TÉCNICO LISBOA

Reward and Punishment in Climate Change Dilemmas

António Miguel da Rocha Gonçalves Góis

Thesis to obtain the Master of Science Degree in

Information Systems and Computer Engineering

Supervisors: Prof. Francisco João Duarte Cordeiro Correia dos Santos

Dr. Fernando Pedro Pascoal dos Santos

Examination Committee

Chairperson: Prof. Daniel Jorge Viegas Gonçalves

Supervisor: Prof. Francisco João Duarte Cordeiro Correia dos Santos

Member of the Committee: Prof. Alexandre Paulo Lourenço Francisco

May 2018

Resumo

Acordos climáticos podem ser formulados como problemas de bem comum que envolvem grupos de indivíduos. Através desta abordagem, pode-se requerer um número mínimo de cooperadores dentro do grupo para preservar o meio ambiente com sucesso. Além disso, as decisões individuais podem depender da perceção de risco associado a perdas futuras. Neste trabalho é investigado o impacto da implementação de recompensas e sanções neste tipo de empreendimento coletivo — batizado como Dilema de Risco Coletivo — através de uma abordagem dinâmica e evolutiva. Os nossos resultados são gratificantes, dadas as limitações à priori dos mecanismos de sanções neste contexto global: Por um lado mostramos que incentivos positivos (recompensas) são essenciais na promoção da cooperação, principalmente quando a perceção de risco é baixa e o número de indivíduos cooperadores é reduzido. Por outro lado, incentivos negativos (sanções) podem servir para manter a cooperação, depois de esta já estar instalada. Por fim, mostramos como o Dilema de Risco Coletivo é alterado através da sinergia entre recompensas e sanções, quando estas atuam em simultâneo.

Palavras-chave: Alterações Climáticas, Cooperação, Recompensas, Teoria de Jogos Evolutiva, Sanções

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.

Keywords: Climate Change, Cooperation, Evolutionary Game Theory, Punishment, Reward

Agradecimentos

Agradeço...

Ao Prof. Francisco Santos, pela imensa disponibilidade que teve no decorrer deste trabalho. A sua preocupação genuína e profunda com os seus alunos, tanto a nível profissional como pessoal, torna o trabalho numa caminhada muito mais leve.

Ao Prof. Jorge Pacheco e Fernando Santos pelas contribuições valiosas para o desenvolvimento deste projeto.

Aos meus pais, que sempre me mostraram a importância do papel da ciência na sociedade (ainda que por vezes em detrimento de outros assuntos igualmente centrais na minha vida). Sem dúvida que eles são os principais responsáveis pela conclusão desta etapa, por acreditarem tanto nas minhas capacidades e me motivarem para o seu uso.

À prof. Ana Paiva por se ter interessado pelo meu percurso menos comum e me ter posto em contacto com o Prof. Francisco Santos em consequência disso.

À Fundação para a Ciência e Tecnologia (FCT) e ao INESC-ID pela possibilidade de participar no projeto eCoop e de ser apoiado com uma bolsa de investigação no decorrer deste trabalho.

Às pessoas com quem me cruzei até hoje, que, de modo informal, tiveram um papel fundamental no desenvolvimento do meu sentido crítico. Se os meus pares fossem creditados por toda a educação que me deram, alguns inesperadamente iam ganhar o título de Doutor. O saber também existe fora das instituições, e espero nunca me esquecer disso.

Table of Contents

List of figures

List of Acronyms

EGT	Evolutionary Game Theory
ESS	Evolutionarily Stable Strategy
TFT	Tit-for-Tat
PD	Prisoner’s Dilemma
PGG	Public Good Game
CRD	Collective Risk Dilemma

1. Introduction

1.1. Have you seen Homo Economicus?

The quest to understand cooperation is an ambitious one. In trying to model human behaviour, classical economics has assumed throughout history that people are utility-maximizing agents. Dating back to Bernoulli [1] and von Neumann [2], the concept of utility has been present in countless analysis of human behaviour. Such approach can seem appealing due to its simplicity and the intuitive axioms that it lies upon, such as the transitivity of preferences ($A \geq B$ and $B \geq C$ implying $A \geq C$). Denying that we act according to these simple rules could render us inconsistent and arguably irrational, depending on the definition considered.

However, as Amartya Sen put it, this focus on rational behaviour and self-interest may result in advice that is not necessarily ethical [3]. As Sen noted on a different work, *"The purely economic man is indeed close to being a social moron"* [4], ironically implying that such an utilitarian description of individuals does not match reality, and when it does, it may not translate the most impeccable behaviour. Indeed, utility maximization is often associated with greedy self-centred behaviour that often is not in line with real-world observations [5]. It is therefore a challenge to explain altruistic behaviour through a theoretical framework.

Summing up, it is true that a model is a simplification of reality. In fact, this is a necessary condition for the model to be of any use and a reason models may fail at times. However, the predictive power of assuming selfishness seems to be weak, particularly when trying to analyse real-scenarios of cooperation between so-called rational agents. Motivated by this incongruence, a great deal of effort is being put in explaining altruism from a rational point of view. It could be that humans are not rational after all. Or it could be that it is rational to cooperate. Different definitions of rationality will lead to different answers [3], however that is not the main subject of this thesis. As long as some sort of "internal consistency" exists, then human behaviour can be modelled. In that case, it is possible to analyse the impact of adding simple features to said models. By doing so, research can aim at identifying the main ingredients of cooperation. And if such an endeavour turns out to be successful, our descriptive and prescriptive powers over a vast number of scenarios may be significantly improved.

1.2. Defining Cooperation

Cooperation, in a colloquial sense, can be defined as a group of individuals working for mutual benefit, instead of engaging in selfish competition for individual benefit. In some cases, it may not be possible for an individual to engage in cooperative behaviour without his peers cooperating simultaneously (e.g. two individuals pushing a large piece of furniture that neither could move alone). In these simple situations it may present no challenge to coordinate efforts and trust the action of our peer. However, in many relevant scenarios this is not the case. As we increase group size and face more complex situations it becomes more likely that one must decide whether to cooperate before knowing his peers' choice of action. To illustrate this idea, let us consider a set of communities (that could range in size from neighbourhoods to countries) which must decide whether to stop producing toxic waste or not. There may be an individual incentive (in the form of short-term economic profit) for a community to keep producing it, contrasting with a collective incentive for all communities to stop and protect their common natural resources. Each community will not know the action of the others before enough time has passed for the possible environmental harm to become visible or until the polluters fail in hiding the waste. Other relevant examples include avoiding overharvesting natural resources, recycling or paying taxes.

To account for these situations where one may have to cooperate while risking that others defect, we accept the following definition of cooperation: to incur an individual cost for providing collective benefit. This way we allow for situations where everyone cooperates and ends up better off, but also for those where some cooperate and others do not, free-riding on the efforts of cooperators. Typically, the perceived individual benefit of cooperating is smaller than the cost of one's action, causing a dilemma – individuals are better off if everyone cooperates than if no one does, but each would maximize his own payoff by avoiding cooperation himself. The area of mathematics which studies these dilemmas is called Game Theory, and its most studied problem is the Prisoner's Dilemma, which will be formalized later. Other games include the Ultimatum, the Trust and the Dictator Game [6]. In classical game theory, games are typically played between two participants – or a small group – assumed to be rational.

However, real-world scenarios usually involve larger groups of individuals that resort to adaptation heuristics such as social learning. For this reason, talking about the emergence of cooperation requires the study of population dynamics. When a collective behaviour is observed, it can be seen as the sum of many individual decisions. Even if the assumption is that individual agents oblige to a simple set of rules, it is possible that a more complex and unexpected behaviour arises from these combined interactions. To study this, rather generic frameworks are used, which resemble real scenarios at very different scales. This generalization of game theory to large populations and to adaptive behaviour is known as Evolutionary Game Theory [7-11]. Due to the variety of scales at which population dynamics occur, Evolutionary Game Theory is relevant to many different subjects. Its insights are meaningful to fields ranging from cell biology to human behaviour. Recently it has been applied to study how the cancer phenotype is affected by interactions between normal and malignant cells [12]. It is also used to understand human cooperation problems, for instance in the context of climate change and global governance [13,14]. The dynamics subjacent to an evolutionary game theoretical analysis can be grasped resorting to large-scale agent-based computer simulations. In particular situations – as is the case in this thesis – it is possible to describe the evolutionary dynamics resorting to stochastic processes (e.g., through Markov Chain analysis). Given the ubiquity of population dynamics, the relevance of studying Evolutionary Game Theory becomes clear. The sooner we understand how it works, the sooner this broad range of problems can be tackled in a structured manner.

1.3. Incentives for Climate Cooperation

One paradigmatic example of a global or large-scale cooperation dilemma is the task of designing climate change agreements. Climate change stands as one of our biggest challenges in what concerns the emergence and promotion of cooperation [15,16]. 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 classic case of a Public Good Game [17]: 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 widespread future losses, with profound effects on the expected dynamics of cooperation [13,18]. Another challenge relates to the lack of knowledge about efficient ways of combining 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) [19]. In the larger context of cooperation studies, positive incentives (rewards), negative incentives (punishment) and the combination of both [6,20-28] have been shown to have different impacts, 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. To address these issues, a recently formulated game known as Collective Risk Dilemma (CRD) [13] will serve as the basis for this work. Adding a mechanism of reward and punishment on top of it, we study the impact of each policy and their combined effect under different conditions.

1.4. Outline

In the current chapter we have presented the difficulties of aligning an assumption of selfishness with actual altruistic behaviour. Evolutionary Game Theory was briefly introduced, which can be used as an alternative framework to understand cooperation. Finally, we presented the main focus of this thesis – to understand cooperation in N-person climate change dilemmas, and to study the impact of adding reward and punishment mechanisms to such scenarios. However, to introduce this topic we should take one step back. As if departing from unicellular organisms before studying multicellular ones, Chapter 2 will begin by introducing classical game theory and 2-person games before discussing N-person dilemmas. It will follow with a review of mechanisms known to solve 2-player cooperation dilemmas. Finally, Chapter 2 introduces the open problem of cooperation in N-person games, including the CRD studied here. In Chapter 3, existing literature on reward and punishment is reviewed, both on a theoretical and experimental level. This pair of mechanisms is introduced as a possible tool to solve N-player cooperation. Next, in Chapter 4, CRD is formalized and incentives are added, allowing for reward and punishment to be applied (separately or simultaneously). In Chapter 5 we explain how reward is overall preferable to punishment to achieve collective cooperation, in some cases performing equivalently and in others significantly better. Only when risk perception is low and requirements to meet the collective goal are tight, does the synergy of both policies combined become crucial. Finally, Chapter 6 summarizes our conclusions and presents future work.

2. Games of Cooperation

This chapter will begin with an introduction on games between two individuals. Particularly, the concept of a payoff matrix is explained and the paradigmatic case of the Prisoner's Dilemma is presented, which can be used as a metaphor for a generic cooperation dilemma. Then we transition from classical Game Theory to Evolutionary Game Theory (EGT), where a population of individuals exists, and in each round players are sampled from this population to interact and adapt behaviour based on social learning. Then we follow with a review of mechanisms known to solve the cooperation problem between two players. Finally, we expand these concepts to N-player games. Explaining which mechanisms allow cooperation to emerge in this broader class of games is an open problem for which this thesis aims to contribute.

2.1. Two-Person Games

In this section, the basic concepts of game theory will be formalized. We will begin with traditional game theory, introducing Nash Equilibrium and Pareto Optimal, together with some classical 2-person games. Next, EGT concepts will be formalized, such as the replicator equation, evolutionarily stable strategies and the distinction between finite and infinite populations.

2.1.1. On Classical Game Theory

Game Theory dates back to the beginning of the twentieth century, with von Neumann's work [2] followed by John Nash's [29]. It began by constituting a mathematical set of tools to analyse simple settings, where typically a small group plays one-shot games. Each person has a set of possible actions to choose from and must select one prior to knowing the opponent's option. After both players lock their actions, a reward is attributed to each depending on the pair of chosen actions.

A game is often defined resorting to a payoff matrix. This matrix has *nm* entries, where *n* is the number of possible actions for player A and *m* for player B. Each entry corresponds to a specific pair of actions N and M, and contains two values: the payoff for player A and for player B given that N and M were chosen. Let us define strategy N as S_N, and the payoff of choosing S_N against S_M as $E(S_N, S_M)$. If we consider that player A can choose an action from the set $\{S_U, S_V\}$ and player B from $\{S_X, S_Y, S_Z\}$, we have the following matrix:

Player A \ Player B	X	Y	Z
U	$E(S_U,S_X), E(S_X,S_U)$	$E(S_U,S_Y), E(S_Y,S_U)$	$E(S_U,S_Z), E(S_Z,S_U)$
V	$E(S_V,S_X), E(S_X,S_V)$	$E(S_V,S_Y), E(S_Y,S_V)$	$E(S_V,S_Z), E(S_Z,S_V)$

Figure 2.1 – Payoff matrix for a 2-player game

Various games have been modelled in order to mimic different real-world scenarios. Some examples are the Prisoner's Dilemma (PD), the Stag-Hunt and the Snowdrift Game [30]. The metaphor that originated the name of PD is the following: two prisoners must decide separately whether to confess the crime or not. Individually each of them is better off if he confesses regardless of the other's choice, but confessing implies harming the other prisoner. If it were possible to have a previous agreement, they would both prefer a scenario where neither of them confessed the crime. The payoff matrix for this game can be set as in Figure 2.2, where each player wants to minimize his jail time. Each individual will have his sentence reduced by confessing, but they will be much worse off if both confess. This creates an interesting problem because they will end up locked in a sub-optimal equilibrium, in this 2-person single-shot setting.

Player A \ Player B	Confess	Not Confess
Confess	5, 5	0, 20
Not Confess	20, 0	1, 1

Figure 2.2 – One possible payoff matrix for a Prisoner's Dilemma, displaying jail time in years

It is helpful to introduce two concepts at this point, to help describe this issue. A Nash Equilibrium occurs when no agent can be better off by individually changing his action. This can be formalized in the following way:

$$E(S_k,S_k) \geq E(S_i,S_k), \forall i \neq k \tag{2.1}$$

Or, if we consider a strict Nash Equilibrium:

$$E(S_k,S_k) > E(S_i,S_k), \forall i \neq k \tag{2.2}$$

Pareto Optimality occurs when it is impossible to improve an agent's payoff without harming at least one other. In PD, the only Pareto-optimal state is when both cooperate. The only Nash equilibrium is when both confess. Since the Nash equilibrium does not match the Pareto-optimal in this game, the natural course of action will be to end up in an undesired state, where both confess. Here lies the core of this social dilemma. Conveniently, we can now clarify the discussion about rationality and cooperation that was started in Section 1.1. In fact, the Nash Equilibrium seemingly predicts the pair of actions to be employed by rational players: as the payoff matrix fully translates their preferences, they should both defect. However, reality disproves this prediction [5]. The study of the institutional arrangements and incentives that can both explain cooperation and foster it where defections prevails – while keeping the assumption that individuals seek to maximize their payoff through social learning – is a pressing challenge (the one that here, as evidenced below, we address).

Other stories besides the PD can lead to a similar payoff matrix. One worth mentioning is the 2-person Public Good Game (PGG), represented in Figure 2.3. Note that from now on we will only show player A's payoff, unless the payoff matrix isn't symmetric. In this game, each player decides whether to contribute to a common pool with a fixed amount of money c. After each player locks his decision the total amount in the pool is multiplied by a factor m, with 1<m<2. It is then divided by both players, no matter if they contributed or not. Similarly to PD, each agent is better off if he doesn't help the other, regardless of the opponent's decision.

		Player B	
		Contribute	Don't Contribute
Player A	Contribute	cm - c	cm/2 - c
	Don't Contribute	$cm/2$	0

Figure 2.3 – 2-player Public Good Game

Interestingly, any of these matrixes can be expressed in a more generic form (Figure 2.4), which unifies both stories. From this generic perspective, to cooperate is to accept a cost in order to provide the opponent with a benefit. Regardless of the opponent's choice, defection is the action which maximizes the individual's payoff. This way, "cooperation" can refer specifically to the state where both agents choose to cooperate, inside this 2-person game. Later, when we expand to populations or to N-person games, it will be possible to quantify cooperation by counting how many individuals choose to cooperate.

		Player B	
		Cooperate	Defect
Player A	Cooperate	benefit – cost	– cost
	Defect	$benefit$	0

Figure 2.4 – Generic Prisoner's Dilemma payoff matrix

2.1.2. Transitioning to Evolutionary Game Theory

By only using the concepts introduced so far, one can study a single interaction between two players. However, this is a very limited approach, since the combination of many interactions spread throughout space and time may change the results drastically. In fact, introducing populations also allows studying adaptation mechanisms whereby individuals resort to information provided by others, in order to adapt their behaviour. While in classical game theory individuals are assumed to reason in isolation and, understanding the interaction they face, choose the most seemingly profitable strategy, in a population individuals may be assumed to copy the behaviour others, as in reality people often do [31]. Introducing complexity through large populations of interacting agents is thereby key for understanding real-world scenarios, and the main framework of complex systems research.

Defining complexity or a complex system is hard. One possible definition proposed by Melanie Mitchell [32] is a *"system in which large networks of components with no central control and simple rules of operation give rise to complex collective behaviour, sophisticated information processing, and adaptation via learning or evolution"*. Using this notion, a component may be a single agent acting according to his payoff matrix, while the network can be a population of agents from which a collective behaviour emerges that is harder to understand.

The tool to introduce complexity and population dynamics in game theory is Evolutionary Game Theory (EGT). Using it, it becomes possible to analyse cooperation in larger groups of agents by observing the emergent behaviour that results from their interactions. In order to present EGT some important extensions will be made to the model explained so far. Regarding the number of agents, so far only the agents necessary for one round were considered. In a 2-person game setting, only two agents existed. However, it is possible to have a larger population of agents, from which the necessary number of agents is drawn at random each time a new round will occur. If any pair of agents is allowed to interact with equal probability, the population is called a well-mixed population. Details on topologies and spatial restrictions will be given later.

With a population larger than the number of needed players, different pairs or groups of players can be chosen in each round. Due to this, the payoff of an isolated interaction becomes insufficient as a metric to evaluate the quality of a strategy, since distinct groups composed by players using different strategies may bring different payoffs. A new metric can be introduced to overcome this, which is fitness. Fitness combines the payoff of each possible scenario with the probability that it occurs. It will be formalized in more detail when introducing N-player games, since a general rule can be given for groups of sizes equal or larger than 2. For now, fitness can be regarded as a computation which returns the accumulated or average payoff of a given strategy.

One of the main ideas of EGT is that the choice of strategy is not purely based on maximizing the returns of a single round or of a pre-determined number of rounds. Strategies may be initialized at random, for instance from a uniform distribution. Then, strategies are updated based on what the agents observe others doing. This way, strategies which are more frequent or bring more fitness are more likely to be imitated than others. The details of these update rules vary depending on whether the population has a finite or infinite size. These concepts will be explained in the next sub-sections.

Given the complexity introduced by EGT, Nash Equilibrium becomes insufficient to analyse a given strategy. In fact, now we must distinguish the strategies that can both invade and remain stable in a population. The expansion of Nash Equilibrium to EGT results in the idea of an Evolutionarily Stable Strategy (ESS). A strategy S_k is evolutionarily stable if it obeys one of the following conditions:

$$E(S_k, S_k) > E(S_i, S_k), \forall i \neq k \tag{2.3}$$

or

$$[E(S_k, S_k) = E(S_i, S_k)] \wedge [E(S_k, S_i) > E(S_i, S_i)], \forall i \neq k$$

The first condition represents a strict Nash equilibrium. The second one is valid if strategy S_i is equally effective against S_k, but loses advantage against S_i. This second condition is known as “Maynard Smith’s second ESS condition” [33]. John Maynard Smith, a founding father of EGT, also provides a more intuitive perspective on ESS: *"An ESS or evolutionarily stable strategy is a strategy such that, if all the members of a population adopt it, no mutant strategy can invade"* [34].

With this definition, a new question can be posed: which mechanisms allow for cooperation to become an ESS? This is motive of wide research [5], and the answers naturally depend on the assumptions made about the population structures and the update process of individuals. The next sub-sections will deal with update rules, vital to define the evolution of a model. Then, in section 2.2, cooperation mechanisms for 2-person games will be presented.

Infinite Populations and the Replicator Equation

Surprisingly, infinite populations create a simpler model than finite ones. In the absence of any population structure, and deterministic update rules and behaviours, all individuals can be seen as perfectly equivalent. Moreover, we can assume that the average outcome of all interactions is precisely the expected value. This way, a population can be represented only by its fraction of cooperators instead of using absolute values.

This introduces the idea of the replicator equation [35]. Given the concept of an infinite population, it is natural to ask how it evolves. Two key factors determine the evolution of a strategy: its frequency and its fitness. Frequent strategies are observed more often, so agents have higher chances of imitating them. Strategies yielding higher fitness are more likely to be imitated once they are observed. This intuition is contained in the general definition of the replicator equation, where x_i is the fraction of the population using strategy *i* and $\langle f(x) \rangle$ is the average fitness of the population, given by $\sum_{j=1}^{n} x_j f_j(x)$.

$$\dot{x}_i = x_i(f_i(x) - \langle f(x) \rangle) \tag{2.4}$$

$\dot{x}_i$, the replicator equation for strategy *i*, describes how *i* varies depending on the frequency of all strategies. It is a differential equation obtained from the derivative of $x_i(t)$, the fraction of the population using strategy *i* over time.

It is now possible to study the particular case where defection and cooperation are the only two strategies. Let's define x as the fraction of cooperators in a population, f_C as the fitness of cooperators and f_D as the fitness of defectors. The evolution of cooperators and defectors follows a gradient of selection g(k). In infinite populations, this gradient can be described by the replicator equation $g(k) \equiv \dot{x}$. From the general replicator equation, a more specific rule can be obtained for the case of 2 strategies:

$$\dot{x} = x(1-x)[f_C(x) - f_D(x)] \tag{2.5}$$

This can be interpreted as the speed at which the fraction of cooperators changes for a given value of x. The element x(1-x) can be seen as the probability of drawing two agents with different behaviours to play a 2-person game, which is the only interaction where an agent can switch strategy. $f_C(x) - f_D(x)$ will determine the direction of change in x, as well as contribute to the intensity of this change. Since this value depends on x, it is said that the model is frequency-dependent. By plotting $\dot{x}$ against x in a PD, we obtain the following graphic:

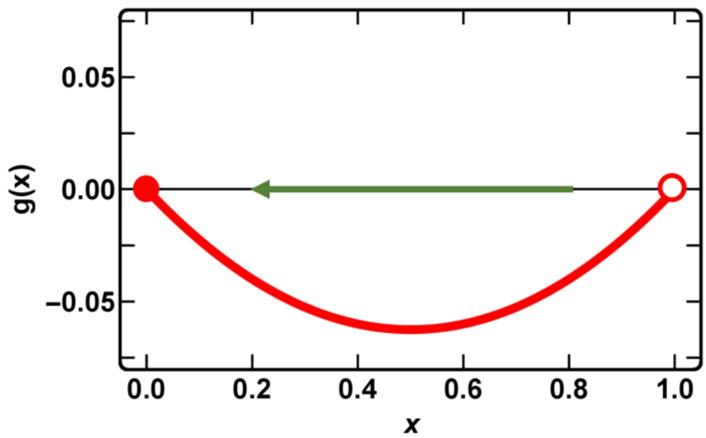

Figure 2.5 – Gradient of Selection for Prisoner's Dilemma with an infinite population

Using this plot, the growth rate of cooperation – for any fraction x of cooperators – can be observed, which follows directly from the replicator equation. In this case, it can be observed that any initial frequency of x will converge to an all-defectors scenario. The only exception is when x=1, since there are no defectors to imitate. Both x=1 and x=0 have zero speed, meaning that, without disturbances, the system will stay still after reaching one of those states. These are called equilibria, represented with circles in the figure. However, x=1 is represented by an empty circle since it isn't stable: it is enough for one defector to appear in order to move the system away from this stationary point. For this reason, when they occur within $x \in]0,1[$, they are called coordination points. It only requires an initial coordination effort to move away from them. It certainly isn't desirable to coordinate an effort in order to move away from x=1, but future examples will provide more meaning to this name. On the other hand, x=0 is represented by a filled circle, since a small perturbation isn't capable of moving the system away from this stationary point. This type of point is called a co-existence point.

Finite Populations and the Fermi Equation

When moving to finite populations, we are not able to use, anymore, the tools that relied on the system's deterministic properties. We must consider stochastic effects. It also becomes impossible to represent the states of the system by using continuous functions. Since there is a finite number of agents, the possible states between 0% and 100% of cooperators correspond to a discrete domain. Without continuity, it is impossible to differentiate $x_i(t)$ in order to obtain the speed of change in the system. For this reason, the logic of the replicator equation no longer applies here. The system may be thought of as a Markov Chain, where each state represents a given amount of cooperators. At each step, a system with k cooperators may stay in the same state, switch to k–1 or to k+1.

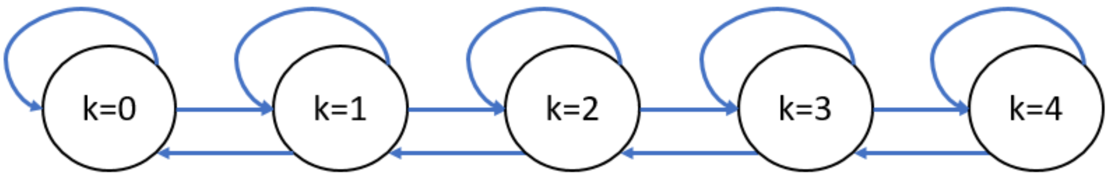

Figure 2.6 – Markov Chain for a population of size 4

There are two options to study these systems: one can analyse the transition probabilities in each state or use computer simulations to estimate the average behaviour of the system. Depending on the complexity of the model, it may not be viable to represent all transition probabilities, even more so with complex spatial structures.

It is now possible to elaborate on this model in order to compute each state's transition probabilities. A player X and a neighbour of his are drawn at random from the population. Player X decides whether to imitate his neighbour depending on the difference between their fitness. This probability can be given by the Fermi function:

$$p \equiv \left[1 + e^{-\beta(f_j - f_i)}\right]^{-1} \tag{2.6}$$

which has the shape of an S-curve and probability 0.5 if both players have the same fitness value. β is a parameter that controls the steepness of the curve (known as an inverse temperature in physics). A β close to zero will approximate a random decision, whereas a high value will be similar to a deterministic rule using a step-function [13,36]. 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.

$$T^{\pm}(k) = \frac{k}{Z}\frac{Z-k}{Z}\left[1 + e^{\pm\beta[f_C(k)-f_D(k)]}\right]^{-1} \quad (2.7)$$

The gradient of selection in infinite populations (using the replicator equation) can now be replaced with the difference $G(k) \equiv T^+(k) - T^-(k)$, resulting in the the following plot:

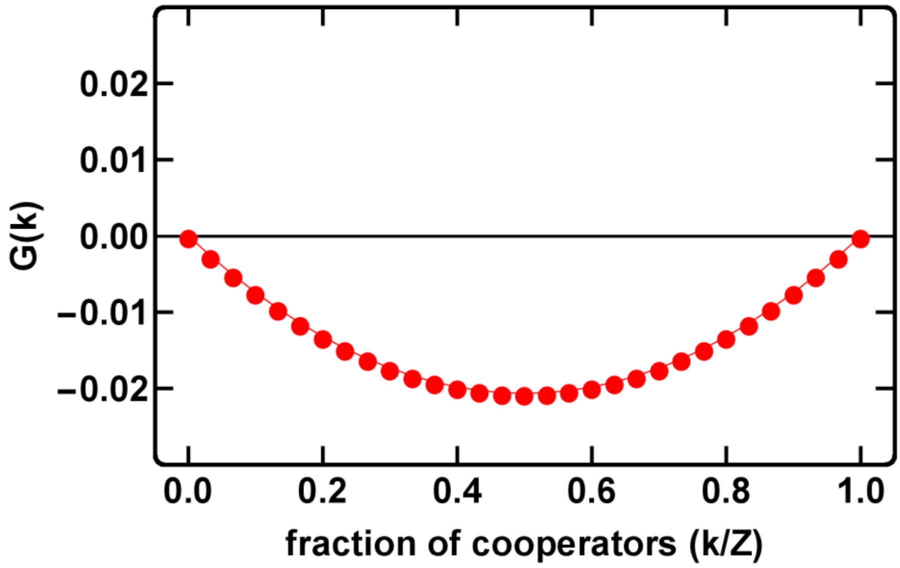

Figure 2.7 – Gradient of selection for Prisoner's Dilemma with a finite population of size Z=30

Under weak selection ($\beta \ll 1$), in the limit when $Z \rightarrow \infty$, the replicator equation is recovered from the transition probabilities [36]. On a final note regarding population structures, finite populations no longer need to allow interactions between all pairs of users. For this reason, structures more restrictive than well-mixed populations can also be studied.

2.2. Mechanisms for Cooperation

As explained in the previous section, the PD or 2-person PGG will converge to defection. One way to study which mechanisms enable cooperation in this class of dilemmas is by simulating artificial societies. There are other ways, for instance with human users and real money playing through computer terminals among each other [5]. This is particularly useful to test whether model assumptions fail or not with humans. Much work has been done is this field, both with synthetic agents and human users, particularly on 2-person games [5]. Certain mechanisms which are analogous to real-world scenarios prove to be effective in improving the outcome of these cooperation games. Here we will review these mechanisms.

To define an evolutionary game, a payoff matrix together with an update rule can be used. A payoff matrix defines the outcome of a single round. An update rule defines how the agent, after observing the outcome of one round, decides whether to mimic his opponent or not. Given this basic model design, certain restrictions or additional rules can be included. This section explains five mechanisms which are known to improve cooperation in 2-person games.

Probably an agent is more inclined to cooperate if he believes that his opponent has cooperated in the past. This is the general notion of reciprocity. Trivers [37] has pioneered work in this area. Generally, by including this information and using it in the decision rule, cooperation can be fostered. Here two ways of obtaining this information are presented.

In many scenarios, repeated encounters between the same pair of individuals is more realistic than considering that the pair will never meet again, or that they won't remember their past encounters. If we consider repeated encounters between the same pair, it is possible to adjust one's strategy based on these past encounters. By recalling past direct encounters, one can use this information to decide whether to cooperate or not. A well-known strategy which uses this information is Tit-for-Tat (TFT). It simply states that an agent should begin by cooperating, and then mimic the strategy used by the opponent on the previous round. A tournament was promoted by Axelrod [38] where participants could submit strategies which would then be used in computer simulations to play against the other participants. TFT was the winner, outperforming much more complex strategies. Later a second edition of the tournament was made, where several strategies were tailor-made to defeat TFT. However, during the tournament all strategies competed between each other and TFT was the winner once again.

Alternatively, information on the opponent can be obtain indirectly. Third-party agents may share their previous experience with the two agents that are about to engage in a new round [39-43]. This can be done publicly or privately. An example of public information is the user rank on eBay. Anyone can know if other users were satisfied after engaging in transactions with a given seller or buyer. An example of private information is when two people gossip about a third person, sharing information which isn't publicly available. Designing strategies which use this information is more complex than using direct reciprocity. Moral dilemmas arise, such as "should one be penalized for not helping an agent if that agent is usually a defector?". In any case, accessing and using this information fosters cooperation in 2-person games.

Up to this point, only well-mixed populations were considered. Any pair of agents is allowed to interact. This is clearly unrealistic in most scenarios, where people are limited by various constraints. In finite populations, it is possible to specify which pairs of agents are allowed to interact. This type of restriction is enforced by imposing a structure on the population, which may be created randomly or by following a rule. Certain simple rules can easily be interpreted as imposing a spatial organization and, due to their simplicity, were among the first to be studied.

Spatial restrictions were first proposed by Nowak and May [44], who observed that this mechanism alone could make cooperation emerge in scenarios which would otherwise lead to full defection. A simple topology called lattice was used. It consists of a regular two-dimensional grid where each node is only allowed to interact with his direct neighbours. As a result, all nodes have exactly the same number of neighbours, typically 4 or 8. The neighbour count of a node is known as its degree. In Nowak's experiments spatial restrictions caused clusters of cooperators to emerge, surrounded by defectors. These clusters would move in unpredictable patterns, creating shapes that became known as Persian carpets.

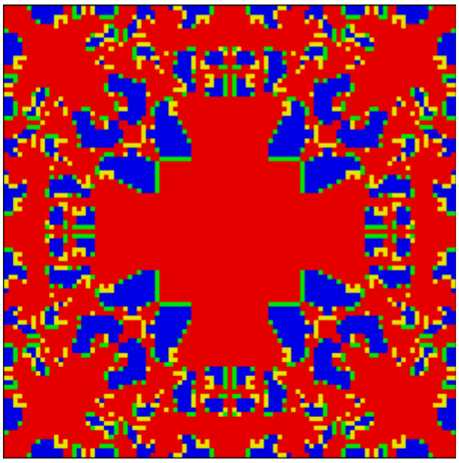

Figure 2.8 – Cooperation in a structured population. Screenshot of Nowak's experiments, replicated from scratch using VPython. A Prisoner's Dilemma game is being played on a 4-neighbourhood lattice. Blue and red represent cooperators and defectors, respectively, which didn't switch behaviour since the previous round. Green and yellow represent cooperators and defectors, respectively, which changed previous round's behaviour.

Since then, network science has taken major steps and more complex topologies were introduced. One of the most important advances was the notion of a scale-free network [45]. This network uses the concept of preferential attachment, which implies that a new node in the network is more likely to connect to nodes which already have many connections, called hubs or high-degree nodes. As a result, its degree distribution – a function which returns the node frequency of a given degree – follows a power law, resulting in a highly right-skewed distribution. This means that there are many low-degree nodes, and very few hubs. This type of network can be more realistic in most scenarios, and fosters cooperation more effectively than lattices [30,46-48].

Another important notion is that agents may compete in groups, rather than individually. In this case defectors may have an advantage inside their groups, but groups of cooperators will clearly outperform groups of defectors. To model this, imitation will occur at group level, instead of occurring at agent level. This type of game will enable cooperative outcomes [49-51].

Finally, Kin Selection occurs when an agent prefers to help his relatives. This idea was first introduced by Hamilton [52], and it is meaningful if an individual has the goal of safeguarding and propagating his genes. To fulfil this goal, an individual is interested in improving his fitness but also the fitness of his relatives. This idea is known as the selfish gene [53]. This is perhaps the most controversial of the 5 mechanisms, due to the strong assumptions that it makes and to real-life examples of cooperation among non-genetically-related individuals.

2.3. N-Person Games

In the previous section, it was concluded that certain simple features can change the course of 2-person games. This brings us one step closer to understanding cooperation. However, many interactions in real life occur between larger groups. Transitioning from 2-person to N-person games isn't obvious, neither is to adapt its cooperation-enhancing mechanisms. This section will introduce N-person games.

2.3.1. N-Person Public Good Game

To introduce this new class of problems, it is useful to recall the 2-person PGG presented in section 2.1. Two agents can decide whether to contribute to a common pool. The total contribution is multiplied by m, with $1<m<(N=2)$. Then, the multiplied contribution is distributed by both players without considering if each cooperated or defected. Given this scenario, it is clear how to extend the game to N players. Each of the N players can decide individually whether to contribute, and the multiplied contribution is distributed by the N players.

It is now possible to define the payoff of a defector in a N-person PGG. For a group size N with j contributors, where each cooperator makes a contribution c that will be multiplied by a factor m, the payoff of a defector is:

$$\Pi_D(j) = \frac{jmc}{N} \tag{2.8}$$

A cooperator will also receive this amount, but he will have spent his initial contribution c in the beginning of the game:

$$\Pi_C(j) = \Pi_D(j) - c \tag{2.9}$$

This poses a dilemma similar to the 2-person PGG, except that now there is more than one opponent who can potentially choose to free-ride by defecting, reducing j and making everyone worse off except for himself.

Many efforts are being conducted to understand this game and prescribe cooperation-enhancing techniques. Some important findings have been made, such as the importance of having several small groups instead of a large one [13], but the overall problem is yet to be solved.

2.3.2. Defining Fitness in N-player Games

The definition of fitness was postponed to this section in order to present a general rule, which can be applied both to 2-player and N-player games. This metric is fundamental to compare an agent's strategy with his partner's. To compute it one needs to consider the probability of forming each possible group composition, depending on the population from which agents are sampled. This is what separates the fitness formula for infinite and finite populations. Given this probability distribution, the payoff of an agent in each group setting is weighted with the probability that it occurs.

Infinite populations imply that agents are sampled with replacement. If a cooperator is picked, the number of cooperators left will not change since the population is infinite, leaving the distribution unchanged as well. This corresponds to a binomial distribution, which is used to calculate the average payoff in this scenario. Naturally, f_C and f_D have different formulas, since the agent's own strategy will affect both the group configuration and his payoff. Let $\Pi_S(k)$ be the payoff of strategy S with k cooperators, x the fraction of cooperators in the population, N the group size and k the number of cooperators in the group, excluding the agent if he is a cooperator himself.

$$f_C = \sum_{k=0}^{N-1} \binom{N-1}{k} x^k (1-x)^{N-1-k} \Pi_C(k+1) \tag{2.10}$$

$$f_D = \sum_{k=0}^{N-1} \binom{N-1}{k} x^k (1-x)^{N-1-k} \Pi_D(k) \tag{2.11}$$

With finite populations, picking one agent will change the set of the remaining ones from which to pick more players. While this may be insignificant in some cases, in others it can be crucial. Given a population with a single cooperator it is impossible to have groups with two or more cooperators, even with the lowest of probabilities. To avoid this problem a hypergeometric distribution is used, modelling a scenario of sampling without replacement. N is the group size, Z the population size, k is total number of cooperators in the population and j the number of cooperators picked for the group.

$$f_C = \binom{Z-1}{N-1}^{-1} \sum_{j=0}^{N-1} \binom{k-1}{j} \binom{Z-k}{N-j-1} \Pi_C(j+1) \tag{2.12}$$

$$f_D = \binom{Z-1}{N-1}^{-1} \sum_{j=0}^{N-1} \binom{k}{j} \binom{Z-k-1}{N-j-1} \Pi_D(j) \tag{2.13}$$

2.3.3. Threshold Games

Other games besides PGG exist to illustrate N-person dilemmas, such as the N-person ultimatum game [54] or the N-person coordination game [55]. Here another class of N-person games is presented, which is central to this thesis and may be more realistic in several real-world situations.

The N-person PGG, which was previously introduced, represents a specific class of real-world scenarios. As the number of cooperators varies, each person's reward varies linearly according to it. However, in real life there are thresholds that must be circumvented before collective cooperation renders some benefit or before some harm can be prevented. Consider coral reefs – they will tolerate a certain amount of perturbation caused by a reduction in quantity and variety of fish [56,57]. However after crossing a certain threshold the system will suffer abrupt changes. Similarly, many issues related to climate change are thought of as varying according to thresholds. A climate tipping point is a conceptual point when global climate changes from one state to another, possibly in an irreversible way [58]. Another relevant example is the particular case of cooperative hunting by groups of three or four lionesses, in Etosha National Park, Namibia [59]. A collective effort of at least three lionesses is required in order for this particular hunting tactic to work, otherwise the prey is able to flee and the lionesses' effort is in vain. Certain situations involving humans also have thresholds. These can literally include hunts such as the whale hunts in Indonesia using traditional methods [60] or the recently discovered hunts of giant sloths [61] (as illustrated in Figure 2.9), which require collective action, but also international relations security dilemmas where a coordinated effort is required to guarantee world peace [62]. Note that these examples don't imply the need of full participation to obtain the best outcome. In fact, in some scenarios returns vary suddenly at certain levels of input, after which additional contributions will not change the outcome [63].

Figure 2.9 – A 3D reconstruction based on recently discovered footprints shows how humans coordinated efforts to hunt a now extinct, much larger animal – the giant sloth. This finding constitutes remarkable evidence that 10,000 years ago humans already engaged in collective action to seek a higher reward than would be expected by hunting alone. Figure reproduced from [64].

Modelling this sort of scenarios calls for a new class of games, where returns vary in a non-linear fashion with respect to the number of cooperators – the class of threshold games. Particularly, a game known as CRD models this type of scenarios [13], and it will serve as the basis for this thesis. Its main characteristic is a non-linear payoff function, unlike in the PGG. More specifically it is a step function with two important parameters – threshold and risk. The threshold represents the minimum amount of cooperators required to obtain a common reward. In case this threshold isn't reached, the risk represents the probability of not obtaining the reward. A risk of 100% implies an expected value of zero if there aren't enough cooperators. Later CRD will be formalized in more detail, introducing the framework to be used for this work.

3. Reward and Punishment

In Games of Cooperation, a review was presented on mechanisms that solve the cooperation dilemma in 2-person interactions, followed by an introduction to the class of N-person games. Indeed, this is not a trivial transition, and most 2-player mechanisms do not fit directly in N-person scenarios. Take for instance the case of reciprocity. If an individual is interacting with a group containing a cooperator and a defector, with whom should he reciprocate? Possible strategies include having a tolerance threshold [65] (i.e. cooperating as long as a minimum number of cooperators exist in the group) or adopting an all-or-none policy [66], where an individual cooperates only if all others cooperate or if no one does. In any case, the idea of reciprocity between two players cannot be directly translated to N-person games. Another example of the increased difficulty in analyzing group games relates to their inner equilibria – as the number of players grows, the number of possible inner equilibria grows linearly with it, or even polynomially if there are more than two strategies [67]. These cases illustrate why explaining the emergence of cooperation in larger groups remains an open problem and is so difficult to analyze.

Reward and Punishment appear as candidates to explain and promote cooperation in these more complex interactions [68]. Typically, each of these mechanisms can appear in two forms – as peer-to-peer or through institutions. Peer-to-peer consists in an individual applying directly a punishment or reward to a peer when he deems suitable, most often incurring in a personal cost for doing it. Usually the individual that decides whether to apply the mechanism will interact with his peer before making a decision, this way suffering directly from his peer's defection or benefiting from his altruism. Peer-punishment is more commonly studied, although peer-reward has also deserved attention [69]. On the other hand, institutions are modelled so that all defectors are punished equally, or all cooperators rewarded, regardless of peer-based decisions. This will require a collective budget before the round begins, which can be taken for granted (e.g. collected in the form of a tax) or not – requiring at least a certain number of individuals to contribute for the institution to exist, before the round takes place. By requiring optional contributions, cooperators become of two types – those who cooperate and contribute for the institution, and those who cooperate but do not contribute. This second strategy raises the problem of second-order free-riding, where cooperators can free-ride on the effort of their peers who contribute. Alternatively to institutions, certain setups, particularly experimental ones, study the possibility of deciding whether to punish/reward after passively observing an interaction as a third party [70]. This is not exactly a peer policy since the individual applying the mechanism hasn't interacted directly with his "peer", but neither an institutional policy in the sense that it is decided case by case, usually with the punisher/rewarder incurring in a personal cost for applying the fine/reward. This setup can test whether people are willing to incur a cost to enforce justice for the benefit of others, although not exploring the deeper reasons for this. [71] offers an interesting discussing on these reasons, namely the idea of "cold glow", which is presented below when reviewing experimental work.

When compared to reciprocity, Reward and Punishment have the advantage of allowing to discriminate actions for each member of the group, instead of addressing the whole group simultaneously by cooperating or refraining from doing so [6]. On the other hand, they have the disadvantage of being costly [72,73]. Much attention has been given to these mechanisms separately, but few work has addressed them combined [21,22]. At first sight reward and punishment may seem mathematically equivalent, since both will work to shorten the gap between the payoff of defectors and cooperators. However, when considering implementation costs these mechanisms cease to be equivalent. This is both due to the possibility of having a different number of cooperators and defectors, and to the possibility of these policies not being equally efficient. This chapter will present an overview of existing work on Reward and Punishment, both on a theoretical and experimental setting.

3.1. Games with Implicit Reward or Punishment

Several games can be designed where the idea of Reward and/or Punishment is present. The following sections will present settings (both theoretical and experimental) where these mechanisms are present one way or another. Different strategies may exist (e.g. allowing for anti-social policies or not [69]) and incentives may be applied in different ways. The most common way to implement incentives consists in having a round of a classical cooperation dilemma, such as the PGG, followed by a round where individuals are rewarded and/or punished according to their previous action [74], be it through their peers [69] or through an institution [14,19]. However, incentives mechanisms can also be present in less obvious ways, which will be clarified now.

Reward may be alternatively represented by a game called “trust game” [69]. In this 2-player game individuals have very different roles: there is a trustor and a trustee. The trustor has an initial endowment and chooses how much to donate to the trustee. If he opts to make no donation, the game is over and only the trustor receives a payoff (his own endowment). Alternatively, if he makes a donation, that amount is multiplied by a factor $r>1$ and then handed to the trustee. As a final step, the trustee may decide to give back part or even all of the donation which was received and multiplied. There is no incentive to do so, meaning that giving back occurs only as a form of reward from the trustee to the trustor, with a cost for the trustee. The size of the initial donation can also be interpreted as the amount of trust, and the retribution as the amount of trustworthiness [70].

Punishment may also exist in a subtler way. It can be implicitly represented in the game of “ultimatum”, another game where two different roles exist: the proposer and the responder. This game actually emerges as a special case of a PD with punishment [75]. Here, the proposer chooses how to split his own initial endowment between himself and the responder. The responder can accept the offer or reject it, in which case no one receives anything. Rejecting has a cost for the responder (higher for more favourable proposals) but will inflict a punishment on the proposer (higher for less favourable proposals). This game can also be expanded for larger groups, with one proposer and several responders [76,77]. In this case a condition to accept the offer must be defined (e.g. majority, unanimity). A simpler version of ultimatum also exists, known as “dictator”, but which excludes any form of punishment. In this game a receiver is obliged to accept any division that a dictator decides to make of his own initial endowment, even if he decides to keep all of it for himself. This can be used experimentally to measure genuine altruism towards strangers, by observing how much is donated to receivers. The value of donations can be used as a benchmark against ultimatum, in that it can measure how donations vary when the threat of rejection (or costly punishment) is introduced in the game.

3.2. Modelling Reward and Punishment

To theoretically study Reward and Punishment, extensive work has been done using EGT frameworks that build on top of well-known linear games such as PGG by adding separately Reward or Punishment and observing which behaviors emerge under different conditions. A comprehensive review of existing games on lattices can be found in [69], comparing different schemes of reward and punishment on top of these homogeneously structured populations. Homogeneous spatial structures alone are known to make cooperation emerge [44], as shown in Figure 2.8. In [69] this result is used as a baseline – for a 5-person PGG on a square lattice, a multiplicative factor r>3.74 is required for a mixed equilibrium of Cooperators and Defectors to emerge. Effective incentive mechanisms will require a lower r or allow for a pure equilibrium without Defectors to emerge. For instance, peer punishment on lattices can create an equilibrium without defectors or cooperators, only cooperating punishers, as long as the cost of punishing is not too high, or the impact of the fine too low (Figure 3.1).

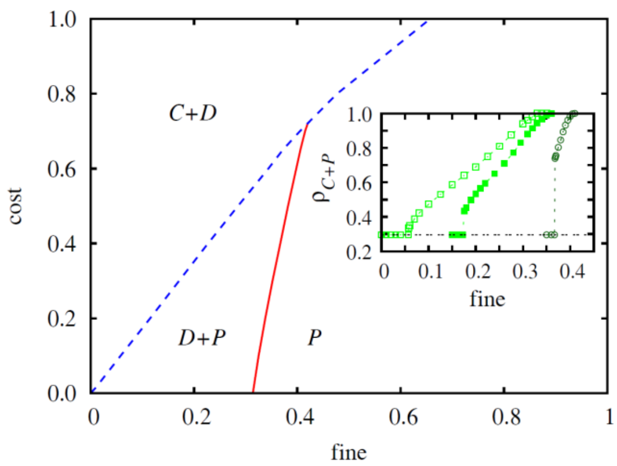

Figure 3.1 – Peer punishment on a square lattice. Mixed or pure equilibria are displayed for different values of cost and fine, obtained for r=3.8. C stands for cooperator, D for defector and P for cooperating punisher. Inset displays overall fraction of cooperation (both cooperators and cooperating punishers) as a function of the impact of fine. Reproduced from [78].

This review also includes results for peer reward, institutionalized reward or punishment and other less usual cases. One example is an adaptive behavior, where cooperators become rewarders or punishers only in low cooperative scenarios, allowing them to save costs when incentives are unnecessary. A clear winner does not emerge among these mechanisms, although there appears to be a trade-off between efficiency (lower cost) in the case of peer strategies and stability in the case of institutions, which is in accordance with [14]. Recall however that all results presented in this review rely on the existence of a homogeneous spatial structure where players interact, known to enhance cooperation by itself [44].

In [14] the mechanisms of peer punishment and institutional (or pool) punishment are theoretically compared using a PGG in which participation is optional. Both policies compete along with defection, cooperation and non-participation. The outcome depends on the possibility of punishing second-order free riders (cooperators who do not punish). Pool punishers win if second-order punishment is included, and peer punishers win otherwise. Simulation results show that peer punishment is more unstable than pool punishment, although less costly. When peer punishment dominates (without second-order punishment), the population is occasionally invaded by non-participants or defectors. On the other hand, when pool punishment dominates (with second-order punishment) these occasional invasions stop occurring. One possible cause for this difference is that second-order free riding is easier to detect with pool punishment, since free-riders are observable even in a population free of defectors. This keeps non-punishers from invading through neutral drift. Interestingly, defectors overtake if we remove the possibility to opt out of the PGG.

Another comparison of peer and institutional punishment can be found in [79], through a model called "gun for hire". Although this theoretical model only studies a classical Nash Equilibrium, it is useful to shed light on the experimental results that the same work presents afterwards – subjects prefer to pay beforehand for an institution than to apply peer-punishment after playing PGG. The PGG used here contains a slight variation: each player can decide how much to contribute, instead of the typical binary option (to contribute or not). The setup where both policies compete consists of a traditional peer-punishment mechanism and a less usual form of pool-punishment: players have the chance to pay for a hired gun before playing the PGG, and this hired gun will punish the lowest contributor, leaving him only slightly worse than the second-lowest one.

Also providing theoretical insight before presenting experimental results, a model studying third-party punishment as costly signaling can be found in [70]. Both Nash equilibria and agent-based simulations show that punishing can be used as a signal of trustworthiness, but individuals prefer to signal through the act of helping when both options are available. The parameters in this model are representative of certain social heuristics (actions that are typically beneficial, which become instinctive due to repetition [80]), making it particularly important to validate these conclusions through an experimental setup. More comments on what makes an experiment relevant are provided in the beginning of the next section.

Finally, theoretical work exists exploring the simultaneous interaction between Reward and Punishment. In [22] a 2-player game is studied, where in each round a helping stage is followed by an incentives stage. Reputations with partial information are also included, to help predict if one's peer will reward (punish) after (not) being helped. Under this setting of peer Reward and Punish, Reward becomes important in early stages of cooperation, especially if the information level on reputations is low. On an institutionalized approach of Reward and Punishment, [21] studies an N-person linear PGG where Reward should be applied first and then switched to Punishment, which acts as a "booster stage". As will be later explained, unlike Threshold Games, PGG does not have internal equilibria that cause cooperation to naturally grow when the population has an intermediate number of cooperators. This way, it is not obvious whether Punishment is also required as a lever for obtaining the critical mass of cooperators, especially relevant when the low fraction of cooperators is not yet sufficient to frequently reap the benefits of cooperation in Threshold dilemmas.

3.3. Incentives in the Laboratory and on the Street

Designing a relevant experiment is not trivial, especially since one may be tempted to simply have a group of humans recreate the conditions that a previous theoretical model has addressed. As argued in [69], this can lead to a common pitfall. Having humans behave in a payoff-maximizing way in the laboratory does not help to reveal hidden mechanisms of cooperation. Whichever behaviour they may exhibit, it may simply be the consequence of rationalizing a given game in order to obtain the highest amount of money possible. On the other hand, if a hidden mechanism exists in the real world, humans may act accordingly (using it unconsciously) even if it isn't payoff-maximizing to do so in a given laboratory setting. This way, it is interesting to instead design games where humans can act according to a proposed mechanism, but doing so will not lead to payoff maximization. The most classical example of this is having humans play one-shot anonymous interactions where they choose to make a donation to a stranger instead of taking the whole reward home, be it through a Dictator game or a round of PD. An excellent example of a more complex scenario can be found in [81]. Although not related with reward or punishment, it ingeniously reveals our hidden propension to cooperate with future generations and argues for our need to be regulated by democratic institutions. A group plays a one-shot game where each player can decide how much money to take from a common pool. If collectively the money taken isn't beyond a given threshold, the pool can continue to exist and another group of unrelated individuals can play this one-shot game. Clearly it is payoff maximizing to extract as much money as allowed, but most individuals choose not to. Instead, it is a small number of individuals in each group that overexploit this common resource and prevent it from continuing to the next group. By changing the rules of the game and introducing a democratic process, where each one votes for or against the overexploitation of the common pool, that pool is successfully preserved for the future generations. This illustrates how an experiment can shed light on these hidden mechanisms known as "social heuristics" [80], which run counter to selfish predictions of classical economics.

Returning to the issue of Reward and Punishment, several experiments have addressed it, many on a peer setup but also some using an institutional setting. Two have already been mentioned in the previous section. Beginning with costly signalling, [70] chooses a setup where one round of "dictator with third party punishment" is followed by one round of trust game. Adding third party punishment to the dictator game means that a third individual, which observes the dictator game, can incur a personal cost to punish the dictator if he observes an unfair interaction. What is original about this experiment is that the trustee is the same person that previously played the role of punisher, and the trustor observes his action during the dictator game, with the punisher/trustee aware that he is being observed. This way, Jordan shows that punishment is used as a costly signal of trustworthiness – trustors trust more trustees who punished, and indeed they are more trustworthy during the trust game. If the option to help the receiver is added, punishers prefer to use help as a costly signalling. In the other mentioned experiment, "gun for hire" [79] by Andreoni, a perhaps less surprising experimental result is displayed – in accordance with the predicted Nash equilibrium, indeed individuals prefer an institutional policy to a peer policy under this particular setting.

Andreoni also provides an example which contemplates both punishment and reward simultaneously [28], tackling a problem conceptually very similar to the one discussed in this thesis but using a radically different setup. This experimental work uses a 2-person game and a mechanism of peer policies, with very different implications from both the N-person threshold game and institutional policy that we will present in Model and Methods. Here cooperation is measured by how much a single individual donates to another (unlike in the setting of this thesis, where cooperation is measured by the number of individuals which binarily choose to cooperate). Such an opposite setup justifies conclusions which are contradictory to the ones presented in this thesis.

Two other relevant experiments argue against peer-policy mechanisms and in favour of legitimate institutions. The first suggests that humans possess a social heuristic called "cold glow" [71]. Analogously to the older concept of "warm glow" [82], in which people have an intrinsic feeling of happiness in helping others, cold glow is proposed as the intrinsic satisfaction in punishing someone for being unfair. Three experiments are successfully designed to unveil this heuristic, with strong implications for mechanism design. One of them tests the lack of crowding-out. Due to cold glow, if two individuals A and B observe an unfair action they are not indifferent between which one applies the punishment. For this reason, if A has already applied an optimal amount of punishment, B is still willing to punish further, moving the global punishment beyond the optimal. Overall, cold glow can cause over punishment or under punishment depending on each situation. This calls for a regulatory institution when applying punishment, instead of letting each person take justice into her own hands. Complementing this idea, and also confirming the importance of a democratic institution (as [81] does), in [83] the effect of random and legitimate authorities is put to test – if an individual is elected based on his previous actions to monitor punishment, cooperation is enhanced further than using a randomly picked individual as a monitor. Still, both cases improve cooperation when compared to a scenario without any central monitoring.

Finally, one particularly original experiment decides to move away from the laboratory and take the concept of peer punishment to the street – particularly to the subway of Athens [84]. The hypothesis behind this study is that laboratory settings do not represent the full cost of peer punishment – for instance, in real life people may risk serious retaliation [85]. Crowded subway stations are chosen to approximate one-shot anonymous interaction, since hundreds of thousands of passengers use it daily. Two actions were repeatedly performed by an actor – to block the left side of an escalator, ignoring requests to step aside, and to throw plastic bottles to the floor. Indeed, the levels of observed punishment were extremely low. One could argue that different cultures (consider Germany or Scandinavian countries for instance) would react differently to this experiment, which could possibly help justify differences between this study and the hypothesis of "cold glow". In any case, this experiment provides evidence that at least in some cultures, peer punishment mechanisms are not expected to work, calling alternatively for reward or institutional policies.

4. Model and Methods

4.1. Public goods and collective risks

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 CRD [13,18,19,86-88]. 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*≤*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≤*r*≤1) stands as the risk of collective failure. Otherwise, everyone will keep whatever he or she has. Such a formulation of a CRD has been shown to capture some of the key features discovered in recent experiments [18,86,89-91], 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, as previously illustrated in section 2.3.3 (Threshold Games), the adoption of non-linear returns mimics situations common to many human and non-human endeavors [55,59,92-98], 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. To define the payoff functions of this dilemma, let $\theta(x)$ be a Heaviside step-function distribution, where $\theta(x) = 0$ if $x < 0$ and $\theta(x) = 1$ if $x \geq 0$. As stated, each player can contribute with a fraction c of his endowment B, (with 0≤c≤1), and in case a group contains less than *M* cooperators (with *M*≤*N*) there is a risk *r* of failure (with 0≤r≤1), in which case no player obtains his remaining endowment. The payoff of a defector ($\Pi_D(j)$) and the payoff of a cooperator ($\Pi_C(j)$) in a group with *j* cooperators, before incorporating any policy, can be written as:

$$\Pi_D(j) = B\{\theta(j - M) + (1 - r)[1 - \theta(j - M)]\} \tag{4.1}$$

$$\Pi_C(j) = \Pi_D(j) - cB \tag{4.2}$$

One possible payoff function of a CRD is shown in Figure 4.1.

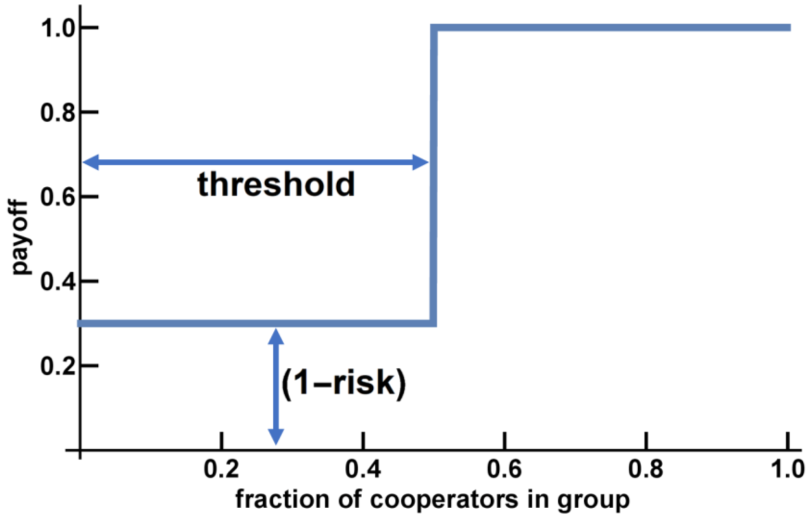

Figure 4.1 – An example of a payoff function in CRD (see Eq. (4.1))

4.2. Including Reward and Punishment

Following [21], we include both reward and punishment mechanisms in this model. A fixed group budget $N \cdot \delta$ (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 (with $0 \leq w \leq 1$). Parameters a and b correspond to the efficiency of Reward and Punishment. For all figures it was assumed that $a=b=1$, except when stated otherwise. Both policies having an efficiency of 1 implies that their effective impact is equivalent, and that each unit spent will directly increase/decrease the payoff of a cooperator/defector by the same amount.

$$\Pi_D^P(j) = \Pi_D(j) - \frac{b(1-w)N\delta}{N-j} \tag{4.3}$$

$$\Pi_C^R(j) = \Pi_C(j) + \frac{awN\delta}{j} \tag{4.4}$$

As Sigmund points out [6], 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 [13]. As discussed below, these will interact in a non-trivial way with positive and negative incentives.

4.3. Evolutionary dynamics in finite and infinite populations

As previously stated, each round of this N-person game occurs by sampling N players from a finite population of size Z. Naturally, parameter Z is dropped when considering infinite populations. To account for processes of sampling with or without replacement (in the case of infinite or finite populations), a binomial or hypergeometric distribution is used, respectively. By including the payoff functions $\Pi_D^P(j)$ and $\Pi_C^R(j)$ (Eqs. (4.3) and (4.4)) in the fitness functions described in (Eqs. (2.10), (2.11), (2.12) and (2.13)), we obtain the fitness values of a cooperator and defector, for a given population state.

Instead of a collection of rational agents engaging in one-shot PGGs [90,99], we adopt an evolutionary description of the behavioral dynamics [13], in which individuals tend to copy others whenever these appear to be more successful (see also Section 2.1). 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 [7,36], 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 [100] (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) [36]. The probability of X imitating Y, depending on their fitness difference, is given by the Fermi function (Eq. (2.6)) for finite populations, or the Replicator equation (Eq. (2.5)) for infinite populations. In the case of finite populations, a mutation rate can be introduced by using the following transition probabilities based on equation (Eq. (2.7)):

$$T_\mu^+(k) = (1-\mu)T^+(k) + \mu\frac{Z-k}{Z} \tag{4.5}$$

$$T_\mu^-(k) = (1-\mu)T^-(k) + \mu\frac{k}{Z} \tag{4.6}$$

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 – $a_G(k)$. This measure of average group success for a given population configuration can be computed in the following way: $a_G(k) = \binom{Z-1}{N-1}^{-1} \sum_{j=M}^{N} \binom{k-1}{j} \binom{Z-k}{N-j-1}$. 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: $\eta_G = \sum_{k=0}^{Z} \bar{p}_k a_G(k)$. This value represents the average fraction of groups that will overcome the CRD, successfully preserving the public good.

5. Results

5.1. Collective Risk Dilemma without incentives

First, a study of CRD without any incentives is presented. As previously stated, the behaviour of infinite populations can be characterized by its frequency-dependent gradient. In line with previous work [13], we show that risk plays a major role in the evolution of the population.

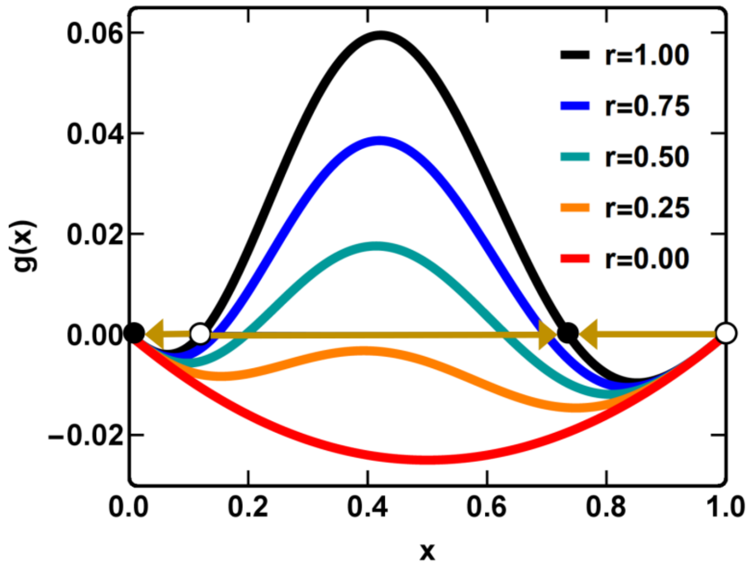

Figure 5.1 – Gradient of Selection of an infinite population playing CRD, for different risk values. Filled circles are co-existence points and empty circles are coordination points. Parameters: Group size N=6, threshold M=3, cost of cooperation c=0.1, initial endowment B=1.

With no risk, the gradient behaves as in a PGG, with no escape from the full defection scenario. However, with a certain minimum value of risk, two additional stationary points appear – a coordination point on the left and a co-existence point on the right. As risk increases, both points move away from each other. This means that the coordination point becomes easier to overcome, and the equilibrium after overcoming the coordination point is a scenario with more cooperators. On the other hand, increasing the threshold moves both points to the right [13]. This causes the co-existence point to be a more favourable equilibrium, but the coordination point becomes harder to overcome. The effect of reward and punishment in these points is not trivial, and naturally cannot be unveiled by resorting to PGG as in [21].

In the case of finite populations, the system's states can be represented by a Markov chain, where each state is a specific count of cooperators. The probabilities to increase or decrease the number of cooperators by one rely on the Fermi update rule (Eq. (2.6)). By plotting the difference

$G(k) = T^{+}(k) - T^{-}(k)$ we obtain the graphic in Figure 5.2, whose meaning is equivalent to the gradient in infinite populations.

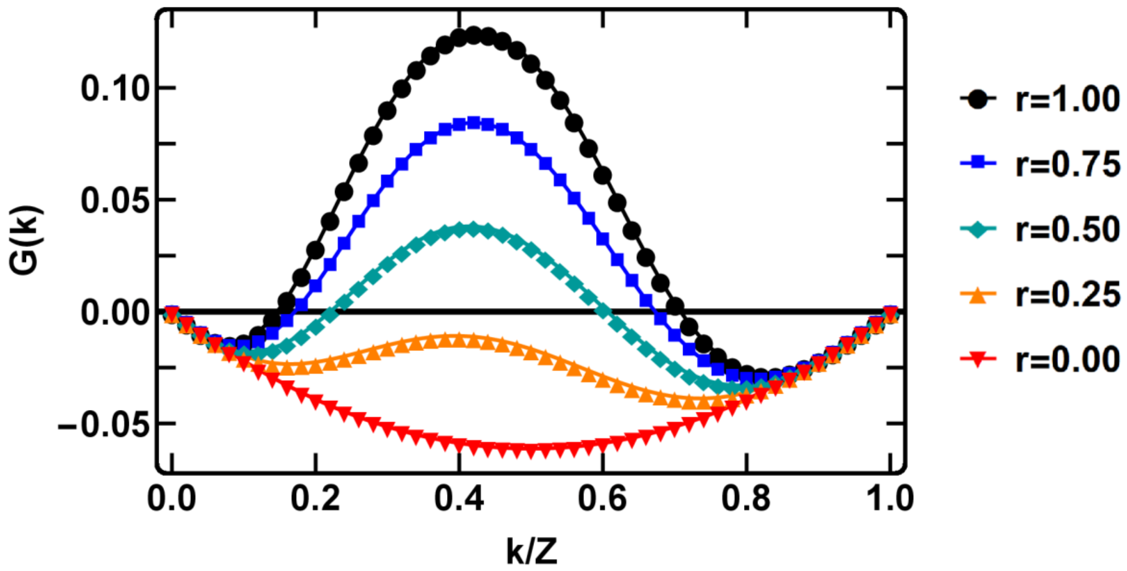

Figure 5.2 – Gradient of selection of a finite population playing CRD, for different risk values. Parameters: population size Z=50, group size N=6, threshold M=3, cost of cooperation c=0.1, initial endowment B=1, intensity of selection=5.

Given this Markov model, it is possible to evaluate cooperation differently than when using infinite populations. One way to do this is through the stationary distribution. Technically, stationary distributions are eigenvectors whose eigenvalues are 1 of the transition matrix S associated with this Markov chain. S is defined by the tridiagonal matrix $S=[p_{ij}]^T$, such that $p_{k,k\pm1}$ is given by $T_{\mu}\pm$(see Eqs. (4.5) and (4.6)), and $p_{k,k}=1- p_{k,k+1} -p_{k,k-1}$. From a more intuitive perspective, a stationary distribution is a vector containing the probabilities for the system to be in each state in any given time. This type of approach can show whether the system will spend more time in cooperative or non-cooperative states.

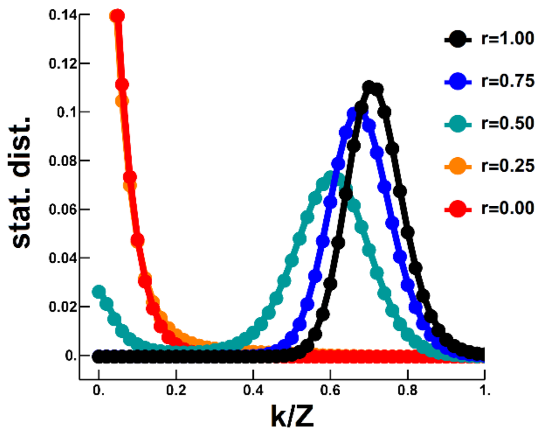

Figure 5.3 – Stationary distributions of CRD for different risk values. As risk increases, cooperative behaviour becomes the norm. Parameters: population size Z=50, group size N=6, threshold M=3, cost of cooperation c=0.1, initial endowment B=1, intensity of selection=5.

As explained in the previous section, an even more succinct metric can be obtained from the stationary distribution – the average group achievement (η_G). This way, it is possible to characterize a whole population through one value, which represents its overall probability of sampling a group which successfully cooperates. This metric will be essential when analysing scenarios with an added incentives scheme.

5.2. Adding Reward and Punishment to the Dilemma

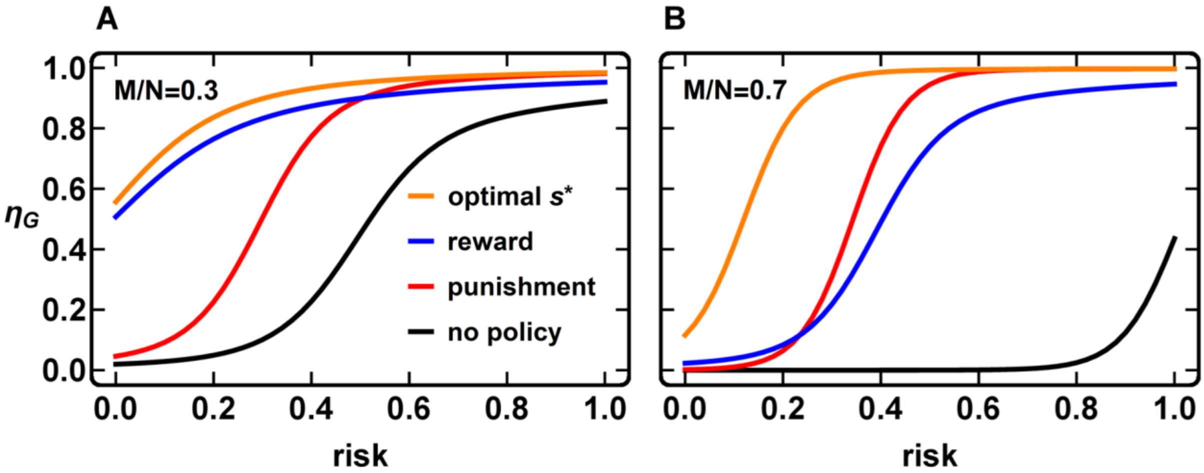

*Figure 5.4 – Average group achievement η_G as a function of risk. **Left**: Group relative threshold M/N=3/10. **Right**: Group relative threshold M/N=7/10. In both panels, the black line corresponds to a reference scenario where no policy is applied. The red line shows η_G in the case where all available 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 [21] (detailed below), 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.*

In Figure 5.4 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 5.4), rewards play the key role, with sanctions outperforming them only for very high values of risk. For high coordination thresholds (right panel, Figure 5.4) 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 5.5 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 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 ($\delta = 0$), 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 [13]. As the budget for incentives increases, using either Reward or Punishment leads to very different outcomes, as depicted in Figure 5.5.

Reward is especially 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 5.5A and 5.5C). 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 only slightly affected, once this barrier is overcome the population will evolve towards a more favourable equilibrium. Thus, while Reward seems to be particularly effective to bootstrap cooperation, Punishment seems effective in sustaining high levels of cooperation. Note that the same effects also occur in the fixed points of infinite populations, as shown in **Appendix**, confirming the impact of these policies in the population's internal equilibria.

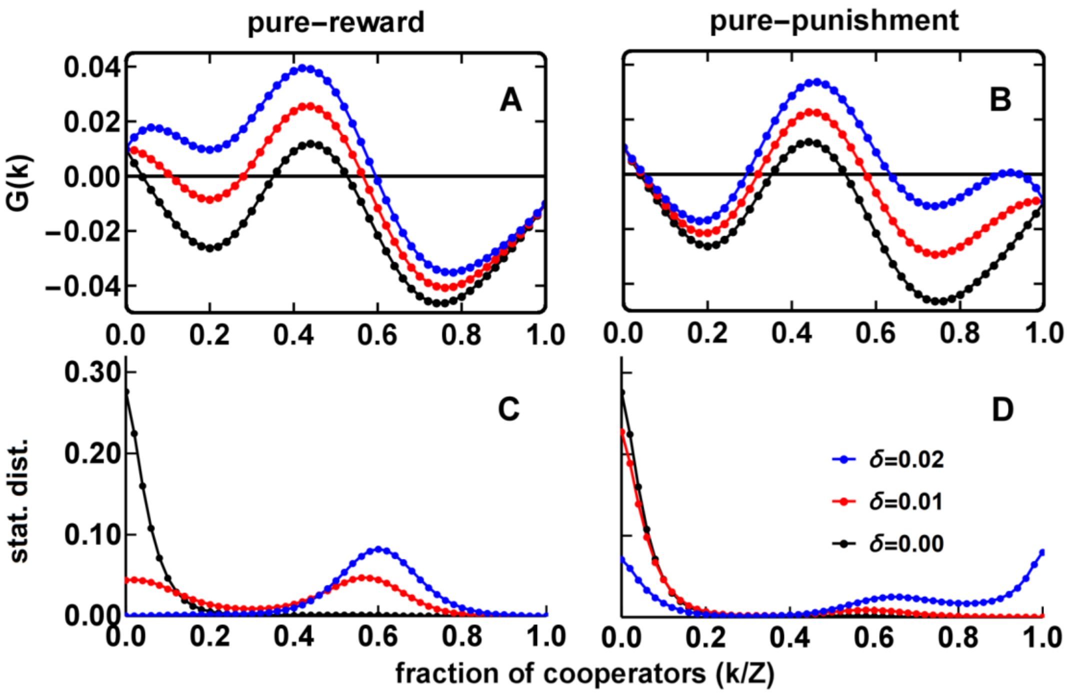

Figure 5.5 – Gradient of selection (top panels, A and B) and stationary distribution (bottom panels, C and D) for the different values of per-capita budget δ indicated, using either pure-Reward (w=1, left panels) or pure-Punishment (w=0, right panels). The black curve is equal on the left and right panels, 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 5.5C and 5.5D), 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 $\delta = 0.02$ (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., $\delta = 0.02$ in Figure 5.5) 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 5.5 show that Reward can be instrumental in fostering pro-social behaviour, 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 ($w = 0$). Similar to PGGs [21], in CRDs this is indeed the policy which minimizes the advantage of the defector, even if we consider the alternative possibility of applying both policies simultaneously. 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 [21], but using finite populations and therefore a hypergeometric distribution (see Eqs. (2.12), (2.13), (4.3), and (4.4)), to account for sampling without replacement. It can be shown that minimizing $F_D - F_C$ is equivalent to maximizing the following expression:

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

where j represents the number of ***C***s in a group of size *N*, sampled without replacement from a population of size *Z* containing k ***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.

By using such policy that we denote by s*, we obtain the best results shown with an orange line in Figure 5.4. 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 5.6 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 each panel in Figure 5.6. Figure 5.6A suggests that, for low thresholds, an optimal policy switching (which under these conditions occurs for $s = 50\%$, as shown in **Appendix**, Figure 7.3) is only marginally better than a policy solely based on rewards ($s = 1$). Figure 5.6A 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*, Figure 5.6B) are combined with a low perception of risk, the adoption of a combined policy becomes necessary (see also right panel of Figure 5.4).

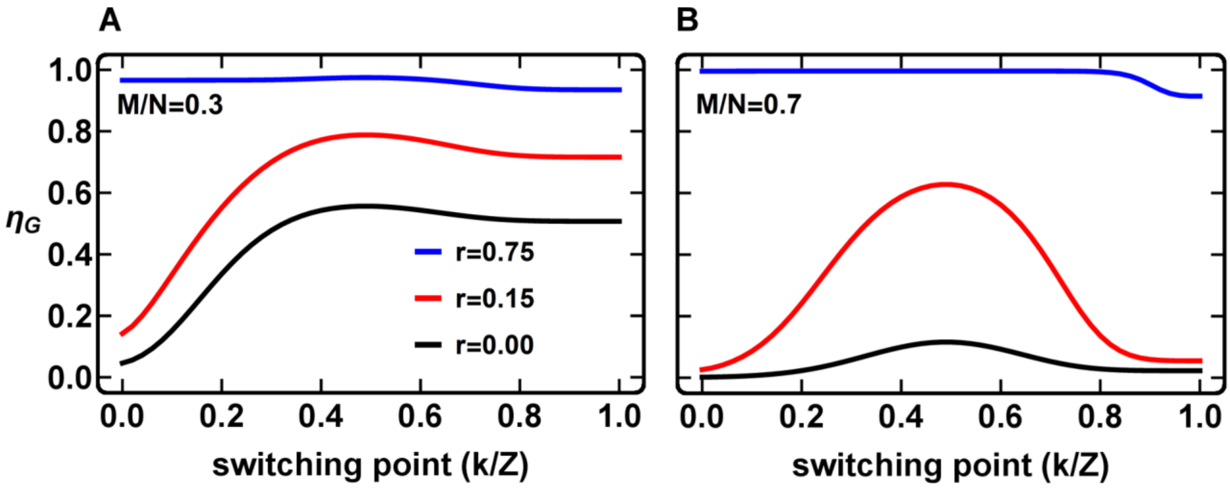

Figure 5.6 – Average group achievement η_G as a function of the switching point. ***Left****: Group relative threshold M/N=3/10.* ***Right****: Group relative threshold M/N=7/10. The switching point corresponds to the configuration of the population above which w suddenly switches from 1 to 0, changing 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 of each panel. Assuming both policies are equally efficient, the optimal switching point occurs at 50% of cooperators. In both panels, 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 of each figure (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 a low threshold (left panel), under low risk it is significantly less harmful to have a late switch from Reward to Punishment than an early one. This implies that when the population configuration cannot be precisely observed, it is preferable to keep rewarding for longer. On the other hand, with a high threshold (right panel) the synergy of both policies becomes crucial. Other parameters: Population size Z=50, initial endowment B=1, mutation rate μ=0.01, intensity of selection=5, group size N=10, cost of cooperation c=0.1, budget δ=0.025.*

6. Conclusion

In this work, we proposed to study the impact of Reward and Punishment in Climate Change Dilemmas. The tools of Evolutionary Game Theory were described, along with existing work on Reward and Punishment, both on a theoretical and experimental level. The game of CRD was then introduced, along with an analysis of its overall behaviour before adding any incentives. A minimum number of contributors is required before collective success is guaranteed, and risk plays a major role in the outcome of this game. Using this framework as a metaphor, a broad class of real-world scenarios including Climate Change Dilemmas can be represented. By adding Reward and Punishment mechanisms on top of it, their effects can be studied under different conditions, such as various levels of risk.

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 dynamic combination of the single policies, Reward and Punishment, in the sense that the relative weight of each policy 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.

It is interesting to note that the optimal policy uses only Punishment or only Reward, depending on the configuration of the population in each moment. Even though Punishment and Reward are allowed to be used simultaneously in one game round, the mathematical result of the optimal policy never contemplates such a scenario. Relatedly, recent experimental work [101,102] has provided evidence against the strong reciprocity hypothesis [103]. According to this hypothesis positive and negative reciprocity are inseparable, implying that reward and punish should be applied together, as pointed out in [69]. The result of this optimal policy, in line with [21], can be seen as further evidence against strong reciprocity, since it never uses Reward and Punishment simultaneously.

Regarding the way these incentives are built, it may seem unfamiliar to have a situation where the size of individual incentive varies according to the number of individuals being targeted. However, it is worth to reflect on this exotic incentives mechanism. Simply because it is not commonly seen in practice does not imply it would not work. Consider start-ups that pay holidays to their excelling employees after they helped secure a new client, or big firms which attract top talent by providing perks in the form of expensive courses or payed trips to conferences. Still, since as more "cooperators" emerge the per-capita reward decreases, it would be interesting to considering the possibility of an anti-social strategy to emerge - where an individual would cooperate but convince his peers not to, or undermine their efforts, in order to receive a larger reward.

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 [16,99]. 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 CRDs. 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 [16,99].

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 [19,87]. 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 [88,91]. Finally, several other effects may further influence and/or affect the present results. Among others, if intermediate tasks are considered [91], or if individuals have the opportunity to pledge their contribution before their actual action [97,104,105], it is likely that pro-social behavior may be enhanced. Our model is minimalistic in the sense that it leaves out these important issues, but it constitutes a convenient theoretical baseline to test more complex (or alternative) incentive mechanisms. Work along these lines is in progress.

7. Appendix

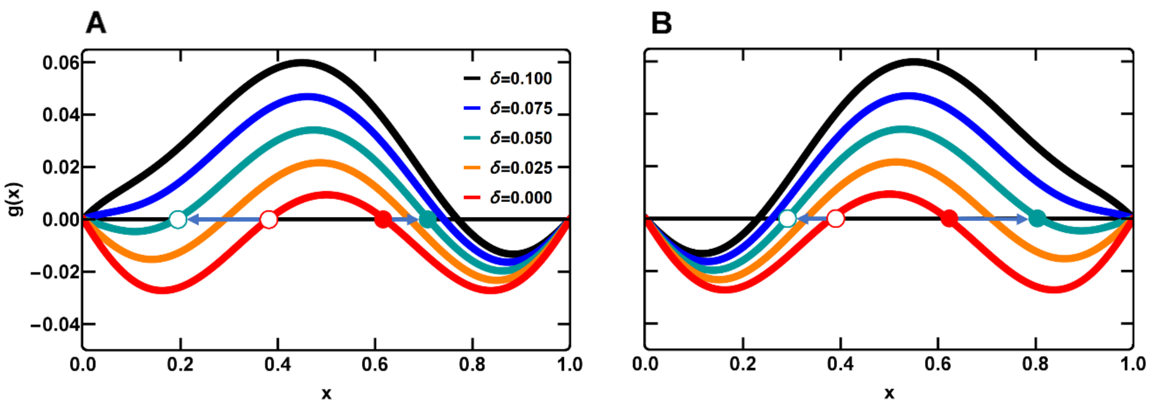

Figure 7.1 – Effects of Reward and Punishment in the internal equilibria of an infinite population's gradient. ***Left****: pure Reward (w=1).* ***Right****: pure Punishment (w=0). Each curve represents one value of budget δ. Filled circles are co-existence points and empty circles are coordination points. Parameters: Group size N=6, threshold M=3, cost of cooperation c=0.3, initial endowment B=1, risk=0.9.*

Deriving the optimal policy

Let us consider the problem of finding the optimal weight *w*, with respect to *k*, that minimizes the defector's advantage $(F_D - F_C)$, formulated following [21]. Following Eqs. (2.12), (2.13), (4.3) and (4.4) the fitness functions in CRD with Reward and Punishment are as follows:

$$F_D = \sum_{j=0}^{N-1} \frac{\binom{Z}{j}\binom{Z-1-k}{N-1-j}}{\binom{Z-1}{N-1}} (\Pi_D(j) - \frac{b(1-w)N\delta}{N-j}) \tag{7.1}$$

$$F_C = \sum_{j=0}^{N-1} \frac{\binom{k-1}{j}\binom{Z-1-(k-1)}{N-1-j}}{\binom{Z-1}{N-1}} (\Pi_C(j+1) - c + \frac{awN\delta}{j+1}) \tag{7.2}$$

Since the impact of cost *-c* and payoffs $\Pi_D(j)$ and $\Pi_C(j+1)$ on $F_D - F_C$ does not depend on *w*, they won't affect the choice of the optimal *w* and can be dropped, leaving us with the problem of minimizing the following:

$$F' = -N\delta \sum_{j=0}^{n-1} \frac{\binom{k}{j}\binom{Z-1-k}{N-1-j}}{\binom{Z-1}{N-1}} \left(\frac{b(1-w)}{N-j}\right) - N\delta \sum_{j=0}^{N-1} \frac{\binom{k-1}{j}\binom{Z-k}{N-1-j}}{\binom{Z-1}{N-1}} \left(\frac{aw}{j+1}\right)$$

Let us consider the following auxiliary steps: $\binom{k}{j} = \binom{k-1}{j}\frac{k}{k-j}$; $\binom{Z-1-k}{N-1-j} = \binom{Z-k}{N-1-j}\frac{Z-k-(N-1-j)}{Z-k}$

$$F' = -N\delta \sum_{j=0}^{N-1} \frac{\binom{k-1}{j}\binom{Z-k}{N-1-j}}{\binom{Z-1}{N-1}} \left(\frac{aw}{j+1} + \frac{b(1-w)}{N-j}\frac{k}{k-j}\frac{Z-k-N+1+j}{Z-k}\right) =$$

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

$$- N\delta \sum_{j=0}^{N-1} \frac{\binom{k-1}{j}\binom{Z-k}{N-1-j}}{\binom{Z-1}{N-1}} \frac{b}{N-j}\frac{k}{k-j}\frac{Z-k-N+1+j}{Z-k}$$

The impact of the second summation on $F_D - F_C$ does not depend on *w* either, therefore it can be dropped as well:

$$F'' = -N\delta * \sum_{j=0}^{N-1} \frac{\binom{k-1}{j}\binom{Z-k}{N-1-j}}{\binom{Z-1}{N-1}} \left(w \left[\frac{a}{j+1} - \frac{b}{N-j}\frac{k}{k-j}\frac{Z-k-N+1+j}{Z-k} \right] \right)$$

Since *N* and δ are always positive, the whole expression can be divided by *N*δ without changing the optimization problem (equivalent to dropping their first occurrences). Moreover, by multiplying the expression by (-1), the problem becomes to maximize the following:

$$w \sum_{j=0}^{N-1} \frac{\binom{k-1}{j}\binom{Z-k}{N-1-j}}{\binom{Z-1}{N-1}} \left(\left[\frac{a}{j+1} - \frac{b}{N-j}\frac{k}{k-j}\frac{Z-k-N+1+j}{Z-k} \right] \right) \tag{7.3}$$

It can be shown that this sum decreases as *k* increases, containing only one root (an example can be found in Figure 7.2). Therefore, 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 value of *k* causes the sum to be negative. We call the root of this expression the optimal switching point s^*.

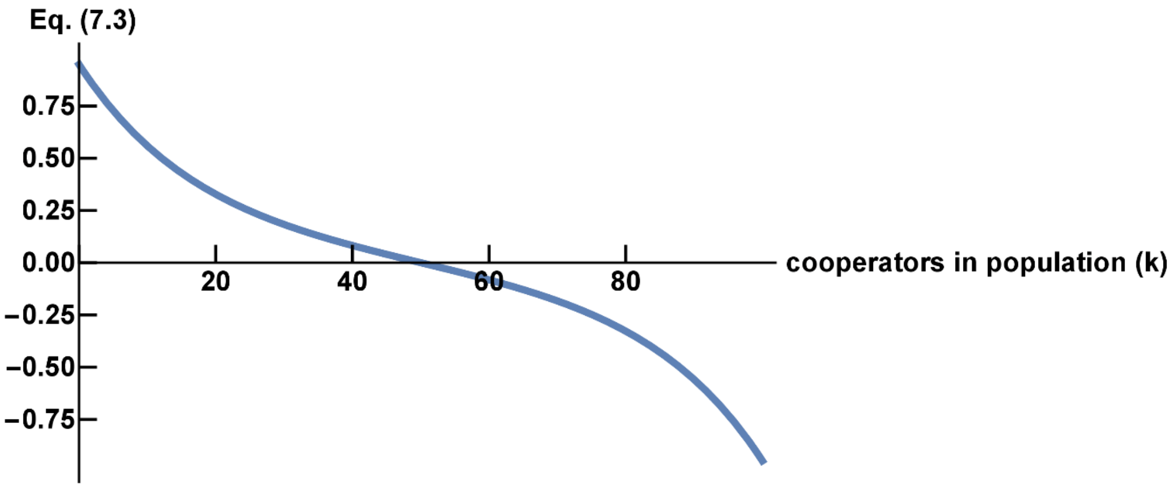

Figure 7.2 – The optimization problem of minimizing the defector's advantage. Eq (7.3) is displayed for different values of k. The remaining parameters are fixed as follows: Z=100, N=10, a=1, b=1

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) on *s** is illustrated in the following figure:

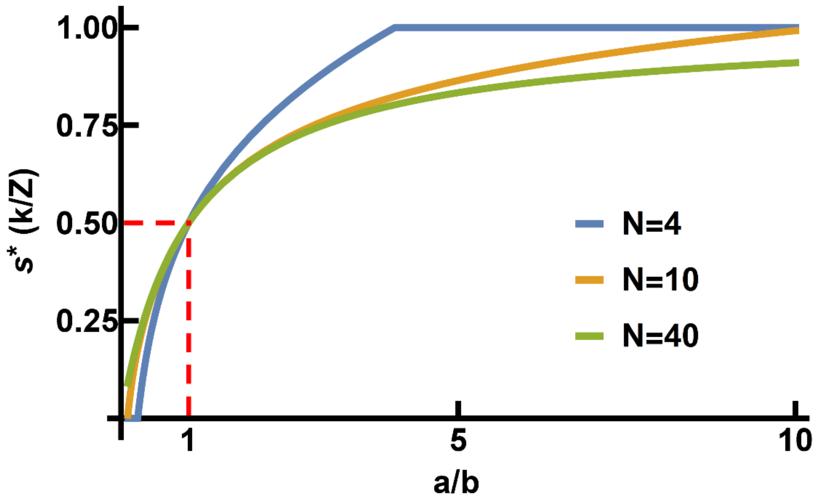

Figure 7.3 – Optimal switching point s as a function of the ratio a/b, for different values of N. s* corresponds to the optimal instant when there is a sudden switch between Reward and Punishment, given by the population configuration (k/Z) where it should occur. For all values of N displayed, the optimal switch point occurs at 0.5 if a=b. Population size Z=100.*

8. References

1. Bernoulli D (1954) Exposition of a New Theory on the Measurement of Risk. Econometrica 22: 23-36.
2. Von Neumann J, Morgenstern O (1944) Theory of games and economic behavior. Princeton, NJ, US: Princeton University Press. xviii, 625-xviii, 625 p.
3. Sen A (1990) Rational Behaviour. In: Eatwell J, Milgate M, Newman P, editors. Utility and Probability. London: Palgrave Macmillan UK. pp. 198-216.
4. Sen A (1977) Rational Fools: A Critique of the Behavioral Foundations of Economic Theory. Philosophy and Public Affairs 6: 317-344.
5. Rand DG, Nowak MA (2013) Human cooperation. Trends in Cognitive Sciences 17: 413-425.
6. Sigmund K (2007) Punish or perish? Retaliation and collaboration among humans. Trends in ecology & evolution 22: 593-600.
7. Sigmund K (2010) The Calculus of Selfishness: Princeton Univ Press.
8. Nowak MA (2006) Evolutionary Dynamics: Belknap/Harvard.
9. Weibull J (1997) Evolutionary Game Theory: The MIT Press.
10. Sigmund K, Nowak MA (1999) Evolutionary game theory. Current Biology 9: R503-R505.
11. Hofbauer J, Sigmund K (1998) Evolutionary Games and Population Dynamics. Cambridge, UK: Cambridge Univ. Press.
12. Dingli D, Chalub F, Santos F, Van Segbroeck S, Pacheco J (2009) Cancer phenotype as the outcome of an evolutionary game between normal and malignant cells. Br J Cancer 101: 1130-1136.
13. Santos FC, Pacheco JM (2011) Risk of collective failure provides an escape from the tragedy of the commons. Proc Natl Acad Sci USA 108: 10421-10425.
14. Sigmund K, De Silva H, Traulsen A, Hauert C (2010) Social learning promotes institutions for governing the commons. Nature 466: 861.
15. Barrett S (1994) Self-enforcing international environmental agreements. Oxford Economic Papers 46: 878-894.
16. Barrett S (2007) Why cooperate?: the incentive to supply global public goods: Oxford UP.
17. Dreber A, Nowak MA (2008) Gambling for Global Goods. Proc Natl Acad Sci USA 105: 2261-2262.
18. Milinski M, Sommerfeld RD, Krambeck HJ, Reed FA, Marotzke J (2008) The collective-risk social dilemma and the prevention of simulated dangerous climate change. Proc Natl Acad Sci USA 105: 2291–2294.
19. Vasconcelos VV, Santos FC, Pacheco JM (2013) A bottom-up institutional approach to cooperative governance of risky commons. Nat Clim Change 3: 797-801.
20. Sigmund K, Hauert C, Nowak MA (2001) Reward and punishment. Proc Natl Acad Sci U S A 98: 10757-10762.
21. Chen X, Sasaki T, Brännström Å, Dieckmann U (2015) First carrot, then stick: how the adaptive hybridization of incentives promotes cooperation. Journal of The Royal Society Interface 12: 20140935.
22. Hilbe C, Sigmund K (2010) Incentives and opportunism: from the carrot to the stick. Proceedings of the Royal Society of London B: Biological Sciences 277: 2427-2433.
23. Gneezy A, Fessler DM (2011) Conflict, sticks and carrots: war increases prosocial punishments and rewards. Proceedings of the Royal Society of London B: Biological Sciences: rspb20110805.
24. Sasaki T, Uchida S (2014) Rewards and the evolution of cooperation in public good games. Biology letters 10: 20130903.
25. Fehr E, Gächter S (2002) Altruistic punishment in humans. Nature 415: 137-140.

26. Masclet D, Noussair C, Tucker S, Villeval M-C (2003) Monetary and nonmonetary punishment in the voluntary contributions mechanism. American Economic Review 93: 366-380.

27. Charness G, Haruvy E (2002) Altruism, equity, and reciprocity in a gift-exchange experiment: an encompassing approach. Games and Economic Behavior 40: 203-231.

28. Andreoni J, Harbaugh W, Vesterlund L (2003) The carrot or the stick: Rewards, punishments, and cooperation. The American economic review 93: 893-902.

29. Nash JF (1950) Equilibrium points in n-person games. Proceedings of the National Academy of Sciences 36: 48.

30. Santos FC, Pacheco JM, Lenaerts T (2006) Evolutionary dynamics of social dilemmas in structured heterogeneous populations. Proc Natl Acad Sci USA 103: 3490-3494.

31. Rendell L, Boyd R, Cownden D, Enquist M, Eriksson K, et al. (2010) Why Copy Others? Insights from the Social Learning Strategies Tournament. Science 328: 208-213.

32. Mitchell M (2009) Complexity: A guided tour: Oxford UP.

33. Paul H. Harvey PJGJMSMS, Smith JM, Greenwood PJ, Harvey PH, Slatkin M (1985) Evolution: Essays in Honour of John Maynard Smith: Cambridge University Press.

34. Maynard-Smith J (1982) Evolution and the Theory of Games. Cambridge: Cambridge University Press.

35. Taylor PD, Jonker LB (1978) Evolutionary stable strategies and game dynamics. Mathematical biosciences 40: 145-156.

36. Traulsen A, Nowak MA, Pacheco JM (2006) Stochastic dynamics of invasion and fixation. Phys Rev E 74: 011909.

37. Trivers RL (1971) The Evolution of Reciprocal Altruism. Quarterly Review of Biology 46: 35-57.

38. Axelrod R (1984) The Evolution of Cooperation. New York: Basic Books. 256 p.

39. Nowak MA, Sigmund K (1998) Evolution of indirect reciprocity by image scoring. Nature 393: 573-577.

40. Nowak MA, Sigmund K (2005) Evolution of indirect reciprocity. Nature 437: 1291-1298.

41. Ohtsuki H, Iwasa Y (2006) The leading eight: social norms that can maintain cooperation by indirect reciprocity. Journal of Theoretical Biology 239: 435-444.

42. Pacheco JM, Santos FC, Chalub FA (2006) Stern-judging: A simple, successful norm which promotes cooperation under indirect reciprocity. PLoS Comput Biol 2: e178.

43. Santos FP, Santos FC, Pacheco JM (2016) Social norms of cooperation in small-scale societies. PLoS Comput Biol 12: e1004709.

44. Nowak MA, May RM (1992) Evolutionary Games and Spatial Chaos. Nature 359: 826-829.

45. Barabási AL, Albert R (1999) Emergence of scaling in random networks. Science 286: 509-512.

46. Santos FC, Santos MD, Pacheco JM (2008) Social diversity promotes the emergence of cooperation in public goods games. Nature 454: 213-216.

47. Santos FC, Pacheco JM (2005) Scale-free networks provide a unifying framework for the emergence of cooperation. Phys Rev Lett 95: 098104.

48. Santos FC, Pinheiro FL, Lenaerts T, Pacheco JM (2012) The role of diversity in the evolution of cooperation. J Theor Biol 299: 88-96.

49. Wilson DS (1975) A theory of group selection. Proceedings of the National Academy of Sciences 72: 143-146.

50. Boyd R, Richerson PJ (1990) Group selection among alternative evolutionarily stable strategies. Journal of Theoretical Biology 145: 331-342.

51. Traulsen A, Nowak MA (2006) Evolution of cooperation by multilevel selection. Proc Natl Acad Sci U S A 103: 10952-10955.

52. Hamilton WD (1964) The genetical evolution of social behaviour (I and II). . Journal of Theoretical Biology 7: 1-16 and 17-52.

53. Dawkins R (1989) The Selfish Gene: Oxford University Press.

54. Santos FP, Santos FC, Melo F, Paiva A, Pacheco JM. Learning to be fair in multiplayer ultimatum games; 2016. IFAAMS. pp. 1381-1382.

55. Pacheco JM, Santos FC, Souza MO, Skyrms B (2009) Evolutionary dynamics of collective action in N-person stag hunt dilemmas. Proc R Soc Lond B 276: 315-321.

56. Scheffer M (2009) Critical transitions in nature and society: Princeton University Press.

57. A. KK, Rod F, S. HB, V. KC, Larry C, et al. (2015) Thresholds in Caribbean coral reefs: implications for ecosystem-based fishery management. Journal of Applied Ecology 52: 402-412.

58. Meyer L, Brinkman S, van Kesteren L, Leprince-Ringuet N, van Boxmeer F (2014) IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of theIntergovernmental Panel on Climate Change. Geneva, Switzerland. 3-87 p.

59. Stander PE (1992) Cooperative hunting in lions: the role of the individual. Behavioral ecology and sociobiology 29: 445-454.

60. Bedding B (2007) The stone-age whale hunters who kill with their bare hands.

61. Bustos D, Jakeway J, Urban TM, Holliday VT, Fenerty B, et al. (2018) Footprints preserve terminal Pleistocene hunt? Human-sloth interactions in North America. Science Advances 4.

62. Jervis R (1978) Cooperation Under the Security Dilemma World Politics 30: 167-214.

63. Hardin R (1982) Collective Action: RFF Press.

64. McClelland A (2018) How to hunt a giant sloth – according to ancient human footprints.

65. Van Segbroeck S, Pacheco JM, Lenaerts T, Santos FC (2012) Emergence of Fairness in Repeated Group Interactions. Phys Rev Lett 108: 158104.

66. Pinheiro FL, Vasconcelos VV, Santos FC, Pacheco JM (2014) Evolution of All-or-None strategies in repeated public goods dilemmas. PLoS Comput Biol 10: e1003945.

67. Gokhale CS, Traulsen A (2010) Evolutionary games in the multiverse. Proceedings of the National Academy of Sciences 107: 5500-5504.

68. Ostrom E (1990) Governing the commons: The evolution of instituitions for collective action: Cambridge Univ. Press.

69. Perc M, Jordan JJ, Rand DG, Wang Z, Boccaletti S, et al. (2017) Statistical physics of human cooperation. Physics Reports 687: 1-51.

70. Jordan JJ, Hoffman M, Bloom P, Rand DG (2016) Third-party punishment as a costly signal of trustworthiness. Nature 530: 473.

71. Ouss A, Peysakhovich A (2015) When punishment doesn’t pay: Cold glow and decisions to punish. The Journal of Law and Economics 58: 625-655.

72. Dreber A, Rand DG, Fudenberg D, Nowak MA (2008) Winners don't punish. Nature 452: 348-351.

73. Ohtsuki H, Iwasa Y, Nowak MA (2009) Indirect reciprocity provides only a narrow margin of efficiency for costly punishment. Nature 457: 79-82.

74. Fehr E, Gachter S (2002) Altruistic punishment in humans. Nature 415: 137-140.

75. Sigmund K, Hauert C, Nowak MA (2001) Reward and punishment. Proceedings of the National Academy of Sciences of the United States of America 98: 10757-10762.

76. Santos FP, Pacheco JM, Paiva A, Santos FC (2017) Structural power and the evolution of collective fairness in social networks. PLoS One in press.

77. Santos FP, Santos FC, Paiva A, Pacheco JM (2015) Evolutionary dynamics of group fairness. J Theor Biol 378: 96-102.

78. Dirk H, Attila S, Matjaž P, György S (2010) Punish, but not too hard: how costly punishment spreads in the spatial public goods game. New Journal of Physics 12: 083005.

79. Andreoni J, Gee LK (2012) Gun for hire: delegated enforcement and peer punishment in public goods provision. Journal of Public Economics 96: 1036-1046.

80. Jordan J, Peysakhovich A, Rand DG (2015) Why We Cooperate. The moral brain: A multidisciplinary perspective: 87.

81. Hauser OP, Rand DG, Peysakhovich A, Nowak MA (2014) Cooperating with the future. Nature 511: 220.

82. Andreoni J (1990) Impure altruism and donations to public goods: A theory of warm-glow giving. The economic journal 100: 464-477.

83. Baldassarri D, Grossman G (2011) Centralized sanctioning and legitimate authority promote cooperation in humans. Proceedings of the National Academy of Sciences 108: 11023-11027.

84. Balafoutas L, Nikiforakis N (2012) Norm enforcement in the city: a natural field experiment. European Economic Review 56: 1773-1785.

85. Edwards R (2008) Police officers beaten by mob after asking girl to pick up litter. https://www.telegraph.co.uk/news/uknews/2433514/Police-officers-beaten-by-mob-after-asking-girl-to-pick-up-litter.html.

86. Milinski M, Semmann D, Krambeck HJ, Marotzke J (2006) Stabilizing the Earth's climate is not a losing game: Supporting evidence from public goods experiments. Proc Natl Acad Sci USA 103: 3994-3998.

87. Pacheco JM, Vasconcelos VV, Santos FC (2014) Climate change governance, cooperation and self-organization. Phys Life Rev 11: 595-597.

88. Vasconcelos VV, Santos FC, Pacheco JM, Levin SA (2014) Climate policies under wealth inequality. Proc Natl Acad Sci U S A 111: 2212-2216.

89. Barrett S (2011) Avoiding disastrous climate change is possible but not inevitable. Proc Natl Acad Sci USA 108: 11733.

90. Barrett S, Dannenberg A (2012) Climate negotiations under scientific uncertainty. Proc Natl Acad Sci U S A 109: 17372-17376.

91. Milinski M, Röhl T, Marotzke J (2011) Cooperative interaction of rich and poor can be catalyzed by intermediate climate targets. Climatic change: 1-8.

92. Boesch C (2002) Cooperative hunting roles among Tai chimpanzees. Human Nature 13: 27-46.

93. Creel S, Creel NM (1995) Communal hunting and pack size in African wild dogs, Lycaon pictus. Animal Behaviour 50: 1325-1339.

94. Black J, Levi MD, De Meza D (1993) Creating a good atmosphere: minimum participation for tackling the'greenhouse effect'. Economica: 281-293.

95. Alvard MS, Nolin DA, Brosius P, Kuznar LA, Rousseau J, et al. (2002) Rousseau's whale hunt? Coordination among big-game hunters. Current anthropology 43: 533-559.

96. Souza MO, Pacheco JM, Santos FC (2009) Evolution of cooperation under N-person snowdrift games. J Theor Biol 260: 581-588.

97. Pacheco JM, Vasconcelos VV, Santos FC, Skyrms B (2015) Co-evolutionary dynamics of collective action with signaling for a quorum. PLoS Comput Biol 11: e1004101.

98. Skyrms B (2004) The Stag Hunt and the Evolution of Social Structure: Cambridge Univ Press.

99. Barrett S (2005) Environment and statecraft: the strategy of environmental treaty-making: Oxford UP.

100. Traulsen A, Hauert C, De Silva H, Nowak MA, Sigmund K (2009) Exploration dynamics in evolutionary games. PNAS 106: 709-712.

101. Yamagishi T, Horita Y, Mifune N, Hashimoto H, Li Y, et al. (2012) Rejection of unfair offers in the ultimatum game is no evidence of strong reciprocity. Proceedings of the National Academy of Sciences 109: 20364.

102. Egloff B, Richter D, Schmukle SC (2013) Need for conclusive evidence that positive and negative reciprocity are unrelated. Proceedings of the National Academy of Sciences 110: E786-E786.

103. Gintis H (2000) Strong reciprocity and human sociality. Journal of theoretical biology 206: 169-179.

104. Tavoni A, Dannenberg A, Kallis G, Löschel A (2011) Inequality, communication and the avoidance of disastrous climate change in a public goods game. Proc Natl Acad Sci USA 108: 11825-11829.

105. Santos FC, Pacheco JM, Skyrms B (2011) Co-evolution of pre-play signaling and cooperation. J Theor Biol 274: 30-35.