Reward and Punishment in Climate Change Dilemmas

MSc Thesis
Extended Summary

António R. G. Góis*
* Departamento de Engenharia Informática and INESC-ID
IST-Taguspark, Universidade Técnica de Lisboa,
2744-016 Porto Salvo, Portugal
antoniogois@tecnico.ulisboa.pt

Abstract – **Climate Change agreements may be conveniently formulated in terms of Public Goods Dilemmas involving groups of individuals. In this framework, one may require a minimum number of cooperators in the group before effective collective action ensues. Furthermore, decision-making may be contingent on the risk associated with future losses. Here we investigate the impact of Reward and Punishment in this type of collective endeavors — coined as collective-risk dilemmas — by means of a dynamic, evolutionary approach. Our results are gratifying, given the a-priori limitations of sanctioning in collective-risk dilemmas: On the one hand, we show that positive incentives (rewards) are essential to foster cooperation, mostly when both the perception of risk is low and the overall number of engaged parties is small. On the other hand, we find that negative incentives (sanctions) act to maintain cooperation, after it has been installed. Finally, we identify the conditions in which cooperation benefits from synergistically combining rewards and sanctions into a single policy.**

Index Terms – Climate Change, Cooperation, Evolutionary Game Theory, Punishment, Reward.

INTRODUCTION

Climate change stands as one of our biggest challenges in what concerns the emergence and promotion of cooperation [1, 2]. Indeed, world citizens build up high expectations every time a new International Environmental Summit is settled, unfortunately with few palpable solutions coming out. This calls for the development of more effective levels of discussion, agreements and coordination. The problem can be conveniently framed resorting to the mathematics of game theory, being a paradigmatic example of a public goods game [3]: a global good from which every single person profits, whether she contributes or not to maintain it. In this case, parties may free-ride on the efforts of others, avoiding any effort themselves, while driving the population into the tragedy of the commons.

One of the multiple shortcomings identified to such agreements is a deficit in perceiving the actual risk of future losses, with profound effects on the expected dynamics of cooperation [4, 5]. Another problem relates to the lack of institutional sanctions (punishment) upon those who do not contribute to the welfare of the planet, and positive incentives (reward) to those who subscribe to green policies [6]. In the larger context of cooperation studies, positive incentives (rewards), negative incentives (punishment) and the combination of both [7-16] have been shown to have an impact, depending on the dilemma in place. Up to now, however, assessing the impact and efficiency of reward and punishment (isolated or combined) in the context of climate change dilemmas remains an open problem.

Here we theoretically address the role of both institutional reward and punishment in the context of climate change agreements, by describing the problem in terms of a N-player collective-risk dilemma (CRD) [4-6, 17-19]. We consider a population of size Z, where each individual can be either a Cooperator (***C***) or a defector (***D***). Everyone starts with the (same) initial endowment B. A cooperator accepts to incur a cost corresponding to a fraction c of his initial endowment B, in order to help prevent a collective failure. On the other hand, a defector chooses not to incur any cost, hoping to free-ride on the contributions of others. We require that a minimum number of individuals $M \leq N$ in a group of size N actually cooperates before collective action is realized: If a group of size N does not contain at least M ***C***s, all members lose their remaining endowments with a probability r, where r ($0 \leq r \leq 1$) stands as the risk of collective failure. Otherwise, everyone will keep whatever he or she has. Such a formulation of a collective-risk dilemma has been shown to capture some of the key features discovered in recent experiments [4, 17, 20-22], while highlighting the importance of risk. In addition, it allows one to explore in detail the consequences of varying in arbitrary ways the model parameters. Moreover, the adoption

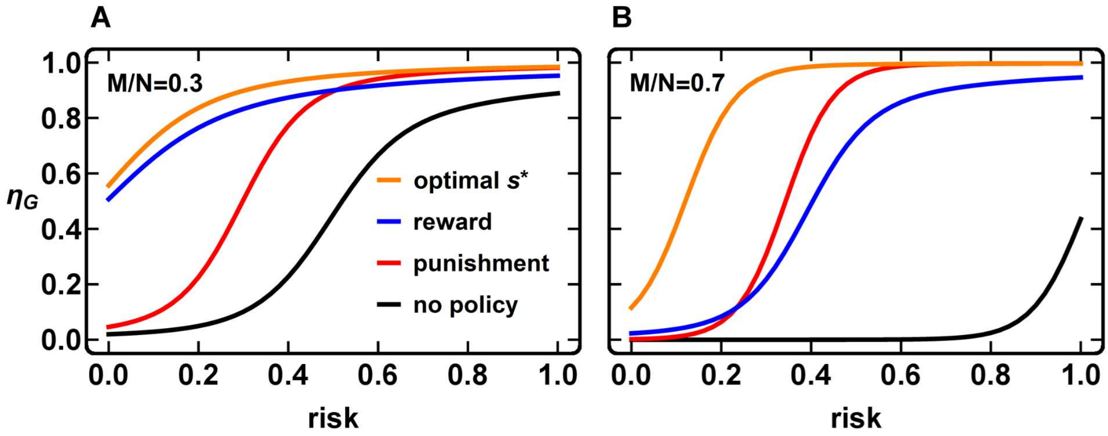

Figure 1. Average group achievement η_G as a function of risk. **Left**: Group relative threshold M/N=3/10. **Right**: Group relative threshold M/N=7/10. In both panels, the black line corresponds to a reference scenario where no policy is applied. The red line shows η_G in the case where all available budget is applied to pure-Punishment (*w*=0), whereas the blue line shows results for pure-Reward (*w*=1). The orange line shows results using the optimal combined policy s* following [8] (detailed in **Methods**), leading (naturally) to the best results. Pure-Reward is most effective at low risk values, while pure-Punishment is marginally most effective at high risk. These features are more pronounced for low relative thresholds (left panel), and only at high thresholds does pure-Punishment lead to sizeable improvement with respect to pure-Reward. Other parameters: Population size Z=50, group size N=10, cost of cooperation c=0.1, initial endowment B=1, budget δ=0.025, reward efficiency a=1, punishment efficiency b=1, intensity of selection=5, mutation rate μ=0.01.

of non-linear returns mimics situations common to many human and non-human endeavors [23-31], where a minimum collective effort is required to achieve a collective goal. Thus the applicability of this framework extends well beyond environmental governance, given the ubiquity of such type of social dilemmas in nature and societies.

As Sigmund points out [13], previous work depicting punishment as more efficient than reward considers the public good as a linear function of the number of contributors. We depart from this linear regime by modelling the public good as a threshold problem, combined with an uncertain outcome, represented by a risk of failure. As a result, our model encompasses a richer structure, exhibiting new internal equilibria [5]. As discussed below, these will interact in a non-trivial way with positive and negative incentives.

Following [8], we include both reward and punishment mechanisms in this model. A fixed group budget $N\delta$ (where $\delta \geq 0$ stands for a per-capita incentive) is assumed to be available, of which a fraction w is applied to a reward policy and the remaining $1-w$ to a punishment policy. We assume the effective impact of both policies to be equivalent, meaning that each unit spent will directly increase/decrease the payoff of a cooperator/defector by the same amount. For details on policies with different efficiencies, see **Methods**.

Instead of a collection of rational agents engaging in one-shot Public Good Games [21, 32], we adopt an evolutionary description of the behavioral dynamics [5], in which individuals tend to copy others whenever these appear to be more successful. Success (or fitness) of an individual is here associated with his average payoff. All individuals are equally likely to interact with each other, causing all cooperators and defectors to be equivalent, on average, and only distinguishable by the strategy they adopt. Therefore, and considering only two strategies are available, the number of cooperators is sufficient to describe any composition of the population. The number of individuals adopting a given strategy (either C or D) evolves in time according to a stochastic birth–death process [33, 34], which describes the time evolution of the social learning dynamics (with exploration): At each time-step each individual (X, with fitness f_X) is given the opportunity to change strategy; with probability μ, X randomly explores the strategy space [35] (a process similar to mutations in a biological context that precludes the existence of absorbing states. With probability (1-μ), X may adopt the strategy of a randomly selected member of the population (Y, with fitness f_Y), with a probability that increases with the fitness difference ($f_Y - f_X$) [34]. This renders the stationary distribution (see **Methods**) an extremely useful tool to rank the most visited states given the ensuing evolutionary dynamics of the population. Indeed, the stationary distribution provides the prevalence of each of the population's possible compositions, in terms of the number of ***C***s (k) and ***D***s ($Z-k$). Combined with the probability of success of each composition, the stationary distribution can be used to compute the overall success probability of a given population – the average group achievement, η_G. This value represents the average fraction of groups that will overcome the Collective Risk Dilemma, successfully preserving the public good.

RESULTS

In **Figure 1** we compare the average group achievement η_G (as a function of risk), for a reference scenario without any policy (i.e., no reward or punishment, in black), and for three other scenarios where a budget is provided and applied to a particular policy. Naturally η_G improves whenever a policy is applied. Less obvious is the difference between the various policies. Applying only rewards (blue curve) is more effective than only punishment (red curve) for low values of risk. The opposite happens when risk is high. On scenarios with a low relative threshold (left panel in **Figure 1**), rewards play the key role, with sanctions outperforming them only for very high values of risk. For high coordination thresholds (right panel, Figure 1) reward and punishment portray comparable efficiencies in the promotion of cooperation, with pure-Punishment (w=0) performing slightly better than pure-Reward (w=1).

The origins of these differences are difficult to grasp solely from the analysis of η_G. Thus, to better understand the behavior of pure-Reward and pure-Punishment, we show in **Figure 2** the gradients of selection (top panels) and stationary distributions (lower panels) for different budget values. Each gradient of selection represents, for each discrete state k/Z, the difference $G(k) = T^{+}(k) - T^{-}(k)$ among the probability to increase ($T^{+}(k)$) and decrease ($T^{-}(k)$) the number of cooperators (see **Methods**) by one. Whenever $G(k) > 0$ the fraction of ***C***s is likely to increase (in a stochastic sense); whenever $G(k) < 0$ the opposite is expected to happen. The stationary distributions show how likely it is to find the population in each (discrete) configuration of our system. The panels on the left-hand side show the results obtained for the CRD under pure-Reward; on the right-hand side, we show the results obtained for pure-Punishment.

Naturally, both mechanisms are inoperative whenever the per-capita incentives are inexistent (δ=0), creating a natural reference scenario in which to study the impact of Reward and Punishment on the CRD. In this case, above a certain value of risk (r), decision-making is characterized by two internal equilibria (i.e., adjacent finite population states with opposite gradient sign, representing the analogue of fixed points in a dynamical system characterizing evolution in an infinite population). Above a certain fraction of cooperators the population overcomes the coordination barrier and naturally self-organizes towards a stable co-existence of cooperators and defectors. Otherwise, the population is condemned to evolve towards a monomorphic population of defectors, leading to the tragedy of the commons [5]. As the budget for incentives increases, using either Reward or Punishment leads to very different outcomes, as depicted in **Figure 2**.

Reward is particularly effective when cooperation is low (small k/Z), showing a particular impact on the location of the finite population analogue of an unstable fixed point. Indeed, with increasing δ the minimum number of cooperators required to reach the cooperative basin of attraction and the co-existence point becomes lower, ultimately disappearing as δ continues to increase (**Figures 2A** and **2C**). This means that a smaller coordination effort is required before the population dynamics starts to naturally favour the increase of cooperators. Once this initial barrier is surpassed, the population will naturally tend towards an equilibrium state which does not improve appreciably under Reward. The opposite happens under Punishment – whereas the location of the coordination point is little affected, once this barrier is overcome the population will evolve towards a more favorable equilibrium. Thus, while Reward seems to be particularly effective to bootstrap cooperation, Punishment seems effective in sustaining high levels of cooperation.

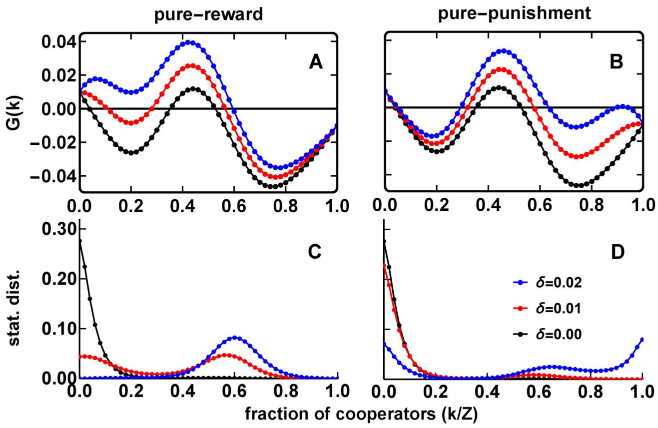

Figure 2. Gradient of selection (top panels, A and B) and stationary distribution (bottom panels, C and D) for the different values of per-capita budget δ indicated, using either pure-Reward (w=1, left panels) or pure-Punishment (w=0, right panels). The black curve is equal on the left and right panels, since in this case δ=0. As δ increases, the behaviour under Reward and Punishment is qualitatively similar, by displacing the (unstable) coordination equilibrium towards lower values of k/Z, while displacing the (stable) coexistence equilibrium towards higher values of k/Z. This happens, however, only for low values of δ. Indeed, by further increasing δ one observes very different behaviours under Reward and Punishment: Whereas under Punishment the equilibria are further moved apart (in accord with what happened for low δ) under Reward the coordination equilibrium disappears, and the overall dynamics becomes characterized by a single coexistence equilibrium which consistently shifts towards higher values of k/Z with increasing δ. This difference in behaviour, in turn, has a dramatic impact in the overall prevalence of configurations achieved by the population dynamics, as shown by the stationary distributions: On the left panel (pure-Reward) the population spends most of the time on intermediate states of cooperation. On the right panel (pure-Punishment) the population spends most of the time on both extremes (high and low cooperation) but especially on low cooperation states. Other parameters: Population size Z=50, group size N=10, threshold M=5, cost of cooperation c=0.1, initial endowment B=1, risk r=50%, reward (a) and punishment (b) efficiency, a=b=1 (see Methods), intensity of selection β=5, mutation rate μ=0.01.

As a consequence, the most frequently observed population compositions are very different when using each of the policies. As shown by the stationary distributions (**Figures 2C** and **2D**), under Reward the population visits more often states with intermediate values of cooperation. Intuitively this happens because the coordination effort is eased by the rewards, causing the population to effectively overcome it and reach the coexistence point (the equilibrium state with an intermediate amount of cooperators) thus spending most of the time near it. On the other hand, Punishment will not ease the coordination effort, and thus the population will spend most of the time in states of low cooperation, failing to overcome this barrier. Notwithstanding, once surpassed, the population will stabilize on higher states of cooperation. This is especially true for high budgets, as shown with δ=0.02 (blue line). Moreover, since $N{\cdot}\delta$ corresponds to a fixed total amount which is distributed by the existing cooperators/defectors, this causes the per-cooperator/defector budget to vary depending on the number of existing cooperators/defectors (i.e., each of the j cooperators receives $w\delta N/j$ and each defector $(1-w)\delta N/(N-j)$). In other words, positive (negative) incentives become very profitable (or severe) if defection (cooperation) prevails within a group. In particular, whenever the budget is significant (see, e.g., δ=0.02 in **Figure 2**) the punishment becomes so high when there are few defectors within a group, that a new equilibrium emerges close to full cooperation.

The results in **Figure 2** show that Reward can be instrumental in fostering pro-social behavior, while Punishment can be used for its maintenance. This suggests that, to combine both policies synergistically, pure-Reward (w=1) should be applied at first when there are few cooperators (low k/Z), up to a certain (critical) point (k/Z=s), above which one should switch to pure-Punishment (w=0). In the **Methods** section, we demonstrate that, similar to Public Goods Games [8], in collective risk dilemmas this is indeed the policy which minimizes the advantage of the defector, even if we consider the alternative possibility of applying both policies simultaneously. In **Methods**, we also compute a general expression for the optimal switching point s^*. By using such policy that we denote by s^*, we obtain the best results shown with an orange line in **Figure 1**. We propose, however, to explore what happens in the context of a CRD when s^* is not used. How much cooperation is lost when we deviate from s^* to either of the pure policies, or to a policy which uses a switching point different from the optimal one? **Figure 3** illustrates how the choice of the switching point s impacts the overall cooperation, as evaluated by η_G, for different values of risk. For a switching point of s=k/Z=1.0 (0.0) a static policy of always pure-Reward (pure-Punishment) is used. This can be seen on the far right (left) of **Figure 3**. **Figure 3** suggests that, for low thresholds, an optimal policy switching (which, for the parameters shown, occurs for s=50%, see **Methods**) is only marginally better than a policy solely based on rewards (s=1). **Figure 3** also allows for a comparison of what happens when the switching point occurs too late (excessive rewards) or too early (excessive sanctions) in a low-threshold scenario. A late switch is significantly less harmful than an early one. In other words, our results suggest that when the population configuration cannot be precisely observed, it is preferable to keep rewarding for longer. This said, whenever the perception of risk is high (an unlikely situation these days) an early switch is slightly less harmful than a late one. In the most difficult scenarios, where stringent coordination requirements (large M) are combined with a low perception of risk, the adoption of a combined policy becomes necessary (see right panel of **Figure 1**).

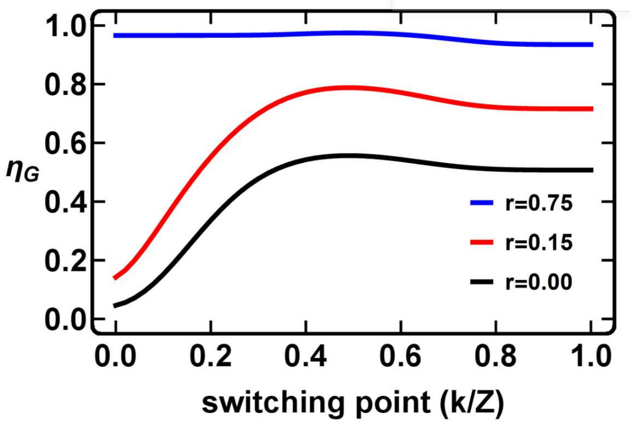

Figure 3. Average group achievement η_G as a function of the switching point. The switching point corresponds to the configuration of the population above which w suddenly switches from 1 to 0, changing from pure-Reward to pure-Punishment. This configuration is defined by the fraction of cooperators in the population (k/Z), which can vary from 0 to 1, as seen on the x-axis. Assuming both policies are equally efficient, the optimal switching point occurs at 50% of cooperators. The far-left values of η_G correspond to a static policy of always pure-Punishment – the switch from pure-Reward to pure-Punishment occurs immediately at 0% of cooperators. On the far-right (switching point = 100%) a pure-Reward policy is depicted. We can also see what happens when the switch occurs too late or too early, for different values of risk. For low values of risk, it is significantly less harmful to have a late switch from Reward to Punishment than an early one, meaning that when the population configuration cannot be precisely observed, it is preferable to keep rewarding for longer. Other parameters: Population size Z=50, initial endowment B=1, mutation rate μ=0.01, intensity of selection=5, group size N=10, threshold M=3, cost of cooperation c=0.1, budget δ=0.025.

DISCUSSION

One might expect the impact of Reward and Punishment to lead to symmetric outcomes – Punishment would be effective for high-cooperation the same way that Reward is effective for low-cooperation. In low-cooperation scenarios (under low risk, threshold or budget) Reward alone plays the most important role. However, in the opposite scenario, Punishment alone does not have the same impact. Either a favourable scenario occurs, where any policy yields a satisfying result, or Punishment cannot improve outcomes on its own. In the latter case, the synergy between both policies becomes essential to achieve cooperation. Such optimal policy involves a combination of the single policies, Reward and Punishment, which is dynamic, in the sense that the combination does not remain the same for all configurations of the population. It corresponds to employing pure Reward at first, when cooperation is low, switching subsequently to

Punishment whenever a pre-determined level of cooperation is reached.

The optimal procedure, however, is unlikely to be realistic in the context of Climate Change agreements. Indeed, and unlike other Public Goods Dilemmas, where Reward and Punishment constitute the main policies available for Institutions to foster cooperative collective action, in International Agreements it is widely recognized that Punishment is very difficult to implement [2, 32]. This has been, in fact, one of the main criticisms put forward in connection with Global Agreements on Climate Mitigation: On one hand, they lack sanctioning mechanisms; but on the other hand, the reason for that is that it is practically impossible to enforce any type of sanctioning at a Global level. In this sense, the results obtained here by means of our dynamical, evolutionary approach, are gratifying, given these a-priori limitations of sanctioning in collective-risk dilemmas. Not only do we show that Reward is essential to foster cooperation, mostly when both the perception of risk is low and the overall number of engaged cooperators is small (low k/Z), but also we show that Punishment mostly acts to sustain cooperation, after it has been installed. Given that low-risk scenarios are more common and harmful to cooperation than high-risk ones, our results in connection with Reward provide a viable way to explore in the quest for establishing Global cooperative collective action. Finally, Reward policies may also be very relevant in scenarios where one couples Climate Agreements with other International agreements from which parties are not interested to deviate from [2, 32].

The model used takes for granted the existence of an institution with a budget available to implement either Reward or Punishment. New behaviours may emerge once individuals are called to decide whether or not to contribute to such an institution, allowing for a scenario where this institution fails to exist [6, 18]. At present, and under the Paris agreement, we are witnessing the potential birth of an informal funding institution, whose goal is to finance developing countries to help them increase their mitigation capacity. Clearly, this is just an example pointing out to the fact that the prevalence of local and global institutional incentives may depend and may be influenced by the distribution of wealth available among parties, in the same way that it influences the actual contributions to the public good [19, 22]. Finally, several other effects may further influence and/or affect the present results. Among others, if intermediate tasks are considered [22], or if individuals have the opportunity to pledge their contribution before their actual action [30, 36, 37], it is likely that pro-social behavior may be enhanced. Work along these lines is in progress.

METHODS

Public goods and collective risks. Let us consider a population with Z individuals, where each individual can be a cooperator (***C***) or a defector (***D***). For each round of this game, a group of N players is sampled from the original finite population of size Z, which corresponds to a process of sampling without replacement. The probability of a group comprising any possible combination of ***C***s and ***D***s is given by the hypergeometric distribution. In the context of a given group, a strategy is associated with a payoff value corresponding to an individual's earnings in that round, which depend on the action of the rest of group. Fitness is the expected payoff of an individual in a population, before knowing to which group he was assigned. This way, for a population with k out of Z ***C***s and each group containing j out of N ***C***s, the fitness of a ***D*** and a ***C*** can be written as:

$$f_D = \binom{Z-1}{N-1}^{-1} \sum_{j=0}^{N-1} \binom{k}{j} \binom{Z-k-1}{N-j-1} \Pi_D(j) \quad (1)$$

$$f_C = \binom{Z-1}{N-1}^{-1} \sum_{j=0}^{N-1} \binom{k-1}{j} \binom{Z-k}{N-j-1} \Pi_C(j+1) \quad (2)$$

where $\Pi_C(j)$ and $\Pi_D(j)$ stand for the payoff or a ***C*** and a ***D*** in a single round, in a group with N players and j ***C***s. To define the payoff functions, let $\theta(x)$ be a Heaviside step-function distribution, where $\theta(x) = 0$ if $x < 0$ and $\theta(x) = 1$ if $x \geq 0$. Each player can contribute with a fraction c of his endowment B, (with $0 \leq c \leq 1$), and in case a group contains less than M cooperators (with $M \leq N$) there is a risk r of failure ($0 \leq r \leq 1$), in which case no player obtains his remaining endowment. The payoff of a defector ($\Pi_D(j)$) and the payoff of a cooperator ($\Pi_C(j)$), before incorporating any policy, can be written as:

$$\Pi_D(j) = B\{\theta(j-M) + (1-r)[1-\theta(j-M)]\} \quad (3)$$

$$\Pi_C(j) = \Pi_D(j) - cB \quad (4)$$

Reward and Punishment. To include a Reward or a Punishment policy, let us follow [8] and consider a group budget $N \cdot \delta$ which can be used to implement any type of policy. The fraction of $N \cdot \delta$ applied to Reward is represented by the weight w, with $0 \leq w \leq 1$. Parameters a and b correspond to the efficiency of Reward and Punishment (for all Figures above it was assumed that $a = b = 1$).

$$\Pi_D^P(j) = \Pi_D(j) - \frac{b(1-w)N\delta}{N-j} \quad (5)$$

$$\Pi_C^R(j) = \Pi_C(j) + \frac{awN\delta}{j} \quad (6)$$

Naturally, these new payoff functions can be included into the previous fitness functions (Π_D^P replaces Π_D and Π_C^R replaces Π_C), letting fitness values account for the different policies.

Evolutionary dynamics in finite populations. The fitness functions written above allow us to setup the (discrete time) evolutionary dynamics. Indeed, the configurations of the entire population may be used to define a Markov Chain, where each state is characterized by the amount of cooperators. To decide in which direction the system will evolve, at each step a player i and a neighbour j of his are drawn at random from the population. Player i decides whether to imitate his neighbour j with a probability depending on the difference between their fitness. This way, a system with k cooperators may stay in the same state, switch to $k-1$ or to $k+1$. The probability of player i imitating player j can be given by the Fermi function:

$$p \equiv \left[1 + e^{-\beta(f_j - f_i)}\right]^{-1} \quad (7)$$

where β is the intensity of selection. Using this update rule it is possible to write down the one-step transition probabilities between adjacent configurations (i.e., that only vary by one cooperator). Let k be the total number of cooperators in the population and Z the total size of the population. $T^{+}(k)$ and $T^{-}(k)$ are the probabilities to increase and decrease k by one, respectively:

$$T^{\pm}(k) = \frac{k}{Z}\frac{Z-k}{Z}\left[1 + e^{\pm\beta[f_C(k) - f_D(k)]}\right]^{-1} \quad (8)$$

The most likely direction can be computed using the difference:

$$G(k) \equiv T^{+}(k) - T^{-}(k) \quad (9)$$

A mutation rate can be introduced by using the following transition probabilities:

$$T_{\mu}^{+}(k) = (1-\mu)T^{+}(k) + \mu\frac{Z-k}{Z} \quad (10)$$

$$T_{\mu}^{-}(k) = (1-\mu)T^{-}(k) + \mu\frac{k}{Z} \quad (11)$$

In all cases we used a mutation rate $\mu = 0.01$, this way avoiding the population to fixate in a monomorphic configuration. In this context, the stationary distribution becomes a very useful tool to analyse the overall population dynamics, providing the probability $\bar{p}_k = P(\frac{k}{Z})$ for each of the Z+1 states of this Markov Chain to be occupied. For each given population state k, the hypergeometric distribution can be used to compute the average fraction of groups that obtain success – $a_G(k)$. Using the stationary distribution and the average group success, the average group achievement (η_G) can then be computed, providing the overall probability of achieving success:

$$\eta_G = \sum_{k=0}^{Z} \bar{p}_k a_G(k) \quad (12)$$

Combined policies. Allowing the weight w to depend on the frequency of cooperators, we can derive the optimal balance s^* between positive and negative incentives by minimizing the defector's advantage ($F_D - F_C$). This is done similarly to [8], but using finite populations and therefore a hypergeometric distribution (see Eqs. (1), (2), (5), and (6)), to account for sampling without replacement. It can be shown that minimizing $F_D - F_C$ is equivalent to maximizing the following expression:

$$w \sum_{j=0}^{N-1} \frac{\binom{k-1}{j}\binom{Z-k}{n-1-j}}{\binom{Z-1}{n-1}} \left(\left[\frac{a}{j+1} - \frac{b}{N-j}\frac{k}{k-j}\frac{Z-k-N+1+j}{Z-k}\right]\right) \quad (13)$$

where j represents the number of $\boldsymbol{C}$s in a group of size N, sampled without replacement from a population of size Z containing k $\boldsymbol{C}$s. We let s* depend on k. Since this sum decreases as k increases, containing only one root, the solution to this optimization problem corresponds to having w set to 1 (pure Reward) for positive values of the sum, suddenly switching to w=0 (pure Punishment) once the sum becomes negative. The optimal switching point s* depends on the ratio a/b, group size N and population size Z. The effect of population size (Z), and group size (N) on s* is limited, while the impact of the efficiency of reward (a) and punishment (b) is illustrated in the following figure:

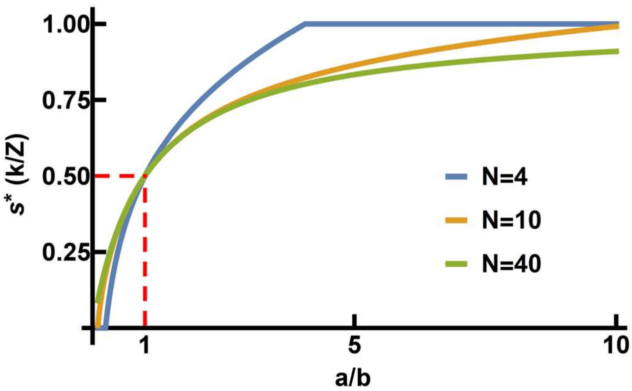

Figure 4. Optimal switching point s* as a function of the ratio a/b, for different values of N. s* corresponds to the optimal instant when there is a sudden switch between Reward and Punishment, given by the population configuration (k/Z) where it should occur. For all values of N displayed, the optimal switch point occurs at 0.5 if a=b. Population size Z=100.

In Text References

1. Barrett, S., *Self-enforcing international environmental agreements.* Oxford Economic Papers, 1994. **46**: p. 878-894.
2. Barrett, S., *Why cooperate?: the incentive to supply global public goods*. 2007: Oxford UP.

3. Dreber, A. and M.A. Nowak, *Gambling for Global Goods.* Proc Natl Acad Sci USA, 2008. **105**(7): p. 2261-2262.
4. Milinski, M., et al., *The collective-risk social dilemma and the prevention of simulated dangerous climate change.* Proc Natl Acad Sci USA, 2008. **105**(7): p. 2291–2294.
5. Santos, F.C. and J.M. Pacheco, *Risk of collective failure provides an escape from the tragedy of the commons.* Proc Natl Acad Sci USA, 2011. **108**(26): p. 10421-10425.
6. Vasconcelos, V.V., F.C. Santos, and J.M. Pacheco, *A bottom-up institutional approach to cooperative governance of risky commons.* Nat. Clim. Change, 2013. **3**(9): p. 797-801.
7. Sigmund, K., C. Hauert, and M.A. Nowak, *Reward and punishment.* Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10757-10762.
8. Chen, X., et al., *First carrot, then stick: how the adaptive hybridization of incentives promotes cooperation.* Journal of The Royal Society Interface, 2015. **12**(102): p. 20140935.
9. Hilbe, C. and K. Sigmund, *Incentives and opportunism: from the carrot to the stick.* Proceedings of the Royal Society of London B: Biological Sciences, 2010. **277**(1693): p. 2427-2433.
10. Gneezy, A. and D.M. Fessler, *Conflict, sticks and carrots: war increases prosocial punishments and rewards.* Proceedings of the Royal Society of London B: Biological Sciences, 2011: p. rspb20110805.
11. Sasaki, T. and S. Uchida, *Rewards and the evolution of cooperation in public good games.* Biology letters, 2014. **10**(1): p. 20130903.
12. Fehr, E. and S. Gächter, *Altruistic punishment in humans.* Nature, 2002. **415**(6868): p. 137-140.
13. Sigmund, K., *Punish or perish? Retaliation and collaboration among humans.* Trends in ecology & evolution, 2007. **22**(11): p. 593-600.
14. Masclet, D., et al., *Monetary and nonmonetary punishment in the voluntary contributions mechanism.* American Economic Review, 2003. **93**(1): p. 366-380.
15. Charness, G. and E. Haruvy, *Altruism, equity, and reciprocity in a gift-exchange experiment: an encompassing approach.* Games and Economic Behavior, 2002. **40**(2): p. 203-231.
16. Andreoni, J., W. Harbaugh, and L. Vesterlund, *The carrot or the stick: Rewards, punishments, and cooperation.* The American economic review, 2003. **93**(3): p. 893-902.
17. Milinski, M., et al., *Stabilizing the Earth's climate is not a losing game: Supporting evidence from public goods experiments.* Proc Natl Acad Sci USA, 2006. **103**(11): p. 3994-3998.
18. Pacheco, J.M., V.V. Vasconcelos, and F.C. Santos, *Climate change governance, cooperation and self-organization.* Phys Life Rev, 2014. **11**: p. 595-597.
19. Vasconcelos, V.V., et al., *Climate policies under wealth inequality.* Proc Natl Acad Sci U S A, 2014. **111**(6): p. 2212-2216.
20. Barrett, S., *Avoiding disastrous climate change is possible but not inevitable.* Proc Natl Acad Sci USA, 2011. **108**(29): p. 11733.
21. Barrett, S. and A. Dannenberg, *Climate negotiations under scientific uncertainty.* Proc Natl Acad Sci U S A, 2012. **109**(43): p. 17372-17376.
22. Milinski, M., T. Röhl, and J. Marotzke, *Cooperative interaction of rich and poor can be catalyzed by intermediate climate targets.* Climatic change, 2011: p. 1-8.
23. Boesch, C., *Cooperative hunting roles among Tai chimpanzees.* Human Nature, 2002. **13**(1): p. 27-46.
24. Creel, S. and N.M. Creel, *Communal hunting and pack size in African wild dogs, Lycaon pictus.* Animal Behaviour, 1995. **50**(5): p. 1325-1339.
25. Black, J., M.D. Levi, and D. De Meza, *Creating a good atmosphere: minimum participation for tackling the'greenhouse effect'.* Economica, 1993: p. 281-293.
26. Stander, P.E., *Cooperative hunting in lions: the role of the individual.* Behavioral ecology and sociobiology, 1992. **29**(6): p. 445-454.
27. Alvard, M.S., et al., *Rousseau's whale hunt? Coordination among big-game hunters.* Current anthropology, 2002. **43**(4): p. 533-559.
28. Pacheco, J.M., et al., *Evolutionary dynamics of collective action in N-person stag hunt dilemmas.* Proc R Soc Lond B, 2009. **276**(1655): p. 315-21.
29. Souza, M.O., J.M. Pacheco, and F.C. Santos, *Evolution of cooperation under N-person snowdrift games.* J Theor Biol, 2009. **260**(4): p. 581-588.
30. Pacheco, J.M., et al., *Co-evolutionary dynamics of collective action with signaling for a quorum.* PLoS Comput Biol, 2015. **11**(2): p. e1004101.
31. Skyrms, B., *The Stag Hunt and the Evolution of Social Structure*. 2004: Cambridge Univ Press.
32. Barrett, S., *Environment and statecraft: the strategy of environmental treaty-making*. 2005: Oxford UP.
33. Sigmund, K., *The Calculus of Selfishness*. 2010: Princeton Univ Press.
34. Traulsen, A., M.A. Nowak, and J.M. Pacheco, *Stochastic dynamics of invasion and fixation.* Phys. Rev. E, 2006. **74**(1 Pt 1): p. 011909.
35. Traulsen, A., et al., *Exploration dynamics in evolutionary games.* PNAS, 2009. **106**(3): p. 709-12.
36. Tavoni, A., et al., *Inequality, communication and the avoidance of disastrous climate change in a public goods game.* Proc Natl Acad Sci USA, 2011. **108**: p. 11825-11829.
37. Santos, F.C., J.M. Pacheco, and B. Skyrms, *Co-evolution of pre-play signaling and cooperation.* J Theor Biol, 2011. **274**(1): p. 30-35.