Simulation of Dynamic Behavior in an Extended Public Goods Game

Per Moosavi* 21970

Petter Nilsson† 22009

659 Degree Project in Economics
Department of Economics, Stockholm School of Economics

Supervisor: Prof. Jörgen W. Weibull
Examinator: Dr. Maria C. Perrotta
Discussants: Joel Wållgren and Yi Zhang

June 4, 2013

Abstract

Establishment and upholding of norms in a society is studied by constructing an agent-based model for the public goods game with punishments and limited information. The model includes players employing one of three strategies: defection, enforcement or cooperation. Game dynamics are introduced based on evolutionary game theory with poor performing players being more likely to revise their strategy.

The model is studied numerically with the use of computer simulations. Questions addressed include the necessary conditions to attain a cooperative state and the impact of including cost-sharing institutional structures. An analytical study is also included, leading to the derivation of an approximate differential equation describing the dynamics of the game.

It was found that for sufficiently severe punishments, enforcers are able to establish and sustain states of cooperation. This ability was observed to depend on the mobility of strategies and could be improved by including institutional structures.

*Electronic address: 21970@student.hhs.se
†Electronic address: 22009@student.hhs.se

JEL Classification: C73, Z13

Contents

1 Introduction **1**

2 Public Goods Game **5**
2.1 Formulation 5
2.2 Variations 7
2.2.1 Local versus Global Considerations 7
2.2.2 Institutions 8
2.2.3 Norm Dependence 8
2.3 Common Property Resources 9

3 Model **11**
3.1 Static Component 11
3.2 Game Dynamics 14
3.3 Model Summary 14
3.4 Theoretical Study 15
3.5 Variations 18
3.5.1 Institutional Structure 18
3.5.2 Norm Dependent Strategies 18

4 Experiments **21**
4.1 Random Initial Distributions 21
4.2 Stability Analysis 21
4.3 Variations 22

5 Results **23**
5.1 Random Initial Distributions 23
5.1.1 Specific Examples 23
5.1.2 Effectiveness of Punishment 27
5.1.3 Local-Global Imitation Delimiter 29
5.2 Stability Analysis 31
5.3 Variations 33
5.3.1 Institutional Structure 33
5.3.2 Norm Dependent Strategies 34

6 Conclusions **35**
6.1 Model Mechanisms 35
6.2 Theoretical versus Experimental Results 37
6.3 Effects of an Institutional Structure 38
6.4 Further Extensions 38
6.5 Closing Remarks 39

Appendices **41**

A Analytical Treatment **41**

B Model Implementation **45**
B.1 Model Complexity and Oscillations 45
B.2 Ergodicity and Randomness 45

References **48**

1 Introduction

Modern research in economics often looks beyond the classical assumption that individuals seek to maximize a utility which only depends on personal consumption. Indeed, evidence for non-selfish behavior is seen every day when humans interact in societies, and similar observations have also been made in laboratory experiments. Several additional factors affecting personal utility have been identified. Examples include, among others, altruism and desires for equality and fairness. It has been observed that these preferences also vary across cultures, but that an explanation based solely on the selfish assumption normally lacks support [10]. These cross-cultural differences are often explained by the social norms prevalent in the societies [13]. How these social norms change and impact a society has an important connection with public policy decisions. The objective of the latter concerns the study of cooperation among the citizens in a society. Formulated as a question, the discussion essentially seeks to answer: how should taxes, benefits and institutions be designed in order to favor cooperation and coordination between citizens?

A remedy for classical economics is to modify the utility function associated with each agent. Instead of having a utility function depending only on personal gains and losses, other variables can be included to model additional preferences. Empirical research is required to understand the effects of these social preferences on personal utilities and how they vary across different societies. Empirically, these issues can be investigated by letting humans participate in laboratory experiments, where they interact and make economic decisions. These experiments often consist of simple games specifically designed to compare different utility factors. One example is the dictator game, which is a two-player game where one player (the dictator) is offered a certain amount of credits and given the option to donate a share to the other player. It turns out that, contrary to the classical assumption of rational behavior, many dictators end up giving substantial shares to the other player. This yields support for the notion that the gains and losses of other individuals indeed matter in economic decisions. A large number of such experiments have been undertaken, for different games, different stakes and in various places. Changing outcomes have also been observed for example when participants are allowed to communicate, which indicates that there is also a psychological factor present [2].

The family of games discussed above can also be studied theoretically within the framework of game theory, which is a well-established field of mathematics. Such tools can be used to predict the outcome of games and the interaction between players. One of the most well-known theoretical concepts is that of *Nash equilibria*, which are strategy profiles in games where no single player can increase her personal gains by a unilateral change of strategy. In multi-stage games a refined class of

equilibria is often considered, so called *subgame perfect equilibria*. For a strategy profile to be subgame perfect, its restriction to any subgame must represent a Nash equilibrium also for that particular subgame. It has been proven that Nash equilibria and subgame perfect equilibria exist in a wide class of games. Other important results include the so called Folk theorems. These theorems essentially state that any strategy profile, for which every player receives at least as much as she can guarantee herself, can constitute an equilibrium in a repeated game. The reader is referred to [14] for an account of current results in this field.

Evolutionary game theory is an active research field which combines game theory with some dynamic mechanism through which the players can change strategies. Typically, better-performing players are assumed to replicate faster, representing a kind of Darwinian selection. Sometimes so called *multi-level selection* is also included, where evolution takes place simultaneously both at an individual level and at the level of the group. An example of this can be found in [23]. Under certain simplifying assumptions problems like these can be treated using tools from statistical physics; see for instance [4, 9]. Applications of evolutionary game theory can be found in as diverse fields as microbiology, social sciences and economics.

In real societies there are often opportunities for agents to act reciprocally, i.e. favor or disfavor certain behavior by engaging other agents. Reciprocity can be negative or positive, corresponding to punishing or rewarding agents based on their behavior, respectively. For example, a union might call for a strike, or a community of fishermen might destroy the equipment of someone fishing in an unsustainable way [18]. By including the opportunity to engage in reciprocal interactions, it is possible to study how punishments, or the threat of punishments, affect cooperation in games [6]. It has been found that when game participants are allowed to punish each other for acting selfishly, it can be easier to establish norms of cooperation [7, 8]. However, success is contingent on the presence of agents who are willing to punish "undesirable" behavior, a presence which varies across cultures [5]. In general, reciprocity may take various forms and in [12] different kinds of retaliation are formalized and shown to have the ability to facilitate the development of states of cooperation.

In this study we introduce an agent-based mathematical model of the *public goods game* where punishments are allowed for. In this game, personal and group interests are in conflict as players can choose whether or not to contribute to a common pool which benefits everybody equally. Numerical simulations constitute our main tool of studying this model, although a brief analytical treatment is also included. Simulating a game numerically is a recurring instrument for studying these types of games, see for example [20, 21], and has several benefits. For instance, it is easy to run large simulations, the cost associated with the study is

low, and the game is easy to modify. The challenge is to connect the simulation results to aspects of reality and draw economically sound conclusions, a discussion which is postponed to the concluding section.

The text is organized as follows. Section 2 provides a description of the public goods game and various extensions. The specific model used in this study is presented in Section 3, which also includes an analytical treatment of the model. In Section 4 we motivate and explain the proposed experiments and the results of these are presented in Section 5. Lastly, in Section 6 the implications of the experiments are discussed and analyzed along with the analytical results.

2 Public Goods Game

A good is considered to be public, in the true meaning of the word, if it is non-exclusively accessible to everyone and non-rival in nature [15]. In other words, individuals cannot be prevented from using the good and one individual's use does not impair others from using the exact same good*. In the public goods game, the public good is dependent on contributions from the participants. In that respect it centers on a specific aspect of societal interactions, namely the possible non-alignment of the interests of a group and those of its members. Stated more succinctly, all participants want to enjoy the benefits of the public good, but they would prefer if others paid for it. As previously mentioned, under the classical economic assumption where players only seek to maximize their own personal gains, all players would act selfishly. However, as stressed in the introduction, we often see evidence of altruistic or cooperative behavior in every day life and similar findings have also been confirmed in laboratory experiments, including those involving the public goods game. Such observations serve to motivate the study of more elaborate versions of the basic public goods game, to be specified below.

2.1 Formulation

The basic framework of the (single round) public goods game is a follows. A fixed number n of players form the group playing the game. Each player $i \in \{1, 2, \ldots, n\}$ is given the option to contribute $x_i \in X$ credits to a common pool, where X is the set of possible contributions. This common pool could for example represent an optional tax in the society, with the money intended to be used in a manner that will benefit everybody. Examples include for instance the construction of a road or a school, or more intangible purposes such as funding basic research. Players can refuse to contribute, in which case $x_i = 0$. In general, the payoff for player i can be expressed as

$$\pi_i(x_1, x_2, \ldots, x_n) = B_i(x_1, x_2, \ldots, x_n) - C_i(x_1, x_2, \ldots, x_n), \tag{2.1}$$

where B_i is the benefit function and C_i is the cost function associated with player i. These functions depend on the contributions of all players and both are assumed to be non-negative and monotonically increasing, with B_i being concave and C_i convex, for all $i = 1, 2, \ldots, n$.

In this study, the benefit functions will be restricted to the special case where they are all equal and depend only the size of the common pool, i.e. the sum

*For instance, the night sky is accessible to, and can be viewed at the same time by, everyone at night. For this reason the beauty of the night sky can be considered a public good.

of the contributions. Specifically, let the amount of credits in the common pool grow with a fixed factor $\kappa > 1$ and be redistributed equally among all players, without regard to who actually contributed*. In other words, the benefit function associated with player i could be expressed as

$$B_i(x_1, x_2, \ldots, x_n) = \frac{\kappa}{n}\left(\sum_{j=1}^{n} x_j\right), \tag{2.2}$$

a function which is linear in all the arguments†. Furthermore, if the cost for a player only consists of the individual contribution, the cost function associated with player i is given by

$$C_i(x_1, x_2, \ldots, x_n) = x_i. \tag{2.3}$$

Suppose $\kappa/n < 1$, which is reasonable since n is normally assumed to be a large number. Then, as previously asserted, if all players are faced with a utility function based solely on their own personal payoff the game has a unique Nash equilibrium where all players contribute zero credits to the common pool: $x_i = 0$ for all $i = 1, 2, \ldots, n$. The underlying reason is that the optimal strategy, given any possible set of actions undertaken by the other players, is to not contribute [13, 17].

A repeated game is created by introducing multiple rounds. Such games are motivated since real life interactions often occur repeatedly, for instance as a society finances research that will eventually benefit all citizens. Also in this case, assuming the game is played finitely many times, there is a unique Nash equilibrium where the contributions in each round are zero [13]. Yet contributions are often seen in laboratory experiments, but they typically decline over time as the contributing players realize that others are free riding [3].

A modification of the game is to allow players to punish each other after each round. This is introduced by including the possibility for a player to cause a reduction in the payoffs of other players at a cost to herself. The decision to punish a fellow player is based on the target having been deemed to play unfairly. Specifically, the type of punishment envisioned here is the ability of contributing players to reduce the payoffs of non-contributors. This would alter the cost function

*A different version of the public goods game sometimes seen in the literature is one where players have the option to withdraw (or use) a limited amount from a given common pool before the remaining resources grow [10]. This is closely related to is what is referred to as the *tragedy of the commons*, i.e. that people tend to use up common resources in an unsustainable way. These two versions are equivalent in terms of payoff structure, and while there may be a psychological difference between voluntarily contributing and withdrawing, such a distinction is irrelevant in a computer simulation.

†Notice that this type of benefit function does not correspond to a true public good according to the definition. However, the name is retained since it is the title conventionally used for this kind of game.

in (2.3) by adding a term representing the cost incurred by punishing others and a term containing the cost of being punished. These additional terms must be modeled such that they depend on the behavior of the players. This includes the propensity of individual players to punish unfair behavior and the preceding contributions made by the players in the game. The reader is referred to Section 3 for further details concerning the specific dependence used in this study.

Allowing for punishments sometimes induces cooperation in experiments: punishers can force selfish players to contribute and in some sense change the social norms in the population through their actions [6]. An empirical experiment of this kind has been carried out in [5], where the public goods game was played repeatedly in USA and Romania. Contrary to the Romanians, the Americans managed to establish a state where significant contributions were made in each round. The authors explained this difference by the presence of so called "leaders" in the American population who were ready to undertake personal costs in order to punish selfish behavior. In addition to negative reciprocity in the form of punishment, individuals can also be given the option to reward "good" behavior. This has been studied empirically in [17]. The results suggest that positive reciprocity may have a better effect on cooperation when the game is played repeatedly with the same participants, compared with the use of negative reciprocity alone.

2.2 Variations

In the following collection of paragraphs, various extensions of the public goods game are presented along with economic interpretations. These are later introduced in the model described in Section 3 in order to study their implications on our specific version of the public goods game.

2.2.1 Local versus Global Considerations

Real world public goods typically have a large number of users. However, in empirical studies of the public goods game the number of participants is often limited to a small number due to practical reasons. Having a large number of participants introduces the option of allowing the game to be played locally, globally or both, where local public goods represent resources which are accessible only by a subset of the players. One way to visualize the local versus the global nature of a public good is by considering a fresh water resource. In real life such a resource may consist of a vast number of lakes and connecting streams. The use of the fresh water resource is therefore divided into regional areas where users can only register the use of the resource by others in their immediate vicinity. On the other hand, using the water for toxic waste disposal will affect users for many hundreds of kilometers, if not the entire subsystem downstream. Those users which are not close to the

source of the toxic waste have no means of knowing who is contaminating their resource. It may even be the case that they are not even aware that their fresh water source is being contaminated by other users upstream.

2.2.2 Institutions

According to the above presentation of the public goods game, a player who decides to punish a non-contributor has to personally bear the cost of punishment. However, more elaborate forms of punishment behavior can also be envisioned. For instance, it could be possible for the punishers to share the burden of punishing non-contributors by splitting the costs between each other. This type of cooperation can be formalized by the creation of an institution overseeing the punishments and managing the costs. This institution can be seen as being financed by a voluntary tax, which brings the costs back to those players who are willing to punish non-contributors*. For previous results involving cost-sharing between punishers in a public goods game setting see for instance [19].

2.2.3 Norm Dependence

Up until now players have only had a binary option concerning punishments; they are either punishers or not. However, it is reasonable to expect that the propensity to punish a discovered non-contributor depends on the number of non-contributors in the game. If there are few contributors, it might be the case that players are less inclined to punish non-contributors simply because they deem the practice futile. The opposite behavior can also be motivated, for instance players letting a few non-contributors go unpunished since they pose no threat to the prevailing norm of contributing to the common pool.

Both cases have rather clear-cut equivalents in a real society. The former case allows for forceful punishments of individuals whose shortcomings make a group of people "look bad". Examples include exclusion from a group or physical abuse due to deviant behavior or an atypical dress code compared with the majority of the group members. On the other hand, the latter case is more pragmatic in the sense that the punishers do not expect all norms to be followed to their fullest extent. For instance, traffic police officers usually do not issue tickets for exceeding the speed limit by a few kilometers per hour. However, more serious or numerous infractions would be punished to a greater extent, possibly even by incarceration.

In a similar manner, contributors themselves can be less inclined to contribute to the common pool if non-contribution is pervasive. For this reason, the propen-

*It is also possible to envision a similar type of cooperation among the non-contributors where they share the effects of the punishments. Examples include organizations providing free riding insurances in the sense that they compensate their (paying) members for fines obtained for instance when getting caught using public transportation without a valid ticket.

sity to contribute can be made directly dependent on the average contribution in previous rounds. Also in this case the economic interpretations are rather straightforward. For instance, residents of an apartment building can be assumed to be less prone to participate in an annual cleaning of a common garden if the likelihood is low that other residents will show up. This is different from having a disposition toward not wanting to contribute. Instead the residents are merely disillusioned and would likely participate if the participation norm was reestablished.

2.3 Common Property Resources

The previously mentioned water resource example is related to the study presented in [18] concerning so called *common property resources*. These resources can be viewed as extensions of public goods and the objective in the cited paper was to examine the usage of commonly owned renewable resources and possible overexploitation from an evolutionary-game-theoretic point of view. Essential questions include the manner in which norms of restraint and punishment can induce a cooperative state such that the resource is harvested in a sustainable manner. To this end, the authors considered a common renewable resource, for example the fish stock in a lake or a local forest. A fixed number of agents with different strategies have access to and the right to harvest the resource. However, in contrast to the public goods game, harvesting the resource under consideration requires an input in the form of labor.

Suppose the utility function associated with each agent consists exclusively of her personal payoff. Furthermore, let the aggregate harvest be considered as a function of aggregate labor input. Under certain requirements on this functional dependence, a model can be created where a certain level of aggregate labor corresponds to sustainable consumption of the resource at the same time as each agent has a reason to overconsume by increasing their input of labor. This leads to contrasting interests between the group and its members. Without further restrictions, the unique Nash equilibrium of the game is a state where each agent uses an unsustainable amount of labor to harvest the resource, which results in overexploitation and the eventual depletion of the resource. This is the basic outline of the so called *common pool resource* (CPR) game. In a repeated version of the CPR game punishment can be included which, similar to the public goods game, can help to establish a norm of sustainable resource use. It is shown in [18] that the inclusion of agents willing to incur personal costs in order to punish those who break the norm can lead to a stable state of sustainable consumption.

A link between the CPR game and the public goods game considered in this study can be provided by assuming that the punishment of defectors is supervised by an external authority. The responsibility of this organization is to detect defectors and punish them. If this authority is financed by a voluntary tax, the

authority itself can be viewed as a type of public good since it allows for a sustainable use of the common resource (assuming the organization is run with a sufficient level of efficiency). Thus, its existence would benefit everybody equally, independent of who payed the voluntary tax [18]. This extension of the CPR game also has a close resemblance to the institutional extension of the public goods game presented above.

3 Model

The proposed extended public goods game is an n-player game with the following basic outline. (i) Each player has a set of characteristics detailing if she contributes to the common pool and whether or not she is inclined to engage in negative reciprocal behavior. (ii) The model is endorsed with a certain network structure in order to define the concept of two players being neighbors. (iii) The game itself is played for T rounds, with each round consisting of three distinct stages: *the contribution stage*, *the punishment stage* and *the revision stage*. In the contribution stage the players decide if they want to contribute to the common pool. In the punishment stage players have the possibility to punish others based on behavior observed during the contribution stage. Lastly, the third and final stage provides the dynamics of the game, i.e. it allows the player configuration to change from one round to the next.

3.1 Static Component

The word *contributors* will be used to describe players who contribute to the common pool, while *punishers* will be used to denote players who are prepared to incur a cost in order to punish non-contributors. Each player can possess both, either or neither of these properties. We will however restrict the set of strategies to the following*:

1. *Defectors* – Players who neither contribute nor punish.
2. *Enforcers* – Players who both contribute and punish.
3. *Cooperators* – Players who contribute but do not punish.

Each of these strategies is denoted using a positive integer k, with 1 corresponding to defection, 2 to enforcement and 3 to cooperation. The terminology follows the one used in [18]. As a first step in each game, n players are generated and equipped with strategies drawn from a certain probability distribution. This can be viewed as randomly picking players from a large population pool, where the prevalences of different strategies define the probability distribution. Specifically, let μ_k denote the probability that a generated player employs strategy k, where the additional requirement $\mu_1 + \mu_2 + \mu_3 = 1$ is imposed.

*There is one possible strategy missing: players who do not contribute, but who would punish other non-contributors. They would therefore have a rather destructive effect on the game and are excluded here, even though it should be possible to find similar malevolent behavior in a real society. It is also possible to envision other types of behavior One example is non-contributing players who punish other contributors, so called *anti-social punishment*. Such counterproductive behavior has actually been observed in experiments [11].

As previously mentioned, the game has an underlying network structure in the form of a lattice with n lattice sites. The purpose of the network is to limit reciprocity to neighbors only, i.e. a player can only respond to the behavior of those situated closest to her. This models a lack of information in the game: players can only observe, and thus only react to, actions undertaken by their neighbors. The types of lattices considered will be restricted to one-dimensional lattices with periodic boundary conditions. The periodic boundary conditions are imposed in order to remove unwanted effects along the boundaries of the lattice. This can be viewed as taking an open string and closing it to form a circle; see Figure 1.

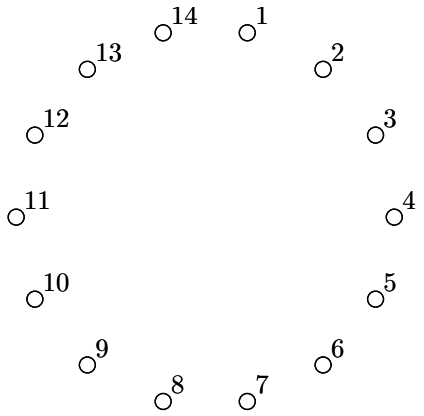


Figure 1: *Example of a one-dimensional lattice with* 14 *lattice sites and periodic boundary conditions.*

Since the lattice is one-dimensional, each player has exactly two nearest neighbors, i.e. neighbors who are situated one step away from the player along an edge in the lattice. Consider player i situated at the i:th lattice site, as depicted in Figure 2. The set of nearest neighbors is the only group visible to i and consequently comprises those players which may be subject to reciprocal engagements undertaken by i*.

Consider the first stage of the game and let x_i denote the contribution from player i to the common pool. The choice is restricted to be binary, i.e. $x_i \in \{0, 1\}$.

* It is possible to also allow for reciprocal behavior between next nearest neighbors and so on, e.g. by introducing an "observation probability" which decreases with distance. This would model a situation where players occasionally also observe the actions of players situated further away. There are also other possibilities, e.g. letting the information gathering from neighbors be plagued by some stochastic noise which intensifies with increasing distance. Thus, trying to determine the behavior of players far away from oneself would be marred by uncertainty. Further developments include altering the lattice, for instance by considering two- or higher-dimensional lattices. A more general model would remove the lattice altogether and allow for arbitrary networks. One could also envision a model where the connections between players are allowed to change from round to round, in that way allowing players to relocate geographically.

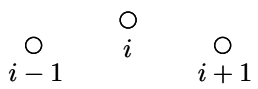


Figure 2: *Illustration of player i and her nearest neighbors.*

The assets in the common pool grow with a factor $\kappa > 1$ and are distributed equally between all players. The initial payoff $\pi_{i,\text{init}}$ for player $i = 1, 2, \ldots, n$ in each round is thus

$$\pi_{i,\text{init}} = \kappa\bar{x} - x_i, \tag{3.1}$$

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{3.2}$$

is the average contribution to the common pool, $0 \leq \bar{x} \leq 1$.

In the second stage, enforcers have the ability to punish at most one nearest-neighboring defector. This is implemented by letting an enforcer i randomly select one nearest neighbor with equal probability. If a defector is found, the enforcer punishes. The *cost of punishment* incurred by the enforcer is denoted by c and the *effectiveness of punishment* $\theta > 0$ is defined as the punishment to cost-ratio, meaning that the payoff of a discovered defector is reduced by θc. The final payoffs for each strategy in a given round can thus be written:

$$\pi_i = \kappa\bar{x} - x_i - \theta c\sum_{j=1}^{n} s_{ji}, \tag{3.3a}$$

$$\pi_i = \kappa\bar{x} - x_i - c\sum_{j=i}^{n} s_{ij}, \tag{3.3b}$$

$$\pi_i = \kappa\bar{x} - x_i, \tag{3.3c}$$

for defectors, enforcers, and cooperators, respectively. The punishment matrix s_{ij} is given by

$$s_{ij} = \begin{cases} 1 & \text{if } i \text{ punishes } j, \\ 0 & \text{otherwise,} \end{cases} \tag{3.4}$$

for $i, j = 1, 2, \ldots, n$. Since an enforcer can punish at most one player per round the sum of the elements in each row in s_{ij} can be at most 1, while the sum of the elements in each column can be at most equal to the number of nearest neighbors.

3.2 Game Dynamics

In the revision stage of each round a certain number n_R of players are randomly selected for strategy revision, a so called *revision draw*. The selected players are then each assigned a probability p_i, which decreases as a function of individual payoff. A random Bernoulli variable with parameter p_i is sampled to determine if the player revises her strategy or not. To be more specific, let λ be a positive parameter. The probability p_i that player i revises her strategy is given by an exponential distribution,

$$p_i = \lambda e^{-\lambda y_i}, \tag{3.5}$$

where

$$y_i = 2 \cdot \frac{\pi_i - \pi_{min}}{\pi_{\max} - \pi_{\min}} \geq 0 \tag{3.6}$$

are relative payoffs, with $\pi_{\max}$ and $\pi_{\min}$ denoting maximal and minimal payoffs in the current round, respectively*.

The strategy revision itself is implemented via mutation or imitation, where imitation can be local or global. Let ξ_1 and ξ_2 be two parameters such that $0 \leq \xi_1 \leq \xi_2 \leq 1$. A player i that has been selected for strategy revision imitates either the strategy of a nearest neighbor with probability ξ_1, the strategy of any of the n players in the game with probability $\xi_2 - \xi_1$, or mutates to a randomly selected strategy with probability $(1 - \xi_2)$. Notice that in the case of mutation it is not necessary for the target strategy to be adopted, or to have been previously adopted, by a player in the game. With global imitation or mutation present, a certain strategy may appear in a cluster of dissimilar strategies, a phenomenon which is not possible in the local imitation case ($\xi_1 = \xi_2 = 1$). In the case of imitation, the player who is imitated is selected using uniform probability distributions: for local imitation each neighbor is imitated with equal probability, while for global imitation any player in the game is imitated with equal probability. On the other hand, in the case of mutation, any strategy is adopted with probability 1/3 by the player revising strategy.

3.3 Model Summary

There are a number of exogenous parameters, some of which will be specified, while others will be subject to parameter studies. In particular, the growth factor of the common pool is chosen as $\kappa = 1.6$ since it does not affect the relative payoffs between players; see equations (3.3) and (3.6). Also, for the sake of simplicity, the

*The factor 2 is included in order to increase the spread, making the numerical simulations converge more rapidly. The distribution is thus relative and always disfavors the worst-performing players. In the special case where $\pi_{\max} = \pi_{\min}$ the number of revision draws is set to $n_R = 0$ in order to avoid division by zero.

cost of punishment is specified as $c = 1$ and the probability parameter is set to $\lambda = 0.75$. In Table 1 all parameters are listed for ease of reference along with their ranges or fixed values.

Table 1: *Exogenous parameters in the model.*

Parameter	Description	Parameter range
n	Number of players	$n \in \{2, 3, \dots\}$
T	Number of rounds	$T \in \{1, 2, \dots\}$
μ_k	Initial probability for strategy k	$\mu_k \in [0, 1]$
κ	Growth factor	$\kappa = 1.6$
c	Cost of punishment	$c = 1$
θ	Effectiveness of punishment	$\theta > 0$
n_R	Number of revision draws	$n_R \in \{0, 1, \dots n\}$
λ	Imitation probability parameter	$\lambda = 0.75$
ξ_1	Local-global imitation delimiter	$\xi_1 \in [0, 1]$
ξ_2	Imitation-mutation delimiter	$\xi_2 \in [\xi_1, 1]$

3.4 Theoretical Study

The model presented above can be studied theoretically. The underlying idea behind the analytical treatment is in analogy with replicator dynamics [22] with inspiration collected from [23]. Here, only a summary of the main results are given, and the reader is referred to Appendix A for the derivation. This includes the statement of the so called master equation for the model [16]; see Proposition 3.1 below. For simplicity we will also assume $n_R = n$, meaning that the probability (3.5) is calculated for all n individual players in the game.

Let k and h be positive integers denoting strategies. Their possible values are 1, 2 or 3, with 1 corresponding to defection, 2 to enforcement and 3 to cooperation. At time t, denote by $\zeta^k(t)$ the fraction of players employing strategy k and let $p^k(t)$ be the probability that a player employing strategy k revises strategy. The latter is given by

$$p^k(t) = \lambda e^{-\lambda y^k(t)}, \tag{3.7}$$

where the relative payoffs $y^k(t)$ are defined as

$$y^k(t) = 2 \cdot \frac{\pi^k(t) - \pi_{\min}(t)}{\pi_{\max}(t) - \pi_{\min}(t)}. \tag{3.8}$$

Using a so called mean field approximation [16], the payoffs $\pi^k(t)$ for each respec-

tive strategy can be derived and shown to be:

$$\pi^1(t) = \kappa\left(\zeta^2(t) + \zeta^3(t)\right) - \theta c\zeta^2(t), \tag{3.9a}$$
$$\pi^2(t) = \kappa\left(\zeta^2(t) + \zeta^3(t)\right) - 1 - c\zeta^1(t), \tag{3.9b}$$
$$\pi^3(t) = \kappa\left(\zeta^2(t) + \zeta^3(t)\right) - 1, \tag{3.9c}$$

for defection, enforcement and cooperation, respectively. Essentially, the mean field approximation implies a global mixing of players between each round. In other words, the probability that a given neighbor employs strategy k is only dependent on the current prevalence of strategy k, and not on the outcomes in previous rounds. In contrast, the complete model exhibits a certain dependence on the positioning of players. This dependence is lost in the mean field approximation, impeding any direct comparisons between the two.

Using the above relations, we are able to state the master equation for the model in the mean field approximation:

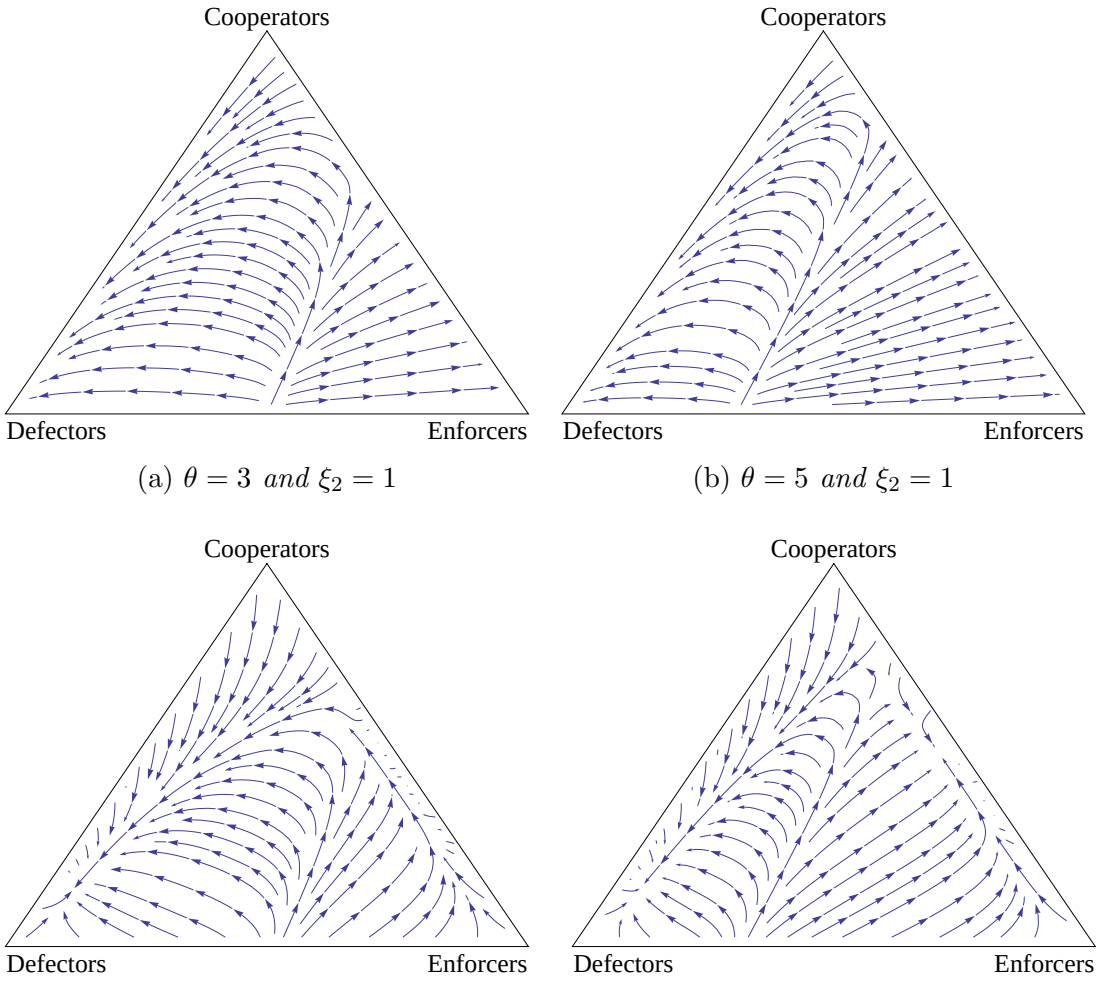
Proposition 3.1. (Mean Field Master Equation) *In the mean field approximation, assuming $\zeta^k > 0$ for all $k = 1, 2, 3$, the evolution of the model is governed by the master equation*

$$\dot{\zeta}^k(t) = -p^k(t)\zeta^k(t) + \sum_{h=1}^{3}\left(\xi_2\zeta^k(t) + \frac{1}{3}(1-\xi_2)\right)p^h(t)\zeta^h(t), \tag{3.10}$$

where $\dot{\zeta}^k(t)$ is the time derivative of $\zeta^k(t)$ and $p^k(t)$ is given by (3.7) *along with* (3.8) *and* (3.9).

Remark 3.2. The parameter ξ_1 is not present in (3.10) since the distinction between local and global imitation disappears in the mean field approximation due to the global mixing of players.

In Figure 3 flows of the master equation (3.10) are plotted in a simplex. These flows are calculated by evaluating the right-hand side for different values of θ and ξ_2 and the diagrams essentially describe the evolution of the model under the rather crude simplifications in the mean field approximation. Clearly, in each of the depicted cases the cooperator strategy is unstable. For the upper two cases (where $\xi_2 = 1$), this means that even the slightest deviation from the pure cooperator equilibrium is enough to cause a divergence from the state. In Section 6 a longer discussion is included along with a comparison between these analytical results and the results from the numerical simulations.



(a) $\theta = 3$ *and* $\xi_2 = 1$ (b) $\theta = 5$ *and* $\xi_2 = 1$

(c) $\theta = 3$ *and* $\xi_2 = 0.5$ (d) $\theta = 5$ *and* $\xi_2 = 0.5$

Figure 3: *Flows of the master equation for different configurations. In the two upper flow diagrams, all three pure states constitute equilibria, but only the pure enforcer and pure defector states are stable. This is not the situation when strategy revision through mutation is present, i.e. when* $\xi_2 < 1$*, in which case neither of the pure states are stable.*

3.5 Variations

The public goods game can be extended by adding a variety of features. Below, two variations of our specific version of the game are presented. Each of these serves to investigate the effects of certain social structures or agent behaviors which can be added to the existing model.

3.5.1 Institutional Structure

In the first variation of the model we consider the situation where the enforcers are able to share the costs of punishing defectors among each other. This is implemented by the creation of an institutional structure supervising the punishment process. In each round, all enforcers are obliged to give an equal amount of credits to the institution, which then covers the costs of the punishments carried out in the round. To this end, let η denote the *institutional inefficiency*, meaning that the total cost of punishment shared by all enforcers is multiplied by a factor $1+\eta$. Using an institutional structure of this kind, the payoffs for each strategy in a given round can be written:

$$\pi_i = \kappa\bar{x} - x_i - \theta c \sum_{j=1}^{n} s_{ji}, \tag{3.11a}$$

$$\pi_i = \kappa\bar{x} - x_i - (1+\eta)\frac{c}{n_E} \sum_{j=1}^{n}\sum_{l=1}^{n} s_{jl}, \tag{3.11b}$$

$$\pi_i = \kappa\bar{x} - x_i, \tag{3.11c}$$

for defectors, enforcers, and cooperators, respectively, where n_E denotes the number of enforcers in the round.

3.5.2 Norm Dependent Strategies

The second variation concerns the possibility of enforcers to base their willingness to punish defectors both on the actions of their nearest neighbors and on the number of contributors, i.e. on the size of the common pool*. The relation between the two sources of information in the decision making process is specified by letting the propensity to punish to depend linearly on the average contribution $\bar{x}$ in each round. To this end, let a and b be parameters such that $a, b \in [0, 1]$ and define the propensity r^p of an enforcer to punish an observed non-contributing nearest

*Each player can calculate the average level of contribution (ex post) using (3.1) provided she knows the value of κ, since she has access to both the size of her own contribution x_i and the initial payoff $\pi_{i,\text{init}}$ received.

neighbor by

$$r^p = a + (b - a)\bar{x}. \tag{3.12}$$

Hence, $r^p = a$ if $\bar{x} = 0$ and $r^p = b$ if $\bar{x} = 1$; see Figure 4 for a graphical illustration. Enforcers with $(a, b) \neq (1, 1)$ are referred to as *norm dependent enforcers* while those with $(a, b) = (1, 1)$ are referred to as *non norm dependent enforcers*. Notice that letting $(a, b) = (0, 0)$ implies that the enforcer strategy becomes the same as the cooperator strategy.

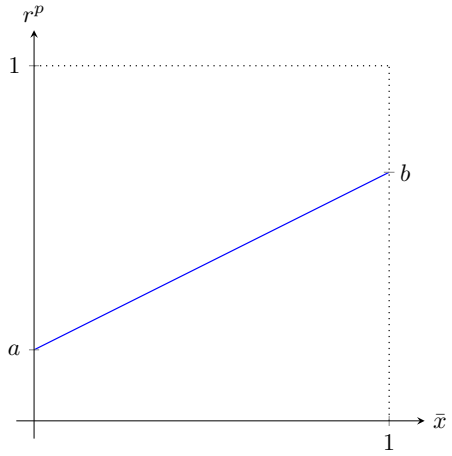


Figure 4: *Propensity r^p of an enforcer to punish an observed non-contributing nearest neighbor as a function of the average contribution $\bar{x}$ to the common pool.*

We can distinguish three qualitatively different enforcer strategies:

(i) $b > a$ – In this case, enforcers are more prone to punish defectors the higher the average contribution. In other words, a few defectors are punished harder, while if there are plenty of defectors, the enforcers become "demoralized" and refrain from punishing. For a real society this would correspond to individuals punishing, more severely or to a greater extent, those who break a given norm if the number of norm breakers is small.

(ii) $b = a$ – This case corresponds to enforcers being indifferent to the size of the average contribution. In particular, the propensity to punish a discovered defector need not be equal to 1. However, this is determined based on external criteria, and not as a consequence of the average contribution.

(iii) $b < a$ – In this case enforcers may overlook the existence of a few defectors, in a sense viewing it as a minor problem. Nevertheless, if the number of defectors grows the enforcers become "agitated" and more eager to punish. This latter behavior can be viewed as the enforcers recognizing that without tougher actions the public good might collapse.

In a similar manner one can introduce an analogue strategy option for contributors in general, defined by two parameters α and β such that $\alpha, \beta \in [0, 1]$. In this case, the propensity r^c to contribute is given by

$$r^c = \alpha + (\beta - \alpha)\bar{x}. \tag{3.13}$$

Correspondingly, *norm dependent contributors* are defined as contributors having $(\alpha, \beta) \neq (1, 1)$, while *non norm dependent contributors* as those with $(\alpha, \beta) = (1, 1)$. Notice that since both enforcers and cooperators are contributors, they are both affected by the choice of the parameters α and β.

4 Experiments

In the following section the simulations included in this study are presented and motivated. Since the model described in Section 3 has a large number of exogenous parameters, we are only able to study a limited number of configurations. The results of the simulation experiments can be found in Section 5.

4.1 Random Initial Distributions

We first consider random initial player distributions in an attempt to assess how the effectiveness of punishment affects the asymptotic outcome of the game. This is motivated by questions concerning, for instance, how severe punishments must be in order to enforce cooperation. As a first step, a number of specific simulations are performed in order to illustrate the dynamics of the model before we continue with a more rigorous parameter study. The latter is carried out by varying the parameter θ between $\theta = 1$ and $\theta = 9$ for a configuration with uniform initial strategy distribution and only local imitations.

As mentioned in Section 3.2, the game dynamics are potentially different with global imitation present, which warrants a study of the local-global imitation delimiter ξ_1. To this end, the effect of local and global imitations is investigated by letting ξ_1 vary between $\xi_1 = 0$ and $\xi_1 = 1$. The simulations are carried out using a configuration with uniform initial strategy distribution, effectiveness of punishment $\theta = 3$ and imitation-mutation delimiter $\xi_2 = 1$.

4.2 Stability Analysis

Another interesting kind of experiment concerns the ability of a group with a deviant strategy to affect an otherwise homogeneous population of players. If the imitation-mutation delimiter is chosen as $\xi_2 = 1$, no mutation occurs and each homogeneous state constitutes an equilibrium. The question we seek to investigate concerns the stability of these equilibria to perturbations. The experiment can be thought of as the effect of immigration on a society with established social norms. When players with a differing social norm arrive in a homogeneous society three things can happen: (i) assimilation, where the immigrants adopt the prevailing norm; (ii) penetration, where the natives convert to the new norm; or (iii) coexistence, where both norms remain. We will look at three cases where a small portion of players with a deviant strategy is inserted into a homogeneous group. We place the newcomers next to each other, which can be thought of as immigrants settling in the same neighborhood. In case (1), we insert enforcers among defectors in order to investigate the ability of the enforcers to establish norms of contribution in a non-cooperative society. In case (2), a seed of defectors is introduced into a

homogeneous population of pure cooperators. Lastly, in case (3), we let defectors settle in a population of enforcers in order to assess the ability of the enforcers to halt the spread of defection.

4.3 Variations

In addition to the original model, a variation which includes institutional punishment is also studied numerically. In this case the enforcers share the costs of punishing defectors equally among each other through an institutional structure. This structure has an institutional inefficiency parameter η which is varied between $\eta = 0$ and $\eta = 10$. The simulations are carried out for a configuration with effectiveness of punishment $\theta = 2.5$, uniform initial strategy distribution and only local imitations.

Finally, the manner in which the presence of norm dependent enforcers and contributors affects the evolution of the game is studied using a parameter study of the effectiveness of punishment. For simplicity, the parameters specifying the norm dependence of the enforcer and cooperator strategies are chosen as $(a, b) = (0, 1)$ and $(\alpha, \beta) = (1/2, 1)$. The parameter study is carried out by varying the parameter θ between $\theta = 1$ and $\theta = 12$ for a configuration with uniform initial strategy distribution and only local imitations.

5 Results

This section contains the results from the various experiments detailed in Section 4. The results are presented using time-plots of certain descriptive outputs, such as accumulated payoff and number of contributors, and circular representations of the distribution of strategies and payoffs in each simulation. For a brief description of issues related to the numerical implementation of the model, see Appendix B.

5.1 Random Initial Distributions

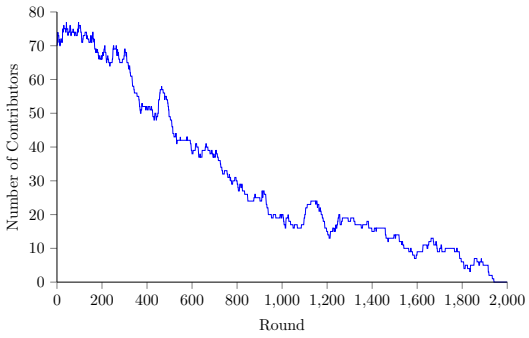
Initially, two examples are given in order to illustrate the behavior of the model. Some caution is advised since the model is stochastic, meaning that conclusions should not be drawn based on the outcomes of individual simulations. However, different mechanisms in the model can still be investigated using this approach. The examples are followed by parameter studies investigating the effects of θ and ξ_1 on the asymptotic outcome of the game.

5.1.1 Specific Examples

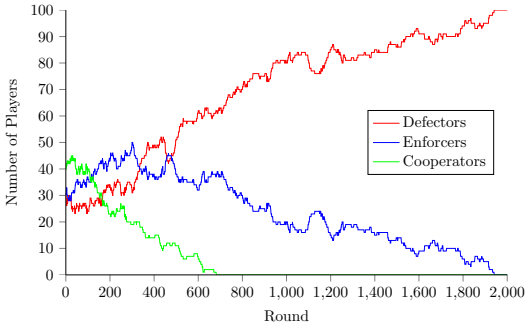
In Examples 1 and 2 below, the exogenous parameters are largely the same except for the value of θ. The first example has been run for 2,000 rounds and the second for 3,000 rounds, using a setup consisting of 100 players, uniform initial strategy distribution, only non norm dependent enforcer and cooperator strategies, and strategy revision restricted to local imitations*.

Example 1 In the following example the effectiveness of punishment is set to $\theta = 2$. Figure 5 includes time-plots of the number of contributors, the number of players employing each strategy, and the accumulated payoff for each strategy. As can be seen, the defectors make huge gains while enforcers and cooperators are losing credits in each round. The players employing strategies other than defection gradually convert to defectors and the game ends up in a non-cooperative equilibrium after approximately 1,000 rounds.

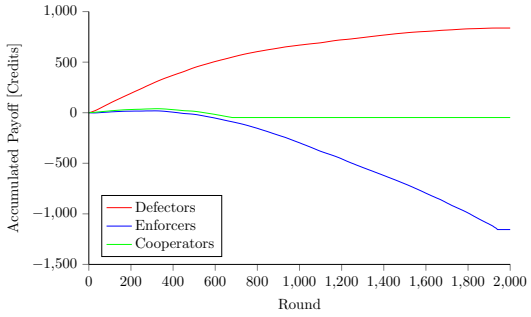
*Specifically, the parameters were set to: $n = 100$, $n_R = 10$, T=2,000 (T=3,000), $\xi_1 = \xi_2 = 1$, $a = b = \alpha = \beta = 1$, $\mu_1 = \mu_2 = 0.33$ and $\mu_3 = 0.34$.



(a) *Number of contributors*



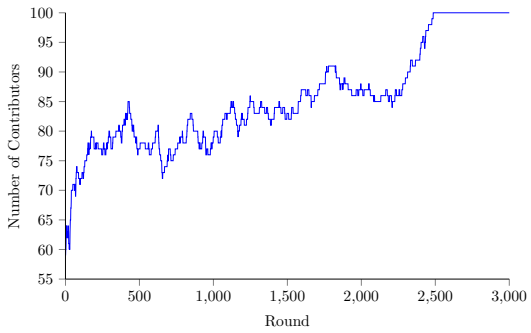
(b) *Number of players for each strategy*



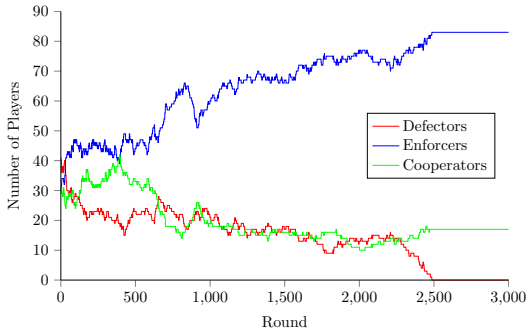
(c) *Accumulated payoff*

Figure 5: *Time-plots of the results from a simulation with* $\theta = 2$.

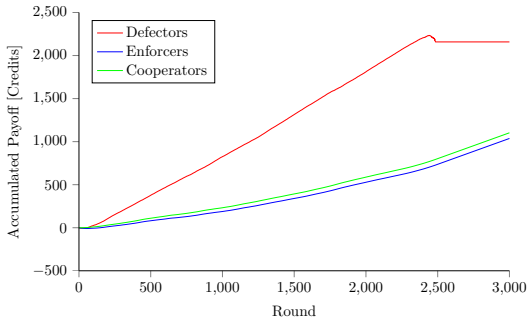
Example 2 In this example the effectiveness of punishment θ is increased to 5. The results are presented using time-plots in Figure 6 and circular representations of the outcomes in each round in Figure 7. Contrary to the previous example, the enforcers prevail and a cooperative equilibrium is established. When comparing the two representations in Figure 7 it is clear that all effects are confined to the boundaries between groups adhering to different strategies. In particular, the defectors situated next to enforcers are punished with such severity that they revise strategy and become enforcers themselves. This process continues until the defector strategy has been completely eliminated from the game.



(a) *Number of contributors*

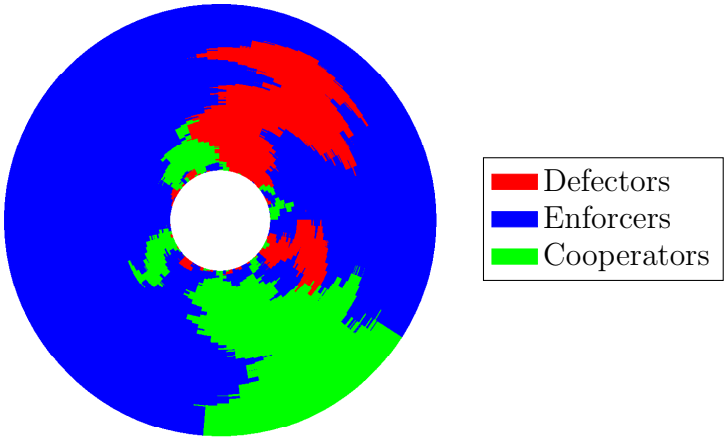


(b) *Number of players for each strategy*

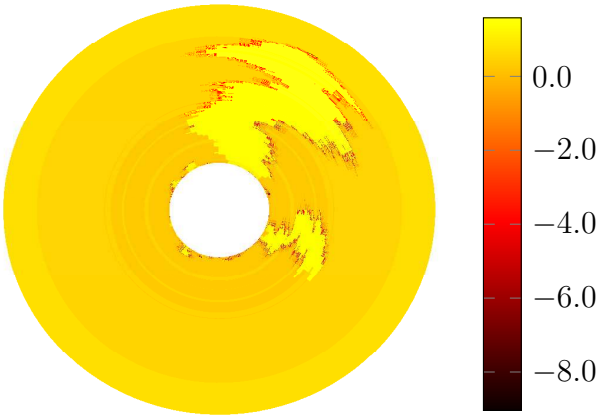


(c) *Accumulated payoff*

Figure 6: *Time-plots of the results from a simulation with* $\theta = 5$.



(a) *Circular representation of the game portraying the distribution of strategies in each round.*



(b) *Circular representation of the game portraying the payoffs in each round. Dark colors correspond to low payoffs and bright colors to high payoffs.*

Figure 7: *Diagrams detailing the evolution of strategies and payoffs in a simulation with* $\theta = 5$. *The radial direction corresponds to the time axis with each concentric circle portraying the state of the game in a particular round. In other words, the innermost circle represents the initial distribution and the outermost circle the final distribution.*

5.1.2 Effectiveness of Punishment

The results from the parameter study of θ detailed in Section 4.1 are presented in Figure 8. For each value of θ, a total of 100 simulations were carried out with the results recorded after 10,000 rounds. As can be seen, the number of contributors increases with θ, with a rapid transition in the interval between $\theta = 2.2$ and $\theta = 4.5$. In Figure 9 simulation outcomes for certain values of θ are presented in simplices. Each point in a simplex represents a unique state, for example the corner points represent homogeneous populations while the other represent mixed states. These simplex diagrams thus recapitulate the distribution of outcomes for the corresponding simulations. Clearly, as the effectiveness of punishment increases, the outcomes transition from ending up around the pure defector state to positions along the line connecting the pure enforcer state with the pure cooperator state.

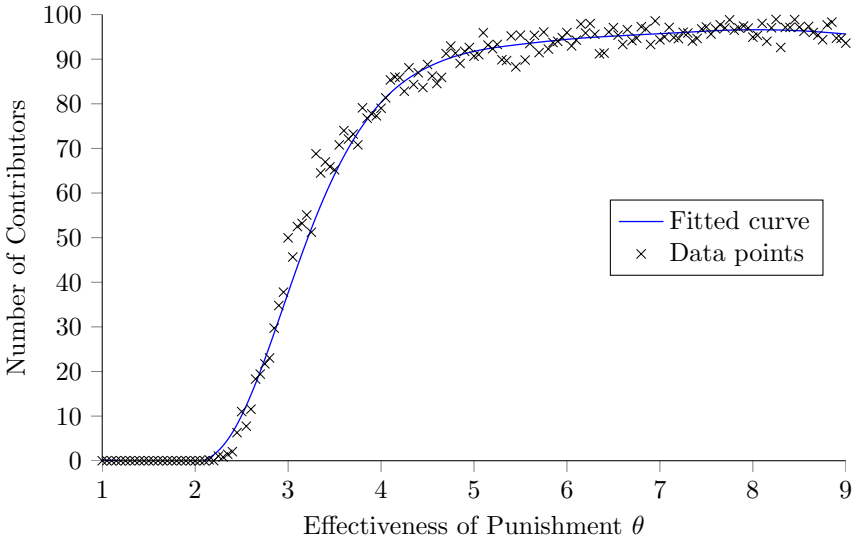
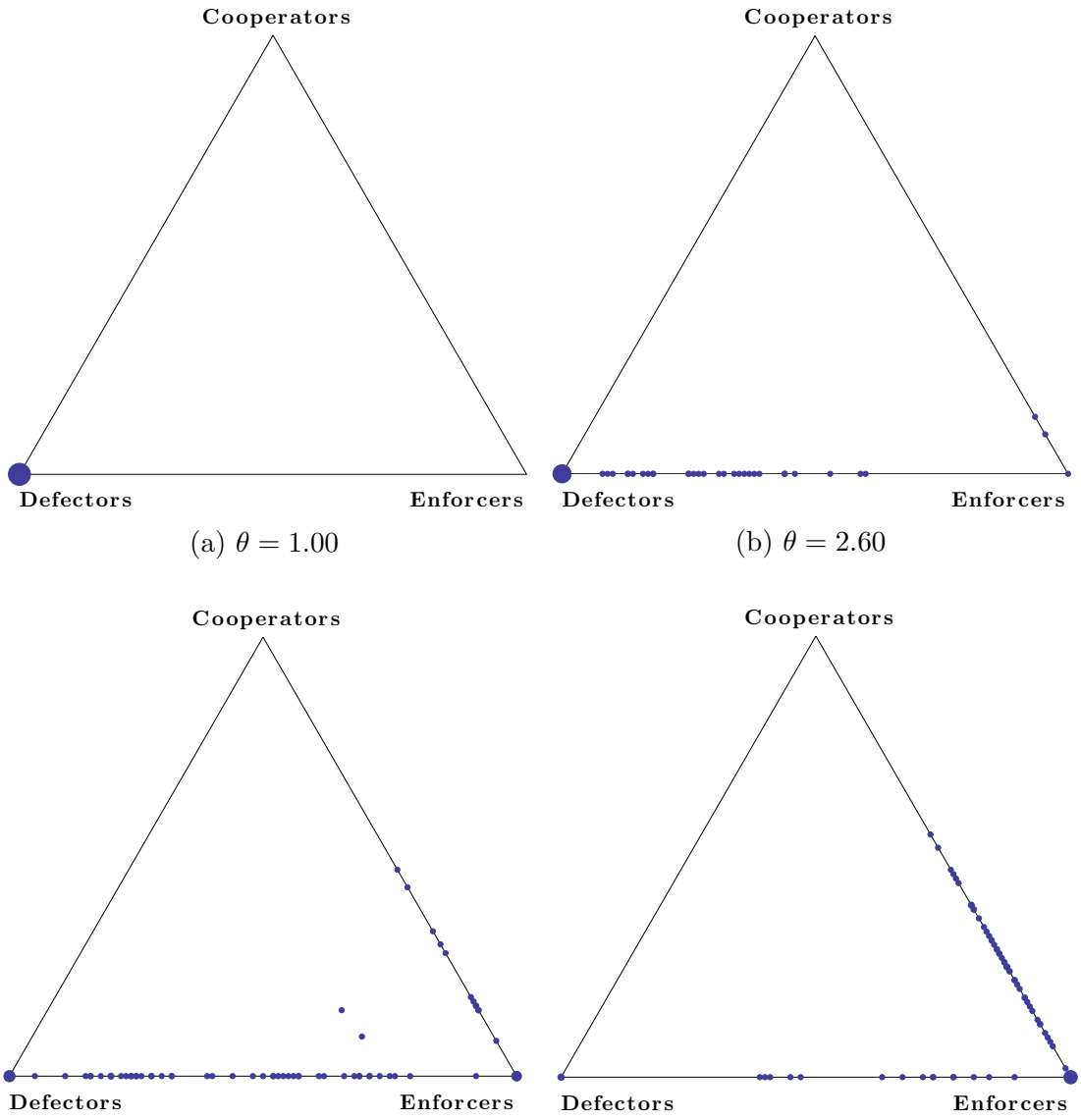


Figure 8: *Number of contributors for different values of the effectiveness of punishment θ. The data points are averages over* 100 *simulations with the results recorded after 10,000 rounds.*



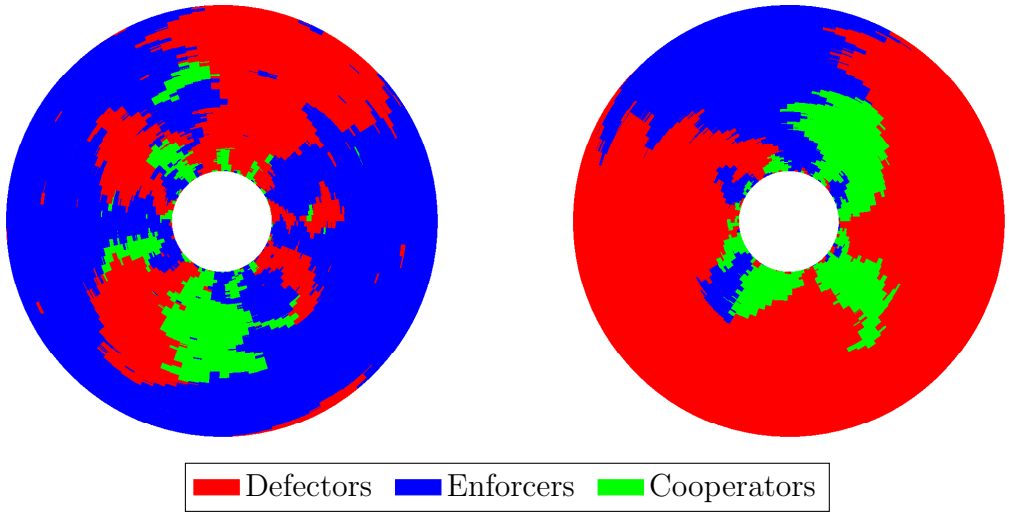
(a) $\theta = 1.00$ (b) $\theta = 2.60$

(c) $\theta = 3.00$ (d) $\theta = 5.00$

Figure 9: *Distribution of states after 10,000 rounds for certain values of θ. The size of a blue dot corresponds to the number of simulations which ended up in that state. Also, the center of a simplex corresponds to the starting point since the simulations were initiated using a uniform strategy distribution.*

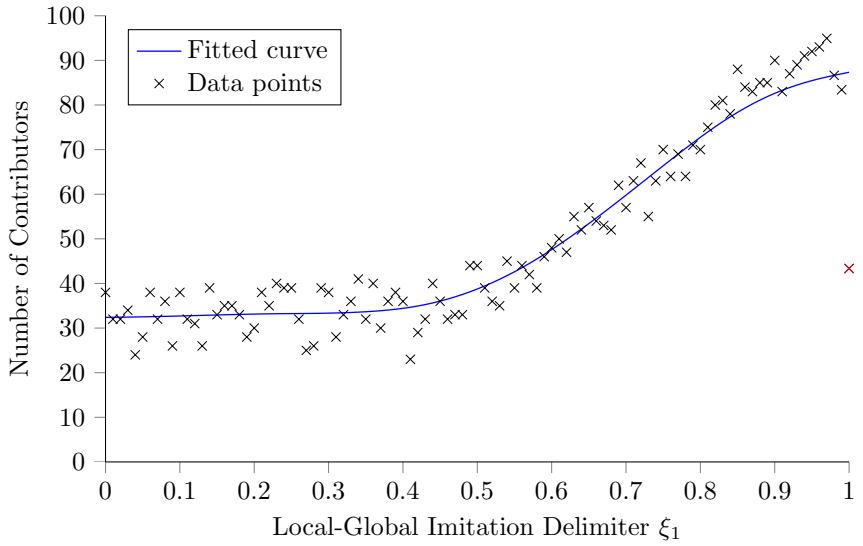
5.1.3 Local-Global Imitation Delimiter

The results from the parameter study of ξ_1 outlined in the last paragraph of Section 4.1 are depicted in Figure 11. From the definition of the local-global imitation delimiter it is clear that a higher value of ξ_1 implies that a player revising strategy is more likely to imitate a neighbor than a random player in the game. As can be seen in Figure 11a, the average level of contribution rises when the probability of local imitation increases from 0.5 to 0.97. Nevertheless, for $\xi_1 = 1$ the average level of contribution goes down abruptly to around 45%, a value which is consistent with the results in Figure 8. This decline is further investigated in Figure 11b which reveals a rapid decline for values just below $\xi_1 = 1$. The number of imitations in a typical game played for 10,000 rounds, similar to the one under consideration, lies between 10,000 and 30,000. Hence, the results in Figure 11 indicate that even if only a few hundred of these are global imitations, the outcome can change significantly. An example of this effect is presented in Figure 10, where a small change in ξ_1 from 0.97 to 1 drastically changes the qualitative behavior in favor of the defectors. As expected, in Figure 10a deviant strategies sometimes suddenly appear in previously homogeneous clusters, which results in a more fragmented strategy distribution.

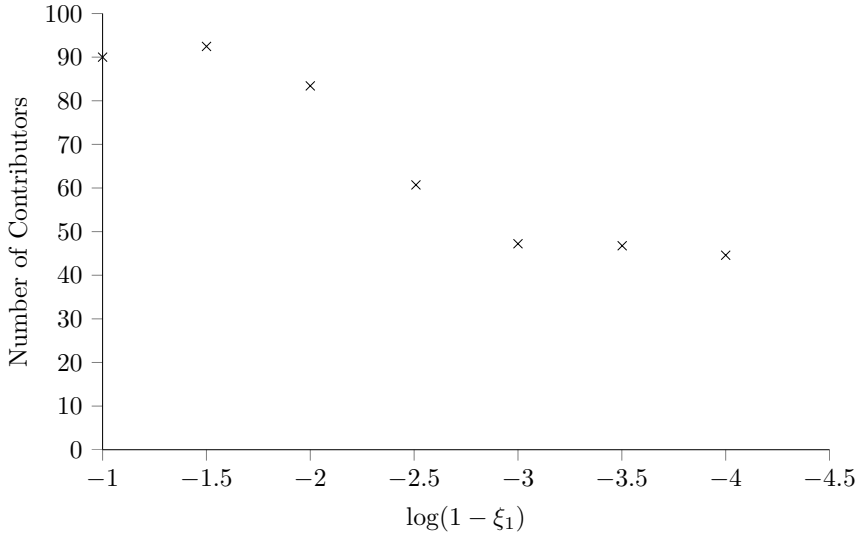


(a) $\xi_1 = 0.97$ (b) $\xi_1 = 1$

Figure 10: *Diagrams detailing the evolution of strategies during 2,000 rounds in two simulations with $\theta = 3$ and different values of ξ_1. The radial direction corresponds to the time axis with each concentric circle portraying the state of the game in a particular round.*



(a) *Simulation results*



(b) *Decline around* $\xi_1 = 1$

Figure 11: *Number of contributors for different values of* ξ_1 *for a configuration with* $\theta = 3$. *The data points are averages over* 100 *simulations with the results recorded after 10,000 rounds. In* (a) *the rightmost data point* (red cross), *corresponding to* $\xi_1 = 1$, *has an average level of contribution which is significantly lower than that of its neighbors. In* (b) *the decline in the average level of contribution around* $\xi_1 = 1$ *is further investigated. The transition from an average level of contribution of* 90% *down to* 45% *is seen to take place between* $\xi_1 = 0.97$ *and* $\xi_1 = 0.999$.

5.2 Stability Analysis

In this section the outcomes of the stability analyses described in Section 4.2 are presented. The results are illustrated using circle diagrams like the ones presented above.

In cases (1) and (2) the deviant strategy is able to penetrate the entire population; see Figures 12 and 13. In the first case, this outcome is due to a relatively high effectiveness of punishment ($\theta = 6$). This exemplifies the strength of the enforcer strategy if the punishments are severe enough, and thus the instability of the pure defector state for sufficiently high values of θ. However, as is clear from Figure 8, the outcome would be very different for lower values of the effectiveness of punishment. The second case (2) displays the weakness of the pure cooperator strategy if faced with defectors. In particular, the pure cooperator state will always be unstable since the inclusion of even a single defector may be enough to initiate a transition away from the pure state.

In case (3), the deviant defector strategy is unable to spread throughout the population of enforcers; see Figure 14. Again this is due to a sufficiently high effectiveness of punishment ($\theta = 3$). As previously asserted, the outcome depends heavily on the value of θ, with increasingly different outcomes expected for lower values of the parameter. However, this particular outcome illustrates the stability of the pure enforcer state for high enough values of the effectiveness of punishment.

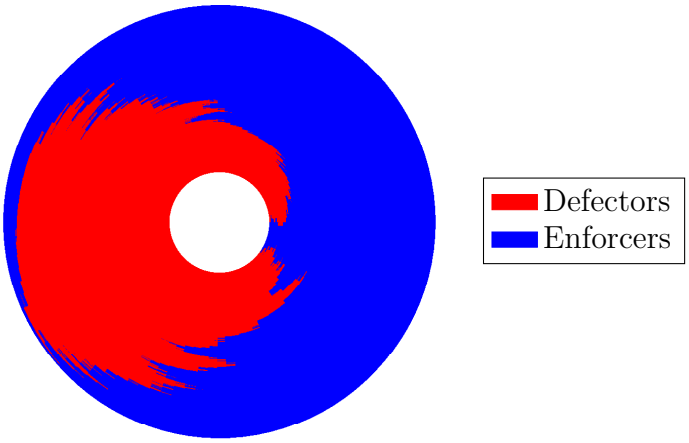


Figure 12: *Case* (1). *A minority group of enforcers are inserted into an otherwise homogeneous population of defectors. Eventually the enforcers manage to establish a norm of contribution due to a high effectiveness of punishment* ($\theta = 6$). *The simulation was run for 6,000 rounds.*

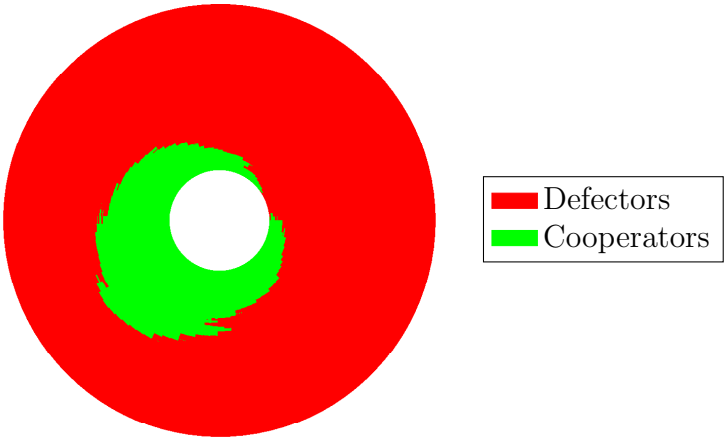


Figure 13: *Case* (2). *A minority group of defectors are inserted into an otherwise homogeneous population of cooperators. The defectors eventually eliminate the cooperators and transform the system into a non-cooperative equilibrium. The simulation was run for 3,000 rounds.*

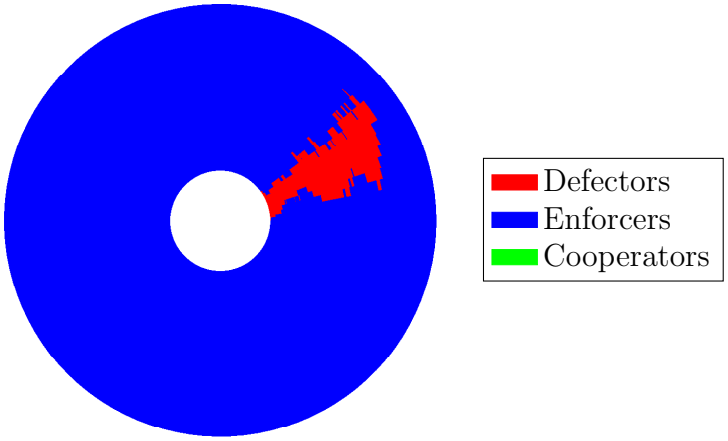


Figure 14: *Case* (3). *A minority group of defectors are inserted into an otherwise homogeneous population of enforcers. The defectors disappear entirely from the population after about 800 rounds. The simulation was run for 1,000 rounds with an effectiveness of punishment* $\theta = 3$.

5.3 Variations

Below, the results from the two variations described in Section 4.3 are presented. In the first one, an institutional structure allowing enforcers to share the costs of punishing defectors is introduced. The second variation explores how norm dependent strategies affect the outcome of the game.

5.3.1 Institutional Structure

The results from the simulations with cost-sharing are illustrated in Figure 15. The first data point ($\eta = 0$) corresponds to a game where the enforcers are able to efficiently share the costs of punishments. As can be seen, with this alteration the enforcers are able to establish a state of contribution for sufficiently low η, contrary to the case without an institution; see the result for $\theta = 2.5$ in Figure 8. As the institutional inefficiency increases the ability of the enforcers to instill a state of contribution is reduced and averages to zero for η larger than 9.

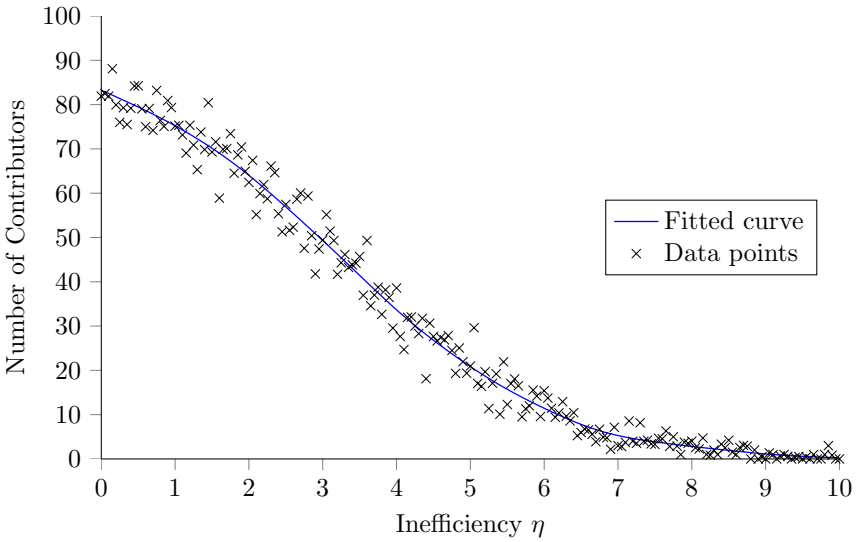


Figure 15: *Number of contributors for different values of the institutional inefficiency η for a configuration with effectiveness of punishment $\theta = 2.5$. The data points are averages over* 100 *simulations with the results recorded after 10,000 rounds.*

5.3.2 Norm Dependent Strategies

The results from this set of simulations are depicted in Figure 16. The inclusion of norm dependent enforcers and contributors is seen to have an adverse effect on the ability of enforcers to establish a state of contribution. Besides having a more outdrawn rise in the number of contributors as a function of the effectiveness of punishment, the fitted curve levels out to an average level of contribution between 85% and 90%. This can be compared with the original model in Figure 8 where the average level of contribution was at a high stable level for all θ larger than 5.

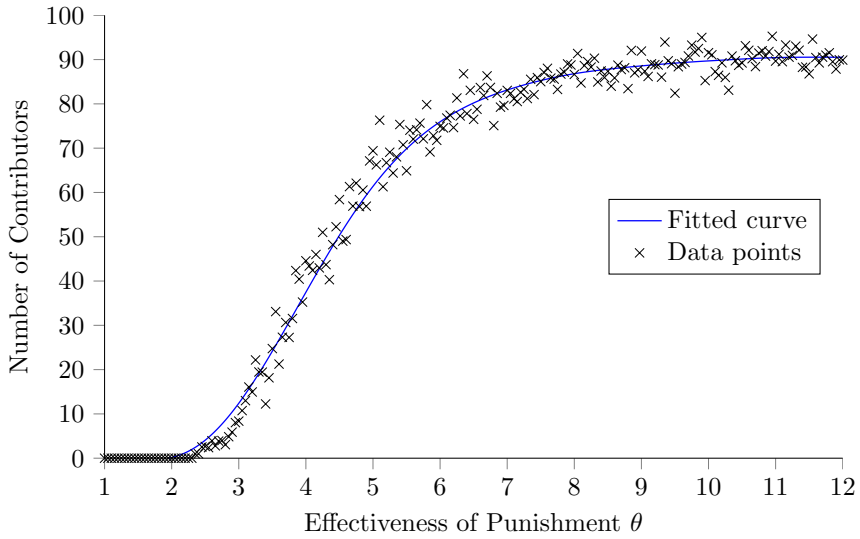


Figure 16: *Number of contributors for different values of the effectiveness of punishment θ when considering norm dependent strategies, with $(a, b) = (0, 1)$ and $(\alpha, \beta) = (1/2, 1)$. The data points are averages over* 100 *simulations with the results recorded after 10,000 rounds. The transition from a non-cooperative state to a cooperative state occurs between $\theta = 2$ and $\theta = 10$.*

6 Conclusions

This concluding section provides a discussion of the experimental results and attempts to investigate the validity of the model with respect to real world considerations. Firstly, the model mechanisms are discussed based on the results from the specific examples provided in Section 5.1 and the subsequent parameter studies of θ and ξ_1. Secondly, we compare the simulation outcomes with the results from the analytical study in Section 3.4. Thirdly, the results from the experiment involving an institutional structure are discussed. Lastly, we outline further extensions to the model and conclude the section with a summary.

6.1 Model Mechanisms

From the single simulation in Figure 5, a pattern often seen in real experiments is recognized, namely that contributions tend to decline over time [3, 13]. In our model these results stem from contributing players who change strategy due to a comparatively lower payoff. This may bear some resemblance to reality, but at the same time the important psychological factor is omitted in the model [2]. Specifically, in constructing the model one crucial simplification has been made: the strategies are independent of previous actions, and of expectations of future outcomes. In other words, the strategies in the model are unable to predict the implications of their actions, or base them on previous observations. This weakness of the model is addressed by the choice of game dynamics. In particular, the dynamics provide a mechanism for comparing the outcomes for different strategies. This creates a framework on which strategy revision can be based so that different strategies can be favored or disfavored. However, this mechanism is qualitatively not the same as the ability to value past and future outcomes in the decision making process.

Considering Figure 6, one crucial observation can be made: tougher punishments can establish a state of contribution in the game, in accordance with the experimental results in for instance [6]. The plots in Figure 7 from the same simulation also reveal a tendency of similar strategies to form clusters, as a result of players imitating their neighbors. This produces a "frontier battle" where payoffs are significantly lower at the boundaries between enforcers and defectors. These are the exact regions where punishments occur and the advances of different strategies depend directly on the relation between the cost and effect of the punishments, i.e. on the effectiveness of punishment θ. With respect to these interactions, the cooperators function as a kind of insulation: defectors who are surrounded by cooperators are protected from being punished and can peacefully enjoy their relatively higher payoffs. In this sense, cooperators are harmful since they stymie enforcement and therefore facilitate the transition to a state of non-contribution.

The parameter studies revealed that the effectiveness of punishment θ indeed is important: tougher punishment increases the probability of attaining a state of cooperation. When the willingnesses to contribute and to punish are dependent on the overall contribution level, as in Figure 16, the same pattern emerges. However, in this case it is harder for the enforcers to establish a state of contribution. Assuming that no strategy mutation takes place ($\xi_2 = 1$), the effectiveness of punishment also has a fundamental impact on the stability of the pure strategy states. For low values of θ the asymptotic behavior of the game tends toward the pure defector state. On the other hand, for sufficiently high values of θ a segment of the line between the pure enforcer state and the pure cooperator state accumulates a vaster share of the outcomes.

The weakness of the cooperator strategy was demonstrated in the stability experiments where a minority of defectors could turn a cooperative society into one of defection. On the other hand, enforcers were able to do the opposite in a non-cooperative society, which again highlights the importance of reciprocal actions in order to establish and sustain norms of contribution.

As long as the cost of being punished is higher than the combined cost of contributing and punishing, an enforcer will have an evolutionary advantage over her target, given that she successfully carries out the punishment. However, a punishment can only be carried out after a successful detection. Thus, the advantage of the enforcer strategy is diminished since information in the model is limited: an enforcer can only observe one (randomly chosen) nearest neighbor per round. Two different changes could potentially increase the strength of the enforcers: (i) increased surveillance where enforcers have more information, or (ii) a network structure where players have more than two neighbors, which implies that defectors could run a higher risk of getting punished.

The depiction of the ξ_1-dependence in Figure 11 indicates that the local-global imitation delimiter has an important impact on the dynamics in the game. When lowering the value of ξ_1 from 1, the average level of contribution is increased drastically. A possible explanation for this observation stems from the increased possibility of enforcers to be injected into clusters of defectors. For moderate levels of global imitation, the defectors who are punished by the injected enforcers will likely imitate one of their neighbors, meaning that there is a possibility that they will adopt the enforcer strategy. On the other hand, for smaller values of ξ_1 it is more likely that players revising strategy imitate random players in the game. This implies that the ability of the injected enforcers to create groups of enforcers inside clusters of defectors is reduced. However, the weakening of this ability is only gradual, potentially explaining the upward trend seen in Figure 11. Another related effect of global imitation is a more fragmented distribution of strategies. As a consequence, defectors are less protected by each other, which

causes the average number of punishments per round to increase. This in turn has an immediate negative effect on the defectors. However, the long term effects depend on the success of cooperators, as explained below.

6.2 Theoretical versus Experimental Results

When comparing the simulation results in Figure 9 with the approximate solutions in Figure 3, a number of similarities can be observed. Notably, punishments with a higher degree of severity facilitate norm enforcement. The asymptotic behavior of the model is expected to be similar to the approximation, although not entirely the same. The main difference between the two is the loss of local mechanisms such as strategy clustering, a direct consequence of the mean field approximation. Therefore, the approximation is more valid in the beginning of a simulation (with random initial distributions) but becomes increasingly invalid as strategies start to form clusters. Furthermore, there is unfortunately no bound on the approximation error, as in for example [1].

Comparing Figure 9c and Figure 3a, which were created using the same value of the effectiveness of punishment, the relative strength of enforcers versus defectors is seen to be similar in the two cases. However, the strength of the enforcer strategy is observed to be slightly lower in the mean field approximation. A reasonable explanation for this observation stems from the global mixing of players caused by the mean field approximation. The advantage depends on the effectiveness of punishment, but for certain values it is probable that global mixing could benefit defectors. One possible mechanism is due to the defectors no longer being contained, which implies that the possibility of a defector being close to an enforcer is increased. In the ensuing "all-out war" both defectors and enforcers will see a reduction in their payoffs. On the other hand, cooperators do not suffer from this struggle, meaning they will have comparatively higher payoffs. Thus, due to strategy revision and local imitation, a transition toward a state with a higher number of cooperators is likely. However, such states have an intrinsic weakness in the presence of defectors. This would alter the transition such that the number of defectors grows. As the game progresses this transition would become more rapid and in the end the state of the game settles in, or close to, the pure defector state. The diagrams in Figure 3 yield support for this notion by the presence of curved flow lines leading from mixed states, via states with a higher number of cooperators, toward the pure defector state. Since no global mixing is present in the simulated model, the above mechanism could potentially explain the comparatively weaker enforcer strategy in the mean field approximation.

6.3 Effects of an Institutional Structure

By allowing the enforcers to share the costs of punishing defectors through an institutional structure, norms of contribution could be upheld for lower values of the effectiveness of punishment. For instance, given an effectiveness of punishment $\theta = 2.5$, sharing the costs of the punishments increased the average contribution from around 10% to approximately 80%; see Figure 15. By gradually increasing the institutional inefficiency it was found that an inefficiency of up to 600% can be tolerated before the average contribution becomes 10% once more. This can be interpreted as the value of an institution, such as a police force. In this case, as long as the institution can use at least one sixth of its resources to punish defectors, its existence can be motivated. Under such circumstances it is reasonable to expect that contributors would be willing to finance the institution through for instance a voluntary tax.

6.4 Further Extensions

The simulations presented in this report represent just a fraction of the number of possible parameter configurations. Additional simulations can be carried out in order to investigate a variety of questions. There are also several possible extensions to the model that could be included for the purpose of studying different issues.

One such extension is to include budget constraints. This can be introduced by providing each player with an initial sum of credits and forbidding negative personal wealth. Implications of this include situations where enforcers are only able to act as long as they can afford to bear the costs of punishing defectors. It would also be possible for the game to simply end in a stalemate if the players who were contributing to the common pool run out of credits. One possible economic interpretation of imposing budget constraints include modeling differences between rich and poor countries. As previously mentioned, the ability of enforcers to establish or maintain norms in a society is contingent on them having enough credits available for punishment. Hence, punishers in a poor country may be willing but economically unable to bring the society into a cooperative state.

Another extension involves altering the dimension of the lattice. In a multi-dimensional setting the number of nearest neighbors would be larger, which in turn would change the punishment situation. The effect this would have, for instance on the behavior of the boundaries between different strategies, is not clear. However, based on the mean field master equation in Proposition 3.1, which is applicable also in a higher-dimensional model, the outcomes are expected to be similar. Furthermore it is also possible to envision more general network structures. This includes networks where the process of gathering information is disturbed by

some stochastic noise, or where the connections between players are allowed to change from round to round. In particular, random networks would be a desirable case to study since this would better emulate real world interactions.

6.5 Closing Remarks

The purpose of the study was to investigate the establishment of states of cooperation and requirements for attaining these in a model with limited information. The inclusion of players willing to incur personal costs in order to enforce a norm of contribution has been seen to accomplish exactly this. Furthermore, in the presence of defectors, contributors not willing to punish unwanted behavior were found to have a detrimental effect on the possibility of establishing states of cooperation. A similar negative effect was observed when contributors were unwilling to contribute or to punish due to low average levels of contribution. The results also suggested that the enforcers have an advantage when the strategy distribution is more fragmented. Lastly, the creation of an institution supervising the punishment process was shown to facilitate the enforcement process, even at considerable rates of inefficiency. This observation underlines the value of such institutions in order to maintain norms in a society. In particular, sharing the costs of punishing defectors, even inefficiently, has a substantial positive effect on the ability of a group to establish norms of cooperation.

A Analytical Treatment

In this appendix the derivation of the mean field master equation presented in Proposition 3.1 is provided.

Let the letters k and h be one of the positive integers $1, 2$ or 3 denoting strategies, with 1 corresponding to defection, 2 to enforcement and 3 to cooperation. Consider a continuous time variable t and an infinitesimal element $\Delta t > 0$. Denote by $n^k(t)$ the number of players using strategy k at time t. In this case the total number of players

$$n(t) = \sum_{k=1}^{3} n^k(t) \tag{A.1}$$

is identically equal to n since it is constant over time. The density of strategy k at time t is defined as

$$\zeta^k(t) \stackrel{\text{def}}{=} \frac{n^k(t)}{n}. \tag{A.2}$$

Furthermore, let $q_j^{k_j}(t)$ denote the probability that player j employs strategy k_j at time t. Consider player i and her nearest neighbors $i-1$ and $i+1$ as presented in Figure 2 on page 13. For this group of players:

$$q_i^{k_i}(t) = \frac{n^{k_i}(t)}{n} = \zeta^{k_i}(t), \tag{A.3a}$$

$$q_{i-1}^{k_{i-1}}(t) = \frac{n^{k_{i-1}}(t) - \delta^{k_{i-1}k_i}}{n-1}, \tag{A.3b}$$

$$q_{i+1}^{k_{i+1}}(t) = \frac{n^{k_{i+1}}(t) - \delta^{k_{i+1}k_i} - \delta^{k_{i+1}k_{i-1}}}{n-2}, \tag{A.3c}$$

where the latter probabilities are conditional with respect to the previous ones and δ^{hk} is the Kronecker delta*.

Denote by $\pi_j^{k_j}(t)$ the payoff of player j at time t if strategy k_j is employed. Using a mean field approximation† the different payoffs can be written:

$$\pi_i^1(t) = \kappa\bar{x}(t) - \theta c \cdot \frac{1}{2}\left(q_{i-1}^2(t) + q_{i+1}^2(t)\right), \tag{A.4a}$$

$$\pi_i^2(t) = \kappa\bar{x}(t) - 1 - c \cdot \frac{1}{2}\left(q_{i-1}^1(t) + q_{i+1}^1(t)\right), \tag{A.4b}$$

$$\pi_i^3(t) = \kappa\bar{x}(t) - 1, \tag{A.4c}$$

*The Kronecker delta is defined as $\delta^{hk} = 1$ if $h = k$ and zero otherwise.

†We treat player i and the interactions with her nearest neighbors exactly while each nearest neighbor is treated approximatively, cf. the Bethe approximation in statistical physics [16].

where the factor 1/2 stems from the equal probability that the player looks in either direction. Since only strategies 2 and 3 contribute we have

$$\bar{x}(t) = \zeta^2(t) + \zeta^3(t), \tag{A.5}$$

and assuming that each strategy has a large number of adhering players, i.e. $n^k(t) \gg 1$ for $k = 1, 2, 3$, we can make the approximations:

$$q_{i-1}^{k_{i-1}}(t) \approx \zeta^{k_{i-1}}(t), \quad \text{and} \quad q_{i+1}^{k_{i+1}}(t) \approx \zeta^{k_{i+1}}(t), \tag{A.6}$$

for the probabilities in (A.3). Hence, in the mean field approximation, the payoffs in (A.4) are characteristic for each respective strategy and can be expressed as:

$$\pi^1(t) = \kappa\left(\zeta^2(t) + \zeta^3(t)\right) - \theta c\zeta^2(t), \tag{A.7a}$$
$$\pi^2(t) = \kappa\left(\zeta^2(t) + \zeta^3(t)\right) - 1 - c\zeta^1(t), \tag{A.7b}$$
$$\pi^3(t) = \kappa\left(\zeta^2(t) + \zeta^3(t)\right) - 1. \tag{A.7c}$$

If player j revises strategy, she can do so in three different ways. Either she imitates another player (locally or globally) or mutates, the latter meaning that a random strategy is selected with probability 1/3. These three options are related via the parameters ξ_1 and ξ_2. Denote by $r_j^{h\to k}(t)$ the transition probability that player j with strategy h, if selected for revision, adopts strategy k. Then,

$$r_i^{h\to k}(t) = \xi_1 \cdot \frac{1}{2}\left(q_{i-1}^k(t) + q_{i+1}^k(t)\right) + (\xi_2 - \xi_1)\cdot\zeta^k(t) + (1-\xi_2)\cdot\frac{1}{3}. \tag{A.8}$$

Under the approximation in (A.6), this can be written

$$r^{h\to k}(t) \approx \xi_1\zeta^k(t) + (\xi_2 - \xi_1)\zeta^k(t) + \frac{1}{3}(1-\xi_2) = \xi_2\zeta^k(t) + \frac{1}{3}(1-\xi_2), \tag{A.9}$$

independent of the player. The number of players employing strategy k at a later time $t + \Delta t$ can therefore be expressed as*

$$n^k(t+\Delta t) \approx \left(1 - \Delta t\cdot p^k(t)\right) n^k(t) + \Delta t \cdot \sum_{h=1}^{3} r^{h\to k}(t)p^h(t)n^h(t). \tag{A.10}$$

In the above equation, $p^k(t)$ is the probability that a player employing strategy k

*To realize this, observe that the fraction of players employing strategy k who are selected for strategy revision between the times t and $t+\Delta t$ is given by $\int_t^{t+\Delta t} p^k(\tau)\,\mathrm{d}\tau \approx \Delta t\cdot p^k(t)$. The fraction of players with strategy k not revising strategy is therefore approximately $1 - \Delta t\cdot p^k(t)$.

revises strategy and is given by

$$p^k(t) = \lambda e^{-\lambda y^k(t)}, \tag{A.11}$$

where the relative payoffs $y^k(t)$ are defined as

$$y^k(t) = 2 \cdot \frac{\pi^k(t) - \pi_{\min}(t)}{\pi_{\max}(t) - \pi_{\min}(t)}; \tag{A.12}$$

compare with (3.5) and (3.6).

Rearranging the terms in (A.10), dividing by Δt and using the approximation in (A.9) yields

$$\frac{n^k(t+\Delta t) - n^k(t)}{\Delta t} \approx -p^k(t)n^k(t) + \sum_{h=1}^{3} \left(\xi_2 \zeta^k(t) + \frac{1}{3}(1-\xi_2) \right) p^h(t) n^h(t). \tag{A.13}$$

Taking the limit $\Delta t \searrow 0$, the left hand side in (A.13) becomes the time derivative $\dot{n}^k(t)$. Using the relation $\dot{n}^k(t) = n\dot{\zeta}^k(t)$ we have derived a master equation [16] for the stochastic variables $\zeta^k(t)$ in the model:

$$\dot{\zeta}^k(t) \approx -p^k(t)\zeta^k(t) + \sum_{h=1}^{3} \left(\xi_2 \zeta^k(t) + \frac{1}{3}(1-\xi_2) \right) p^h(t) \zeta^h(t). \tag{A.14}$$

The derivation can be put on a more stable ground by starting out from a more general master equation and thereafter making the necessary approximations; see for instance Chapter 8 in [16] for a general presentation of the master equation and its use when considering stochastic processes. However, the (more heuristic) arguments in the above derivation have the benefit of illustrating the features of the different components in (A.14) in a way that might otherwise be lost. Moreover, the generalization to a d-dimensional cubic lattice ($d = 1, 2, \ldots$) is straightforward using the same types of approximations as in the derivation for the one-dimensional case. The result is the same mean field master equation as in (A.14), which thus generalizes to a d-dimensional setting as well.

B Model Implementation

In this appendix, some issues concerning the implementation of the model in order to run numerical simulations are presented. The implementation itself was done in C++ programming language since it allows for object-oriented programming. This essentially means that the players in the game can be generated as objects and thus treated, more or less, as actual players with intrinsic properties in the form of their choice of strategy.

B.1 Model Complexity and Oscillations

The model presented in Section 3 has a favorable computational complexity when implemented on a computer. Since player-to-player interactions occur sparsely, the complexity is linear both in number of rounds and number of players. This allows for the possibility to run fairly large simulations.

A potential issue with the kind of simulations used in this study is the introduction of oscillations into the system. This can for instance occur if too many players change strategies in each round, thereby causing the system to oscillate around an equilibrium point. To overcome this problem the dynamics can be slowed down, in this case for example by adjusting the number of revision draws n_R in each round. To ensure that a simulation does not suffer from such oscillations, its state should be measured at different points in time, which are not multiples of each other.

B.2 Ergodicity and Randomness

A similar issue to the oscillation problem mentioned above is ergodicity and the level of randomness present in the system. A system is ergodic if all states are reachable from each other and no periodic paths exist. If these conditions are fulfilled, any simulation will eventually converge to the ergodic state, which is a desirable property. For a system containing multiple stable states, a simulation may converge to different equilibria due to the inherent randomness present in the model. A certain level of stochasticity is required in order to simplify the model, but these negative effects have to be taken into account when evaluating the results. Hence, the results should be validated by running many experiments of the same kind, thereby minimizing the risk of drawing conclusions based on low-probable outcomes.

References

[1] Benaim, M., and Weibull, J. W. Deterministic approximation of stochastic evolution in games. *Econometrica 71*, 3 (2003), 873–903.

[2] Bohnet, I., and Frey, B. S. The sound of silence in prisoner's dilemma and dictator games. *Journal of Economic Behavior & Organization 38*, 1 (1999), 43–57.

[3] Burlando, R., and Guala, F. Heterogeneous agents in public goods experiments. *Experimental Economics 8*, 1 (2005), 35–54.

[4] Castellano, C., Fortunato, S., and Loreto, V. Statistical physics of social dynamics. *Reviews of Modern Physics 81*, 2 (2009), 591–646.

[5] Ellingsen, T., Herrmann, B., Nowak, M. A., Rand, D. G., and Tarnita, C. E. Civic capital in two cultures: The nature of cooperation in Romania and USA. Working paper, 2012.

[6] Fehr, E., and Gächter, S. Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives 14*, 3 (2000), 159–181.

[7] Fehr, E., and Gachter, S. Altruistic punishment in humans. *Nature 415*, 6868 (2002), 137–140.

[8] Gintis, H. Strong reciprocity and human sociality. *Journal of Theoretical Biology 206*, 2 (2000), 169–179.

[9] Hauert, C., and Szabó, G. Game theory and physics. *American Journal of Physics 73*, 5 (2005), 405–414.

[10] Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Henrich, N., Hill, K., White, F., Gurven, M., Marlowe, F., Patton, J., and Tracer, D. "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences 28*, 06 (2005), 795–815.

[11] Herrmann, B., Thöni, C., and Gächter, S. Antisocial punishment across societies. *Science 319*, 5868 (2008), 1362–1367.

[12] Nowak, M. A. Five rules for the evolution of cooperation. *Science 314*, 5805 (2006), 1560–1563.

[13] Ostrom, E. Collective action and the evolution of social norms. *Journal of Economic Perspectives 14*, 3 (2000), 137–158.

[14] Peters, H. *Game Theory: A Multi-Leveled Approach*, 1st ed. Springer, 2008.

[15] Pindyck, R. S., and Rubinfeld, D. L. *Microeconomics*, 7th ed. Pearson Education, 2009.

[16] Plischke, M., and Bergersen, B. *Equilibrium Statistical Physics*, 3rd ed. World Scientific, 2006.

[17] Rand, D., Dreber, A., Ellingsen, T., Fudenberg, D., and Nowak, M. A. Positive interactions promote public cooperation. *Science 325*, 5945 (2009), 1272–1275.

[18] Sethi, R., and Somanathan. The evolution of social norms in common property resource use. *The American Economic Review 86*, 4 (1996), 766–788.

[19] Sigmund, K., Silva, H. D., Traulsen, A., and Hauert, C. Social learning promotes institutions for governing the commons. *Nature 466* (2010), 861–863.

[20] Szabó, G., and Hauert, C. Phase transitions and volunteering in spatial public goods games. *Physical Review Letters 89*, 11 (2002), 118101.

[21] Szolnoki, A., Szabó, G., and Perc, M. Phase diagrams for the spatial public goods game with pool punishment. *Physical Review E 83*, 3 (2011), 036101.

[22] Taylor, P. D., and Jonker, L. B. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences 40*, 1 (1978), 145–156.

[23] Weibull, J. W., and Salomonsson, M. Natural selection and social preferences. *Journal of Theoretical Biology 239* (2006), 79–92.