universität
wien

DISSERTATION

Titel der Dissertation

Public Good Games With Incentives

Verfasser

Christian Hilbe

angestrebter akademischer Grad

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien, im März 2011

Studienkennzahl lt. Studienblatt:	A 091 405
Dissertationsgebiet lt. Studienblatt:	Mathematik
Betreuer:	Univ.-Prof. Karl Sigmund

Acknowledgements

First of all, I would like to thank my supervisor Prof. Karl Sigmund. He introduced me to the fascinating world of science in general and to the world of evolutionary game theory in particular. His passion for the subject and his support for his students is second to none. He opened doors I did not even know they existed.

I would also like to thank the Center of Rationality at the Hebrew University, Jerusalem, for hosting me for five exciting months and the IIASA, Laxenburg, for inviting me to their summer program. Both were wonderful and instructive experiences. Moreover, I would like to give thanks to Prof. Sergiu Hart, Prof. Josef Hofbauer, Dr. Arne Traulsen, Dr. Åke Brännström and Dr. Ulf Dieckmann for valuable advices and discussions.

It is needless to say that this thesis would not exist without the help, the support and the constant coffee supply of my friends and my family. Thank you, Renate, Andreas, Thomas, Claudia, Pia, Johannes, Verena, Markus, Anna, Bernhard, Stephan, Thomas, Ulrich, Alban, Susanne, Marianne, Katharina, Satoshi, Hannes, Katrin, Hildegard, Marieke, Ilan, Aviv, Einav, Anat, Tali, Gali, Yoni, Uri, Ziv, Michele, Ed, Regina, Hubert, Claudia, Florian, Gregor, and – above all – thank you, Beate.

Abstract

Public good games reflect the common conflict between group interest and self interest: While collaborative projects depend on joint efforts of all participants, each individual performs best by free-riding on the others' contributions. In this thesis I use evolutionary game dynamics to study the interplay of cooperation and incentives. I demonstrate that rewards may act as a catalyst for individual contributions, while punishment is needed to maintain mutual cooperation. In this process, reputation plays a key role: It helps to mitigate the second-order free-rider problem and prevents subjects from being spiteful. In contrast to previous studies, I do not find that punishment can promote any behaviour (as long as deviations from that norm are punished). Instead, sanctions are targeted at non-cooperators only, and lead to stable cooperation.

Furthermore, this thesis provides some mathematical tools for the study of public good games with incentives. It extends the theory of role games and it introduces a modified replicator dynamcis that allows to investigate the consequences of local competition. Under this local replicator dynamics, even dominated strategies may prevail if they lead to a relative payoff advantage – which can be considered as a basic model for the evolution of spite.

Contents

Introduction

I find it useful to distinguish between two different branches of evolutionary game theory. The first branch has its focus on evolution itself and the relevant question under this perspective are:

How can we describe learning processes in strategic environments mathematically?
What can we say about the properties of such learning rules?

Research articles that have dealt with this kind of questions include the papers on fictitious play (Robinson, 1951), the best response dynamics (Matsui, 1992) and the replicator dynamics (Taylor and Jonker, 1978). In recent years, researchers are also increasingly interested in the converse type of problem:

Are there learning rules that have a specific property?

Typical examples include the question whether learning processes for sequential games can lead to the backward induction solution (for a positive answer see Hart, 2002) or whether there is a reasonable evolutionary dynamics that leads to the extinction of dominated strategies (for a negative answer, see Hofbauer and Sandholm, in press).

The second approach has its focus on the application of evolutionary game theory to relevant economic or biological settings. One of the pioneering research articles of this kind is also considered to be the birth of evolutionary game theory: *The logic of animal conflict* of Maynard Smith and Price (1973) provides a rationale for ritual fighting in deers. Maybe the most popular application of evolutionary game theory deals with the evolution of cooperation (Axelrod, 1984), considering the following type of questions:

Why do we find so much cooperation in nature?
Which mechanisms enhance individual cooperativeness?

The first question is motivated by several empirical findings: Vampire bats regurgitate blood to one another to increase the conspecific's chances of survival (Wilkinson, 1984) and in cleaning symbioses, host fishes refrain from eating their cleaner fishes even if cleaning is over (see Trivers, 1971). Since evolutionary forces are predicted to select against altruistic traits, scientists are interested in the mechanisms that uphold cooperation (see for example Nowak, 2006b).

The title of the thesis suggests that this monograph belongs to the second category of evolutionary game theory and indeed, Chapter 2 and 4 investigate the question whether positive and negative incentives can promote cooperative behaviour. However, in order to respond to these questions, I needed to develop some mathematical tools that rather belong to the first branch of evolutionary game theory and these tools are presented as Chapter 1 and 3. At first sight, Chapter 5 seems to be a bit out of place because it neither deals with the evolution of cooperation, nor does it provide a mathematical framework to respond to such questions. Instead, this chapter emerged as a product of my research on this thesis. When I tried to apply standard tools of evolutionary game theory to public good games with incentives, I got several counterintuitive results that left me puzzled. Chapter 5 is my attempt to resolve this puzzle.

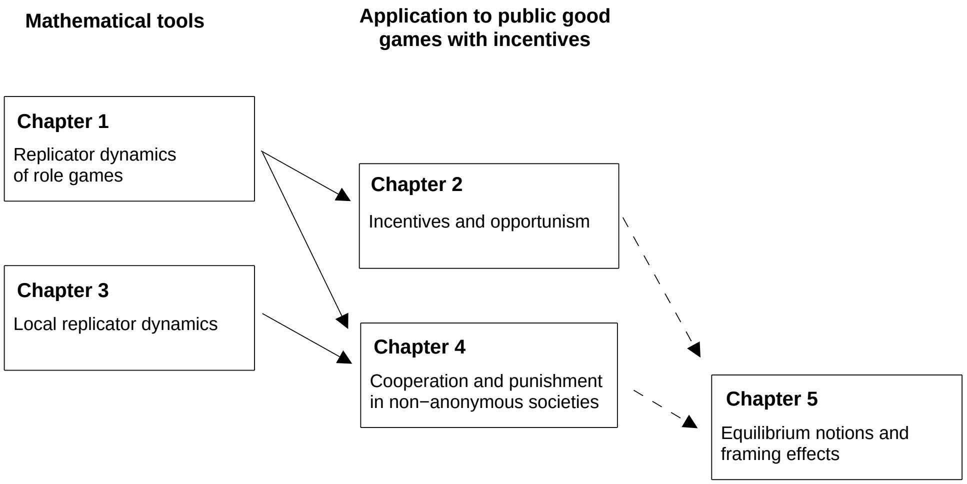


Figure 1: A schematic overview of the thesis

Public good games with incentives

Microeconomics defines public goods as goods that are *nonrival* and *non-excludable*. Due to non-rivalry, the consumption of the good by one individual does not affect the consumption possibilities of all others and due to non-excludability, no one can be precluded from using the good. Especially because of the latter property, public goods are vulnerable to the free-rider problem: Since non-contributors cannot be excluded, there is no incentive to engage in the provision of public goods. This may lead to the overuse of common resources such as fishing grounds (Hardin, 1968), to delayed actions against climate change (e.g. Milinski et al., 2006) or to non-compliance to social norms (see Fehr and Gächter, 2000).

In order to test these predictions in the laboratory, behavioural economists have designed a simple experimental situation that mimics the incentive structure of public goods. Before the experiment starts, each of the four participants is endowed with a starting capital C. Then the subjects may decide anonymously which amount of their money they want to spend to a public pool, knowing that all contributions to the pool are doubled and shared equally among the participants. A theoretical analysis of this decision problem is straightforward: If the contributions of a player are denoted by x_i, then the player's payoff π_i can be calculated as

$$\pi_i = C - x_i + \frac{1}{4}\sum_{k=1}^{4} 2x_k \tag{1}$$

The optimal strategy of a player depends on the player's target function. A player i aiming to maximize the group payoff $\sum_{j=1} \pi_j$ may compute

$$\frac{\partial}{\partial x_i}\big(\sum_{j=1} \pi_j\big) = 1 > 0. \tag{2}$$

Therefore, the group optimum is attained if all players contribute their full starting capital C to the public pool (in which case all players obtain a payoff of $2C$). On the other hand, a player aiming to maximize the individual payoff π_i needs to calculate

$$\frac{\partial \pi_i}{\partial x_i} = -1/2 < 0. \tag{3}$$

Consequently, the individual optimum is attained if the player withholds all contributions. However, in a group of self-interested subjects, each player only yields the starting capital C instead of the *pareto efficient* outcome of $2C$. Public good games thus describe a situation where individual interests are at odds with group welfare – a *social dilemma*. In the laboratory, the result of such a game may be affected by a variety of parameters (for a survey, see Ledyard, 1995): In fact, contributions depend on the starting capital, group size, communication possibilities, whether the public pool is doubled or tripled or on the (economic) education and gender of the participants. As a robust finding, most of the subjects start by contributing around half of their endowment; however, if the public good game is repeated for several rounds, the players learn to behave selfishly. In the last round, 73% of all players contribute nothing (see Fehr and Schmidt, 1999).

Such a negative outcome can be prevented if contributions are incentivized. Indeed, if we assume that contributors yield subsidies that are proportional to their contributions, we obtain the following payoff function:

$$\pi_i = C - (1-r)x_i + \frac{1}{4}\sum_{k=1}^{4} 2x_k, \tag{4}$$

where r denotes the subsidy rate. According to this modified payoff function, self-interested individuals will contribute their full endowment if

$$\frac{\partial \pi_i}{\partial x_i} = r - 1/2 > 0. \tag{5}$$

Incentives thus may solve the social dilemma; however, without proper institutions providing these subsidies, it is unclear how such an incentive scheme can be implemented. In a path-breaking experiment, Fehr and Gächter (2000) have shown that no exterior institution is needed, since subjects seem to be able to govern themselves. In this experiment, players had the opportunity to punish their co-players after the public good game. It was shown that players are willing to punish free-riders even if any personal long-run benefit for the punisher is excluded by the experimental design. As a consequence, contributions to the public good were significantly higher if sanctioning mechanisms were available.

From a game-theoretic perspective, this result is puzzling: Since punishment is costly, there is no obvious incentive for the punisher to engage in retribution. A theoretical model of the co-evolution of cooperation and punishment must therefore explain:

1) Which mechanisms ensure the maintenance of a punishment regime?

This question addresses a second-order free-rider problem: Since a sanctioning mechanism is costly, but beneficial for the group as a whole, it is a public good by itself. How can players be incentivized to engage in the punishment of free-riders? This problem is intensified in groups that fully consist of free-riders:

2) Which mechanisms promote the implementation of a punishment regime in a population of free-riders?

Finally, it needs to be clarified how a punishment regime can be controlled since several experiments indicate that a non-negligible fraction of subjects abuse sanctioning opportunities for spite (see, for example Herrmann et al., 2008; Nikiforakis, 2008).

3) Which mechanisms prevent the evolution of anti-social punishment?

In the last ten years, several evolutionary explanations for the emergence of punishment have been proposed – a short (and non-representative) survey is shown in Tab. 1. While there are numerous mechanisms that guarantee the fixation of punishment, it seems to be more difficult to identify mechanisms that allow the initial implementation of a sanctioning regime. Most surprisingly, there is only one study (Rand et al., 2010) that allows for anti-social punishment, and there the co-evolution of cooperation and punishment is fully suppressed if subjects have the opportunity to be spiteful.

It is the aim of this thesis to demonstrate how non-anonymity can resolve all these three questions. In the model presented in Chapter 2 and Chapter 4, individuals adapt to the reputation of their co-players and learn to cooperate against punishers. This in turn provides an incentive to engage in costly punishment when necessary. It is shown in Chapter 2 that in a population of non-cooperators, rewards may promote the emergence of an efficient punishment regime when punishment alone cannot establish stable cooperation. In Chapter 4 we investigate how reputation can prevent subjects from abusing sanctions for spite and anti-social punishment.

In order to simplify the analysis, we consider public good games between two players only. Such games can be represented as a role game, where players may have different actions depending on their respective role in the current interaction. Role games are formally introduced in Chapter 1, where we also discuss the basic properties of the replicator dynamics of such games. Chapter 3 investigates the influence of local competition on the evolution of strategies. For this reason we derive a modified replicator dynamics, assuming that players interact in finite groups of size N. According to this local replicator dynamics, dominated strategies do not necessarily go extinct. Instead, a dominated strategy may spread among the population if it gives a relative advantage to the focal player. One may therefore interpret the local replicator dynamics as a model for the evolution of spite.

When I applied the Moran process (Nowak et al., 2004) to public good games with incentives, I noticed a remarkable framing effect: Adding a new strategy to a game may affect the game's result even if this new strategy is identical to a strategy that was already

Article	**Mechanism promoting the invasion of punishment**	**Mechanism promoting the fixation of punishment**	**Antisocial punishment included?**
Boyd and Richerson (1992) *Punishment allows the evolution of cooperation (or anything else) in sizeable groups*	In repeated interactions, punishment may invade if it makes the co-players more cooperative	Higher-order punishment	Not explicitly
Gintis (2000) *Strong reciprocity and human sociality*	Punishment cannot invade when rare	Cooperative groups play more rounds of the PGG and hence obtain a higher payoff	No
Sigmund et al. (2001) *Reward and Punishment*	Punishment cannot invade when rare	Due to reputation, subjects are more likely to cooperate against punishers	No
Boyd et al. (2003) *The evolution of altruistic punishment*	Punishment cannot invade when sufficiently rare (within their group)	Group selection: Cooperative groups are more likely to replace other groups	No
Gardner and West (2004a) *Cooperation and punishment, especially in humans*	Punishment can invade if there is a correlation between punishment and received cooperation	Punishment is stable if there is a correlation between punishment and received cooperation	No
Fowler (2005) *Altruistic punishment and the origin of cooperation*	Voluntary participation	Second-order punishment	No
Nakamaru and Iwasa (2006) *The coevolution of altruism and punishment: Role of the selfish punisher*	In finite populations, severe punishment leads to a relative fitness advantage	In finite populations, severe punishment leads to a relative fitness advantage	No
Hauert et al. (2007) *Via freedom to coercion: The emergence of costly punishment*	Voluntary participation	Due to stochastic fluctuations, no strategy is stable	No
Eldakar and Wilson (2008) *Selfishness as second-order altruism*	Punishment cannot invade when rare	Coexistence of non-punishing cooperators and punishing defectors in finite groups	No
Boyd et al. (2010) *Coordinated punishment of defectors sustains cooperation and can proliferate when rare*	Only possible if there is some assortment in group formation	Punishers coordinate their activities, punishment costs decrease overproportionally with the number of punishers	No
Rand et al. (2010) *Anti-social punishment can prevent the co-evolution of punishment and cooperation*	Punishment cannot invade	Punishment is not stable	Yes

Table 1: A short history of evolutionary models on the co-evolution of cooperation and punishment

available to the players. Chapter 5 discusses solution concepts with this kind of framing effect and gives an unexpected condition for such inconsistencies: If a game-theoretic solution concept depends differentiably on the payoffs, then it also depends on the representation of the game.

Chapter 1

The replicator dynamics of role games

Abstract

In a role game, two players interact in an asymmetric game with both players having equal chance to be the row-player or the column-player. After reviewing previous results on role games and their replicator dynamics, we proceed by analyzing the special case where the row-player can choose among three pure actions and the column-player among two. We identify cases where the dynamics degenerates to the dynamics of a reduced game, but we also give a numerical example where the system exhibits non-trivial cycles and where strategies are sustained in the population that are not part of any Nash equilibrium of the game. In addition, we prove a general result for mixed equilibria in non-generic two-player games: If the row-player randomizes between k pure actions, then so does the column-player.

1.1 Introduction

In game theory, it is common to distinguish between symmetric and asymmetric games. In symmetric games, players are assumed to be undistinguishable and all agents have the same set of available strategies and face the same monetary consequences, as for example in the prisoner's dilemma. In asymmetric games, on the other hand, strategies and feasible payoffs may differ across agents. For example, traders have different strategies than their customers, and the possible payoffs of a hunting predator differ from the payoffs of its prey.

Role games are somewhat between these two categories (Gaunersdorfer et al., 1991; Hofbauer and Sigmund, 1998; Berger, 2001). While the actual game itself (the so-called base game) is generally asymmetric and the payoffs and strategies depend on the player's role in the game, it is assumed that each player has equal chances to act in a given role, making the interaction implicitly symmetric. For example, in online trading systems such as ebay, a subject may be in the role of the seller at one time and in the role of the buyer at another time. Similarly, in innerspecific territory conflicts (Maynard Smith, 1982), an animal may sometimes be in the role of the owner and sometimes in the role of the intruder. In such games, players can condition their strategies on their present role.

The prime example for a role game is the ultimatum game (Güth et al., 1982), which was one of the first experimental games demonstrating the irrelevance of backward induction in human decision making. The protocol of the ultimatum game assumes that subjects are assigned in pairs. In each pair, one of the subjects is randomly determined to be in the role of the proposer, whereas the other subject is in the role of the responder. The proposer is asked to make an offer how to split a certain amount of money between them. The responder can then either accept this split or reject it, in which case both subjects obtain nothing. Although payoff-maximizing responders should accept any positive amount, offers below 20 % of the total are frequently rejected (see for example Camerer and Fehr, 2006). There have been several evolutionary approaches to rationalize this outcome. Nowak et al. (2000) interpret this result as the subject's attempt to build up a strict reputation (which can be beneficial in future encounters). Such a behaviour is, of course, ill-suited for the anonymous conditions in the laboratory, but it might have played an important role in our evolutionary history. Gale et al. (1995) explain the deviation from backward induction by introducing noise in the decision making process.[1] If

---

[1] Note, however, that only Nowak et al. (2000) treat the ultimatum game as a role game where players may find themselves in both roles. Gale et al. (1995), on the other hand, interpret the game as an asymmetric game with two distinct populations, a population of proposers and and a population of responders.

responders are more often subject to noise, then replicator dynamics can lead to a Nash equilibrium that predicts zero-offers to be rejected.

The replicator dynamics of role games was also used to explain cooperative behaviour under the presence of incentives (Sigmund et al., 2001; De Silva and Sigmund, 2009). In these studies, the player in the role of the donor needs to decide whether or not to transfer a benefit b to the recipient (at own costs c). If the donor refuses to cooperate, the recipient may engage in punishment, by reducing the donor's payoffs by some amount β. Recently, this cooperation game was extended to include the possibility of rewards (Hilbe and Sigmund, 2010). This additional option increases the recipient's strategy space and therefore requires to study the dynamics of higher dimensional systems. It is therefore the aim of this chapter to extend the general theory of role games. As a special case, we study the dynamics of games where one player may choose among two actions (for example, to cooperate or to defect) and where the other player can choose among three actions (to punish a defector, to reward a cooperator or to do nothing). We find that in most cases, the replicator dynamics of such a role game degenerates to a game of lower dimension (where at least one pure action of the subjects is never used). However, we also give a numerical example for a game where all actions persist in the population, even those that are not part of any Nash equilibrium of the game.

Parts of our analysis of role games will somewhat resemble the studies of Cressman et al. (2000) and Chamberland and Cressman (2000) on the dynamics of games with two independent decisions. In their model, two subjects play one of two possible symmetric games against each other, with nature determining randomly which game they play. While the players in role games condition their actions on their present role in an asymmetric game, the subjects in Cressman et al. (2000) and Chamberland and Cressman (2000) can condition their strategies on the symmetric game that was chosen by nature. The mathematical treatment of these two distinct approaches is quite similiar and we will transfer some of their results to our setting.

The remainder of this chapter is organized as follows: In the next section, we introduce role games formally and study their Nash equilibria. As a corollary, we prove that equilibria for asymmetric games require the two players to randomize between the same number of strategies. Thereafter, in Section 1.3, we define the replicator equation of a role game and establish some properties of the resulting dynamics. These results are then applied to the case of 2×2-games (Section 1.4) and 3×2-games (Section 1.5). In Section 1.6, we generalize the replicator equation to study the effect of recombination and in the last section, we give a brief discussion of our results.

1.2 Role games and equilibrium behaviour

Role games can be described as follows: Before the game starts, one of the two players is randomly determined to be the row-player, whereas the other player is determined to be the column-player. Thereafter, both subjects play a simultaneous game, which will be called the base game. The row-player may choose an action among the finite set $\{e_1, \dots, e_m\}$ and the column-player may choose among the finite set $\{f_1, \dots, f_n\}$. Then, depending on their respective choices, each player yields a payoff that is given by the following bimatrix (A, B). As usual, the first entry in the bimatrix denotes the payoff of the row-player, whereas the second entry gives the payoff for the column-player.

$$\left(\begin{array}{c|ccc} & f_1 & \dots & f_n \\ \hline e_1 & (a_{11}, b_{11}) & \dots & (a_{1n}, b_{1n}) \\ \vdots & \vdots & \ddots & \vdots \\ e_m & (a_{m1}, b_{m1}) & \dots & (a_{mn}, b_{mn}) \end{array}\right) \tag{1.1}$$

In the base game, players are allowed to randomize between their pure actions. Let $p \in \Delta^m$ denote a probability vector, where p_i is the probability that the row-player chooses the pure action e_i. Let $q \in \Delta^n$ be the corresponding probability vector for the column-player. In this case, the row-player's payoff can be calculated as $p \cdot Aq$, whereas the column-player obtains $p \cdot Bq$.

In the role game, each player needs to specify how to act in each role. Therefore, a pure strategy of the role game is a pair $[e_i, f_j]$. A mixed strategy for the role game is represented by a matrix $(x_{ij}) \in \Delta^{mn}$. There is a natural correspondence between mixed strategies x in the role game and pairs of mixed actions $[p(x), q(x)]$ in the base game, by defining $p_i(x) = \sum_j^n x_{ij}$ and $q_j(x) = \sum_i^m x_{ij}$. This means that $p(x)$ and $q(x)$ are just the marginal probabilities of x. Of course, this correspondence is not injective; however, it can be easily seen that the correspondence is surjective, for example by defining $x_{ij} = p_i \cdot q_j$. In order to specify payoffs for the role game, we assume that each player has equal chance to be in one of the two roles. Then, a subject with strategy x, interacting with a co-player with strategy y yields the payoff

$$\pi(x, y) := \frac{1}{2} p(x) \cdot Aq(y) + \frac{1}{2} p(y) \cdot Bq(x). \tag{1.2}$$

In the special case that both players use pure strategies, $x = [e_i, f_j]$ and $y = [e_k, f_l]$, expression (1.2) simplifies to

$$\pi_{ij,kl} := \frac{1}{2}a_{il} + \frac{1}{2}b_{kj}. \tag{1.3}$$

In this section, we want to examine equilibrium behaviour in role games. It is well-known that the Nash equilibria of a role game are strongly related to the Nash equilibria of the underlying base game (see for example Berger (2001)). The proof of the following Proposition follows the lines of Sigmund (2010).

Proposition 1 (Nash equilibria in role games)
The mixed strategy $\hat{x}$ is a symmetric Nash equilibrium of the role game if and only if $[\hat{p}, \hat{q}] = [p(\hat{x}), q(\hat{x})]$ is a Nash equilibrium of the base game.

Proof. Suppose $\hat{x}$ is a symmetric Nash equilibrium of the role game. This means that $\pi(y, x) \leq \pi(x, x)$ for all $y \in \Delta^{mn}$. By Eq. (1.2) this implies

$$p(y) \cdot A\hat{q} + \hat{p} \cdot Bq(y) \leq \hat{p} \cdot A\hat{q} + \hat{p} \cdot B\hat{q} \quad \text{for all } y \in \Delta^{mn} \tag{1.4}$$

Because the map $y \mapsto [p(y), q(y)]$ is surjective, it follows that

$$p \cdot A\hat{q} + \hat{p} \cdot Bq \leq \hat{p} \cdot A\hat{q} + \hat{p} \cdot B\hat{q} \quad \text{for all } p \in \Delta^{m}, q \in \Delta^{n} \tag{1.5}$$

For $p = \hat{p}$, this yields

$$\hat{p} \cdot Bq \leq \hat{p} \cdot B\hat{q} \quad \text{for all } q \in \Delta^{n}, \tag{1.6}$$

whereas $q = \hat{q}$ leads to

$$p \cdot A\hat{q} \leq \hat{p} \cdot A\hat{q} \quad \text{for all } p \in \Delta^{m}. \tag{1.7}$$

Hence, $\hat{p}$ is a best reply to $\hat{q}$ in the underlying base game, and vice versa. To show the converse direction, suppose that $\hat{x}$ is not a symmetric Nash equilibrium of the role game. This means that there is a $y \in \Delta^{mn}$ such that $\pi(y, x) > \pi(x, x)$, or again by Eq. (1.2):

$$p(y) \cdot A\hat{q} + \hat{p} \cdot Bq(y) > \hat{p} \cdot A\hat{q} + \hat{p} \cdot B\hat{q} \tag{1.8}$$

In this case, it follows that either $p(y) \cdot A\hat{q} > \hat{p} \cdot A\hat{q}$ or that $\hat{p} \cdot Bq(y) > \hat{p} \cdot B\hat{q}$, which proves that $[\hat{p}, \hat{q}]$ is not a Nash equilibrium of the underlying base game. $\square$

This Proposition can be used to prove existence of Nash equilibria for general 2-player games. For example, in Sigmund (2010) it is shown that every symmetric 2-player game has a symmetric equilibrium. If this result is applied to a role game it follows by Proposition 1, that also the (potentially asymmetric) base game has a Nash equilibrium.

We next want to study under which conditions we may expect Nash equilibria where both players make use of all their strategies. For a vector v we write $v > 0$ if all entries $v_i > 0$. Furthermore, we say that a game (A, B) is *generic* if the rows of A and the columns of B are linearly independent. It is easily seen that the set of generic games has full measure with respect to the Lebesgue measure on $\mathbb{R}^{mn} \times \mathbb{R}^{mn}$. The following Proposition is, to the best of my knowledge, new.

Proposition 2 (A necessary condition for a Nash equilibrium with full support)
Let $[\hat{p}, \hat{q}] \in \Delta^m \times \Delta^n$ *be a Nash equilibrium for a generic game* (A, B)*. If* $\hat{p} > 0$ *and* $\hat{q} > 0$ *then* $m = n$.[2]

Proof. Assume to the contrary that (without loss of generality) $m > n$. Because $[\hat{p}, \hat{q}]$ is a Nash equilibrium with full support, it follows that each of the row-player's strategies yields the same payoff:

$$(A\hat{q})_i - (A\hat{q})_1 = \sum_{l=1}^{n} (a_{il} - a_{1l}) \cdot \hat{q}_l = 0 \qquad \text{for all } \; 2 \leq i \leq m. \tag{1.9}$$

Together with the condition $\sum_{l=1}^{n} \hat{q}_l = 1$, Eqs. (1.9) form a system of m linear equations in n unknowns. Since the rows of A are linearly independent, this system has no solution. □

Both previous Propositions together imply that, generically, a role game can only have an equilibrium in the interior of the state space if the underlying base game is square. Note that the converse result to Proposition 2 is also true: If $m = n$, then the set of games (A, B) with an interior equilibrium has positive Lebesgue measure. Indeed, the base game $(\hat{A}, -\hat{A}^T)$, with $\hat{a}_{ij} = 1$ if $j - i = 1 \bmod n$ and $\hat{a}_{ij} = 0$ otherwise, has a unique equilibrium $(\hat{p}, \hat{p})$, where $\hat{p} = (1/n, \ldots, 1/n)$. By continuity, also games (A, B) that are in a sufficiently small neighborhood of $(\hat{A}, -\hat{A}^T)$ have an equilibrium in the interior of the state space.

Furthermore, we have the following corollary.

Corollary 3 (On the number of strategies supported by mixed equilibria)
Consider a generic two-player game (A, B)*. If, in equilibrium, the row-player mixes be-*

---

[2] It seems to me that this result is a game-theoretic analogon to the exclusion principle from population ecology (see, for example Hofbauer and Sigmund, 1998). Consider a situation where m different species feed from n different resources. According to the exclusion principle there is no equilibrium with $m > n$. Such a principle can be deduced from the Lotka-Volterra equations and the proof is in fact similar to the proof we give for this Proposition.

tween k pure strategies, then so does the column-player.

Proof. Follows immediately by the previous Proposition by considering the reduced game $(\tilde{A}, \tilde{B})$, which originates from (A, B) by deleting all rows and columns that are not played in equilibrium. $\square$

1.3 Replicator dynamics for role games

In the remainder of this chapter, we follow the evolutionary interpretation of game theory. We consider an infinite population of individuals, playing a role game (A, B). Each member of the population is assumed to play a pure strategy. Under this interpretation, x_{ij} denotes the fraction of players who use the strategy $[e_i, f_j]$. As before, we denote the marginal distributions by $p(x)$ and $q(x)$, that is $p_i(x) = \sum_j^n x_{ij}$ denotes the fraction of e_i-players in the population, whereas $q_j(x) = \sum_i^m x_{ij}$ gives the fraction of f_j-players. If $x = (x_{ij})$ is the current state of the population, then the expected payoff for a player with strategy $[e_i, f_j]$ becomes

$$\pi_{ij} = \pi_{ij}(x) = \sum_{k,l} \pi_{ij,kl} \cdot x_{kl} = \frac{1}{2}\big(Aq(x)\big)_i + \frac{1}{2}\big(p(x)^T B\big)_j. \tag{1.10}$$

Summing up over all strategies yields the following average payoff of the whole population:

$$\bar{\pi} = \bar{\pi}(x) = \sum_{i,j} \pi_{ij} \cdot x_{ij} = \frac{1}{2} \cdot p(x)^T (A + B) q(x) \tag{1.11}$$

We assume that the evolution of the system is governed by the replicator dynamics (Taylor and Jonker, 1978; Zeeman, 1980). That is, we assume that the growth rate of strategy $[e_i, f_j]$ is proportional to the payoff differential $\pi_{ij} - \bar{\pi}$. Thus, we consider the ordinary differential equation (ODE)

$$\dot{x_{ij}} = x_{ij} \cdot [\pi_{ij} - \bar{\pi}]\,. \tag{1.12}$$

The analytical properties of ODE (1.12) are well understood, standard references are Weibull (1995), Hofbauer and Sigmund (1998) and Hofbauer and Sigmund (2003). The following Lemma is a straightforward application of these known results to our specific setting:

Lemma 4 (Some analytical properties of the replicator dynamics for role games)
For a role game (A, B), let $x(t)$ be a solution of the replicator equation.

(i) If the row-player's strategy e_i is dominated, then $p_i(t) \to 0$ (i.e., $x_{ij}(t) \to 0$ for all j). Analogously, if the column-player's strategy f_j is dominated, then $q_j(t) \to 0$.

(ii) If (A, B) is generic and $m \neq n$, then all orbits starting in the interior of the state space Δ^{mn} converge to the boundary.

(iii) Adding a constant c to a column of A or to a row of B does not alter the dynamics.

Proof. (i) If e_i is dominated then there is a strategy e_k such that $a_{il} < a_{kl}$ for all l. It follows that

$$\dot{\left(\frac{x_{ij}}{x_{kj}}\right)} = \frac{x_{ij}}{x_{kj}} \cdot [\pi_{ij} - \pi_{kj}] = \frac{x_{ij}}{x_{kj}} \cdot \left[\sum_{l}^{n} \frac{a_{il} - a_{kl}}{2} q_l\right] \leq 0, \tag{1.13}$$

with equality if and only if $x_{ij} = 0$. This proves that $x_{ij}(t)/x_{kj}(t) \to 0$ monotonically, and since $x_{kj}(t)$ is bounded, the result follows.

(ii) According to Hofbauer and Sigmund (1998), all orbits $x(t)$ converge to the boundary if there is no rest point in $\mathrm{int}(\Delta^{mn})$. Because rest points in the interior are Nash equilibria, the result follows from Proposition 2.

(iii) Adding a constant c to each entry of the j-th column of A increases the payoff of all players uniformly by $cq_j(x)$, and thus leaves the payoff differential $\pi_{ij} - \bar{\pi}$ unchanged. □

We next want to study the geometry of the state space under replicator dynamics. Thereby, we closely follow the approach of Cressman et al. (2000). First, we note that for any initial population $x(0) \in \mathrm{int}(\Delta^{mn})$,

$$Z_{ij} = \frac{x_{11} \cdot x_{ij}}{x_{1j} \cdot x_{i1}} \tag{1.14}$$

is an invariant of motion for all $1 \leq i \leq m$ and $1 \leq j \leq n$. Indeed, by deriving expression (1.14), we obtain

$$\begin{aligned} \dot{Z}_{ij} &= Z_{ij} \cdot [\pi_{11} + \pi_{ij} - \pi_{1j} - \pi_{i1}] \\ &= Z_{ij} \cdot \textstyle\sum_{k,l} (a_{1l} + b_{k1} + a_{il} + b_{kj} - a_{1l} - b_{kj} - a_{il} - b_{k1})/2 \cdot x_{kl} = 0. \end{aligned} \tag{1.15}$$

Furthermore, for $1 < i \leq m$ and $1 < j \leq n$, all Z_{ij} are independent: Given the frequencies in the first row x_{1l} and in the first column x_{k1}, each Z_{ij} uniquely determines the value of x_{ij}. Therefore, if we fix an ordered set Z of $(m-1)(n-1)$ initial conditions, $Z = \big\{Z_{ij} > 0 \;\big|\; 1 < i \le m,\ 1 < j \le n\big\}$, then $W_Z \subset \Delta^{mn}$ is an invariant submanifold, where

$$W_Z = \big\{x \in \mathrm{int}(\Delta^{mn}) \;\big|\; x_{1j}x_{i1} = Z_{ij}x_{11}x_{ij},\ Z_{ij} \in Z\big\}\,. \tag{1.16}$$

Thus, the interior of the state space Δ^{mn} foliates into $(n+m-2)$–dimensional manifolds invariant under (1.12). Next we define a map Φ that gives the marginal distributions for each $x \in W_Z$, that is $\Phi : W_Z \to \mathrm{int}(\Delta^m \times \Delta^n)$ is given by $\Phi(x) = (p(x), q(x))$. Without a proof, we mention the following result from Cressman et al. (2000):

Proposition 5 (Geometry of the state space)
Φ is a diffeomorphism, implying that each invariant submanifold W_Z is diffeomorph to int$(\Delta^m \times \Delta^n)$.

In particular, it follows that if there is an ordered set $\hat{Z}$ such that $W_{\hat{Z}}$ contains a rest point of Eq. (1.12), then all manifolds W_Z contain such a rest point. In the case where all $Z_{ij} = 1$, we obtain the *Wright manifold*

$$W_1 = \big\{x \in \mathrm{int}(\Delta^{mn}) \;\big|\; x_{1j}x_{i1} = x_{11}x_{ij},\ 1 < i \le m, 1 < j \le n\big\}\,. \tag{1.17}$$

On the Wright manifold, the decisions in the two roles are statistically independent, $x_{ij} = p_i(x) \cdot q_j(x)$. In this case, if the strategy $[e_i, f_j]$ becomes extinct and if the fraction of players using e_i is bounded away from zero, then $q_j(x) \to 0$. This result can be extended to all manifolds W_Z:

Lemma 6 (Extinction Lemma)
Let $x(t)$ be a solution of the replicator equation on some manifold W_Z. Suppose there is an $\varepsilon > 0$ such that for all t we have $q_j\big(x(t)\big) > \varepsilon$ for some j. Then $x_{ij}(t) \to 0$ if and only if $p_i\big(x(t)\big) \to 0$.

Proof. Because $x_{ij} \le p_i(x)$, the "if"-direction is trivial. To show the converse direction, we renumber the strategies such that $i = 1$ and $j = 1$. Let

$$z := \min\big\{Z_{kl} \;\big|1 \le k \le m, 1 \le j \le n\big\} > 0. \tag{1.18}$$

By Eq. (1.14), we have $x_{11}x_{kl} = Z_{kl}x_{1l}x_{k1}$ for all k, l. Summing up over $1 \le k \le m$ and

$1 \leq j \leq n$ thus leads to

$$x_{11} = \sum_{k,l} Z_{kl} x_{1l} x_{k1} \geq z \cdot p_1 q_1. \tag{1.19}$$

Because $z \cdot q_1$ is bounded away from zero, the result follows. □

In the following sections, we apply these results to give a classification of two special cases. We start by reviewing the case of 2×2-games (see also Gaunersdorfer et al. (1991); Hofbauer and Sigmund (1998); Berger (2001)) and then we proceed to 3×2-games.

1.4 The replicator dynamics of 2×2 games

In the simplest case, each player has two pure actions in the base game. The state space of the corresponding role game takes the form of a tetrahedron and each invariant submanifold

$$W_Z = \big\{x \in \operatorname{int}(\Delta^4) \;\big|\; x_{12}x_{21} = Z \cdot x_{11}x_{22},\ Z > 0\big\} \tag{1.20}$$

is diffeomorph to a square (see Fig. 1.1). Without loss of generality (see lemma 4iii) we may assume that the game (A, B) has the following payoff matrix:

$$\begin{pmatrix} (\ 0\ ,\ 0\ ) & (\ 0\ , b_{12}) \\ (a_{21},\ 0\ ) & (a_{22}, b_{22}) \end{pmatrix} \tag{1.21}$$

Note that this payoff matrix is chosen such that the first strategy, $[e_1, f_1]$ obtains a constant payoff of zero. Depending on the values a_{21}, a_{22}, b_{12} and b_{22}, we can distinguish between three different scenarios:

(1) *Dominance:* If $a_{21}a_{22} > 0$ or if $b_{12}b_{22} > 0$, then one of the pure actions in the base game is dominated. By lemma 4i, replicator dynamics leads to the extinction of this action and the game degenerates to a 2×1 game. In this case, there is a unique strategy $[e_i, f_j]$ that is globally stable (see Fig. 1.2i, 1.2ii).

If both, $a_{21}a_{22} < 0$ and $b_{12}b_{22} < 0$, then the underlying base game has a unique Nash equilibrium $(\hat{p}, \hat{q})$ in the interior of the state space, where

$$\hat{p} = \frac{1}{b_{22} - b_{12}} \cdot \big(b_{22}, -b_{12}\big), \qquad \hat{q} = \frac{1}{a_{22} - a_{21}} \cdot \big(a_{22}, -a_{21}\big). \tag{1.22}$$

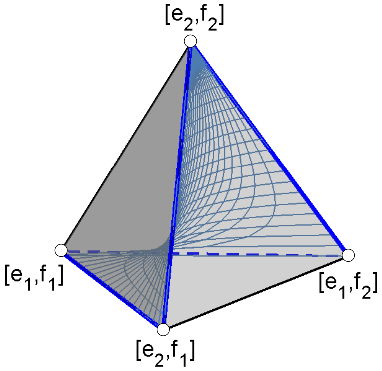


Figure 1.1: The state space of a role game involving the four strategies $[e_1, f_1]$, $[e_1, f_2]$, $[e_2, f_1]$ and $[e_2, f_2]$. The corners of the three-dimensional simplex correspond to the homogeneous populations using that strategy, the interior points denote mixed populations. For each initial state, the evolution of the system is restricted to a two-dimensional saddle-like manifold that can be represented by a square.

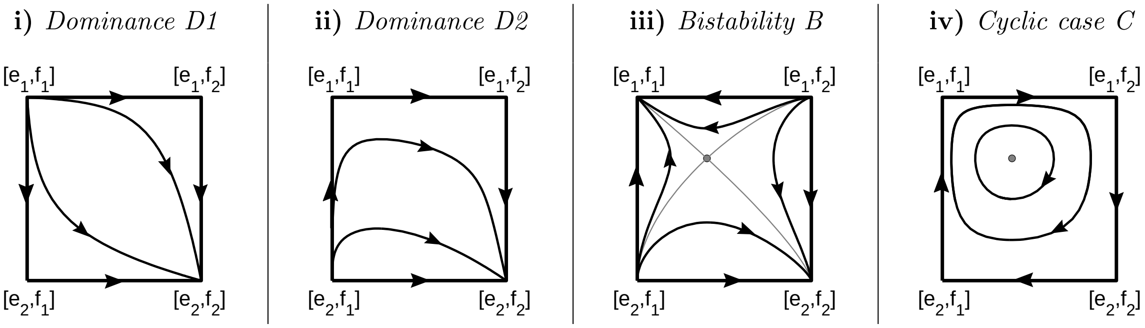


Figure 1.2: Possible scenarios for the replicator dynamics of 2×2 role games. **i)** All orbits converge monotonically to the only undominated strategy $[e_2, f_2]$. **ii)** Replicator dynamics leads to the extinction of f_1, leading to $[e_2, f_2]$ thereafter. **iii)** There are two evolutionarily stable strategies and an unstable equilibrium in the interior. **iv)** Depending on the manifold W_Z, orbits are periodic, converge to the fixed point or converge to the boundary. All graphs show the dynamics on the Wright manifold, $Z = 1$.

Therefore, for the corresponding role game there is a unique interior Nash equilibrium $\hat{x}_Z$ on each invariant manifold W_Z which can be written as

$$\hat{x}_Z = \frac{1}{(a_{22} - a_{21})(b_{22} - b_{12})} \cdot \begin{pmatrix} a_{22}b_{22} + \lambda_Z & -a_{21}b_{22} - \lambda_Z \\ -a_{22}b_{12} - \lambda_Z & a_{21}b_{12} + \lambda_Z \end{pmatrix}, \tag{1.23}$$

where the parameter λ_Z is chosen such that $\hat{x}_Z \in W_Z$. Note that $\lambda_Z = 0$ implies that $\hat{x}_z$ is on the Wright-manifold. If there is such an interior equilibrium, we have the following two additional scenarios:

(2) *Bistability:* If $a_{21}b_{12} > 0$, then the interior fixed point $\hat{x}$ is a saddle and each orbit converges to one of the two asymptotically stable fixed-points at the boundary, see Fig. 1.2iii.

(3) *Cyclic case:* If $a_{21}b_{12} < 0$, then the interior fixed point $\hat{x}$ is a focus and there is no fixed point at the boundary. The asymptotic behaviour of the system depends on the initial conditions. On the Wright-manifold W_1 all orbits are periodic and $\hat{x}$ is stable but not asymptotically stable (Fig. 1.2iv). For $Z > 1$ the fixed point is a spiral sink whereas for $Z < 1$, orbits converge to the heteroclinic cycle at the boundary (or vice versa), see Gaunersdorfer et al. (1991).

For the cyclic case it will prove useful to determine an expression for the time averages of the orbits. If the orbits do not converge to the boundary, such an expression can be derived easily (see also Hofbauer and Sigmund, 1998, for an analogous expression in the case of the Lotka-Volterra equation).

Proposition 7 (Time averages of interior orbits)
Consider a base game (A, B) with a unique Nash equilibrium $[\hat{p}, \hat{q}]$ in the interior of the state space $\Delta^n \times \Delta^n$. Let $x(t)$ be a solution of the replicator equation for the corresponding role game. If there is an ε such that $x_{ij}(t) > \varepsilon$ for all i, j and t, then

$$\lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau p\big(x(t)\big)dt = \hat{p} \tag{1.24a}$$

$$\lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau q\big(x(t)\big)dt = \hat{q} \tag{1.24b}$$

Proof. Let us first rewrite the replicator equation (1.12) as $(\log x_{ij})^{\dot{}} = \pi_{ij}(x) - \bar{\pi}(x)$. Integration from 0 to τ and division by τ leads to

$$\frac{\log x_{ij}(\tau) - \log x_{ij}(0)}{\tau} = \frac{1}{\tau} \int_0^\tau \Big(\pi_{ij}\big(x(t)\big) - \bar{\pi}\big(x(t)\big)\Big)dt \quad \forall i, j \ \ \forall \tau. \tag{1.25}$$

Since x_{ij} is bounded away from zero, we may pass to the limit $\tau \to \infty$,

$$0 = \lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau \Big(\pi_{ij}\big(x(t)\big) - \bar{\pi}\big(x(t)\big)\Big)dt \quad \forall i, j. \tag{1.26}$$

In particular, because we may normalize payoffs such that $\pi_{11}\big(x(t)\big) = 0$ for all t, we may conclude that

$$\lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau \bar{\pi}\big(x(t)\big)dt = 0. \tag{1.27}$$

Eq. (1.26) thus implies that

$$\lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau \pi_{ij}\big(x(t)\big)dt = \lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau \pi_{kl}\big(x(t)\big)dt \qquad \forall i,j,k,l, \tag{1.28}$$

which is, due to the linearity of payoffs according to (1.10), equivalent to

$$\begin{array}{rcl} \left(A\bar{q} \right)_i & = & \left(A\bar{q} \right)_k \\ \left(\bar{p}^T B\right)_j & = & \left(\bar{p}^T B\right)_l \end{array} \qquad \forall i,j,k,l \tag{1.29}$$

where $\bar{p} = \lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau p\big(x(t)\big)dt$ and $\bar{q} = \lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau q\big(x(t)\big)dt$. Eqs. (1.29) are exactly the conditions for Nash equilibria and since $x(t)$ is bounded away from zero, it follows that $(\bar{p}, \bar{q}) \in \mathrm{int}(\Delta^n \times \Delta^n)$. Therefore, $\bar{p} = \hat{p}$ and $\bar{q} = \hat{q}$. $\square$

However, in the case that orbits converge to a heteroclinic cycle, we cannot expect to obtain similar expressions for the time averages. In fact, as shown in Gaunersdorfer (1992) and Sigmund (1992), the time averages fail to converge to a single point because the time that an orbit spends near a fixed point increases exponentially for such orbits. Instead the accumulation points of the time averages form a polygon, as described in Gaunersdorfer (1992). We will return to this point in the following section.

1.5 The replicator dynamics of 3×2 games

For $m = 3$ and $n = 2$, the interior of the state space Δ^6 foliates into the 3–dimensional invariant manifolds

$$W_Z = \big\{x \in \mathrm{int}(\Delta^6) \;\big|\; x_{12}x_{21} = Z_{22} \cdot x_{11}x_{22},\ x_{12}x_{31} = Z_{32} \cdot x_{11}x_{32},\ Z_{22}, Z_{32} > 0\big\} \tag{1.30}$$

By Proposition 5, each W_Z is diffeomorph to $\mathrm{int}(\Delta^3 \times \Delta^2)$ and the invariant manifolds can be represented by a prism (see Fig. 1.3). The boundary of the prism consists of three squares (which correspond to the case where one of the row-player's strategies is absent, that is we have a 2×2–game) and of two triangles (which correspond to the case where one of the two column-player's strategies is absent in the population). By lemma 4ii, generically there is no fixed point in the interior of this prism and all orbits converge to the boundary. Without loss of generality, we suppose that the payoff matrix is given by

$$\begin{pmatrix} (\,0\,,\,0\,) & (\,0\,, b_{12}) \\ (a_{21},\,0\,) & (a_{22}, b_{22}) \\ (a_{31},\,0\,) & (a_{32}, b_{32}) \end{pmatrix} \tag{1.31}$$

The dynamics at the edges of the prism depends on the relative magnitudes of the a_{ij} and on the signs of the b_{ij}, respectively (as depicted in Fig. 1.3).

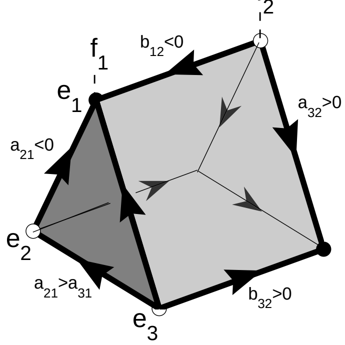


Figure 1.3: The invariant manifolds W_K can be represented by a prism. The corners of this prism correspond to homogenous populations $[e_i, f_j]$ and the boundary surfaces correspond to 2×2–games and 3×1–games, respectively. Each arrow depicts an inequality for the payoff values; the present example represents a game with $a_{31} < a_{21} < 0,\ \ 0 < a_{22} < a_{32}$ and $b_{12} < 0,\ \ b_{22} > 0,\ \ b_{32} > 0$.

We characterize the dynamics of 3×2 role games by considering all possible combinations for the directions at the edges. With respect to the actions of the column-player, f_1 and f_2, we may distinguish between two cases:

(i) One of the column-player's actions is dominated (in which case we may assume that the dominated action is f_1, that is $b_{12}, b_{22}, b_{32} > 0$). Furthermore, possibly after renumbering of the strategies of the row-player, we may assume that $a_{31} < a_{21} < 0$. This case is depicted in Fig. 1.4a. For the remaining payoff values $a_{12} = 0$, a_{22} and a_{32} (which determine the orientation of the back triangle), there are 3! possible orderings, leading to 6 possible subcases.

(ii) If none of the column-player's actions is dominated, we may assume that $b_{12} < 0$ whereas $b_{22} > 0$ and $b_{32} > 0$, see Fig. 1.4b–d. For the front triangle, there are – up to symmetry – three possible orientations. Because of the 6 possible orientations of

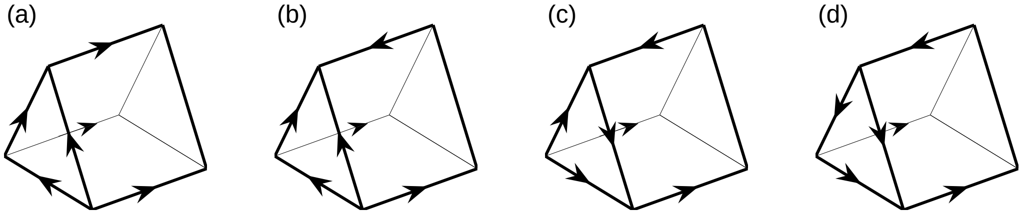


Figure 1.4: Possible orientations of the prism. There are two main cases: Either all arrows that are not part of a triangle point in the same direction, or one of the three arrows points in the opposite direction. In the first case, we may assume that the front triangle is oriented as in (a), whereas in the second case there are, up to symmetry, 3 different orientations of the front triangle (b-d).

the back triangle, this yields 18 possible subcases.

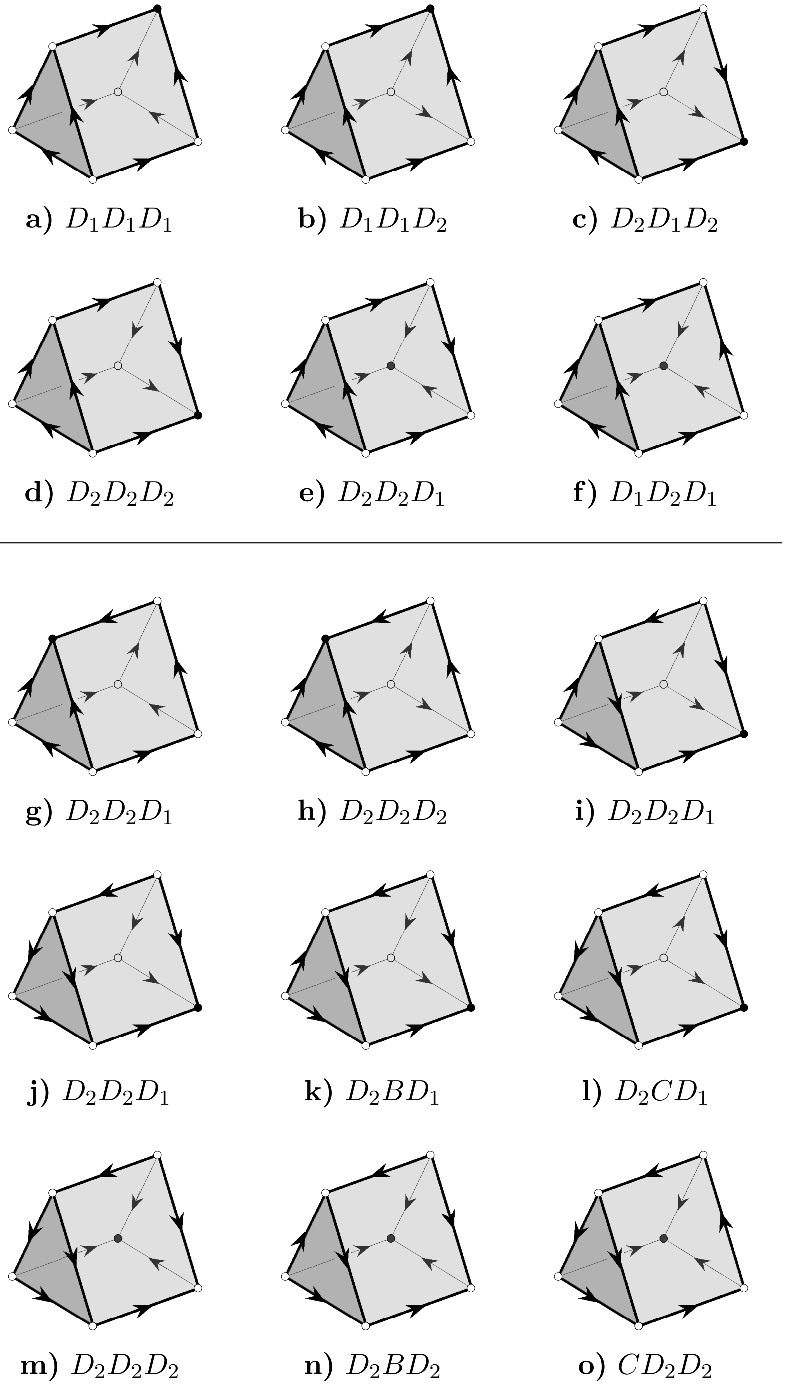
Overall, this leads to 24 different cases and a rough classification is shown in Fig. 1.5. In 21 of the 24 cases, at least one of the actions in the base game is dominated and the 3×2-game degenerates to a game with lower dimension. In particular, in the cases (a)–(o), any orbit starting in $\mathrm{int}(\Delta^3 \times \Delta^2)$ converges to a uniquely determined vertex of the prism. In the cases (p)–(r), every orbit in the interior converges to a boundary face with bistable competition. Therefore, almost all orbits are attracted by one of the two asymptotically stable vertices. Eventually, in the cases (s)–(u), every inner orbit converges to a unique boundary face with cyclic behaviour.

Let us now turn to the non-degenerated cases without dominated strategies. In the first such case (see Fig. 1.5v) we have a bistable competition between two evolutionarily stable strategies, $[e_1, f_1]$ and $[e_3, f_2]$. In the other two cases, Fig. 1.5w,x, we need to determine the stability of boundary faces with cyclic orbits. Let us therefore fix an invariant manifold W_Z and denote by F^i the corresponding boundary face where the row-player's strategy e_i is absent, that is $F^i = \mathrm{bd}(W_Z) \cap \{p_i(x) = 0\}$. Graphically, F_i corresponds to the square opposite to the e_i-edge. In order to determine the stability of such a boundary face, we call F^i saturated if

$$\lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau \pi_{ij}\big(x(t)\big)dt < \lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau \bar{\pi}\big(x(t)\big)dt \tag{1.32}$$

for all j and all orbits $x(t) \in \mathrm{int}(F^i)$. Condition (1.32) ensures that the absent strategy e_i yields a lower long-run payoff than the residents for any initial population in $\mathrm{int}(F^i)$. The following proposition establishes a relation between the stability of a boundary face F^i and the stability of fixed points $\hat{x}^i \in \mathrm{int}(F^i)$:

There is a dominated f_i.



a) $D_1D_1D_1$ **b)** $D_1D_1D_2$ **c)** $D_2D_1D_2$

d) $D_2D_2D_2$ **e)** $D_2D_2D_1$ **f)** $D_1D_2D_1$

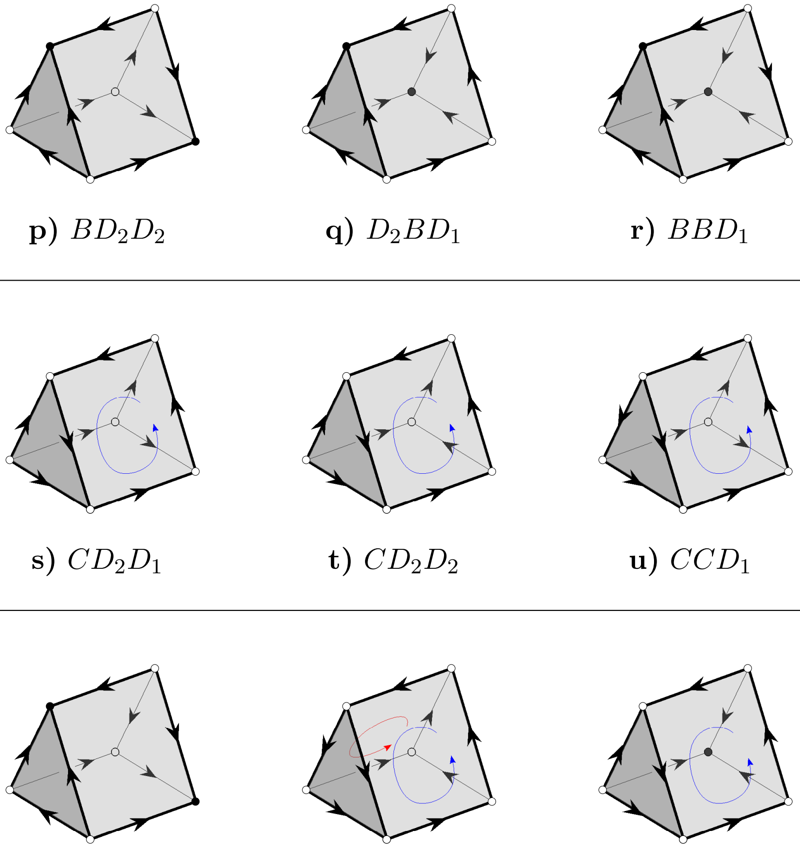
g) $D_2D_2D_1$ **h)** $D_2D_2D_2$ **i)** $D_2D_2D_1$

j) $D_2D_2D_1$ **k)** D_2BD_1 **l)** D_2CD_1

m) $D_2D_2D_2$ **n)** D_2BD_2 **o)** CD_2D_2

There is a dominated e_i and the remaining 2×2-game has a unique globally stable equilibrium.

Figure 1.5: A classification of 3×2 role games (Part 1): The 15 cases where one pure strategy, which is marked by a black circle, is globally stable. The letters D_1, D_2, B and C denote the dynamics on the 2×2-boundary-faces (Dominance, bistability and cycles, see Fig. 1.2).



There is a dominated e_i and the remaining 2×2-game exhibits bistable competition.

p) BD_2D_2 **q)** D_2BD_1 **r)** BBD_1

There is a dominated e_i and the remaining 2×2-game exhibits evolutionary cycles.

s) CD_2D_1 **t)** CD_2D_2 **u)** CCD_1

There is no dominated strategy.

v) Bistability BBD_2

w) Two Cycles CCD_2

x) ESS vs. Cycle CBD_2

Figure 1.5: A classification of 3×2 role games (Part 2).

Proposition 8 (A condition for saturated boundary faces)
Suppose there is a unique fixed point $\hat{x}^i \in int(F^i)$. If $\hat{x}^i$ is a (Lyapunov) stable Nash equilibrium, then F^i is saturated.

Proof. Follows immediately from Proposition 7. □

In particular, it follows for the cases of Fig. 1.5w,x, that orbits starting sufficiently close to a cyclic boundary face F_i will eventually converge to F^i if the corresponding fixed point $\hat{x}^i$ is a stable Nash equilibrium. However, the previous proposition does not make a statement about boundary faces F^i with heteroclinic orbits. In such a case, the following numerical example suggests that the dynamics may be non-trivial:[3]

Example 9 (The case of unstable fixed points)

Consider the following game M (which is of the type depicted as Fig. 1.5w) with initial population x_0:

$$M = \begin{pmatrix} 0,0 & 0,-0.18 \\ 0.23,0 & -0.09,0.96 \\ 0.92,0 & -0.37,0.27 \end{pmatrix}, \qquad x_0 = \begin{pmatrix} 0.26 & 0.5 \\ 0.06 & 0.14 \\ 0.02 & 0.02 \end{pmatrix} \tag{1.33}$$

Plugging the values of x_0 into expression (1.14) leads to $Z_{22} \approx 0.82$ and $Z_{32} \approx 1.92$. The respective fixed points on the boundary surfaces can be calculated according to Eq. (1.23):

$$\hat{x}^2 \approx \begin{pmatrix} 0.14 & 0.46 \\ 0 & 0 \\ 0.15 & 0.25 \end{pmatrix}, \qquad \hat{x}^3 \approx \begin{pmatrix} 0.24 & 0.60 \\ 0.04 & 0.12 \\ 0 & 0 \end{pmatrix} \tag{1.34}$$

None of these fixed points is stable and orbits on F^2 and F^3 converge to the boundary. By the index theorem (see Hofbauer and Sigmund, 1998), only one of the fixed points is saturated (i.e., transversally stable) and a straightforward calculation verifies that this unique Nash equilibrium is $\hat{x}^3$ (Both absent strategies $[e_3, f_1]$ and $[e_3, f_2]$ yield the payoff $-0.0072 < 0$). Although the initial population x_0 is close to the saturated fixed point $\hat{x}^3$, simulations suggest that the orbit does not converge to F^3; instead e_3 is regularly reintroduced, (see Fig. 1.6). We will further discuss the cause and the consequences of these non-trivial oscillations in the last section.

1.6 Recombination

Classical replicator dynamics does not introduce new strategies (see Hofbauer and Sigmund, 1998). For example, $[e_1, f_1]$ remains absent if it is not played in the initial population even if both pure actions e_1 and f_1 exist (e.g. if the initial population consists of $[e_1, f_2]$ and $[e_2, f_1]$ players). In this section we introduce recombination, thereby extending

---

[3] In general, one needs to be cautious with numerical simulations of heteroclinic cycles, where orbits converge to the boundary and the relevant fractions $x_{ij}(t)$ may well fall below the machine epsilon ε_M. However, in this case, the results seem to be sufficiently robust to be worth presenting.

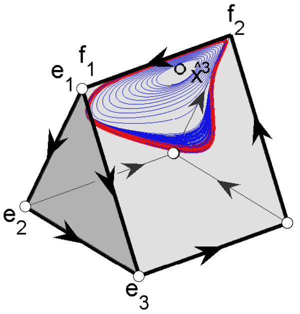


Figure 1.6: Numerical simulations for the game M without stable fixed points. The blue orbit shows the solution of the replicator equation over a time interval of 20,000 units, the red orbit gives the state of the population for the last 5% of this time interval (which must not be confused with a limit cycle). Note that the dynamics suggests that none of the pure actions e_i or f_j goes extinct.

the work of Gaunersdorfer et al. (1991).

We consider a general $m \times 2$ role game (A, B) with the actions $e_1, \ldots, e_m$ for the row-player and the actions f_1 and f_2 for the column-player. Now we modify the replicator dynamics (1.12) by adding recombination:

$$\dot{x}_{ij} = x_{ij} \cdot \left[\pi_{ij} - \bar{\pi}\right] - \rho D_{ij}. \tag{1.35}$$

Here $\rho > 0$ is the recombination fraction and $D_{ij} := x_{ij} - p_i(x) \cdot q_j(x)$ is the linkage disequilibrium (In an infinitesimal time interval of length Δt, the combination $[e_i, f_j]$ is broken up with probability $\rho x_{ij} \Delta t$, and formed anew with probability $\rho p_i(x) \cdot q_j(x) \Delta t$).

Proposition 10 (Convergence to the Wright manifold)
Under the modified replicator equation (1.35), all orbits starting in the interior of the state space converge to the Wright manifold. On this manifold, the modified and the classical replicator dynamics coincide.

Proof. In order to simplify notation we write Z_k for the expressions Z_{k2} defined by (1.14). We will show $Z_2 \to 1$ under the modified replicator dynamics (1.35); all other Z_k follow by a symmetry argument. We have

$$\frac{\dot{Z}_2}{Z_2} = -\rho \cdot \left[D_{11}/x_{11} + D_{22}/x_{22} - D_{12}/x_{12} - D_{21}/x_{21}\right] \tag{1.36}$$

It is easy to see that

$$\begin{aligned} D_{11}/x_{11} &= 1 - \tfrac{p_1 q_1}{x_{11}} = 1 - \tfrac{(x_{11}+\cdots+x_{1n})(x_{11}+x_{21})}{x_{11}} = \\ &= 1 - x_{11} - \cdots - x_{1n} - x_{21} - \tfrac{x_{12}x_{21}}{x_{11}} - \cdots - \tfrac{x_{1n}x_{21}}{x_{11}} = \\ &= x_{22} + \cdots + x_{2n} - \tfrac{1}{Z_2}x_{22} - \cdots - \tfrac{1}{Z_n}x_{2n} = \\ &= (1 - 1/Z_2)x_{22} + (1 - 1/Z_3)x_{23} + \cdots + (1 - 1/Z_n)x_{2n}, \end{aligned} \tag{1.37}$$

hence

$$D_{11}/x_{11} = \textstyle\sum_{k=1}^{n}(1 - \tfrac{Z_1}{Z_k})x_{2k} \tag{1.38}$$

and analogously

$$\begin{aligned} D_{12}/x_{12} &= \textstyle\sum_{k=1}^{n}(1 - \tfrac{Z_2}{Z_k})x_{2k} \\ D_{22}/x_{22} &= \textstyle\sum_{k=1}^{n}(1 - \tfrac{Z_k}{Z_2})x_{1k} \\ D_{21}/x_{21} &= \textstyle\sum_{k=1}^{n}(1 - \tfrac{Z_k}{Z_1})x_{1k} \end{aligned} \tag{1.39}$$

Plugging these identities into Eq. (1.36) yields

$$\frac{\dot{Z}_2}{Z_2} = -\rho(Z_2 - 1)[\sum_{k=1}^{n}(\frac{x_{2k}}{Z_k} + \frac{x_{1k}Z_k}{Z_1 Z_2})] = -\rho(Z_2 - 1)[p_1\frac{x_{21}}{x_{11}} + p_2\frac{x_{12}}{x_{22}}], \tag{1.40}$$

which shows that Z_2 (and therefore all Z_k) converge to 1 monotonically. In this case it follows from Eqs. (1.38) and (1.39) that $D_{ij} \to 0$. □

1.7 Discussion

In this introductory chapter, we have given an overview of the equilibria and the replicator dynamics of role games. In particular, we have seen that away from the Wright manifold, role games can induce non-trivial evolutionary oscillations. A similarly peculiar result has been previously reported by Chamberland and Cressman (2000), who were investigating a symmetric game with payoff matrix:

$$A = \begin{pmatrix} 0 & 6 & -4 \\ -4 & 0 & 4 \\ 2 & -2 & 0 \end{pmatrix}. \tag{1.41}$$

For this 3×3-base game of rock-scissors-paper type, all orbits of the replicator dynamics converge to the unique fixed point in the interior. However, for the corresponding role game (A, A^T), the set of Nash equilibria is unstable. Or, as interpreted by Sandholm (2006): If players observe commonly the toss of a coin before the game begins, then the long-run dynamics differs from a game without such a coin toss.

To understand this counterintuitive result, it is important to recall that the replicator dynamics of a role game promotes strategies that yield a high *aggregate* payoff. Even if the row-player's strategy e_i performs well taken by itself, the fraction of e_i-players does not necessarily increase. Instead, if e_i-players tend to use an inferior strategy f_j in the role of the column-player (which is, of course, only possible away from the Wright-manifold), then $p_i(x)$ may decrease. Therefore, the dynamics and the stability of fixed points does not only depend on the game itself but in general also on correlations in the initial population.

Chamberland and Cressman (2000) offer three possible resolutions to avoid these inconsistencies. The first suggestion is to restrict attention to the dynamics on the Wright-manifold. The authors consider this restriction as problematic: In a population of independent individuals, each playing a pure strategy, there is no reason to assume that strategy choices are uncorrelated (however, see our convergence result if we allow recombination). Secondly, they propose to consider evolutionarily stable strategies (ESS) instead of Nash equilibria because in the case of ESS, inconsistencies can be excluded (see Cressman et al., 2000, and note that game (1.41) does not have an ESS). As their last resolution, Chamberland and Cressman (2000) question the applicability of replicator dynamics (or any aggregate monotonic selection dynamics) to games with two independent decisions. Instead, they suggest to use *subgame monotone selection dynamics*, where $p_i(x)/p_k(x)$ increases if and only if e_i yields a higher payoff than e_k.

This requirement is fulfilled if we assume that strategy revisions for the two roles of the base game occur independently:

$$\begin{aligned} \dot{p}_i &= p_i \cdot \left[(Aq)_i - p^T A q\right] \\ \dot{q}_j &= q_j \cdot \left[(p^T B)_j - p^T B q\right] \end{aligned} \tag{1.42}$$

However, one can easily verify that such a dynamics is equivalent to the replicator dynamics of the corresponding role game, restricted to the Wright-manifold:

$$\begin{aligned} \dot{p}_i(x) &= \sum_{j=1}^n x_{ij}\left[\pi_{ij}(x) - \bar{\pi}(x)\right] \\ &= \sum_{j=1}^n p_i(x) q_j(x) \Big[\tfrac{1}{2}\big(Aq(x)\big)_i + \tfrac{1}{2}\big(p(x)^T B\big)_j - \tfrac{1}{2} \cdot p(x)^T (A+B) q(x)\Big] \\ &= \tfrac{1}{2} p_i(x) \cdot \Big[\big(Aq(x)\big)_i - p(x)^T A q(x)\Big]. \end{aligned} \tag{1.43}$$

An analogous calculation holds for $\dot{q}_j(x)$. In the following applications we will therefore mostly focus on the dynamics on the Wright-manifold.

Chapter 2

Incentives and Opportunism: From the carrot to the stick

Abstract

Cooperation in public good games is greatly promoted by positive and negative incentives. In this paper, we use evolutionary game dynamics to study the evolution of opportunism (the readiness to be swayed by incentives) and the evolution of trust (the propensity to cooperate in the absence of information on the co-players). If both positive and negative incentives are available, evolution leads to a population where defectors are punished and players cooperate, except when they can get away with defection. Rewarding behavior does not become fixed, but can play an essential role in catalyzing the emergence of cooperation, especially if the information level is low.

2.1 Introduction

Social dilemmas are obstacles to the evolution of cooperation. Examples such as the Prisoner's Dilemma show that self-interested motives can dictate self-defeating moves, and thus suppress cooperation. Positive and negative incentives (the carrot and the stick) can both induce cooperation in a population of self-regarding agents (see e.g. Olson, 1965; Ostrom and Walker, 2003; Sigmund, 2007). The provision of such incentives is costly, however, and therefore raises a second-order social dilemma. This issue has been addressed in many papers, particularly for the case of negative incentives. We mention for example Yamagishi (1986); Boyd and Richerson (1992); Fehr and Gächter (2002); Walker and Halloran (2004); Bowles and Gintis (2004); Gardner and West (2004a); Nakamaru and Iwasa (2006); Sefton et al. (2007); Carpenter (2007); Lehmann et al. (2007a); Kiyonari and Barclay (2008).

It is easily seen that the efficiency of the two types of incentives relies on contrasting and even complementary circumstances. Indeed, if most players cooperate, then it will be costly to reward them all, while punishing the few defections will be cheap: often, the mere threat of a sanction suffices (Boyd et al., 2003; Gächter et al., 2008). On the other hand, if most players defect, then punishing them all will be a costly enterprise, while rewarding the few cooperators will be cheap. Obviously, therefore, the best policy for turning a population of defectors into a population of cooperators would be, first, to use the carrot, and at some later point, the stick.

In the absence of a proper institution to implement such a policy, members of the population can take the job onto themselves. But what is *their* incentive to do so? It pays only if the threat of a punishment, or the promise of a reward, should turn a co-player from a defector into a cooperator. Hence, the co-players must be opportunistic, i.e., prone to be swayed by incentives. In order to impress a co-player, the threat (or promise) of an incentive must be sufficiently credible. In the following model, we shall assume that the credibility is provided by the players' reputation, i.e. by their history, and thus assume several rounds of the game, not necessarily with the same partner (see e.g. Sigmund et al., 2001; Fehr and Fischbacher, 2003; Barclay, 2006). Credibility could alternatively be provided by a verbal commitment, for example. Since mere talk is cheap, however, such commitments need to be convincing; ultimately, they must be backed up by actions, and hence again rely on reputation. Whether a player obtains information about the co-players' previous actions from direct experience, or from witnessing them at a distance, or hearing about them through gossip, can be left open at this stage. In particular, we do not assume repeated rounds between the same two players, but do not exclude them either. Basically, the carrot or the stick will be applied after the cooperation, or defection, and hence are forms of targeted reciprocation (while conversely, of course, the promise to return good with good and bad with bad, can act as an incentive).

In the following, we present a simple game theoretic model to analyze the evolution of opportunism, and to stress the smooth interplay of positive and negative incentives. The model is based on a previous paper (Sigmund et al., 2001), which analyses punishment and reward separately and which *presumes* opportunistic agents. Here, we show how such opportunistic agents evolve via social learning, and how first rewards, then punishment lead to a society dominated by players who cooperate, except when they expect that they can get away with defection. Rewards will not become stably established; but they can play an essential role in the transition to cooperation, especially if the information level is below a specific threshold. Whenever the benefit-to-cost ratio for the reward is larger than one, the eventual demise of rewarders is surprising, since a homogeneous population of rewarding cooperators would obtain a higher payoff than a homogeneous population of punishing cooperators. We first analyze the model by means of the replicator dynamics, then by means of a stochastic learning model based on the Moran process. Thus both finite populations and the limiting case of infinite populations will be covered. In the discussion, we study the role of errors, compare our results with experiments and point out the need to consider a wider role for incentives.

2.2 The model

Each round of the game consists of two stages, a helping stage and an incentive stage. Individuals in the population are randomly paired. A dice decides who is in the role of the (potential) Donor, and who is Recipient. In the first stage, Donors may transfer a benefit b to their Recipients, at their own cost c, or they may refuse to do so. These two alternatives are denoted by **C** (for cooperation) resp. **D** (for defection). In the second stage, Recipients can reward their Donors, or punish them, or refuse to react. If rewarded, Donors receive an amount β; if punished, they must part with that amount β; in both cases, Recipients must pay an amount γ, since both rewarding and punishing is costly. As usual, we assume that $c < b$, as well as $c < \beta$ and $\gamma < b$. Using the same parameter values β and γ for both types of incentives is done for convenience only: basically, all that matters are the inequalities. They ensure that Donors are better off by choosing **C**, if their

Recipients use an incentive; and that in the case of rewards, both players have a positive payoff. But material interests speak against using incentives, since they are costly; and in the absence of incentives, helping behavior will not evolve.

The four possible moves for the second stage will be denoted by **N**, to do nothing; **P**, to punish defection; **R**, to reward cooperation; and **I**, to provide for both types of incentives, i.e. to punish defection *and* to reward cooperation. For the first stage, next to the two unconditional moves **AllC**, to always cooperate, and **AllD**, to always defect, we also consider the opportunistic move: namely to defect except if prodded by an incentive. We shall, however, assume that information about the co-player may be incomplete. Let μ denote the probability to know whether the co-player provides an incentive or not, and set $\bar{\mu} = 1 - \mu$. We consider two types of opportunists, who act differently under uncertainty: players of type $\mathbf{O_C}$ defect only if they know that their co-player provides no incentive, and players of type $\mathbf{O_D}$ defect except if they know that an incentive will be delivered. Hence in the absence of information, $\mathbf{O_C}$ players play **C** and $\mathbf{O_D}$-players **D**. This yields sixteen strategies, each given by a pair $[i, j]$, with $i \in M_D := \{\mathbf{AllC}, \mathbf{O_C}, \mathbf{O_D}, \mathbf{AllD}\}$ specifying how the player acts as a Donor and $j \in M_R := \{\mathbf{N}, \mathbf{P}, \mathbf{R}, \mathbf{I}\}$ how the player acts as Recipient. If player I is Donor and player II Recipient, the pair (p_I, p_{II}) of their payoff values is determined by their moves in the corresponding roles. Hence we can describe these pairs by a 4×4 matrix $(a_{[ij]}, b_{[ij]})$, see Tab. 2.1.

$*$	**N**	**P**	**R**	**I**
AllC	$(-c, b)$	$(-c, b)$	$(\beta - c, b - \gamma)$	$(\beta - c, b - \gamma)$
$\mathbf{O_C}$	$(-\bar{\mu}c, \bar{\mu}b)$	$(-c, b)$	$(\beta - c, b - \gamma)$	$(\beta - c, b - \gamma)$
$\mathbf{O_D}$	$(0, 0)$	$(-\mu c - \bar{\mu}\beta, \mu b - \bar{\mu}\gamma)$	$(\mu(\beta - c), \mu(b - \gamma))$	$(-(1 - 2\mu)\beta - \mu c, \mu b - \gamma)$
AllD	$(0, 0)$	$(-\beta, -\gamma)$	$(0, 0)$	$(-\beta, -\gamma)$

Table 2.1: Payoffs for the helping game with punishment.

This specifies the payoff values for the corresponding symmetrized game, which is given by a 16×16-matrix. A player using $[i, j]$ against a player using $[k, l]$ is with equal probability in the role of the Donor or the Recipient and hence obtains as payoff $(a_{[i,l]} + b_{[k,j]})/2$. The state of the population $\mathbf{x} = (x_{[i,j]})$ is given by the frequencies of the 16 strategies.

There exist a wealth of possible evolutionary dynamics, describing how the frequencies of the strategies change with time under the influence of social learning (Hofbauer and Sigmund, 1998). We shall consider only one updating mechanism, but stress that the results hold in many other cases too. For the learning rule, we shall use the familiar Moran-like 'death-birth' process (Nowak, 2006a): we thus assume that occasionally, players can update their strategy by copying the strategy of a 'model', i.e., a player chosen at random with a probability which is proportional to that player's fitness. This fitness in turn is assumed to be a convex combination $(1-s)B+sP$, where B is a 'baseline fitness' (the same for all players), P is the payoff (which depends on the model's strategy, and the state of the population), and $0 \leq s \leq 1$ measures the 'strenght of selection', i.e. the importance of the game for overall fitness. (We shall always assume s small enough to avoid negative fitness values). This learning rule corresponds to a Markov process. The rate for switching from strategy $[k,l]$ to strategy $[i,j]$ is $(1-s)B+sP_{[i,j]}$, independently of $[k,l]$.

2.2.1 Large populations

The learning rule leads, in the limiting case of an infinitely large population, to the replicator equation for the relative frequencies $x_{[ij]}$: the growth rate of any strategy is given by the difference between its payoff and the average payoff in the population (Hofbauer and Sigmund, 1998). This yields an ordinary differential equation which can be analyzed in a relatively straightforward way, despite being 15-dimensional.

Let us first note that **I** is weakly dominated by **P**, in the sense that **I**-players never do better, and sometimes less well, than **P**-players. Hence, no state where all the strategies are played can be stationary. The population always evolves towards a region where at least one strategy is missing. Furthermore, **AllC** is weakly dominated by $\mathbf{O_C}$, and **All D** by $\mathbf{O_D}$. This allows to reduce the dynamics to lower dimensional cases. Of particular relevance are the states where only two strategies are present, and where these two strategies prescribe the same move in one of the two stages of the game. The outcome of such pairwise contests is mostly independent of the parameter values, with three exceptions:

(i) In a homogeneous $\mathbf{O_C}$-population, **R** dominates **N** if and only if $\mu > \frac{\gamma}{b}$.

(ii) In a homogeneous $\mathbf{O_D}$-population, **P** dominates **N** if and only if $\mu > \frac{\gamma}{b+\gamma}$.

(iii) In a homogeneous $\mathbf{O_D}$-population, **P** dominates **R** if and only if $\mu > 1/2$.

In each case, it is easy to understand why higher reputation will have the corresponding effect. Owing to our assumption $\gamma < b$, all these thresholds for μ lie in the open interval (0,1).

One can obtain a good representation of the dynamics by looking at the situations where there are two alternatives for the first stage (namely **AllD** and $\mathbf{O_C}$, or **AllD** and $\mathbf{O_D}$, or $\mathbf{O_C}$ and $\mathbf{O_D}$), and the three alternatives **N**, **P** and **R** for the second stage. In each such case, the state space of the population can be visualized by a prism. Here, each of its 'square faces' stands for the set of all mixed populations with only four strategies present. For instance, if the population consists only of the four strategies $[\mathbf{O_C}, \mathbf{N}], [\mathbf{O_C}, \mathbf{R}], [\mathbf{O_D}, \mathbf{N},]$ and $[\mathbf{O_D}, \mathbf{R}]$, then the state corresponds to a point in the three dimensional simplex spanned by the corresponding four monomorphic populations. But since the double ratios $x_{[ij]}x_{[kl]}/x_{[il]}x_{[kj]}$ are invariant under the replicator dynamics (see Hofbauer and Sigmund, 1998, pp. 122-125), the state cannot leave the corresponding two-dimensional surface, which may be represented by a square. For several pairs of strategies (such as $[\mathbf{O_C}, \mathbf{P}]$ and $[\mathbf{AllC}, \mathbf{P}]$, or $[\mathbf{AllD}, \mathbf{N}]$ and $[\mathbf{O_D}, \mathbf{N}]$), all populations which are mixtures of the corresponding two strategies are stationary. There is no selective force favouring one strategy over the other. We shall assume that in this case, small random shocks will cause the state to evolve through neutral drift. This implies that evolution then leads ultimately to $[\mathbf{O_C}, \mathbf{P}]$, and hence to a homogeneous population which stably cooperates in the most efficient way. Indeed, it is easy to see that no other strategy can invade a monomorphic $[\mathbf{O_C}, \mathbf{P}]$-population through selection. The only flaw is that $[\mathbf{AllC}, \mathbf{P}]$ can enter through neutral drift. Nevertheless, $[\mathbf{O_C}, \mathbf{P}]$ is a Nash equilibrium.

But how can $[\mathbf{O_C}, \mathbf{P}]$ get off the ground? Let us first consider what happens if the possibility to play **R**, i.e. to reward a cooperative move, is excluded. The asocial strategy $[\mathbf{AllD}, \mathbf{N}]$ is stable. It can at best be invaded through neutral drift by $[\mathbf{O_D}, \mathbf{N}]$. If $\mu > \gamma/(b + \gamma)$, this can in turn be invaded by $[\mathbf{O_D}, \mathbf{P}]$, which then leads to $[\mathbf{O_C}, \mathbf{P}]$. If μ is smaller, however, that path is precluded and the population would remain in an un-cooperative state. It is in this case that the **R**-alternative plays an essential role. By neutral drift, $[\mathbf{AllD}, \mathbf{R}]$ can invade $[\mathbf{AllD}, \mathbf{N}]$. More importantly, $[\mathbf{O_D}, \mathbf{R}]$ dominates $[\mathbf{O_D}, \mathbf{N}]$, $[\mathbf{AllD}, \mathbf{R}]$ and $[\mathbf{AllD}, \mathbf{N}]$. From $[\mathbf{O_D}, \mathbf{R}]$, the way to $[\mathbf{O_C}, \mathbf{R}]$ and then to $[\mathbf{O_C}, \mathbf{P}]$ is easy.

The essential step of that evolution occurs in the transition from $\mathbf{O_D}$ to $\mathbf{O_C}$, when players start cooperating by default, i.e., in the absence of information (see the third column in Fig. 2.1). If the **R**-alternative is not available, then for small values of μ, the population can be trapped in $[\mathbf{O_D}, \mathbf{N}]$. But if the **R**-alternative can be used, it can switch from $[\mathbf{O_D}, \mathbf{N}]$ to $[\mathbf{O_D}, \mathbf{R}]$. In a population where the first move is either $\mathbf{O_D}$ or $\mathbf{O_R}$, and the second move either **N** or **R**, there is a (four-membered) Rock-Paper-Scissors cycle, one strategy is superseded by the next. There exists a unique stationary state where these

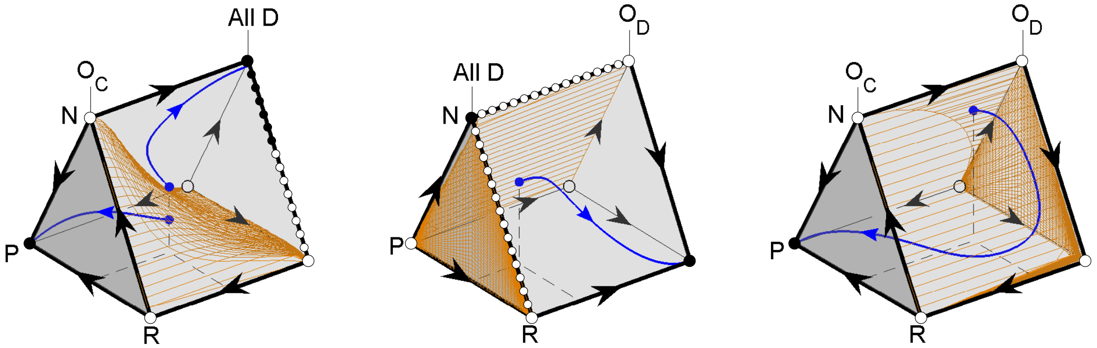


Figure 2.1: Dynamics of a population consisting of $\mathbf{O_C}$ and **AllD** (left column), **AllD** and $\mathbf{O_D}$ (middle column) resp. $\mathbf{O_C}$ and **AllD** (right column). Black points represent Nash-equilibria, white points indicate unstable fixed points. The arrows on the edges indicate the direction of the dynamics if all other strategies are absent. The orange grid is the manifold that separates initial values with different asymptotic behaviour. The blue curves represent the typical dynamics for a given initial population. Parameter values: $b = 4,\ c = 1,\ \beta = \gamma = 2$ and $\mu = 30\%$ (hence $\frac{\gamma}{2b} < \mu < \frac{\gamma}{\gamma+b}$).

four alternatives are used. We show in the appendix that for $\mu < \gamma/2b$, this stationary state cannot be invaded by any strategy using $\mathbf{P}$. But due to the Rock-Paper-Scissors dynamics, it is inherently unstable. The population will eventually use mostly strategy $[\mathbf{O_C}, \mathbf{R}]$. There, the strategy $[\mathbf{O_C}, \mathbf{P}]$ can invade and become fixed.

In the competition between $[\mathbf{O_D}, \mathbf{N}]$ and $[\mathbf{O_C}, \mathbf{P}]$, the latter is dominant if and only if $\mu > (c+\gamma)/(c+\gamma+b)$ (a condition which is independent of β). If not, then the competition is bistable, meaning that neither strategy can invade a homogeneous population adopting the other strategy. An equal mixture of both strategies converges to the pro-social strategy $[\mathbf{O_C}, \mathbf{P}]$ if and only if $\mu(\beta - 2c - 2b - \gamma) < \beta - 2c - \gamma$. In the case $\gamma = \beta$, this simply reduces to $\mu > c/(c+b)$. We thus obtain a full classification of the replicator dynamics in terms of the parameter μ. The main bifurcation values are $\frac{\gamma}{2b} < \frac{\gamma}{b+\gamma} < \frac{\gamma}{b}$ and $\frac{1}{2}$. These can be arranged in two ways, depending on whether $b < 2\gamma$ or not. But the basic outcome is the same in both case (see Fig. 2.1 and the appendix).

It is possible to modify this model by additionally taking into account the recombination of the traits affecting the first and the second stage of the game. Indeed, recombination does not only occur for genetic transmission of strategies, but also for social learning. A modification of an argument from Gaunersdorfer et al. (1991) allows to show that in this case, the double ratios $x_{[ij]}x_{[kl]}/x_{[il]}x_{[kj]}$ converge to 1, so that the traits for the first and the second stage of the game become statistically independent of each other. Hence the previous analysis still holds. In Lehmann et al. (2007b) and Lehmann and Rousset (2009) it is shown, in contrast, that recombination greatly affects the outcome in a lattice and in a finite population model without reputational effects.

2.2.2 Small mutation rates

In the case of a finite population of size M, the learning process corresponds to a Markov chain on a state space which consists of the frequencies of all the strategies (which sum up to M). The absorbing states correspond to the homogeneous populations: in such a homogeneous population, imitation cannot introduce any change. If we add to the learning process a 'mutation rate' (or more precisely, an exploration rate), by assuming that players can also adopt a strategy by chance, rather than imitation, then the corresponding process is recurrent (a chain of transitions can lead from every state to every other) and it admits a unique stationary distribution. This stationary distribution describes the frequencies of the states in the long run. It is in general laborious to compute, since the number of possible states grows polynomially in M. However, in the limiting case of a very small exploration rate (the so-called adiabatic case), we can assume that the population is mostly in a homogeneous state, and we can compute the transition probabilities between these states (Nowak, 2006a). This limiting case is based on the assumption that the fate of mutant (i.e., whether it will be eliminated or fixed in the population) is decided before the next mutation occurs. We can confirm the results from the replicator dynamics. For simplicity, we confine ourselves to the non-dominated strategies $\mathbf{O_C}$, $\mathbf{O_D}$, resp. $\mathbf{N}$, $\mathbf{P}$ and $\mathbf{R}$; similar results can be obtained by considering the full strategy space.

In the stationary distribution, the population is dominated by the strategy $[\mathbf{O_C}, \mathbf{P}]$, but for smaller values of μ, it needs the presence of the $\mathbf{R}$-alternative to emerge. This becomes particularly clear if one looks at the transition probabilities (see appendix). Except for large values of μ, only the strategy $[\mathbf{O_D}, \mathbf{R}]$ can invade the asocial $[\mathbf{O_D}, \mathbf{N}]$ with a fixation probability which is larger than the neutral fixation probability $1/M$.

If $[\mathbf{O_C}, \mathbf{P}]$ dominates $[\mathbf{O_D}, \mathbf{N}]$, or when it fares best in an equal mixture of both strategies, then it needs not the help of $\mathbf{R}$-players to become the most frequent strategy in the long run (i.e. in the stationary distribution). But for smaller values of μ, rewards are essential. In Fig. 2.2, it is shown that that the existence of rewarding strategies allows the social strategy $[\mathbf{O_C}, \mathbf{P}]$ to supersede the asocial $[\mathbf{O_D}, \mathbf{N}]$ even in cases in which the players have hardly any information about their co-players. The time until the system leaves $[\mathbf{O_D}, \mathbf{N}]$ is greatly reduced if rewarding is available (see Fig. 2.3). In the appendix it is shown that the state $[\mathbf{O_C}, \mathbf{P}]$ is usually reached from $[\mathbf{O_C}, \mathbf{R}]$, while the strategy most likely

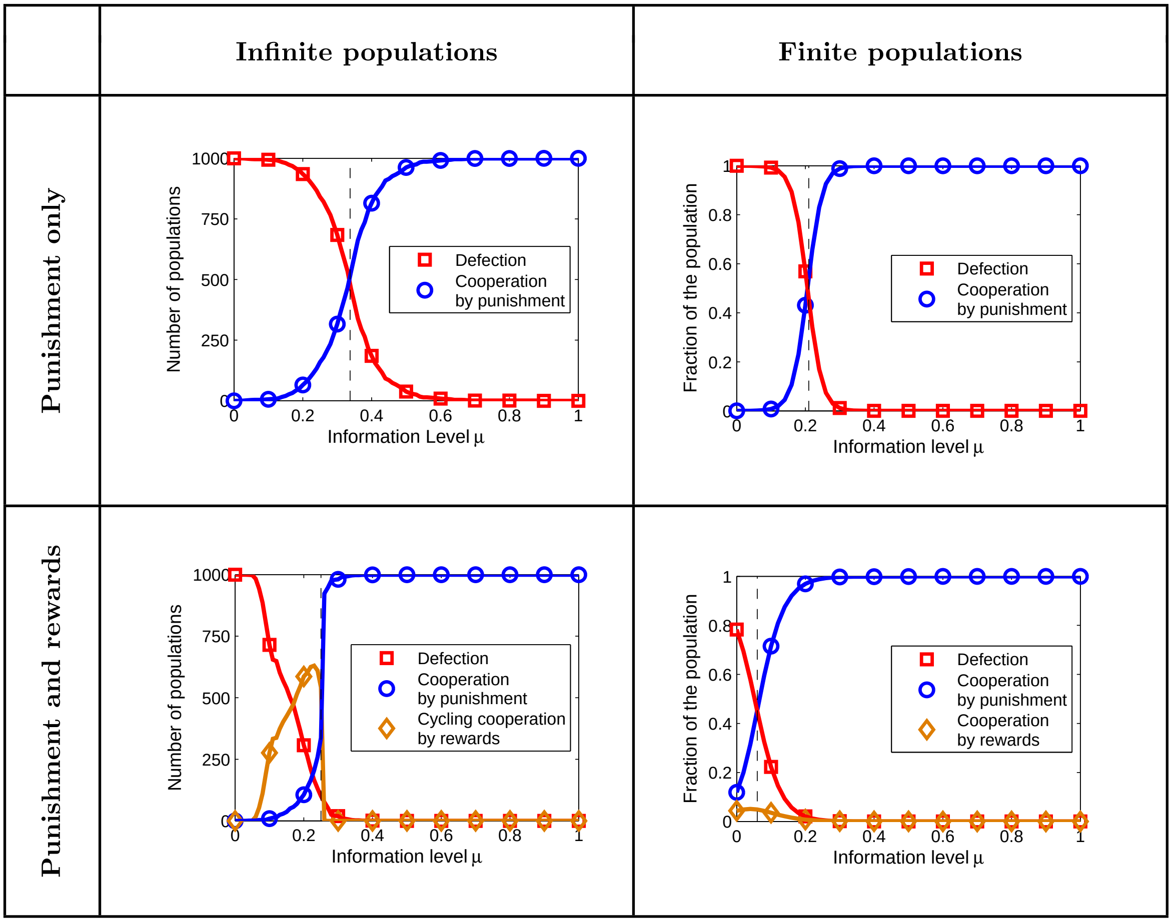


Figure 2.2: Strategy selection in finite and infinite populations, depending on the information parameter μ. The left column shows the outcome of a simulation of the replicator equation for 1,000 randomly and independently chosen initial populations. If only punishment is available to sway opportunistic behaviour then cooperative outcomes become more likely if μ exceeds roughly 1/3 (in which case $[\mathbf{O_C}, \mathbf{P}]$ becomes fixed). As soon as rewards are also allowed, punishment enforced cooperation becomes already predominant for $\mu > \gamma/2b = 1/4$. Additionally, for smaller values of μ the population may tend to cycle between the strategies $[\mathbf{O_C}, \mathbf{R}]$, $[\mathbf{O_C}, \mathbf{N}]$, $[\mathbf{O_D}, \mathbf{N}]$ and $[\mathbf{O_D}, \mathbf{R}]$, represented by the orange line in the lower left graph.

The right column shows the stationary distribution of strategies in a finite population. Again, whitout rewards a considerably higher information level μ is necessary to promote punishment-enforced cooperation (either $[\mathbf{O_C}, \mathbf{P}]$ or $[\mathbf{O_D}, \mathbf{P}]$; note that the both opportunists become indistinguishable in the limit case of complete information as they only differ in the way they play against unknown co-players). In finite populations, rewarding states only act as a catalyzer for punishing populations; even for small μ, the outcomes $[\mathbf{O_C}, \mathbf{R}]$ resp. $[\mathbf{O_D}, \mathbf{R}]$ are never most abundant.

Parameter values: $b = 4, c = 1, \beta = \gamma = 2$. For finite populations we used a population size $M = 100$ and selection strength $s = 1/10$.

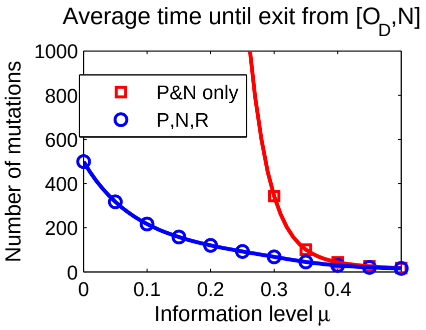


Figure 2.3: Average number of necessary mutants until a population of $[\mathbf{O_D}, \mathbf{N}]$ players is successfully invaded. Adding the possibility of rewards decreases the exit time considerably (for $\mu = 0\%$ it takes 500 mutations with rewards and almost 500,000 mutations without). As the information level increases, this catalytic effect of rewards disappears. Parameter values: Population size $M = 100$, selection strength $s = 1/10$; $b = 4, c = 1, \beta = \gamma = 2$.

to invade the asocial $[\mathbf{O_D}, \mathbf{N}]$ is $[\mathbf{O_D}, \mathbf{R}]$. These outcomes are robust, and depend little on the parameter choices. Moreover, they are barely affected by the mutation structure. If, instead of assuming that all mutations are equally likely, we only allow for mutations in the behavior in one of the two stages (i.e., no recombination between the corresponding traits), the result is very similar. Apparently, if it is impossible to mutate directly from $[\mathbf{O_D}, \mathbf{N}]$ to $[\mathbf{O_C}, \mathbf{P}]$, then the detour via $[\mathbf{O_D}, \mathbf{P}]$ works almost as well.

Even for the limiting case $\mu = 0$ (no reputation effects), the role of rewards is strongly noticeable. Without rewards, the stationary probability of the asocial strategy $[\mathbf{O_D}, \mathbf{N}]$ is close to 100 percent; with the possibility of rewards, it is considerably reduced.

2.3 Discussion

We have analyzed a two-person, two-stages game. It is well-known that it corresponds to a simplified version of the Ultimatum game (Güth et al., 1982), in the punishment case, or of the Trust game (Berg et al., 1995), in the reward case (De Silva and Sigmund, 2009; Sigmund, 2010). Similar results also hold for the N-person public good game with reward and punishment (e.g. Hauert et al., 2004). However, the many-person game offers a wealth of variants having an interest of their own (as, for instance, when players decide to mete out punishment only if they have a majority on their side; see Boyd et al., 2010). In this paper, we have opted for the simplest set-up and considered pairwise interactions only.

In classical economic thought, positive and negative incentives have often been treated on equal footing, so to speak (Olson, 1965). In evolutionary game theory, punishing is studied much more frequently than rewarding. The relevance of positive incentives is sometimes queried, on the ground that helping behavior makes only sense if there is an asymmetry in resource level between donor and recipient. If A has a high pile of wood, and B has no fuel, A can give some wood away at little cost, and provide a large benefit to B. This is the cooperative act. Where is the positive incentive? It would be absurd to imagine that B rewards A by returning the fuel. But B can reward A by donating some other resource, such as food, or fire, which A is lacking.

In experimental economics, punishing behavior has been studied considerably more often than rewarding behavior (Yamagishi, 1986; Fehr and Gächter, 2002; Barclay, 2006; Dreber et al., 2008). In the last few years, there has been a substantial amount of empirical work on the interplay of the two forms of incentives (Andreoni et al., 2003; Rockenbach and Milinski, 2006; Sefton et al., 2007). The results, with two exceptions to be discussed presently, confirm our theoretical conclusion: punishment is the more lasting factor, but the combination of reward and punishment works best. This outcome is somewhat surprising, because in most experiments, players are anonymous and know that they cannot build up a reputation. One significant exception is the investigation, in Fehr and Fischbacher (2003), of the Ultimatum game, which has essentially the same structure as our two-stage game with punishment. In that case, the treatment without information on the co-player's past behavior yields a noticeably lower level of cooperation than the treatment with information. Nevertheless, even in the no-information treatment, both the level of cooperation (in the form of fair sharing) and of punishment (in the form of rejection of small offers) are remarkably high.

A serious criticism of the model presented in this paper is thus that it does not seem to account for the pro-social behavior shown by players who know that reputation-building is impossible. We believe that this effect is due to a mal-adaptation. Our evolutionary past has not prepared us to expect anonymity. In hunter-gatherer societies and in rural life, it is not often that one can really be sure to be unobserved. Even in modern life, the long phase of childhood is usually spent under the watchful eyes of parents, educators or age-peers. Ingenious experiments uncover our tendency to over-react to the slightest cues indicating that somebody may be watching (for instance, the mere picture of an eye, see Haley and Fessler (2005) and Bateson et al. (2006), or three dots representing eyes and mouth, see Rigdon et al. (2009)). The idea of personal deities scrutinizing our behavior, which seems to be almost universal, is probably a projection of this deep-seated conviction (Johnson and Bering, 2006). The concept of conscience was famously described, by Mencken, as 'the inner voice that warns us somebody may be looking' (see Levin, 2009).

In several experimental papers, however, the role of reputation is very explicit. In Rand et al. (2009), players are engaged in fifty rounds of the public good game with incentives, always with the same three partners. Hence they know the past actions of their co-players. In this case, we can be sure that $\mu > \gamma/b$. Thus in a homogeneous $\mathbf{O_C}$-population, **R** should dominate **N**. Moreover, as the leverage for both punishment and reward is 1:3 in this experiment (as in many others), an $[\mathbf{O_C}, \mathbf{R}]$-population obtains a payoff $b - c + \beta - \gamma$ which is substantially larger than that of an $[\mathbf{O_C}, \mathbf{P}]$-population. In the experiment, rewarding performs indeed much better than punishing, and Rand *et al.* conclude that 'Positive reciprocity should play a larger role than negative reciprocity in maintaining public cooperation in repeated situations.'

Nevertheless, according to our model, **P**-players ought to invade. This seems counterintuitive. Punishers do not have to pay for an incentive (since everyone cooperates), but they will nevertheless be rewarded, since they cooperate in the public goods stage. Thus $[\mathbf{O_C}, \mathbf{P}]$ should take over, thereby lowering the average payoff. By contrast, in the repeated game considered by Rand *et al.*, it is clear that cooperative players who have not been rewarded by their co-player in the previous round will feel cheated, and stop rewarding that co-player. They will not be impressed by the fact that the co-player is still providing an incentive by punishing defectors instead. In other words, in this experiment rewards are not only seen as incentives, but as contributions in their own right, in a repeated prisoner's dilemma game. Players will reciprocate not only for the public goods behavior, but for the 'mutual reward game' too. In fact, if there had been two players only in the experiment by Rand *et al*, it would reduce to a repeated prisoner's dilemma game with 100 rounds.

This aspect is not covered in our model, where the incentives are only triggered by the behavior in the public goods stage, but not by previous incentives. In particular, rewarding behavior cannot be rewarded, and fines do not elicit counter-punishment. This facilitates the analysis of incentives as instruments for promoting cooperation, but it obscures the fact that in real life, incentives have to be viewed as economic exchanges in their own right.

A similar experiment as in Rand *et al.* was studied by Milinski et al. (2002), where essentially the public goods rounds alternate with an indirect reciprocity game (see also Panchanathan and Boyd, 2004). Helping, in such an indirect reciprocity game, is a form of reward. In Milinski's experiment, punishment was not allowed, but in Rockenbach and

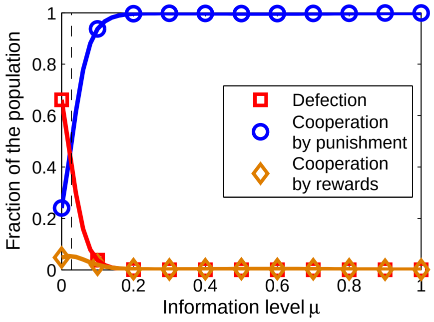


Figure 2.4: Stationary distribution in the case of ”cheap” incentives. If $\gamma < \beta$, a homogeneous population of rewarding cooperators would obtain the maximum payoff. While cooperation can easily be established, players clearly prefer to use the stick instead of the carrot. Thus, there is only a marginal difference to the case with $\gamma = \beta$ (Fig. 2.2, right column). Parameter values: Population size $M = 100$, selection strength $s = 1/10$; $b = 4, c = 1, \beta = 2, \gamma = 3/4$.

Milinski (2006), both types of incentives could be used. Groups were rearranged between rounds, as players could decide whether to leave or to stay. Players knew each other's past behavior in the previous public goods rounds and the indirect reciprocity rounds (but not their punishing behavior). It was thus possible to acquire a reputation as a rewarder, but not as a punisher. This treatment usually led to a very cooperative outcome, with punishment focused on the worst cheaters, and a significant interaction between reward and punishment.

In our numerical examples, we have usually assumed $\gamma = \beta$, but stress that this does not affect the basic outcome (see Fig. 2.4 for the case $\gamma < \beta$). In most experiments, the leverage of the incentive is assumed to be stronger. Clearly, this encourages the Recipients to use incentives (Carpenter, 2007; Egas and Riedl, 2008; Vyrastekova and van Soest, 2008). But it has been shown (Carpenter, 2007; Sefton et al., 2007) that many are willing to punish exploiters even if it reduces their own account by as much as that of the punished player. In the Trust game, it is also usually assumed that the second stage is a zero-sum game. In most of the (relatively few) experiments on rewarding, the leverage is 1:1 (Walker and Halloran, 2004; Sefton et al., 2007); in Rockenbach and Milinski (2006) and Rand et al. (2009) it is 1:3. In Vyrastekova and van Soest (2008), it is shown that increasing this leverage makes rewarding more efficient. In our view, it is natural to assume a high benefit-to-cost ratio in the first stage (the occasion for a public good game is precisely the situation when mutual help is needed), but it is less essential that a high leverage also applies in the second stage. Punishment, for instance, can be very costly it the other player retaliates, as seems quite natural to expect (at least in pairwise interactions; in N-person games, sanctions can be inexpensive if the majority punishes a single cheater).

For the sake of simplicity, we have not considered the probability of errors in implementation. But it can be checked in a straightforward way that the results are essentially unchanged if we assume that with a small probability $\epsilon > 0$, an intended donation fails (either due to a mistake of the player, or to unfavorable conditions). The other type of errors in implementation (namely helping without wanting it) seems considerably less plausible. We note that in a homogeneous $[\mathbf{O_C}, \mathbf{P}]$-population, usually there is no need to punish co-players, and hence no way of building up a reputation as a punisher. But if errors in implementation occur, there will be opportunities for punishers to reveal their true colours. In Sigmund (2010), it is shown that if there are sufficiently many rounds of the game, occasional errors will provide enough opportunities for building up a reputation.

A Classification of the dynamics in large populations

Let us first analyze a reduced system where all donors are opportunists and where recipients either reward cooperators, punish non-cooperators or do not provide any incentive. Hence we consider the following extract of the payoff matrix:

$*$	$\mathbf{N}$	$\mathbf{P}$	$\mathbf{R}$
$\mathbf{O_C}$	$(-\bar{\mu}c, \bar{\mu}b)$	$(-c, b)$	$(\beta - c, b - \gamma)$
$\mathbf{O_D}$	$(0, 0)$	$(-\mu c - \bar{\mu}\beta, \mu b - \bar{\mu}\gamma)$	$(\mu(\beta - c), \mu(b - \gamma))$

Table 2.2: Payoff matrix for the strategies $\mathbf{O_C}$, $\mathbf{O_D}$, respectively $\mathbf{N}$, $\mathbf{P}$ and $\mathbf{R}$.

From this payoff matrix we may draw the following conclusions:

(i) In a homogeneous $\mathbf{O_C}$ population, $\mathbf{P}$ always dominates $\mathbf{N}$ and $\mathbf{R}$. Additionally, an $\mathbf{R}$ player obtains a higher payoff than an $\mathbf{N}$ player if and only if $b - \gamma > \bar{\mu}b$, i.e. iff

$$\mu > \gamma/b \tag{2.1}$$

(ii) Similarly, in a homogeneous $\mathbf{O_D}$ population, $\mathbf{N}$ is always dominated by $\mathbf{R}$. $\mathbf{N}$ is also dominated by $\mathbf{P}$ iff $\mu b - \bar{\mu}\gamma > 0$, i.e. iff

$$\mu > \gamma/(b + \gamma) \tag{2.2}$$

$\mathbf{P}$ also dominates $\mathbf{R}$ iff $\mu b - \bar{\mu}\gamma > \mu(b - \gamma)$, i.e. iff

$$\mu > 1/2 \tag{2.3}$$

(iii) If incentives are used ($\mathbf{P}$ or $\mathbf{R}$), $\mathbf{O_C}$ dominates $\mathbf{O_D}$; in the absence of incentives ($\mathbf{N}$), this relation is reversed.

In particular, it follows that $\mathbf{N}$ is strictly dominated by $\mathbf{P}$ if $\mu > \gamma/(b + \gamma)$. Iterated elimination of strictly dominated strategies then leads to the conclusion that the only possible outcome of the dynamics is $[\mathbf{O_C}, \mathbf{P}]$.

However, if $\mu < \gamma/(b + \gamma)$ we find a more interesting behaviour. There is a uniqe fixed point M in the interior of the square spanned by the strategies $[\mathbf{O_C}, \mathbf{N}]$, $[\mathbf{O_C}, \mathbf{R}]$, $[\mathbf{O_D}, \mathbf{N}]$ and $[\mathbf{O_D}, \mathbf{R}]$. This fixed point is surrounded by spiralling orbits. The asymptotic behaviour of these orbits depends on the initial value of

$$Z_{\mathbf{NR}} := \frac{x_{[\mathbf{O_C},\mathbf{N}]} \cdot x_{[\mathbf{O_D},\mathbf{R}]}}{x_{[\mathbf{O_C},\mathbf{R}]} \cdot x_{[\mathbf{O_D},\mathbf{N}]}}.$$

If $Z_{\mathbf{NR}}(0) < 1$ all orbits converge to the boundary of the square, whereas $Z_{\mathbf{NR}}(0) > 1$ implies that M is a global attractor. Finally, if $Z_{\mathbf{NR}}(0) = 1$, then M is surrounded by periodic orbits (see Hofbauer and Sigmund, 1998). Independently of the initial condition, a population in this fixed point M yield the payoff

$$\mu \left(-b + c - c^2/\beta + b^2/\gamma\right),$$

whereas both absent strategies, $[\mathbf{O_C}, \mathbf{P}]$ and $[\mathbf{O_D}, \mathbf{P}]$ would obtain the payoff

$$-\gamma + \mu \left(b + c - c^2/\beta + b^2/\gamma\right).$$

Hence, a population in M can be invaded by punishers if and only if

$$\mu > \gamma/(2b). \tag{2.4}$$

If we also allow for the strategies **AllC** and **AllD** we do not see any additional bifurcations as the payoffs of these strategies do not depend on μ. Hence, the dynamics of the game is fully described by the thresholds (2.1) - (2.4), which can be arranged in two possible ways:

(i) If $\gamma < b/2$, we have $\frac{\gamma}{2b} < \frac{\gamma}{\gamma+b} < \frac{\gamma}{b} < 1/2$

(ii) Otherwise we obtain $\frac{\gamma}{2b} < \frac{\gamma}{\gamma+b} < 1/2 < \frac{\gamma}{b}$

However, both cases induce the same long-term dynamics (see also Fig. 2.5, which depicts the borderline case $\gamma = b/2$): For any value of μ, a population consisting of $\mathbf{O_C}$ and **AllD** evolves either towards punishment enforced cooperation, $[\mathbf{O_C}, \mathbf{P}]$, or towards a purely selfish regime, $[\mathbf{AllD}, \mathbf{N}]$ resp. $[\mathbf{AllD}, \mathbf{R}]$ (Fig. 2.5, first column). As **AllD** is always weakly dominated by its opportunistic counterpart, $\mathbf{O_D}$ might invade (Fig. 2.5, second column). In this case the eventual outcome depends on the information level: If $\mu > \gamma/(b + \gamma)$, $[\mathbf{O_C}, \mathbf{P}]$ mutants succeed in an $\mathbf{O_D}$ population and lead to stable cooperation, independently of the additional inequalities $\mu > 1/2$ or $\mu > \gamma/b$ (Fig. 2.5, last two graphs in the last column). For $\mu < \gamma/(b + \gamma)$, however, only the possibility of rewards allows an $\mathbf{O_C}$ minority to invade a homogeneous $\mathbf{O_D}$ population, which may lead to oscillations between reward-driven cooperation and defection. If $\mu > \gamma/(2b)$ these cycles are unstable and open the way for $[\mathbf{O_C}, \mathbf{P}]$ (Fig. 2.5, upper two graphs in the last column).

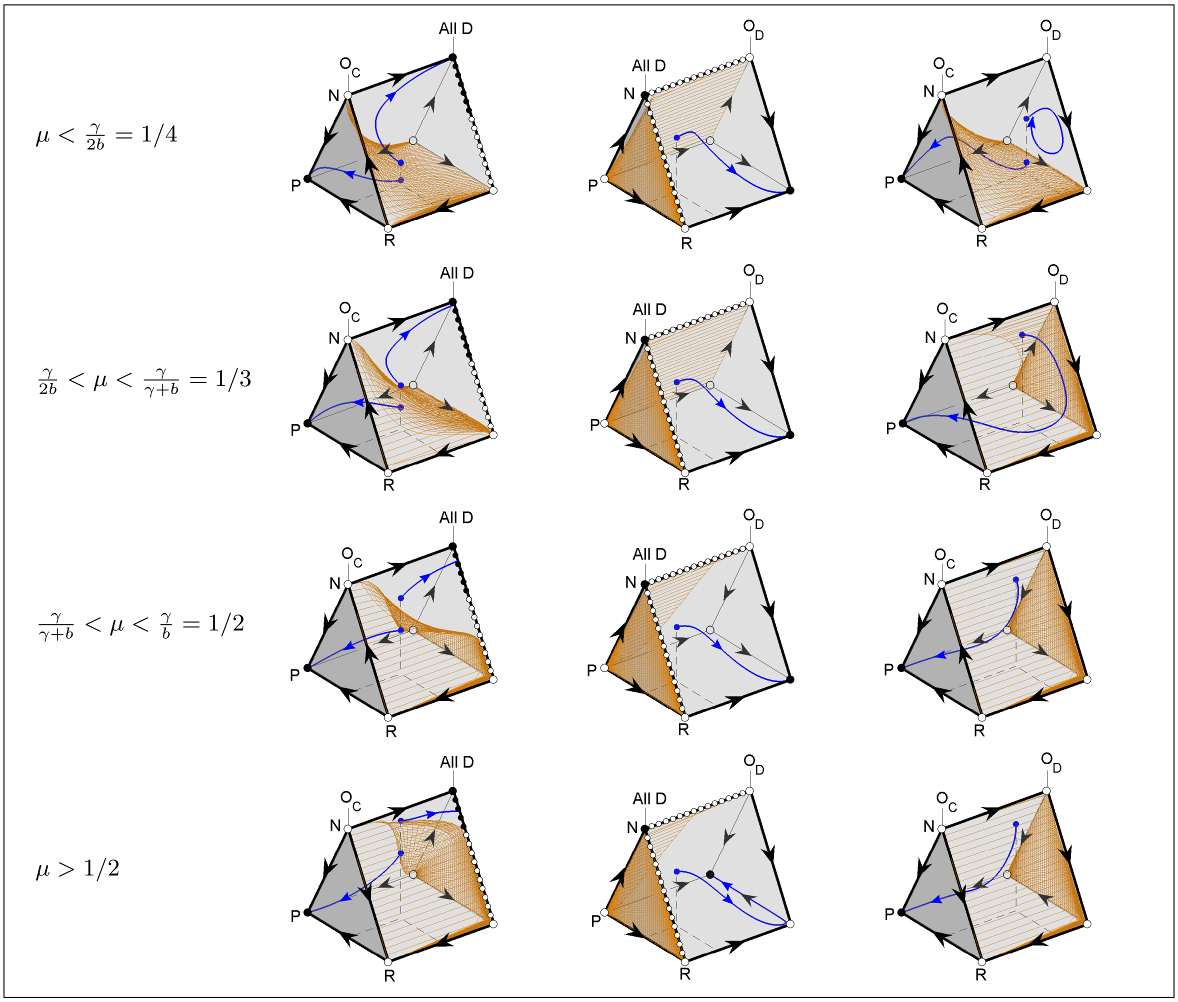


Figure 2.5: Replicator dynamics for $b = 4,\ c = 1,\ \beta = \gamma = 2$ and $\mu = 15\%, 30\%, 45\%$ resp. $\mu = 60\%$. The arrows on the edges indicate the direction of the dynamics if all other strategies are absent. Black points represent Nash-equilibria, white points indicate unstable fixed points. The orange grid is the manifold that separates initial values with different asymptotic behaviour (i.e., a separatrix). The blue curves represent the typical dynamics for a given initial population.
Note that this choice of parameters implies $\gamma/b = 1/2$; therefore there occur two bifurcations between the third and the fourth row (in the last row $\mathbf{R}$ dominates $\mathbf{N}$ in a homogeneous $\mathbf{O_C}$ population and $\mathbf{P}$ dominates $\mathbf{R}$ in a homogeneous $\mathbf{O_D}$ population). The dynamics in the interior of the prism, however, is the same no matter whether $\gamma/b < 1/2$ or $\gamma/b > 1/2$.

B Numerical data for the case of small mutation rates

B.1 Transition probabilities

The transition probabilities for the adiabatic case are calculated according to Nowak (2006a), see also Sigmund (2010). The following tables (a_{ij}) show the fixation probabilities of a mutant i, invading a resident strategy j for different values of μ. For low information rates, $[\mathbf{O_D}, \mathbf{R}]$ is the only strategy which can invade the asocial $[\mathbf{O_D}, \mathbf{N}]$ with a higher probability than $1/M$ (the neutral fixation probability).
Parameter values: $M = 100, s = 1/10, b = 4, c = 1, \beta = \gamma = 2$.

$\mu = 10\%$	$[\mathbf{O_C}, \mathbf{N}]$	$[\mathbf{O_C}, \mathbf{R}]$	$[\mathbf{O_C}, \mathbf{P}]$	$[\mathbf{O_D}, \mathbf{N}]$	$[\mathbf{O_D}, \mathbf{R}]$	$[\mathbf{O_D}, \mathbf{P}]$
$[\mathbf{O_C}, \mathbf{N}]$	0.010	0.162	0.001	**0.000**	0.068	0.204
$[\mathbf{O_C}, \mathbf{R}]$	0.000	0.010	0.000	**0.000**	0.088	0.229
$[\mathbf{O_C}, \mathbf{P}]$	0.041	0.202	0.010	**0.000**	0.014	0.085
$[\mathbf{O_D}, \mathbf{N}]$	0.094	0.074	0.003	**0.010**	0.003	0.138
$[\mathbf{O_D}, \mathbf{R}]$	0.000	0.000	0.000	**0.023**	0.010	0.158
$[\mathbf{O_D}, \mathbf{P}]$	0.004	0.001	0.000	**0.000**	0.000	0.010

$\mu = 30\%$	$[\mathbf{O_C}, \mathbf{N}]$	$[\mathbf{O_C}, \mathbf{R}]$	$[\mathbf{O_C}, \mathbf{P}]$	$[\mathbf{O_D}, \mathbf{N}]$	$[\mathbf{O_D}, \mathbf{R}]$	$[\mathbf{O_D}, \mathbf{P}]$
$[\mathbf{O_C}, \mathbf{N}]$	0.010	0.082	0.000	**0.000**	0.015	0.001
$[\mathbf{O_C}, \mathbf{R}]$	0.000	0.010	0.000	**0.005**	0.069	0.078
$[\mathbf{O_C}, \mathbf{P}]$	0.120	0.202	0.010	**0.005**	0.036	0.066
$[\mathbf{O_D}, \mathbf{N}]$	0.073	0.018	0.000	**0.010**	0.002	0.022
$[\mathbf{O_D}, \mathbf{R}]$	0.006	0.000	0.000	**0.060**	0.010	0.078
$[\mathbf{O_D}, \mathbf{P}]$	0.177	0.049	0.000	**0.004**	0.000	0.010

$\mu = 50\%$	$[\mathbf{O_C}, \mathbf{N}]$	$[\mathbf{O_C}, \mathbf{R}]$	$[\mathbf{O_C}, \mathbf{P}]$	$[\mathbf{O_D}, \mathbf{N}]$	$[\mathbf{O_D}, \mathbf{R}]$	$[\mathbf{O_D}, \mathbf{P}]$
$[\mathbf{O_C}, \mathbf{N}]$	0.010	0.011	0.000	**0.000**	0.000	0.000
$[\mathbf{O_C}, \mathbf{R}]$	0.009	0.010	0.000	**0.048**	0.049	0.001
$[\mathbf{O_C}, \mathbf{P}]$	0.201	0.202	0.010	**0.071**	0.072	0.047
$[\mathbf{O_D}, \mathbf{N}]$	0.052	0.001	0.000	**0.010**	0.000	0.000
$[\mathbf{O_D}, \mathbf{R}]$	0.051	0.000	0.000	**0.100**	0.010	0.009
$[\mathbf{O_D}, \mathbf{P}]$	0.246	0.137	0.001	**0.102**	0.011	0.010

B.2 Typcial transitions between states

For different information levels, the following table displays the distribution of strategies which succeed in invading a homogeneous [$\mathbf{O_D}$,**N**] population. For small information levels especially, [$\mathbf{O_D}$,**R**] is best at overcoming this asocial state. The table shows the result of a simulation of the stochastic process with 10^7 mutations, population size $M = 100$, selection strength $s = 1/10$, $b = 4, c = 1, \beta = \gamma = 2$.

	$[\mathbf{O_C}, \mathbf{N}]$	$[\mathbf{O_C}, \mathbf{R}]$	$[\mathbf{O_C}, \mathbf{P}]$	$[\mathbf{O_D}, \mathbf{N}]$	$[\mathbf{O_D}, \mathbf{R}]$	$[\mathbf{O_D}, \mathbf{P}]$
$\mu = 0\%$	0.000	0.001	0.000	–	**0.999**	0.000
$\mu = 10\%$	0.000	0.003	0.001	–	**0.996**	0.000
$\mu = 20\%$	0.000	0.017	0.011	–	**0.970**	0.002
$\mu = 30\%$	0.000	0.068	0.040	–	**0.858**	0.034
$\mu = 40\%$	0.000	0.191	0.191	–	**0.381**	0.238

The next table displays the distribution of strategies which immediately preceded a homogeneous [$\mathbf{O_C}$;**P**] populations. If there is little information about co-players, the state of punishing cooperators is mostly reached via rewards (Simulation with 10^7 mutations, $M = 100, s = 1/10, b = 4, c = 1, \beta = \gamma = 2$).

	$[\mathbf{O_C}, \mathbf{N}]$	$[\mathbf{O_C}, \mathbf{R}]$	$[\mathbf{O_C}, \mathbf{P}]$	$[\mathbf{O_D}, \mathbf{N}]$	$[\mathbf{O_D}, \mathbf{R}]$	$[\mathbf{O_D}, \mathbf{P}]$
$\mu = 0\%$	0.228	**0.644**	–	0.000	0.128	0.000
$\mu = 10\%$	0.394	**0.457**	–	0.002	0.144	0.004
$\mu = 20\%$	0.234	**0.479**	–	0.014	0.212	0.062
$\mu = 30\%$	0.098	**0.419**	–	0.018	0.211	0.254
$\mu = 40\%$	0.014	0.169	–	0.008	0.205	**0.604**

Chapter 3

Local replicator dynamics: A simple link between deterministic and stochastic models of evolutionary game theory

Abstract

Classical replicator dynamics assumes that individuals play their games and adopt new strategies on a global level: Each player interacts with a representative sample of the population and if a strategy yields a payoff above the average, then it is expected to spread. In this article we connect evolutionary models for infinite and finite populations: While the population itself is infinite, interactions and reproduction occurs in random groups of size N. Surprisingly, the resulting dynamics simplifies to the traditional replicator system with a slightly modified payoff matrix. The qualitative results, however, mirror the findings for finite populations, in which strategies are selected according to a probabilistic Moran process. In particular we derive a one-third law that holds for any population size. In this way, we show that the deterministic replicator equation in an infinite population can be used to study the Moran process in a finite population and vice versa. We apply the results to three examples to shed light on the evolution of cooperation in the iterated prisoner's dilemma, on risk aversion in coordination games and on the maintenance of dominated strategies.

3.1 Introduction

Evolutionary game theory is the appropriate mathematical framework whenever the fitness of an individual does not only depend on its own phenotype but also on the composition of the whole population (Maynard Smith, 1982; Nowak and Sigmund, 2004). Such biological interdependencies are wide-spread. Game theoretic models help us to understand molecular biological processes (Pfeiffer and Schuster, 2005), animal behaviour (Dugatkin and Reeve, 1998) or plant growth (Falster and Westoby, 2003). Even the evolution of abstract cultural entities, such as languages (Niyogi, 2006) or norms (Gintis et al., 2005) can be captured with equations. In this way, mathematical arguments reveal the logic of indirect speech (Pinker et al., 2008) as well as the origin of moral systems (Uchida and Sigmund, 2010).

The classical approach to evolutionary games is replicator dynamics (Taylor and Jonker, 1978; Hofbauer and Sigmund, 1998): When a certain strategy leads to a fitness above the average, then this behaviour is expected to spread. More formally, suppose the individuals of a population can choose among n strategies. A player with strategy i, interacting with a j-player, obtains a payoff a_{ij}. The relative frequency of i-individuals within the population is denoted by x_i. If we collect these frequencies to a vector $x = (x_1, \dots, x_n)^T$ and the payoffs to a matrix $A = (a_{ij})$, the expected fitness of strategy i can be written as $\pi_i(x) := (Ax)_i = a_{i1}x_1 + \ldots + a_{in}x_n$. Over the whole population, this results in an average payoff of $\bar{\pi}(x) := x \cdot Ax$. According to the replicator equation, the fraction of i-players grows whenever $\pi_i(x)$ exceeds $\bar{\pi}(x)$:

$$\dot{x}_i = x_i \cdot [(Ax)_i - x \cdot Ax] \tag{3.1}$$

The dot above x_i on the left hand's side denotes the time derivative, dx_i/dt.

There are several reasons for the prominent role of replicator dynamics: Firstly, equation (3.1) is relatively simple and mathematically well understood (Hofbauer and Sigmund, 2003). It is equivalent to the famous Lotka-Volterra model in population ecology (Hofbauer, 1981). Additionally, there are beautiful connections between replicator dynamics and the concepts of classical game theory (Fudenberg and Tirole, 1991; Weibull, 1995). For example, if strategy i is *dominated*, meaning that there is another strategy which is always better, then replicator dynamics will lead to the extinction of i. Furthermore, *strict Nash equilibria*, i.e. strategies that are optimal against themselves, are asymptotically stable rest points of equation (3.1).

The conceptual beauty of replicator dynamics has its price. To be valid, several conditions have to be met: (i) The population is *well mixed*, meaning that any two players interact with the same probability. (ii) Before reproduction, individuals play against a representative sample of the population. (iii) Players may choose among a finite set of strategies, and (iv) the population needs to be infinite.

The game-theoretical toolbox, however, is evolving too. In the last two decades there have been several attempts to relax the above assumptions. Instead of considering well mixed populations, various models explore games in which players only interact with their neighbors on a lattice (Nowak and May, 1992; Durrett and Levin, 1994; Hauert and Döbeli, 2004; Helbing and Yu, 2009) or on a random graph (Abramson and Kuperman, 2001; Lieberman et al., 2005; Ohtsuki et al., 2006; Lehmann et al., 2007a; Szabó and Fáth, 2007; Santos et al., 2008). Additionally, some authors investigated the effect of a non-representative sample of interaction partners (Sánchez and Cuesta, 2005; Roca et al., 2006; Traulsen et al., 2007; Wölfing and Traulsen, 2009). The introduction of continuous strategy spaces has led to adaptive dynamics theory (Hofbauer and Sigmund, 1994; Geritz et al., 1998; Dieckmann et al., 2006).

To study the effect of finite populations, Nowak et al. (2004) consider a group of N individuals who may choose among two strategies, R and M. Evolution takes place according to a Moran process: In each time step, one randomly chosen individual dies. Its place is filled with the copy of another member of the population, whereby players with a high payoff have a better chance to be copied. The resulting dynamics is a stochastic birth-death process. Evolutionary success is measured with fixation probabilities. For example, in a state with R-types only, one may calculate the likelihood that one M-mutant can invade and take over the whole population. If this probability $\rho_{M \to R}$ exceeds $1/N$, then *selection favours M replacing R*.

There are interesting connections between replicator dynamics and the Moran-process. For example, in the limit of large populations, the diffusion approximation of the Moran-process converges to a variant of the replicator equation (Traulsen et al., 2005). A more subtle connection is called the one-third law (Taylor et al., 2004; Ohtsuki et al., 2007a; Bomze and Pawlowitsch, 2008): Consider a game with payoff matrix

$$\begin{array}{c|cc} & R & M \\ \hline R & a & b \\ M & c & d \end{array} \tag{3.2}$$

For $a > c$ and $d > b$, replicator dynamics predicts that both pure strategies, R and M, are evolutionarily stable. Additionally, there is an interior fixed point at $x^* = (a-c)/(a-b-c+d)$, whereby x denotes the frequency of M-players. For large population sizes N, it can be shown that selection favours M replacing R if this fixed point is below one third, i.e.

$$\rho_{M\to R} > 1/N \Leftrightarrow x^* < 1/3. \tag{3.3}$$

A similar result holds for *risk dominance* (Harsanyi and Selten, 1988): Strategy M has the bigger basin of attraction if M has a higher chance to replace R than vice versa:

$$\rho_{M\to R} > \rho_{R\to M} \Leftrightarrow x^* < 1/2. \tag{3.4}$$

Even if replicator dynamics and the Moran process are closely related for large populations, they can lead to very different conclusions if the population size N is small. Dominated strategies may spread in finite populations (Taylor et al., 2004) and regimes that are stable according to replicator dynamics may be invaded under the Moran process and vice versa (Nowak et al., 2004). Therefore, several recent studies have compared the dynamics in finite respectively infinite populations separately, e.g. Imhof et al. (2005) or Hilbe and Sigmund (2010).

In this paper, we aim to connect both approaches: While the population itself is infinite, individuals interact in randomly formed groups of size N. In each group, one member may adopt the strategy of a co-player. Strategies that yield a high payoff are more likely to be adopted. After reproduction, groups are formed anew. This scenario was used by Nakamaru et al. (1997, 1998), Nakamaru and Iwasa (2006) and Rand et al. (2010) to approximate games on a lattice, in which each player interacts with its $N-1$ neighbors. Here we show that the resulting dynamics can be described by a modified replicator equation. Instead of payoff matrix A, the modified matrix $\tilde{A}_N = A-(A+A^T)/N$ governs the evolutionary outcome. We refer to this modified system as the *local replicator dynamics*. Despite its deterministic nature, the local replicator dynamics strongly mirrors the findings for finite populations. We will show that there is a one-to-one map from the fixation probabilities according to the Moran process to the fixed points of the local replicator equation. In particular the one-third rule holds for any population size N. Thus, a simple transformation of the payoff matrix allows us to use classical replicator dynamics to analyze games in finite populations.

The remainder of this article is organized as follows: In Section 3.2 we derive the local replicator equation and in Section 3.3 we discuss its properties. In Section 3.4, we present three examples for the resulting dynamics: First, we show how the evolution of cooperation in the iterated prisoner's dilemma is influenced by the size of the group. Then we provide a simple model for the emergence of risk aversion in a coordination game. In the last example, we explain how dominated strategies can succeed in finite interaction neighborhoods. In Section 3.5, we extend our model: Instead of considering a fixed N, we assume that the size of a given group is a random variable. Still, the dynamics can be described with the local replicator equation by simply replacing N with a proper mean of all possible group sizes. In Section 3.6, finally, we give a summary and discussion of our results.

3.2 Derivation of the local replicator dynamics

We consider an infinite population whose individuals are randomly assigned to groups of equal size. Within their groups, players engage in pairwise interactions that determine the payoff of each individual. More specifically, let k be the number of possible strategies and $A = (a_{ij})$ the payoff matrix. For a given group of size N, denote by n_i the number of players with strategy i and by $\vec{n} = (n_1, ..., n_k)$ the vector of those numbers, with $|\vec{n}| := n_1 + \ldots + n_k = N$. If individuals interact with every other group member and self-interactions are excluded, an individual with strategy i yields the average payoff

$$\pi_i(\vec{n}) = \frac{1}{N-1}\left[n_1 a_{i1} + \ldots + (n_i - 1)a_{ii} + \ldots + n_k a_{ik}\right]. \tag{3.5}$$

After these interactions, one group member is chosen at random to compare its payoff with another randomly chosen group member. We follow the update rule in Traulsen et al. (2005, 2006a) and assume that an i-player adopts the role model's strategy j with probability

$$p_{ij} = \frac{1}{2} + \frac{s}{2}\frac{\pi_j(\vec{n}) - \pi_i(\vec{n})}{\Delta\pi_{\max}}, \tag{3.6}$$

where $\Delta\pi_{\max}$ is the maximum possible payoff difference and $0 < s \leq 1$ measures the strength of selection. For weak selection ($s \approx 0$), a coin toss essentially determines whether the role model is imitated; if the selection parameter approaches 1, this decision is increasingly biased towards the strategy with higher payoff. Since players are selected randomly for the imitation process, this leads to an expected change in the number of i-players of

$$\Delta n_i = \sum_{j=1}^{k} \frac{n_i}{N}\frac{n_j}{N}(p_{ji} - p_{ij}) = \frac{s}{\Delta\pi_{\max}}\frac{n_i}{N}[\pi_i(\vec{n}) - \bar{\pi}(\vec{n})]. \tag{3.7}$$

The term $\bar{\pi}(\vec{n})$ denotes the average payoff of all players in the group, $\bar{\pi}(\vec{n}) = \sum_{j=1}^{k} n_j\pi_j(\vec{n})/N$. By using n_j/N instead of $n_j/(N-1)$, we have implicitly assumed that players may imitate themselves (as for example in Ohtsuki et al., 2006; Ohtsuki and Nowak, 2006). However, if an individual chooses itself as role model, then nothing will change. Overall, this simplification only changes the velocity of evolution but not its direction.

According to Eq. (3.7), the expected change of strategies within a certain group can be described through a replicator equation with local frequencies n_i/N. In order to calculate the aggregate outcome, we have to sum up over all possible groups with $|\vec{n}| = N$. If x_i denotes the global frequency of strategy i and if groups are formed according to a multinomial distribution, this results in the following equation for the expected change of global strategy abundances,

$$\dot{x}_i = \sum_{|\vec{n}|=N} \frac{N!}{n_1! \,\ldots\, n_k!}\, x_1^{n_1} \cdot \ldots \cdot x_n^{n_k} \cdot \frac{n_i}{N}\big[\pi_i(\vec{n}) - \bar{\pi}(\vec{n})\big], \tag{3.8}$$

where the constant factor $s/\Delta\pi_{\max}$ has been ommited since it only affects the timescale of evolution. We call Eq. (3.8) the local replicator dynamics of the system. In general, the expression on the right hand's side is laborious to compute: The polynomial has degree N and the number of summands increases exponentially in the number of strategies k. As it turns out, this expression can be simplified considerably:

Proposition 1 (A simple representation for the local replicator dynamics)
The local replicator dynamics (3.8) can be rewritten as

$$\dot{x}_i = x_i \cdot \big[(\tilde{A}x)_i - x \cdot \tilde{A}x\big], \tag{3.9}$$

with $\tilde{A} := A - \frac{A+A^T}{N}$. *As usual,* A^T *denotes the transposed payoff matrix.*

Proof: See appendix.

This representation of the local replicator equation can be interpreted by slightly rearranging the terms to

$$\dot{x}_i = x_i \cdot \big[(Ax)_i - \frac{(N-2)\cdot x^T Ax + (Ax)_i + (A^T x)_i}{N}\big], \tag{3.10}$$

meaning that a player with strategy i still gets an expected payoff of $(Ax)_i$. However, being an i-player in a finite group of size N leads to a bias in the average payoff of the group, on the one hand due to the own payoff $(Ax)_i$ and on the other hand due to the expected influence on the others' payoffs $(A^T x)_i$.

Consequently, there are two alternative mechanisms that promote the evolution of a certain strategy under the local replicator dynamics: Strategy i may spread either because it leads to a high own payoff $(Ax)_i$, or because it leads the co-players to have a low payoff, i.e. it decreases $(A^T x)_i$. In particular, it may pay off to be spiteful and to diminish the fitness of other group members (see also Nowak et al., 2004). As group size increases, the impact of spiteful behaviour disappears, since the own influence on the co-players' payoffs becomes negligible. In the limit of infinite group sizes, spite becomes completely irrelevant.

In this section we have derived the local replicator dynamics for an imitation process, where individuals with higher payoff are imitated more often than others. One could also investigate a model in which sucessful players produce more offspring; such an attempt is shown in the appendix, leading to the same dynamics. Irrespective of its derivation, the local replicator equation has interesting properties: It connects the results of the Moran process in finite populations with the classical replicator dynamics in infinite populations. This is shown in the next section.

3.3 Properties of the local replicator dynamics

Representation (3.9) makes it straightforward to derive the properties of the local replicator dynamics: We only have to plug in the values of $\tilde{A}_N = A - (A + A^T)/N$, to transfer the well-known results of the classical replicator equation (see, for example, in Hofbauer and Sigmund, 1998). This is done in the following two propositions:

Proposition 2 (Analytical properties of the local replicator equation)

(i) *The simplex* $S_k = \big\{(x_1, ..., x_k) : \ x_1 + ... + x_k = 1\big\}$ *and all its subfaces are invariant under the local replicator dynamics.*

(ii) *Quotient rule:* $\dot{\left(\frac{x_i}{x_j}\right)} = \frac{x_i}{x_j} \cdot \left[(N-1)\frac{(Ax)_i - (Ax)_j}{N} + \frac{(A^T x)_j - (A^T x)_i}{N}\right]$

(iii) *In the special case of* $N = 2$*, the values of the diagonal entries of* A *do not influence the dynamics.*

(iv) *For zero sum games (i.e. if* $A = -A^T$*) or in the case of infinite group sizes (*$N \to \infty$*), the local replicator equation coincides with the classical version.*

In applications, it is often important to know which states can be stably established. To this end, one considers a homogeneous population with strategy R, that is invaded by a rare mutant with an alternative strategy M. If the evolutionary dynamics leads to the extinction of any such mutant, then R is called *evolutionarily stable* or an *ESS*.

Proposition 3 (Evolutionary Stability)
Consider a game with 2 *strategies and payoff matrix*

$$\begin{array}{c|cc} & R & M \\ \hline R & a & b \\ M & c & d \end{array} \tag{3.11}$$

Let x *denote the fraction of* M*-players in the population. Then the local replicator dynamics can be written as*

$$\dot{x} = x(1-x) \cdot \left(\frac{(N-1)c - b - (N-2)a}{N} + \frac{(N-2)(a-b-c+d)}{N} x \right). \tag{3.12}$$

In particular, the resident strategy R *is (strictly) evolutionarily stable, if*

$$(N-2)a + b > (N-1)c. \tag{3.13}$$

Again, this proposition is simply an adapted version of the analogous result for classical replicator dynamics. Obviously, the above ESS-condition reduces to the usual Nash condition $a > c$ for infinite groups. For finite populations, however, several definitions have been suggested to capture evolutionary stability, see Wild and Taylor (2004). The condition $(N-2)a + b > (N-1)c$ is included, for example, in the definitions of Maynard Smith (1988) and Schaffer (1988). Nowak et al. (2004) suggest an additional ESS-criterion for finite populations. Beside the above condition that rare mutants have a lower fitness than the resident type, they require the mutant strategy to have a fixation probability lower than $1/N$. When selection is weak (i.e. the impact of the payoff in the game has a small effect on the reproductive success of a strategy), they find that this fixation probability is approximately given by

$$\rho_{M\to R} = \frac{1}{N} + \frac{1}{6N}(\alpha N - \beta)s, \tag{3.14}$$

with $\alpha = d + 2c - b - 2a$ and $\beta = 2d + b + c - 4a$. Again, s measures the strength of selection. Approximation (3.14) is valid if $Ns \ll 1$ and will be called the weak selection fixation probability. The reverse probability $\rho_{R\to M}$ that a single R-mutant reaches fixation in an otherwise homogeneous population of M-players can be derived with a symmetry argument. Our next aim is to show that there is a strong connection between these fixation probabilities and the local replicator dynamics.

Proposition 4 (One-third law)
Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *and* $\tilde{A}_N = A - \frac{A+A^T}{N} = \begin{pmatrix} \tilde{a} & \tilde{b} \\ \tilde{c} & \tilde{d} \end{pmatrix}$ *be payoff matrices for a game with two strategies, R and M.*

(i) The weak selection fixation probabilities of the Moran process with payoff matrix A can be infered from the replicator dynamics of the game with matrix $\tilde{A}_N$*:*

$$\rho_{M\to R} = \tfrac{1}{N} + \tfrac{\tilde{D}}{2}(\tilde{x}^* - \tfrac{1}{3})s, \qquad \rho_{R\to M} = \tfrac{1}{N} + \tfrac{\tilde{D}}{2}(\tfrac{2}{3} - \tilde{x}^*)s. \tag{3.15}$$

Here, $\tilde{x}^*$ *denotes the fixed point of the replicator equation,* $\tilde{x}^* = (\tilde{c}-\tilde{a})/(\tilde{c}-\tilde{a}-\tilde{d}+\tilde{b})$*, and* $\tilde{D}$ *denotes the denominator of* $\tilde{x}^*$*. In particular, for a bistable competition (*$\tilde{a} > \tilde{c}$ *and* $\tilde{d} > \tilde{b}$*), we recover the one-third law for any group size N:*

$$\begin{aligned} \rho_{M\to R} > 1/N \quad &\Leftrightarrow \quad \tilde{x}^* < 1/3 \\ \rho_{M\to R} > \rho_{R\to M} \quad &\Leftrightarrow \quad \tilde{x}^* < 1/2 \end{aligned} \tag{3.16}$$

(ii) Conversely, if the replicator equation for the game with matrix $\tilde{A}_N$ *has a fixed point* $\tilde{x}^* \in (0,1)$*, then it can be calculated using the weak selection fixation probabilities of the Moran process for game A:*

$$\tilde{x}^* = \frac{1}{3}\left(1 + \frac{\rho_{M\to R} - 1/N}{\rho_{M\to R} + \rho_{R\to M} - 2/N}\right) \tag{3.17}$$

In this case, the dynamics is bistable if $\rho_{M\to R} + \rho_{R\to M} < 2/N$ *and it leads to coex-*

istence if $\rho_{M\to R}+\rho_{R\to M} > 2/N$*. If there is no interior fixed point, the strategy with larger fixation probability dominates the other.*

Proof. (i) The formula for $\rho_{M\to R}$ is verified by a straight calculation: Inserting

$$\tilde{a} = \frac{N-2}{N}a, \quad \tilde{b} = \frac{N-1}{N}b - \frac{1}{N}c, \quad \tilde{c} = \frac{N-1}{N}c - \frac{1}{N}b \text{ and } \tilde{d} = \frac{N-2}{N}d$$

into the right hand's side leads to the same result as Eq. (3.14). The expression for $\rho_{R\to M}$ follows from a symmetry argument, as the equilibrium frequency of R is $1-\tilde{x}^*$ instead of $\tilde{x}^*$.

(ii) Adding up the two equations in (3.15) leads to the relation

$$\frac{\tilde{D}}{2}s = 3\left(\rho_{M\to R} + \rho_{R\to M} - 2/N\right). \tag{3.18}$$

Inserting this expression into the formula for $\rho_{M\to R}$ in (3.15) gives Eq. (3.17). Additionally, it follows that $\tilde{D} = \tilde{c}-\tilde{a}-\tilde{d}+\tilde{b}$ has the same sign as $\rho_{M\to R}+\rho_{R\to M}-2/N$. Since the fixed point $\tilde{x}^* \in (0,1)$ exists if and only if $(\tilde{a}-\tilde{c})(\tilde{d}-\tilde{b}) > 0$, the inequality $\rho_{M\to R}+\rho_{R\to M} > 2/N$ is equivalent to $\tilde{c} > \tilde{a}$ and $\tilde{b} > \tilde{d}$, i.e. to coexistence under replicator dynamics.

□

Simply speaking, and under the assumption of weak selection, this means playing a game A in a finite population of size N is the same as playing the game $\tilde{A}_N$ in an infinite population. As soon as one knows the evolutionary outcome in one regime, one can deduce the results of the other. Thus one may use replicator dynamics to study the behaviour of finite populations (by simply using the transformation $A \mapsto \tilde{A}_N$, see Fig. 3.1). Conversely, the replicator dynamics of game $\tilde{A}_N$ can be examined by studying the finite population dynamics of game A. As a consequence, we will see in examples that the local replicator dynamics, despite having the form of the classical replicator system, mirrors the findings in finite populations. In particular, dominated strategies can prevail.

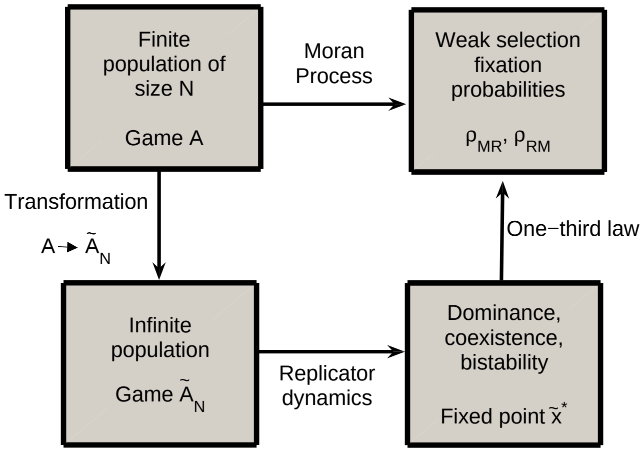


Figure 3.1: Connection between replicator dynamics and the Moran process. To study the game A in a finite population, one may analyze the replicator dynamics of game $\tilde{A}_N$, which we have called the local replicator dynamics of A.

3.4 Examples for the local replicator dynamics

3.4.1 The iterated prisoner's dilemma

In this section we apply our previous results to the iterated prisoner's dilemma (Axelrod, 1984). Suppose that two players may independently decide whether to transfer a benefit b at own costs c to the co-player or not (and $b > c$). Assume that this game is repeated with a constant probability ω. We distinguish three strategies: **All D** always refuses to help, whereas **All C** always cooperates. The strategy **TFT** (*Tit for Tat*) starts with cooperation and imitates the co-player's previous move in all subsequent rounds. Thus we get the following average payoffs per round (for a detailed derivation see Sigmund, 2010):

A	**All C**	**All D**	**TFT**
All C	$b-c$	$-c$	$b-c$
All D	b	0	$b(1-\omega)$
TFT	$b-c$	$-c(1-\omega)$	$b-c$

Obviously, the selfish strategy **All D** is always evolutionarily stable. Because $0 < b-c$, all players in a population of defectors would be better off if everybody cooperated. **All C** is unstable; but if the game lasts sufficiently long ($\omega > c/b$), then **TFT** can resist the

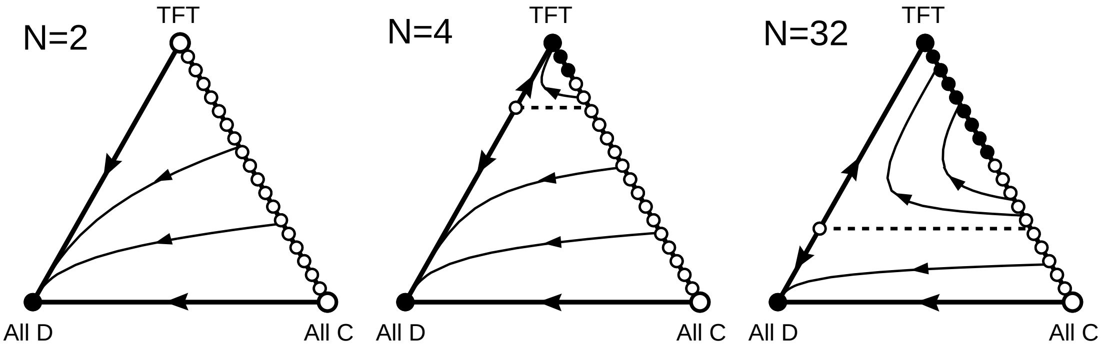


Figure 3.2: The iterated prisoner's dilemma for different group sizes. If $N = 2$, then any player will adopt the **All D** strategy in the long run: As both other strategies perform worse than **All D** in the direct competition, they will usually adopt the co-player's strategy if this co-player is a defector. Increasing group size N diminishes this dilemma. As a consequence, cooperation in the iterated prisoner's dilemma is most likely to evolve in large populations (Parameter values: $b = 3$, $c = 1$ and $\omega = 2/3$).

invasion of **All D**. However, to study the effect of playing in local groups of size N, we have to consider the modified matrix $\tilde{A}_N$ which is up to a factor $1/N$:

$\tilde{A}_N$	**All C**	**All D**	**TFT**
All C	$(N-2)(b-c)$	$-(N-1)c-b$	$(N-2)(b-c)$
All D	$(N-2)b$	0	$(1-\omega)\big[(N-1)b+c\big]$
TFT	$(N-2)(b-c)$	$-(1-\omega)\big[(N-1)c+b\big]$	$(N-2)(b-c)$

According to the local replicator equation, **TFT** is stable against invasion of defectors if and only if

$$(1-\omega)\big[(N-1)b+c\big] < (N-2)(b-c). \tag{3.19}$$

Therefore, even if the game is frequently repeated (ω close to 1), **TFT** might be unstable if its interaction neighborhood N is small (note that TFT is never stable for $N = 2$, see Fig. 3.2).

3.4.2 Risk aversion in a coordination game

Experiments indicate that people tend to risk-averse decisions (Kahnemann and Tversky, 1979; Holt and Laury, 2002; Platt and Huettel, 2008): A bird in the hand is worth two in the bush. As we will see, local replicator dynamics promotes risk-aversion in coordination games. Consider the following stag hunt game (with $r > 0$):

$$A = \begin{pmatrix} 1+r & 0 \\ r & 1 \end{pmatrix}$$

Obviously this game has two pure Nash-equilibria. Owing to the higher variance of the payoffs of the first strategy, we will interpret the first strategy as risk-prone and the second as risk-averse. As the sum of each row is $1 + r$, rational players are exactly indifferent if they know that their co-player chooses each strategy with probability 1/2. Therefore classical replicator dynamics predicts equal basins of attraction. However, for the local replicator dynamics we compute

$$\tilde{A}_N = \frac{1}{N} \begin{pmatrix} (N-2)(1+r) & -r \\ (N-1)r & N-2 \end{pmatrix}.$$

According to this matrix, the risk-averse strategy dominates the risk-prone strategy if $r > N - 2$ and it has the bigger basin of attraction for $r \leq N - 2$. Hence the local replicator dynamics has a clear bias towards the emergence of risk-averse strategies. The intution behind this result is as follows: Individuals can only imitate a different strategy, if both types of players are present in the group. Therefore a successful strategy has to perform well in mixed groups, i.e., it has to yield a notable payoff against any co-player. Overall, this effect may promote the emergence of risk-aversion, although a homogeneous population of risk-prone subjects would yield the highest possible welfare.

In general, coordination to the risk-prone equilibrium is only partly observed in experiments (see, for example Straub, 1995; Friedman, 1996; Battalio et al., 2001) and human behaviour in coordination games seems to be very sensitive on details of the experimental setup. For example, Cooper et al. (1990) have shown that even the presence of a third, dominated strategy can alter the decisions of the subjects considerably.

3.4.3 The dominance paradox

Consider a cooperation game with payoffs

$$A = \left(\begin{array}{c|cc} & C & D \\ \hline C & 2+r & 1 \\ D & 1+r & 0 \end{array} \right)$$

Obviously, there is no social dilemma as the cooperative strategy dominates defective motives. Replicator dynamics thus predicts a stable population with the highest possible payoff, $2+r$. However, this result may be overturned for the local replicator equation:

$$\tilde{A}_N = \frac{1}{N}\begin{pmatrix} (N-2)(2+r) & N-2-r \\ N-2+(N-1)r & 0 \end{pmatrix}$$

According to this matrix, the defective strategy dominates cooperation for $r > N-2$. Although this result seems highly counterintuitive, it is a direct consequence of the assumed locality of interaction and replication. To see this, consider the case $r=8$, $N=4$, for which the (original) payoff matrix becomes

$$A = \begin{pmatrix} 10 & 1 \\ 9 & 0 \end{pmatrix}$$

If a group consists of cooperators only, everybody yields a payoff of 10; however, there is no defector in this group who could adopt this highly profitable strategy. In a group consisting of one defector and three cooperators, the average defector's payoff sums up to $\pi_D = 9$, compared to a cooperator's payoff of $\pi_C = (10+10+1)/3 = 7$. As players imitate strategies with higher payoffs, it is likely that a cooperator of this group dies and is substituted by a defector. Similarly, groups with 2 resp. 3 defectors lead to payoffs $\pi_D = 6$ and $\pi_C = 4$, resp. to $\pi_D = 3$ and $\pi_C = 1$. Having this in mind, the paradoxical emergence of defection becomes unsurprising.

It is important to note that this paradox is not restricted to the local replicator dynamics. Instead it is a general phenomenon for payoff-dependent imitation processes in finite populations. To see this, we calculate the weak selection fixation probabilities under the Moran process according to Eq. (3.14) and find (for $N=4$, $r=8$ and selection strength $s=0.1$):

$$\begin{aligned} \rho_{D\to C} &= 0.357 > 1/4 \\ \rho_{C\to D} &= 0.192 < 1/4 \end{aligned} \tag{3.20}$$

Hence, imitation processes in finite populations might induce social dilemmas where the payoff matrix itself does not reveal any problematic issue at first sight. Note that this example holds for any population size N as long as $r > N-2$.

3.5 Variable group sizes

So far we have assumed that all groups are of the same size. Now we focus on variable group sizes instead and consider N as a random variable. Hence, in any moment of time there may be groups of different size. The population is *well-mixed* in the sense, that any player has the same probability to become a member of a given group, i.e. the average frequency of strategy i in a group does neither depend on the group size N nor on the strategy of the co-players. Within any group, strategy updating occurs as described in the previous sections. In particular, we assume for the moment that in every group there is exactly one player who updates its strategy. The group size N is distributed according to the probability mass function $f(N)$. We exclude atomic groups of one player only, i.e. $f(1) = 0$. Let ν be the harmonic mean of the group sizes N, i.e. $\nu := \big(\sum_{N=1}^{\infty} f(N)/N\big)^{-1}$. Then we get the following generalization for random group sizes:

Proposition 5 (Local replicator dynamics for variable group sizes)
Consider a well-mixed population of players, engaged in pairwise interactions within groups of random size N. If $\nu < \infty$, then the local replicator equation is given by

$$\dot{x}_i = x_i \cdot \big[(\tilde{A}_\nu x)_i - x \cdot \tilde{A}_\nu x\big]. \tag{3.21}$$

Proof. Due to Proposition 1 and our assumptions, the local replicator equation of the system is given by

$$\dot{x}_i = \sum_{N=2}^{\infty} x_i \left[(\tilde{A}_N x)_i - x \cdot \tilde{A}_N x\right] \cdot f(N), \tag{3.22}$$

which can be rewritten as

$$\begin{aligned}
\dot{x}_i &= \textstyle\sum_{N=2}^{\infty} x_i \left[\left(\left(A - \frac{A+A^T}{N}\right) x\right)_i - x \cdot \left(A - \frac{A+A^T}{N}\right) x\right] \cdot f(N) \\
&= x_i \left[(Ax)_i - x \cdot Ax\right] + \textstyle\sum_{N=2}^{\infty} x_i \left[\left(-\frac{A+A^T}{N} x\right)_i + x \cdot \frac{A+A^T}{N} x\right] \cdot f(N) \\
&= x_i \left[(Ax)_i - x \cdot Ax\right] - x_i \left[\left(\left(A + A^T\right) x\right)_i - x \cdot \left(A + A^T\right) x\right] \textstyle\sum_{N=2}^{\infty} f(N)/N \\
&= x_i \left[(Ax)_i - x \cdot Ax\right] - x_i \left[\left(\frac{A+A^T}{\nu} x\right)_i - x \cdot \frac{A+A^T}{\nu} x\right] \\
&= x_i \cdot \big[(\tilde{A}_\nu x)_i - x \cdot \tilde{A}_\nu x\big].
\end{aligned}$$

□

To obtain this result, we have assumed that exactly one player per group is allowed to imitate others. As a consequence, a given player in a large group is less likely to update its strategy than a player in a small group. Alternatively we could also consider a scenario where each player is equally likely to compare its payoff with others, implying that larger groups have on average more updating events. A straightforward adaption of the proof then shows that the harmonic mean needs to be replaced by the arithmetic mean, i.e. in this case we have $\nu = \sum_{N=1}^{\infty} Nf(N)$.

Note that while N had to be an integer, ν may be any positive real number with $\nu \geq 2$. As a consequence, we may use derivatives to study the effect of group size on the dynamics of the system. For example, by calculating $\partial \tilde{x}^*/\partial \nu$, we may quantify the influence of ν on equilibrium frequencies.

3.6 Discussion

Instead of the traditional assumption of a well-mixed and infinite population, we have studied a process in which individuals interact and update their strategies in randomly formed finite groups. On the one hand, such a setting can be interpreted as the mean-field approximation of a game on a lattice, where each player is connected with its $N-1$ neighbors only (as for example in Nakamaru and Iwasa, 2006; Rand et al., 2010). On the other hand, similar procedures are used to test human behaviour in laboratory experiments, e.g. the stranger protocol in Fehr and Gächter (2000) or the control treatment in Traulsen et al. (2010). We have shown that this process can be modelled through a replicator dynamics with a slightly modified payoff matrix.

Recent models for the evolution of cooperation in so-called *public good games* rather use a different setup. Typically, individuals are assumed to interact in randomly formed groups, but strategy updating occurs on a global level (e.g in Brandt et al., 2006; Boyd et al., 2010), leading to classical replicator dynamics. For models of cultural learning, such a setting can be problematic: While individuals may be able to assess the performance of their everyday interaction partners, it remains unclear why they compare themselves with an abstract population average or with random role models they have never encountered before. From this perspective, local updating seems to be more intuitive. Related issues for games on graphs are discussed in Ohtsuki et al. (2007b,c).

Irrespective of its derivation, we have shown that local replicator dynamics connects the Moran process with the classical replicator equation. In the frequently considered case of weak selection, any 2x2 game A that is played in a finite population of N players can be treated *as if* it was game $\tilde{A}_N$, played in an infinite population. In this way, local replicator dynamics provides a comfortable tool to study games in finite populations, especially if the game under consideration allows more than two strategies. Although there has been much effort to extend the Moran process to general nxn games (e.g. Fudenberg and Imhof, 2006; Antal et al., 2009; Tarnita et al., 2009), these generalizations typically require the specification of mutation rates between strategies. Local replicator dynamics, on the other hand, is immediately defined for any number of strategies and does not involve mutation. It is important to note, however, that we have focused on games that are played pairwise. Instead one could also consider games where more than two players are involved. Such systems show a fascinating behaviour; at the same time they are highly complex. A first important step to explore such multi-player games is made by Gokhale and Traulsen (2010).

The modified matrix that we have studied, $\tilde{A}_N = A - (A + A^T)/N$, was also obtained by Lessard (2005) and Lessard and Ladret (2007), who generalized the finite population results of Nowak et al. (2004). Instead of focusing on interactions between two pure strategies, R and M, they allowed mixed strategies p_R and p_M (i.e., strategies that randomize between several pure strategies). For the Moran process they found that selection opposes p_M replacing p_R for every $p_M \neq p_R$ close enough to p_R if and only if p_R is evolutionarily stable for the game with payoff matrix $\tilde{A}_N$ (with respect to the replicator dynamics). This can be regarded as another clue for the strong relationship between the game A in a finite population and the game $\tilde{A}_N$ in an infinite population.

Interestingly, also games that are played on a graph can be represented by a modified replicator equation (Ohtsuki and Nowak, 2006; Ohtsuki et al., 2007c,b; Ohtsuki and Nowak, 2008; Nowak et al., 2010). There it is typically assumed that players only interact with their direct neighbors. According to birth-death updating, individuals with a high payoff are likely to reproduce. The offspring replaces a randomly chosen neighbor. If each player is surrounded by N neighbors, then Ohtsuki and Nowak (2006) find that the dynamics follows a replicator equation with a modified payoff matrix $A + B$, where B encapsulates the impact of the population structure. The entries of B are given by the formula

$$b_{ij} := \frac{a_{ii} + a_{ij} - a_{ji} - a_{jj}}{N - 2}. \tag{3.23}$$

To compare this result with our findings, we multiply the modified matrix $\tilde{A}_N$ with the constant factor $N/(N-2)$ (which only affects the timescale of the dynamics). With this trick we can write the modified matrix for the local replicator dynamics as the sum of the payoff matrix A and a structure matrix B with

$$b_{ij} := \frac{a_{ij} - a_{ji}}{N-2}. \tag{3.24}$$

Thus we get the same structure matrix B, up to the diagonal terms a_{ii} and a_{jj}. As Ohtsuki and Nowak (2006) point out, these terms characterize the effect of assortativeness. In our model we have assumed a well-mixed population and groups that are formed anew after each updating event; hence the diagonal terms are missing in the matrix transformation for the local replicator dynamics. It remains an interesting endeavour, however, to find a general rule under which conditions games in structured population can be described with a modified replicator equation.

Since the correction term in the definition of $\tilde{A}_N$ is of order $1/N$, local replicator dynamics only results in a minor correction if players interact with a large sample of the population. However, for small groups there can be considerable differences, compared to the classical replicator dynamics. As we have seen, imitation processes in finite populations may lead to the evolution of dominated strategies. In particular, for the payoff matrix

$$A = \left(\begin{array}{c|cc} & C & D \\ \hline C & 10 & 1 \\ D & 9 & 0 \end{array}\right),$$

a homogeneous population of defectors is evolutionarily stable if the population size N is small. This result may be attributed to spite: Playing D instead of C implies a small cost for oneself, but harms a cooperative co-player considerably. Overall, this may transform a win-win situation into a lose-lose situation. However, if this game is played in the laboratory, I would rather predict a cooperative equilibrium, as not only the whole population benefits from switching to cooperation but also each player individually. This raises the question whether the local replicator equation (and equivalently, the Moran process) are good approximations on human strategy updating (for an interesting experiment on that issue see Traulsen et al., 2010). The dominance paradox, as described above, relies on the assumption that individuals blindly copy the strategy of co-players with a higher payoff, irrespective of the consequences this might have on the own future payoff. Personally, I believe that individuals have more foresight when they opt for a new strategy. Human strategy updating involves imitation *and* calculation and if evolutionary models should also account for cultural learning, then the game theoretical toolbox has to evolve further.

Acknowledgements

I would like to thank Karl Sigmund, Arne Traulsen and Hisashi Ohtsuki for valuable comments on earlier drafts of this article, as well as two anonymous referees for their various suggestions for improvements. Any remaining errors are my own. This work was funded by EUROCORES TECT I-104-G15.

A An alternative derivation of the local replicator equation

Local replicator dynamics can also be derived from a birth-death process, instead of imitation updating. Again, we assume that players interact within groups of size N. After those interactions, one player of each group is chosen to reproduce, proportional to its fitness. The offspring replaces a randomly chosen member of the same group.

If $d_i(\vec{n})$ denotes the probability that a player with strategy i is chosen to die and $b_j(\vec{n})$ is the corresponding probability that an individual with stratey j is born, then the expected change of players with strategy i within this group $\vec{n}$ is is given by

$$\Delta n_i = \big(1 - d_i(\vec{n})\big) \cdot b_i(\vec{n}) - d_i(\vec{n}) \cdot \big(1 - b_i(\vec{n})\big) = b_i(\vec{n}) - d_i(\vec{n}). \tag{3.25}$$

The birth probability of an individual depends on its fitness $f_i(\vec{n})$, which is a sum of baseline fitness (normalized to 1) and the payoff of this strategy,

$$f_i(\vec{n}) = 1 + s\pi_i(\vec{n}). \tag{3.26}$$

Hence, if $\bar{f}(\vec{n})$ denotes the average fitness in the group, i.e. $\bar{f}_i(\vec{n}) = \big(n_1 f_1(\vec{n}) + \ldots + n_k f_k(\vec{n})\big)/N$, the birth probability is given by

$$b_i(\vec{n}) = \frac{n_i f_i(\vec{n})}{N\bar{f}(\vec{n})}. \tag{3.27}$$

We note that $\sum_{i=1}^{k} b_i(\vec{n}) = 1$. Players die randomly, irrespective of their payoff, and therefore we may write the death probability as $d_i(\vec{n}) = \frac{n_i}{N}$. Thus, we may conclude that

$$\Delta n_i = b_i(\vec{n}) - d_i(\vec{n}) = \frac{n_i}{N} \frac{f_i(\vec{n}) - \bar{f}(\vec{n})}{\bar{f}(\vec{n})} = \frac{n_i}{N}\big[\pi_i(\vec{n}) - \bar{\pi}(\vec{n})\big] \cdot s + \mathcal{O}(s^2) \tag{3.28}$$

Again, we have to sum up these expected changes within one group over all possible groups of size $|\vec{n}| = N$. In the limit of weak selection, this reproduces the local replicator equation in the form of Eq. (3.8):

$$\dot{x}_i = \sum_{|\vec{n}|=N} \frac{N!}{n_1! \, \ldots \, n_k!} \, x_1^{n_1} \cdot \ldots \cdot x_n^{n_k} \cdot \frac{n_i}{N}\big[\pi_i(\vec{n}) - \bar{\pi}(\vec{n})\big]. \tag{3.29}$$

B Proof of Proposition 1

Proposition 1 (A simple representation for the local replicator dynamics)
The local replicator dynamics (3.29) can be rewritten as

$$\dot{x}_i = x_i \cdot \big[(\tilde{A}x)_i - x \cdot \tilde{A}x\big],$$

with $\tilde{A}_N = A - \frac{A+A^T}{N}$.

Proof. First we note that the local replicator equation (3.29) can be written as

$$\begin{aligned}
\dot{x}_i &= \sum_{\substack{|\vec{n}|\,=\,N \\ n_i\,\geq\,1}} \tfrac{N!}{n_1!\ \dots\ n_k!}\ x_1^{n_1} \dots x_n^{n_k} \cdot \tfrac{n_i}{N}\big[\pi_i(\vec{n}) - \bar{\pi}(\vec{n})\big] \\
&= x_i \cdot \sum_{\substack{|\vec{n}|\,=\,N \\ n_i\,\geq\,1}} \tfrac{(N-1)!}{n_1!\dots(n_i-1)!\dots n_k!}\ x_1^{n_1} \dots x_i^{n_i-1} \dots x_n^{n_k}\big[\pi_i(\vec{n}) - \bar{\pi}(\vec{n})\big] \qquad (3.30) \\
&= x_i \cdot \sum_{|\vec{n}|=N-1} \tfrac{(N-1)!}{n_1!\dots n_i!\dots n_k!}\ x_1^{n_1} \dots x_i^{n_i} \dots x_n^{n_k}\big[\pi_i(\vec{n}_{+i}) - \bar{\pi}(\vec{n}_{+i})\big]
\end{aligned}$$

In the last line, the symbol $\vec{n}_{+i}$ denotes a group with n_i+1 players with strategy i and n_j players with strategy $j \neq i$. Accordingly, the last sum is taken over all possible co-players of an i-player (i.e., over all possible groups with $|\vec{n}| = N-1$ players), while the first sum was taken over all possible groups with N individuals. If individuals play against everybody else in their group, but not against themeselves, the corresponding average payoffs are given by:

$$\pi_i(\vec{n}_{+i}) = \tfrac{1}{N-1}\Big(a_{i1}n_1 + ... + a_{ii}n_i + ... + a_{ik}n_k\Big) = \textstyle\sum_{l=1}^k a_{il} \cdot \tfrac{n_l}{N-1}$$

$$\begin{aligned}
\pi_j(\vec{n}_{+i}) &= \tfrac{1}{N-1}\Big(a_{j1}n_1 + ... + a_{ji}(n_i+1) + ... + a_{jj}(n_j-1)... + a_{ik}n_k\Big) \\
&= \tfrac{a_{ji}-a_{jj}}{N-1} + \textstyle\sum_{l=1}^k a_{jl} \cdot \tfrac{n_l}{N-1} \qquad \text{for } j \neq i,\ n_j > 0
\end{aligned}$$

$$\begin{aligned}
\bar{\pi}(\vec{n}_{+i}) &= \tfrac{1}{N}\Big(n_1\pi_1(\vec{n}_{+i}) + ... + (n_i+1)\pi_i(\vec{n}_{+i}) + ... + n_k\pi_k(\vec{n}_{+i})\Big) \\
&= \tfrac{\pi_i(\vec{n}_{+i})}{N} + \textstyle\sum_{j=1}^k \pi_j(\vec{n}_{+i})\tfrac{n_j}{N} \\
&= \textstyle\sum_{l=1}^k \tfrac{a_{il}n_l}{N(N-1)} + \sum_{l,j=1}^k \tfrac{a_{jl}n_l n_j}{N(N-1)} + \sum_{j=1}^k \tfrac{a_{ji}n_j}{N(N-1)} - \sum_{j=1}^k \tfrac{a_{jj}n_j}{N(N-1)}
\end{aligned}$$

Owing to the properties of the multinomial distribuition, we have

$$\begin{aligned}
&\sum_{|\vec{n}|=N-1} \frac{(N-1)!}{n_1!...n_k!}\, x_1^{n_1} \ldots x_n^{n_k} \cdot n_j = (N-1)x_j \\
&\sum_{|\vec{n}|=N-1} \frac{(N-1)!}{n_1!...n_k!}\, x_1^{n_1} \ldots x_n^{n_k} \cdot n_j^2 = (N-1)(N-2)x_j^2 + (N-1)x_j \\
&\sum_{|\vec{n}|=N-1} \frac{(N-1)!}{n_1!...n_k!}\, x_1^{n_1} \ldots x_n^{n_k} \cdot n_j n_l = (N-1)(N-2)x_j x_l \quad \text{for } j \neq l.
\end{aligned} \tag{3.31}$$

Thus we obtain, due to the linearity of the payoffs:

$$\begin{aligned}
\sum_{|\vec{n}|=N-1} \tfrac{(N-1)!}{n_1!...n_k!}\, x_1^{n_1} \ldots x_n^{n_k} \cdot \pi_i(\vec{n}_{+i}) &= \sum_{|\vec{n}|=N-1} \tfrac{(N-1)!}{n_1!...n_k!}\, x_1^{n_1} \ldots x_n^{n_k} \cdot \Big(\textstyle\sum_{j=1}^k a_{ij} \cdot \tfrac{n_j}{N-1} \Big) \\
&= \textstyle\sum_{j=1}^k \tfrac{a_{ij}}{N-1} \Big(\displaystyle\sum_{|\vec{n}|=N-1} \tfrac{(N-1)!}{n_1!...n_k!}\, x_1^{n_1} \ldots x_n^{n_k} \cdot n_j \Big) \\
&= \textstyle\sum_{l=1}^k a_{ij} x_j = \big(Ax\big)_i
\end{aligned} \tag{3.32}$$

Analogously we may derive

$$\begin{aligned}
\sum_{|\vec{n}|=N-1} \tfrac{(N-1)!}{n_1!...n_k!}\, x_1^{n_1} \ldots x_n^{n_k} \cdot \bar{\pi}(\vec{n}_{+i}) &= \tfrac{N-2}{N} \textstyle\sum_{l,j=1}^k x_l a_{lj} x_j + \tfrac{1}{N} \sum_{j=1}^k a_{ij} x_j + \tfrac{1}{N} \sum_{j=1}^k a_{ji} x_j \\
&= \tfrac{N-2}{N} \cdot x^T A x + \tfrac{1}{N} \big(Ax\big)_i + \tfrac{1}{N} \big(A^T x\big)_i
\end{aligned} \tag{3.33}$$

Eventually we obtain

$$\dot{x}_i \quad = x_i \cdot \Big[\big(Ax\big)_i - \tfrac{N-2}{N} \cdot x^T A x - \tfrac{1}{N} \big(Ax\big)_i - \tfrac{1}{N} \big(A^T x\big)_i \Big] = x_i \cdot \big[(\tilde{A}x)_i - x^T \tilde{A} x \big]. \tag{3.34}$$

□

Chapter 4

Evolution of cooperation and punishment in non-anonymous societies

Abstract

Empirical evidence suggests that the mere opportunity to punish others can act as catalyst for cooperative behaviour. It has been argued that humans are equipped with a taste for punishment, but what is the evolutionary cause for such a taste? Moreover, recent experiments reveal a problematic side of retribution. A non-negligible fraction of subjects abuse sanctioning opportunities to engage in spiteful acts (which can harm cooperators, or everyone) or in revenge (as a response to being punished). Most evolutionary models neglect these issues, but recently it has been shown that the co-evolution of punishment and cooperation is dampened, or fully suppressed, if spiteful actions are available. Here we present a simple game-theoretic approach to show that these negative results hinge on the unrealistic yet common assumption of anonymous interactions. In our model, individuals adapt to the reputation of their co-players and learn to behave opportunistically, by cooperating against players who punish defectors but saving the cooperation costs otherwise. This in turn provides an incentive to engage in costly punishment without being spiteful. To unfold this positive effect, sanctions must neither be too soft nor too severe. Between these two extremes, punishment is targeted at noncooperators only, and leads to stable cooperation. These findings are in line with theories that emphasize the strategic role of emotions: In non-anonymous societies, anger or vengefulness may partly have evolved to serve as a credible signal to bystanders.

4.1 Introduction

Behavioural experiments provide overwhelming evidence that many people are willing to engage in punishment (Fehr and Gächter, 2002; de Quervain et al., 2004; Rockenbach and Milinski, 2006; Henrich et al., 2006; Egas and Riedl, 2008; Ostrom et al., 1992). Subjects punish non-cooperators even if the punishment is costly and even if any direct or indirect personal benefit for the punisher is excluded. Although such experiments are typically run in completely anonymous environments, reputation effects may play a subtle role. The mere picture of an eye (Haley and Fessler, 2005) or the physical presence of the experimenter (Kurzban et al., 2007) can affect the participants' behaviour, often making them more cooperative or increasing their willingness to punish non-cooperators. It has been shown that players in the ultimatum game reject unfair offers even when this decision increases rather than decreases inequity (Yamagishi et al., 2009). While this finding is counterintuitive under the anonymous conditions of the laboratory, the subjects' behaviour may be perfectly adapted to the conditions of everyday life, where reputation plays a key role. Emotions such as anger and vengefulness seem to act as a commitment device; they help to credibly demonstrate that one is not willing to accept unfair treatments (Frank, 1988).

However, previous research on the co-evolution of cooperation and punishment gives little weight to reputation. Instead, most evolutionary models tend to mimic the impersonal conditions in the laboratory (Fowler, 2005; Nakamaru and Iwasa, 2006; Lehmann et al., 2007b; Eldakar and Wilson, 2008; Sigmund et al., 2010; Janssen and Bushman, 2008; Rand et al., 2010). In these models, punishment acts by reducing the non-cooperators' payoffs, thereby preventing the evolutionary spread of defection. We call this the payoff mechanism of punishment. In a recent study, Rand et al. (2010) argue that the payoff mechanism alone cannot maintain cooperation. Indeed, the same conditions that favour punishment may also promote the evolution of spite, once this possibility is considered. In our model, punishment does not only affect payoffs, but also the reputation of a player. Being known as a person who is willing to pay a cost to take revenge on non-cooperators may be of advantage in future interactions. As we will show, it is this reputation mechanism that socializes punishment, keeping spite and antisocial punishment at bay.

4.2 The model

We consider the following pairwise game with two stages (see also Sigmund et al., 2001). Before the game starts, a coin toss determines which player is in the role of the donor and which one is in the role of the recipient. In the subsequent helping stage, donors may transfer a benefit b to their recipients, at their own cost c, or they may refuse to do so. These two alternatives are denoted by C (for cooperation) and D (for defection), respectively. We make the standard assumption that cooperation is the efficient outcome, $b > c$. Then, in the punishment stage, recipients decide whether or not to punish the donor. Depending on the outcome of the helping stage, there are four possible reactions: Punishing defectors only (denoted by P for social punishment), punishing cooperators only (A for antisocial punishment), punishing everybody (S for spiteful punishment) or nobody (N). If punished, donors have to pay a fine β whereas the recipient is charged an amount γ. Because sanctions are costly, immediate self-interest speaks against either form of punishment.

In order to incorporate reputation, we assume that donors know their co-player's strategy with an exogenous probability μ. Whether this information comes from previous encounters, from observation or gossip can be left open. We can therefore distinguish four different types of donors: $AllC$-players always cooperate, whereas $AllD$-players never do so, regardless of the opponent's reputation. The third type of donor is the opportunistic cooperator, O_C. Donors with this strategy generally cooperate; but if they find out that their co-player is not a social punisher P, they decide to save the costs of helping. Finally, players with strategy O_D are non-cooperative except they know they are interacting with a P-recipient.

We use evolutionary dynamics to study the above game. There exists a wealth of possible dynamics and each specific choice may affect the result in subtle ways (Nowak et al., 2010). The traditional approach is to consider a virtually infinite population of players. Each individual acts according to its strategy $[i, j]$, where $i \in \{AllC, O_C, O_D, AllD\}$ prescribes how to play in the role of a donor and $j \in \{P, N, A, S\}$ prescribes how to react as a recipient. Classical replicator dynamics(see ref. Sigmund, 2010) assumes that the fraction of $[i, j]$-players, x_{ij}, evolves according to the following differential equation:

$$\dot{x}_{ij} = x_{ij} \cdot [\pi_{ij} - \pi^e]. \tag{4.1}$$

In this equation, π_{ij} denotes the average payoff of an $[i, j]$-player, whereas π^e denotes the mean payoff over the whole population. Replicator dynamics thus describes a situation where strategies that yield a payoff above the global average will spread among the population. However, global competition is an inappropriate model for the evolution of spite (Gardner and West, 2004b). Instead, such antisocial behaviour can only emerge when players compete locally and harming others gives a relative advantage to the focal player (Rand et al., 2010). In order to allow for spite, we therefore assume that players interact within small and randomly formed groups of size $n+1$. The dynamics of such a scenario can be captured by the following adapted replicator equation (Hilbe, 2011):

$$\dot{x}_{ij} = x_{ij} \cdot [\pi_{ij} - \frac{(n-1)\pi^e + \pi_{ij} + \pi_{ij}^T}{n+1}]. \tag{4.2}$$

According to the *local replicator dynamics* (4.2), a strategy propagates if players adopting this strategy yield a payoff above the group average. Being an $[i,j]$-player has an influence on the mean group payoff for two reasons: On the one hand because there is at least one individual with payoff π_{ij}. On the other hand the $[i,j]$-player affects its co-players' payoffs by an expected amount denoted by π_{ij}^T. By taking these two biases into account we obtain the above dynamics (see also appendix). The most important difference between (4.1) and (4.2) is that the latter allows for spite: Under the local replicator dynamics, the frequency x_{ij} increases either by rising the own payoff π_{ij} or by diminishing the others' payoffs via π_{ij}^T. As group sizes become large, the impact of spite disappears.

4.3 Analysis and results

We first analyze under which conditions social punishment P is able to maintain cooperation. Traditional analysis suggests that punishment acts as a deterrent if fines exceed the costs of cooperation, $\beta > c$. However, players who compete locally also consider the effect on the co-players' payoffs and the respective condition becomes $\beta > \beta_{\min} = c + (b+\gamma)/n$. We can use this condition to eliminate dominated strategies. If $\beta > \beta_{\min}$ it is easy to verify that opportunists, O_C and O_D, always perform better than their unconditional counterparts *AllC* and *AllD*: Once information is available, it is optimal to adapt to the co-player's reputation. On the recipient's side it is safe to reject antisocial-punishment A: Because players react opportunistically, the reputation mechanism disfavours players who punish cooperators. On the other hand, the payoff mechanism might promote antisocial behaviour, if punishment is sufficiently harmful (that is, if $\beta/n > \gamma$). However, in this case it is optimal to harm everybody and not cooperators only (Rand et al., 2010), leading

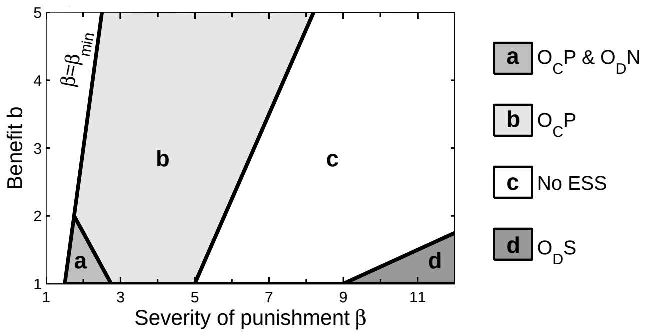
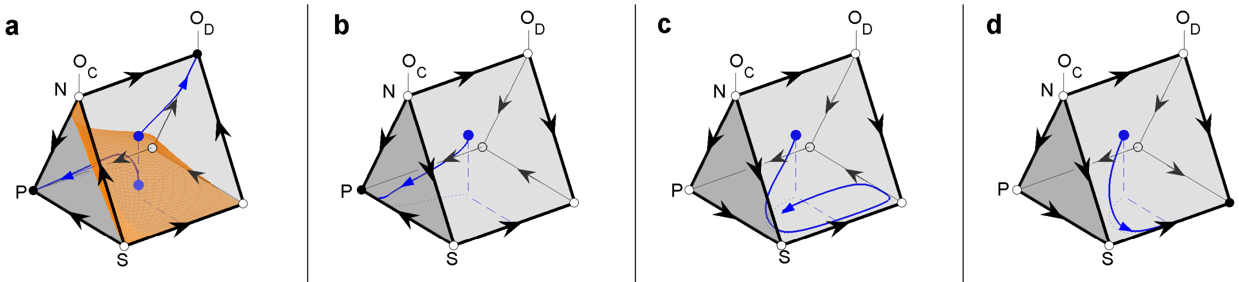


Figure 4.1: Evolutionarily stable states (ESS) for the helping game with punishment. We consider which strategies form a stable equilibrium of the evolutionary dynamics, given that punishment is severe enough to act as a deterrent. Depending on the parameters b and β there are four different scenarios: **a)** Low values of b and β allow two stable states, either the welfare maximizing [O_C,P] or the selfish [O_D,N]. **b)** If punishment becomes more severe or if the benefit of cooperation rises then [O_C,P] is the only ESS. **c)** Increasing the fine even further destroys the stabilizing effect on cooperation and no strategy is stable. **d)** For very large fines, spiteful punishment takes over (compare these results with Fig. 4.2).
Parameters: $c = \gamma = 1$, $\mu = 20\%$ and $n = 4$. For bigger group sizes n, area **a** and **b** grow at the expense of area **c** and **d** since spite becomes too costly in large groups.

to spite instead of antisocial punishment.

Let us therefore consider a reduced system, with all donors being opportunists and recipients being either of type P, N or S. Depending on the intensity of sanctions, there are four possible evolutionary regimes (Fig. 4.1). A moderate level of benefits b and comparably soft penalties β lead to bistability: If most of the individuals are not willing to use sanctions initially, then the population will move to a selfish state without punishment [O_D,N]. However, a population containing a certain baseline number of punishers is able to maintain mutual cooperation, moving the state to [O_C,P] (Fig. 4.2a). Note that all players in a homogeneous [O_C, P] population cooperate, either by default or because they anticipate that they would be punished otherwise. Increasing the fine β makes the inferior equilibrium [O_D,N] disappear and the social [O_C,P] population becomes globally attracting. In this parameter region sanctions work best, leading to the optimal welfare even if both, cooperation and punishment, are initially rare (Fig. 4.2b). A further increase of β destroys the positive effect of punishment, resulting either in instability (Fig. 4.2c) or spite (Fig. 4.2d). Here, sanctions cannot maintain cooperation since it is no longer guaranteed that only selfish individuals are punished. Antisocial behaviour can invade.

To verify this outcome also for the full strategy set and for other evolutionary dy-



a $\beta > \beta_{\min}$, $\beta < n(\gamma - \frac{\mu}{\bar\mu} b) - \frac{\mu}{\bar\mu} c$

b $\beta > n(\gamma - \frac{\mu}{\bar\mu} b) - \frac{\mu}{\bar\mu} c$, $\beta < n(\gamma + \mu b) + \mu c$

c $\beta > n(\gamma + \mu b) + \mu c$, $\beta < n(\gamma + b) + c$

d $\beta > n(\gamma + b) + c$

Figure 4.2: Evolutionary dynamics for the four different parameter regions. For each set of parameters, the strategy space can be represented by a prism. The corners of the prism mark homogeneous populations, whereas points in the interior correspond to mixed populations. Blue orbits indicate the dynamics for a given initial population. **a)** Both equilibria, $[O_C, P]$ and $[O_D, N]$ are stable. The orange grid separates initial populations that converge to the social state $[O_C, P]$ from those that move to the inferior equilibrium $[O_D, N]$. **b)** $[O_C, P]$ is globally stable. **c)** Spiteful punishers can invade a homogeneous $[O_C, P]$ population, making punishment enforced cooperation unstable. The dynamics exhibits cycles between cooperation and defection, respectively between spite and social punishment. **d)** Recipients become fully spiteful, moving the population to the detrimental $[O_D, S]$ equilibrium.
Parameters: $c = \gamma = 1$, $b = 3/2$, $\mu = 20\%$, $\bar\mu = 1 - \mu = 80\%$, $n = 4$ and **a)** $\beta = 2$ **b)** $\beta = 5$ **c)** $\beta = 8$ **d)** $\beta = 12$.

namics, we have run computer simulations that implement the well-known Moran process (Nowak et al., 2004). We obtain similar results: When reputation effects are sufficiently strong, opportunism soon takes over, which in turn promotes the evolution of social punishment (Fig. 4.3). In the end, a homogeneous population of $[O_C,P]$ players evolves. These simulation results are largely independent of the exact modelling of selection (how players switch to strategies with higher payoffs) and mutation (how they experiment with new strategies). For a detailed discussion of the influence of the parameters, we refer to the appendix.

4.4 Discussion

Previous evolutionary models cannot explain why individuals learn to deal responsibly with sanctions. Instead, it is either presumed that punishment is targeted at defectors only (Boyd et al., 2003; Gardner and West, 2004a; Fowler, 2005; Nakamaru and Iwasa,

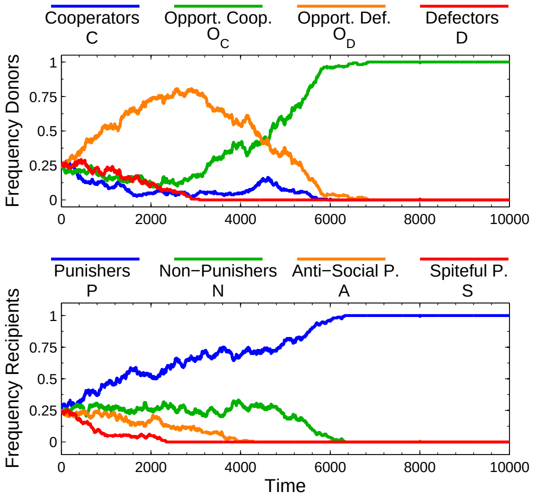


Figure 4.3: Co-evolution of cooperation and punishment over time. A typical individual-based simulation run of the Moran process shows that opportunists soon dominate the population (upper graph), which in turn promotes the steady evolution of social punishment P (lower graph). When the fraction of social punishers exceeds a certain threshold it pays off to cooperate when in doubt, favouring O_C over O_D in the long run.
Parameters: $c = \gamma = 1$, $b = 4$, $\beta = 3$, $\mu = 30\%$ and population size 160. In each time step, one individual was selected to imitate another player of the population and updating occurred according to the so-called Fermi-rule with medium selection strength, $s = 0.5$ (see appendix).

2006; Lehmann et al., 2007b; Eldakar and Wilson, 2008; Sigmund et al., 2010; Boyd et al., 2010), or it is predicted that evolution leads to non-punishing defectors (if $\beta < \gamma n$) or spite (if $\beta > \gamma n$), respectively (Rand et al., 2010). Here we have shown that reputation can resolve this puzzle. Under non-anonymity, there is little appeal in harming those who help. If targeted punishment evolves, then it is systematically directed at non-cooperators. Hence, we also contradict the popular "folk theorem" that any behaviour, even if abstruse, can become a common norm as long as deviations are punished (Boyd and Richerson, 1992). Opportunistic individuals will stop to impose sanctions on pro-social activities, simply because it is in their own interest to let cooperative outcomes evolve.

As we have shown, opportunism emerges automatically if individuals are able to evaluate the reputation of their opponents. Of course this implies some cognitive requirements on the subjects: They have to monitor their co-players and need to process and remember this information properly. For humans, this does not seem to be too restrictive. For example, subjects in laboratory experiments show an enhanced memory for faces of cheaters

(Mealey et al., 1996) and although not tested empirically, one may expect similar results for punishers. Moreover, it is well-known that humans anticipate the importance of reputation and therefore behave differently when being observed (Seemmann et al., 2004). But also in some animal species, individuals are able to distinguish each other's type and to behave accordingly, as reported in studies of indirect reciprocity - for example for cleaner fishes (Bshary and Grutter, 2006) or for great apes (Russel et al., 2008).

The model that we have proposed is admittedly minimalistic, assuming pairwise interactions and neglecting the possibility of errors. Both assumptions are made to allow an intuitive analysis and do not qualitatively affect the results. It seems to be a more serious shortcoming that the model does not incorporate the possibility of retaliation. Especially in non-anonymous societies it may pay to engage in counter-punishment, thereby demonstrating that one does not accept opposition. However, a way to address this issue without increasing the complexity of the model is to ask whether individuals are still willing to punish others, even if counter-punishment is a sure event (that is, even if $\gamma = \beta$). Surprisingly, we find that counter-punishment prevents spite but still allows for social punishment. Intuitively, individuals engage in costly vendettas only if this yields some positive return in form of increased cooperation in future.

Interesting effects are to be expected if we relax the assumption of exogenous information and if we consider reputation as a strategic variable. In many circumstances, subjects may have the choice between explicit and subtle forms of punishment. While explicit penalties serve as a signal to bystanders, more inconspicuous sanctions minimize the risk of retaliation. Both emergence and form of punishment are important issues that can only be understood if their reputational implications are taken into account.

A Calculation of the payoffs

We have considered the following four strategies for recipients:

P	Social punishment (to punish defectors only)
N	No punishment (to punish nobody)
A	Anti-social punishment (to punish cooperators only)
S	Spite (to punish everybody)

Furthermore, we have distinguished between four different strategies for donors:

AllC	Unconditional cooperators
O_C	Opportunistic cooperators (they generally cooperate, except they know they are not interacting with a P-recipient)
O_D	Opportunistic defectors (they generally refuse to cooperate, except they know they are interacting with a P-recipient)
AllD	Unconditional defectors

With this specification of the strategies, it is straightforward to calculate the payoff for each player. For example, consider an O_C-donor who encounters an N-recipient. With probability μ, the donor knows that the recipient does not punish others, in which case it is safe to refuse cooperation. In the other case, if the donor does not know the recipient's strategy, O_C-donors cooperate by default (which happens with probability $\bar{\mu} := 1-\mu$). In total, this yields an average payoff of $-\bar{\mu}c$ for the O_C-donor and $\bar{\mu}b$ for the N-recipient. Repeating this computation for all other strategy pairs yields the following bimatrix $(A, B) = (a_{ij}, b_{ij})$. In this bimatrix, the first entry denotes the payoff of the donor whereas the second entry denotes the corresponding payoff of the recipient:

$*$	**P**	**N**	**A**	**S**
AllC	$\big(-c, b\big)$	$\big(-c, b\big)$	$\big(-c-\beta, b-\gamma\big)$	$\big(-c-\beta, b-\gamma\big)$
$\mathbf{O_C}$	$\big(-c, b\big)$	$\big(-\bar{\mu}c, \bar{\mu}b\big)$	$\big(-\bar{\mu}(c+\beta), \bar{\mu}(b-\gamma)\big)$	$\big(-\bar{\mu}c-\beta, \bar{\mu}b-\gamma\big)$
$\mathbf{O_D}$	$\big(-\mu c-\bar{\mu}\beta, \mu b-\bar{\mu}\gamma\big)$	$\big(0, 0\big)$	$\big(0, 0\big)$	$\big(-\beta, -\gamma\big)$
AllD	$\big(-\beta, -\gamma\big)$	$\big(0, 0\big)$	$\big(0, 0\big)$	$\big(-\beta, -\gamma\big)$

Table 4.1: Payoff matrix for the helping game with punishment

Each individual acts according to its strategy $[i, j]$, where $i \in \{AllC,\ O_C,\ O_D,\ AllD\}$ prescribes how to play in the role of a donor and $j \in \{P,\ N,\ A,\ S\}$ prescribes how to react as a recipient. We assume that individuals have an equal chance to be the donor or the recipient in a given interaction. Consequently, a player with strategy $[i, j]$ who encounters an opponent $[k,l]$ obtains on average

$$\pi_{ij,kl} := (a_{il} + b_{kj})/2. \tag{4.3}$$

B Analysis of the local replicator dynamics

We consider an infinite population and denote the fraction of players with strategy $[i,j]$ by x_{ij}. Of course, all these frequencies add up to one, $\sum_{i,j} x_{ij} = 1$. We assume that players interact in finite groups. Groups are formed randomly and each individual plays against every other group member. In this case, the expected (normalized) payoff of a player with strategy $[i,j]$ can be calculated as:

$$\pi_{ij} := \sum_{k,l} x_{kl} \cdot \pi_{ij,kl}. \tag{4.4}$$

Furthermore, we compute the average payoff of the whole population by summing up over all individual payoffs, yielding:

$$\pi^e := \sum_{i,j} x_{ij} \cdot \pi_{ij} \tag{4.5}$$

Classical replicator dynamics asserts that a strategy $[i,j]$ spreads whenever it leads to a payoff above the global average, that is if π_{ij} exceeds π^e. However, replicator dynamics cannot explain why individuals engage in spiteful acts. Spiteful subjects pay a cost (decreasing their own payoff) to harm others; but since competition is global and because the population is arbitrarily large, a spiteful subject has a negligible impact on the mean payoff of the population. Global competition is therefore an inappropriate model for the evolution of spite. Thus we consider the case where players compete locally, assuming that a strategy spreads if its payoff is above the group average. Being an $[i,j]$-player in a finite group affects the mean payoff of the group for two reasons: On the one hand, there is at least one player that obtains a payoff π_{ij}. On the other hand, all other group members play at least once against an $[i,j]$-player, leading to an expected influence on the co-players' payoffs of

$$\pi_{ij}^T := \sum_{k,l} x_{kl} \cdot \pi_{kl,ij}. \tag{4.6}$$

Taking these two effects into account, the average payoff of a group containing at least one $[i, j]$-player and n randomly chosen other group members is given by

$$\frac{n-1}{n+1}\pi^e + \frac{1}{n+1}\pi_{ij} + \frac{1}{n+1}\pi_{ij}^T \tag{4.7}$$

Overall, the assumption of local competition leads to a modified replicator dynamics. The time evolution of x_{ij} is governed by the following equation:

$$\dot{x}_{ij} = x_{ij}\left[\pi_{ij} - \frac{n-1}{n+1}\pi^e - \frac{1}{n+1}\pi_{ij} - \frac{1}{n+1}\pi_{ij}^T\right]. \tag{4.8}$$

For a detailed discussion of the analytical properties of this differential equation, we refer the reader to Hilbe (2011). There it is also shown that a homogeneous population is evolutionarily stable with respect to the local replicator dynamics, if and only if it is stable with respect to the viability model of Nakamaru and Iwasa (2006). This allows us to compare our results under non-anonymity with the results of Rand et al. (2010).

It is possible to simplify Eq. (4.8). Denote by $A = (a_{ij})$ the matrix that contains the donor's payoffs and by $B = (b_{ij})$ the corresponding matrix for the recipient's payoffs. Next we define the two modified payoff matrices

$$\begin{aligned}\tilde{A} &= A - \tfrac{A+B}{n+1}\\ \tilde{B} &= B - \tfrac{A+B}{n+1}\end{aligned} \tag{4.9}$$

and compute modified payoffs $\tilde{\pi}_{ij}$ resp. $\tilde{\pi}^e$ according to Eqs. (4.3) – (4.5). Then a straightforward calculation (presented at the end of this section) shows that the local replicator dynamics (4.8) is equivalent to

$$\dot{x}_{ij} = x_{ij}[\tilde{\pi}_{ij} - \tilde{\pi}^e]. \tag{4.10}$$

Consequently, the local replicator equation can be written in the form of a classical replicator equation. The system evolves *as if* the payoffs were given by $\tilde{A}$ and $\tilde{B}$.

Elimination of dominated strategies

When we apply the transformation (4.9) to the helping game with punishment, we obtain – up to a factor $1/(n+1)$ – the modified payoff matrices $\tilde{A}$ and $\tilde{B}$ in (4.13a) resp. (4.13b). In order to simplify the further analysis, we first eliminate dominated strategies. By considering the donor's modified payoffs in (4.13a) we see that O_C always gets, at least, the payoff of *AllC*. The dominance relation is not that clear if it comes to the two defective strategies *AllD* and O_D. According to (4.13a), *AllD* is weakly dominated by O_D if

$$-\mu nc - \bar{\mu} n\beta - \mu b + \bar{\mu}\gamma > -n\beta + \gamma. \tag{4.11}$$

This condition is met if punishment is sufficiently severe, i.e. if β exceeds the threshold

$$\beta_{\min} := (\gamma + b)/n + c. \tag{4.12}$$

For large group sizes the condition $\beta > \beta_{\min}$ simplifies to $\beta > c$. On the other hand, if $\beta < \beta_{\min}$, unconditional defection dominates all other strategies. Hence we interpret $\beta_{\min}$ as the minimum fine to act as a deterrent.

Modified payoffs for a donor (row player)

$\tilde{\mathbf{A}}$	P	N	A	S
AllC	$-nc-b$	$-nc-b$	$-nc-n\beta-b+\gamma$	$-nc-n\beta-b+\gamma$
O_C	$-nc-b$	$\bar{\mu}(-nc-b)$	$\bar{\mu}(-nc-n\beta-b+\gamma)$	$-\bar{\mu}nc-n\beta-\bar{\mu}b+\gamma$
O_D	$-\mu nc-\bar{\mu}n\beta-\mu b+\bar{\mu}\gamma$	0	0	$-n\beta+\gamma$
AllD	$-n\beta+\gamma$	0	0	$-n\beta+\gamma$

(4.13a)

Modified payoffs for a recipient (column player)

$\tilde{\mathbf{B}}$	P	N	A	S
AllC	$nb+c$	$nb+c$	$nb-n\gamma+\beta+c$	$nb-n\gamma+\beta+c$
O_C	$nb+c$	$\bar{\mu}(nb+c)$	$\bar{\mu}(nb-n\gamma+\beta+c)$	$\bar{\mu}nb-n\gamma+\bar{\mu}c+\beta$
O_D	$\mu nb-\bar{\mu}n\gamma+\mu c+\bar{\mu}\beta$	0	0	$\beta-n\gamma$
AllD	$\beta-n\gamma$	0	0	$\beta-n\gamma$

(4.13b)

On the recipient's side, it is safe to eliminate the anti-social punishment strategy A. According to (4.13b), A is weakly dominated by N if $n\gamma > \beta$ and weakly dominated by S whenever $n\gamma < \beta$.

Evolutionarily stable strategies

Following the observations in the last section we restrict our attention to a system with all donors being opportunists and all recipients being either of type P, N or S. Our next aim is to calculate all evolutionarily stable states (ESS) of the reduced system. A homogeneous population is called evolutionarily stable if rare mutants have a strictly lower payoff than the residents. Since by (4.3) the payoff of a strategy $[i, j]$ is a linear combination of the payoff as a donor and the payoff as a recipient, strategy $[i, j]$ is an ESS if and only if it is componentwise stable (that is, neither an $[i, l]$-mutant nor a $[k, j]$-mutant can invade).

Let us first consider when $[O_C,P]$ is evolutionarily stable. According to (4.13a), $[O_D,P]$ always has a lower payoff than $[O_C,P]$, provided that $\beta > \beta_{\min}$. According to (4.13b), $[O_C, P]$ can be invaded by $[O_C,S]$ if

$$nb + c < \bar{\mu}nb - n\gamma + \bar{\mu}c + \beta. \tag{4.14}$$

Rearranging the terms leads to the conclusion that $[O_C,P]$ is evolutionarily stable if the punishment fines are in the interval $\beta_{\min} < \beta < n(\gamma+\mu b)+\mu c$. With similar computations for all other strategies we get a list of possible ESS, see Tab. 4.2.

The conditions in Tab. 4.2 are presented as Fig. 4.1 in the main text. Note that the thresholds in the right column of Tab. 4.2 can be ordered as

$$n(\gamma - \mu/\bar{\mu}b) - \mu/\bar{\mu}c \leq n(\gamma + \mu b) + \mu c \leq n(\gamma + b) + c \tag{4.15}$$

We can use these conditions to discuss the influence of the parameters. Let us start with the group size n. An increasing group size increases all thresholds in the right column of Tab. 4.2 and therefore shifts all boundary lines in Fig. 4.1 to the right, except for the line for $\beta = \beta_{\min}$. In effect, this leads to an increase of the areas **a** and **b**, at the expense of areas **c** and **d**. In the limit $n \to \infty$, which corresponds to the result of the classical replicator equation, areas **c** and **d** disappear completely. In this case, either $[O_C, P]$ is the unique ESS (if $\gamma < \mu/\bar{\mu}b$) or there are two ESS, namely $[O_C, P]$ and $[O_D, N]$ (if $\gamma > \mu/\bar{\mu}b$). Spiteful behaviour has vanished, since the costs of a spiteful punisher increase linearly in n, whereas the detrimental effect on the co-players remains γ. If groups are sufficiently large, it becomes too expensive to be spiteful. Similarly we can explore the effect of all other parameters; the results are presented in Tab. 4.3.

Strategy $[i, j]$	Donor's strategy i is stable if	Recipient's strategy j is stable if
Cooperation through punishment $[O_C, P]$	$\beta > \beta_{\text{min}}$	$\beta < n(\gamma + \mu b) + \mu c$
Defection through laissez-faire $[O_D, N]$	$\beta > \beta_{\text{min}}$	$\beta < n(\gamma - \mu/\bar{\mu} b) - \mu/\bar{\mu} c$
Defection through spite $[O_D, S]$	$\beta > \beta_{\text{min}}$	$\beta > n(\gamma + b) + c$

Table 4.2: Conditions for evolutionary stability

Parameter	Influence on the dynamics
Group size n	Bigger group sizes diminish spiteful behaviour. Areas **a** and **b** increase on the expense of **c** and **d**.
Information level μ	The higher the information level, the more it pays off to punish defectors to get opportunistic donors to cooperate. Hence area **b** increases on the expense of **a** and **c**, whereas **d** remains unchanged.
Costs of cooperation c	Increasing costs shift all areas to the right, i.e. it takes higher fines β to guarantee the same outcome.
Costs of punishing γ	Increasing punishing fees reduce the danger of spiteful punishment, shifting **c** and **d** to the right. Additionally, individuals may tend to avoid the costs γ completely, i.e. area **a** increases on the expense of **b**.

Table 4.3: Influence of the parameters. Bold letters refer to the respective parameter regions in Fig. 4.1 of the main text.

Analysis of the global dynamics

In the previous section we have identified the stable states of the dynamics. Once an ESS is reached, the population will stay there. In this section we study the evolution of populations that are not yet in such a stable state. As it turns out, also the global dynamics can be classified by means of the conditions in Tab. 4.2. In the following we assume that punishment is severe enough to act as a deterrent, $\beta > \beta_{\text{min}}$.

If we first consider the case that $\beta < n(\gamma + \mu b) + \mu c$, then it follows from payoff matrix (4.13b) that S is dominated by P. Indeed, $[O_C, P]$ yields a higher payoff than $[O_C, S]$

and also $[O_D, P]$ performs better than $[O_D, S]$. Therefore, replicator dynamics leads to the extinction of spite (see Hofbauer and Sigmund, 1998). In the Figures 4.2a and 4.2b of the main text, this is graphically indicated by the arrows pointing from S to P for both possible roles as a donor. On the remaining plane spanned by $[O_C,P]$, $[O_C,N]$, $[O_D,N]$ and $[O_D,P]$, there is either a bistable competition (Fig. 4.2a) or punishment enforced sociality is globally attracting (Fig. 4.2b).

The other case, in which $\beta > n(\gamma+\mu b)+\mu c$, can be treated similarly. For this parameter region, the recipient's strategy N is dominated by S and all evolutionary trajectories approach the plane that is spanned by $[O_C,P]$, $[O_C,S]$, $[O_D,S]$ and $[O_D,P]$. There, we either observe cyclic behaviour (Fig. 4.2c) or convergence to the detrimental $[O_D,S]$-equilibrium (Fig. 4.2d). The cyclic case is particularly interesting: $[O_C,P]$-societies can be subverted by spiteful individuals with strategy $[O_C,S]$. But without the specific punishment of defection, the population evolves towards a state in which individuals defect by default, using strategy $[O_D,S]$. In this situation, however, it pays off to care for a strict reputation to guarantee the cooperation of the co-players and $[O_D\ ,P]$ evolves. As soon as a sufficient number of individuals switches to social punishment P, it is cheaper to cooperate when in doubt, leading back to $[O_C,P]$. Depending on initial conditions, this results either in infinite heteroclinic cycles or it eventually leads to a stable mixture of these four strategies (Gaunersdorfer et al., 1991).

The case of counter-punishment

Experiments indicate that the option to punish others may lead to long and costly vendettas between players (Nikiforakis, 2008; Denant-Boemont et al., 2007; Herrmann et al., 2008). The option to retaliate punishment is not explicitly included in our model. A possible way to address this issue is to add a third stage to the game in which donors may choose whether or not to punish their recipients. Such an additional counter-punishment-stage allows for a wealth of new strategies and leads to a considerable increase in complexity. However, a basic way to incorporate counter-punishment is to assume that punishment is equally costly for both parties. This assumption models a situation in which counter-punishment is a sure event. If $\gamma = \beta$, the minimum fine that acts as a deterrent (given by Eq. (4.12)) becomes

$$\beta_{\min} = \frac{nc+b}{n-1}. \tag{4.16}$$

Only if fines exceed this threshold, it is optimal to cooperate in a population of social punishers, P. Again, the condition $\beta > \beta_{\min}$ simplifies to the traditional condition $\beta > c$

Strategy $[i, j]$	Donor's strategy i is stable if	Recipient's strategy j is stable if
Cooperation through punishment $[O_C, P]$	$\beta > \beta_{\min}$	Always
Defection through laissez-faire $[O_D, N]$	$\beta > \beta_{\min}$	$\beta > \frac{\mu}{\bar{\mu}(n-1)}(nb + c)$

Table 4.4: Conditions for evolutionary stability if $\gamma = \beta$.

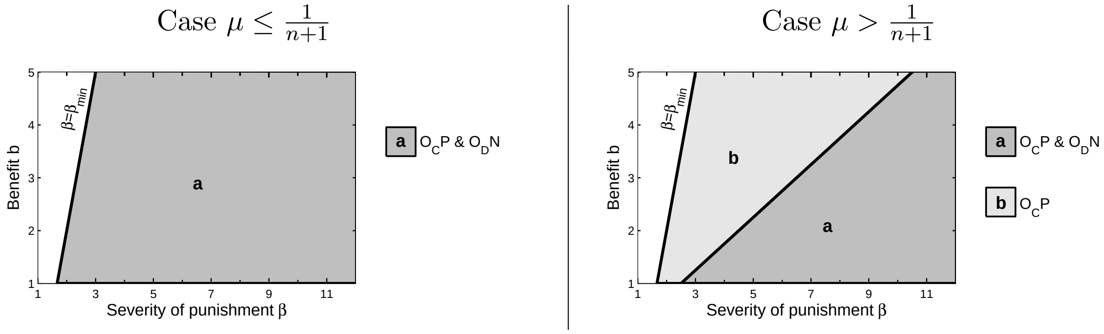


Figure 4.4: Evolutionary stability in the case of $\gamma = \beta$. The left graph depicts the case of low information ($\mu = 15\%$), where both $[O_C, P]$ and $[O_D, N]$ are stable if $\beta > \beta_{\min}$. The right graph describes the case of high information ($\mu = 60\%$); here we can find an area where $[O_C, P]$ is the only ESS. Other parameter values: $c = 1$, $n = 4$.

for large group sizes. The respective conditions for evolutionary stability in the case of $\gamma = \beta$ are given in Tab. 4.4. Surprisingly, while counter-punishment still allows for a stable $[O_C, P]$-population, the spiteful state $[O_D, S]$ becomes unstable. Intuitively, spiteful punishment can only evolve because of the payoff mechanism, that is if punishment is more costly for the target than it is for oneself. If counter-punishment is a sure event, then spite loses its evolutionary advantage.

Fig. 4.4 presents the conditions for evolutionary stability in a $\beta - b-$diagram. We can distinguish between two cases: If $\mu \leq 1/(n+1)$ (i.e. if on average there is no player whose reputation is known), then the area where $[O_D, N]$ is an ESS coincides with the area where $[O_C, P]$ is an ESS. Only if $\mu > \frac{1}{n+1}$ these two areas differ. The more information about the co-player's strategy is available, the more easily a population evolves towards the social $[O_C, P]$-state.

Local versus classical replicator dynamics

It remains to show that the two representations for the local replicator equation, (4.8) and (4.10), are equivalent. More specifically, we need to prove that the local replicator dynamics for the game with payoff matrices A and B coincides with the classical replicator dynamics for the game with modified payoff matrices $\tilde{A} = A - (A+B)/(n+1)$ and $\tilde{B} = B - (A+B)/(n+1)$.

First, we calculate the modified payoff of a player with strategy $[i,j]$ against a co-player using strategy $[k,l]$:

$$\begin{aligned} \tilde{\pi}_{ij,kl} &= \tfrac{\tilde{a}_{il}+\tilde{b}_{kj}}{2} \\ &= \tfrac{\left(a_{il}-\frac{a_{il}+b_{il}}{n+1}\right)+\left(b_{kj}-\frac{a_{kj}+b_{kj}}{n+1}\right)}{2} \\ &= \tfrac{a_{il}+b_{kj}}{2} - \tfrac{a_{il}+b_{kj}}{2(n+1)} - \tfrac{a_{kj}+b_{il}}{2(n+1)} \\ &= \pi_{ij,kl} - \tfrac{1}{n+1}\pi_{ij,kl} - \tfrac{1}{n+1}\pi_{kl,ij} \end{aligned} \tag{4.17}$$

Summing up over all possible co-players $[k,l]$ thus leads to

$$\begin{aligned} \tilde{\pi}_{ij} &= \textstyle\sum_{k,l} \tilde{\pi}_{ij,kl} \cdot x_{kl} \\ &= \textstyle\sum_{k,l} \pi_{ij,kl} \cdot x_{kl} - \sum_{k,l} \tfrac{1}{n+1}\pi_{ij,kl} \cdot x_{kl} - \sum_{k,l} \tfrac{1}{n+1}\pi_{kl,ij} \cdot x_{kl} \\ &= \pi_{ij} - \tfrac{1}{n+1}\pi_{ij} - \tfrac{1}{n+1}\pi^T_{ij} \end{aligned} \tag{4.18}$$

For the average payoff in the population we finally sum up over all strategies $[i,j]$, leading to

$$\begin{aligned} \tilde{\pi}^e &= \textstyle\sum_{i,j} \tilde{\pi}_{ij} \cdot x_{ij} \\ &= \textstyle\sum_{i,j} \pi_{ij} \cdot x_{ij} - \sum_{i,j} \tfrac{1}{n+1}\pi_{ij} \cdot x_{ij} - \sum_{i,j} \tfrac{1}{n+1}\pi^T_{ij} \cdot x_{ij} \\ &= \pi^e - \tfrac{1}{n+1}\pi^e - \tfrac{1}{n+1}\pi^e. \end{aligned} \tag{4.19}$$

Hence we indeed end up with

$$x_{ij}[\tilde{\pi}_{ij} - \tilde{\pi}^e] = x_{ij}\left[\pi_{ij} - \frac{n-1}{n+1}\pi^e - \frac{1}{n+1}\pi_{ij} - \frac{1}{n+1}\pi^T_{ij}\right], \tag{4.20}$$

which proves that the two representations for the local replicator dynamics, (4.8) and (4.10), are equivalent.

C Numerical simulations

The previous analysis suggests that under appropriate conditions, opportunism is able to socialize punishment, which in turn promotes the evolution of cooperation. In order to verify this prediction, we have run extensive numerical simulations. The protocol of the simulations follows the lines of the well-known Moran process (Nowak et al., 2004). We consider a finite population of size $n = 160$. Each individual of the population plays a strategy $[i, j]$ with $i \in \{AllC,\ O_C,\ O_D,\ AllD\}$ and $j \in \{P,\ N,\ A,\ S\}$. In the initial population all 16 strategies are equally abundant: Each strategy is played by 10 individuals. We assume that every individual plays against any other member of the population and thereby obtains a payoff that is specified by (4.3) and (4.4).

Afterwards, an evolutionary strategy updating process takes place. Two players, say x and y with payoffs π_x and π_y, are chosen randomly. Player x adopts the strategy of player y with a probability that is an increasing function of the payoff difference $\pi_y - \pi_x$. A frequently used parametrization of this transition probability $p_{x\to y}$ is the so-called Fermi-rule (Blume, 1993; Szabó and Toke, 1998; Traulsen et al., 2006b):

$$p_{x\to y} = \frac{1}{1 + \exp\left[-s(\pi_y - \pi_x)\right]} \tag{4.21}$$

The parameter $s \geq 0$ denotes the imitation strength: For small s, a coin toss essentially decides whether or not to imitate the role model. In the other limit $s \to \infty$, player x always imitates co-players that have a higher payoff. These two limits are usually referred to as the case of weak and of strong selection, respectively.

Additionally we allow for mutations. In each time step, there is a positive probability $m > 0$ that one member of the population is chosen randomly. This member is then allowed to switch to a different strategy, with each of the 16 strategies having an equal chance to be selected. In the numerical simulations we have assumed a mutation rate $m = 0.001$.

We have presented a typical simulation run as Fig. 4.3 of the main text, showing the first $10,000$ time steps of the stochastic process. Additionally, we may calculate with which frequency a given strategy is played in the long run. For this reason, we have simulated the above selection-mutation process over a period of $2 \cdot 10^9$ time steps (i.e., each of the 160 individuals may implement more than 10^7 strategy changes). Fig. 4.5 shows the resulting average frequencies as a function of the selection strength s. We can roughly distinguish between three different scenarios:

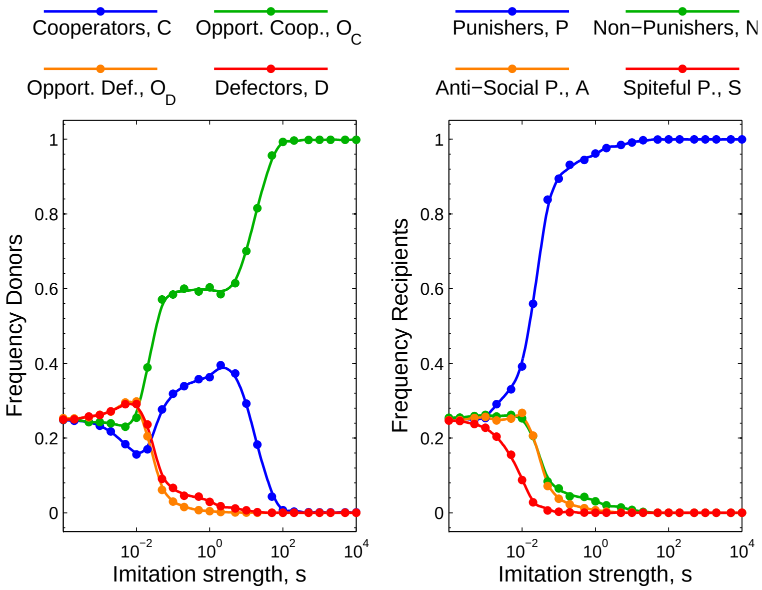


Figure 4.5: Long-run frequencies for the Moran process with mutations. The left graph shows which strategies are selected by the donors, whereas the right graph presents the strategy choices of the recipients. Parameter values: $c = \gamma = 1$, $b = 4$, $\beta = 3$, $\mu = 30\%$, $m = 0.001$ and population size 160.

(i) If selection is strong ($s \gg 1$), then nearly all of the donors are opportunistic cooperators, whereas almost all recipients are social punishers.

(ii) For medium selection strength, ($s \approx 1$), we still find that social punishment is the most abundant strategy among recipients. For donors, we find a stable coexistence of unconditional cooperators and opportunistic cooperators. In a population without defectors, both strategies, *AllC* and O_C, perform equally well. However, due to mutations, a minority of defectors may persist in the population. In such a mixed population, O_C has a slight payoff advantage compared to *AllC*, but the selection pressure is not strong enough to eliminate *AllC* completely. On average, the population consists to 2/3 of opportunistic cooperators and to 1/3 of unconditional altruists.

(iii) Only if selection is weak ($s \ll 1$) and game payoffs play a subordinate role on the strategies that are played, cooperation falls behind: In such a case, all strategies are played with almost equal shares, but only one out of four of the recipients' strategies supports cooperation (namely P), whereas the other three actions N, A and S implicitly promote defection.

The influence of the parameters b, c, γ, β and μ on the results of the Moran process is similar to their impact on the local replicator dynamics and we therefore omit a detailed analysis. The mutation rate m has only a minor influence on the long-run frequencies as long as mutations are sufficiently rare ($m \ll 1$). High mutation rates lead to a situation where all strategies are almost equally abundant; in this case, the results mirror the findings that we have obtained for weak selection ($s \ll 1$).

Overall, the results of the numerical simulations strongly support our analytical conclusions from the previous sections. If players have the opportunity to gain a reputation, then social punishment is predominant among the population. This in turn promotes the evolution of cooperation, provided that neither selection is too weak nor that mutations are too frequent.

Chapter 5

A short note on equilibrium notions and framing effects

Abstract

Empirical evidence suggests that the Nash equilibrium leads to inaccurate predictions of human behaviour in a vast set of games. This observation has promoted the development of new solution concepts like the quantal response equilibrium (QRE, see McKelvey and Palfrey, 1995) or evolutionary equilibria that are based on the long-run performance of a strategy (Antal et al., 2009; Ohtsuki, 2010). However, it is well-known that the QRE is subject to framing effects: Duplicating a strategy affects the equilibrium predictions. Here we show that the above mentioned evolutionary equilibria exhibit the same inconsistency. Furthermore, we prove that such framing effects are inevitable if a game theoretic solution concept depends differentiably on the payoffs. As a consequence, we argue that differentiable equilibrium notions, while being of great help in analyzing well-specified games, are unsuitable for theoretical modeling, where it is not clear which payoff matrix gives the true representation of an economic interaction.

5.1 Introduction

For any class of games, static or dynamic, with complete or incomplete information, it is a simple task to create an example where the Nash equilibrium mispredicts human behaviour, as shown for example in Goeree and Holt (2001). This holds true even for the most simple games with only one rationalizable equilibrium, like in the traveler's dilemma introduced by Basu (1994). Suppose that two travelers, returning home from their vacation, discover that the airline has lost their luggage. The airline asks both travelers independently to make claims for compensation and, in order to prevent excessive claims, determines that only the lower of both amounts will be paid. Additionally, it is announced that if the claims are different, the person with the lower claim obtains some reward $R > 1$, whereas the same amount R will be deducted from the other traveler's reimbursement as a penalty. In case that only claims between \$180 and \$300 are accepted, the Nash prediction is straightforward: In order to rake in the reward, it is always optimal to undercut the co-player's claim by one dollar. Consequently, the lower bound of \$180 is the unique equilibrium. While this analysis holds true for any $R > 1$, simple intuition suggests that subjects in the laboratory may try to coordinate on a higher claim if R is comparably low.[1] Indeed, this intuition is confirmed by experiments: For $R = 5$, around 80 % of the subjects opt for the maximum claim; only if R is sufficiently increased, claims approach the Nash equilibrium outcome (Goeree and Holt, 2001). Seemingly, subjects in these experiments do not strictly stick to best responses and do not necessarily eliminate dominated strategies.

These observations are the starting point for several alternative equilibrium notions. In this article we will review two distinct examples, the quantal response equilibrium (QRE) of McKelvey and Palfrey (1995) and the evolutionary equilibrium described in Ohtsuki (2010). Instead of considering traditional steplike best response correspondences, these equilibrium notions assume that strategy choices are positively but imperfectly related to payoffs.[2] As a consequence, also dominated strategies may be played from time to time, which in turn may affect equilibrium behaviour. Ironically, *because* these alternative

---

[1] As Kaushik Basu (1994) puts it, the strategy pair ("large","large") is a Nash equilibrium in ill-defined categories; if a player is told that the other player will choose a large number and if the reward R is neglectable, then the best reply is to choose a large number as well. This explanation bears some similarity with the examples in Camerer and Fehr (2006), who describe under which conditions a minority of irrational agents can trigger a majority of rational individuals to mimic the minority's behaviour.

[2] The same idea has also been applied to some learning models, for example smooth fictitious play, see Fudenberg and Levine (1998).

equilibrium notions allow a more realistic description of human behaviour, they also have a serious drawback: These equilibrium notions *themselves* are subject to framing effects. Different representations of the same economic situation result in different predictions. In particular, giving a strategy a second alias may affect the position of the equilibrium.

We proceed as follows: In the next section, we review the QRE and the evolutionary equilibrium described in Ohtsuki (2010). We show how two seemingly equivalent games can lead to diametrically opposed equilibrium predictions. While such framing effects are well-known in the case of the QRE, they have not been previously reported for the evolutionary equilibria. In Section 5.3 we give an unexpected sufficient condition for such framing effects: If an equilibrium concept depends differentiably on the payoffs then inconsistencies are inevitable.[3] As a consequence it is argued in Section 5.4 that the QRE and other differentiable equilibrium notions, although being of great help in analyzing already specified strategic games, might be unsuitable for doing theory, where the true representation of an economic problem is far from being clear.

5.2 Examples of equilibrium notions with framing effects

5.2.1 The quantal response equilibrium

The QRE was introduced by Richard D. McKelvey and Thomas R. Palfrey, first for games in normal form (1995) and later also for extensive form games (1998). Goeree et al. (2005) provide an axiomatic foundation. Since then, this concept was applied to various economic settings, including the traveler's dilemma (Capra et al., 1999) or coordination games (Anderson et al., 2001). Typically, the QRE outplays the Nash equilibrium by far when it comes to predict human behaviour in laboratory experiments.[4] Remarkably, the QRE can also be used to estimate the rationality of the subjects (McKelvey and Palfrey, 1995) and to which extent they believe in their co-player's rationality (Weizsäcker, 2003).

For our purposes, it will be sufficient to consider the simplest case, a finite normal form game between two players. The R-player chooses a row of the matrix $M = \big(a_{kl}, b_{kl}\big)$,

---

[3] Roughly speaking, differentiability means that small changes of the payoffs lead to a small and predictable change of the equilibrium. Note that the Nash equilibrium concept does not satisfy this condition, since small changes in the payoffs may completely change the best response correspondences.

[4] The overwhelming success is illustrated by the following quote of Camerer et al. (2004): *Quantal response equilibrium, a statistical generalization of Nash, almost always explains the deviations from Nash and should replace Nash as the static benchmark to which other models are routinely compared.*

whereas the C-player chooses a column. As usual, players are allowed to randomize between their pure actions; we denote by p^R and p^C the respective mixed strategy vectors. For each player $K \in \{R, C\}$, we denote by u_i^K the expected payoff of K's pure action i, which of course depends on the co-player's strategy p^{-K}, that is $u_i^K = u_i^K(p^{-K})$. A main aspect of the QRE is that choice probabilities are positively but imperfectly related to payoffs. According to the most commonly used parametrization of the QRE, the logit rule, the probability to play action i is determined by the following stochastic reaction function σ:

$$p_i^K = \sigma(u_i^K) = \frac{\exp(\lambda \cdot u_i^K)}{\sum_j \exp(\lambda \cdot u_j^K)} \tag{5.1}$$

The sum in the denominator ensures that the probabilities sum up to one. The parameter λ can be interpreted as a measure of rationality: $\lambda = 0$ means that actions are chosen randomly from the set of possible alternatives, whereas for large λ the choice is increasingly biased towards the strategy with the highest payoff. Note that as long as $\lambda < \infty$, even dominated strategies get a positive weight. For analyzing data, the parameter λ is typically estimated using the maximum likelihood method. A logit equilibrium is then defined as a fixed point of the map σ: A pair of mixed strategies $\hat{p} = (\hat{p}^R, \hat{p}^C)$ is an equilibrium if for both players $K \in \{R, C\}$ and all their strategies i the following condition holds:

$$\hat{p}_i^K = \sigma\big(u_i^K(\hat{p}^{-K})\big). \tag{5.2}$$

Such equilibria always exist but need not to be unique. As λ goes to infinity, logit equilibria approach Nash equilibria. Furthermore, the graph of all fixed points $\hat{p}$ contains a unique branch, starting at the centroid of the strategy simplex for $\lambda = 0$ and converging to a unique Nash equilibrium as λ approaches infinity, implying that the logit equilibrium can be applied to the problem of equilibrium selection. Since the stochastic reaction function σ depends differentiably on the payoffs for $\lambda < \infty$, by the implicit function theorem the same holds true for each branch of the graph of the logit equilibria.[5]

Let us illustrate the logit equilibrium with an example taken from Goeree and Holt (2001). Consider the following coordination game in which players receive \$1.80 for coordinating on the high equilibrium and \$0.90 if they coordinate on the low equilibrium. Additionally, the column-player has an outside option that guarantees a safe payoff of

---

[5] As we will see in Section 5.3, the smooth dependence on the payoffs plays a key role. It is valid not only for the logit equilibrium but for the QRE in general, since stochastic reaction functions are generally assumed to be differentiable, see McKelvey and Palfrey (1995) resp. Goeree et al. (2005).

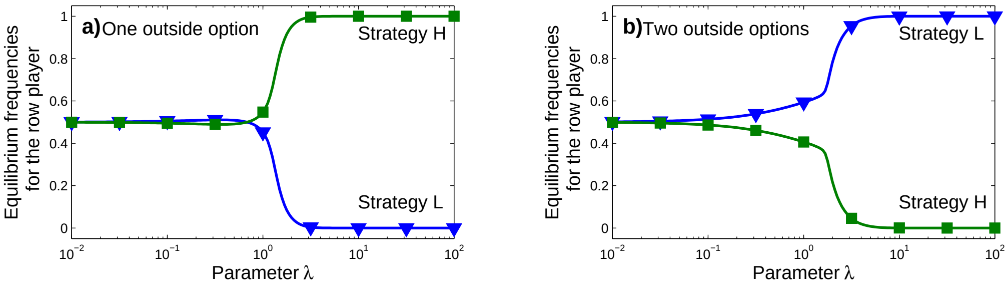


Figure 5.1: The unique branch of the logit equilibrium for the coordination game with one, respectively with two outside options ($x = 160$). a) In the game with one outside option, the value of x is small enough to enable subjects to coordinate on the high equilibrium. b) However, if the problem is represented differently, splitting the safe option in two, players are predicted to coordinate on the low equilibrium.

\$0.40:

	L	H	S
L	$90, 90$	$0, 0$	$x, 40$
H	$0, 0$	$180, 180$	$0, 40$

This game has two pure Nash equilibria, (H, H) and (L, L), and the safe option is never part of an equilibrium. Nevertheless, as shown in Goeree and Holt (2004), the outside option has a deciding influence on coordination behaviour in behavioural experiments. In particular, the exact value of x controls which strategies are chosen, with sufficiently low values of x preferring the (H, H) equilibrium. Such an effect is correctly predicted by the logit equilibrium but not by the Nash equilibrium, see Fig. 5.1a for an example with $x = 160$: The unique branch of the logit equilibrium, starting in the center for $\lambda = 0$ converges to the high equilibrium in the limit of rational agents, $\lambda \to \infty$. To illustrate that the QRE is subject to framing effects, we consider the same game, but with the second player having two (identical) outside options:

	L	H	S_1	S_2
L	$90, 90$	$0, 0$	$x, 40$	$x, 40$
H	$0, 0$	$180, 180$	$0, 40$	$0, 40$

(5.3)

While this additional strategy has no effect on the Nash equilibria, it alters the set of logit equilibria (Fig. 5.1b): Giving the outside option a second *name* leads the logit equilibrium to select the low instead of the high equilibrium. Thus, the prediction of the logit equilibrium depends sensitively on the exact formulation of the alternatively possible strategies.[6]

5.2.2 Evolutionary equilibria

Ohtsuki (2010) considers the following model of an evolutionary dynamics for an asymmetric game: There are two populations, a population R of row-players and a population C of column-players, with population size N_R resp. N_C. Each player in R chooses a row i of the matrix $M = (a_{kl}, b_{kl}) \in \mathbb{R}^{mn}$, whereas each player in C chooses a column j. Then, every subject in population R plays against every subject in the other population, leading to the payoffs u_i^R and u_j^C, respectively. Subsequently, the fitness f_i of a player with strategy i is defined by an exponential transformation of its payoff, i.e. $f_i^K = \exp(\delta u_i^K)$ with $K \in \{R, C\}$. The factor δ measures the importance of the game for the fitness of a player and is usually called the strength of selection. If $\delta \to 0$, each agent has approximately the same fitness, a case which is termed the weak selection limit.

After those interactions, one subject (of any of the two populations) is chosen at random. This agent is allowed to change its strategy by imitating the strategy of another player of the same population. It is assumed that strategies with higher fitness are more likely to be adopted. More specifically, if N_k^R denotes the current number of row-players with strategy k, then the probability that a randomly chosen row-player imitates an agent with strategy i is given by

$$p_i^R = \frac{N_i^R f_i^R}{\sum_k N_k^R f_k^R}. \tag{5.5}$$

Additionally, one allows mutations: With probability u, the agent does not imitate others, but chooses randomly any of the available strategies.

---

[6] In the above example one might argue that the inconsistency can be avoided if identical columns are omitted by definition. However, if the game is marginally modified such that there are no identical columns, elimination of the additional column seems unjustified:

$$\begin{array}{c|c|c|c|c|}
 & L & H & S_1 & S_2 \\ \hline
L & 90,90 & 0,0 & x,40 & x+\varepsilon,40 \\ \hline
H & 0,0 & 180,180 & 0,40 & 0,40 \\ \hline
\end{array} \tag{5.4}$$

More fundamentally, it is typically not the subjects who construct payoff matrices to help them with their decisions, but it is the researcher who uses such tools to describe the decision maker's behaviour. How should one decide which matrix gives the true representation of the decision problem? Similarly, to adapt the argumentation of Kohlberg and Mertens (1986) on a related issue, *elementary transformations*, [like giving a column a second alias], *are irrelevant for correct decision making: after all, the transformed matrix is just a different representation of the same decision problem, and decision theory should not be misled by representation effects. To hold the opposite point of view is to admit that decision theory is useless in real-life applications, where problems present themselves without a special formalism.*

Overall, this evolutionary dynamics results in a stochastic selection-mutation process without absorbing states. In some special cases, it is possible to explicitly calculate the invariant distribution of the process. One important case is the limit of weak selection, in which the fitness of each individual is largely independent of its payoff in the game. As a consequence, each strategy for population R is approximately played with probability $1/m$, only slightly truncated by a term φ_i, which reflects the impact of the respective strategy. If both populations are of equal size this term is, up to a multiplicative constant, given by

$$\varphi_i = \bar{a}_i - \bar{a}, \tag{5.6}$$

where $\bar{a}_i$ denotes the average of all feasible payoffs for a player with strategy i, that is $\bar{a}_i = \sum_j a_{ij}/n$, and $\bar{a}$ denotes the average of all feasible payoffs for individuals in population R, $\bar{a} = \sum_{i,j} a_{ij}/(mn)$.[7] Since this mutation-selection process does not settle down on any stable state, the deviation terms φ_i take the role of the major characteristic of the system. It is said that *selection favors strategy i* if φ_i is positive (Nowak, 2006a). Furthermore, one may compare two different strategies with each other: Antal et al. (2009) call *strategy i more abundant than strategy k* if $\varphi_i > \varphi_k$. In effect, this approach allows a ranking of the strategies - based on the long run performance of each strategy in the above described evolutionary process.

However, it is easy to show that this evolutionary equilibrium exhibits the same framing effects as the QRE. In fact, calculating φ for the example in the previous section (for $x = 60$) yields $\varphi_L^R = -5$ and $\varphi_H^R = 5$ in the case of representation (5.3), respectively $\varphi_L^R = 3.75$ and $\varphi_H^R = -3.75$ in the case of representation (5.4). Hence, neither does the absolute value of φ_L^R allow a consistent assessment across the different treatments, nor is the order of H and L left invariant. In the case of weak selection and uniform mutations, one can easily determine the reason for this inconsistency: While in the first representation, the outside option is played by roughly 1/3 of all column-players, this fraction increases to approximately 1/2 of the C-population if there are two outside options, which in turn encourages row-players to choose strategy L.

---

[7] This expression for φ_i resembles the well-known replicator dynamics, where it is assumed that the frequency of players with strategy i increases if the payoff u_i exceeds the average payoff in the population $\bar{u}$ (see, for example Weibull, 1995). However, while the payoffs u_i and $\bar{u}$ may vary over time, depending on the current state of population, the term φ_i is constant and does only depend on the payoff matrix.

5.3 An Impossibility Result

Let us turn to the question whether it is possible to construct other equilibrium notions (or other parametrizations of the QRE, respectively more general evolutionary equilibria) that avoid the inconsistencies shown in the previous section. For simplicity, we focus on two-player normal form games and identify each game Γ with its payoff matrix. For this reason, denote by $\mathcal{M} = \big\{(a_{ij}, b_{ij}) :\ a_{ij}, b_{ij} \in \mathbb{R}\big\}$ the set of all payoff matrices for normal form games, and let $\mathcal{M}_k \subset \mathcal{M}$ be the set of all payoff matrices that have exactly k rows (i.e., that admit k pure strategies for the row-player).

Definition 1 (Assessment formula)
A game theoretic assessment is a function $f : \mathcal{M}_k \to \mathbb{R}^k$.

One may interpret each entry of $f(M)$ as the predicted equilibrium frequencies for the row-player in the normal form game defined by the payoff matrix M.[8] In particular, note that the above definition of game theoretic assessments includes the logit equilibrium $\hat{p}$ and the evolutionary assessment φ from the previous section.

In order to exclude framing effects, we demand that equivalent representations of a game lead to the same assessments. Up to a renumbering of the strategies of the column-player, we say that two matrices $M, \hat{M} \in \mathcal{M}_k$ are equivalent if they result in the same matrix after deleting all columns that are a copy of a previous column. More formally, $M \sim \hat{M}$ if for all columns j of M there is a column l in $\hat{M}$ such that $(a_{ij}, b_{ij}) = (\hat{a}_{il}, \hat{b}_{il})$ for all rows i (and vice versa, for all columns l in $\hat{M}$ there is such a column j in M). Obviously, this defines an equivalence relation on the set $\mathcal{M}_k$ for all k.

Definition 2 (Consistent assessment formulas)
Fix a $k \geq 2$. *An assessment formula* $f : \mathcal{M}_k \to \mathbb{R}^k$ *is called consistent if it has the following properties:*

(i) Non-manipulability: If $M, K \in \mathcal{M}_k$ *and* $K \sim M$ *then* $f(K) = f(M)$.[9]

(ii) Validity: If the row-player's strategy i *is strictly dominated, then* i *cannot be optimal,* $f_i(M) < \max_j f_j(M)$.

---

[8] In this case one can restrict the image of the game theoretic assessment to the unit simplex Δ^k instead of $\mathbb{R}^k$. Since equilibria need not to be unique, this interpretation requires that for each game one particular equilibrium is selected out of the set of possible equilibria.

[9] A similar condition can be found, for example, in Milnor's famous work on *games against nature*, see Milnor (1951).

Non-manipulability means that a consistent assessment is *well-defined* with respect to the above equivalence relation, i.e. it respects that two matrices $M, \hat{M}$ with $M \sim \hat{M}$ represent the same game. Validity excludes constant assessment formulas from being considered. After these preparations, we are able to formulate the main result:

Theorem 3 (An impossibility theorem)
There is no assessment formula that is both, consistent and differentiable.

Proof. We show the case $k = 2$: Suppose there is such a consistent and differentiable assessment f and consider the arbitrary but fixed matrix

$$M = \begin{bmatrix} (a_{11}, b_{11}) & (a_{12}, b_{12}) \\ (a_{21}, b_{21}) & (a_{22}, b_{22}) \end{bmatrix}.$$

Let $\partial f / \partial \alpha_{ij}$ denote the marginal change of the assessment if the row-player's payoff in the i-th row and the jth column is varied. The idea of the proof is as follows: Non-manipulability implies that all partial derivatives $\partial f / \partial \alpha_{ij}(M)$ are zero, which suggests that the value of $f(M)$ is independent of the values of a_{ij}. This in turn contradicts validity. To show that the derivatives equal zero, we define the two matrices

$$M_1(t) = \begin{bmatrix} (a_{11} + t, b_{11}) & (a_{12}, b_{12}) \\ (a_{21}, b_{21}) & (a_{22}, b_{22}) \end{bmatrix} \quad \text{and}$$

$$M_2(t) = \begin{bmatrix} (a_{11} + t, b_{11}) & (a_{11} + t, b_{11}) & (a_{11} + t, b_{11}) & (a_{12}, b_{12}) \\ (a_{21}, b_{21}) & (a_{21}, b_{21}) & (a_{21}, b_{21}) & (a_{22}, b_{22}) \end{bmatrix}.$$

$M_2(t)$ is obtained from $M_1(t)$ by doubling the first column two times. Note that $M_1(0) = M$. Next we define two functions that measure how the respective strategy assessments vary with t, i.e. we define $u_i : \mathbb{R} \to \mathbb{R}^k$ with $u_i(t) = f\big(M_i(t)\big)$ for $i = 1, 2$. Since $M_1(t) \sim M_2(t)$ for all t, non-manipulability implies that $u_1(t) = u_2(t)$. In particular, the derivatives for $t = 0$ coincide:

$$u_1'(0) = \frac{\partial f}{\partial \alpha_{11}}(M) = \frac{\partial f}{\partial \alpha_{11}}\big(M_2(0)\big) + \frac{\partial f}{\partial \alpha_{12}}\big(M_2(0)\big) + \frac{\partial f}{\partial \alpha_{13}}\big(M_2(0)\big) = u_2'(0). \tag{5.7}$$

Therefore, since we want to show $\partial f / \partial \alpha_{11}(M) = 0$, we have to compute the expression on the right hand's side of (5.7). For this reason, we define two new matrices:

$$M_3(t) = \begin{bmatrix} (a_{11}, b_{11}) & (a_{11}, b_{11}) & (a_{11}+t, b_{11}) & (a_{12}, b_{12}) \\ (a_{21}, b_{21}) & (a_{21}, b_{21}) & (a_{21}, b_{21}) & (a_{22}, b_{22}) \end{bmatrix} \quad \text{and}$$

$$M_4(t) = \begin{bmatrix} (a_{11}, b_{11}) & (a_{11}+t, b_{11}) & (a_{11}+t, b_{11}) & (a_{12}, b_{12}) \\ (a_{21}, b_{21}) & (a_{21}, b_{21}) & (a_{21}, b_{21}) & (a_{22}, b_{22}) \end{bmatrix}.$$

Note that these two matrices have the same reduced normal form and hence are equivalent. Additionally, they fulfill $M_3(0) = M_4(0) = M_2(0)$. If we again define functions $u_i(t) = f\big(M_i(t)\big)$ for $i = 3, 4$, we may conclude that

$$u_3'(0) = \frac{\partial f}{\partial \alpha_{13}}\big(M_2(0)\big) = \frac{\partial f}{\partial \alpha_{12}}\big(M_2(0)\big) + \frac{\partial f}{\partial \alpha_{13}}\big(M_2(0)\big) = u_4'(0), \tag{5.8}$$

and therefore $\partial f/\partial\alpha_{12}\big(M_2(0)\big) = 0$. With a similar calculation one can show that the other two expressions on the right hand's side of (5.7) , $\partial f/\partial\alpha_{11}\big(M_2(0)\big)$ and $\partial f/\partial\alpha_{13}\big(M_2(0)\big)$, vanish as well. Therefore, we indeed end up with $\partial f/\partial\alpha_{11}(M) = 0$. A symmetry argument then immediately implies that $\partial f/\partial\alpha_{ij}(M) = 0$ for all i and j. As a consequence, the assessment $f(M)$ does not vary in the row-player's payoffs, which leads to a contradiction with the validity of the assessment. □

Therefore, we must conclude that there is no reasonable equilibrium concept that is both, non-manipulable and smooth. If we interpret the value of $f_i(M)$ slightly differently, as an indicator of the performance of strategy i, then the previous theorem states that is impossible to measure the success of a strategy with a differentiable formula.[10]

In particular, the inconsistencies of the evolutionary assessment φ cannot be simply attributed to the assumption of weak selection. Even in the case of some positive but finite selection pressure δ, the stationary distribution in Ohtsuki (2010) depends differentiably on the entries of the payoff matrix.

5.4 Discussion

Explaining human behaviour with game theoretic models faces at least two difficulties. Firstly, the modeller does usually not know the exact subjective utilities of the agents;

---

[10] Note that the above theorem is sharp in the sense that there exist assessment formulas that are consistent and (Lipschitz-)continuous. As an example, consider the assessment that assigns to each strategy i its maximum attainable payoff, $f_i(M) = \max_j a_{ij}$.

instead there might be only some rough estimates. In order to obtain robust results, one might therefore require that the output of the model depends differentiably on the input data. Secondly, in order to set up the model, the researcher needs to choose one specific description of reality, out of many alternatively possible descriptions. One such choice might entail, for example, to determine whether a certain player has only one outside option or several similar options. In the best case - if the methods are consistent in the sense defined above - the exact representation of the game does not affect the qualitative results.

However, as we have shown, the two requirements of consistency and differentiability are incompatible. If the results of a game theoretic equilibrium notion depend differentiably on the payoffs, then these results also depend on the representation. In this sense, solution concepts for games in strategic form are necessarily imperfect. Therefore any game theoretic concept that can be applied to normal form games faces the choice whether it violates one requirement or the other. The matrix presented as Tab. 5.1 attempts to give an overview over some choices that were made. It classifies some popular game theoretic tools according to whether they violate the smooth dependence on payoffs condition or the non-manipulability condition. Of course, such a list is notoriously incomplete and each cell of this matrix might contain several other elements - with the exception of the cell that corresponds to the differentiable and consistent concepts.

A natural question is then to ask which of the two requirements is the more indispensable one. Differentiable equilibrium notions, and in particular the QRE, are quite successful in predicting human behaviour for normal form games - once it is known which representation of the game the subjects choose. In laboratory experiments this is certainly no problem, since it may be assumed that the players' internal model of the game is close to the instructions that are provided by the experimenter (in particular it is likely that all subjects have a similar internal representation). From a behavioural point of view, the framing effects presented in the previous sections even seem to be a desirable feature - after all it is well documented that humans are subject to framing effects as well.[11] Psychologically, it is not unreasonable to expect that a duplication of the outside option increases the number of L players in game (5.3). The outside options may act as a coordination device: Because both options point to the low equilibrium, this equilibrium may be interpreted as a focal point (Schelling, 1960).

However, if it comes to explain human behaviour in the field it is not at all clear how

---

[11] For the related question whether subjects in dynamic games play differently if confronted with different game trees that represent formally equivalent games, see McKelvey and Palfrey (1998).

	Differentiable concepts	Non-differentiable concepts
Consistent concepts		Nash equilibrium (Nash, 1950) **Refinements of the Nash equilibrium** Perfect equilibrium (Selten, 1975) Proper equilibrium (Myerson, 1978) **Approaches that apply the Nash equilibrium to transformed utilities** Fairness model of Fehr and Schmidt (1999) **Learning processes for which Nash equilibria are rest points** Fictitious play (Brown, 1951) Replicator dynamics (Taylor and Jonker, 1978) Best response dynamics (Gilboa and Matsui, 1991)
Inconsistent concepts	**Behavioural equilibrium notions** QRE (McKelvey and Palfrey, 1995) Level-k reasoning model (Stahl and Wilson, 1995) Noisy introspection (Goeree and Holt, 2004) **Smooth learning processes** Exponential fictitious play (Fudenberg and Levine, 1998) **Long run equilibria for evolutionary processes with uniform mutations and smooth selection** Moran process (Antal et al., 2009; Ohtsuki, 2010)	**Long run equilibria for evolutionary processes with uniform mutations and best-reply selection** Moran process with strong selection (Fudenberg et al., 2006)

Table 5.1: A classification of game theoretic concepts

individuals perceive their interactions, let alone that these perceptions are comparable across subjects. For theoretical modeling, the above described framing effects are undesirable (or even dangerous). If an equilibrium concept leads to predictions that depend on the representation of the game (which is chosen by the modeler himself), then the results will be somewhat arbitrary in the best case and manipulable in the worst.

A possible solution to avoid framing effects in differentiable equilibrium notions is to consider the equivalence class of a game. That is, instead of calculating the logit equilibrium $\hat{p}_\lambda(M)$ of a game M one may calculate the set of possible logit equilibria $\hat{p}_\lambda([M])$ for all games that are equivalent to M,

$$[M] = \left\{ \hat{M} \in \mathcal{M}_k \;\middle|\; \hat{M} \sim M \right\}. \tag{5.9}$$

However, in this case, the logit equilibrium loses its ability to select a unique Nash equilibrium in the limit of rational agents, $\lambda \to \infty$. Instead, most of the Nash equilibria of a game M (including all strict Nash equilibria) are predictable by the unique branch of $\hat{p}_\lambda(\hat{M})$ - if only the game is appropriately framed. Therefore, it seems to me that the solution concept of the Nash equilibrium is (almost, see Selten, 1975; Myerson, 1978) as good as it gets.

Bibliography

G. Abramson and M. Kuperman. Social games in a social network. *Phys. Rev. E*, 63: 030901, 2001.

S. P. Anderson, J. K. Goeree, and C. A. Holt. Minimum-effort coordination games: Stochastic potential and logit equilibrium. *Games Econ. Behav.*, 34:177–199, 2001.

J. Andreoni, W. Harbaugh, and L. Vesterlund. The carrot or the stick: Rewards, punishments, and cooperation. *Am. Econ. Rev.*, 93:893–902, 2003.

T. Antal, A. Traulsen, H. Ohtsuki, C. E. Tarnita, and M. A. Nowak. Mutation-selection equilibrium in games with multiple strategies. *J. Theor. Biol.*, 258:614–622, 2009.

R. Axelrod. *The Evolution of Cooperation*. Basic Books, New York, 1984.

P. Barclay. Reputational benefits for altruistic punishment. *Evol. Hum. Behav.*, 27:325–344, 2006.

K. Basu. The traveler's dilemma: Paradoxes of rationality in game theory. *Am. Econ. Rev.*, 84:391–395, 1994.

M. Bateson, D. Nettle, and G. Roberts. Cues of being watched enhance cooperation in a real-world setting. *Biol. Lett.*, 2:412–414, 2006.

R. Battalio, L. Samuelson, and J. Van Huyck. Optimization incentives and coordination failure in laboratory stag hunt games. *Econometrica*, 69:749–764, 2001.

J. Berg, J. Dickhaut, and K. McCabe. Trust, reciprocity, and social history. *Games Econ. Behav.*, 10:122–142, 1995.

U. Berger. Best response dynamics for role games. *Int. J. Game Theory*, 30:527–538, 2001.

L. E. Blume. The statistical mechanics of strategic interaction. *Games Econ. Behav.*, 5: 387–424, 1993.

I. Bomze and C. Pawlowitsch. One-third rules with equality: Second-order evolutionary stability conditions in finite populations. *J. Theor. Biol.*, 254:616–620, 2008.

S. Bowles and H. Gintis. The evolution of strong reciprocity: Cooperation in heterogeneous populations. *Theor. Popul. Biol.*, 65:17–28, 2004.

R. Boyd and P. J. Richerson. Punishment allows the evolution of cooperation (or anything else) in sizeable groups. *Ethol. Sociobiol.*, 13:171–195, 1992.

R. Boyd, H. Gintis, S. Bowles, and P. J. Richerson. The evolution of altruistic punishment. *Proc. Natl. Acad. Sci. USA*, 100:3531–3535, 2003.

R. Boyd, H. Gintis, and S. Bowles. Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science*, 328:617–620, 2010.

H. Brandt, C. Hauert, and K. Sigmund. Punishing and abstaining for public goods. *Proc. Natl. Acad. Sci. USA*, 103:495–497, 2006.

G. Brown. Iterative solutions of games by fictitious play. In T. Koopmans, editor, *Activity Analysis of Production and Allocation*. Wiley, New York, 1951.

R. Bshary and A. S. Grutter. Image scoring and cooperation in a cleaner fish mutualism. *Nature*, 441:975–978, 2006.

C. F. Camerer and E. Fehr. When does "economic man" dominate social behavior? *Science*, 311:47–52, 2006.

C. F. Camerer, T.-H. Ho, and J. K. Chong. Behavioral game theory: Thinking, learning, and teaching. In S. Huck, editor, *Advances in Understanding Strategic Behavior, Essays in Honor of Werner Güth*. Palgrave MacMillan, London, 2004.

M. C. Capra, J. K. Goeree, R. Gomez, and C. A. Holt. Anomalous behavior in a traveler's dilemma? *Am. Econ. Rev.*, 89:678–690, 1999.

J. P. Carpenter. The demand for punishment. *J. Econ. Behav. Organ.*, 62:522–542, 2007.

M. Chamberland and R. Cressman. An example of dynamic (in)consistency in symmetric extensive form evolutionary games. *Games Econ. Behav.*, 30:319–326, 2000.

R. W. Cooper, D. V. DeJong, R. Forsythe, and T. W. Ross. Selection criteria in coordination games: Some experimental results. *Am. Econ. Rev.*, 80:218–233, 1990.

R. Cressman, A. Gaunersdorfer, and J.-F. Wen. Evolutionary and dynamic stability in symmetric evolutionary games with two independent decisions. *Int. Game Theory Rev.*, 2:67–81, 2000.

D. J. F. de Quervain, U. Fischbacher, V. Treyer, M. Schellhammer, U. Schnyder, A. Buck, and E. Fehr. The neural basis of altruistic punishment. *Science*, 305:1254–1258, 2004.

H. De Silva and K. Sigmund. Public good games with incentives: The role of reputation. In S. A. Levin, editor, *Games, groups and the global good.* Springer, New York, 2009.

L. Denant-Boemont, D. Masclet, and C. N. Noussair. Punishment, counterpunishment and sanction enforcement in a social dilemma experiment. *Econ. Theory*, 33:145–167, 2007.

U. Dieckmann, M. Heino, and K. Parvinen. The adaptive dynamics of function-valued traits. *J. Theor. Biol.*, 241:370–389, 2006.

A. Dreber, D. A. Rand, D. Fudenberg, and M. A. Nowak. Winners don't punish. *Nature*, 452:348–351, 2008.

L. A. Dugatkin and H. K. Reeve, editors. *Game Theory and Animal Behaviour.* Oxford University Press, Oxford, 1998.

R. Durrett and S. A. Levin. Stochastic spatial models: A user's guide to ecological applications. *Philos. T. Roy Soc. B*, 343:329–350, 1994.

M. Egas and A. Riedl. The economics of altruistic punishment and the maintenance of cooperation. *Proc. R. Soc. Lond. B.*, 275:871–878, 2008.

O. T. Eldakar and D. S. Wilson. Selfishness as second-order altruism. *Proc. Natl. Acad. Sci. USA*, 105:6982–6986, 2008.

D. S. Falster and M. Westoby. Plant height and evolutionary games. *Trends Ecol. Evol.*, 18:337–343, 2003.

E. Fehr and U. Fischbacher. The nature of human altruism. *Nature*, 425:785–791, 2003.

E. Fehr and S. Gächter. Cooperation and punishment in public goods experiments. *Am. Econ. Rev.*, 90:980–994, 2000.

E. Fehr and S. Gächter. Fairness and retaliation: The economics of reciprocity. *J. Econ. Persp.*, 14:159–181, 2000.

E. Fehr and S. Gächter. Altruistic punishment in humans. *Nature*, 415:137–140, 2002.

E. Fehr and K. M. Schmidt. A theory of fairness, competition, and cooperation. *Q. J. Econ.*, 114:817–868, 1999.

J. H. Fowler. Altruistic punishment and the origin of cooperation. *Proc. Natl. Acad. Sci. USA*, 102:7047–7049, 2005.

R. H. Frank. *Passions within reason: The strategic role of emotions.* WW Norton, New York, 1988.

D. Friedman. Equilibrium in evolutionary games: Some experimental results. *Econ. J.*, 106:1–25, 1996.

D. Fudenberg and L. A. Imhof. Imitation processes with small mutations. *J. Econ. Theory*, 131:252–262, 2006.

D. Fudenberg and D. Levine. Learning in games. *Eur. Econ. Rev.*, 42:631–639, 1998.

D. Fudenberg and J. Tirole. *Game Theory.* MIT Press, Cambridge, MA, 1991.

D. Fudenberg, M. A. Nowak, C. Taylor, and L. A. Imhof. Evolutionary game dynamics in finite populations with strong selection and weak mutation. *Theor. Popul. Biol.*, 70: 352–363, 2006.

S. Gächter, E. Renner, and M. Sefton. The long-run benefits of punishment. *Science*, 322: 1510–1512, 2008.

J. Gale, K. G. Binmore, and L. Samuelson. Learning to be imperfect: The ultimatum game. *Games Econ. Behav.*, 8:56–90, 1995.

A. Gardner and S. A. West. Cooperation and punishment, especially in humans. *Am. Nat.*, 164:753–764, 2004a.

A. Gardner and S. A. West. Spite and the scale of competition. *J. Evol. Biol.*, 17: 1195–1203, 2004b.

A. Gaunersdorfer. Time averages for heteroclinic attractors. *SIAM J. Appl. Math.*, 52: 1476–1489, 1992.

A. Gaunersdorfer, J. Hofbauer, and K. Sigmund. On the dynamics of asymmetric games. *Theor. Popul. Biol.*, 39:354–357, 1991.

S. Geritz, E. Kisdi, G. Meszéna, and J. Metz. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol*, 12:35–57, 1998.

I. Gilboa and A. Matsui. Social stability and equilibrium. *Econometrica*, 59:859–867, 1991.

H. Gintis. Strong reciprocity and human sociality. *J. Theor. Biol.*, 206:169–179, 2000.

H. Gintis, S. Bowles, R. T. Boyd, and E. Fehr, editors. *Moral Sentiments and Material Interests - The Foundations of Cooperation in Economic Life.* MIT Press, Cambridge, MA, 2005.

J. K. Goeree and C. A. Holt. Ten little treasures of game theory and ten intuitive contradictions. *Am. Econ. Rev.*, 91:1402–1421, 2001.

J. K. Goeree and C. A. Holt. A model of noisy introspection. *Games Econ. Behav.*, 46: 365–382, 2004.

J. K. Goeree, C. A. Holt, and T. R. Palfrey. Regular quantal response equilibrium. *Exp. Econ.*, 8:347–367, 2005.

C. Gokhale and A. Traulsen. Evolutionary games in the multiverse. *Proc. Natl. Acad. Sci. USA*, 107:5500–5504, 2010.

W. Güth, R. Schmittberger, and B. Schwarze. An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.*, 3:367–388, 1982.

K. J. Haley and D. M. T. Fessler. Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evol. Hum. Behav.*, 26:245–256, 2005.

G. Hardin. The tragedy of the commons. *Science*, 162:1243–1248, 1968.

J. Harsanyi and R. Selten. *A General Theory of Equilibrium Selection in Games.* MIT Press, Cambridge, MA, 1988.

S. Hart. Evolutionary dynamics and backward induction. *Games Econ. Behav.*, 41:227–264, 2002.

C. Hauert and M. Döbeli. Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature*, 428:643–646, 2004.

C. Hauert, N. Haiden, and K. Sigmund. The dynamics of public goods. *Discrete Contin. Dyn. Syst.*, B 4:575–585, 2004.

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

D. Helbing and W. Yu. The outbreak of cooperation among success-driven individuals under noisy conditions. *Proc. Natl. Acad. Sci. USA*, 106:3680–3685, 2009.

J. Henrich, R. McElreath, A. Barr, J. Ensminger, C. Barrett, A. Bolyanatz, J. C. Cardenas, M. Gurven, E. Gwako, N. Henrich, C. Lesorogol, F. Marlowe, D. Tracer, and J. Ziker. Costly punishment across human societies. *Science*, 312:1767–1770, 2006.

B. Herrmann, C. Thöni, and S. Gächter. Antisocial punishment across societies. *Science*, 319:1362–1367, 2008.

C. Hilbe. Local replicator dynamics: A simple link between deterministic and stochastic models of evolutionary game theory. *Bull. Math. Biol.*, 2011.

C. Hilbe and K. Sigmund. Incentives and opportunism: From the carrot to the stick. *Proc. R. Soc. Lond. B.*, 277:2427–2433, 2010.

J. Hofbauer. On the occurence of limit cycles in the Volterra-Lotka equation. *Nonlinear Anal. Theor.*, 5:1003–1007, 1981.

J. Hofbauer and W. H. Sandholm. Survival of dominated strategies under evolutionary dynamics. *Theor. Economics*, in press.

J. Hofbauer and K. Sigmund. Adaptive dynamics and evolutionary stability. *Appl. Math. Letters*, 7:65–70, 1994.

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

J. Hofbauer and K. Sigmund. Evolutionary game dynamics. *B. Am. Math Soc.*, 40: 479–519, 2003.

C. A. Holt and S. K. Laury. Risk aversion and incentive effects. *Am. Econ. Rev.*, 92: 1644–1655, 2002.

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

M. A. Janssen and C. Bushman. Evolution of cooperation and altruistic punishment when retaliation is possible. *J. Theor. Biol.*, 254:541–545, 2008.

D. Johnson and J. Bering. Hand of god, mind of man: Punishment and cognition in the evolution of cooperation. *Evol. Psychol.*, 4:219–233, 2006.

D. Kahnemann and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47:263–291, 1979.

T. Kiyonari and P. Barclay. Cooperation in social dilemmas: Free riding may be thwarted by second-order reward rather than by punishment. *J. Pers. Soc. Psychol.*, 95:826–842, 2008.

E. Kohlberg and J.-F. Mertens. On the strategic stability of equilibria. *Econometrica*, 54: 1003–1037, 1986.

R. Kurzban, P. DeScioli, and E. O'Brien. Audience effects on moralistic punishment. *Evol. Hum. Behav.*, 28:75–84, 2007.

J. O. Ledyard. Public goods: A survey of experimental research. In J. H. Kagel and A. E. Roth, editors, *Handbook of Experimental Economics*. Princeton University Press, Princeton, 1995.

L. Lehmann and F. Rousset. Perturbation expansions of multilocus fixation probabilities for frequency-dependent selection with applications to the Hill-Robertson effect and to the joint evolution of helping and punishment. *Theor. Popul. Biol.*, 76:35–51, 2009.

L. Lehmann, L. Keller, and D. Sumpter. The evolution of helping and harming on graphs: The return of the inclusive fitness effect. *J. Evol. Biol.*, 20:2284–2295, 2007a.

L. Lehmann, F. Rousset, D. Roze, and L. Keller. Strong reciprocity or strong ferocity? A population genetic view of the evolution of altruistic punishment. *Am. Nat.*, 170:21–36, 2007b.

S. Lessard. Long-term stability from fixation probabilities in finite populations: New perspectives for ESS theory. *Theor. Popul. Biol.*, 68:19–27, 2005.

S. Lessard and V. Ladret. The probability of fixation of a single mutant in an exchangeable selection model. *J. Math. Biol.*, 54:721–744, 2007.

S. A. Levin, editor. *Games, groups and the global good*, chapter Preface. Springer, New York, 2009.

E. Lieberman, C. Hauert, and M. A. Nowak. Evolutionary dynamics on graphs. *Nature*, 433:312–316, 2005.

A. Matsui. Best response dynamics and socially stable strategies. *J. Econ. Theory*, 57: 343–362, 1992.

J. Maynard Smith. *Evolution and the Theory of Games.* Cambridge University Press, Cambridge, UK, 1982.

J. Maynard Smith. Can a mixed strategy be stable in a finite population? *J. Theor. Biol.*, 130:209–221, 1988.

J. Maynard Smith and G. R. Price. The logic of animal conflict. *Nature*, 246:15–18, 1973.

R. McKelvey and T. Palfrey. Quantal response equilibria for normal form games. *Games Econ. Behav.*, 10:6–38, 1995.

R. McKelvey and T. R. Palfrey. Quantal response equilibria for extensive form games. *Exp. Econ.*, 1:9–41, 1998.

L. Mealey, C. Daood, and M. Krage. Enhanced memory for faces of cheaters. *Behav. Ecol. Sociobiol.*, 17:119–128, 1996.

M. Milinski, D. Seemmann, and H.-J. Krambeck. Reputation helps solve the 'tragedy of the commons'. *Nature*, 415:424–426, 2002.

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

J. Milnor. Games against nature. In R. M. Thrall, C. H. Coombs, and R. L. Davis, editors, *Decision Processes.* Wiley, New York, 1951.

R. B. Myerson. Refinements of the Nash equilibrium concept. *Int. J. Game Theory*, 15: 133–154, 1978.

M. Nakamaru and Y. Iwasa. The coevolution of altruism and punishment: Role of the selfish punisher. *J. Theor. Biol.*, 240:475–488, 2006.

M. Nakamaru, H. Matsuda, and Y. Iwasa. The evolution of cooperation in a lattice-structured population. *J. Theor. Biol.*, 184:65–81, 1997.

M. Nakamaru, H. Nogami, and Y. Iwasa. Score-dependent fertility model for the evolution of cooperation in a lattice. *J. Theor. Biol.*, 194:101–124, 1998.

J. Nash. Equilibrium points in n-person games. *Proc. Natl. Acad. Sci. USA*, 36:48–49, 1950.

N. Nikiforakis. Punishment and counter-punishment in public goods games: Can we really govern ourselves? *J. Pub. Econ.*, 92:91–112, 2008.

P. Niyogi. *The Computational Nature of Language Learning and Evolution.* MIT Press, Cambridge, MA, 2006.

M. A. Nowak. *Evolutionary Dynamics.* Harvard University Press, 2006a.

M. A. Nowak. Five rules for the evolution of cooperation. *Science*, 314:1560–1563, 2006b.

M. A. Nowak and R. M. May. Evolutionary games and spatial chaos. *Nature*, 359:826–829, 1992.

M. A. Nowak and K. Sigmund. Evolutionary dynamics of biological games. *Science*, 303: 793–799, 2004.

M. A. Nowak, K. Page, and K. Sigmund. Fairness versus reason in the ultimatum game. *Science*, 289:1773–1775, 2000.

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

M. A. Nowak, C. E. Tarnita, and T. Antal. Evolutionary dynamics in structured populations. *Proc. R. Soc. Lond. B.*, 365:19–30, 2010.

H. Ohtsuki. Stochastic evolutionary dynamics of bimatrix games. *J. Theor. Biol.*, 264: 136–142, 2010.

H. Ohtsuki and M. A. Nowak. The replicator equation on graphs. *J. Theor. Biol.*, 243: 86–97, 2006.

H. Ohtsuki and M. A. Nowak. Evolutionary stability on graphs. *J. Theor. Biol.*, 251: 698–707, 2008.

H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441:502–505, 2006.

H. Ohtsuki, P. Bordalo, and M. A. Nowak. The one-third law of evolutionary dynamics. *J. Theor. Biol.*, 249:289–295, 2007a.

H. Ohtsuki, M. A. Nowak, and J. M. Pacheco. Breaking the symmetry between interaction and replacement in evolutionary dynamics on graphs. *Phys. Rev. Lett.*, 98:108106, 2007b.

H. Ohtsuki, J. M. Pacheco, and M. A. Nowak. Evolutionary graph theory: Breaking the symmetry between interaction and replacement. *J. Theor. Biol.*, 246:681–694, 2007c.

M. Olson. *The logic of collective action*. Harvard University Press, Cambridge MA, 1965.

E. Ostrom and J. Walker. *Trust and reciprocity: Interdisciplinary lessons from experimental research*. Russel Sage Foundation, New York, 2003.

E. Ostrom, J. Walker, and R. Gardner. Covenants with and without a sword: Self-governance is possible. *Am. Polit. Sci. Rev.*, 86:404–417, 1992.

K. Panchanathan and R. Boyd. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature*, 432:499–502, 2004.

T. Pfeiffer and S. Schuster. Game-theoretical approaches to studying the evolution of biochemical systems. *Trends Biochem. Sci.*, 30:20–25, 2005.

S. Pinker, M. A. Nowak, and J. J. Lee. The logic of indirect speech. *Proc. Natl. Acad. Sci. USA*, 105:833–838, 2008.

M. L. Platt and S. A. Huettel. Risky business: The neuroeconomics of decision making under uncertainty. *Nat. Neurosi.*, 11:398–403, 2008.

D. G. Rand, A. Dreber, D. Fudenberg, T. Ellingsen, and M. A. Nowak. Positive interactions promote public cooperation. *Science*, 325:1272–1275, 2009.

D. G. Rand, J. J. Armao IV, M. Nakamaru, and H. Ohtsuki. Anti-social punishment can prevent the co-evolution of punishment and cooperation. *J. Theor. Biol.*, 265:624–632, 2010.

M. Rigdon, K. Ishii, M. Watabe, and S. Kitayama. Minimal social cues in the dictator game. *J. Econ. Psychol.*, 30:358–367, 2009.

J. Robinson. An iterative method of solving a game. *Ann. Math.*, 54:296–301, 1951.

C. P. Roca, J. A. Cuesta, and A. Sánchez. Time scales in evolutionary dynamics. *Phys. Rev. Lett.*, 97:158701, 2006.

B. Rockenbach and M. Milinski. The efficient interaction of indirect reciprocity and costly punishment. *Nature*, 444:718–723, 2006.

Y. I. Russel, J. Call., and R. I. M. Dunbar. Image socoring in great apes. *Behav. Proc.*, 78:108–111, 2008.

A. Sánchez and J. A. Cuesta. Altruism may arise from individual selection. *J. Theor. Biol.*, 235:233–240, 2005.

W. H. Sandholm. Book review of "Evolutionary dynamics and extensive form games" by Ross Cressman. *Int. Rev. Econ. Finance*, 15:136–140, 2006.

F. Santos, M. Santos, and J. Pacheco. Social diversity promotes the emergence of cooperation in public goods games. *Nature*, 454:213–216, 2008.

M. E. Schaffer. Evolutionarily stable strategies for a finite population and a variable contest size. *J. Theor. Biol.*, 132:469–478, 1988.

T. C. Schelling. *The Strategy of Conflict.* Harvard University Press, Cambridge MA, 1960.

D. Seemmann, H.-J. Krambeck, and M. Milinski. Strategic investment in reputation. *Behav. Ecol. Sociobiol.*, 56:248–252, 2004.

M. Sefton, R. Shupp, and J. M. Walker. The effects of rewards and sanctions in provision of public goods. *Econ. Inquiry*, 45:671–690, 2007.

R. Selten. Re-examination of the perfectness concept for equilibrium points in extensive games. *Int. J. Game Theory*, 4:25–55, 1975.

K. Sigmund. Time averages for unpredictable orbits of deterministic systems. *Ann. Oper. Res.*, 37:217–228, 1992.

K. Sigmund. Punish or perish? Retaliation and collaboration among humans. *Trends Ecol. Evol.*, 22:593–600, 2007.

K. Sigmund. *The Calculus of Selfishness.* Princeton University Press, Princetion, NJ, 2010.

K. Sigmund, C. Hauert, and M. A. Nowak. Reward and punishment. *Proc. Natl. Acad. Sci. USA*, 98:10757–10762, 2001.

K. Sigmund, H. De Silva, A. Traulsen, and C. Hauert. Social learning promotes institutions for governing the commons. *Nature*, 466:861–863, 2010.

D. O. Stahl and P. W. Wilson. On players' models of other players: Theory and experimental evidence. *Games Econ. Behav.*, 10:218–254, 1995.

P. G. Straub. Risk dominance and coordination failures in static games. *Q. Rev. Econ. Financ.*, 35:339–363, 1995.

G. Szabó and G. Fáth. Evolutionary games on graphs. *Phys. Rep.*, 446:97–216, 2007.

G. Szabó and C. Toke. Evolutionary prisoners dilemma game on a square lattice. *Phys. Rev. E*, 58:69–73, 1998.

C. Tarnita, T. Antal, and M. A. Nowak. Mutation-selection equilibrium in games with mixed strategies. *J. Theor. Biol.*, 261:50–57, 2009.

C. Taylor, D. Fudenberg, A. Sasaki, and M. A. Nowak. Evolutionary game dynamics in finite populations. *B. Math. Biol.*, 66:1621–1644, 2004.

P. Taylor and L. Jonker. Evolutionary stable strategies and game dynamics. *Math. Biosci.*, 40:145–156, 1978.

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

A. Traulsen, J. C. Claussen, and C. Hauert. Coevolutionary dynamics in large, but finite populations. *Phys. Rev. E*, 74:0119901, 2006a.

A. Traulsen, M. A. Nowak, and J. M. Pacheco. Stochastic dynamics of invasion and fixation. *Phys. Rev. E*, 74:011909, 2006b.

A. Traulsen, M. A. Nowak, and J. M. Pacheco. Stochastic payoff evaluation increases the temperature of selection. *J. Theor. Biol.*, 244:349–356, 2007.

A. Traulsen, D. Semmann, R. D. Sommerfeld, H.-J. Krambeck, and M. Milinski. Human strategy updating in evolutionary games. *Proc. Natl. Acad. Sci. USA*, 107:2962–2966, 2010.

R. L. Trivers. The evolution of reciprocal altruism. *Quart. Rev. Biol.*, 46:35–57, 1971.

S. Uchida and K. Sigmund. The competition of assessment rules for indirect reciprocity. *J. Theor. Biol.*, 263:13–19, 2010.

J. Vyrastekova and D. P. van Soest. On the (in)effectiveness of rewards in sustaining cooperation. *Exp. Econ.*, 11:53–65, 2008.

J. M. Walker and W. A. Halloran. Rewards and sanctions and the provision of public goods in one-shot settings. *Exp. Econ.*, 7:235–247, 2004.

J. Weibull. *Evolutionary Game Theory.* MIT Press, Cambridge, MA, 1995.

G. Weizsäcker. Ignoring the rationality of others: Evidence from experimental normal-form games. *Games Econ. Behav.*, 44:145–171, 2003.

G. Wild and P. D. Taylor. Fitness and evolutionary stability in game theoretic models of finite populations. *Proc. R. Soc. Lond. B.*, 271:2345–2349, 2004.

G. S. Wilkinson. Reciprocal food sharing in the vampire bat. *Nature*, 308:181–184, 1984.

B. Wölfing and A. Traulsen. Stochastic sampling of interaction partners versus deterministic payoff assignment. *J. Theor. Biol.*, 257:689–695, 2009.

T. Yamagishi. The provision of a sanctioning system as a public good. *J. Pers. Soc. Psychol.*, 51:110–116, 1986.

T. Yamagishi, Y. Horta, H. Takagishi, M. Shinada, S. Tanida, and C. S. Cook. The private rejection of unfair offers and emotional commitment. *Proc. Natl. Acad. Sci. USA*, 106: 11520–11523, 2009.

E. C. Zeeman. Population dynamics from game theory. In Z. Nitecki and C. Robinson, editors, *Global Theory of Dynamical Systems, Lecture Notes in Mathematics.* Springer, 1980.

Zusammenfassung

Public good games dienen als Modell für den Konflikt zwischen Allgemeinwohl und individuellem Vorteil: Während der Erfolg eines gemeinschaftlichen Projekts oft vom Einsatz aller Beteiligten abhängt, kann für den Einzelnen der Anreiz zum Trittbrettfahren bestehen. In dieser Dissertation untersuche ich das Zusammenspiel von Kooperation und Anreizsystemen mit Hilfe der evolutionären Spieltheorie. Es wird gezeigt, dass Belohnungen zwar individuelle Kooperation anstoßen können, dass aber Bestrafungsmöglichkeiten notwendig sind um die Zusammenarbeit aufrechtzuerhalten. Dabei liefert die individuelle Reputation der Spieler einen Anreiz, die Einhaltung von Normen zu überwachen und Abweichungen zu sanktionieren. Im Gegensatz zu früheren Studien werden Bestrafungsmechanismen jedoch nicht zur Stabilisierung von beliebigen Normen und Verhaltensvorschriften verwendet. Stattdessen werden Sanktionen gezielt dazu eingesetzt um die soziale Wohlfahrt zu verbessern.

In dieser Dissertation stelle ich auch einige mathematische Werkzeuge und methodische Konzepte vor, die bei der Untersuchung von public good games hilfreich sind. Dazu wird die Theorie der Rollenspiele erweitert und eine modifizierte Replikatorgleichung eingeführt. Unter dieser lokalen Replikatordynamik können sich selbst dominierte Strategien durchsetzen, falls diese zu einem relativen Vorteil führen.

Curriculum Vitae

Personal Information

Name:	Christian Hilbe
Date of birth:	December 18th, 1981
Place of birth:	Dornbirn, Austria
Citizenship:	Austrian

Higher Education

Oct. 2008 – Mar. 2011	Doctoral studies in mathematics, University of Vienna.
Sep. 2008	Master exam with distinction in mathematics, University of Vienna.
Mar. 2004 – Sep. 2008	Undergraduate studies of mathematics, University of Vienna.

Scientific Employment

Since Oct. 2008	Research assistant at the Department of Mathematics (University of Vienna), fully supported by EUROCORES TECT I-104-G15.
Feb. 2008 – Jun. 2008	Research assistant at the Industriewissenschaftliches Institut, Vienna.

Research Stays

Jun. 2009 – Aug. 2009	Participation in the Young Scientists Summer Program of the International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria
Oct. 2010 – Feb. 2011	Research stay at the Center of Rationality of the Hebrew University, Jerusalem.

Awards

Oct. 2009	Würdigungspreis of the Austrian Ministry of Science and Research
Mar. 2010	Mikhalevich Scholarship of the IIASA

List of Publications

Articles

(i) Christian Hilbe (2009). Contrition does not ensure cooperation in the iterated prisoner's dilemma. *Int. J. Bifurcat. Chaos* 19, 3877-3885.

(ii) Christian Hilbe and Karl Sigmund (2010). Incentives and opportunism: From the carrot to the stick. *Proc. Roy. Soc. B* 277, 2427–2433.

(iii) Christian Hilbe (2011). Local replicator dynamics: A simple link between deterministic and stochastic models of evolutionary game theory. *Bull. Math. Biol.*, in press.

Diploma thesis

Wiederholte Spiele mit Fehlern. Advised by Prof. Karl Sigmund.

Selected talks and presentations

(i) Incentives and opportunism: From the carrot to the stick.
Talk at the "Mini-Symposium on Evolutionary Game Theory", August 9th, 2010, Plön, Germany.

(ii) Incentives and opportunism: From the carrot to the stick.
Talk at the TECT Final Conference, 15–17 September 2010, Budapest.

(iii) Equilibrium notions and framing effects.
Talk at the game theory seminar of Prof. Sergiu Hart, December 16th, 2010, Jerusalem.