is critical research question. In this work we focus on the question of how to optimize the budget used for provided incentives.

There are several simplifications made for the theoretical analysis to be possible. First, in order to derive analytical formula for the frequency of cooperation, we assumed the small mutation limit. Despite the simplified assumption, this small mutation limit approach has been shown to be widely applicable to scenarios which go well beyond the strict limit of very small mutation rates (Rand et al., 2013; Sigmund et al., 2010b; Zisis et al., 2015). Relaxing this assumption would make the derivation of a close form for the frequency of cooperation intractable.

Second, we focused in this paper on two important cooperation dilemmas, the DG and the PGG. They have in common a useful property that the difference in (average) payoff between a cooperator and a defector, $\delta = \Pi_C(i) - \Pi_D(i)$, does not depend on i, the number of cooperators in the population. This property allows us to simplify the fundamental matrix to a tridiagonal form and apply techniques matrix analysis to obtain a close form of its inverse matrix (see SI). In games with more complex payoff matrices such as the Prisoner's dilemma in its general form and the collective risk game (Santos and Pacheco, 2011), the difference δ depends on i and our technique in this paper can not be directly applied. We might consider other approaches to approximate the inverse matrix exploiting its block structure.

Data Accessibility. The article contains supporting information.

Authors' Contributions. MHD and TAH designed the researched, performed the research and wrote the paper. All authors gave final approval for publication and agree to be held accountable for the work performed therein

Competing Interests. We declare we have no competing interests.

Funding. T.A.H. was supported by Leverhulme Research Fellowship (RF-2020-603/9).

References

Alberto Antonioni and Alessio Cardillo. Coevolution of synchronization and cooperation in costly networked interactions. *Physical review letters*, 118(23):238301, 2017.

Delia Baldassarri and Guy Grossman. Centralized sanctioning and legitimate authority promote cooperation in humans. *Proceedings of the National Academy of Sciences*, 108(27):11023–11027, 2011.

Pranab Bardhan. Institutions matter, but which ones? *Economics of transition*, 13(3):499–532, 2005.

Samuel Bowles. *Microeconomics: behavior, institutions, and evolution*. Princeton University Press, 2009.

Samuel Bowles and Herbert Gintis. Social capital and community governance. *The economic journal*, 112(483):F419–F436, 2002.

Robert Boyd, Herbert Gintis, Samuel Bowles, and Peter J. Richerson. The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, 100(6):3531–3535, March 2003. doi: 10.1073/pnas.0630443100. URL http://dx.doi.org/10.1073/pnas.0630443100.

Robert Boyd, Herbert Gintis, and Samuel Bowles. Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science*, 328(5978):617–620, 2010.

Xiaojie Chen, Tatsuya Sasaki, Åke Brännström, and Ulf Dieckmann. First carrot, then stick: how the adaptive hybridization of incentives promotes cooperation. *Journal of The Royal Society Interface*, 12(102):20140935, 2015.

Theodor Cimpeanu, The Anh Han, and Francisco C Santos. Exogenous rewards for promoting cooperation in scale-free networks. In *The 2018 Conference on Artificial Life: A Hybrid of the European Conference on Artificial Life (ECAL) and the International Conference on the Synthesis and Simulation of Living Systems (ALIFE)*, pages 316–323. MIT Press, 2019.

Theodor Cimpeanu, Cedric Perret, and The Anh Han. Promoting fair proposers, fair responders or both? cost-efficient interference in the spatial ultimatum game. In *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems*, pages 1480–1482, 2021.

Elias Fernández Domingos, Jelena Grujić, Juan C Burguillo, Georg Kirchsteiger, Francisco C Santos, and Tom Lenaerts. Timing uncertainty in collective risk dilemmas encourages group reciprocation and polarization. *Iscience*, 23(12):101752, 2020.

Yali Dong, Tatsuya Sasaki, and Boyu Zhang. The competitive advantage of institutional reward. *Proceedings of the Royal Society B: Biological Sciences*, 286(1899):20190001, 2019.

Ernst Fehr and Simon Gachter. Cooperation and punishment in public goods experiments. *American Economic Review*, 90(4):980–994, 2000.

D. Fudenberg and L. A. Imhof. Imitation processes with small mutations. *Journal of Economic Theory*, 131:251–262, 2005.

Julián García and Arne Traulsen. Evolution of coordinated punishment to enforce cooperation from an unbiased strategy space. *Journal of the Royal Society Interface*, 16(156):20190127, 2019.

António R Góis, Fernando P Santos, Jorge M Pacheco, and Francisco C Santos. Reward and punishment in climate change dilemmas. *Scientific reports*, 9(1):1–9, 2019.

Özgür Gürerk, Bernd Irlenbusch, and Bettina Rockenbach. The competitive advantage of sanctioning institutions. *Science*, 312(5770):108–111, 2006.

W.D. Hamilton. The genetical evolution of social behaviour. i. *Journal of Theoretical Biology*, 7 (1):1 – 16, 1964. ISSN 0022-5193. doi: 10.1016/0022-5193(64)90038-4. URL http://www.sciencedirect.com/science/article/pii/0022519364900384.

T. A. Han, L. M. Pereira, and F. C. Santos. Corpus-based intention recognition in cooperation dilemmas. *Artificial Life*, 18(4):365–383, 2012.

T. A. Han, L. M. Pereira, F. C. Santos, and T. Lenaerts. Good agreements make good friends. *Scientific reports*, 3(2695), 2013.

The Anh Han. *Intention Recognition, Commitments and Their Roles in the Evolution of Cooperation: From Artificial Intelligence Techniques to Evolutionary Game Theory Models*, volume 9. Springer SAPERE series, 2013. ISBN 978-3-642-37511-8.

The Anh Han and Long Tran-Thanh. Cost-effective external interference for promoting the evolution of cooperation. *Scientific reports*, 8(1):1–9, 2018.

The Anh Han, Luís Moniz Pereira, and Tom Lenaerts. Evolution of commitment and level of participation in public goods games. *Autonomous Agents and Multi-Agent Systems*, pages 1–23, 2016. ISSN 1573-7454. doi: 10.1007/s10458-016-9338-4. URL http://dx.doi.org/10.1007/s10458-016-9338-4.

The Anh Han, Simon Lynch, Long Tran-Thanh, and Francisco C. Santos. Fostering cooperation in structured populations through local and global interference strategies. In *IJCAI-ECAI'2018*, pages 289–295, 2018.

The Anh Han, Luis Moniz Pereira, Tom Lenaerts, and Francisco C. Santos. Mediating Artificial Intelligence Developments through Negative and Positive Incentives. *PLOS ONE*, 16(1): e0244592, 2021. doi: 10.1371/journal.pone.0244592.

C. Hauert, A. Traulsen, H. Brandt, M. A. Nowak, and K. Sigmund. Via freedom to coercion: The emergence of costly punishment. *Science*, 316:1905–1907, 2007a.

Christoph Hauert, Arne Traulsen, Hannelore Brandt, Martin A Nowak, and Karl Sigmund. Via freedom to coercion: the emergence of costly punishment. *science*, 316(5833):1905–1907, 2007b.

Benedikt Herrmann, Christian Thöni, and Simon Gächter. Antisocial Punishment Across Societies. *Science*, 319(5868):1362–1367, March 2008.

Christian Hilbe, Arne Traulsen, Torsten Röhl, and Manfred Milinski. Democratic decisions establish stable authorities that overcome the paradox of second-order punishment. *PNAS*, 111(2):752–756, 2014.

J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998.

L. A. Imhof, D. Fudenberg, and Martin A. Nowak. Evolutionary cycles of cooperation and defection. *Proc. Natl. Acad. Sci. U.S.A.*, 102:10797–10800, 2005.

S. Karlin and H. E. Taylor. *A First Course in Stochastic Processes*. Academic Press, New York, 1975.

J. Kemeny and J. Snell. *Finite Markov Chains*. Undergraduate Texts in Mathematics. Springer, 1976.

Marcus Krellner and The Anh Han. Putting oneself in everybody's shoes-pleasing enables indirect reciprocity under private assessments. In *Artificial Life Conference Proceedings*, pages 402–410. MIT Press, 2020.

Luis A Martinez-Vaquero, The Anh Han, Luis Moniz Pereira, and Tom Lenaerts. When agreement-accepting free-riders are a necessary evil for the evolution of cooperation. *Scientific reports*, 7(1):1–9, 2017.

R. M. Nesse. *Evolution and the capacity for commitment*. Foundation series on trust. Russell Sage, 2001. ISBN 9780871546227.

M. A. Nowak. *Evolutionary Dynamics: Exploring the Equations of Life*. Harvard University Press, Cambridge, MA, 2006a.

M. A. Nowak and K. Sigmund. Evolution of indirect reciprocity. *Nature*, 437(1291-1298), 2005.

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

Martin A. Nowak. Five rules for the evolution of cooperation. *Science*, 314(5805):1560, 2006b.

Hisashi Ohtsuki and Yoh Iwasa. The leading eight: Social norms that can maintain cooperation by indirect reciprocity. *Journal of Theoretical Biology*, 239(4):435 – 444, 2006. ISSN 0022-5193. doi: DOI:10.1016/j.jtbi.2005.08.008. URL http://www.sciencedirect.com/science/article/B6WMD-4H4T62M-1/2/fda30e013d634e1c0b30a4559c2342bb.

Isamu Okada. A review of theoretical studies on indirect reciprocity. *Games*, 11(3):27, 2020.

Elinor Ostrom. *Governing the commons: The evolution of institutions for collective action*. Cambridge university press, 1990.

Jorge Peña, Bin Wu, Jordi Arranz, and Arne Traulsen. Evolutionary games of multiplayer cooperation on graphs. *PLoS computational biology*, 12(8):e1005059, 2016.

Matjaž Perc, Jesús Gómez-Gardeñes, Attila Szolnoki, Luis M Floría, and Yamir Moreno. Evolutionary dynamics of group interactions on structured populations: a review. *Journal of The Royal Society Interface*, 10(80):20120997, 2013.

Matjaž Perc, Jillian J Jordan, David G Rand, Zhen Wang, Stefano Boccaletti, and Attila Szolnoki. Statistical physics of human cooperation. *Phys Rep*, 687:1–51, 2017.

Simon T Powers, Anikó Ekárt, and Peter R Lewis. Modelling enduring institutions: The complementarity of evolutionary and agent-based approaches. *Cognitive Systems Research*, 52: 67–81, 2018.

David G. Rand, Corina E. Tarnita, Hisashi Ohtsuki, and Martin A. Nowak. Evolution of fairness in the one-shot anonymous ultimatum game. *Proc. Natl. Acad. Sci. USA*, 110:2581–2586, 2013.

F. C. Santos, J. M. Pacheco, and T. Lenaerts. Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proceedings of the National Academy of Sciences of the United States of America*, 103:3490–3494, 2006. ISSN 0027-8424.

Francisco C Santos and Jorge M Pacheco. Risk of collective failure provides an escape from the tragedy of the commons. *Proceedings of the National Academy of Sciences*, 108(26):10421–10425, 2011.

Tatsuya Sasaki, Åke Brännström, Ulf Dieckmann, and Karl Sigmund. The take-it-or-leave-it option allows small penalties to overcome social dilemmas. *Proceedings of the National Academy of Sciences*, 109(4):1165–1169, 2012.

Tatsuya Sasaki, Isamu Okada, Satoshi Uchida, and Xiaojie Chen. Commitment to cooperation and peer punishment: Its evolution. *Games*, 6(4):574–587, 2015.

Suzanne Scotchmer. *Innovation and incentives*. MIT press, 2004.

K Sigmund, C Hauert, and M Nowak. Reward and punishment. *P Natl Acad Sci USA*, 98(19): 10757–10762, 2001.

K. Sigmund, H. De Silva, A. Traulsen, and C. Hauert. Social learning promotes institutions for governing the commons. *Nature*, 466:7308, 2010a.

Karl Sigmund. *The Calculus of Selfishness*. Princeton University Press, 2010.

Karl Sigmund, Hannelore De Silva, Arne Traulsen, and Christoph Hauert. Social learning promotes institutions for governing the commons. *Nature*, 466(7308):861–863, Aug 2010b. ISSN 1476-4687. doi: 10.1038/nature09203. URL https://doi.org/10.1038/nature09203.

A. Traulsen and M. A. Nowak. Evolution of cooperation by multilevel selection. *Proceedings of the National Academy of Sciences of the United States of America*, 103(29):10952, 2006.

A. Traulsen, M. A. Nowak, and J. M. Pacheco. Stochastic dynamics of invasion and fixation. *Phys. Rev. E*, 74:11909, 2006.

Arne Traulsen, Dirk Semmann, Ralf D Sommerfeld, Hans-Jürgen Krambeck, and Manfred Milinski. Human strategy updating in evolutionary games. *Proceedings of the National Academy of Sciences*, 107(7):2962–2966, 2010.

Vitor V Vasconcelos, Francisco C Santos, and Jorge M Pacheco. A bottom-up institutional approach to cooperative governance of risky commons. *Nature Climate Change*, 3(9):797, 2013.

Shengxian Wang, Xiaojie Chen, and Attila Szolnoki. Exploring optimal institutional incentives for public cooperation. *Communications in Nonlinear Science and Numerical Simulation*, 79:104914, 2019.

S.A. West, A.A. Griffin, and A. Gardner. Evolutionary explanations for cooperation. *Current Biology*, 17:R661–R672, 2007.

Jia-Jia Wu, Cong Li, Bo-Yu Zhang, Ross Cressman, and Yi Tao. The role of institutional incentives and the exemplar in promoting cooperation. *Scientific reports*, 4:6421, 2014.

Ioannis Zisis, Sibilla Di Guida, The Anh Han, Georg Kirchsteiger, and Tom Lenaerts. Generosity motivated by acceptance - evolutionary analysis of an anticipation games. *Scientific reports*, 5 (18076), 2015.