Exclusion and cooperation in networks

Ule, A.

Citation for published version (APA):
Chapter 6

Dynamics of exclusion and cooperation

Ever since Robert Axelrod initiated his computer tournament aimed at finding the winning strategy in a repeated prisoner's dilemma game (see Axelrod, 1984) have simulations and tournaments been the preferred methodology of social scientists investigating the phenomena of reciprocity and cooperation. Simulations are attractive because they can often be applied when theoretical deduction is not possible. Even simple models of social interaction often produce complex dynamics that are hard to characterize with mathematical analysis. This is especially likely when the system under study has many elements that interact with each other in many different ways. The dynamics of such a system usually depends on the emerging interaction pattern as well as on the private experiences of each agent. Simulations may yield paths of system dynamics given by any set of behavioral rules and thus provide an insight into effects of any particular modelling assumption. Approaching the same task mathematically is often much more difficult.

In this chapter we use computer simulations to study the dynamics of social exclusion among agents that follow simple behavioral rules in a stylized model of a social dilemma embedded in an endogenous network. We identify the conditions under which a norm can emerge that prescribes individual exclusion of players that do not cooperate, and the consequences of this norm for social cooperation. The modelling variables are the inclination to fairness among the players and their comprehension of the behavioral strategies used by other players. Even though our model considers only basic behavioral rules it produces rich dynamics that vary to a great extent across distinct simulation conditions.

In the following section we discuss how an evolutionary view to behavior may give an important alternative approach to that of modelling rational agents in economics. We then shortly describe our dynamic model and discuss some related approaches in the literature, and our contribution. The model is formally introduced and analyzed in section 6.2. Medium and long-run simulations are
described in detail in section 6.4, and their application is demonstrated for the case of selfish players. In section 6.5 we use simulations to study the dynamics of behavior among non-selfish players. In section 6.6 we discuss the results of our simulations and some possible extensions of our dynamic model. We briefly summarize the chapter in section 6.7.

6.1 Introduction

In his seminal tournament Axelrod asked a selection of researchers to submit strategies that would compete against each other in a finitely repeated prisoner’s dilemma game. Even though many sophisticated strategies were submitted the winning one was simple. Called tit-for-tat, it was characterized by cooperating in the first period and subsequently repeating the action that the opponent played in the previous period. The strengths of this strategy were unconditional reciprocation and simplicity, so that other strategies were quickly able to learn to cooperate with it. Reciprocity and simplicity also characterized the winner of a computer simulation of the evolution of strategies in which strategies were not submitted but randomly generated by the computer (see Axelrod, 1987).

Two kinds of evolutionary approaches are most prominent among the numerous simulation studies that have since been implemented by social scientists, computer scientists, mathematicians, economists and others. The first approach is to consider a limited set of strategies, usually a convex hull of those that describe intuitive modes of behavior. The dynamics of proportions of agents playing different strategies is then analyzed in a simulated environment under the assumption that the change in proportion of each strategy is a function of its relative success. In particular, the use of successful strategies is assumed to grow and of unsuccessful ones to decline, although they never quite disappear. This assumption can be motivated by conjecturing that successful strategies are imitated more often than unsuccessful ones, or that agents playing successful strategies have more offspring than those playing unsuccessful ones. See Friedman (1991) and Nowak and Sigmund (1998) for examples of this approach.

The second approach is to set complexity bounds on prospective strategies and specify the rules by which new strategies evolve from previous ones. Strategies may, for example, be structured as cellular automata in which case complexity can be bounded by limiting the maximal number of automata states. Only a small subset of strategies is present in the population at any certain time but new strategies may appear by way of random mutations of the most successful past strategies. This potentially permits a very large set of prospective strategies that is not limited by the intuition of the researcher. Evolutionary simulations of this kind may yield unexpected strategies and are useful exploratory studies but may be difficult to interpret. A number of recent papers using various evolutionary approaches for the study of the emergence of cooperation are surveyed in Axelrod (2000).

The main reason to adopt computer simulations as a methodology of research in economics is to facilitate the study of models that reach beyond the tradi-
tional assumption of fully rational, informed, computationally capable and selfish economic agents. This assumption gives rise to accurate predictions about aggregate market behavior (Friedman, 1953) but it offers an oversimplified view of human behavior and its predictions at the micro-level often dramatically fail. One example is its failure to explain high levels of cooperation in experimental finitely repeated two-player prisoner’s dilemma games. Many other examples can be found in laboratory experiments in game theory (see e.g. Goeree and Holt, 2001). Empirical evidence as well as casual experience yield further phenomena that the standard theory cannot explain, such as altruism, reciprocity, and conformism with social norms (Fehr and Fischbacher, 2003).

The postulates about rationality, perfect foresight and selfishness are contested by the models of bounded rationality. Although people may try to act rationally the requirements of information and computational abilities that standard models assume are rarely met (Simon, 1955). Customers have strong feelings about the perceived fairness of firms and worker’s perception of what constitutes a fair wage constrains the wage setting of firms (Kahneman et al., 1986; Fehr et al., 1993). It is striking that when people have incomplete information or use behavioral rules that are constrained in complexity they may reach far higher levels of cooperation than suggested by standard theory (Kreps et al., 1982; Neyman, 1985).

A common approach to modelling bounded rationality is to consider agents that follow simple behavioral rules. By way of learning and adaptation they occasionally change their heuristics. The updating dynamics may lead to a stable constellation of rules in the population of such agents. Rather than being the consequence of rational deliberation the emerging behavior constitutes the outcome of adaptation of agents to a complex environment. Using this approach one can contribute to an explanation of reciprocity, conformism with social norms, and cooperation. Because of analytical constraints these models are often analyzed with the help of computer simulations.

We introduce in this chapter such a model to study conditions under which a norm that prescribes exclusion of defective players can survive, and the consequences of this norm for the dynamics of cooperation. Social exclusion of defectors may be seen as a form of punishment. It may, however, have additional effects on the evolution of behavior, especially because it affects the social structure and the pattern of future interactions. We therefore discuss our approach in relation to the literature on the evolution of norms and punishment behavior as well as the literature on spatial evolution of cooperation.

That cooperation may be supported by reciprocation was already demonstrated by the success of the tit-for-tat strategy in Axelrod’s tournament (although some evolutionary support for this conjecture already existed in the biological literature, see e.g. Trivers, 1971). This strategy, however, is much less successful in multi-player games and some kind of retaliation that is directed solely at non-cooperators is needed to sustain cooperation. Boyd and Richardson (1992, 1988) refer to such directed retaliation as retribution and study its evolutionary success. Retribution, if costly, can be seen as an altruistic act. Punishment is beneficial to the group, but costly to the individual, and
selection favors individuals that do not punish. This is commonly referred to as the "second order" free riding problem (Oliver, 1980). Several solutions to this problem have been suggested. One may, for example, consider punishers who punish any order free riders, thus following a metanorm (Axelrod, 1986), cooperators who always retaliate (Yamagishi and Hayashi, 1994), higher order punishments (Henrich and Boyd, 2001), and inter-group competition (Bowles and Gintis, 2003).

The structure of social interaction, on the other hand, may also crucially affect the evolution of cooperation. This has first been demonstrated by Nowak and May (1992). Large number of models of adaptive play on different interaction structures, in what became known as models of local interaction, have since been analyzed using simulations (Eshel et al., 1998a; Cohen et al., 2001), tournaments (Burt, 1999) and theoretical analysis (Eshel et al., 1998b; Outkin, 2003; Tieman et al., 2000). A particularly thorough investigation is reported by Cohen et al. (1998) who use simulations to compare dynamics of cooperation across different structures of interaction, adaptive processes and strategy spaces.

A common assumption in models of the evolution of retaliation and local interaction is that successful strategies grow by way of imitation or conformism. In particular, best response dynamics are not considered. Eshel et al. (1998b) claim that "altruism has no hope in the world of best responders", as defection is the unique best response. This, however, is true only in the world of myopic best responders, that is, players who look for strategies that maximize only their immediate payoff. We demonstrate in this chapter that to sustain cooperation under best response dynamics it is sufficient to assume that players understand the reactions of other players and take them into account when calculating their payoff.

Players that are sufficiently sophisticated to include their knowledge about strategies used by other players when calculating their best response are usually referred to as level-2 players (see Stahl and Wilson, 1995), as opposed to level-1 players who play best response to previous period actions and level-0 players who play randomly. That many people exhibit what is called strategic sophistication was demonstrated in experiments by Stahl and Wilson (1994) and Costa-Gomez et al. (2001) who classify between 45% and 53% of their subjects as level-2 players. Such players may have significant impact on the dynamics of play, but have so far been largely neglected in dynamic models of adaptive behavior. Only the simplest modes of behavior, such as imitation or myopic best response, or most sophisticated ones, such as perfect foresight, are usually considered. We propose in this chapter a simple way of modelling level-2 players via limited forward looking.

Players with forward looking of limited scope $r \geq 0$ choose strategies that maximize their utility across $r$ periods. Players with $r = 2$ consider reactions of other players to their own behavior in their utility calculations in accordance with level-2 sophistication. Players with $r = 1$ choose myopic best response in accordance with level-1 sophistication. Players with $r = 0$ do not maximize their future payoff but make random actions, in accordance with level-0 sophistication. To study the effect of increased forward looking we also study players
with $r = 4$, although not with level-4 sophistication. A suggestion on how higher strategic sophistication may be simulated is given in section 6.6.2. As far as we know models of players with limited forward looking and limited, but not trivial, strategic sophistication have not been studied before. A reference to a few possible models of limited foresight, with a short discussion of their shortcomings, is given by Rubinstein (1998). Jehiel (1995) proposes an intricate notion of an equilibrium with limited foresight, which we discuss in section 6.6.2.

Consideration of best response dynamics permits us to make another contribution, by modelling the dynamics of behavior among players with non-selfish preferences. That people may have sentiments for the well-being of other people, and may not necessarily be selfish, has been debated by philosophers as well as economists (a historical reference is Adam Smith, e.g. Smith, 1975, see also Binmore, 2000). It has been observed empirically and demonstrated via laboratory experiments (see e.g. the dictator game experiments by Eckel and Grossman, 1996). Recently a number of approaches to model non-selfish players have been suggested (e.g. Rabin, 1993; Levine, 1998; Fehr and Schmidt, 1999; Cox et al., 2004). In this chapter we consider two simple models from this literature, the model of altruism by Levine (1998) and the model of inequality aversion by Fehr and Schmidt (1999), and study the dynamics of norms and cooperation among altruistic or inequity averse players. Because these models require comparisons of prospective payoffs they are incompatible with assumptions of imitation or conformism made in other evolutionary studies. In our model, in which we assume best-response behavior, they can be implemented in a straightforward manner.

Best response has previously been used in the framework of the dynamics of conventions. Ellison (1993) and Blume (1993) both studied the evolution of coordination among myopic best responders embedded in a local interaction network. They noted that non-perturbed dynamics may converge to several equilibria, but if the process is perturbed by small errors the risk dominant equilibrium is most likely to be selected. Their approach is based on observations made by Fudenberg and Maskin (1990) and Young and Foster (1991) that the dynamics in games with noise may drastically differ from that in games without noise. Small perturbations may result in a process that, in the long run, spends considerably more time in some states than in the others. In this sense we may say that the dynamics select among the equilibria in the long run. We adopt the approach of Young and Foster (1991) to detect the long-run equilibria in our model.

This is how we set up our model. We assume that a relatively small group of players is playing the prisoner’s dilemma game embedded in an endogenous network. In each period players simultaneously choose with whom to play and what action to play. Each pair of players play the prisoner’s dilemma game when both mutually chose each other, and earn the outside option otherwise. Our aim is to study the dynamics of the norm that prescribes cooperation and exclusion of players that defected in the previous period. We conjecture that full cooperation will emerge if sufficient number of players follows such a norm. Whether such a situation can be sustained in the long run may depend on forward looking and
social preferences of the players as well as on the outside option value. We run computer simulation across different scopes of forward looking, different outside options, and different constellations of social preferences in the group. Our main conclusions are that (i) sufficient forward looking is as important for sustaining cooperation as is sufficient regard for others, (ii) cooperation may be sustained among sufficiently forward looking selfish players, and (iii) even though assuming other-regarding preferences increases the set of cooperative equilibria, there are no qualitative differences in the long-run dynamics between groups of other-regarding and groups of selfish players. Cooperation increases considerably only when extremely high altruism or inequality aversion is assumed.

A few other studies exist that discuss the evolution of cooperation when one’s interaction partners may be endogenously chosen or refused. Schuessler (1989) and Vanberg and Congleton (1992) explore the replicator dynamics within a small set of simple strategies in repeated two person prisoner’s dilemma game with options of exit or partner change. Batali and Kitcher (1995) and Hauert et al. (2002) study the evolution of altruism in Public Goods games with exit possibility. Morikawa et al. (1995) and Orbell et al. (1996) look at the evolutionary success of optimistic cooperators who often play the game, facing pessimistic defectors who are more likely to take the outside option. Mailath et al. (2001), Skyrms and Pemantle (2000) and Outkin (2003) describe probabilistic models in which players are matched according to endogenously chosen probability distribution. Flache (2001) studies the effects of risk preferences on migration and cooperation in a model with heterogeneous players. Vega-Redondo (2002) describes a model of network formation with exogenously defined linking rules. Yamagishi et al. (1994) and Hayashi and Yamagishi (1998) use computer tournaments to study the success of various strategies in a finitely repeated prisoner’s dilemma with endogenous matching. Smucker et al. (1994) and Ashlock et al. (1996) use genetic algorithms to study the evolution of cellular automata playing prisoner’s dilemma in a partner choice and refusal environment and Hauk (2001) and Hanaki et al. (2004) consider a similar environment to study adaptive dynamics within a small set of intuitive strategies.

All of these studies assume that behavior spreads via imitation or a related adaptive model and that players are constrained in the number of partners they may choose. In chapter 4 we show that if linking is constrained then cooperation can be achieved in subgame-perfect equilibria of finitely repeated network dilemma games even among rational and selfish players with perfect foresight. In contrast, we study in this chapter a repeated network dilemma game with costless and unconstrained linking. Our intuition that boundedly rational players may achieve cooperation in such setting is confirmed in our dynamic model.
6.2 Adaptive play

In this chapter we consider network dilemma games with unconstrained and
costless linking, introduced in chapter 3. We refer to such games as basic games.

Given a basic game \( \Gamma = \langle N, J, \pi \rangle \) with \( n \) players, the following dynamic
model is considered. Let \( t = 0, 1, 2, \ldots \) denote successive time periods. The
basic game is played once each period. Let \( a_i(t) \) be the action and let \( p_i(t) \)
be the linking choice of player \( i \) at time \( t \). The profile of moves at time \( t \) is
\( (a(t), p(t)) = ((a_1(t), \ldots, a_n(t)), (p_1(t), \ldots, p_n(t))) \in J \). The history of moves up
to time \( t \) is the sequence \( h(t) = (a(0), p(0), a(1), p(1), \ldots, a(t), p(t)) \). The set
\( H[t] = J^t \) consists of all possible time \( t \) histories.

At each time \( t \) player \( i \) is characterized by a strategy, defined as a mapping
\( s_i[t] : H[t-1] \rightarrow A_i \times P_i \)
assigning a move \( (a_i(t), p_i(t)) = s_i[t](h(t-1)) \) to each possible history \( h(t-1) \).
For convenience we omit the time parameter \( [t] \) whenever we refer to generic
moves, strategies, and histories. We denote by \( p_i(s_i \mid h) \) the linking choices and
by \( a_i(s_i \mid h) \) the action chosen by the strategy \( s_i \) given history \( h \).

We assume that players exhibit a fair amount of inertia in choice of their
strategies. With high probability they keep the same strategy across several
periods. Only occasionally a player realizes what strategies are played by the
other players and updates her own strategy by choosing one that maximizes her
utility given the strategies of the other players. An updating player assumes
that the other players will continue to play the same strategies for a few more
periods, and that she herself will be playing her new strategy for a few following
periods. These assumptions may be viewed as a boundedly rational behavior
of the players facing the difficult task of predicting the future behavior of the
other players.

When maximizing her future payoffs the player may consider only the payoffs
in the coming period, in which case we say that player chooses a myopic best
response. A standard result is that defection is the unique myopic best response
in a multi player prisoner's dilemma game. We show here that in our basic
game a player following the myopic best response never cooperates over an
established link. The aim of this chapter is to extend the analysis by considering
players which, to a limited extent, understand the future consequences of their
actions. This is modelled by assuming that players behave in order to maximize
their payoff across several periods of play. We discuss when cooperation among
players with such "limited forward looking" can be sustained.

6.2.1 Strategy set

Below we introduce the set of strategies considered in the simulations. Rather
than allowing for all feasible strategies at each time \( t \) we consider a fixed set
of possible strategies. The strategies are independent of the time period. This
is because we assume that players have a limited memory and consider only
the most recent part of history, even though the length of history increases with
time. In particular, we assume for simplicity that the strategies are independent of all past periods except the most recent one. We also assume that strategies depend only on past actions, but are independent of past linking choices.

Further, we assume action consistency and linking consistency. Action consistency requires that strategies are either fully cooperative or fully defective. That is, a player cooperates in each period she is playing a cooperative strategy and defects in each period she is playing a defective strategy. Linking consistency requires that a player proposes a link to another player playing strategy $s$ if and only if she also proposes links with all other players playing strategy $s$. We do not consider mixed strategies.

The resulting set consists of four intuitively plausible strategies. The benefit of a small set of feasible strategies is the tractability and clarity of interpretation of the simulation dynamics.

We assume that each player is of one of the following types:

(CA) (Non-exclusive) cooperator. If player $i$ is of this type she proposes links to all other players and cooperates in the prisoner’s dilemma game. For any time $t$ the corresponding strategy $s^C_A(t) = (C, p^C_A)$ is defined by

$$s^C_A(t-1) = (C, p^C_A),$$

and the vector of proposed links $p^C_A$ by

$$p^C_A = \begin{cases} 1 & \text{if } j \neq i \\ 0 & \text{if } j = i \end{cases}.$$

(CE) Exclusive cooperator. If player $i$ is of this type she proposes links only to those other players that cooperated in the previous period, and cooperates in the prisoner’s dilemma game. For any time $t$ the corresponding strategy $s^C_E(t) = (C, p^C_E)$ is defined by

$$s^C_E(t-1) = (C, p^C_E),$$

and the vector of proposed links $p^C_E$ by

$$p^C_E = \begin{cases} 1 & \text{if } a_j(t-1) = C \\ 0 & \text{if } a_j(t-1) = D \text{ or } i = j \end{cases}.$$

(DA) (Non-exclusive) defector. If player $i$ is of this type she proposes links to all other players and defects in the prisoner’s dilemma game. For any time $t$ the corresponding strategy $s^D_A(t) = (D, p^D_A)$ is defined by

$$s^D_A(t-1) = (D, p^D_A),$$

and the vector of proposed links $p^D_A$ by

$$p^D_A = \begin{cases} 1 & \text{if } j \neq i \\ 0 & \text{if } j = i \end{cases}.$$
always link   exclude defectors

cooperate

<table>
<thead>
<tr>
<th></th>
<th>$s^{CA}$</th>
<th>$s^{CE}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>defect</td>
<td>$s^{DA}$</td>
<td>$s^{DE}$</td>
</tr>
</tbody>
</table>

Table 6.1: Classification of strategies according to cooperation and exclusiveness.

(DE) Exclusive defector. If player $i$ is of this type she proposes links only to those other players that cooperated in the past period, but herself defects in the prisoner’s dilemma game. For any time $t$ the corresponding strategy $s^{DE}_i$ is defined by

$$s^{DE}_i(h(t-1)) = (D, p^{DE}_i),$$

and the vector of proposed links $p^{DE}_i$ by

$$p^{DE}_{ij} = \begin{cases} 1 & \text{if } a_j(t-1) = C \\ 0 & \text{if } a_j(t-1) = D \text{ or } i = j \end{cases}.$$

We refer to a player of type CA or CE as cooperative, and to a player of type DA or DE as defective. We refer to a player of type CE or DE as exclusive, and to a player of type CA or DA as non-exclusive. In Table 6.1 we use the same nominations to classify the corresponding strategies. The non-exclusive strategies are independent of the history, while the exclusive strategies depend on the most recent actions only. Let $S_i = \{s^{CA}_i, s^{CE}_i, s^{DA}_i, s^{DE}_i\}$ and let $S = \{(s_1, ..., s_n) | s_i \in S_i\}$ be the space of strategy vectors.

We are interested in the long run dynamics of norms prescribing exclusion of defectors, and the impact of the exclusion in the long run success of cooperation. The four strategies defined above represent the modal types of behavior for our analysis. A notable omission is the absence of tit-for-tat type strategies, that is, strategies which prescribe cooperation as long as certain number of other players cooperate, and defection otherwise. In $n$-player prisoner’s dilemma games defection of a single player negatively affects all other players, those who defect as well as those who cooperate. Even if the intent of defection is to punish a subgroup of defectors it may induce only more defections by the remaining cooperators, and the play may slide into the uniform defection. Exclusion of defectors, on the other hand, is a form of retaliation and does not risk being misinterpreted for defection by the remaining players. Rather than tit-for-tat strategy, we study in this chapter the success of the out-for-tat strategy $s^{CE}_i$.

6.2.2 Updating rule

We describe next the strategy updating rule. Players use the same rule in all periods in which they update their strategy. For convenience we omit the time index $"[t]"$ whenever it is not essential. Let $a$ be the vector of actions at time $t-1$ and let $s = (s_1, ..., s_n) \in S$ be the vector of strategies that players use at
time $t$. The profile of moves at time $t$ is then determined by $s$ and $a$. Let
$a_i(s_i) \equiv a_i(s_i | h(t-1))$ be the action, and let $p_i(s_i | a) \equiv p_i(s_i | h(t-1))$ be the
linking choice induced at time $t$ by strategy $s_i$, given actions $a$ at time $t-1$. Let
$a_i(s) = (a_1(s_1), ..., a_{i-1}(s_{i-1}), a_i(s_i), ..., a_n(s_n))$ be the vector of actions
and let $p(s | a) = (p_1(s_1 | a), ..., p_{i-1}(s_{i-1} | a), p_i(s_i | a), ..., p_n(s_n | a))$ be
the vector of linking choices of players $N/i$ induced by $s$. The vector of strategies
that players use at time $t$ is denoted by $s[t] = (s_1[t], ..., s_n[t])$.

Players begin with an initial vector of strategies $s[1] \in S$ at time $t = 1$. At
each time $t \geq 2$, before making her move, player $i$ with some fixed probability
$q_i \in (0, 1)$ updates her own strategy, by observing the vector of strategies at
time $t - 1$ and maximizing her payoff across the next $r_i$ periods, under the
assumption that the strategies of the other players do not change in that time
span. Formally, let $s = (s_1, ..., s_n) \in S$ be the vector of strategies at time $t - 1,$ and let

$$s' = (s_1, ..., s_{i-1}, s_i', s_{i+1}, ..., s_n)$$

be the vector of strategies during time periods $[t, t + r_i]$ assuming that player
$i$ changes her strategy from $s_i$ to $s_i'$ at $t$, keeps it during $[t, t + r_i]$, and that
players $N/i$ do not change their strategies during $[t - 1, t + r_i]$. Under these
assumptions, the payoff of player $i$ across the $r_i$ periods following her update is

$$x_i(s_i' | s) = \pi_i(a_i(s_i'), p_i(s_i' | a(s)) , a_{-i}(s), p_{-i}(s | a(s)))$$

$$+ (r_i - 1) \cdot \pi_i(a_i(s_i'), p_i(s_i' | a(s')) , a_{-i}(s'), p_{-i}(s' | a(s'))). \quad (6.1)$$

The first row of equation (6.1) is the payoff player $i$ receives at time $t$, that is,
when the remaining players believe she is playing the old strategy $s_i = s_i[t - 1],$
but in fact she is playing the new strategy $s_i'$. The second row of equation (6.1)
is the payoff player $i$ receives during the subsequent $r_i - 1$ time periods, that
is, when the remaining players realize she is playing the new strategy $s_i'$. This
calculation is correct if no updates take place during $[t - 1, t + r_i]$, aside from
the initial update by player $i$.

If player $i$ is updating at time $t$ she chooses a strategy $s_i[t]$ such that

$$s_i[t] \in \arg\max_{s_i' \in \{s_{C'A}, s_{C'E}, s_{D'A}, s_{D'E}\}} x_i(s_i' | s[t - 1]). \quad (6.2)$$

If this payoff is maximized by the current strategy of the player, the player keeps
her current strategy. If not, she randomly selects one of the optimal strategies.
The parameter $r_i$ is called the scope of forward looking of player $i$. If player $i$
does not update her strategy, which happens with probability $(1 - q_i)$, we say
that she exhibits 'inertia', i.e. she maintains her current strategy.

The assumption that strategies of other players can be observed is not very
restrictive. To see this, note that (i) actions of all players are always observable

---

1In our simulations we will consider different initial strategy profiles. We specify them later.
and (ii) whenever at least two players defect the linking behavior of all players in the next period reveals their type. This means that whenever at least two players defect strategies of all players are revealed in the following period. For simplicity we assume that a player observes all strategies whenever she is updating her strategy.

6.3 The basic process

When at some time $t \geq 2$ a player updates her strategy her choice depends only on the vector of strategies played a time $t - 1$. The probability that players will play strategies $s[t]$ at time $t$ depends therefore only on the vector of strategies $s[t - 1]$ played at time $t - 1$, and not on the strategies played during the previous time periods. Furthermore this probability is stationary, that is, it is independent of $t$. The updating rule outlined above thus generates a Markov chain with the finite state space $S$. This Markov chain is completely characterized by its transition matrix $M$ in which for each $s, s' \in S$ the entry

$$ M_{ss'} = \text{prob}[s[t] = s' \mid s[t - 1] = s] $$

(6.3)

gives the probability that the process is in state $s'$ at time $t$, given that it was in state $s$ at time $t - 1$. The value of $M_{ss'}$ is also called the probability of transition from state $s$ into state $s'$. The value of $M_{ss}$ is the probability that the process will not make any transition from state $s$. For each state $s$ the corresponding row of the matrix represents the probability distribution over transitions from state $s$, and satisfies

$$ \sum_{s' \in S} M_{ss'} = 1. $$

(6.4)

We refer to this particular discrete-time Markov process as the basic process.

6.3.1 Absorbing states of the basic process

Let $s = (s_1, \ldots, s_n)$ be an absorbing state. By definition the set $\{s\}$ is recurrent, which implies that $M_{ss'}^\tau = 0$ for every state $s' \neq s$ and all $\tau \in \mathbb{N}$. It then follows from (6.4) that $M_{ss} = 1$, from (6.3) that $s[t] = s$ implies $s[t + 1] = s$, and, indeed, that $s[t + \tau] = s$ for all $\tau \in \mathbb{N}$. This may hold only when

$$ s_i \in \arg \max_{s'_i \in \{s^{CA}, s^{CE}, s^{DA}, s^{DE}\}} x(s'_i \mid s), $$

(6.5)

If no player defects then in the following period all players propose all links and do not reveal whether they are exclusive or not. If only one player defects then this player, regardless of her type, proposes all links in the following period.

We assume that time starts with $t = 1$, but the first update may take place only after the first round, that is at $t = 2$. This is because players need to observe at least one period of play before they can update their strategies.

The definition (6.3) requires $t \geq 2$. Because of the stationarity of the process this probability is well defined, that is, it is independent of $t$. 

163
for each player $i$. It is easy to verify that the opposite argument is also true, which proves the following Proposition.

**Proposition 6.1** State $s$ is absorbing if and only if (6.5) holds for each player $i$.

In other words, in an absorbing state, each player is playing a strategy that is a best response with limited forward looking to the strategies played by the other players. In this sense we may say that the process is in equilibrium when it reaches an absorbing state.

In each state the set of possible transitions depends only on the numbers of players of each type. This is shown in the proof of Proposition 6.2, below. An intuitive explanation is that the strategies do not discriminate among the players on the basis of their identities. If players $j$ and $k$ play the same strategy at time $t-1$, then each player $i \neq j, k$, using one of the strategies outlined above, will react to players $j$ and $k$ in the same way at time $t$. Formally, for any type $T \in \{CA, CE, DA, DE\}$ let

$$n^T(s) = |\{i \in N | s_i = s_i^T\}|$$

be the number of players of type $T$ given the vector of strategies $s$.\(^5\) For each $s \in S$ we refer to the 4-tuple

$$z(s) = (n^{CA}(s), n^{CE}(s), n^{DA}(s), n^{DE}(s))$$

as the *type distribution* in state $s$. Let

$$Z = \{(n^{CA}, n^{CE}, n^{DA}, n^{DE}) |$$

$$n^{CA} + n^{CE} + n^{DA} + n^{DE} = n, n^T \geq 0 \text{ for each type } T\}$$

be a set of possible type distributions. For each type distribution $z \in Z$ let

$$S_z = \{s \in S | z(s) = z\}$$

be the set of all states with type distribution $z$. Sets $S_z$ partition $S$ into groups of states with equal type distribution.

Let us introduce a short notation for the states in which all players play the same strategy. For a type $T \in \{CA, CE, DA, DE\}$ let $s^T = (s_1^T, ..., s_n^T)$ be the state in which each player $i$ plays the strategy $s_i^T$. The corresponding type distributions are

- (CA) $z(s^{CA}) = (n, 0, 0, 0),$
- (DA) $z(s^{CE}) = (0, n, 0, 0),$
- (CE) $z(s^{DA}) = (0, 0, n, 0),$  
- (DE) $z(s^{DE}) = (0, 0, 0, n).$

\(^5\)Cardinality $|A|$ gives the number of elements in the set $A$. 

164
The following proposition states that for each \( z \in Z \) either all or none of the states with type distribution \( z \) are absorbing. The proof is given in the appendix to this chapter.

**Proposition 6.2** Let \( z \in Z \) be a type distribution. State \( s \in S_z \) is absorbing if and only if each of the states in \( S_z \) is absorbing. State \( s \in S_z \) is transient if and only if each of the states in \( S_z \) is transient.

Propositions 6.1 and 6.2 facilitate characterization of absorbing states of the basic process. To prove that states with type distribution \( z \) are absorbing it suffices to show that in some state \( s \in S_z \) strategies of all players satisfy condition (6.5). To prove that states with type distribution \( z \) are not absorbing it suffices to show that the strategy of some player in some state \( s \in S_z \) violates condition (6.5), that is, her strategy is not a best response with limited forward looking to strategies played by the other players.

A non-absorbing state is either transient or belongs to a recurrent set. If the absorbing states are known, characterization of recurrent sets reduces to the characterization of the set of transient states. For this we use the following result.

**Proposition 6.3** A state \( s \) is transient if some transient or absorbing state is accessible from \( s \). In particular, a state \( s \) is transient if \( M_{ss'} > 0 \) for some transient or absorbing state \( s' \neq s \).

**Proof.** By definition, state \( s \) belongs to a recurrent set if all states accessible from \( s \) also belong to the same recurrent set. Also by definition, a transient or absorbing state never belongs to a (non-singleton) recurrent set. If a transient or absorbing state is accessible from \( s \), this implies that \( s \) does not belong to any recurrent set, and is therefore transient. \( \blacksquare \)

We implement the assertion of Proposition 6.3 to detect the transient states of the basic process via the following algorithm.

**Algorithm 1**

1. Calculate the transition matrix \( M \).
2. Determine the absorbing states \( S^A \) of the process.
3. Let \( S^0 \) be the set of states \( s \in S^0 \) such that \( M_{ss'} > 0 \) for some absorbing state \( s' \neq s \).
4. For \( \tau = 1, 2, ..., \tau \), repeat:
   - let \( S^\tau \) be the set of states \( s \in S^\tau \) such that \( M_{ss'} > 0 \) for some \( s' \in S^{\tau-1} \), until \( S^\tau = S^{\tau-1} \).

**Proposition 6.4** Let, for \( \tau = 0, 1, ..., \tau \), the sets \( S^\tau \) be yielded by the Algorithm 1. The process has no recurrent sets, aside from the absorbing states, if and only if \( S^0 \cup S^\tau = S \).
Proof. An iterative application of Proposition 6.3 implies that for each \( \tau \) all states in \( S^\tau \) are transient. By the definition of the updating process, regardless of the state, the event that no player is updating has positive probability. This is to say, \( M_{ss} > 0 \) for each state \( s \). Consequently, \( S^{\tau-1} \subseteq S^\tau \) for each \( \tau \geq 2 \). The algorithm thus makes at most \( |S| \) iterations. Since the set \( S \) is finite, the algorithm stops in finite time.

Clearly, if \( S^0 \cup S^\tau = S \) there are no recurrent sets aside from the absorbing states in \( S^0 \). To prove the opposite argument, assume that there are no recurrent sets aside from the absorbing states in \( S^0 \). At least one such absorbing state must be accessible from each state \( s \). Therefore, for some \( \tau \leq \tau \) and some absorbing state \( s' \), \( M_{ss'}^\tau > 0 \), which implies that \( s \in S^\tau \), and thus \( s \in S^\tau \).

As it turns out every implementation of Algorithm 1 in this chapter yields the sets of absorbing states \( S^0 \) and of transient states \( S^\tau \), which satisfy \( S^0 \cup S^\tau = S \). Using proposition 6.4 we conclude that, across all the processes considered, the recurrent sets consist solely of absorbing states.

Throughout this chapter we assume that \( q_i^c = q \) and \( r_i = r \) for all \( i \). For reference we include the payoff matrix of the prisoner's dilemma below.

<table>
<thead>
<tr>
<th>Player ( j )</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Player ( i )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>( c, c )</td>
<td>( e, f )</td>
</tr>
<tr>
<td>D</td>
<td>( f, e )</td>
<td>( d, d )</td>
</tr>
</tbody>
</table>

Without the loss of generality we may assume that all payoffs in the prisoner’s dilemma game are positive, \( f > c > d > e > 0 \). Using Proposition 6.2 it is straightforward to verify that the state \( s^{CE} \), in which all links are established, all players cooperate, and threaten to exclude any defective player, is an absorbing state of the basic process whenever

\[
rc \geq f + (r - 1)o, \tag{6.6}
\]

i.e. whenever the payoff from \( r \) periods of mutual cooperation exceeds the payoff from one period of free-riding followed by \( r - 1 \) periods of isolation. The complete characterization of absorbing states is, however, less straightforward. Below we analytically characterize absorbing states of the basic process with \( o = 0 \). In general, however, we use Algorithm 1 to characterize absorbing states of other processes considered in this chapter.

Let the value of the outside option be \( o = 0 \). For each player \( i \) the choice profile \( s_i^{DA} \) constitutes the unique dominant strategy in the basic game. Similarly, with any scope of forward looking \( r \) the state \( s_i^{DA} \), in which all links are established and all players defect, is an absorbing state of the basic process. For \( r = 1 \) absorbing states consist only of defective players. Proposition 6.5 below states that there might, however, exist absorbing states with cooperative players, when their scope of forward looking is sufficiently long. Note that in any state with type distribution \((p, n-p, 0, 0)\), such that \( 0 \leq p \leq n \), all \( n \) players are cooperative and \( n - p \) are exclusive.
Proposition 6.5 Consider the basic process with $o = 0$. State $s^{DA}$ is absorbing for all $r$. State $s^{DE}$ is absorbing for $r = 1$. States $s$ with type distribution $z(s) = (n - 1, 0, 0, 1)$ are absorbing if and only if

$$c > f - \frac{e}{n - 2} + \frac{d}{r(n - 2)}. \quad (6.7)$$

States $s$ with type distribution $z(s) = (p, n - p, 0, 0)$ are absorbing if and only if

$$c > \left(1 - \frac{(n - p - 1)(r - 1)}{(n - 1)r}\right)f. \quad (6.8)$$

In particular, state $s^{DE}$ is absorbing if and only if $r$ satisfies (6.6). There are no other absorbing states.

The proof is given in the appendix to this chapter. Several general implications may be drawn from Proposition 6.5. Condition (6.6) with $o = 0$ is necessary for (6.8) and thus for the existence of any fully cooperative absorbing state. Condition

$$c \geq f - \frac{e}{n - 2} \quad (6.9)$$

is necessary for the existence of absorbing states with type distribution $(n - 1, 0, 0, 1)$. There is always an $r$ such that (6.6) is satisfied, but this is not the case with (6.9). To summarize, when their scope of foresight is short (e.g. $r = 1$, myopic best response) all players are defecting in an absorbing state. Only when players have a sufficiently long scope of foresight do there exist absorbing states in which all players cooperate.

6.3.2 Stability of absorbing states

The dynamics of the process may crucially depend on the initial conditions. In general, if a Markov chain starts in an absorbing state it will never leave that state. On the other hand, if it starts in a non-absorbing state it will leave that state and, in the long run, reach a recurrent set (or an absorbing state). The probability that the Markov chain is in a state that does not belong to a recurrent set is thus asymptotically zero. In this sense only the states in recurrent sets can be selected by the process in the long run.

It turns out that our basic process may have several absorbing states but no other recurrent sets. The basic process reaches, with probability one, one of its absorbing states in a finite time and remains in that state forever. To get an indication about which of the absorbing states are most likely to be reached we introduce the notion of their stability. In short, some absorbing states may be more persistent than others, when the process is subject to small perturbations. If the players occasionally violate the updating rule, e.g. due to mistakes or experimentation, the process is most likely to spend most periods in an absorbing state that is robust against these violations.
This notion of stability has been formalized in concepts of stochastic stability and the long-run equilibrium by Young (1993) and Kandori et al. (1993), respectively. Both concepts have been invoked as solution criteria in evolutionary models with best response dynamics, e.g. in Samuelson (1994) and Kandori and Rob (1995). Here, we use stochastic stability as the selection mechanism among the absorbing states of the basic process. An introduction to the concept of stochastic stability and discussion on the relation with concepts of asymptotic stability and evolutionary stability may be found in Foster and Young (1990). Essentially, stochastic stability requires robustness against persistent or regular mistakes, while evolutionary stability requires only robustness against isolated mutations. Young (1998) argues in favor of its implementation in the studies of social phenomena and provides a series of applications of the concept.

Consider again the Markov chain defined by the best response dynamics introduced in section 6.2. Assume also that the players occasionally make mistakes in the choice of their strategy. Let \( \varepsilon \in (0,1) \) be the probability that an updating player chooses her new strategy at random. The player otherwise follows the best response rule (6.2). This modified updating rule generates a perturbed Markov process with state space \( S \) and transition matrix \( M^\varepsilon \). Because of random error there is a positive probability \( M^\varepsilon_{ss'} > 0 \) of a transition between any pair of states \( s, s' \in S \), which implies that all entries in the transition matrix \( M^\varepsilon \) are positive. Proposition 2.1 describes a number of properties of a process with such transition matrix. In particular, the perturbed process is ergodic for each \( \varepsilon > 0 \), and has a unique stationary distribution \( \mu^\varepsilon(s) \). This distribution fully characterizes the asymptotic properties of the process: independently of the initial state \( s^0 \), the asymptotic proportion of time that the state spends in state \( s \) is given by \( \mu^\varepsilon(s) \), and the probability that the process is in state \( s \) at time \( t \) approaches \( \mu^\varepsilon(s) \) as \( t \) grows large.

**Definition 6.1** A state \( s \) is stochastically stable if

\[
\lim_{\varepsilon \to 0} \mu^\varepsilon(s) > 0.
\]

The following Theorem is due to Young (1993).

**Theorem 6.1** For each \( s \in S \), \( \mu^0(s) = \lim_{\varepsilon \to 0} \mu^\varepsilon(s) \) exists and the limiting distribution \( \mu^0 \) is a stationary distribution of the non-perturbed process.

It follows, in particular, that the basic process has at least one stochastically stable state. Intuitively, these are the states that are most likely to be observed over the long run when random mistakes are rare. The distribution \( \mu^0 \) is one of solutions of the system (2.2). The stochastically stable sets therefore form a union of recurrent sets. However, it may not put positive probability on all recurrent sets. For the basic process this means that no other states than some of the absorbing states are stochastically stable.
Table 6.2: The parameters considered in the simulations of the basic dynamic process: (a) the prisoner’s dilemma game, and (b) the updating rule.

6.4 Simulations of the basic process

To study dynamics of the process we employ two classes of computer simulations. Simulations of the first class are an implementation of Algorithm 1 and designed to detect the direction of transitions between the states of the Markov chain. Simulations of the second class are designed to study the asymptotic properties of the process, such as the stabilities of different absorbing states. The first class considers the non-perturbed basic process, while the second class considers its perturbed form. We consider two basic games that coincide with the experimental games $T_l$ and $T_h$ in chapter 5. The corresponding prisoner’s dilemma game and the updating parameters considered in the simulations are shown in Table 6.2. We provide the details of each class of simulations, and demonstrate their application, for the basic process with low outside option $o = 0$. We proceed by using the simulations to detect the dynamics and asymptotic properties of the basic process with high outside option $o = 40$.

6.4.1 Low outside option, $o = 0$

Consider $o = 0$. The parameters given in Table 6.2 do not comply with (6.9), but do comply with (6.6) for any $r \geq 2$. That is, there is no $r \in \mathbb{N}$ such that states with type distribution $(n-1,0,0,1)$ are absorbing, while fully cooperative absorbing states, such as $s^{CE}$, do exist for $r \geq 2$. According to Proposition 6.5 the absorbing states of the basic process, under these parameters, are

- $r = 1$ : $s^{DA}$ and $s^{DE}$,
- $r \in \{2,3\}$ : $s^{DA}$, $s^{CE}$, and $s$ such that $z(s) \in \{(1,5,0,0),(2,4,0,0)\}$,
- $r \geq 4$ : $s^{DA}$, $s^{CE}$, and $s$ such that $z(s) \in \{(1,5,0,0),(2,4,0,0),(3,3,0,0)\}$.

Below we discuss the dynamics of the basic process with scope of forward looking of 1, 2, or 4. These values of $r$ were chosen because they are the smallest values representing all different sets of absorbing states.

6.4.2 Detecting dynamics of the process

The direct application of Algorithm 1 requires calculation of the transition matrix $M$ of the non-perturbed process. In the worst case it will have $|S|$ iterations.
With \( n = 6 \) and with 4 possible types the state space \( S \) consists of \( 4^6 = 4096 \) states and the transition matrix \( M \) of \( 4096^2 > 16,000,000 \) entries. Algorithm 1 would in the worst case require 4096 operations on the matrix \( M \).

A simplification of Algorithm 1 is possible, however. According to Proposition 6.2 it is sufficient to classify one of the states with type distribution \( z \in Z \) as absorbing (transient) to prove that every state with type distribution \( z \) is absorbing (transient). Further, to prove that a state \( s \in S \) is transient it is sufficient to simulate the updating process, initiated with \( s[1] = s \), until an absorbing state or another transient state is reached. If, on the other hand, the process makes no transitions for many periods from the state \( s \), it is likely to be absorbing.

We employ these observations and use computer simulations for detection of the absorbing states and of the direction of transitions from the transient states. For each possible initial distribution \( z \in Z \) we choose a state \( s_z \in S_z \) and simulate 100 iterations of the non-perturbed basic process initiated with \( s[1] = s_z \). Additionally, for each \( s_z \) we repeat the simulation 100 times.

If no transition from the initial state is observed during the 100 periods then we classify it as absorbing. The probability of misclassification is below \( 10^{-4} \). To see this, note that in a non-absorbing state at least one player would change her type when updating. For the process to remain in the non-absorbing state, such player should never update. The probability that a player never updates during 100 periods is \((1 - q)^{100} = 0.9^{100} < 10^{-4}\). Further, by repeating each such simulation 100 times, we assure observations of many transitions from every non-absorbing state. The probability that we observe less than 100 such transitions is \((1 - 10^{-4})^{100} < 0.01\).

In the course of simulations we progressively classify the states as follows. We begin by classification of absorbing states. We classify any remaining state as transient whenever the process initiated in that state reaches an absorbing state. We proceed by iteration and classify any state as transient whenever the process initiated in that state reaches another transient state. If, after a number of such iterations, all states are classified as either absorbing or transient we conclude that there are no recurrent sets other than absorbing states.

By following the transitions of the processes across all initial states \( s_z \) we can also get an impression of the dynamics of cooperation and exclusion in the basic process. To turn this impression into a two-dimensional graphical presentation we must aggregate the information about the observed transitions. For each state \( s \in S \) let \( y_1(s) = n^CA(s) + n^CE(s) \) be the number of cooperative players and let \( y_2(s) = n^CE(s) + n^DE(s) \) be the number of exclusive players in the state \( s \). Each realization of the basic process \( s[t] \) implies a realization of the truncated process

\[
y[t] = (y_1(s[t]), y_2(s[t]))
\]

representing the corresponding evolution of cooperation and exclusion. The process \( y[t] \) is the projection of the basic process onto the two dimensional state space \( Y = \{0, 1, ..., 6\}^2 \).

For each \( y \in Y \) we characterize how cooperation and exclusion change in one
period, starting from a state with \( y_1 \) cooperative and \( y_2 \) exclusive players. For this we estimate, for each \( y \), the average value of \( y[t + 1] - y[t] \) conditional on \( y[t] = y \). Let this value be given for each \( y \in Y \) by the mapping \( \chi : Y \rightarrow \mathbb{R}^2 \). For each \( y \in Y \) the value of \( \chi(y) \) gives the average observed direction of transitions from \( y \). Specifically, the value of \( \chi_1(y) \) gives the average increase in cooperation, and the value of \( \chi_2(y) \) the average increase in exclusion, observed between a period in which there are \( y_1 \) cooperative and \( y_2 \) exclusive players and the next period. In other words, the states with \( y_1 \) cooperative and \( y_2 \) exclusive players induce an average increase in cooperation of \( \chi_1(y) \) and an average increase in exclusion of \( \chi_2(y) \).

For completeness we include the formal definition of the mapping \( \chi \). Let \( \Psi \) be the set of observed realizations of the basic process. For each \( y \in Y \) and realization \( s[t]^{100}_{t=1} \in \Psi \), let

\[
\mu(y \mid s[t]^{100}_{t=1}) = \left| \{ t \leq 100 \mid (y_1(s[t]), y_2(s[t])) = y \} \right|
\]

be the number of periods \( t \) during which the realization \( s[t]^{100}_{t=1} \) was in a state with \( y_1 \) cooperative and \( y_2 \) exclusive players. The mapping \( \chi \) is then defined for each \( y \in Y \) by

\[
\chi(y) = \sum_{s[t]^{100}_{t=1} \in \Psi, y(s[t]) = y} (y(s[t + 1]) - y(s[t])) / \sum_{s[t]^{100}_{t=1} \in \Psi} \mu(y \mid s[t]^{100}_{t=1}).
\]

The mapping \( \chi \) does not accurately describe the basic process. The main shortcoming of using a projection of the process is that it might aggregate over states with very different dynamic properties. It does, however, offer a fair idea of the dynamics within the process.

Results of this class of simulation are presented in Figure 6.1 for \( r = 1, 2, 4 \). The dark spots represent absorbing states of the process. The arrows represent the mapping \( \chi \). For each \( y \in Y \), \( \chi_1(y) \) determines the vertical dimension and \( \chi_2(y) \) the horizontal dimension of the corresponding arrow. In this respect the length and the direction of the vectors represent the "speed" and "direction" of the dynamics. If for some \( y \in Y \) all corresponding states are absorbing, only a dark spot is shown but no arrow.

For all values of \( r = 1, 2, 4 \) the sets of absorbing states yielded by the simulations coincide with those given by Proposition 6.5. All the remaining states are transient which implies that there are no recurrent sets, other than absorbing states. An inspection of Figure 6.1 suggests that the dynamics in the basic process overwhelmingly favors state \( s^{DA} \) over alternative absorbing states. The figure also gives an indication about the dynamics of the perturbed process that we discuss in detail in the following section. For \( r = 2 \) and \( r = 4 \) a single mistake in the choice of best response may lead the process away from absorbing states with full cooperation and into the state \( s^{DA} \) with full defection. In contrast, once in \( s^{DA} \), the process returns to \( s^{DA} \) even after several simultaneous mistakes. In this sense the state \( s^{DA} \) is the only stable among the absorbing
states. The direction of the dynamics suggests that, starting from a random initial state, $s^{DA}$ is the most likely absorbing state to be selected by the process over the long run, for all considered values of $r$.

### 6.4.3 Detecting the asymptotic behavior of the process

Using the first class of simulations we characterized the absorbing states of the basic process, and the dynamics between the remaining states. The process considered was non-perturbed and assumed that players always follow the updating rule (6.2) precisely. This assumption allows us to obtain a useful benchmark for
the asymptotic properties of the process, such as its absorbing states. To select among the absorbing states, and to study the basic process in the presence of perturbations, we consider the second class of simulations. We still assume that players follow their best response most of the time, but allow occasional deviations from the updating rule (6.2). That is, when updating a player chooses a random strategy with small probability $\varepsilon$, and follows the rule (6.2) otherwise.

The asymptotic properties of the slightly perturbed process may be of interest for various reasons. Firstly, human players are often prone to errors in their decision making. Deviations may be unintentional, e.g. due to imperfect information or limited computing abilities, or intentional, such as when players choose to play a non-optimal strategy hoping to induce a transition from one equilibrium to another. In the setting of the repeated basic game, human players may, for example, realize that the absorbing state $s^{DA}$ is payoff dominated by the absorbing state $s^{CE}$ and choose to cooperate and exclude defectors with intention to induce a transition between the two absorbing states. The perturbed process may give us an impression on how successful players can be in inducing such transitions.

Secondly, even though the recurrent sets are the only candidates for the long run selection by the non-perturbed process, the perturbed process may spend only a small fraction of time in those sets. For instance, we see below that the perturbed basic process with $\omega = 0$ spends most of the periods in the absorbing state $s^{DA}$. It spends most of the remaining periods in the states with type distributions $\{(1,0,5,0), (0,1,5,0), (0,0,5,1)\}$, which arise by a single deviation from the absorbing state $s^{DA}$ and have a high number of defective players. We denote a set of these states by $\{s^{DA} - 1\}$.

However, in some instances even a small perturbation of the process results in dynamics that only rarely stabilizes in an absorbing state. In section 6.4.4, for example, we show that the perturbed basic process with $\omega = 40$ and $r = 2$ spends a significant number of periods in states with many cooperative players, even though all absorbing states are fully defective.

Finally, because the considered error probability is small, the stationary distribution $\mu^\varepsilon(s)$ gives us an indication about the limit distribution $\mu^0(s) = \lim_{\varepsilon \to 0} \mu^\varepsilon(s)$ and the stochastically stable states. In section 6.3.2 we argue that only the absorbing states of the non-perturbed process may be stochastically stable. The comparison of the values of $\mu^\varepsilon(s)$ between the absorbing states of the non-perturbed process thus allows us to speculate about their stability.

In section 6.3.2 we argue that the perturbed basic process is ergodic, with a unique stationary distribution $\mu^\varepsilon(s)$. Analytical computation of $\mu^\varepsilon(s)$ requires computation of matrix $M^\varepsilon$ and finding the solution of system (2.2), given this matrix. If the process is observed for a large number of periods $\tau$, however, the observed frequency distribution $\mu^\varepsilon(s|s^0)$ gives with high probability a close approximation of $\mu^\varepsilon(s)$, independently of the initial state $s^0$. Instead of computing $\mu^\varepsilon(s)$ analytically we may therefore follow the perturbed process, initiated in an arbitrary initial state, and, after some large number of periods $\tau$, consider $\mu^\varepsilon(s|s^0)$ as a good approximation of $\mu^\varepsilon(s)$. The calculation of the convergence rate of the frequency distribution toward the stationary distribution requires
Table 6.3: BASIC PROCESS WITH \( o=0 \). In this and subsequent tables the relative frequencies (a) of absorbing and most frequent other states, (b) of different types, obtained from long-run simulations of various perturbed processes, are shown for \( r = 1, 2, 4 \). In this table "6 coop." stands for all states with 6 cooperative players.

<table>
<thead>
<tr>
<th>states</th>
<th>( r ) :</th>
<th>1</th>
<th>2</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 coop.</td>
<td></td>
<td>.01</td>
<td>.01</td>
<td>.01</td>
</tr>
<tr>
<td>( s^{DA} )</td>
<td></td>
<td>.65</td>
<td>.65</td>
<td>.64</td>
</tr>
<tr>
<td>{( s^{DA} - 1 }</td>
<td></td>
<td>.29</td>
<td>.28</td>
<td>.29</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>type</th>
<th>( r ) :</th>
<th>1</th>
<th>2</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( CA )</td>
<td></td>
<td>.02</td>
<td>.02</td>
<td>.02</td>
</tr>
<tr>
<td>( CE )</td>
<td></td>
<td>.02</td>
<td>.02</td>
<td>.02</td>
</tr>
<tr>
<td>( DA )</td>
<td></td>
<td>.93</td>
<td>.93</td>
<td>.93</td>
</tr>
<tr>
<td>( DE )</td>
<td></td>
<td>.03</td>
<td>.03</td>
<td>.03</td>
</tr>
</tbody>
</table>

(a) (b)

equal computing capacity as the analytical calculation of \( \mu^\varepsilon(s) \).\(^6\) Here, we rely rather on an intuitive criterion to decide how long to follow the process.

We simulate the process sufficiently long to observe a sufficient number of transitions between the most distant states. The states which, between them, require the maximal number of transitions are the corner states \( \{s^{CA}, s^{CE}, s^{DA}, s^{DE}\} \). The criterion we implement is to continue the simulations until at least 5 transitions are detected into each of the corner states, from another corner state. The process is initiated in the state \( s^0 = s^{CA} \). We consider \( \varepsilon = 0.1 \).\(^7\) The values of \( \mu^r(s|s^0) \) for the most frequent states of the non-perturbed basic process, as well as the frequencies of different types, are reported in Table 6.3. For any value of \( r \) the process spends around 65 percent of periods in state \( s^{DA} \) and around 29 percent of periods in states \( \{s^{DA} - 1\} \) that can be reached from state \( s^{DA} \) by a deviation of a single player. We conclude that for all considered values of \( r \) the state \( s^{DA} \) appears to be unique stochastically stable state. Consequently, \( DA \) is by far the most frequently observed type.

### 6.4.4 High outside option, \( o = 40 \)

To study the dynamics and the long run behavior of the basic process with \( o = 40 \) we employ the same classes of simulations as in the case with \( o = 0 \). We do not provide an analytical characterization of the absorbing states for \( o = 40 \). Rather, we rely on the simulations of the first class to accurately classify the

\(^6\) Convergence in probability may be evaluated as follows, see Bhattacharya and Waymire (1990) for details. For any two distributions \( \rho \) and \( \sigma \) over the states of the process consider \( ||\rho - \sigma|| = \max_{x \in X} |\rho_x - \sigma_x| \), the maximal difference in probabilities between two distributions. For an ergodic Markov chain on a finite state space, with transition matrix \( M \) and stationary distribution \( \pi \), the convergence of \( \rho M^t \) awards \( \mu \) is at an exponential rate for any initial distribution \( \rho \), that is, \( ||\rho M^t - \mu|| \leq c(\theta^t) \) for some \( 0 < \theta < 1 \) and \( c > 0 \). To find \( \theta \) one needs to calculate the largest non-unit eigenvalue of \( M \), calculation of which is comparable to solving the system \( \mu M = \mu \). Ellison (1993) uses this property to compare rates of convergence between local and global interaction processes.

\(^7\) Smaller error probability may give a better indication about stochastic stability of absorbing states but also substantially increases the duration of simulations.
sets of absorbing and transient states. We have argued above that an inaccurate classification of the states is very unlikely.

Again, we consider parameters given in Table 6.2, with \( o = 40 \). The absorbing states of the basic process, as classified by the first class of simulations under these parameters (see also Figure 6.2), are

- \( r = 1 \) : \( s^{DE} \), and \( s \) such that \( z(s) \in \{(0, 0, 1, 5), (0, 1, 0, 5), (0, 1, 1, 4)\} \),
- \( r = 2 \) : \( s^{DE} \), and \( s \) such that \( z(s) = (0, 0, 1, 5) \),
- \( r = 4 \) : \( s^{CE}, s^{DE} \), and \( s \) such that \( z(s) = (0, 0, 1, 5) \).

All the remaining states were classified as transient, which implies that there are no recurrent sets, other than absorbing states. In the state \( s^{DE} \) all players defect and exclude other defective players, which results in an empty network.
with isolated defective players. For convenience we denote by \( \{s^{DE}+1\} \) the set of states with type distributions in \( \{(0,0,1,5), (0,1,0,5), (0,1,1,4)\} \).

Results of this class of simulation are presented in Figure 6.2 for \( r = 1, 2, 4 \). The direction of the dynamics for \( r = 1 \) suggests that, starting from a random initial state, the process quickly reaches an absorbing state. For \( r = 2 \) only fully defective absorbing states exist but the dynamics suggests that the process may need a large number of periods to reach one of them (state \( (0,3,0,3) \) is not absorbing). More interesting, the illustration suggests that small perturbations may lead the process to spend a substantial proportion of periods in states with 2, 3 or 4 cooperative players. For \( r = 4 \) Figure 6.2(c) does not obviously suggest which of the absorbing states is most likely to be selected by the process. Judging by the dynamics close to the absorbing states, state \( s^{CE} \) seems to be most attractive.

The observations above can be verified using the second class of simulations. Again, the perturbed process with \( \varepsilon = 0.1 \) is initiated in the state \( s^0 = s^{CA} \). For convenience we refer to the absorbing states of the non-perturbed process simply as absorbing states. The values of \( \mu^+(s|s^0) \) for absorbing states and most frequent transient states, as well as the frequencies of different types, are reported in Table 6.4. The statistics about the asymptotic behavior of the process confirm our conjectures inferred from Figure 6.2. For all considered values of \( r \) exclusion is very high. For \( r = 1 \) the perturbed process spends most of the time in the absorbing states implying low frequency of cooperation and low density of links. For \( r = 2 \), however, the perturbed process spends almost half of the time in the non-absorbing states with all players exclusive and some of them cooperative. For \( r = 4 \) the absorbing state \( s^{CE} \) in which all players cooperate and threaten to exclude defectors, appears the unique stochastically stable state.

It is interesting to conclude that with \( o = 0 \) the process remains in state \( s^{DA} \) with full defection for most of the time even when there is a small probability that players experiment with sub-optimal strategies. This contrasts with the behavior of the process with \( o = 40 \) where some cooperation appears as long as the players have a non-trivial scope of forward looking \( r \geq 2 \). In this case a single cooperator may disturb the defective absorbing state \( s^{DE} \) by inducing even more players to cooperate. When their scope of forward looking is \( r = 4 \) players remain in fully cooperative and exclusive state \( s^{CE} \) for much of the time.
6.5 Social preferences

In the previous section players were assumed to consider only their own payoff when choosing their strategy. In this sense we may say they were selfish for not considering the consequences of their choices for the payoffs of the remaining players in the group. When shortsighted such players do not fare well in terms of cooperation. With scope of forward looking \( r = 1 \), and players choosing their strategy by maximizing their own immediate payoff, cooperation could not be sustained in the long run.

This is not surprising in view of the theoretical results that cooperation in social dilemma games is unlikely if selfish players behave according to the myopic best response. In chapter 4, for example, we argue that there should be no cooperation in the finitely repeated basic game if it is (common knowledge that it is) played by fully rational selfish players. Under these assumptions the same holds also for general public goods games with or without punishment, see e.g. Fehr and Schmidt (1999).

These predictions, however, are in stark contrast with experimental evidence. In chapter 5 we discuss an experimental study of the finitely repeated basic game. The observed average cooperation rates were above 80% for low and high outside options. Similarly, positive contributions are commonly observed in experiments with public goods games (see e.g. Davis and Holt, 1993). If a costly punishment possibility is included with the public goods game the contribution rates even increase, may be sustained until the last periods, and may reach 100% (see Fehr and Gächter, 2000).

One explanation provided by several authors is that a non-negligible proportion of people are not really selfish. The utility of people may depend on their own payoff as well as on the payoff of other players in their reference group. We say that people with such utilities are other-regarding and exhibit social preferences rather than selfish ones. In this section we look at two models of social preferences: the model of altruism by Levine (1998) and the model of inequity aversion by Fehr and Schmidt (1999). Each has been shown to explain positive contributions in public goods games with punishment, even when assuming myopic best response behavior. In chapter 5 we study equilibria of our basic game if played by other-regarding, but rational players. In this section we look at the dynamics of cooperation and exclusion in the repeated basic game, played by other-regarding players with limited forward looking. For each set of parameters considered we characterize dynamic and asymptotic properties of the corresponding processes using the simulations introduced in section 6.4. For convenience we repeat the essential definitions from chapter 5.

In Tables 6.5-6.8 below we use the following notation:

- "6 (5) coop." stands for the set of all states with 6 (or 5) cooperative players;
- "other: 6 (5) excl." denotes the set of states with 6 (or 5) exclusive players (not already included in other categories);
• \( \{s^{DA} - 1\} \) is the set of states with type distribution \((1,0,5,0), (0,1,5,0)\) or \((0,0,5,1)\);\(^8\)

• \( \{s^{DE} + 1\} \) is the set of states with type distribution \((0,1,0,5), (0,0,1,5)\), or \((0,1,1,4)\).\(^9\)

6.5.1 Process with altruism

To explain, i.a., positive contributions in public goods games Levine (1998) proposes a model of altruism, which assumes that players gain additional utility from payoffs to those other players they consider nice, and lose utility from payoffs to those other players they consider aggressive. Each player is associated with a coefficient of altruism \(\gamma_i \in (-1,1)\) which determines her other-regarding inclination. If payoffs of players are given by the vector \((x_1, ..., x_n)\), player \(i\) receives the (adjusted) utility

\[
u_i^A = x_i + \sum_{j \neq i} \frac{\gamma_i + \lambda \gamma_j}{1 + \lambda} x_j,
\]

where the coefficient \(\lambda \in [0,1]\) reflects the fact that players have a higher regard for altruistic opponents than for the spiteful ones.

In the basic process all players were assumed to consider only their own payoff, which is the special case of the model with altruism with trivial coefficients \(\gamma_i = 0\) for all \(i\). Levine (1998), however, estimates that data from ultimatum game experiments of Roth et al. (1991) are better explained if \(\lambda = 0.45\). Levine shows that the data from the public goods game experiments of Isaac and Walker (1988) are roughly consistent with the following distribution of coefficients of altruism:

\[
\gamma_i = \begin{cases} 
-0.9 & ; \text{pr} = 0.20 \\
-0.22 & ; \text{pr} = 0.52 \\
0.133 & ; \text{pr} = 0.04 \\
0.243 & ; \text{pr} = 0.17 \\
0.453 & ; \text{pr} = 0.07 
\end{cases} \quad (6.10)
\]

In the following we study the adaptation of the basic process by considering players, whose coefficients of altruism are non-trivial. We consider a group of \(n\) other-regarding players with a uniform scope of forward looking \(r\) repeatedly playing the basic game. At each time \(t\) each player \(i\) plays one of the strategies \(\{s_i^{CA}, s_i^{CE}, s_i^{DA}, s_i^{DE}\}\). As in the basic process, at each time \(t \geq 2\), before making her move, player \(i\) with some fixed probability \(q_i \in (0,1)\) updates her own strategy, by observing the vector \(s_{t-1}\) of strategies at time \(t-1\). A player updates by choosing a strategy which maximizes her utility across the next \(r_i\) periods, under the assumption that the strategies of the other players do not change in those periods.

\(^8\){\(s^{DA} - 1\)} represents states that can be reached from state \(s^{DA} = (0,0,6,0)\) by a deviation of one player.

\(^9\){\(s^{DE} + 1\)} represents absorbing states of the processes with high outside option and \(r = 1\).
Formally, let \( s = (s_1, \ldots, s_n) \in S \) be the vector of strategies at time \( t - 1 \), and let \( s' = (s_1, \ldots, s_{i-1}, s_i', s_{i+1}, \ldots, s_n) \) be the vector of strategies at time \( t \). Player \( i \) changes her strategy from \( s_i \) to \( s'_i \) at \( t \) and assumes no other updates in the group during \( [t - 1, t + r_i] \). Let \( x_i(s' \mid s) \), defined as in (6.1), be the payoff of player \( i \) across the \( r_i \) periods following her update, and let

\[
x_{ij}(s'_i \mid s) = \pi_j(a_j(s_j), p_j(s_j \mid a(s)), a_{-j}(s'), p_{-j}(s' \mid a(s))) + (r_i - 1) \cdot \pi_j(a_j(s_j), p_j(s_j \mid a(s')), a_{-j}(s'), p_{-j}(s' \mid a(s'))),
\]

be the payoff to player \( j \) across the same periods. The utility of player \( i \), for changing her strategy \( s_i \) to strategy \( s'_i \) is then given by

\[
u_i^A(s'_i \mid s) = x_i(s'_i \mid s) + \sum_{j \neq i} \frac{\gamma_i + \gamma_j}{1 + \lambda} x_{ij}(s'_i \mid s).
\]

If player \( i \) is updating at the start of period \( t \) she chooses a strategy \( s_i[t] \) such that

\[
s_i[t] \in \arg\max_{s'_i \in \{s_1^{A,k}, s_2^{A,k}, \ldots, s_n^{A,k}\}} u_i^A(s'_i \mid s[t - 1]).
\]

If this payoff is maximized by the current strategy of the player, the player keeps her current strategy. If not, she randomly selects one of the optimal strategies.

The updating rule outlined above generates a Markov chain with the finite state space \( S \) and transition matrix \( M^A \). We refer to this Markov chain the process with altruism. The basic process can be seen as a special case of the process with altruism when all coefficients of altruism are trivial. Using the same argument that proved Proposition 6.1 it can be shown that in the process with altruism a state \( s \) is absorbing if and only if

\[
s_i \in \arg\max_{s'_i \in \{s_1^{A,k}, s_2^{A,k}, \ldots, s_n^{A,k}\}} u_i^A(s'_i \mid s),
\]

for each player \( i \). Propositions 6.3 and 6.4 hold for any Markov chain and thus for the process with altruism. The following proposition is a generalization of Proposition 6.2 for the process with altruism. Its proof is given in the appendix to this chapter.

**Proposition 6.6** Consider the process with altruism and let there be \( \gamma \) such that \( \gamma_i = \gamma \) for each player \( i \). Let \( z \in Z \) be a type distribution. State \( s \in S_z \) is absorbing if and only if each of the states in \( S_z \) is absorbing. State \( s \in S_z \) is transient if and only if each of the states in \( S_z \) is transient.

In section 6.4.2 we argue that Proposition 6.2 permits us to use the first class of simulations, instead of Algorithm 1, to detect the absorbing and transient states of the basic process. With Proposition 6.6 we can argue in the same way that, whenever players are homogeneous with respect to their coefficients of altruism, we may implement the first class of simulations to detect absorbing and transient states of the process with altruism.
We study two cases with players having homogeneous coefficients of altruism, and one case with players whose coefficients of altruism are distributed according to distribution (6.10). The three processes are characterized by:

A1 $\gamma_i = 1$ for each player $i$,

A2 $\gamma_i = 0.453$ for each player $i$, or

A3 for each player $i$, $\gamma_i$ is independently drawn from (6.10).

We assume $r_i = r$ and $q_i = q$ for each $i$. We consider all the stage game constellations given by Table 6.2 to achieve consistency with parameters considered in the basic process.

We want to study the process with altruism both with realistic coefficients of altruism, as in A3, and with those that give the best chance of cooperation. We conjecture that players with high coefficients of altruism are more likely to achieve cooperation than those with low coefficients of altruism. With this reason we consider the highest feasible coefficient of altruism in A1 and the highest coefficient of altruism allowed by (6.10) in A2.

Processes A1 and A2 consider homogeneous coefficients of altruism and may be studied via simulations of the first, non-perturbed, class or the second, perturbed, class. Process A3 considers heterogeneous coefficients of altruism and may be studied only via the simulations of the second class. Moreover, (6.10) induces a distribution over possible allocations of coefficients in the group. To study the asymptotic behavior of process A3 we repeat 50 runs of the following procedure. In each run we independently draw from (6.10) a coefficient of altruism for each player, and simulate 1.000.000 periods of the perturbed process using simulations of the second class. We assume that coefficients of altruism are common knowledge. To report the results of these simulations we aggregate the frequencies observed across all 50 runs.

**Low outside option**

Let $a = 0$. For each value of $r$ the dynamics of A1 is shown in the left column and the dynamics of A2 in the right column of Figure 6.3. In Table 6.5 we report the results of long run simulations of perturbed processes A1, A2 and A3. For each of the processes we note the frequencies of the most frequent states and the frequency distribution across different types.

By looking at Figure 6.3 we may conclude that cooperation is likely only among sufficiently altruistic players. This seems certainly true when players are altruistic at the extreme, as in the process A1. However, players with high, but not extreme, coefficients of altruism also need to be sufficiently forward looking to achieve cooperation. As is evident from Figure 6.3(b) players whose coefficients of altruism are maximal among those suggested by (6.10) fail to attain cooperation when $r = 1$, that is, if they behave according to the myopic best response.

These observations are confirmed by the second class of simulations, see Table 6.5. On the long run cooperation is sustained only in the process A1 with players
Figure 6.3: PROCESSES WITH ALTRUISM WITH o=0. Non-perturbed processes A1 and A2.
Table 6.5: Processes with altruism with $o=0$. Perturbed processes A1, A2 and A3.

that have extremely high coefficients of altruism, e.g. if $\gamma_i = 1$ for all players. Consider now the process A2 with players whose coefficients of altruism are the highest among those suggested by the theory. Cooperation can be achieved among such players with positive probability only if they are sufficiently forward looking, that is, if $r > 2$. If $r = 1$ then the state $s^{DA}$, in which all players defect, seems to be the unique stochastically stable state and $DA$ is the most frequently observed type. The results for the process A3 show that cooperation cannot be sustained if coefficients of altruism are distributed as suggested by the model of altruism. The process A3 spends around 65 percent of periods in state $s^{DA}$ and around 29 percent of periods in states $\{s^{DA} - 1\}$. The state $s^{DA}$ seems to be the unique stochastically stable state and $DA$ is the most frequently observed type for all values of $r$.

**High outside option**

Let $o = 40$. For each value of $r$ the dynamics of A1 is shown in the left column and the dynamics of A2 in the right column of Figure 6.4. In Table 6.6 we report the results of the long run simulations of processes A1, A2 and A3, respectively. For each of the processes we note the frequencies of the most frequent states
Figure 6.4: PROCESSES WITH ALTRUISM WITH $o=40$. A1 and A2.
and the frequency distribution across different types.

Again, by looking at Figure 6.4 and at the frequencies given in Table 6.6, we may conclude that cooperation is likely only among sufficiently altruistic or sufficiently forward looking players. In the process A1 with players with extremely high coefficients of altruism cooperation always prevails. In the process A2 with players whose coefficients of altruism are the highest among those suggested by Levine’s model complete cooperation can be sustained if \( r > 2 \) but not if \( r = 1 \). Finally, in the process A3 where coefficients of altruism are distributed as suggested by the model of altruism, cooperation increases with forward looking. If \( r = 4 \) then 52 percent of observed types are cooperative.

### 6.5.2 Process with inequity aversion

Experiments with finitely repeated public goods games by Fehr and Gächter (2000) have shown that contribution rates increase by a large magnitude if the possibility for costly individual punishment is provided.\(^{10}\) At odds with

---

\(^{10}\)That punishment may increase cooperation in social dilemma type games was already observed in earlier experiments. Ostrom et al. (1992), for example, report that possibility for punishment significantly increased yield in the repeated common-pool resource games if the number of periods is unknown to players. The experiment of Fehr and Gächter (2000), however, was the first in which punishment possibility was shown to increase cooperation in
experimental evidence, the standard theory assuming selfish players predicts no punishment in the last period, if punishment is costly. To better account for this evidence Fehr and Schmidt (1999) propose a model of inequity aversion, which assumes that players dislike uneven payoffs. The model assumes that people loose utility from being payoff advantaged or disadvantaged, with the loss from an advantage being smaller than the loss from a disadvantage.

Each player is associated with a coefficient of aversion to disadvantageous inequality \( \alpha_i \geq 0 \) and a coefficient of aversion to advantageous inequality \( \beta_i \geq 0 \) satisfying \( \beta_i \leq \min\{\alpha_i, 1\} \). If payoffs of players are given by the vector \((x_1, ..., x_n)\), player \( i \) receives the (adjusted) utility

\[
u^F_i = x_i - \frac{\alpha_i}{n-1} \sum \max\{x_j - x_i, 0\} - \frac{\beta_i}{n-1} \sum \max\{x_i - x_j, 0\}.
\]

Players in the basic process are assumed to be selfish and can be described by having trivial coefficients \( \alpha_i = \beta_i = 0 \). However, in Fehr and Schmidt (1999) the distribution of coefficients of inequity aversion is estimated as

\[
\alpha_i = \begin{cases} 
0 & \text{if } \Pr = 0.3 \\
0.5 & \text{if } \Pr = 0.3 \\
1 & \text{if } \Pr = 0.3 \\
4 & \text{if } \Pr = 0.1 
\end{cases}
\]

\[
\beta_i = \begin{cases} 
0 & \text{if } \Pr = 0.3 \\
0.25 & \text{if } \Pr = 0.3 \\
0.6 & \text{if } \Pr = 0.4 
\end{cases}
\]

This distribution is shown to be consistent with the data from the experimental public goods game with punishment reported in Fehr and Gächter (2000). In the repeated basic game the exclusion of a defector can be viewed as a form of punishment. In the following we therefore consider an adaptation of the basic process by assuming players whose coefficients of inequity aversion are non-trivial.

We consider a group of \( n \) other-regarding players with a uniform scope of forward looking \( r \) repeatedly playing the basic game. At each time \( t \) each player \( i \) plays one of the strategies \( \{s_i^{CA}, s_i^{CE}, s_i^{DA}, s_i^{DE}\} \). As in the basic process, at each time \( t \geq 2 \), before making her move, player \( i \) with some fixed probability \( q \in (0, 1) \) updates her own strategy, by observing the vector \( s[t-1] \) of strategies at time \( t - 1 \). Player updates to a strategy maximizing her utility across the next \( r_i \) periods, under the assumption that the strategies of the other players do not change in those periods.

Formally, let \( s = (s_1, ..., s_n) \in S \) be the vector of strategies at time \( t - 1 \), and let \( s'_i = (s_1, ..., s_{i-1}, s'_i, s_{i+1}, ..., s_n) \) be the vector of strategies at time \( t \). Player \( i \) changes her strategy from \( s_i \) to \( s'_i \) at \( t \) and assumes no other updates in the group during \([t-1, t+r] \). Let \( x_i(s'_i | s) \), defined as in (6.1), be the payoff of player \( i \) across the \( r_i \) periods following her update, and let \( x_{ij}(s'_i | s) \), defined as in (6.11), be the payoff to player \( j \) across the same periods. The utility of

the one-shot and finitely repeated public goods games among strangers.
player $i$ for changing her strategy $s_i$ to strategy $s'_i$ is then given by

$$
\begin{align*}
&u_i^F(s'_i \mid s) = x_i(s'_i \mid s) - \frac{\alpha_i}{n-1} \sum_{j \neq i} \max \{x_{ij}(s'_i \mid s) - x_i(s'_i \mid s), 0 \} \\
&\quad - \frac{\beta_i}{n-1} \sum_{j \neq i} \max \{x_i(s'_i \mid s) - x_{ij}(s'_i \mid s), 0 \},
\end{align*}
$$

(6.15)

If player $i$ is updating at the start of period $t$ she chooses a strategy $s_i[t]$ such that

$$
s_i[t] \in \arg\max_{s'_i \in \{s_{CA}, s_{DA}, s_{CE}, s_{DE}\}} u_i^F(s'_i \mid s[t-1]),
$$

If this payoff is maximized by the current strategy of the player, the player keeps her current strategy. If not, she randomly selects one of the optimal strategies.

The updating rule outlined above generates a Markov chain with the finite state space $S$ and transition matrix $M^F$. We refer to this Markov chain the process with inequity aversion. The basic process can be seen as a special case of the process with inequity aversion when all coefficients of inequality aversion are trivial. Using the same argument that proved Proposition 6.1 it can be shown that in the process with inequity aversion a state $s$ is absorbing if and only if

$$
s_i \in \arg\max_{s'_i \in \{s_{CA}, s_{DA}, s_{CE}, s_{DE}\}} u_i^F(s'_i \mid s),
$$

(6.16)

for each player $i$. Propositions 6.3 and 6.4 hold for any Markov chain and thus for the process with inequity aversion. The proof of the following generalization of Proposition 6.2 for the process with altruism is given in the appendix to this chapter.

**Proposition 6.7** Consider the process with inequity aversion and let there be $\alpha$ and $\beta$ such that $\alpha_i = \alpha$ and $\beta_i = \beta$ for each player $i$. Let $z \in Z$ be a type distribution. State $s \in S_z$ is absorbing if and only if each of the states in $S_z$ is absorbing. State $s \in S_z$ is transient if and only if each of the states in $S_z$ is transient.

Consequently, whenever players are homogeneous with respect to their coefficients of inequality aversion, we may implement the first class of simulations to detect absorbing and transient states of the process with inequity aversion.

We study two cases with players having homogeneous coefficients of inequality aversion, and one case with players whose coefficients of inequality aversion are distributed according to distribution (6.14). The three processes are characterized by:

F1 $\alpha_i = 40$ and $\beta_i = 1$ for each player $i$,

F2 $\alpha_i = 4$ and $\beta_i = 0.6$ for each player $i$, or

F3 for each player $i$, $\alpha_i$ and $\beta_i$ are randomly and independently drawn from (6.14).
Figure 6.5: PROCESSES WITH INEQUITY AVERSION WITH o=0. Non-perturbed processes F1 and F2.
Table 6.7: PROCESSES WITH INEQUITY AVERSION WITH O=0. Perturbed processes F1, F2 and F3.

We assume \( r_i = r \) and \( q_i = q \) for all \( i \). We consider all the stage game constellations given by Table 6.2 to achieve consistency with parameters considered in the basic process.

As in the model with altruism we conjecture that players with high coefficients of inequality aversion are more likely to achieve cooperation than those with low ones. With this reason we consider the highest feasible coefficients of inequality aversion in F1, and the highest coefficients suggested by (6.14) in F2. Because F1 and F2 consider homogeneous coefficients they may be studied via simulations of first or second class. To study the asymptotic behavior of process F3 we follow the same procedure as outlined in section 6.5.1 for the process A3.

Low outside option

Let \( o = 0 \). For each value of \( r \) the dynamics of F1 is shown in the left column and the dynamics of F2 in the right column of Figure 6.5. In Table 6.7 we report the results of the long run simulations of perturbed processes F1, F2 and F3, respectively. For each of the processes we note the frequencies of the most frequent states and the frequency distribution across different types.

Comparing the dynamics in processes with inequity aversion F1 and F2, shown in the Figure 6.5, with that of the basic process for \( o = 0 \), we can see
Figure 6.6: PROCESSES WITH INEQUITY AVERSIO WITH O=40. F1 and F2.
Table 6.8: PROCESSES WITH INEQUITY AVERSION WITH $o=40$. F1, F2 and F3.

A surprising similarity. The absorbing state $s^{DA}$ seems to be the unique stable state for every $r$ considered in both F1 and F2, just as it is in the basic process, even though for the process F1 all fully cooperative states are absorbing. These observations are confirmed by comparing the asymptotic behavior of processes F1, F2 and F3, given in Table 6.7, with that given by Table 6.3 for the basic process. The long-run behavior statistics are virtually identical for all these processes, suggesting that state $s^{DA}$ is the unique stochastically stable state for any distribution of coefficients of inequality aversion in the group. In this sense, we may say that, for $o = 0$, the model of inequality aversion gives the same prediction as the model with selfish players.

**High outside option**

Let $o = 40$. For each value of $r$ the dynamics of F1 is shown in the left column and the dynamics of F2 in the right column of Figure 6.6. In Table 6.8 we report the results of the long run simulations of perturbed processes F1, F2 and F3, respectively. For each of the processes we note the frequencies of the most frequent states and the frequency distribution across different types.

We can compare the results of the simulations of perturbed processes with and without inequity aversion by looking at Tables 6.4 and 6.8. By comparing the frequencies observed in F1, F2, F3 and in the basic process with the high
outside option we conclude that cooperation increases with inequity aversion and with \( r \). However, in F2 and F3 cooperation is increased only marginally, and only if players are sufficiently forward looking, that is, if \( r \geq 2 \).

6.6 Discussion

6.6.1 Results of our simulation study

We review and discuss in this section the main results of the chapter. Many conclusions may be drawn simply by examination of the dynamics of the processes, illustrated in figures 6.1 - 6.6. Inference about the long-run behavior of the perturbed processes and about the stability of absorbing states solely on the basis of the figures may, however, be misleading. We therefore analyze and compare the statistics of the long-run behavior of the processes, given by tables 6.3 - 6.8, but refer to corresponding figures for explanation.

Recall that the process is said to be in equilibrium if it reaches an absorbing state. For reference below we call an absorbing state cooperative if at least one link is established with at least one cooperative player. An absorbing state is called defective if all cooperative players are isolated. The process is in a long-run equilibrium if it reaches a stochastically stable state. When \( r = 1 \), that is, when players update their strategies according to the myopic best response, we say that their scope of forward looking is trivial.

Result 6.1 Outside option

A high outside option induces more cooperation than a low outside option.

This is intuitively plausible. When the outside option is \( o = 40 \) players earn more by excluding defectors than by linking to them. In contrast, with low outside option \( o = 0 \) exclusion yields no earning, while establishing a link always yields a positive earning, regardless of the actions of linked players. In this sense, exclusion is costly when the outside option is low and cheap when the outside option is high. Since more exclusion, as a form of retribution, yields more cooperation, more cooperation should be observed with a high outside option.

The above result may be verified by comparing the long-run cooperation levels, for each of the processes considered, between the low and high outside options. Long-run cooperation levels are computed by summing the frequencies of types \( CA \) and \( CE \), and are given in Table 6.9. For each particular process and fixed scope of forward looking \( r \) the long-run cooperation under \( o = 40 \) is at least as high as the long-run cooperation under \( o = 0 \), with the exception of the process A1 with \( r = 1 \) where the difference is negligible.

Result 6.2 Forward looking

Non-trivial forward looking is necessary and sufficient for the existence of cooperative equilibria in the basic process. In particular, under myopic best response
Table 6.9: Cooperation in perturbed processes.

<table>
<thead>
<tr>
<th>r = 1</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>o = 0</td>
<td>.94</td>
<td>.04</td>
<td>.04</td>
<td>.04</td>
<td>.04</td>
<td>.04</td>
<td>.04</td>
</tr>
<tr>
<td>o = 40</td>
<td>.92</td>
<td>.13</td>
<td>.13</td>
<td>.34</td>
<td>.13</td>
<td>.13</td>
<td>.13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>r = 2</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>o = 0</td>
<td>.94</td>
<td>.37</td>
<td>.04</td>
<td>.04</td>
<td>.04</td>
<td>.04</td>
<td>.04</td>
</tr>
<tr>
<td>o = 40</td>
<td>.94</td>
<td>.94</td>
<td>.40</td>
<td>.88</td>
<td>.76</td>
<td>.50</td>
<td>.36</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>r = 4</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>o = 0</td>
<td>.94</td>
<td>.37</td>
<td>.05</td>
<td>.04</td>
<td>.04</td>
<td>.04</td>
<td>.04</td>
</tr>
<tr>
<td>o = 40</td>
<td>.94</td>
<td>.94</td>
<td>.52</td>
<td>.88</td>
<td>.86</td>
<td>.76</td>
<td>.69</td>
</tr>
</tbody>
</table>

no cooperative absorbing states exist. The same is true for processes with altruism or with inequity aversion unless extreme altruism or inequality aversion is assumed.

For the basic process with o = 0 the above result is implied by Proposition 6.5. It follows from Figure 6.2 that, for the basic processes with o = 40, there exists a fully cooperative absorbing state when r = 4, while all (if any) cooperative players are isolated in any absorbing state when r = 1. The absorbing states of the process with altruism A2 are given in the right-hand columns of Figures 6.3 and 6.4. The absorbing states of the process with inequity aversion F2 are given in the right-hand columns of Figures 6.5 and 6.6. For both processes A2 and F2 and both outside option values considered cooperative absorbing states exist for r ≥ 2, but not for r = 1.

Cooperative equilibria under myopic best response do, however, exist if we assume that the coefficients of altruism or inequity aversion are higher than those estimated from the empirical data. The absorbing states of processes A1 and F1 are shown in the left-hand columns of Figures 6.3 - 6.6. For both outside option values and all scopes of forward looking considered all fully cooperative states are absorbing.

This suggests that high coefficients of altruism or inequity aversion induce more cooperative equilibria than low ones. Further support for this may be found by comparing the sets of absorbing states between the basic process and the processes with altruism A2 and A1, or between the basic process and the processes with inequity aversion F2 and F1, for fixed o and r. If we do this for all fixed pairs of o and r we may conclude:

**Result 6.3 Equilibria induced by social preferences**

*The set of cooperative equilibria of the process with altruism increases with the*
altruism of players. The set of cooperative equilibria of the process with the inequity aversion increases with inequity aversion of players.

Support for this result is given in Figures 6.3-6.6. The set of cooperative equilibria of the basic process is contained in that of process A2, which is further contained in that of process A1. Similarly, the set of cooperative equilibria of the basic process is contained in that of process F2, which is further contained in that of process F1.

A large set of cooperative absorbing states, however, does not imply that the updating dynamics are likely to direct the process toward these states. A case in point is given by Figure 6.1 which illustrates the dynamics of the basic process with $o = 0$. It is evident from the figure that while increasing the scope of forward looking $r$ increases the set of cooperative absorbing states the dynamics remain mostly unaffected and direct the basic process toward the defective absorbing state $s^{DA}$ for all considered $r$. Table 6.3 confirms that the perturbed basic process with $o = 0$ spends most of the time in the state $s^{DA}$. This indicates that $s^{DA}$ is the unique stochastically stable state of this process.

In contrast, even if all equilibria are defective, cooperation may be substantial in the presence of persistent errors. The basic process with $o = 40$ and $r = 2$ is one such example. It is evident from Figure 6.2(b) and Table 6.4 that no cooperative absorbing states of the non-perturbed process exist, but yet the slight perturbation of the process leads to a substantial level of cooperation.

The following two results summarize the outcomes of the long-run simulations of perturbed processes.

**Result 6.4 LONG-RUN EQUILIBRIA, LOW OUTSIDE OPTION**
Consider $o = 0$. State $s^{DA}$, in which all players defect and form a complete network, is the unique stochastically stable state of the basic process, as well as of the process with inequity aversion, for any distribution of inequity aversion and for all considered $r$. It is also the unique stochastically stable state of the process with altruism when, as suggested by the model of altruism, the distribution of altruism is given by (6.10). Cooperative absorbing states may be stochastically stable only when all players have (extremely) high coefficients of altruism, such as in processes with altruism A1 and A2.

**Result 6.5 LONG-RUN EQUILIBRIA, HIGH OUTSIDE OPTION**
Consider $o = 40$. When $r \leq 2$ there are no cooperative stochastically stable states of the basic process. When $r = 2$, however, a small perturbation of the basic process may lead the dynamics away from defective states and, consequently, medium levels of cooperation may be achieved. When $r = 4$ the state $s^{CE}$, in which all players cooperate and exclude defectors, is the unique stochastically stable state of the basic process. If, as suggested by the corresponding theories, coefficients of altruism or inequity aversion are drawn randomly from the distributions (6.10) or (6.14) the processes with altruism and inequity aversion exhibit the same dynamics as the basic process. Cooperation, however, increases when these coefficients are (extremely) high, such as in processes A1, A2, F1, or F2.
The stochastically stable states and the long-run behavior of each process may be inferred from the frequencies of visits of different absorbing states by the process, given in the left-hand columns of Tables 6.3 - 6.8. The final result summarizes the comparison of the basic process with processes with altruism and inequity aversion from the discussion above.

**Result 6.6 Cooperation Induced by Social Preferences**

If coefficients of altruism are distributed by (6.10), as suggested by the model of altruism, the process of altruism exhibits qualitatively the same behavior as the basic process assuming selfish players. Similarly, if coefficients of inequity aversion are distributed by (6.14), as suggested by the model of inequality aversion, the process of inequity aversion exhibits qualitatively the same behavior as the basic process. For \( o = 0 \) the process of inequity aversion exhibits qualitatively the same behavior as the basic process for any distribution of coefficients of inequity aversion.

In chapter 5 we discuss the dynamics of behavior of human subjects playing the finitely repeated basic game in a laboratory experiment. The game was repeated for 60 periods. The behavior in the experiment differed significantly between the last 5 periods and the periods 1-55, which is in accordance with the assumption, made in this chapter, that people look only a few periods ahead. With both low and high outside options high average levels of cooperation were observed during most of the periods 1-55, while a strong tendency toward defection was evident in the final 5 periods, the effect known in experimental repeated games as the "end-game effect".

To see that this is consistent with the limited forward looking assumption, consider players whose scope of forward looking \( r \) is sufficiently long to induce cooperation during initial periods. When the play approaches the final periods of the repeated game players adjust their behavior to the shorter scope of the future. A player, for example, who realizes that there are only \( T < r \) periods left may behave as if having a shorter scope of forward looking \( r \). For sufficiently short \( r \) players may thus defect, which would lead to the last period decline in cooperation.

Among the processes considered in this chapter only the process with altruism A2 reaches substantial cooperation with \( r \geq 2 \) and low cooperation with \( r = 1 \), for both outside options considered. This suggests both (the possibility of) high altruism as well as the non-trivial forward looking as determinants of behavior in our experiments.

### 6.6.2 Discussion of the model assumptions

Some caution may be in order in the interpretation of our results. Firstly, our set of strategies was restricted by assumptions of limited memory and consistency. Enlarging the set of strategies may affect the results. However, it is not obvious which set of strategies would best represent those used in real life. One way to circumvent this problem would be to allow for all feasible strategies via the methodology of genetic algorithms (this approach is endorsed by...
Ashlock et al., 1996). The nature of this approach, however, means that the evolution of strategies is not easily tractable and the interpretation of (success of) the complex surviving strategies is left to speculation (see Smucker et al., 1994). Redundant sophistication yielded by genetic algorithms may sometimes be eliminated by adding the cost of complexity, e.g. as in Abreu and Rubinstein (1988). The manner in which cost is related to complexity may, however, bias the evolution. See, for example, the debate between Binmore and Samuelson (1992) and Cooper (1996).

Another approach is to consider relaxed assumptions about the memory length and consistency of strategies. However, the relaxed bounds are just as artificial as any other, while the resulting set of all possible strategies soon becomes too large to permit straightforward analysis or interpretation. An occasional compromise is to consider a large set of intuitive strategies inferred from real life or experimental evidence (Hauk, 2001 advocates this approach). A set of strategies obtained in this way is, however, far from complete within the considered boundaries, as well as restricted by a subjective judgment.

Best response and imitation models usually assume that players know the strategies played by the other players. This, however, becomes a strong assumption if players use relatively complex strategies. It may be reasonable to assume that actions are perfectly observable, but they need not reveal the strategies that yield these actions. In a world of complex strategies the dynamics of players’ beliefs about strategies of the other players could therefore be better described by a model of learning from observed actions such as Bayesian learning (see e.g. Kalai and Lehrer, 1993 and Nachbar, 1997, Cheung and Friedman, 1997, or Fudenberg and Levine, 1998 for overviews and tests of the models of learning in games).

A model of adaptive play that may allow for complex strategies without such strong assumptions is that of reinforcement learning, suggested by Roth and Erev (1995). In the future we may complement our results with a reinforcement learning model that allows limited forward looking. In this chapter, however, we opt for a best response model within a small but complete set of strategies within clearly defined and intuitive restrictions. The resulting dynamics appears complex yet sufficiently tractable to allow straightforward interpretation.

The second caution, common to all simulation studies, concerns the sensitivity of results to the specific choice of parameters in the simulations. In particular, we considered a single set of prisoner’s dilemma game parameters, group size \( n = 6 \), updating probability \( q = 0.2 \), and error probability \( \varepsilon = 0.1 \). We next discuss how changing these parameters may affect our results.

The error probability does affect the process dynamics, and its impact was described in section 6.3.2. In particular, Theorem 6.1 describes how the asymptotic behavior of the process changes as \( \varepsilon \) decreases toward zero. It states that for sufficiently small \( \varepsilon \) the process is most likely to be observed in one of its stochastically stable states. We, however, have shown that if perturbation is small but not negligible the long-run behavior of the process may differ greatly from this limiting prediction. One such example is given by Result 6.5 for our basic process with \( a = 40 \) and \( r = 2 \).
Figure 6.7: Average cooperation in simulations of the basic process, the process with altruism A3, and the process with fairness F3, for different values of $f$, with $r = 2$ and (a) low outside option $o = 0$, and (b) high outside option $o = 40$. Note that $f = 70$ is chosen in all simulation in this chapter.

The impact of the prisoner's dilemma game parameters for the process dynamics requires more thorough elaboration. For this we perform a test of the robustness of our main results to a change in the value of $f = v(D, C)$, the prisoner's dilemma game payoff when defecting on a cooperator. This value may be seen as a proxy for the incentive to defect. We use the simulations of the second class to detect the asymptotic behavior of the basic process, the process with altruism A3, and the process with inequity aversion F3, for 26 different values of $f$, equidistantly taken from the interval $[50, 100]$. We assume $r = 2$ in all these simulations. For each value of $f$ and each of the two outside options considered we register the average numbers of cooperative players in each of the three processes. Figure 6.7 shows the graphs of average cooperation depending on the value of $f$, for each of the processes and outside option values. For low outside option cooperation levels remain negligible for all processes and values of $f$. For high outside option, however, the value of $f$ does have an effect on cooperation. It is interesting to note that for low $f$, close to $c = 50$, cooperation is lower in the process with altruism A3 than in the basic process. This is most likely due to the presence of spiteful players with negative coefficients of altruism, which gain utility by decreasing the payoff of other players.

Finally, our definition of forward looking is a relatively simple way of modeling the relation between one's own present behavior and the future behavior of the others. We assume that players believe, when they are updating their own strategy, that no other players will update in the periods covered by their scope of forward looking. This assumption seems a fair approximation of the (short span) future behavior if, as is the case in our model, the possibilities for updates are rare, and it is correct once the play reaches a long-run equilibrium. It would nevertheless be worthwhile to relax it in a model that allows for players who understand that (and how) other players may revise their strategies in the future. Such players might, for example, strictly prefer exclusion from non-exclusion, even when it is costly to exclude, because they realize that threatening with exclusion induces other players to abandon defective strategies.
Jehiel (1995) suggested such a model when introducing the concept of limited foresight. Players with limited foresight in each period choose an action which maximizes their payoff across a finite number of coming periods, believing that play is going to follow a particular sequence of action profiles. The game is said to be in a limited foresight equilibrium if, at each period, beliefs of all players coincide with the actual path of play. Jehiel (1998) describes a model of learning about the strategy used by the opponent, that leads to a limited forecast equilibrium in two-player games. Such an equilibrium considers players who at each period use the truncated history of play to update their belief about the future paths of play. In Jehiel (2001) cooperation in repeated two-player prisoner’s dilemma game is shown to be sustainable in such an equilibrium. Adjusting this concept to the setting of our $n$-player stage game may, however, prove difficult, among other things due to the large number of possible action paths.

A simpler approach, arising from our model of limited forward looking, may be to consider players that know the transition matrix of the basic process and assume that their opponents behave according to the underlying updating process. We may call such players level-3 players, as opposed to level-2 players which understand the strategies of their opponents but assume that they never update. When updating, a level-3 player chooses a strategy that maximizes her expected payoff across several coming periods, under the assumption that other players change strategies according to the basic process. We may similarly define level-4 players, or indeed, level-$n$ players for any $n$. Finally, we may consider the dynamics of the process that emerges as $n$ goes to infinity. Players in such a process would correctly predict the behavior of their opponents, as opposed to finite-level players. It would be interesting to compare the model of players with such extreme comprehension to the existing models of rational players with perfect foresight (Selten, 1965, van Damme, 1989). For a similar iterative definition of the levels of reasoning see Stahl (1993) and Nagel (1995).

**6.7 Conclusions**

In this chapter we study the dynamics of social exclusion and cooperation among boundedly rational players that to a limited extent understand the consequences of their own actions and take them into account when adapting their behavior. We show that the assumption of limited forward looking induces cooperative equilibria even if players are best responding to their neighborhood. These equilibria, however, are often not stable with respect to occasional deviations from the best response behavior that we assume players make. In particular, when exclusion is costly cooperation remains low. It turns out that limited forward looking increases cooperation only when exclusion is sufficiently cheap.

Assuming that players are other-regarding, rather than selfish, does not substantially change these results. We consider in this chapter two theories of social preferences: the model of altruism and the model of inequality aversion. Assuming altruistic or inequity averse players further increases the set of co-
operative equilibria. However, if altruism or inequity aversion is distributed according to the distributions proposed by the corresponding models the dynamics of behavior among other-regarding players does qualitatively not differ from that among selfish players. This suggests that, at least in the setting of the repeated prisoner’s dilemma game played on an endogenous network, a limited comprehension of the reactions to one’s own actions may be as important for sustaining cooperation as is a concern for others.

6.8 Appendix: Proofs

Proof of Proposition 6.2. We begin by characterizing the payoff (6.1) of player $i$ in terms of the distribution of types. Let $z(s) = (n_{CA}(s), n_{CE}(s), n_{DA}(s), n_{DE}(s))$ be the distribution of types in state $s$. For each $j$ let $T_j$ be the type of player $j$ in state $s$. Consider a player $i$ that is updating her current strategy $s_i = s_{T_i}$ with a new strategy $s'_i = s'_{T'_i}$ of type $T'_i$. For each player $i$, let $\delta_i^T(s) = 1$ if her type in state $s$ is $T$, and $\delta_i^T(s) = 0$ otherwise. The four-tuple

$$z(s) - \delta_i(s) = (n_{CA}(s) - \delta_i^{CA}(s), n_{CE}(s) - \delta_i^{CE}(s), n_{DA}(s) - \delta_i^{DA}(s), n_{DE}(s) - \delta_i^{DE}(s))$$

gives the distribution of types among the players $N/i$ in state $s$.

The definition of the stage game payoff (6.2) may be rewritten as

$$\pi_i(a_i, p_i, a_{-i}, p_{-i}) = \sum_{j \in N/i} v(a_i, a_j) \cdot p_{ij} \cdot p_{ji} \quad (6.17)$$

for the case that the outside option value is $o = 0$, and as

$$\pi_i(a_i, p_i, a_{-i}, p_{-i}) = \sum_{j \in N/i} [v(a_i, a_j) \cdot p_{ij} \cdot p_{ji} + (1 - p_{ij} \cdot p_{ji}) \cdot o] \quad (6.18)$$

for general $o$. Let $p^{TT'}$ be 1 if players of type $T$ propose links to players that were of type $T'$ in the previous period, and 0 otherwise. Let $a^T$ be the action chosen by strategy $s^T$. Combine the definition of payoff (6.1) with the definition of the stage game payoffs (6.17) to obtain

$$x_i(s'_i | s) = \sum_{j \in N/i} v(a_i(s'_i), a_j(s_j)) \cdot p_{ij}(s'_i | a(s)) \cdot p_{ji}(s_j | a(s))$$

$$+ (r - 1) \sum_{j \in N/i} v(a_i(s'_i), a_j(s_j)) \cdot p_{ij}(s'_i | a(s')) \cdot p_{ji}(s_j | a(s'))$$

$$= \sum_{T \in \{CA, CE, DA, DE\}} (n^T(s) - \delta_i^T(s)) \cdot v(a^T, a^T) \cdot p^{TT'} \cdot [p^{TT_i} + (r - 1)p^{TT'_i}] \quad (6.19)$$
for the case $o = 0$. One can show in the similar fashion that for general $o$,

\[
x_i(s'_i \mid s) = \sum_{T \in \{CA, CE, DA, DE\}} (n^T(s) - \delta^T_i(s)) .
\]

\[
[v(a^T_i, a^T) \cdot p^{T'_iT} \cdot [p^{TT_i} + (r - 1)p^{TT_i}]] + o \cdot \left(1 - p^{T'_iT} \cdot [p^{TT_i} + (r - 1)p^{TT_i}]\right)
\]

(6.20)

Let

\[
\hat{v}[T_i, T'_i \mid T] = v(a^T_i, a^T) \cdot p^{T'_iT} \cdot [p^{TT_i} + (r - 1)p^{TT_i}]
\]

\[
+ o \cdot \left(1 - p^{T'_iT} \cdot [p^{TT_i} + (r - 1)p^{TT_i}]\right)
\]

(6.21)

be the $r$-period payoff to any player $i$ that changed her type from $T_i$ to $T'_i$ from the interaction with any player $j \neq i$ of type $T$. This payoff is independent of the identities $i$ and $j$. The equation (6.20) rewrites as

\[
x_i(s'_i \mid s) = \sum_{T \in \{CA, CE, DA, DE\}} (n^T(s) - \delta^T_i(s)) \cdot \hat{v}[T_i, T'_i \mid T]
\]

The best response of each player thus depends on her own previous type and on the distribution of types among the remaining players. Consequently, players who were playing the same strategy in the previous period have the same set of best responses, which depends only on the distribution of the types in the previous state. If state $s = (s_1, ..., s_n)$ is absorbing, than the state $s^\phi = (s_{\phi(1)}, ..., s_{\phi(n)})$, for any permutation of the player indices $\phi : N \to N$, is also absorbing. To conclude the proof, note that $S_z$ consists exactly of all states $s^\phi$ that can be obtained by permuting the player indices via any permutation $\phi$, from any state $s \in S_z$.

**Proof of Proposition 6.5.** Several observations can be made on the basis of (6.19). First, non-exclusive strategy always earns at least as much as its exclusive counterpart: for all $s$,

\[
x_i(s^C_i \mid s) \geq x_i(s^E_i \mid s) \quad \text{and} \quad x_i(s^D_i \mid s) \geq x_i(s^E_i \mid s).
\]

(6.22)

This is because $p^{CA,T} \geq p^{CE,T}$ and $p^{DA,T} \geq p^{DE,T}$ for all $T$, with strict inequalities for $T = DA, DE$. The equality is attained only in the border cases: $x_i(s^C_i \mid s) = x_i(s^E_i \mid s)$ implies that $n^D = n^E = 0$. Similarly, $x_i(s^D_i \mid s) = x_i(s^E_i \mid s)$ implies $n^D = 0$.

[1] Strategy $s^{DE}$ may thus be a part of an absorbing state only when $n^D = n^E = 0$. Consider a state with type distribution $(n - p, 0, 0, p)$. If $r = 1$ cooperation is not a best response, implying $p = n$. For $r \geq 2$ exclusion is not best response in presence of another defector, implying $p \leq 1$. For $p = 1$ cooperation is a best response if and only if condition (6.7) holds.
When there are no exclusive players, that is when $n_{CE} = n_{DE} = 0$, the utility (6.19) rewrites as

$$x_i(s'_i \mid s) = \sum_{T \in \{CA, DA\}} n^T \cdot v(a^{T'_i}, a^T) \cdot p^{T'_i T} \cdot r$$

in which case the strategy $s^{DA}$ is the best response. To see this note that $v(D, a^T) > v(C, a^T)$ for any $T$, which in case that in $s$ there are no exclusive players implies $x_i(s^{DA}_i \mid s) \geq x_i(s^{CA}_i \mid s)$. Combine this with (6.19) to conclude that $s^{DA}$ maximizes the utility.

When $n_{CE} \geq 1$, there are no defective players, $n_{DA} = n_{DE} = 0$. A state with type distribution $(p, n-p, 0, 0)$ is absorbing whenever players playing strategy $s^{CA}$ do not receive higher utility (6.19) by updating to strategy $s^{DA}$. A straightforward calculation yields that for a player $i$ with $s_i = s^{CA}$, $x_i(s^{CA}_i \mid s) \geq x_i(s^{DA}_i \mid s)$ if and only if condition (6.8) holds.

Absorbing states may therefore have only one of the following type distributions: $(0, 0, 0, n)$ only for $r = 1$, $(n-1, 0, 0, 1)$ whenever (6.7), $(0, 0, n, 0)$, and $(p, n-p, 0, 0)$ whenever (6.8). Constraint $f > c > d > e$ implies that conditions (6.7) and (6.8) do not hold for $r = 1$.

Finally, note that for $p = 0$ condition (6.8) rewrites to (6.6). □

**Proof of Propositions 6.6 and 6.7.**

Let $\bar{v}[T, T'_i \mid T]$ be defined as in (6.21) and let

$$\bar{v}[T \mid T_i, T'_i] = v(a^T, a^{T'_i}) \cdot p^{T'_i T} \cdot [p^{T T'_i} + (r-1)p^{T T'_i}]$$

be the $r$-period payoff to any player $j$ of type $T$ from interaction with player $i \neq j$ that changed her type from $T_i$ to $T'_i$. When $o = 0$, the $r$-period utility of player $j$, $x_{ij}(s'_i \mid s)$, defined as in (6.11), can be rewritten as

$$x_{ij}(s'_i \mid s) = r \sum_{k \in N \setminus \{i,j\}} v(a_j(s_j), a_k(s_k)) \cdot p_{jk}(s_j \mid a(s)) \cdot p_{kj}(s_k \mid a(s))$$

$$+ v(a_j(s_j), a_i(s'_i)) \cdot p_{ji}(s_j \mid a(s)) \cdot p_{ij}(s'_i \mid a(s))$$

$$+ (r-1)v(a_j(s_j), a_i(s'_i)) \cdot p_{ji}(s_j \mid a(s')) \cdot p_{ij}(s'_i \mid a(s'))$$

$$= \sum_{T \in \{CA, CE, DA, DE\}} (n^T(s) - \delta^T_j(s) - \delta^T_i(s)) r \cdot v(a^{T_j}, a^T) \cdot p^{T_j T} \cdot p^{T T_j}$$

$$+ v(a^{T_j}, a^{T'_i}) \cdot p^{T'_i T_j} \cdot [p^{T_j T_i} + (r-1)p^{T_j T'_i}]$$

again using the definition of the stage game payoffs (6.17), which further implies

$$x_{ij}(s'_i \mid s) = \sum_{T \in \{CA, CE, DA, DE\}} (n^T(s) - \delta^T_j(s) - \delta^T_i(s)) \cdot \bar{v}[T_j \mid T, T] + \bar{v}[T'_j \mid T_i, T'_i]$$

200
and

\[
x_i(s'_i | s) - x_{ij}(s'_i | s) = \sum \limits_{T \in \{CA, DA, CE, DE\}} (n^T(s) - \delta_i^T(s)) \cdot \tilde{v}[T, T_i | T]
\]

\[
- \sum \limits_{T \in \{CA, DA, CE, DE\}} (n^T(s) - \delta_i^T(s)) \cdot \tilde{v}[T_j | T, T]
\]

\[
+ \tilde{v}[T_j | T_j, T_j] - \tilde{v}[T_j | T_i, T_i].
\]

One can show in the similar fashion that the above equality holds for general \( o \).

It is important to note that players, distinct from \( i \), of the same type have the same utility: if \( T_j = T_k \) for some \( j, k \neq i \), then \( x_{ij}(s'_i | s) = x_{ik}(s'_i | s) \), and

\[
x_i(s'_i | s) - x_{ij}(s'_i | s) = x_i(s'_i | s) - x_{ij}(s'_i | s). \quad (6.23)
\]

This motivates the use of notation

\[
x_i^T(s'_i | s) = \sum \limits_{T \in \{CA, DA, CE, DE\}} \left( n^T(s) - \delta_j^T(s) - \delta_i^T(s) \right) \cdot \tilde{v}[T | T_i, T_i] + \tilde{v}[T | T_i, T_i]
\]

for each type \( T \). Clearly, \( x_{ij}(s'_i | s) = x_{iT_j}(s'_i | s) \).

[1] Assume the model of altruism of section 6.5.1 and let \( \gamma_i = \gamma \) for each player \( i \). In this case the definition of utility (6.12) reduces to

\[
u_i^A(s'_i | s) = x_i(s'_i | s) + \gamma \sum \limits_{j \neq i} x_{ij}(s'_i | s)
\]

\[
= x_i(s'_i | s) + \gamma \sum \limits_{T \in \{CA, CE, DA, DE\}} \left( n^T(s) - \delta_i^T(s) \right) x_i^T(s'_i | s)
\]

The value of \( u_i^A(s'_i | s) \) therefore depends on the previous and the updated strategies of player \( i \), and on the distribution of types among the remaining players. In particular, it does not depend on the identities of players of particular type. This means that if two players were of the same type in the previous period, the sets of their best responses coincide. Following the final steps outlined in the proof of Proposition 6.2 we prove the claim of Proposition 6.6.

[2] Assume the model of inequity aversion of section 6.5.2 and let \( \alpha_i = \alpha \) and \( \beta_i = \beta \) for each player \( i \). The definition of utility (6.15) reduces to

\[
u_i^F(s'_i | s) = x_i(s'_i | s)
\]

\[
- \frac{\alpha}{n-1} \sum \limits_{T \in \{CA, CE, DA, DE\}} \left( n^T(s) - \delta_i^T(s) \right) \max \{ x_i^T(s'_i | s) - x_i(s'_i | s), 0 \}
\]

\[
- \frac{\beta}{n-1} \sum \limits_{T \in \{CA, CE, DA, DE\}} \left( n^T(s) - \delta_i^T(s) \right) \max \{ x_i(s'_i | s) - x_i^T(s'_i | s), 0 \}.
\]

Again, the value of \( u_i^F(s'_i | s) \) depends on the previous and the updated strategies of player \( i \), and on the distribution of types among the remaining players. In particular, it does not depend on the identities of players of particular type and the claim of Proposition 6.7 follows in the similar steps as above. 

201