Learning and evolution in games and oligopoly models

Possajennikov, A.

Publication date:
2000

Document Version
Publisher's PDF, also known as Version of record

CentER

Learning and Evolution in Games and Oligopoly Models

Alexandre Possajennikov

Learning and Evolution in Games and Oligopoly Models

Learning and Evolution in Games and Oligopoly Models

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Katholieke Universiteit Brabant, op gezag van de rector magnificus, prof. dr. F.A. van der Duyn Schouten, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op woensdag 8 maart 2000 om 16.15 uur door

ALEXANDRE POSSAJENNIKOV

geboren op 20 juni 1971 te Perm, Rusland.

PROMOTOR: prof. dr. E.E.C. van Damme

To Masha

Acknowledgments

This thesis is a collection of papers that were written during my participation in the doctoral program at CentER for Economic Research of Tilburg University. My thanks to CentER for providing me with this possibility. The Dutch Science Foundation (NWO) also provided me with funds to travel to the conferences and supported my three-month visit to University College London in the framework of the ENTER exchange program.

My most sincere thanks go to my wife Maria Montero, who provided me with constant support both in and out the university. She has not only been "the light of my life" but also encouraged me to go forward with my thesis and gave many most valuable suggestions and comments. The domestic atmosphere created by her led to a higher productivity and better output of research.

I would like to thank my supervisor, Eric van Damme, for his (rather unpressing) guidance of my research. His remarks and suggestions were always in the depth of the matter and supplied me with new ideas. I also thank Tilman Börgers for his guidance and support during my visit to UCL.

Tilburg, November 1999.

Contents

Chapter 1

Games, Evolution, and Learning

1.1 Introduction

Imagine yourself in a situation where you have to make a choice among several actions. The actions will bring you some payoff; however, you do not know with certainty what payoff each particular action will bring. What will be your choice?

Of course, the situation above is not completely specified. There may be a lot of extra information available to analyze it. For example, you might remember being in a similar situation before; this may help in making a choice now. It might be known that there are others who have been in a similar situation; you may learn from them. It might be that the actions of others may influence your payoff; then you are in a game situation. Many specifications are possible and several of them will be considered in this thesis. Nevertheless, the structure of the first paragraph (a situation with a given set of actions with uncertain payoffs) is the base on which they all are built.

Examples of such situations are abound. The whole life is, in a sense, a sequence of such problems. Think of buying a yogurt: will you like it? will it be tasty? will it be good for your stomach? For an interactive situation, imagine approaching a street crossing on a bicycle (typical for Dutch life): will the car that is approaching from the other direction go first or let you pass? Or even simpler, passing through a door: will the other person go first or let you pass? These are all examples of (very simple) problems of the character described above. This thesis is a collection of papers that consider such sort of situations from a particular point of view: with limited information and repeated experience, what can and should one do in them?

There is a powerful tool to analyze such situations: game theory (and decision theory as its restriction to the one-player case). It is, indeed, a powerful tool: for almost every situation described above, it provides a solution (or a set of solutions). However, the underlying assumptions of standard game theory are strong: it is required that each individual knows his preferences over possible outcomes, the preferences of all other individuals involved, all the relevant probabilities etc. Observe the difference with the first paragraph: all that is said there is that a set of actions is known and that there is a payoff obtained. Moreover, in classical game theory each individual is very capable: all the multiple interdependent maximization problems can be solved and all the information available is used to that purpose. This is quite a requirement for modest human abilities: can you tell quickly what the 50th digit in the decimal notation of number π is? The use of a less demanding approach suggests itself.

There has been an "as if" defence of this full rationality approach, manifested in Alchian (1950) and Friedman (1953). This defence, in my view, is not exactly valid. First, it says that a theory should be judged by results rather than by assumptions. But nobody yet convincingly demonstrated that what happens in reality corresponds to what the full rationality theory predicts. Second, even if the predictions of the full rationality theory were close enough to reality, in some cases an alternative theory may make predictions not far from those of the full rationality theory (as we will see), and they are not easy to distinguish. Then the use of a "wrong theory" instead of the "right theory" for making further predictions about the development of the system may lead to rather disastrous results. The fact that a theory seemingly worked for one case does not imply that it works for other, even similar, cases. We do not claim that we found an alternative theory that is "true"; what we say is that the models of this thesis are at least no worse than the full rationality model in some cases and that they are closer to the "common sense" decision making. The rational game theory will serve us as a useful benchmark.

There are two basic components that the individual in the situation described knows: the set of actions and the payoff or utility obtained. The second paragraph in the beginning describes how information can be obtained, usually through own or others' past experience. To model this, a dynamic approach should be used. The essence of this approach is that the situation is repeated so that past experience can be used to guide future actions. There is a tool for this as well: the theory of repeated games. Nevertheless, as argued in the previous paragraphs, limited human abilities point to the use of a less demanding approach. Thus the use of information by the individual will be limited: not all past experience will be used. The basic components of the models considered will be a dynamic structure and bounded rationality.

To complete the model a rule of transforming past experience into future actions should be specified. There are great many possibilities to do this; some of them will be discussed briefly below in the introduction, some in detail in the main body of the thesis. Given a rule, the use of actions changes over time. What actions will be used in the long run[1]? These actions are in a sense a solution for the situation. What is the relationship of this solution with the solutions of classical rational game theory? These are two main questions that will be addressed in this thesis.

The theory of dynamic models with bounded rationality in games led to the appearance of several books[2] and numerous journal articles in recent years. The state of the theory currently is rather disperse; there are few general results and a great deal of specific ones. Below we discuss rather general specifications of dynamic models. The main body of the thesis contains more specific models that stem from "common sense" assumptions and lead to interesting results.

1.2 Evolution

One dynamic model of bounded rationality in game theory stems from the biological idea of evolution. Imagine that individuals in a large population are genetically programmed to play a certain strategy (action). Payoffs represent fitness. Fitness may mean higher probability of survival or higher number of offspring. Offspring inherit the strategy of the parent. Thus, the proportion of the population playing the strategy with a higher payoff increases as evolution progresses.

The idea is close to the "as if" explanation of rationality in economics: if an individual uses a strategy that does not maximize his payoff, evolution will drive out such a strategy. The idea is appealing; indeed it has found support both in economic theory and in game theory. The models introduced along evolutionary lines, such as evolutionary stable strategy and replicator dynamics (see, e.g. Weibull (1995)), lead often to a subset of Nash equilibria, the solutions found by the classical game theory.

There are exceptions, however. Instead of converging to the payoff maximizing strat-

[1]Why are we interested in the long run outcome? An answer could be that "in the long run we are all dead" but some are more dead than others! More seriously, first, often asymptotic properties are easier to analyze than finite time properties. Second, the long run may be not that long, and provide good approximation for finite times.

[2]Weibull (1995), Vega-Redondo (1996), Samuelson (1997), Fudenberg and Levine (1998), Young (1998).

egy, the evolutionary process may cycle. Then the state of the population in any given moment does not coincide with what classical game theory predicts. As we will see in chapters 4 and 5 of the thesis, in some interactive situations maximization of own payoff does not necessarily lead to the highest payoff, thus evolution of preferences does not necessarily lead to the classical game theory solutions.

There are also conceptual difficulties with evolutionary interpretations. One needs a large population to apply the model. Instead of focusing on the behavior of individuals, the aggregate behavior of population is analyzed. This is similar to macro models rather than to the analysis of the behavior of micro units. Individuals have no choice: their strategy is genetically programmed. Observe that this changes the problem in the very first paragraph of the thesis: the action set is reduced to the one genetically programmed action. Still, one can speak of the whole population changing actions as evolution progresses. In this case the question of which action to choose stands before the society rather than before a single individual.

1.3 Learning

A more general dynamic approach that has been used in game theory is learning. Learning is a broad notion. It is understood here as individual learning of how to act in the situation described in the very first paragraph, when information is very scarce. There are a lot of ways to model individual learning; some ways may, after aggregation, coincide with the evolutionary models described in the previous section. But there are other ways too. Furthermore, a large population is not needed. Thus, in my view, learning is broader than evolution.

A learning rule is a rule that maps past experience into future actions. There are almost no requirements on a learning rule, except that it uses the payoff obtained and looks plausible from a common sense point of view. Observe that we did not say a word about bounded rationality in learning yet; indeed there are learning models that use all the information that comes along in a rational way. However, there are rules that are plausible and boundedly rational and they often perform not worse than the rational rules in games. In some of the chapters of the thesis we consider such learning rules.

The learning rules can be deterministic, that is the past experience completely determines future actions. The deterministic approach has, however, two major problems. One is theoretical: deterministic models often have too many possible outcomes and different initial conditions lead to different outcomes. Another is empirical: there are many factors in human behavior that are unexplained as yet. To overcome these problems, stochastic learning rules are used. Stochastic rules give a probability distribution over future actions.

The question that we analyze in this thesis is, given a learning rule, what is the outcome of the learning process? What action(s) is(are) played in the long run? That is an interesting question but another question comes up: where does the learning rule come from? Why do we analyze this particular learning rule and not another one? The answer is that we do not know yet what learning rules are used in reality, if any; the empirical and experimental evidence is not convincing in favor of one or another rule. Thus, we do not address this question in the thesis. The first question above has more normative character: what are the rules that perform well? The answer to this question that this thesis tries to provide might help in understanding real learning processes in life in general and in the economy in particular.

1.4 Overview of the Thesis

This thesis is about the problems and models described above. Chapter 1, about half of which you have already read, gives an introduction to the subject in the first half and technical preliminaries in the second half. Chapters 2 to 6 present the analysis of some particular dynamic models for some particular games.

Chapter 2 considers a reinforcement model in decision problems and several games, both from the point of view of convergence to the optimal action and the speed thereof. Reinforcement learning occurs when a strategy is reinforced by a good payoff associated with it. The model is derived from a learning model of Bush and Mosteller (1955) whose relevance for economics was demonstrated in Arthur (1993) and Roth and Erev (1995). Several modifications of the basic reinforcement dynamic are considered. Their convergence properties in decision problems are analyzed. It is shown that there is a trade-off between convergence to the optimal action and the speed of convergence; one modification is chosen to cope with the trade-off. This modification is then applied to certain games. By means of computer simulations we show that an equilibrium that is more central in the set of equilibria, or more egalitarian (has more egalitarian distribution of payoffs) is more likely to be observed in the medium run. There is a connection between the egalitarian equilibrium and the notion of risk-dominance; this connection is also illustrated in the chapter. The chapter is based on Possajennikov (1997).

Chapter 3 analyzes another dynamic process, imitation, in a Cournot-type game. When a player observes a payoff higher than his own, he may be tempted to imitate the strategy that brings this higher payoff. Vega-Redondo (1997) has demonstrated that the rule "imitate the best" in small populations lead to the inefficient (for firms) Walrasian outcome. The chapter considers two modifications of the simple "imitate the best" rule. One modification is imperfect imitation when players cannot observe full strategies of other players. The other one is separating interaction and imitation by enlarging the population. Introducing imperfection in the imitation process allows players to achieve better results than without imperfections; also separating interaction and imitation leads to the same outcome. The chapter is a revised version of Possajennikov (1999a).

Chapters 2 and 3 consider examples of particular learning processes that are not modeled by evolution; chapters 4 and 5 take another view and consider a stability notion that comes from an evolutionary process. Chapters 4 and 5 model evolution of preferences. Evolution works indirectly; it affects the composition of the preferences in the population through equilibrium payoffs. The model can be also understood as an adjustment in players' valuations of different strategy combinations through cultural evolution of values. This indirect evolution approach was initiated in Güth and Yaari (1992). Games stemming from industrial organization are analyzed in chapter 4 while chapter 5 considers 2×2 symmetric games from a more general point of view. Chapter 4 extends the model of Bester and Güth (1998) to spiteful preferences and shows when such preferences are stable. It also demonstrates that the results of indirect evolution may be affected by a change in the set of feasible preferences. As a logical consequence of this result, chapter 5 considers all possible (von Neumann-Morgenstern) preferences in 2×2 symmetric games and analyzes their stability. In some games, like prisoners' dilemma and chicken-type games, selfish preferences (that is, the ones that maximize the material payoff) are not evolutionarily stable while in others, like coordination problems, they are stable. The models in chapters 4 and 5 are extended to an incomplete information setup when players do not know the preferences of the opponent. In both chapters it is shown that with complete information selfish preferences are not necessarily stable under evolution; with incomplete information such preferences are stable more often. Some results of chapter 4 can also be found in Possajennikov (1999b).

Finally, chapter 6 is slightly different in that it does not model explicitly a dynamic process. Instead, equilibrium selection techniques related to dynamic processes are applied to a merger game. The merger game is modeled as a non-cooperative game of coalition formation (see Bloch (1997) for a survey of such games). The non-cooperative merger game in the chapter often has multiple equilibria. The equilibrium selection techniques are applicable for any three-player merger game and their work is illustrated in an asymmetric linear Cournot oligopoly. It is shown that the equilibrium selection approach chooses a unique equilibrium. In this equilibrium a merger that is preferred by its participants to other mergers forms. Thus, the model allows to predict which merger is more likely to occur. The analysis of the asymmetric Cournot triopoly shows how asymmetries affect the selected merger. Roughly speaking, if asymmetries are large, the most and the next efficient firm merge, while if asymmetries are small, the most and the least efficient firms merge. The chapter grew from a joint project with Maria Montero.

All the chapters are almost self-contained. The basic concepts from game theory that are used in the thesis are introduced below.

1.5 Preliminaries

This section describes some basic notions of preference theory, game theory, and stochastic processes that are used in the thesis. For a more thorough treatment of preference theory and game theory the reader is referred to Mas-Colell, Whinston, and Green (1995, Chs. 1, 6, 7, 8, 9) and for stochastic processes to Bhattacharya and Waymire (1990, Chs. 1, 2, 0).

Preferences

The basic idea in economic analysis is that each economic agent has *preferences*. Given a finite set A of possible outcomes, with a being a typical element of A, agent i has a complete and transitive preference relation $\succsim_i$ over outcomes, where $a' \succsim_i a''$ denotes that the agent prefers a' over a'' or is indifferent between them. Such preferences can be represented by a utility function $u_i : A \to \mathbb{R}$, such that $a' \succsim_i a'' \iff u_i(a') \geq u_i(a'')\ \forall a', a'' \in A$. Assume moreover that on the space of probability distributions on A the preferences are continuous and satisfy the independence axiom. Then they can be represented by a von Neumann-Morgenstern expected utility function (see Mas-Colell, Whinston, and Green (1995, Ch.6)). If $\mu(\cdot)$ is a probability distribution on A, where $\mu(a) = \Pr(a)$, then $u(\mu(\cdot)) = \sum_a u(a)\mu(a)$. Thus, for any probability distribution utility is determined by the utilities of the elements of A. In what follows we work with von Neumann-Morgenstern utility functions.

Games

A *game* G in normal form is a tuple $(N, \{S_i\}_{i=1}^n, \{u_i(\cdot)\}_{i=1}^n)$, where $N = \{1, ..., n\}$ is an arbitrary finite set of players, S_i is an arbitrary set of strategies of player i, and $u_i(\cdot)$ is the (von Neumann-Morgenstern) payoff (utility) function of player i from the set of strategy profiles $\bar{S} = \times_{i=1}^n S_i$ to the real line. Denote by $s_i \in S_i$ a typical strategy (action) of player i and by $s = (s_1, ..., s_n) \in \bar{S}$ a typical strategy profile. The strategy sets are often extended to the sets of mixed strategies ΔS_i, which is the sets of probability distributions over S_i. Then the set of (independent) mixed strategy profiles is $\Delta\bar{S} = \times_{i=1}^n \Delta S_i$. A pure strategy s_i corresponds to the probability distribution having unit weight on s_i. A completely mixed strategy is a probability distribution having non-zero weights on all $s_i \in S_i$. Since the payoff function is von Neumann-Morgenstern, it extends to mixed strategies in the straightforward way and is determined by its values on pure strategy profiles. A typical element of $\Delta\bar{S}$ is denoted by $\sigma = (\sigma_1, ..., \sigma_n)$, where $\sigma_i \in \Delta S_i$.

A *decision problem* is a game with one player that differs from the above description in that the payoff function of the player depends not only on the strategy but also on the state of the world. If $\Omega = \{\omega_1, ..., \omega_k\}$ is the finite set of states of the world, and $\mu \in \Delta\Omega$ is a probability distribution over it specifying the probability of each state ($\mu_i = \Pr(\omega_i)$), the payoff function $u : \bar{S} \times \Omega \to \mathbb{R}$ specifies what the player gets in each state ω given strategy s.

Denote by $u_i(\sigma_i, \sigma_{-i})$ the payoff of player i when player i uses strategy σ_i and other players use strategies $(\sigma_1, ..., \sigma_{i-1}, \sigma_{i+1}, ..., \sigma_n) =: \sigma_{-i} \in \Delta S_{-i} := \times_{j=1, j\neq i}^n \Delta S_j$. Strategy σ_i dominates strategy σ_i' if $u_i(\sigma_i, \sigma_{-i}) > u_i(\sigma_i', \sigma_{-i})$ $\forall \sigma_{-i}$. Strategy σ_i weakly dominates strategy σ_i' if $u_i(\sigma_i, \sigma_{-i}) \geq u_i(\sigma_i', \sigma_{-i})$ $\forall\sigma_{-i}$ and $\exists \sigma_{-i}$ such that $u_i(\sigma_i, \sigma_{-i}) > u_i(\sigma_i', \sigma_{-i})$. Strategy σ_i is *dominant* if σ_i dominates any other $\sigma_i' \in \Delta S_i$. Strategy σ_i is *dominated* if $\exists \sigma_i'$ such that σ_i' dominates σ_i. Strategy σ_i is *weakly dominated* if $\exists\sigma_i'$ such that σ_i' weakly dominates σ_i. Strategy σ_i is a *best response* to $\sigma_{-i} \in \Delta S_{-i}$ if $u_i(\sigma_i, \sigma_{-i}) \geq u_i(\sigma_i', \sigma_{-i})$ $\forall\sigma_i' \in \Delta S_i$. The set of all best responses to a given σ_{-i} is denoted by $BR(\sigma_{-i})$. A strategy profile $\sigma = (\sigma_1, ..., \sigma_n) \in \Delta\bar{S}$ is a *Nash equilibrium* of the game if $u_i(\sigma_i, \sigma_{-i}) \geq u_i(\sigma_i', \sigma_{-i})$ $\forall\sigma_i' \in \Delta S_i, \forall i \in N$, that is all players play best responses. If a strategy profile $\sigma = (\sigma_1, ..., \sigma_n)$ is such that for every player i, $\sigma_i \in BR(\sigma_{-i})$ we write $\sigma \in BR(\sigma)$. Thus, σ is a Nash equilibrium if $\sigma \in BR(\sigma)$. A Nash equilibrium σ is *pure* if $\sigma \in \bar{S}$. A Nash equilibrium σ is *strict* if $u_i(\sigma_i, \sigma_{-i}) > u_i(\sigma_i', \sigma_{-i})$ $\forall\sigma_i' \in \Delta S_i, \forall i \in N$.

A game is symmetric if for any players i, j $S_i = S_j$ and for any permutation $\pi : N \to N$ $u_i(s_1, ..., s_n) = u_{\pi(i)}(s_{\pi(1)}, ..., s_{\pi(n)})$. For symmetric games we will often write $G = (N, S, u)$

since $S := S_1$ is the same for all players and $u_1 =: u$ determines u_j for all j.

These are the basic notions of game theory that are used often in the thesis. The following notions are not used often therefore they are given in an informal way, with references to the formal definitions.

A game in normal form assumes that each player chooses a strategy simultaneously with other players. A *game in extensive form* specifies which player moves at any moment of time and what information he has (for a formal definition see Mas-Colell, Whinston, and Green (1995, Ch.7)). For each game in extensive form there exists a normal form representation of it. An equilibrium of this normal form representation is an equilibrium of the extensive form game. An extensive form game is usually represented as a tree. A *subgame perfect equilibrium* is a strategy profile that is an equilibrium in each subgame of the extensive form game, that is, in each subtree satisfying certain properties (see Mas-Colell, Whinston, and Green (1995, Ch.9)).

In a *game with incomplete information*, or a *Bayesian game* each player has a payoff function that depends on the *type* of the player that is a random variable. The probability distribution on types is known to all players. In an equilibrium of a Bayesian game players play mutual best responses for every realization of their types (the formal definition is in Mas-Colell, Whinston, and Green (1995, Ch.8)).

Stochastic processes

The dynamical systems considered in this thesis will consist of the following elements. Let $(\Omega, \mathcal{F}, P)$ be a probability space. A *discrete time stochastic process* is a sequence of random variables $\{X_t\}_{t=0}^{\infty}$ on the probability space, taking values in a set Z. The set Z is the state space of the process. We will work with either finite Z, countable Z, or $Z = \mathbb{R}^k$.

The canonical construction of a stochastic process takes as Ω the space of all sample paths, that is, $\Omega = Z^\infty = \{z = (z_0, z_1, ...), z_t \in Z\}$. If Z is finite, the σ-field $\mathcal{F}$ is the smallest σ-field containing all events $\{z \in \Omega : z_0 \in Z, ..., z_m \in Z\}$. If $Z = \mathbb{R}^k$, the σ-field $\mathcal{F}$ is the smallest σ-field containing all events $\{z \in \Omega : z_0 \in B_0, ..., z_m \in B_m\}$, where B_i's are Borel subsets of $\mathbb{R}^k$. The probability measure P on $(\Omega, \mathcal{F})$ extends from the set of all finite sequences by Kolmogorov's Existence Theorem (Bhattacharya and Waymire (1990, Ch.1)). The stochastic process is given by $X_t(z) = z_t$. Observe that the probability measure defines probabilities for infinite events thus we can speak of the asymptotic behavior of the process.

A sequence of random variables $\{X_t\}_{t=0}^{n}$ on the probability space $(\Omega, \mathcal{F}, P)$ is said to *converge in probability* to a random variable X if $\forall \varepsilon > 0 \lim_{t\to\infty} \Pr(\{|X_t - X| > \varepsilon\}) = 0$. A sequence $\{X_t\}_{t=0}^{n}$ is said to *converge almost surely* to X if $\forall \varepsilon > 0 \Pr(\{\lim_{t\to\infty} |X_t - X| > \varepsilon\}) = 0$. Convergence almost surely is a stronger notion than convergence in probability.

A stochastic process $\{X_t\}_{t=0}^{\infty}$ is a *Markov process,* if the conditional distribution of X_{t+1} given $X_0, ..., X_t$ depends only on X_t. If Z is finite, we have $\Pr(X_{t+1} = z|X_0 = z_0, ..., X_t = z_t) = \Pr(X_{t+1} = z|X_t = z_t)$ for any $z \in Z$ and the process is called *Markov chain.* If the distribution of X_{t+1} given X_t depends only on X_t but not on t, the Markov process is *homogeneous.*

In other words, for a (homogeneous) Markov process, the dynamic is given by a function $f : \Delta Z \to \Delta Z$, where ΔZ is the set of probability distributions on Z. If μ_t is a probability distribution on Z at time t, the probability distribution at time $t+1$ is given by $\mu_{t+1} = f(\mu_t)$. Applying f recursively from a given initial probability distribution over states μ_0, the probability distribution at time t is given by $\mu_t = f^t(\mu_0)$. We will be mainly interested in the long run outcome of the dynamics, which is given by $\lim_{t\to\infty} f^t(\mu_0)$.

If set Z is finite, we have a finite Markov chain. In the following we talk about finite Markov chains. A Markov chain is described by a square matrix P of size $|Z| \times |Z|$, whose elements $p_{ij} = \Pr(z_{t+1} = z^j|z_t = z^i)$ we will also denote as $P_{z_i z_j}$. Denote by μ_t a probability distribution on Z at time t, represented as a row vector of size $|Z|$. Matrix P determines a probability distribution at time $t+1$ in the straightforward way $\mu_{t+1} = \mu_t P$. The probability distribution μ is *stationary* if $\mu P = \mu$. If a stationary probability distribution μ places unit weight on a state z, then z is a *stationary state.* A state z' is *accessible* from state z if $\exists n$ such that $P^n_{zz'} > 0$.

A subset of states $W \subset Z$ is a *recurrent class* if $\forall z, z' \in W$ z' is accessible from z, and $\forall z'' \notin W$ and $\forall z \in W$, z'' is not accessible from z. A stationary state is always a recurrent class. A Markov chain is *irreducible* if $\exists n$ such that P^n does not have any element equal to 0. An irreducible Markov chain has only one recurrent class, the whole space Z. If a Markov chain is irreducible it has a unique stationary distribution $\tilde{\mu}$. Moreover, starting from any initial distribution μ_0, the probability distribution as the process progresses, $\mu_0 P^t$ converges to $\tilde{\mu}$ in probability and $\tilde{\mu}$ then corresponds to the proportion of time the system spends in each state in the long run, that is, the process is *ergodic.*

Chapter 2

Simple Reinforcement Dynamic

2.1 Introduction

This chapter analyses one of the simplest dynamic models of strategy adjustment, the so called reinforcement dynamic. The adjustment is very elementary and makes use only of information about a player's own payoff. The main assumption behind the model is that the agents are ignorant about the precise structure of interaction or, alternatively, they do not use all the information they are required to use by the standard game theory. The main question is whether this simple adjustment process will nevertheless with time converge to one of the standard game or decision theoretic solutions.

The reinforcement dynamic stems from theoretical and experimental research in psychology (e.g. Bush and Mosteller (1955)) and experimental works in economics (e.g. Arthur (1993) and Roth and Erev (1995)). These papers show that real human behavior can be approximated by such an adjustment process. They argue that people playing a (complex) game do not necessarily perform sophisticated calculations and reasoning to find the (game-theoretic) solution. Instead, if the game is played repeatedly many times, agents learn to play the game with a simple learning dynamic. The reinforcement dynamic is an example of such learning.

The behavior of the subjects in the experiments does not necessarily converge to the optimal action or to a Nash equilibrium (or its refinement) in some of the examples Arthur (1993) and Roth and Erev (1995) consider. Together with experimental results, Arthur (1993) offers convergence results for some versions of the reinforcement dynamic in decision problems. In this chapter we report analytical results about convergence to the optimal action for other versions of the dynamic. Although the dynamic is formulated quite simply, the analysis of it requires sophisticated tools of stochastic optimization.

We compare several different versions of the reinforcement dynamic, since modifications of the dynamic lead to different convergence results. The focus is not only on the long run properties (convergence, optimality), but also on the medium run since it can be of more economic importance. The environment in an economy is not likely to be constant indefinitely. Thus the issue of the speed of convergence will play a role in our analysis. There is a trade-off between the long run convergence to the optimal action and the speed of convergence, which will be analyzed by means of computer simulations.

Convergence results for the reinforcement dynamic in decision problems can, and have been, generalized for games. It has been shown for some versions of the dynamic that the dynamic converges to any of the strict Nash equilibria with a positive probability (Posch (1997), Laslier et al. (1999)). In this chapter we focus rather on the selection properties of the dynamic in the case of multiple equilibria. Though any of the equilibria has a non-zero probability to be observed in the long run and in the medium run, some of them are more likely to be observed. Using the version of the dynamic that performs best in decision problems, by means of computer simulations we try to determine which equilibrium is more likely to be observed in the medium run and in the long run.

We demonstrate that there is a relation between the concept of risk dominance between equilibria (Harsanyi and Selten (1988)) and the medium run results of the dynamic in 2×2 games. Risk dominance is defined only for 2×2 games; in larger games we define another concept of central, or egalitarian equilibrium. We show that in some games in the medium run the egalitarian equilibrium is observed rather often though in the long run the subgame perfect equilibrium starts to play a larger role.

The remainder of the chapter is structured as follows. Section 2.2 describes the model, Section 2.3 gives the analysis for the case of one-player decision problem, Section 2.4 reports results for some games and Section 2.5 concludes. Most of the simulations results are gathered in Appendices.

2.2 The Model

The model is formulated for games though it can easily be simplified to one-player decision problems. The description of the model follows Arthur (1993) and Roth and Erev (1995).

There are n players. The set of players is denoted by N. The stage game G, that is, the game played in each period is a game in normal form. Thus $G = (N, \{S_i\}_{i=1}^n, \{\pi_i(\cdot)\}_{i=1}^n)$, where S_i are the (finite) sets of pure strategies for each player i, $\bar{S} = \times_{i=1}^n S_i$, and π_i :

$\bar{S} \to \mathbb{R}$ are the payoff functions. The payoff functions are assumed to be positive for every player and for every profile of pure strategies, that is $\pi_i(s) > 0 \ \forall i \in N, \forall s \in \bar{S}$.

Let player i have k pure strategies, $|S_i| = k$. Time is discrete, $t = 1, 2, \ldots$. The state of player i at time t is described by the vector $q_i^t = (q_{i1}^t, \ldots, q_{ik}^t) \in \mathbb{R}^k$. The vector q_i^t is the vector of *propensities*, that is q_{ij}^t denotes the propensity of player i to play strategy $j \in S_i$ at time t. The propensities are assumed to be strictly positive, $q_{ij}^t > 0 \ \forall t, \forall i, \forall j$.

Let us define Q_i^t as the sum of the propensities for player i at time t, $Q_i^t = \sum_{j=1}^k q_{ij}^t$. Given the vector of the propensities, the *probability* to play strategy j at time t is defined as

$$p_{ij}^t = \frac{q_{ij}^t}{Q_i^t} \tag{2.1}$$

The vector $p_{ij}^t = (p_{i1}^t, \ldots, p_{ik}^t) \in \Delta S_i$, where ΔS_i is the set of mixed strategies of player i. Working with propensities rather than with probabilities is easier since probabilities have to be kept between 0 and 1 while from the only restriction of being positive on the propensities it follows that the probabilities defined in the way described above are between 0 and 1.

The state of the whole process at time t is determined by the state of each player, that is by the vectors $q_1^t, \ldots, q_n^t$. According to the vector of probabilities derived from the q's, each player chooses a pure strategy to play in the present period. The stage game G is then played with the chosen pure strategies. Though the goal of the investigation is the dynamic of probabilities p_i^t, we work with the propensities q_i^t through which the probabilities are determined.

There are two main requirements we pose on the dynamic. First, the dynamic should be *reinforcing*: if a strategy is played and the payoff is "satisfactory", the probability to play this strategy should increase. Second, the dynamic should be in a sense *simple*: no complex functions must be involved. One class of the dynamics satisfying these requirements is following. A player triggers a strategy, observes his payoff and adds the payoff to the propensity of playing this strategy. Then the player renormalizes the propensities by multiplying them by a certain variable.

In the model above all payoffs are positive, which means that the propensity of playing a strategy increases after the strategy has been played. The probability of this strategy increases as well. Thus, all payoffs are satisfactory and all trials are reinforcing. Since the payoff is added to the propensity, the function involved is the simple, linear. Finally, renormalization by multiplication is not complicated either.

The renormalization will play an important role. Since it multiplies all propensities by the same number it does not change the probabilities in the current period but it does change the expected motion of probabilities. This will influence the convergence results and the speed of convergence. We try to answer the following questions: for which forms of normalization the dynamic converges to an equilibrium, and for which forms the speed of the dynamic is fast enough to achieve good results quickly.

Formally, assume that at time $t = 1$ the vectors of initial propensities q_i^1 are given, $q_{ij}^1 > 0 \ \forall i \in N, \forall j \in S_i$. If player i chooses strategy j at time t while the other players choose strategies $s_{-i} \in S_{-i}$, then the state of the player i at time $t+1$ is defined as

$$\begin{aligned} q_{ij}^{t+1} &= (q_{ij}^t + B_{ij}^t) A_i^t \\ q_{ik}^{t+1} &= q_{ik}^t A_i^t, \ k \neq j \end{aligned} \tag{2.2}$$

where $B_{ij}^t = \pi_i(j, s_{-i})$ is the realized payoff at period t, A_i^t is the normalizing multiplier. Since the initial propensities are assumed to be positive, and the payoffs are positive, the propensities at period t are positive as well. The normalizing multiplier can be chosen such as to keep the sum of the propensities equal to a predetermined variable C_i^t, in which case $A_i^t = \frac{C_i^{t+1}}{Q_i^t + B_{ij}^t}$. This type of normalization was used in Arthur (1993), is mathematically tractable, and gives a rich variety of results since C_i^t determines the speed of adjustment.

The variable C_i^t can be deterministic, for example, $C_i^t = Ct^\nu$, where C, ν are given constants. This case will be called the *normalization* case. In the case of *no normalization* $A_i^t = 1$ or, equivalently, $C_i^{t+1} = Q_i^t + B_{ij}^t$, which is a random variable since the payoff B_{ij}^t is generally random. Yet another interesting case is the so called *forgetting* case, when $A_i^t = \delta < 1$, that is, $C_i^{t+1} = \delta(Q_i^t + B_{ij}^t)$. The parameter δ is a forgetting parameter since the payoff that was obtained τ periods ago enters the sum of the propensities with factor δ^τ, that is, such a payoff is partially forgotten.

The use of variable C_i^t in the analysis is justified by the observation, as it will be clear later, that C_i^t will determine the step size of the dynamic, which is important for convergence. In fact, it will be shown in the subsequent section that the inverse of C_i^t has the same order as the step size of the dynamic.

Observe that in equation (2.2) the propensities at time $t+1$ depend only on propensities at time t, on realized payoff at time t, and on A_i^t that in turn depends only on variables at time t as well. Therefore, the dynamic process on propensities is a Markov process, though not necessarily a homogeneous one (for example, in the case $C_i^t = Ct^\nu$ it is not). The state space of the process is $\mathbb{R}^{k_1 + \ldots + k_n}$, where k_i is the number of strategies of player i. In distinction from that, the process on probabilities (2.1) is not a Markov process since probabilities at time t are not sufficient to determine probabilities at time $t+1$, since they depend on propensities.

As usual in evolutionary and learning theories of games, two interpretations of the above model are possible. The first one is that each player i possesses at every moment of time t a mixed strategy, characterized through the propensities q_i^t by the vector of probabilities p_i^t. The other interpretation is that for every role of players $1, ..., n$ there is a large population of agents each possessing a pure strategy but the distribution of pure strategies in each population i is given by the vector p_i^t and than the evolutionary dynamic of changes in the composition of the populations is given by (2.2). We think that in the context of the reinforcement dynamic, the first interpretation is more plausible. That is, though the evolutionary interpretation is possible, we believe that the reinforcement model describes learning by an individual player, rather than evolution of a population.

Though the system is described in terms of a mixed strategy, each period a pure strategy is played. The precise mechanism of this choice is not modeled. In the first interpretation above it can be a randomizing device used by a player. In the second interpretation such a choice can be understood as a random draw of an agent with a pure strategy from a population.

Notice that the updating of propensities and, therefore, of mixed strategies, depends only on a player's own payoff. The model can be applied also to one-player decision problems with random distributions of payoffs. This distribution is unknown by the players. The players behave as if the distribution is stationary and the strategy that was good yesterday will be good today too. Though the formulation of the model does not restrict the dynamic to stationary distributions, in the decision problems we analyze we assume stationarity. In games the distribution of other players' strategies is not stationary but is changing according to the dynamic. The justification of carrying the dynamic to games can be that agents may not know that they are participating in a game, or do not perceive the complex interdependence between their own payoff and other agents' actions, or do not know who the opponents are, or the preferences of the opponents. Of course, in economic situations agents may have some idea of what is going on, or form subjective expectations, but we will consider the extreme case of ignorance of strategic interaction. The model can also be applied easily to extensive form games where not all information sets are reached during the course of the game, because players do not need to know the other players' strategies.

The normalization case with $C_i^t = Ct^\nu$ was considered in Arthur (1993) for decision problems and in Posch (1997) for 2×2 games. These papers analyze the convergence properties of the dynamic and show that if $t < 1$ then the dynamic does not necessarily converge to the optimal action or equilibrium while if $t = 1$ the dynamic does converge to it. The no normalization case and the forgetting case were introduced in Roth and Erev (1995) who did not consider long run convergence. Laslier et al. (1999) consider convergence properties of the no normalization case in decision problems and games showing that the dynamic converges to the optimal action or equilibrium. We supply these results with the result for the forgetting version.

However, the issue of the speed of convergence and medium term results were largely neglected in the papers mentioned above (except Roth and Erev (1995)). In economic situations, when the environment changes, fast convergence may be more important. We address the questions of how the versions of the reinforcement dynamic perform in the medium run for both stochastic problems and games by performing computer simulations.

Roth and Erev (1995) point out that the dynamic captures two important aspects of learning, derived from the psychological literature. The first one, the Law of Effect, states that the choices that have led to good outcomes should be repeated more often in the future. In the model, since every strategy gives a positive payoff, the probability of playing it in the next round increases, hence the law is fulfilled. The second aspect, the Power Law of Practice, says that learning tends to be fast in the beginning and then slows down. In the no normalization case ($A_i^t = 1$) the propensities increase with time. If the forgetting parameter δ is close enough to 1, this is also true in the forgetting case for quite a long initial time. For the normalized case ($C_i^t = Ct^\nu$) with $C, \nu > 0$, C_i^t grows with time too. Payoffs at later stages change the probabilities less than at earlier stages, when the propensities are not yet very large, so learning indeed slows down as time progresses.

Another interpretation of the dynamic can be in the spirit of learning dynamics with an aspiration level. In such dynamics a strategy is regarded as successful and its weight increases if it gives a payoff that is higher than the aspiration level. The aspiration level may be fixed or variable. Our model can be understood as having a fixed aspiration level not larger than 0, hence every strategy is successful, or, at least, not unsuccessful. A similar model for decision problems, employing a variable aspiration level, is considered in Börgers and Sarin (1996), where the probabilities change directly, not through propensities. A description of a model with variable aspirations for 2×2 games can be found in Karandikar et al. (1998).

The model of reinforcement learning can be applied also to learning by automata. An extended analysis of machine reinforcement learning is provided in Narendra and Thathachar (1989). They consider mostly schemes of reward-penalty nature, which require knowledge of the maximal and minimal payoffs. If the realized payoff for an action is close to the maximal one, the probability of playing this action increases (reward) while if the payoff is close to the minimal one, the probability decreases (penalty). Our scheme can be considered as a reward-reward scheme since the probability of playing a strategy increases independently of the outcome.

2.3 One-Player Decision Problems

Though the ultimate goal is to analyze the dynamic in games, it is interesting and useful to consider its behavior in one-player decision problems. In decision problems the distribution of payoffs to an action is random and unknown but is not changing as it happens in games when all players are learning simultaneously. Binmore et al. (1996) expressed the view that it might be wiser to consider human behavior in decision problems first and only then apply it to games. We will proceed in this spirit.

2.3.1 The Long Run Convergence

One of the main questions concerning every dynamic is whether it converges and if so, to what point or set of points. For games the Nash equilibrium concept serves as a benchmark thus it is desirable that a dynamic converges to a Nash equilibrium and, if there are multiple Nash equilibria, to a certain refined equilibrium, for example, to a subgame perfect equilibrium. In one-player decision problems the concept of Nash equilibrium corresponds to the choice of the action that gives the highest expected payoff. However, the lock-in problem of path dependency may arise. Path dependency refers to the phenomenon when initial choices influence the long run outcome. It is not a priori clear that the optimal action will be chosen in the long run.

In this subsection we make an overview of results for different versions of the dynamic. We report recent results regarding convergence of the dynamic to the optimal action in decision problems for the normalization and the no normalization cases, supplied also by a result for the forgetting case. The normalization and no normalization results are taken from Arthur (1993), Posch (1997), and Laslier et al. (1999).

In the model described in the previous section, we have now a decision problem, so $n = 1$ and we can omit subscript i in the formulas. The payoff to strategy j, denoted by $\pi_j(\omega)$, depends on the state of the world ω that is a random variable drawn each period after a strategy was chosen by the decision maker. The environment is the set of the states of the world $\Omega = \{\omega_1, ..., \omega_l\}$, assumed to be finite, and a stationary probability distribution μ on Ω that gives the probability of each state of the world. The decision maker has k strategies, or actions.

Suppose the decision maker plays strategy j at time t. Denote the realization of $\pi_j(\omega)$ by B^t. Let e_j be the unit k-vector with 1 on the jth place and let $b^t = B^t e_j$. Thus, b^t is the vector of payoffs to all k strategies at period t. Since strategies other than j were not tried, they did not get any payoff. The formulas for the dynamic of the propensities from the previous section can be rewritten in vector form with the multiplier expressed in the form of a given sum of the propensities:

$$q^{t+1} = (q^t + b^t)\frac{C^{t+1}}{Q^t + B^t}. \tag{2.3}$$

Notice that after normalization $Q^{t+1} = C^{t+1}$. Rearranging terms,

$$\frac{q^{t+1}}{Q^{t+1}} = \frac{q^t}{Q^t + B^t} + \frac{b^t}{Q^t + B^t} = \frac{q^t}{Q^t} - \frac{B^t q^t}{(Q^t + B^t)Q^t} + \frac{b^t}{Q^t + B^t}. \tag{2.4}$$

Since $\frac{q^t}{Q^t} = p^t$, it can be rewritten as

$$p^{t+1} = p^t + a^t(b^t - B^t p^t), \tag{2.5}$$

where $a^t = \frac{1}{(Q^t + B^t)}$ determines the step size of the process. Assume that the payoffs are bounded and denote by M the upper bound and by m the lower bound. Since the set of states and the set of strategies are finite, M can be taken as the maximal payoff over states and strategies and m as the minimal payoff over states and strategies. Then B^t is bounded for each t. Whether a^t is bounded and its asymptotic (as $t \to \infty$) behavior is determined by $\frac{1}{Q^t}$. In the normalization case $Q^t = C^t = Ct^\nu$, hence $a^t = O(t^{-\nu})$. In the non-normalized dynamic by definition $Q^1 + (t-1)m \leq Q^t \leq Q^1 + (t-1)M$, hence $a^t = O(t^{-1})$. For the case with forgetting, observe that $Q^1\delta^{t-1} + m\sum_{\tau=1}^{t-1}\delta^\tau \leq Q^t \leq Q^1\delta^{t-1} + M\sum_{\tau=1}^{t-1}\delta^\tau$. After calculating the sum of the geometric series, we have $Q^1\delta^{t-1} + m\delta\frac{1-\delta^{t-1}}{1-\delta} \leq Q^t \leq Q^1\delta^{t-1} + M\delta\frac{1-\delta^{t-1}}{1-\delta}$. Hence, $a^t = O(\frac{1}{1-\delta^{t-1}})$.

Assume that the kth action is the unique optimal action, that is, it has the largest expected payoff. To have some idea about the expected motion of probabilities, we can look at $E[p_k^{t+1}|q^t]$. Since the kth action is optimal, one may wonder if, starting from some time, the expected probability of playing the optimal action increases. The actual probability, of course, does not have to increase.

Proposition 2.1 *In the case of normalization ($C^t = Ct^\nu, C, \nu > 0$) and also in the case of no-normalization ($A^t = 1$) it holds that $\exists T$ such that $\forall t > T$ $E[p_k^{t+1}|q^t] > p_k^t$.*

Proof. From the expression for probabilities (2.5) $E[p_k^{t+1}|q^t] = p_k^t + E[(a^t(b^t - B^t p^t))_k]$. The last term can be rewritten as $E[\frac{b_k^t - B^t p_k^t}{Q^t+B^t}] = p_k^t E[\frac{b_k^t - B^t p_k^t}{Q^t+B^t}|\{k\}] + (1-p_k^t)E[\frac{b_k^t - B^t p_k^t}{Q^t+B^t}|\{-k\}]$, where $\{k\}$ denotes the event that action k is played, and $\{-k\}$ denotes the event that any other action is played. By definition of b^t and B^t it can be further rearranged as $E[\frac{b_k^t - B^t p_k^t}{Q^t+B^t}] = p_k^t E[\frac{B^t(1-p_k^t)}{Q^t+B^t}|\{k\}] + (1-p_k^t)E[\frac{-B^t p_k^t}{Q^t+B^t}|\{-k\}] = p_k^t(1-p_k^t)(E[\frac{B^t}{Q^t+B^t}|\{k\}] - E[\frac{B^t}{Q^t+B^t}|\{-k\}])$. The first two terms in the product are positive. The last term can be rewritten as $E[B^t|\{k\}] + E[-B^t\frac{Q^t+B^t-1}{Q^t+B^t}|\{k\}] - E[B^t|\{-k\}] - E[-B^t\frac{Q^t+B^t-1}{Q^t+B^t}|\{-k\}]$, or, equivalently, as $(E[B^t|\{k\}] - E[B^t|\{-k\}]) - (E[B^t\frac{Q^t+B^t-1}{Q^t+B^t}|\{-k\}] - E[B^t\frac{Q^t+B^t-1}{Q^t+B^t}|\{k\}])$. Consider the last parenthesis. Notice that, if Q^t is large enough, $1-\varepsilon \le \frac{Q^t+B^t-1}{Q^t+B^t} < 1$ $\forall \varepsilon > 0$. Then $E[B^t\frac{Q^t+B^t-1}{Q^t+B^t}|\{-k\}] - E[B^t\frac{Q^t+B^t-1}{Q^t+B^t}|\{k\}] < E[B^t|\{-k\}] - E[B^t(1-\varepsilon)|\{k\}] = E[B^t|\{-k\}] - E[B^t|\{k\}] + \varepsilon E[B^t|\{k\}]$. Substituting into the previous expression, we have $E[\frac{B^t}{Q^t+B^t}|\{k\}] - E[\frac{B^t}{Q^t+B^t}|\{-k\} > 2(E[B^t|\{k\}] - E[B^t|\{-k\}]) - \varepsilon E[B^t|\{k\}]$. Since ε is arbitrarily small, and $E[B^t|\{k\}] - E[B^t|\{-k\}] > 0$ by the optimality of action k, the last expression is positive and $E[p_k^{t+1}|p^t] > p_k^t$. We used that Q^t is large enough, which is the case for both non-normalized and normalized models. ■

It can be easily shown that if Q^t is not large, the statement that $E[p_k^{t+1}|q^t] > p_k^t$ does not have to hold.

Example 2.1 *Consider the following decision problem with two states of the world and two actions*

	ω_1	ω_2
s_1	3	$\frac{1}{2}$
s_2	1	1

$, \Pr(\omega_1) = \Pr(\omega_2) = \frac{1}{2}.$

Strategy s_1 is optimal. Consider $q_1^t = q_2^t = \frac{1}{8}$. By straightforward calculations $E[p_1^{t+1}] = \frac{389}{780} \approx 0.4987 < 0.5 = p_1^t$.

In the case with forgetting Q^t is bounded from above by $Q^1\delta^{t-1} + M\delta\frac{1-\delta^{t-1}}{1-\delta}$, thus we cannot say that $\exists T$ such that $\forall t > T$ $E[p_k^{t+1}|q^t] > p_k^t$ since Q^t may not grow enough. Though a useful insight, saying that on average the probability of the optimal action increases, Proposition 2.1 does not necessarily mean that the optimal action is played in the limit as $t \to \infty$ with probability 1.

To establish (or disprove) the last statement we will make use of the notion of almost sure convergence.

Definition 2.1 *A sequence of random variables* $\{p^t\}_{t=0}^{\infty}$ *is said to* ***converge almost surely*** *to a random variable* p *if* $\Pr(\{\lim_{t\to\infty} |p^t - p| < \varepsilon\}) = 1 \ \forall \varepsilon > 0$.

Convergence almost surely also implies convergence in probability. Thus, if we show that $\{p^t\}$ in our model converges to a distribution placing unit weight on the optimal action, it would mean that the optimal action is played in the long run with probability 1.

The following results show when the process (2.5) does converge to the state when the optimal action is played with probability 1 and when it does not. Recall that k is the unique optimal action.

Theorem 2.1 *In the normalization case* ($C^t = Ct^\nu, C, \nu > 0$)*, if* $\nu < 1$ *or* ($\nu = 1$ *and* $C < m$)*,* $\Pr(\{\lim_{t\to\infty} |p_k^t - 1| < \varepsilon\}) < 1$*, that is, a non-optimal action is played in the limit* $t \to \infty$ *with non-zero probability.*

The theorem says that the process does not converge almost surely to the optimal action. The proof of the $\nu < 1$ case is in Arthur (1993), the proof of the ($\nu = 1$ and $C < m$) case is in Posch (1997). The proofs are similar to the one below for Theorem 2.4.

Theorem 2.2 *In the normalization case* ($C^t = Ct^\nu, C, \nu > 0$)*, if* $\nu = 1$ *and* $C \geq m$*,* $\Pr(\{\lim_{t\to\infty} |p_k^t - 1| < \varepsilon\}) = 1$*, that is, the probability of the optimal action converges to* 1 *almost surely.*

The proof is contained in both Arthur (1993) and Posch (1997). In the proof the stochastic dynamic is approximated by a deterministic one, based on the expected motion of the dynamic, whose convergence properties are easier to analyze. These properties carry over to the stochastic dynamic only if the step size a^t becomes small sufficiently fast which is the case when $\nu = 1$.

Theorem 2.3 *In the no normalization case* ($A^t = 1$)*,* $\Pr(\{\lim_{t\to\infty} |p_k^t - 1| < \varepsilon\}) = 1$*, that is, the probability of the optimal action converges to* 1 *almost surely.*

Proof. We have seen that in this case $Q^1 + (t-1)m \leq Q^t \leq Q^1 + (t-1)M$, which can be rewritten as $mt \leq Q^t \leq (Q^0 + M)t$. Then the dynamic is equivalent, in the sense of having the same asymptotic behavior, to the normalization case with $\nu = 1$ and $C \geq m$. By Theorem 2.2 the process converges to the optimal action[1]. ∎

[1] An alternative proof of this theorem can be found in Laslier et al. (1999).

Theorem 2.4 *In the case with forgetting ($A^t = \delta < 1$),* $\Pr(\{\lim_{t\to\infty} |p_k^t - 1| < \varepsilon\}) < 1$, *that is, the probability of playing a non-optimal action is positive in the limit $t \to \infty$.*

Proof. The proof is along the lines of the similar proof for the normalization case given in Arthur (1993). We have seen that $Q^1\delta^{t-1} + m\delta\frac{1-\delta^{t-1}}{1-\delta} \le Q^t \le Q^1\delta^{t-1} + M\delta\frac{1-\delta^{t-1}}{1-\delta}$, that is Q^t has the order of $1-\delta^{t-1}$. To proof the non-optimality consider the event that an inferior action j is triggered from time t on. Denote this event by D^t. We need to show that $\Pr\{D^t\} = \Pi_t^\infty p_j^t > 0$. Let $a_j^t = 1 - p_j^t$. Since $0 < a_j^t < 1$ the convergence of the infinite product $\Pi_t^\infty(1 - a_j^t)$ is necessary for the convergence of the series $\sum_t^\infty a_j^t$. From the process equation (2.5), the dynamic for action j when it is played at all times from t on can be written as $p_j^{t+1} = p_j^t + \frac{B^t(1-p_j^t)}{Q^t+B^t}$. Then $a_j^{t+1} = a_j^t(1 - \frac{B^t}{Q^t+B^t})$. The ratio of two successive terms is then $\frac{a_j^{t+1}}{a_j^t} = 1 - \frac{B^t}{Q^t+B^t}$. Given that Q^t has the order of $1-\delta^t$ the last expression can be rewritten as $\frac{a_j^{t+1}}{a_j^t} = 1 - \frac{B^t}{c_1+c_2(1-\delta^{t-1})+B^t} < 1 - \frac{B^t}{c_1+c_2+B^t}$ for some constants $c_1, c_2 > 0$. It means that a_j^t decreases faster than a geometric series and therefore converges. Therefore, the probability of playing an inferior action j from time t on is positive and the theorem is proven. ■

The long run results are very different for different specifications of the dynamics. Roughly speaking, if the propensities grow fast enough, as in the normalized version with $\nu = 1$, or in the non-normalized version, the optimal action is found in the long run. With forgetting, it is not necessarily so. However, for economic relevance one should look also at the speed of convergence and medium run results. It may well be that the learning algorithm that eventually discovers the optimal action is too slow to achieve good results in the medium run and may be inferior in that respect to a non-optimal learning algorithm. The next subsection presents some simulation results comparing different variations of the model in the medium run.

2.3.2 The Speed of Convergence and the Probability of Convergence to the Optimal Action

In any stochastic decision problem there is a trade-off between the speed of convergence and long run optimality. If the speed of learning is high, there is a probability of locking in in an inferior action, while if learning is slow, the optimal action will be eventually found. For example, in the normalization case, if $\nu < 1$ learning is quick (the step size of the process is large), so the process converges to an action quickly but it does not have to be the optimal action. On the contrary, if $\nu = 1$ learning is slow enough to achieve optimality. However, even in the case with $\nu < 1$, changing the other parameter C can make the probability of playing the optimal action as close to 1 as one desires, that is, the learning scheme can be made ε-optimal[2] (Narendra and Thathachar (1989)).

Instead of pursuing this approach, or deriving the expected rate of convergence analytically, we adopt a simulation method. The method allows to derive some stylized facts on the basis of which certain conclusion can be made. To illustrate the issues of both non-optimal convergence and the speed thereof, we choose plausible parameter values for the dynamics and run simulations[3] for a couple of decision problems. The version that performs better in the sense of average realized payoff over a number of simulations for a given period will be said to be simulation-better.

For a given decision problem and a given initial propensities of the actions, a simulation calculates the propensities for periods $1, ..., T$ according to (2.2) using a pseudorandom numbers generator to determine which action j is played and which state of nature is realized at period t. The associated probabilities are calculated by (2.1). We know the stream of realized payoffs and can compare it with another stream for a different version of the dynamic.

The definition of simulation-better dynamic for a given decision problem and a given number of periods is as follows. Denote two versions of the dynamic by P_1 and P_2. Denote by $x_i^t(P_j)$ the realized payoff for dynamics P_j at period t in simulation i.

Definition 2.2 *Let a decision problem G, a number N, and dynamics P_1 and P_2 be given. Then P_1 is N-**simulation-better** than P_2 for period T if the average (over N simulations) realized payoff up to period T, $\overline{x}_j^T = \frac{1}{N}\sum_{i=1}^N \frac{1}{T}\sum_{t=1}^T x_i^t(P_j)$ is larger for $j = 1$ than for $j = 2$.*

We will illustrate the method on the following two decision problems.

1)

	ω_1	ω_2
s_1	4	1
s_2	1	4

$, \Pr\{\omega_1\} = \frac{1}{3}, \Pr\{\omega_2\} = \frac{2}{3};$

[2] A scheme is said to be ε-optimal if $\forall \varepsilon > 0\ \exists T$ such that $\forall t > T$ the probability of playing the optimal action $p_k^t > 1 - \varepsilon$.

[3] A DOS program for running simulations was written by the author in the C programming language and is available upon request.

2)

	ω_1
s_1	2
s_2	3
s_3	2.5

$, \Pr\{\omega_1\} = 1.$

For the first decision problem the expected payoffs are $E(s_1) = 2, E(s_2) = 3$. Hence the second decision problem is the certain case of the first one plus an additional action. The second decision problem is devised to show how convergence slows down with more actions.

The sum of the initial propensities is set to $Q^1 = 30$. The number is chosen to be close to the estimated from human behavior by Arthur (1993). Roth and Erev (1995) in their simulations also choose this parameter with a similar magnitude, so we follow their path. They also argue that choosing a number of a similar order will not change the results considerably. Initial propensities q^1 for actions are equal to each other, that is $q_i^1 = 15$ in the first problem, and $q_i^1 = 10$ in the second one. This is chosen in order to give all strategies equal chances in the beginning. The simulations are run for three specifications of the dynamic, namely for the normalized version with $C = Q^1 = 30, \nu = 0$, for the non-normalized version ($A^t = 1$), and for the forgetting case with $\delta = 0.999$. The versions of the dynamic are fully specified.

The question remains of what should be taken as the medium run. We have chosen that the long run starts after 100,000 periods. The 100,000th period will be considered as the medium run. This is also one of the periods that in Roth and Erev (1995) distinguishes the medium run, when the process did not necessarily converge, from the long run, when the process already is in (or very close to) a stationary state. A justification for choosing 100,000 could be that if an action is chosen every hour, then 100,000 hours is approximately 11.5 years. This period, almost equal to a decade, in an economy can be roughly considered as having a constant environment. We do not consider the more difficult question of how the dynamic responds to a change in the environment but focus on the medium run results since they can be considered as a first step towards that more difficult question.

The results of the simulations are reported in Tables 2.1 and 2.2 for the problems 1), 2) correspondingly. The columns in Tables 2.1 and 2.2 show results of a hundred simulations for each of the three different specifications of the dynamic. The column "$\nu = 0$" is for the normalized case with $\nu = 0$; the column "no-norm" is for the non-normalized case; and the column "forgetting" is for the case with $\delta = 0.999$. The tables consist of three

Time		$\nu = 0$	no-norm	forgetting
300	Probability	0.998	0.764	0.768
	Mode	100	99	97
	Av.Payoff	2.896	2.706	2.701
10,000	Probability	1	0.914	0.992
	Mode	100	100	100
	Av.Payoff	2.997	2.879	2.937
100,000	Probability	1	0.959	1
	Mode	100	100	100
	Av.Payoff	2.999	2.939	2.994
$N = 100$; $q^1 = (15, 15)$				

Table 2.1: Simulations results for decision problem 1)

parts: for period $T = 300$, for $T = 10,000$, and for $T = 100,000$. In the row labeled "Probability" the average probability of playing the optimal action k (action 2 in both problems) over $N(= 100)$ simulation is given, that is, $\frac{1}{N}\sum_{i=1}^{N} p_k^{T,i}$. In the row labeled "Mode" the number of simulations where the probability of the optimal action is larger than $\frac{1}{2}$ is given, that is, $\#\{i : p_k^{T,i} > \frac{1}{2}\}$. The row labeled "Av.Payoff" gives the average payoff up to time T over N simulations, that is, $\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=1}^{T} x_i^t$. Thus, a dynamic is 100-simulation-better if the number in the row "Av.Payoff" is higher.

From Table 2.1 one can see that if $\nu = 0$, the dynamic learns very quickly: already in period 300 the probability of playing the optimal action is 0.998. Both other modifications are slower. The dynamic with forgetting accelerates after an initial slow period and in period $100,000$ it almost catches up with the fast model $\nu = 0$. All three variations converge to the optimal action for this decision problem. The "$\nu = 0$" case collects a larger average payoff due to the higher speed of learning. Applying the definition of simulation-betterness, we can say that the normalized version of the dynamics with $\nu = 0$ is simulation-best among the three versions for decision problem 1).

The second decision problem shows the non-optimality of the normalization approach. As Table 2.2 shows, not all of the simulations go to the optimal action; some of them lock in in a suboptimal one. Again, the model with $\nu = 0$ collects a larger average payoff in the beginning. However, the two other models find out what the optimal action is and they start to regain the payoff difference in later periods. The non-

Time		$\nu = 0$	no-norm	forgetting
300	Probability	0.841	0.511	0.532
	Mode	85	75	76
	Av.Payoff	2.812	2.630	2.634
10,000	Probability	0.860	0.687	0.904
	Mode	86	90	100
	Av.Payoff	2.926	2.777	2.863
100,000	Probability	0.860	0.774	1
	Mode	86	97	100
	Av.Payoff	2.930	2.846	2.983
$N = 100$; $q^1 = (10, 10, 10)$				

Table 2.2: Simulations results for decision problem 2)

normalized version of the dynamic did not catch up with the "$\nu = 0$" case but the forgetting variation did overcome it. According to definition 2.2 the forgetting variation is simulation-best for the decision problem 2) in the medium run among the three models.

The main conclusion we want to draw from the analysis of the decision problems is that in the medium run a model that has sufficiently high speed of convergence while it does not lock in in an inferior action too often is in a sense optimal. The model with forgetting seems to satisfy this criterion since it accelerates after a slow learning in the beginning. The normalized model with low ν gets locked in in an inferior action rather often; the non-normalized model explores its optimality too slowly to catch up with other models in the medium run. The model with forgetting has a non-zero probability of getting trapped in an inferior action, but the probability is very small. In fact, none of the simulations reported for the forgetting case converged to an inferior action. All in all, though we do not give a formal criterion for choice of a model in all situations, it seems that the model with forgetting performs better than the other two models. It should be noted that the results may depend on the choice of the forgetting parameter and the magnitude of the initial propensities. Nevertheless, for the simulations for games, reported in later sections, we will use the model with forgetting.

2.4 Games

Games provide additional insight into the behavior of the dynamic. The payoffs to a player depend now not on the state of the world but on the actions of the opponents. Since the opponents are also learning and do not always choose their optimal strategy, it is more difficult for the player to learn her optimal strategy. If the opponents play the same strategy all the time, the problem is reduced to the decision problem of learning the best response to this strategy. As we have seen in the previous section, it is possible to learn the best response to a given (but unknown) probability distribution with the simple reinforcement model. Thus, a Nash equilibrium is the most likely outcome of the dynamic since in it all players play mutual best responses. We start the analysis with 2×2 games and proceed to more complex ones.

2.4.1 2×2 Games

The main results about the convergence of the reinforcement dynamic in 2×2 games are stated in Posch (1997) and Laslier et al. (1999).

Theorem 2.5 *(i) If the game has strict Nash equilibria, then the dynamic with normalization ($C^t = Ct^\nu, v \geq 1, C > m$) and the dynamic without normalization ($A^t = 1$) converge to the set of strict Nash equilibria almost surely and all strict equilibria are attained in the limit with positive probability.*

(ii) If the game has no strict Nash equilibria then the dynamics may cycle.

The proof of the normalization case is in Posch (1997). The intuition of it is similar to the decision problem case, only the dimensions of the dynamic system are different since now we have two players instead of one. By asymptotic equivalence of the case with no normalization ($A^t = 1$) to the normalized dynamic the result can be extended to this case. Since other versions (normalization with $\nu < 1$, forgetting) did not converge to the optimal action in decision problems, we cannot expect that they converge to equilibria in games.

The theorem does not say anything about selection among strict Nash equilibria, except that they all have non-zero probability. To get some insight into the problem, we run simulations for some 2×2 games. The simulations showed that there could be a relation between the notion of risk dominance (Harsanyi and Selten (1988, Ch. 3)) and the results of the dynamic.

Consider a 2×2 game with two strict pure Nash equilibria (s_1, s_1) and (s_2, s_2).

	s_1	s_2
s_1	a_{11}, b_{11}	a_{12}, b_{12}
s_2	a_{21}, b_{21}	a_{22}, b_{22}

Let $u_1^1 = a_{11} - a_{21}, u_2^1 = b_{11} - b_{12}$ be the deviation losses of the two players in equilibrium (s_1, s_1) and $u_1^2 = a_{22} - a_{12}, u_2^2 = b_{22} - b_{21}$ the deviation losses in equilibrium (s_2, s_2).

Definition 2.3 *Equilibrium* (s_i, s_i) ***risk-dominates*** *equilibrium* (s_j, s_j) *if* $u_1^i u_2^i > u_1^j u_2^j$.

That is, the risk-dominant equilibrium is the one with the larger product of the losses from the deviations. Intuitively, it is less risky to play the strategy that is a part of the risk-dominant equilibrium since it is more likely that the other player will play her part of the equilibrium since the losses if she does not are higher.

First, we consider two symmetric games. The first game is of pure coordination type, the second one is of stag-hunt type. Therefore, in the first game there is no conflict between efficiency and risk dominance while in the second game one equilibrium is Pareto efficient and the other is risk dominant.

Game 1	s_1	s_2
s_1	$3, 3$	$1, 1$
s_2	$1, 1$	$2, 2$

Game 2	s_1	s_2
s_1	$3, 3$	$\frac{1}{2}, 2$
s_2	$2, \frac{1}{2}$	$2, 2$

Both games have two strict equilibria $(s_1, s_1), (s_2, s_2)$. In the first game the efficient equilibrium (s_1, s_1) is also risk-dominant, in the second one the inefficient equilibrium (s_2, s_2) is risk dominant. Since the model with forgetting has desirable properties of converging fast enough while not locking in a suboptimal action too often, we present the results of the simulations for the forgetting model with the parameters from the previous section in Tables 2.3 and 2.4 in Appendix 2.A. Similar to the decision problems, a simulation calculates the dynamic of propensities for each player according to (2.2), using associated probabilities and pseudorandom numbers generator to determine which strategies are played in a given period. The played strategies and obtained payoff are kept to report statistics after a specified number of periods T.

Table 2.3 shows that the efficient risk-dominant equilibrium is almost exclusively chosen by the dynamic in Game 1. Only one simulation converged to the inferior equilibrium. In Game 2 the inefficient risk-dominant equilibrium is chosen more often than the efficient one. The results suggest that risk-dominance plays a more important role in equilibrium selection than efficiency does. Since the initial propensities are equal, in the beginning the risk-dominant strategy is optimal. It may play a role in the convergence more often to the risk-dominant equilibrium though the results do not differ much for random initial propensities. The result is in line with the one of Kandori et al. (1993) where a different dynamic also favors the risk-dominant equilibrium.

The previous games are symmetric, but the notion of risk-dominance is not restricted to symmetric games. We illustrate than the reinforcement dynamic favors the risk-dominant equilibrium in general 2×2 games by the following game.

Game 3	s_1	s_2
s_1	$x, \frac{3}{2}$	$1, 1$
s_2	$1, 1$	$2, 2$

Depending on the value of x, either equilibrium (s_1, s_1) or equilibrium (s_2, s_2) is dominant. If $x < 3$, (s_2, s_2) is risk-dominant while if $x > 3$, (s_1, s_1) is risk-dominant. Table 2.5 in Appendix 2.A shows results of the simulations for $x = 2\frac{1}{2}, 3, 3\frac{1}{2}$, and 5 for period $T = 100,000$.

The simulations show clearly the attractiveness of the risk-dominant equilibrium. If $x = 2\frac{1}{2}$, the risk-dominant equilibrium (s_2, s_2) is observed more often. If $x = 3\frac{1}{2}$ or 5, the risk-dominant equilibrium (s_1, s_1) is observed more often. This equilibrium is also more likely to be observed when $x = 3$, though neither of the equilibria is risk-dominant in this case. The risk-dominant equilibrium is in a sense more stable (so that the dynamic does not escape once there) because the deviations from it lead to higher losses. Though it is not clear whether the stability is determined by the product of the deviation, like in the concept of risk-dominance, there is a relation between risk-dominance and higher probability of convergence to an equilibrium for the reinforcement dynamic. The exact form of this relation is an open question.

2.4.2 Other Two-Player Games

For general two-player games there is a result of convergence of the dynamic without normalization ($A^t = 1$) to any strict Nash equilibria with positive probability. This result is presented in Laslier et al. (1999).

Theorem 2.6 *If the game has strict Nash equilibria, then the dynamic without normalization ($A^t = 1$) converges to any strict Nash equilibrium with positive probability.*

The result is weaker than for 2×2 games since it does not state that only strict Nash equilibria are attained in the limit. Again, like in 2×2 games we are interested more in the equilibrium selection features of the dynamic. Therefore, we use simulations to get insight into the question of which equilibrium is selected in which game.

Our main hypothesis comes from the analysis of the behavior of the dynamic in the medium run presented in Roth and Erev (1995). They illustrated that the dynamic converges quickly to the subgame perfect equilibrium in the best-shot game but not in the ultimatum game. It happens because though there are multiple equilibria in both games, the structure of the set of the equilibrium outcomes is different in these games. In one of them (the best-shot game) only two outcomes are (pure) equilibrium outcomes. In the other game (the ultimatum game) there are (pure) equilibrium outcomes that lie between the extreme points of the set of the equilibrium outcomes. The hypothesis is that the more "central" Nash equilibrium is more likely to be observed in the medium run. We check this hypothesis by simulations.

The notion of *centrality* can be formalized as following. Consider three equilibria x_1, x_2, x_3 with payoffs $(\pi_1^1, \pi_2^1), (\pi_1^2, \pi_2^2)$, and (π_1^3, π_2^3) respectively.

Definition 2.4 *Equilibrium x_i is **central** with respect to equilibria x_j, x_k if π_1^i lies between π_1^j, π_1^k and π_2^i lies between π_2^j, π_2^k.*

A central equilibrium lies in between the other equilibria in the space of payoffs for both players. The concept of centrality is more of an ordinal concept since it uses only inequalities.

If we are in a cardinal world, and the payoffs (utilities) of the players are comparable, we can define another, somewhat related notion of *egalitarity* of payoffs. Formally, consider two Nash equilibria x_1, x_2 with payoffs (π_1^1, π_2^1) and (π_1^2, π_2^2) respectively (the subindex denotes player, the superindex denotes equilibrium).

Definition 2.5 *Equilibrium x_i is more* ***egalitarian*** *than equilibrium x_j if $|\pi_1^i - \pi_2^i| < |\pi_1^j - \pi_2^j|$.*

For the games that we consider below both notions will be equivalent, in the sense that a central equilibrium will be more egalitarian than the equilibria around it. The medium run results of the dynamic may depend on the magnitude of payoffs or initial propensities, therefore, we need to assume cardinality to analyze them. Thus, we can use the notion of egalitarity and we will mostly talk about it rather than about centrality. Observe that for egalitarity the existence of three equilibria is not needed. Nevertheless, when we talk about egalitarity below, we mean that the distribution of payoffs in an egalitarian equilibrium is more equal than in the other two equilibria with respect to which the egalitarian equilibrium is central.

We want to show that if a game possesses several equilibria then a more egalitarian one has rather high chances to be observed in the medium run. Since learning is simultaneous for both players, they are pressing each other for a compromise and since a central equilibrium is closer to a compromise, it is played more often.

To illustrate this point, we show the results of simulations of the dynamic for several games. The first two games, the ultimatum game and the best shot game, are analyzed also in Roth and Erev (1995). The third game, the oligopoly leadership game, has the same structure as the best shot game but it has a different set of equilibria, among which a central one. Through a comparison of the two last games we can see that the egalitarian equilibrium indeed has almost the same probability of being observed in the medium run as the subgame perfect equilibrium. The fourth game is a kind of property game analyzed in Young (1998, Ch.9). It also possesses an egalitarian equilibrium that is observed more often in the medium run.

The version of the dynamic used for simulations is the one with forgetting since it learns fast enough to achieve convergence in the medium run, and does not lock in in a suboptimal action too often. Like in 2×2 games the simulations calculate the vectors of propensities for each player and for each strategy, using associated vectors of probabilities to find which strategies were played and the payoffs obtained in each period. Though some of the payoffs are 0 instead of being positive, the dynamic is well defined for such payoffs.

The Ultimatum Game

Two players are to divide 10 units. The first player makes a demand $x \in \{1, ..., 9\}$ for himself and accordingly leaves $10 - x$ to the second player. The second player can then accept or reject the demand. If the second player accepts, the payoffs are $(x, 10 - x)$. If the second player rejects, the payoffs are $(0, 0)$.

We restrict the strategy set of the second player to monotone strategies. This means that strategies in which player 2 accepts a demand $x > k$ but rejects a demand $x \leq k$ are ruled out. Then player 2 has nine strategies $\{1, ..., 9\}$. Strategy k means that player 2 accepts any demand $x \leq k$, that leaves $10 - x \geq 10 - k$ for her, and rejects any demand $x > k$. The strategy set for player 1 is the same, $\{1, ..., 9\}$, where strategy j means demanding j. Then the game is reduced to a game in normal form where both players choose a strategy simultaneously. Each player chooses a (pure) strategy according to a vector of propensities. Let player 1 choose strategy j and player 2 strategy k. If $j \leq k$, the demand of player 1 is accepted and the players get $(j, 10 - j)$ correspondingly. If $j > k$ the demand is rejected and the players get $(0, 0)$. The propensities are updated according to the dynamic and the game moves to the next period. Though the original game is in extensive form, we analyze it in normal form since it is easier and does not change essentially the structure. Player 1 knows whether his current demand is accepted or rejected but does not know what would have happened with larger or smaller demands. The reinforcement model allows us to analyze such situations since updating depends only on the obtained payoff and does not depend on the payoffs that might have been obtained.

The ultimatum game has nine equilibria in pure strategies. A pair of strategies (j, j) is an equilibrium $\forall j \in \{1, ..., 9\}$. The subgame perfect equilibrium of the original extensive form game is $(9, 9)$.

Table 2.6 in Appendix 2.B shows the results of the simulations for the 100,000th period. For each pair of strategies (i, j) it reports the number of simulations (out of 100) that have the probability of playing this pair of strategies larger than $\frac{1}{2}$. Thus, 6 in cell $(7, 8)$ means that in six simulations the probability that in period $T = 100,000$ the pair $(7, 8)$ is played, in larger than $\frac{1}{2}$, that is, $p_{1,7}^T p_{2,8}^T > \frac{1}{2}$. Thus, the number in cell (i, j) indicates how many of the simulations converged, or at least started to converge to the pair of strategies (i, j). Observe that not all simulations converge to the subgame perfect equilibrium $(9,9)$. Equilibrium $(8,8)$ attracted more simulations than $(9,9)$ in period $100,000$, while one simulation converged to $(5,5)$. Thus, a more egalitarian equilibrium $(8,8)$ (with payoffs $(8,2)$) is more likely to be observed in the medium run than the subgame perfect equilibrium $(9,9)$ (with payoffs $(9,1)$). The average payoff of about 7 for player 1 also shows that there is some money left on the table. Though the most egalitarian equilibrium $(5,5)$ did not attract many simulations, there is clearly a possibility of convergence to an equilibrium that is more egalitarian than the subgame perfect one.

Roth and Erev (1995) use experiments for the ultimatum game. Their data also favor more egalitarian equilibria than the subgame perfect one. The mean demand in the experiments was between 5 and 6. From the point of view of the dynamic, a high demand is likely to get rejected, at least in the beginning. A more modest demand has a smaller probability to be rejected because of the monotonicity of the strategies of player 2. Therefore, it yields a positive payoff more often thus reinforcing itself. Though a high demand reinforces itself better, it happens less often, hence it is not clear a priori whether the subgame perfect equilibrium (with high demands) will be chosen. The result is also not new in the area of dynamic game theory. A model of noisy replicator dynamic and extended discussion about the convergence of it to an equilibrium that is not subgame perfect in the ultimatum game can be found in Gale et al. (1995).

The Best-Shot Game

Both players have three actions, $S_i = \{s_1, s_2, s_3\}$. Player 1 moves first, after that player 2 observes player 1's move and makes own move. Thus, the first player has 3 strategies, and the second player has $3^3 = 27$ pure strategies in the normal form corresponding to the described in extensive form game. The payoffs, nevertheless, can be described by a 3×3 bimatrix since the payoffs depend only on the first player's strategy and the answer to this strategy by the second player. What player 2 intended to play in response to other strategies of player 1 is irrelevant for this payoff. Payoffs are given by the following bimatrix.

	s_1	s_2	s_3
s_1	$0, 0$	$1.95, 0.31$	$3.70, 0.42$
s_2	$0.31, 1.95$	$0.31, 0.31$	$2.06, 0.42$
s_3	$0.42, 3.70$	$0.42, 2.06$	$0.42, 0.42$

The underlying story for the game is that players choose a contribution towards provision of a public good. The private cost of provision is an increasing function of the quantity provided. The provided level of the public good depends on the maximum of the contributions. Thus the benefit from the public good is an increasing function of the maximum between the two players' contributions. Actions s_1, s_2, s_3 correspond to the low, medium, and high contributions. The exact payoffs in the bimatrix are taken from Roth and Erev (1995) who report about an experiment on an extended version of the best-shot game and also about simulations of various versions of the reinforcement dynamic.

Since player 1 chooses first, the subgame perfect equilibrium strategy for him is to choose s_1, and for player 2 to choose s_3 if player 1 played s_1, and to choose s_1 if player 1 played s_2 or s_3. We will denote strategies of player 2 as $s_i s_j s_k$, where the first symbol is her response to player 1 playing s_1, the second symbol is the response to s_2, and the third symbol is the response to s_3. The subgame perfect equilibrium can be denoted as $(s_1, s_3 s_1 s_1)$. The set of pure strategy equilibria for the game consists of $\{(s_3, s_1 s_1 s_1), (s_3, s_1 s_2 s_1), (s_1, s_3 xx)\}$, where x stands for any action s_1, s_2, s_3. The subgame perfect equilibrium belongs to the last subset. In all equilibria of the last subset the payoffs are 3.7 for player 1 and 0.42 for player 2. In the other two equilibria payoffs are inverse, 0.42 for player 1 and 3.7 for player 2. There is no central, or egalitarian, equilibrium in the game.

The averages of the probabilities of actions over a hundred simulations are reported in Table 2.7 in Appendix 2.B. One can clearly see that the subset of equilibria $(s_1, s_3 xx)$ that contains the subgame perfect equilibrium is chosen with a probability indistinguishably close to 1. The results do not differ much from the ones reported in Roth and Erev (1995, Table II). Player 1 learns to choose strategy 1 rather quickly. Learning by player 2 is slower in the beginning since she has much more strategies to choose from. But in period $100,000$ player 2 learns to reply with s_3 on s_1. The game converges to the subset $(s_1, s_3 xx)$ of equilibria. Although it is not reported, the distribution among xx is such that $(s_1, s_3 s_1 s_1)$ is the most likely outcome though other strategy combinations are also present since player 2 does not have much opportunity to learn what she should play in response to s_2 and s_3.

An explanation for the finding of the subgame perfect equilibrium is that the difference in payoffs between the two equilibria is rather large for player 1. Therefore, player 1 learns not to play other equilibrium rather quickly. In the ultimatum game the difference in payoffs for player 1 between two neighboring equilibria was not that large, thus player 1 could be persuaded (by player 2 rejecting high demands) more easily to move into another equilibrium. The best-shot game does not have a central, of egalitarian, equilibrium, hence we observe fast convergence of the dynamic to the subgame perfect one. In the next subsection we consider a game with the same structure as in the best-shot game but with an egalitarian equilibrium.

The Oligopoly Leadership Game

The structure of the game is as in the best-shot game. There are two players, both have three actions, one moves first. The interpretation of the actions and the payoffs are different. The game is a simplified version of the Stackelberg game. The players are firms; they choose production levels. Firm 1 chooses first, firm 2 follows. The price for the good produced by the firms and therefore the profits received by the firms depend on the aggregate level of production. The demand function is decreasing and quadratic and the cost function, the same for both firms, is increasing and quadratic too. By an appropriate choice of functions the following payoffs bimatrix can be obtained for low, medium, and high production levels.

	s_1	s_2	s_3
s_1	$1, 1$	$1, 2.3$	$1, 4$
s_2	$2.3, 1$	$2, 2$	$0.6, 0.3$
s_3	$4, 1$	$0.3, 0.6$	$0, 0$

The magnitude of the payoffs has the same order as in the best-shot game. Interpreting actions s_1, s_2, s_3 as low, medium, and high levels of production correspondingly, one can see that the subgame perfect equilibrium strategy for firm 1 is to capture the market by choosing the high level of production s_3. Firm 2 is then left with the small remaining fraction of the market. The payoffs are 4 for firm 1 and 1 for firm 2. These actions correspond to subset (s_3, xxs_1) of the Nash equilibria of the game, where x again denotes any of the actions s_1, s_2, s_3. Equilibrium $(s_3, s_3s_2s_1)$, which belongs to this subset is the subgame perfect one. The game possesses also other pure Nash equilibria with two different payoffs. One type of them, as in the best-shot game, is an inversion of the subgame perfect equilibrium. Such equilibria are $(s_1, s_3s_3s_2)$ and $(s_1, s_3s_3s_3)$ with payoffs 1 for firm 1 and 4 for firm 2. The new type of equilibria is the egalitarian one, namely (s_2, xs_2s_2) and (s_2, xs_2s_3), where payoffs are 2 for both firms. By means of simulations we show that this last type of equilibria has not much fewer chances of being observed in the medium run than the subgame perfect equilibrium. Table 2.8 in Appendix 2.B shows the averages of the probabilities of actions over a hundred simulations.

One can see from the table that the egalitarian equilibrium (s_2, xs_2x) is learned faster than the subgame perfect one. In period 1,000 the equilibrium (s_2, xs_2x) has a larger probability of being played than the subgame perfect one. As time progresses, the subgame perfect equilibrium regains its strength. In period 10,000 the probabilities of playing both s_2 and s_3 for player 1 are equal. In period 100,000 the subgame perfect equilibrium finally gets a larger probability. If the egalitarian equilibrium fails to gain a lion's share in probability distribution in the beginning then the subgame perfect equilibrium can regain the probability in later periods. Still, the egalitarian equilibrium succeeds in being observed in about $\frac{1}{3}$ of the simulations (not reported in the table). 23 out of 100 simulations have the probability of playing this equilibrium after 100,000 periods larger than $\frac{1}{2}$. For the set of equilibria containing the subgame perfect one the number of such simulations is 45. This shows that though the subgame perfect equilibrium is more likely to be observed, the difference in probabilities between it and the egalitarian equilibrium is not that large. The egalitarian equilibrium has not much fewer chances to be played in the medium run. Since the subgame perfect equilibrium also has s_2 as the response to s_2, the early recognition of the egalitarian equilibrium helps to discriminate the subgame perfect equilibrium among the set (s_2, xxs_1). According to the simulations, the dynamic converges to the subgame perfect equilibrium in $\frac{2}{3}$ of the cases when it converges to this set (not reported in the table).

The oligopoly leadership game, as well as the ultimatum game, possesses an equilibrium that lies between two extreme equilibria where almost all payoff goes to one of the players. In this central equilibrium the payoffs are divided more equally between the players, that is, the equilibrium is also egalitarian. In contrast with these two games, the best-shot game does not possess such an equilibrium and the subgame perfect equilibrium gains dominance very easily. The implication of the observations presented in this section is that the likelihood of equilibria under the reinforcement dynamic in the medium run may depend on the structure of the set of Nash equilibria of the game. If the game has only extreme equilibria, the players (or at least one of them) quickly learn to play the subgame perfect one. A possible explanation, suggested by the analysis of the best-shot game, might be that the price player 2 has to pay to induce player 1 to play another equilibrium is too high, therefore player 2 has to allow the unfavorable for her subgame perfect outcome. In the presence of an egalitarian equilibrium between the extreme ones, it is easier for player 2 to insist on the egalitarian equilibrium and the subgame perfect one is not learned that quickly. The simulation analysis presented here shows that the central, or egalitarian, equilibrium has rather big chances to attract the process in the medium run, though in the long run the subgame perfect equilibrium prevails.

The non-equilibrium payoffs can also influence the outcome. Notice that the strategy

s_2, which is played in the egalitarian equilibrium, is best response against the uniform distribution over the opponent's strategies. This lead to initial success of strategy s_2. If the payoff are changed in such a way (without changing the set of equilibria) that the player 1 subgame perfect strategy s_3 is optimal against the uniform opponent's strategy, the chances that the egalitarian equilibrium is observed decrease. However, they are still higher than the chances to observe any other equilibria in the best-shot game, even when the player 1 subgame perfect strategy s_1 is not the best one against the uniform strategy of the opponent. Observe also that if the moves of the players are simultaneous, that is, the bimatrix above represents the game rather than the payoff structure, the central equilibrium (s_2, s_2) risk-dominates both other equilibria in pairwise comparison. As the analysis of 2×2 games has shown, such an equilibrium has high chances to be observed in the medium run.

The Property Game

In the analysis of the games above the egalitarian equilibrium was not subgame perfect while there was a subgame perfect equilibrium that regained probability in the long run though the egalitarian equilibrium did not perform too badly. In this subsection we consider a game where all three equilibria seem to be equally plausible and one cannot give preference to any of them by conventional refinements. The game has the following payoff matrix.

	s_1	s_2	s_3
s_1	$1,1$	$1,1$	$4,2$
s_2	$1,1$	$3,3$	$1,1$
s_3	$2,4$	$1,1$	$1,1$

The story behind it, adapted from Young (1998, Ch.9), is that two people are to divide a property of 6. They can agree on three outcomes (4,2),(3,3), and (2,4). In case of disagreement they both get an outside option 1. Thus the game is a simplified version of the Nash demand game. The three equilibria of the game are (s_1, s_3), (s_2, s_2), and (s_3, s_1). The second equilibrium is the egalitarian one since it has a more equal distribution of payoffs than the other two.

The results of the simulations are reported in Table 2.9 in Appendix 2.B. The simulations show that starting from equal initial propensities for all three strategies, the central equilibrium is observed more often than the other two though all three equilibria are observed in the medium run. The probability of playing the pair of strategies (s_2, s_2) is larger than the probabilities of playing the other two equilibria and the dynamic converges to the central equilibrium rather than to another equilibrium in a larger number of simulations as it is seen from the "Modes" column. An explanation for this is that if one calculates the expected probabilities of playing the equilibria for period 2 (for period 1 they are equal since the propensities are equal) then the expected probability of playing the central equilibrium is slightly higher that for the other equilibria. Hence in expectation the dynamic should go more often to the egalitarian equilibrium. Due to the noise the other equilibria also have a chance to be observed in the medium run, which is confirmed by the simulations. Another explanation for the results is that the central equilibrium risk-dominates any of the other two equilibria. Thus, from the results for 2×2 games, it has larger chances to be observed in the medium run.

The simulations support the hypothesis about a larger likelihood of the egalitarian equilibrium in the property game. In Young (1998, Ch.9) the findings for pure coordination games are similar under a different dynamic.

The hypothesis that we formulated in the beginning of the analysis for games is that a central, or egalitarian, equilibrium has high chances to be observed under the reinforcement dynamic in the medium run, even if other refinements, like subgame perfection, do not select it. The games considered above provide some evidence to support the hypothesis. Of course, it is only partial evidence and the exact relationship between the probability of observing a particular equilibrium and the structure of the set of equilibria is an open problem.

2.5 Conclusion

The simple reinforcement dynamic captures certain aspects of human learning such as the Law of Effect and the Power Law of Practice. Hence it may describe the behavior of humans while learning in decision problems and games. The analysis for the case of one-player decision problem shows that some versions of the dynamic selects the optimal action in the long run despite the fact that non-optimal actions are also reinforcing. The long run, however, is too long as the speed of convergence is low. The average payoff of other versions of the dynamic was higher after $100,000$ periods.

The speed of convergence might depend on the difference between payoffs for the optimal and non-optimal actions, and on the number of actions. Since this speed is too low for the non-normalized version of the dynamic, the dynamic was modified in such a way that learning is slow in the beginning (so that the probability of a lock-in in an inferior action is low) but accelerates later (so that the overall speed of convergence is higher). Such a modification is achieved by introducing a forgetting parameter into the dynamic. In our view, the speed of convergence plays an important role in real decision situations and games. For example, chess, in principle, can be solved explicitly but we do not have all the time in the world just to play chess. Therefore we must sometimes admit having non-optimality in actions. The trade-off between the speed and the convergence seems to be best resolved by a model with forgetting, at least for some decision problems.

The applications of the dynamic to games yield some interesting observations. Though it seems that in the long run the dynamic will eventually converge to the subgame perfect equilibrium, in the medium run it often converges to another equilibrium or fails to find an equilibrium at all (as in the ultimatum game). The equilibrium to which the dynamic converges rather often (between $\frac{1}{3}$ and $\frac{1}{2}$ of the simulations for the games we considered) has the feature to be central with respect to the set of Nash equilibria for the game, as in the oligopoly leadership game and in the property game. In some games (the property game, and in a sense in the oligopoly leadership game) it also risk-dominates other equilibria in pairwise comparison. Such an equilibrium is also called egalitarian because of the more equal distribution of payoffs. In the absence of a suitable egalitarian equilibrium, like in the best-shot game, the dynamic finds the subgame perfect equilibrium rather quickly.

The chapter provides a formal analysis of the convergence properties of the dynamic only for the case of one-player decision problems. Through simulations it gives examples of how the dynamic performs in games. Of course, simulations can give only indications for the results. Nevertheless, we believe that by the analysis of simulations enough observations can be made about chances of observing particular equilibria. Namely, the egalitarian, or central, equilibrium can compete with the subgame perfect one in the medium run if it also risk-dominates the subgame perfect equilibrium.

Appendices

2.A Simulations Results for 2×2 Games

A two-player game is given. Player 1 has k_1 actions and Player 2 has k_2 actions. The vectors of initial propensities are given and are the same for all simulations i, $q_1^{1,i} =$

Time	Probabilities	Modes	Average Payoff
300	0.638, 0.062	79, 1	2.136, 2.136
1,000	0.816, 0.023	93, 1	2.410, 2.410
10,000	0.990, 0.010	99, 1	2.882, 2.882
100,000	0.990, 0.010	99, 1	2.979, 2.979
$N = 100$; $q^1 = (15, 15)$			

Table 2.3: Simulation results for Game 1

Time	Probabilities	Modes	Average Payoff
300	0.150, 0.482	7, 52	1.866, 1.862
1,000	0.120, 0.620	9, 69	1.893, 1.887
10,000	0.106, 0.877	11, 88	2.028, 2.027
100,000	0.120, 0.880	12, 88	2.110, 2.109
$N = 100$; $q^1 = (15, 15)$			

Table 2.4: Simulation results for Game 2

$(q_{1,1}^{1,i}, ..., q_{1,k_1}^{1,i}), q_2^{1,i} = (q_{2,1}^{1,i}, ..., q_{2,k_2}^{1,i})$. The result of simulation i is a sequence of vectors of propensities for both players $(q_1^{1,i}, q_2^{1,i}), ..., (q_1^{T,i}, q_2^{T,i})$ determined recursively according to (2.2). Also, the associated sequences of vectors of probabilities p, of played actions s, and of realized payoffs x are kept.

For the 2×2 games of subsection 2.4.1 the tables 2.3, 2.4, and 2.5 reports the results in the following way. The column "Probabilities" reports the average (over N simulations) probabilities of equilibria $(s_1, s_1), (s_2, s_2)$ respectively. That is the first number in the column is $\frac{1}{N}\sum_{i=1}^{N} p_{1,s_1}^{T,i} p_{2,s_1}^{T,i}$ and the second number is $\frac{1}{N}\sum_{i=1}^{N} p_{1,s_2}^{T,i} p_{2,s_2}^{T,i}$. The column "Modes" reports the numbers of simulations in which the probability of the given equilibrium is larger than $\frac{1}{2}$, that is, $\#\{i : p_{1,s_1}^{T,i} p_{2,s_1}^{T,i} > \frac{1}{2}\}$ and $\#\{i : p_{1,s_2}^{T,i} p_{2,s_2}^{T,i} > \frac{1}{2}\}$. The column "Average Payoffs" gives the average (over N simulation) payoffs up to time t for players 1 and 2 respectively, that is, $\frac{1}{N}\sum_{i=1}^{N} \frac{1}{T}\sum_{t=1}^{T} x_1^{t,i}$ and $\frac{1}{N}\sum_{i=1}^{N} \frac{1}{T}\sum_{t=0}^{T-1} x_2^{t,i}$. The rows correspond to time periods $T = 300$, $T = 1,000$, $T = 10,000$, and $T = 100,000$.

For Game 3 the first column is the value of x. The numbers are given for $T = 100,000$.

x	Probabilities	Modes	Average Payoff
2.5	0.380, 0.620	38, 62	2.161, 1.790
3	0.610, 0.390	61, 39	2.568, 1.678
3.5	0.860, 0.140	86, 14	3.231, 1.557
5	0.940, 0.060	94, 6	4.746, 1.520

$N = 100$; $q^1 = (15, 15)$

Table 2.5: Simulation results for Game 3

	1	2	3	4	5	6	7	8	9
1	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0	0	0
6	0	0	0	0	0	4	1	2	0
7	0	0	0	0	0	0	10	6	12
8	0	0	0	0	0	0	0	18	15
9	0	0	0	0	0	0	0	0	11

Average payoffs: Player 1: 7.141, Player 2: 2.789

$N = 100$; $q_i^1 = (10, 10, 10, 10, 10, 10, 10, 10, 10, 10), i = 1, 2.$

Table 2.6: Simulation results for the ultimatum game

2.B Simulations Results for Other Games

The simulations are similar to the ones for 2×2 games. For the ultimatum game we report the results only for the 100,000th period. The number in cell (i, j) represents the number of simulations in which the probability of playing pair (i, j) is larger than $\frac{1}{2}$, $\#\{m : p_{1,i}^{T,m} p_{2,j}^{T,m} > \frac{1}{2}\}$.

For the best shot and oligopoly leadership games the results are reported in the following manner. The tables are divided into three parts, for $T = 1,000$, $T = 10,000$, and $T = 100,000$. In each part, for player 1 the average (over N simulations) probabilities of playing each of the three strategies are reported, that is, $\frac{1}{N} \sum_{i=1}^{N} p_{1,s_k}^{T,i}$. For player 2 the average probabilities of answers to a given strategy of player 1 are reported, that is, for example, in the column s_k and the row "Player 2 on s_1" the number is

Time		s_1	s_2	s_3	Av.Payoffs
1,000	Player 1	0.951	0.036	0.009	2.151
	Player 2 on s_1	0.127	0.331	0.538	0.506
10,000	Player 1	0.999	0.000	0.000	3.195
	Player 2 on s_1	0.000	0.058	0.941	0.411
100,000	Player 1	1	0	0	3.645
	Player 2 on s_1	0	0	1	0.419

$n = 100;\ q_1^1 = (10, 10, 10), q_2^1 = (\frac{30}{27}, ..., \frac{30}{27})$

Table 2.7: Simulation results for the best-shot game

Time		s_1	s_2	s_3	Av.Payoffs
1,000	Player 1	0.052	0.567	0.380	1.975
	Player 2 on s_2	0.164	0.743	0.092	1.445
	Player 2 on s_3	0.467	0.306	0.226	
10,000	Player 1	0.000	0.500	0.500	2.621
	Player 2 on s_2	0.068	0.893	0.037	1.489
	Player 2 on s_3	0.567	0.220	0.175	
100,000	Player 1	0	0.420	0.580	3.079
	Player 2 on s_2	0.069	0.897	0.034	1.436
	Player 2 on s_3	0.623	0.176	0.167	

$N = 100;\ q_1^1 = (10, 10, 10), q_2^1 = (\frac{30}{27}, ..., \frac{30}{27})$

Table 2.8: Simulation results for the oligopoly leadership game

$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{9}\sum_{m=1}^{3}\sum_{j=1}^{3}p_{2,s_k s_j s_m}^{T,i}$. Analogously in rows "Player 2 on s_2" and "Player 2 on s_3". The column "Av.Payoffs" reports the average realized payoffs up to time T, $\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=1}^{T}x_1^{t,i}$ for Player 1 and $\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=1}^{T}x_2^{t,i}$ for Player 2.

For the property game the set of equilibria is $(s_1, s_3), (s_2, s_2), (s_3, s_1)$. The numbers in the table are given correspondingly to the equilibria, that is, the first number in each entry is for equilibrium (s_1, s_3), etc. The column "Probabilities" reports the average (over N simulations) probabilities of equilibria. That is the first number in the column is $\frac{1}{N}\sum_{i=1}^{N}p_{1,s_1}^{T,i}p_{2,s_3}^{T,i}$, the second number is $\frac{1}{N}\sum_{i=1}^{N}p_{1,s_2}^{T,i}p_{2,s_2}^{T,i}$, and the third number is $\frac{1}{N}\sum_{i=1}^{N}p_{1,s_3}^{T,i}p_{2,s_1}^{T,i}$. The column "Modes" reports the numbers of simulations in which the probability of the given equilibrium is larger than $\frac{1}{2}$, that is, $\#\{i : p_{1,s_1}^{T,i}p_{2,s_3}^{T,i} > \frac{1}{2}\}$,

Time	Probabilities	Modes	Average Payoffs
300	0.113, 0.191, 0.143	2, 12, 4	1.777, 1.831
1,000	0.133, 0.260, 0.175	8, 29, 18	1.920, 1.993
10,000	0.289, 0.406, 0.284	30, 41, 29	2.612, 2.651
100,000	0.300, 0.410, 0.290	30, 41, 29	2.969, 2.955

$N = 100;\ q_i^1 = (10, 10, 10), i = 1, 2$

Table 2.9: Simulation results for the property game

$\#\{i : p_{1,s_2}^{T,i} p_{2,s_2}^{T,i} > \frac{1}{2}\}$, and $\#\{i : p_{1,s_3}^{T,i} p_{2,s_1}^{T,i} > \frac{1}{2}\}$. The column "Average Payoffs" gives the average (over N simulation) payoffs up to time t for players 1 and 2 respectively, that is, $\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=1}^{T} x_1^{t,i}$ and $\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=0}^{T-1} x_2^{t,i}$. The rows correspond to time periods $T = 300$, $T = 1,000$, $T = 10,000$, and $T = 100,000$.

Chapter 3

Imitation Dynamic in Cournot Oligopoly

3.1 Introduction

In the previous chapter we considered a dynamic that uses information only about a player's own payoffs. In this chapter we turn to a dynamic that uses information about other players' payoffs too.

People often observe what their neighbors do. For example, when planning to buy a computer, or a car, it is not uncommon to ask around in the circle of friends or colleagues what they advise with respect to what brand is the best one. It is often worthwhile to compare own experience with that of somebody else and do what the other person did if he/she fared better. Imitative behavior is a part of real life and we often employ it, maybe unconsciously.

When observing other people in the same decision problem individuals can see what others have done in the same situation as they are in. Clearly, such observations give some valuable information without incurring the cost of experimentation. Thus, one can learn via mistakes of others rather than via own mistakes[1]. Still, the extra information should be handled with care. Similarly to the reinforcement dynamic of the previous chapter, if imitation is too fast, a lock-in in an inferior action is possible. Banerjee (1992) considers a model with imperfect information that can lead (rationally!) to the selection of a non-optimal action. The population may end up with an inferior technology if the agents disregard their own information in view of the behavior of other agents. Examples

[1] A wise person said: "only fools learn by their own mistakes; wise people learn by mistakes of others".

of boundedly rational sequential decision making with imitation that may lead to the efficient action are given in Ellison and Fudenberg (1993, 1995). The question in decision problems is not whether to imitate but rather how to imitate.

Imitation can occur in games too. Players want to improve their performance, therefore if they observe a payoff higher than their own they might be tempted to imitate the strategy bringing this higher payoff. Doing so may actually decrease their own payoff but at the moment of making the decision players might not realize it. One justification can be that players do not know that they are playing a game, or they do not understand the complex interdependence of their actions. Thus, they may use the same behavioral rule as in decision problems. Other justifications might be such considerations as "other players are more clever, I shall imitate their behavior", or "finding the best strategy is too difficult, let's watch what others do". Imitation can be useful if the environment changes rapidly, as is argued in Rhode and Stegeman (1997). There is some experimental support for imitative behavior in games, including Cournot oligopoly models, nicely presented in Huck et al. (1999). Yet another justification of imitative behavior in games is that players are concerned about relative payoffs. Such an interpretation of imitation appears, for example, in Schaffer (1989). Non-strategic behavior, like imitation, may lead in some games to the outcome that would have appeared had the behavior been strategic. The main question of this chapter is when this would happen for games having the Cournot oligopoly structure.

Specifically, in games interaction and imitation may occur in the same population[2]. This case is of practical interest. For example, firms in an industry can be competing with each other and at the same time be interested in each other's market strategy so that a successful firm might be imitated by others. Another justification for imitation is that a firm with a higher profit increases its market share and expands, establishing new factories. These factories still employ the same market strategy as the original firm. Other firms shrink, thus with time the strategy employed by the firm with higher profit gains more weight in the population. Of course, the presence of imitation does not imply that everybody will be better off in the end; on the contrary, the results might be very inefficient for the firms as in the Cournot oligopoly model of Vega-Redondo (1997).

This chapter focuses on games arising from Cournot oligopoly models. Our model follows closely the model of Vega-Redondo (1997). In his paper he shows that the dynamic "imitate the best" converges to the Walrasian (competitive) equilibrium rather than to

[2]The case when interaction and imitation are in different populations is considered in Schlag (1998).

the Cournot-Nash equilibrium. We introduce variations of the imitation process or the underlying game that allow for the reinstalling of the "as if" rational behavior represented by Nash equilibrium. The variations still have as a base the rule "imitate the best". Gale and Rosenthal (1999) have a model of imitation in Cournot oligopoly that leads to the Cournot-Nash equilibrium but their result was due to the fact that imitators are imitating the aggregate action of the population, disregarding the comparison of payoffs to their current action and the new action. We want to keep as a rule imitation of success.

The first variation concerns imperfection in imitation. Imitation in real life can be imperfect, for example, when one observes only a part of the strategy of another player but not the whole strategy. The reason is that a strategy prescribes what to do in every contingency, but the players observe only the realized play. When trying to imitate a strategy, the absent parts are filled with something else, for instance, with a part of the own old strategy. In games such models are considered in the context of the repeated prisoners' dilemma game by Cooper (1996) and Ruebeck (1999). Depending on the exact specification of the model both a substantial amount of coordination and total defection are obtained as outcomes. In this chapter we extend the one-stage Cournot oligopoly to a two-stage capacity-price game and we demonstrate that imperfect imitation gives some support to the Cournot-Nash outcome in the sense that it is a possible long run outcome of the dynamic process.

The second variation differentiates interaction and imitation. In a large population one can observe (and imitate) a player with whom there was no direct interaction. For example, players interact in different locations, like in the evolutionary local interaction model of Ellison (1993) but the information is spread over the whole population. Alternatively, firms interact in different markets but the information available comes not only from the own market but from other markets too. The possibility of observing another copy of the game allows comparison of different outcomes and restores the Nash equilibrium if it is more efficient that the outcome of the purely imitative process. A similar model is considered in Palomino (1996) who makes the accent on the survival of non-rational outcomes while we make accent on the possibility of the rational outcome.

The chapter is constructed as follows. Section 3.2 presents the general model of imitation and experimentation and an example of Cournot oligopoly game. Section 3.3 extends the example to imperfect imitation and Section 3.4 to a local interaction model. Section 3.5 concludes.

3.2 The Model

3.2.1 The Perfect Imitation Dynamic

There are $N = m \cdot n$ agents in the population. They play a finite n-player symmetric game $G = (n, S, \pi)$ having been matched each period randomly or deterministically according to a prespecified rule; m is interpreted as the number of locations where the game is played and can be infinite. The strategy set S of the game is assumed to be finite.

The dynamic process is in discrete time $t = 1, 2, \ldots$. Agent $i \in N$ is characterized at time t by a pure strategy s_i^t she plays in period t. The state of the population at time t is a vector $s^t \in S^N, s^t = (s_1^t, \ldots, s_N^t)$ of the strategies of all agents. Having been matched and having played the game G with these strategies the players realize payoffs. Let $\pi^t(s^t) = (\pi_1^t, \ldots, \pi_N^t)$ be the vector of the realized payoffs of the players.

The imitation process works as follows. In each period of time each player has the possibility to revise her strategy[3]. To revise her strategy the player samples k other players randomly or according to a prespecified sampling procedure and observes their payoffs and corresponding strategies. A (natural) restriction we impose on the sampling is that the players always observe their own payoff and strategy and those of their direct opponents in the match. The players can also receive information about a play in other copies of the n-player game, though. We assume that each player i has a non-zero probability to observe any player j.

Denote by $IP_i^t = ((\pi_{j_1}^t, s_{j_1}^t), \ldots, (\pi_{j_k}^t, s_{j_k}^t))$ the information player i has obtained at the end of period t and by $P_i^t = (j_1, \ldots, j_k)$ the set of players whose payoffs and strategies player i has observed. The player finds the maximal payoff in his information set IP_i^t and then copies the corresponding strategy. That is,

$$s_i^{t+1} = s_{j_m^*}^t, \; j_m^* \in \arg\max_{j_m \in P_i^t} \{\pi_{j_m}^t\}. \tag{3.1}$$

If there are several strategies that give the maximal payoff, one of them is chosen randomly according to a probability distribution with full support on the set of strategies giving this highest payoff[4]. This kind of imitation is called "imitate the best", since only the

[3]If there is a probability $\tau < 1$ (independent across players) of revising the strategy, the process slows down but the results do not change.

[4]If we assume that the player does not change strategy if the current payoff is among the highest then stationary states other than monomorphic are possible, which makes the analysis more difficult but does not change the qualitative results.

highest payoff in the sample can be imitated. If the initial strategies s_i^1 are given, the dynamic is fully specified.

We call a population state *monomorphic* if all individuals play the same strategy. In such a state all individuals receive the same payoff since the game is symmetric. Observe that the imitation process alone cannot bring a new strategy into the population. Obviously, if the population is in a monomorphic state, it will stay there. Therefore, monomorphic states are stationary. Conversely, consider the case when there are several strategies in the population. Suppose two of the strategies bring different payoffs. Since each player has a positive probability to observe any other player, the player with the lower payoff will eventually observe the player with the higher payoff and will imitate her. Therefore, such a state cannot be stationary. Alternatively, suppose that all present strategies bring the same payoff. By the same reason as above, and by the fact that any of the strategies that bring the maximal payoff is imitated with a positive probability, there is a non-zero probability that players with different strategies will switch to the same strategy. This will reduce the number of strategies present. Thus the only stationary state of this sort is the one where all agents play the same strategy.

Remark 3.1 *All monomorphic states, that is, the states with strategy profile of the population $s^t = (s_1^t, ..., s_N^t)$ with $s_i^t = s_j^t$ $\forall i, j$ are stationary states of the imitation process. These are the only stationary states of the process.*

In order to analyze which stationary states are more stable against small perturbations of the process, the possibility of experimentation is introduced. In each period each player has an independent probability λ of experimentation, that is, a player can switch to a strategy that is not necessarily found by the imitation procedure. The probability distribution over strategies resulting from experimentation will be taken to be uniform across strategies for each player though any not very extreme (in the sense that for any two strategies the ratio of probabilities that the strategies are the result of experimentation remains bounded as $\lambda \to 0$) probability distribution will do. The combined process of imitation and experimentation defines a Markov chain on the space of states of the population. If m is finite, the Markov chain is ergodic and has a unique stationary distribution $\mu(\lambda)$ over states. We are interested in the case when the probability of experimentation is arbitrarily small, that is, we consider $\lim_{\lambda \to 0} \mu(\lambda)$. This limit will be called the limit stationary distribution.

The limit stationary distribution is a distribution over the states of the process. The notion of stochastic stability uses the limit stationary distribution.

Definition 3.1 *A state is **stochastically stable** if it has a non-zero weight in the limit stationary distribution.*

Since the process is ergodic, the weights in the limit stationary distribution correspond to the proportion of time the process spends in each stochastically stable state in the long run, independent of the initial conditions. Thus, if the limit stationary distribution puts a non-zero weight on a state, this state is observed some non-negligible proportion of the time in the long run. A stochastically stable state will be called also a long-run outcome of the process.

Notice that we require very little rationality from the players. First, they do not suspect that they are playing a game; they simply copy actions of others that may be their direct opponents. Second, they condition their behavior only on the last period observation, that is, they have one-period memory. These boundaries on rationality of the players are not uncommon in the literature. For example, Schlag (1998) analyzes behavioral rules in one-player decision problems where players are allowed to use information only about the last period play of oneself and one other individual. He shows that in the class of such models only imitative models never decrease the expected payoff. The closest to our model are the model of Vega-Redondo (1997) where players could observe all other players in the population in the setup of a Cournot oligopoly and the model of Palomino (1996) where players could observe a subpopulation of other players. The aim of this chapter is to show that even with little rationality the players can achieve "as if" rational behavior.

The simplest example of the process described above is a population of two agents, playing a two-player game (thus $m = 1, n = 2$). In this case the long-run outcome of the imitation process is influenced by relative rather than by absolute payoffs[5]. How exactly the outcome depends on relative payoffs in general games is not clear; in section 3.3 we present an example showing that the set of stochastically stable states does not necessarily coincide with the set of Nash equilibria of the relative payoff game.

In what follows we consider the imitation and experimentation dynamic in symmetric Cournot oligopoly type games with a unique pure strategy symmetric equilibrium. The main question is to what state the imitation and experimentation dynamic converges in the long run in such games.

[5] Shubik (1982, Ch.10) was one of the first to notice this. See also Schaffer (1989).

3.2.2 The Cournot Oligopoly Example

One of the results in the class of imitation models described above is due to Vega-Redondo (1997), who shows that the process converges to the non-Nash Walrasian outcome in Cournot oligopoly. We will build on the model of Vega-Redondo (1997).

Consider an industry with n identical firms all competing with each other, thus $m = 1$. Denote by $q_i \geq 0$ the output of firm i and by Q the aggregate output. The demand side is given by the twice differentiable decreasing demand function $D(p)$ that is concave. The inverse demand function is denoted $P(Q) := D^{-1}(Q)$. The inverse demand function is decreasing and concave too. The twice differentiable increasing cost function $C(q)$ is assumed to be convex and $C(0) = 0$. For our purposes we need to define the Walrasian and Cournot production levels since they play a particular role in the analysis.

Definition 3.2 *The* ***Walrasian*** *(competitive) production level q_W is the production level such that it maximizes profit taking price as given, $P(nq_W)q_W - C(q_W) \geq P(nq_W)q' - C(q')$ $\forall q'$. The corresponding price $P(nq_W)$ is the Walrasian price.*

Alternatively, the Walrasian production level is such that price equals marginal cost, $P(nq_W) = C'(q_W)$.

Definition 3.3 *The* ***Cournot*** *(Nash) production level q_N is the production level in symmetric Nash equilibrium, $P(nq_N)q_N - C(q_N) \geq P(q' + (n-1)q_N)q' - C(q')$ $\forall q'$. The corresponding price $P(nq_N)$ is the Cournot (Nash) price.*

Given the assumptions above, the Walrasian and the Cournot equilibria exist and are unique. The Walrasian output level is higher than the Cournot output level.

To bring the setup under the general model considered above, assume that firms can choose quantities from a finite grid $\Gamma_q = \{0, \delta_q, ..., v\delta_q\}$. It is assumed that both the Walrasian output level and the Cournot equilibrium output level belong to the grid[6].

Since all firms are competing with each other, that is, they are all in a single game, they all observe each other. Then the work of the imitation process can be illustrated on the example with a linear cost function. Observe that if price is higher than marginal cost, the profit of firm i is higher than that of firm j if firm i has a higher output, since the difference between price and marginal cost is the same for all firms. If price is below marginal cost, all firms make losses, but the one with the lowest production level makes the

[6]The finite grid can be largely dispensed of, see Schenk-Hoppé (1997). We assume the finite set of strategies to avoid unnecessary technical difficulties.

smallest loss. Suppose that the firms are in a stationary state, that is, they all produce the same quantity. If price is higher than marginal cost, a firm experimenting with a higher quantity (still keeping price above marginal cost) obtains a higher profit that the other firms. The other firms then imitate the new quantity and the old stationary state is upset. If the current price is lower than marginal cost, a firm experimenting with a lower quantity will have a smaller loss than the other firms. Finally, if price is equal to marginal cost no experimentation by one firm can upset this stationary state.

Therefore, whatever the current output in a stationary state is, a firm will get more than the other firms if it decreases or increases output towards the point where price equals marginal cost. If all firms produce the competitive output, such that price equals marginal cost, a unilateral experimentation will not bring a profit higher than the profit of other firms, while at all other production levels a unilateral experimentation upsets the corresponding stationary states. Since we assumed that the probability of experimentation is arbitrarily small, a unilateral experimentation is infinitely more likely than a simultaneous experimentation by more firms. Thus, the competitive production level is more stable than the other stationary states. The limit stationary distribution of the process of imitation and experimentation puts weight one on the Walrasian equilibrium. The result holds for an arbitrary convex cost function.

Theorem 3.1 *The only long-run outcome of the imitation and experimentation dynamic for the Cournot oligopoly described above is the Walrasian equilibrium.*

The formal proof is in Vega-Redondo (1997).

Thus, the long-run outcome of the process for Cournot oligopoly coincides with the Walrasian outcome. The latter is not a Nash equilibrium of the game with absolute profits as payoffs but it is a Nash equilibrium of the game where instead of absolute profit relative profit is maximized (see Vega-Redondo (1997)). Though it is the case here that the outcome of the process coincides with a Nash equilibrium of the relative payoffs game, this does not always hold as we will demonstrate in the example in section 3.3.

The result of the imitation and experimentation dynamic in the simple Cournot oligopoly model is very inefficient for the firms. Profits in the Walrasian equilibrium are smaller than in the Nash equilibrium. If the firms simply imitate each other and occasionally experiment, in the long run they receive a smaller profit than in the one-shot Nash equilibrium. This comes from the fact that a deviation from the Nash equilibrium in Cournot oligopoly hurts the competitors more than it hurts the deviator. This result also holds when a higher payoff is not always imitated. Schlag (1998) argues that proportional imitation, that is, when the probability of imitating is proportional to the difference in payoffs, is optimal (in the sense of payoff increasing in the next round for any decision problem) for a situation when agents play many copies of a decision problem. In our case proportional imitation slows down the process but does not change the long-run outcome since a strategy close to the Walrasian one gets imitated with a probability higher than the probability of imitating a strategy far from the competitive equilibrium.

Despite the inefficiency of the outcome in Cournot oligopoly, imitation does not seem to be an implausible behavioral rule. People do imitate in the real life and experiments (see Huck et al. (1999)). Is it possible to change the setting of the model slightly so that the agents fare better? The next two sections build on the model extending it in some dimensions.

3.3 Imperfect Imitation

Imperfections in imitation can occur, for example, when players can observe only the realized play but not the intended strategy, like in extensive form games. Imitation in the context of a repeated prisoners' dilemma is considered in Cooper (1996) and Ruebeck (1999). The players there can observe only the realized play and therefore they are not able to distinguish between "always cooperate" and "tit-for-tat" strategies if their opponents never defect. To incorporate such a possibility into the oligopoly model, we extend the game to the two-stage game of price competition with capacity precommitment, the model considered in Kreps and Scheinkman (1983) to justify Cournot competition.

Kreps and Scheinkman (1983) consider a game where in the first stage two firms simultaneously choose capacities and in the second stage, after the capacities are known, they simultaneously choose prices. The assumptions on the demand and cost functions are the same as in section 3.2. The rationing rule is efficient, that is, the consumers with higher valuations for the good buy from the cheapest supplier. If the prices are equal then each firm gets a half of the corresponding demand. Kreps and Scheinkman (1983) show that under certain mild conditions the choice of capacities and prices corresponding to the Cournot equilibrium of the one-stage game is an equilibrium of the two-stage game.

We now assume that the firms observe each others' actions and profits. In the two-stage game a firm can observe only the capacity of the other firm and the price announced by the other firm after both firms have chosen capacities, that is, a firm does not know what the other firm's prices are in response to other capacity choices. Therefore imitation is imperfect, that is, a firm can imitate just one price and one capacity it has observed.

There are several different ways to model such a restriction. For example, the firms can keep intact their intended responses on capacity combinations different from the observed one. Alternatively, they can change all intended price choices to the one observed. We adopt the second way of modeling. Though it is quite a restriction on the rationality of the players, it requires the least memory to remember a strategy (only one price to remember instead of the whole vector of intended prices).

Formally, imitation in the two-stage game described above is modeled as follows. There is one location with just two firms. Thus, $m = 1, n = 2$. In addition to the grid on quantities Γ_q, the firms choose prices from a finite grid $\Gamma_p = \{0, \delta_p, ..., w\delta_p\}$ such that $\forall q \in \Gamma_q$ it contains $P(2q)$. Then it contains prices that correspond to the Walrasian and Cournot equilibria. A (full) strategy of firm i in the normal form corresponding to the two-stage game in extensive form can be described by a $(v+1)^2+1$-tuple specifying the capacity choice and the price choice for each combination of the capacities of the two firms. Denote such a strategy in period t by $s_i^t = (q_i; p_i(0,0), ..., p_i(v\delta, v\delta))$ where $p_i(x,y)$ is the intended price choice of firm i if the capacities are x and y. After playing the game firm i has observed own capacity choice q_i^*, the capacity choice q_j^* of firm j, own price $p_i(q_i^*, q_j^*) =: p_i^*$ and firm j's price $p_j(q_i^*, q_j^*) =: p_j^*$. Thus, in distinction from the perfect imitation dynamic of the previous section, the information set of firm i is $IP_i^t = ((\pi_i^t, s_i^t), ((\pi_j^t, c(s_j^t)))$, where $c(s_j^t) = (q_j^t; \cdot, ..., \cdot, p_j(q_i^t, q_j^t), \cdot, ..., \cdot)$, that is, $c(s_j^t)$ is a signal that suppresses information about intended choices for capacities other than the played ones.

By assumption, when imitating, firms change the capacity and all intended price choices to the one observed. Thus, if firm j's profit was larger, firm i's strategy in the next period will be $s_i^{t+1} = (q_j^*; p_j^*, ..., p_j^*)$. Therefore, except possibly in a finite number of periods after each experimentation, a firm's strategy is completely described by just two variables, (q, p). This is equivalent to the situation when firms announce both capacity and price simultaneously. The original two-stage game is reduced to a simultaneous move one-stage game. Observe that this reduction comes from our assumption on the imitation process, not on the game itself.

The resulting one-stage game can have multiple symmetric Nash equilibria. Choosing the Walrasian quantity and price may be an equilibrium (if the market is not too big to accommodate comfortably a firm with a higher price), as well as choosing the Cournot quantity and price. This makes the point that the Cournot-Nash quantity and price will have a non-zero probability in the limit stationary distribution of the process stronger, since the Cournot equilibrium has to compete with other Nash equilibria.

Let us illustrate how the process works on a simple example with linear demand and cost functions and only two possible levels of capacities and of prices, corresponding to the Cournot-Nash and Walrasian competitive outcomes. Then each firm has four strategies. Denote them by $(q_N, p_N), (q_N, p_W), (q_W, p_N), (q_W, p_W)$, where q_N (p_N) stands for the Cournot-Nash capacity (price) and q_W (p_W) stands for the Walrasian capacity (price). The demand and cost functions are $P(Q) = a - bQ$ and $C(q) = cq$, with positive a, b, c and $a > c$. Then the Cournot-Nash output of each firm is $q_N = \frac{a-c}{3b}$, and the corresponding price is $p_N = \frac{a+2c}{3}$. The competitive output of each firm is $q_W = \frac{a-c}{2b}$, the corresponding price is $p_W = c$. One can calculate the profits of the firms. After dividing each profit by $\frac{a-c}{18b}$, which does not change the structure, the game has the following form:

	(q_N, p_N)	(q_N, p_W)	(q_W, p_N)	(q_W, p_W)
(q_N, p_N)	$2a-2c, 2a-2c$	$2a-2c, 0$	$2a-2c, 2a-5c$	$a-4c, 0$
(q_N, p_W)	$0, 2a-2c$	$0, 0$	$0, 2a-5c$	$0, 0$
(q_W, p_N)	$2a-5c, 2a-2c$	$2a-5c, 0$	$2a-5c, 2a-5c$	$a-7c, 0$
(q_W, p_W)	$0, a-4c$	$0, 0$	$0, a-7c$	$0, 0$

The pure (without experimentation) imitative process has four stationary states on the diagonal of the bimatrix. Now let us introduce experimentation. Consider the Cournot-Nash stationary state in which both firms play (q_N, p_N). Experimentation with (q_N, p_W) or (q_W, p_N) cannot bring a higher profit than that of the firm that does not experiment. If $a > 4c$, experimentation with (q_W, p_W) cannot bring a higher payoff either. All other stationary states are upset by experimentation with (q_N, p_N) by one firm. Therefore, if $a > 4c$, the Cournot-Nash outcome has weight one in the limit stationary distribution and is played in the long run.

If $a \leq 4c$, both the Cournot (q_N, p_N) and the competitive (q_W, p_W) outcomes are Nash equilibria of the game above. Experimentation by one firm is enough to upset any of the stationary states, for example (q_N, p_N) is upset by one firm moving to (q_W, p_W). The state (q_W, p_W) itself is upset by one firm experimenting with (q_N, p_W). We need to show that the Cournot-Nash outcome can be achieved from other stationary states with no more experimentations than any other stationary state. Using the graph-theoretic techniques introduced to evolutionary game theory by Kandori et al. (1993) and Young (1993), which are described in the Appendix 3.A, we can show that the Cournot-Nash outcome has a non-zero probability in the limit stationary distribution in the game above[7]. In the

[7]For the game above, the limit stationary distribution puts weights $(\frac{1}{4}, \frac{1}{4}, 0, \frac{1}{2})$ on stationary states $((q_N, p_N), (q_N, p_W), (q_W, p_N), (q_W, p_W))$ respectively.

general case, a similar reasoning gives

Theorem 3.2 *In the model of imitation and experimentation described above the limit stationary distribution put a non-zero probability on the Cournot-Nash outcome if the grids Γ_q and Γ_p are fine enough.*

The formal proof is in the Appendix 3.A.

Notice that in the example above the Walrasian outcome also has a strong position in the limit stationary distribution. In general, if the grid is fine enough, it has a non-zero weight too, since, unless the process is already in the Walrasian equilibrium where marginal benefit equals marginal cost, it is always possible to undercut slightly the other firm in price and expand capacity so that the deviating firm's profit does not change but the remaining firm's demand decreases thus its profit decreases. The remaining firm is then forced to follow the deviation.

There are other states of the process that are stochastically stable. Any state (q, p) that has the property that a firm playing (p, q) has a higher or equal profit than a firm playing (q_N, p_N) is stochastically stable. Such states lie above the $q = \frac{D(p)}{2}$ line, that is, in such states $q > \frac{D(p)}{2}$. Since the Walrasian outcome is also stochastically stable, states that have the property that a firm in such a state has a higher or equal profit than a firm playing the Walrasian strategy, are long run outcomes too. Such states lie below the $q = \frac{D(p)}{2}$ line, that is, in such states $q < \frac{D(p)}{2}$. In the example above, the state (q_N, p_W) is such a state. For any price p between the Cournot-Nash and the competitive price there exists a set of capacity choices q such that (p, q) is stochastically stable.

The two-stage game considered in this section gives also an answer to the question whether the imitation and experimentation dynamic always leads to the outcomes that are Nash equilibria of a game where absolute payoffs are replaced by relative payoffs. Consider the example with two strategies, the Cournot-Nash one and the Walrasian one, and let us take particular values for the parameters, $a = 2, b = c = 1$. The bimatrix of relative payoffs is

	(q_N, p_N)	(q_N, p_W)	(q_W, p_N)	(q_W, p_W)
(q_N, p_N)	$0, 0$	$2, -2$	$3, -3$	$-2, 2$
(q_N, p_W)	$-2, 2$	$0, 0$	$1, -1$	$0, 0$
(q_W, p_N)	$-3, 3$	$-1, 1$	$0, 0$	$-5, 5$
(q_W, p_W)	$2, -2$	$0, 0$	$5, -5$	$0, 0$

Let us denote by $\theta_1, \theta_2, \theta_3, \theta_4$ ($\varphi_1, \varphi_2, \varphi_3, \varphi_4$) the probabilities with which the strategies are played by player 1 (player 2). The Nash equilibria of this game are given by $\theta_1 = \varphi_1 = 0$, $\theta_2 \leq \frac{1}{2}$, $\varphi_2 \leq \frac{1}{2}$, $\theta_3 = \varphi_3 = 0$, $\theta_4 = 1 - \theta_2$, $\varphi_4 = 1 - \varphi_2$. Strategy (q_N, p_N) is never played in an equilibrium while the limit stationary distribution of the imitation and experimentation dynamic places a non-zero weight on it. Therefore, in this game the set of long-run outcomes of the imitation and experimentation dynamic is larger than the set of Nash equilibria of a game where absolute payoffs are replaced by relative payoffs. This shows that it is not generally true that a long-run outcome of the dynamic is a Nash equilibrium of the relative payoffs game.

This section showed that imperfect imitation, in distinction from perfect imitation, may lead to the outcome corresponding to a Nash equilibrium in a Cournot type game. In this outcome firms earn higher profits than in the competitive outcome, thus imperfection in imitation improves efficiency. The next section considers another modification of the imitation dynamic, namely a local interaction model.

3.4 Random Matching and Local Interaction

In this section we consider models with number of locations $m > 1$. Specifically, we assume that there is a large (infinite) population. Thus, there is a large (infinite) number of locations. The individuals in the population are randomly assigned to locations (matched) to play an n-player Cournot oligopoly type game described below.

Let us first consider the imitation process without experimentation. We return to the perfect imitation setup of section 3.2. We focus on the analysis when there are only two strategies present in the population. This is not a strong restriction since with imitation only, the number of strategies in the population cannot increase but can decrease from an arbitrary situation to the situation with no more than two strategies. We call the situation with only two strategies a *two-strategy contest.* If there is a strategy s that wins every two-strategy contest without experimentation, then such a strategy is very likely to be the long-run outcome of the process since any other monomorphic state of the population can be upset by experimentation with s by a small proportion of players, while the state where everybody plays s cannot be upset by experimentation with other strategies.

The Cournot oligopoly type game we consider has the following features. The finite strategy set is ordered. It can be interpreted as the set of possible production levels. There is one strategy, denoted by s_N, that is called the Cournot-Nash strategy. Suppose that there are only two strategies present in the population, s_N and some other strategy

s. Denote by $\pi_N(l)$ the payoff for a player playing the Cournot-Nash strategy when l players in the match play the Cournot-Nash strategy (and thus $n-l$ players play s). Similarly, $\pi(l)$ denote the payoff of a player playing s in the same situation. To represent a Cournot oligopoly type game, for any strategy $s > s_N$, we assume that both $\pi_N(l)$ and $\pi(l)$ are increasing in l, $\pi_N(l) < \pi(l)$ $\forall l < n$ and $\pi_N(n) > \pi(n-1)$. We also assume the worst possible case for the Cournot-Nash strategy, $\pi_N(n-1) < \pi(0)$, that is, one deviator lowers the payoff of the other firms considerably. These assumptions represent general properties of the Cournot oligopoly, namely that an upward deviation from the Cournot-Nash strategy makes a firm worse off but hurts its competitors even more. For any strategy $s < s_N$ we assume that both $\pi_N(l)$ and $\pi(l)$ are decreasing in l, $\pi_N(n) > \pi(n-1)$ and $\pi_N(1) > \pi(0)$, that is, the Cournot-Nash strategy s_N is a better response to s than s itself.

According to the general model, players sample k other players. We assume that the direct opponents in a match are always observed, so $k = n + k'$. We assume that $k' > 0$ and sampling of other players is uniform, i.e. the probability to observe any other player is the same. As usual, after sampling the players copy the strategy that gives the highest payoff among sampled.

Let us first consider strategies $s > s_N$. Consider first the case $k' = 1$. Since we have only two strategies, the state of the system can be described by one variable, the proportion of the population playing one of the strategies. Denote by α_t the proportion of the population playing the Cournot-Nash strategy at time t. If a player playing s_N is matched with other players playing the Cournot-Nash strategy, she gets the highest possible payoff and therefore does not change her strategy. If the player is matched with some players who play the other strategy s, she switches to their strategy (since it gives a higher payoff and direct opponents are always observed) unless she samples a player getting the highest possible payoff. The last event occurs with probability α_t^n under uniform matching and sampling. Correspondingly, a player playing strategy s does not switch to the Cournot-Nash strategy unless he samples a player with the highest payoff. In a large (infinite) population, the population imitation dynamic can be described as

$$\alpha_{t+1} = \alpha_t - \alpha_t(1-\alpha_t^{n-1})(1-\alpha_t^n) + (1-\alpha_t)\alpha_t^n. \qquad (3.2)$$

The last term in the equation is the increase in the proportion of players playing the Cournot-Nash strategy due to the observation of the highest payoff and the second term on the right-hand side is the corresponding decrease. The equation has stationary states $0, 1$, and $\alpha^* \in (0,1)$. It is easily seen that 0 and 1 are stable stationary states in the sense that if the process starts near them, it converges to them. Then α^* is an unstable stationary state. Thus, both states 0 and 1 have non-empty basins of attraction and the process of imitation converges to one of them depending on the initial conditions unless it starts precisely at α^*.

As k' increases, the probability of sampling a player with the highest possible payoff increases, approaching 1 as k' goes to infinity, if $\alpha_t > 0$. The formula for the population dynamic for a general k' becomes

$$\alpha_{t+1} = \alpha_t - \alpha_t(1-\alpha_t^{n-1})(1-\alpha_t^n)^{k'} + (1-\alpha_t)(1-(1-\alpha_t^n)^{k'}). \tag{3.3}$$

For any $\alpha_t > 0$ it is possible to find k' large enough so that $\alpha_{t+1} > \alpha_t$. Then the dynamic converges to the state $\alpha = 1$, that is, everybody plays the Cournot-Nash strategy.

For production levels s lower that the Cournot-Nash one s_N, an analogous reasoning works even if we consider a situation that is worse for the Cournot-Nash strategy that the original assumptions. That is, we consider the situation when $\pi_N(2) < \pi(n-1)$ instead of $\pi_N(n) > \pi(n-1)$. Then strategy s brings a higher payoff unless there is a match where only one player plays s_N, when s_N gives the highest possible payoff. Formally, the population dynamic is then described by the equation

$$\begin{aligned} \alpha_{t+1} &= \alpha_t - \alpha_t(1-(1-\alpha_t)^{n-1}(1-\alpha_t(1-\alpha_t)^{n-1})^{k'} \\ &\quad +(1-\alpha_t)(\alpha_t(1-\alpha_t)^{n-2} + (1-\alpha_t(1-\alpha_t)^{n-2})(1-(1-\alpha_t(1-\alpha_t)^{n-1})^{k'} \end{aligned}$$

where the second term is the decrease in the proportion of the population playing s_N, and the last term is the increase in that proportion. As the number of observations of others k' increases there is a larger chance to observe a player playing s_N in a match with players playing s. Thus, we have

Theorem 3.3 *In the model with infinite population playing the Cournot oligopoly type game described above with uniform random matching and sampling, the Cournot-Nash strategy wins every two-strategy contest if the sample size k is large enough.*

So far we considered the dynamic without experimentation. If experimentation is introduced, a small number of experimentators in the state $\alpha = 0$ is enough to upset the state, since for any $\alpha_t > 0$ it is possible to find k' large enough so that $\alpha_{t+1} > \alpha_t$. Then $\alpha = 1$ is the long run outcome of the process.

Remark 3.2 *The Cournot-Nash strategy is the long-run outcome of the imitation and experimentation process in the game described above starting from the situation when only s_N and one other strategy s are present, if the sample size k is large enough.*

Palomino (1996) considers similar games and shows that the strictly dominated Walrasian strategy survives if players have information about less than the whole population. Our results differs from Palomino (1996) in two respects. First, the order of the limits is changed: we let first the sample size k grow keeping the initial proportions α_1 fixed. In Palomino (1996) it is shown that for any finite k there exists initial proportion α_1 such that the process converges to the strictly dominated strategy. Our result is the opposite: for any initial proportion α_1 there exists k such that the process converges to the Cournot-Nash strategy. Second, by introducing experimentation we look at the stochastic stability of the stationary states of the imitation dynamic starting from a situation with only two strategies. Both monomorphic states are steady states of the dynamic but with experimentation the Cournot-Nash strategy has an upper hand.

For a finite population a similar result can be achieved for a fixed local interaction model. Suppose the locations of the players are fixed. Players in a location interact among themselves but not with the outside world. However, each player also has information about some players in other locations. In the contest of s_N with a strategy $s > s_N$, if there is a location playing the efficient Nash equilibrium $(s_N, ..., s_N)$, the players in the location never change strategies since they have the highest possible payoff. Players in other locations sooner or later sample a player in the efficient location and change their strategy to the Cournot-Nash one. Therefore, the state when the whole population plays the Walrasian strategy can be upset by exactly n experimenting players if they happen to be in the same location. For the state when everybody plays the Cournot-Nash strategy it is not enough. The Nash equilibrium is then the long-run outcome. The contest of s_N with an $s < s_N$ has features of a prisoners' dilemma with s_N being the defecting strategy. Even if n cooperators convert some other player to cooperation, a defector will exploit these cooperators, which might induce the original cooperators to switch to defection too. This comes from the fact than mutual cooperation does not give the highest payoff but is beaten by the payoff of a defector against a cooperator.

Remark 3.3 *In a finite population with fixed local interactions playing the Cournot oligopoly type game above the Cournot-Nash equilibrium is the unique long-run outcome for any $k' > 0$, if initially only s_N and one other strategy s are present.*

In the context of industrial organization the model can be interpreted as one with several identical industries each having the same number of firms, or several identical geographical or other markets. In the random matching model firms are randomly assigned to markets while in the fixed local interaction model their locations are fixed. The firms can observe their own industry or market perfectly and they can also sample firms in other industries or markets.

Our result is similar to the one of Robson and Vega-Redondo (1996). In their model players can imitate the best on average action in the population. In 2×2 coordination games the efficient equilibrium is selected. In our model players can observe other locations and if they observe a more efficient one, they imitate it. In a sense players make their own location relatively efficient by imitation of other locations.

The results presented in this section show that when imitating, it is good to have some external information, that is, not only information about the direct opponents. The intuition for the result is simple and resembles the intuition for imitation in decision problems. If the efficient way to solve the game is never observed, as in the model of section 3.2, it cannot be found by imitation. However, with external information, if the efficient equilibrium exists somewhere in the population, it is found. The dynamic was constrained to two strategy contests; if there are more strategies originally, it can happen that the Cournot-Nash strategy is eliminated before the process goes down to a two strategy contest. Winning every two strategy contest, however, is a feature underlying stability of a strategy against an invasion by one other strategy, and we have demonstrated that the Cournot-Nash strategy is immune to such invasions.

3.5 Conclusion

Several models of imitation in Cournot oligopoly games are considered in this chapter. Though the results cast doubt on (perfect) imitation, imitative behavior is not as unattractive as it may seem: taken with care, it produces plausible results. Actually, one of the most celebrated strategies in the repeated prisoners' dilemma, "tit-for-tat" is nothing else than imitation of what the other player did in the previous period (though it does not take into account the payoff). Applied with care, that is starting from cooperation, "tit-for-tat" "solves" the dilemma. Therefore, imitative behavior should not be rejected without considering the game at hand.

Indeed, in the chapter we demonstrated that imperfection in imitation, and separation of imitation and interaction may have a Nash equilibrium as the outcome. Though for the latter model (separate imitation and interaction) the result is not surprising, since we require that players have more information than in perfect imitation, for the former model the result is surprising since imperfection reduces the information available to the players. This shows that "as if" rational behavior can be achieved by an increase in rationality (using more information) as well as by a decrease in rationality (using less information).

The model of imitation presented here is applicable only to symmetric games. Still, asymmetric games might be considered. For example, in a sequential extensive form game, it is possible to imitate the preceding player move. In chess, for example, there was a (wrong) belief that if Black imitates White, it can achieve a draw. Of course, in chess it is not true, but there might be games where it is. More applicable to economics, particularly to industrial organization, is the situation when there are asymmetries in information and a less informed agent (firm) may find it profitable to imitate a more informed agent (firm). Such situations are not considered in this chapter but the ones that are considered show that imitation does not necessarily lead to bad results.

Appendix

3.A Proof of Theorem 3.2

The technique to work with the notion of stochastic stability was introduced into the evolutionary game theory by Kandori et al. (1993) and Young (1993). First we recall a useful result for general Markov processes with perturbations. In its presentation we follow Young (1998, Ch.3).

Let P^0 is a finite Markov chain on the state space Z. A set $\{P^\varepsilon\}$ of perturbed processes is *regular* if $\forall \varepsilon \in (0, \varepsilon^*]$ P^ε is irreducible, $\forall z, z' \in Z$ $\lim_{\varepsilon \to 0} P^\varepsilon_{zz'} = P^0_{zz'}$, and if $P^\varepsilon_{zz'} > 0$ then $\exists r(z, z')$ such that $0 < \lim_{\varepsilon \to 0} \frac{P^\varepsilon_{zz'}}{\varepsilon^{r(z,z')}} < \infty$. The real number $r(z, z')$ is the *resistance* of the transition from z to z'. Each P^ε has a unique stationary distribution μ^ε. A state z is *stochastically stable* if $\lim_{\varepsilon \to 0} \mu^\varepsilon(z) > 0$.

Suppose that P^0 has recurrent classes $E_1, ..., E_k$. A *path* from E_i to E_j is a sequence of states $\zeta = (z_1, .., z_q), z_1 \in E_i, z_q \in E_j$. The resistance of the path is the sum of the resistances of transitions (z_i, z_j): $r(\zeta) = r(z_1, z_2) + ... + r(z_{q-1}, z_q)$. Let $r_{ij} = \min\, r(\zeta)$, where the minimum is over all possible paths from E_i to E_j.

For a given recurrent class E_i, a i-tree is a directed graph with recurrent classes as vertices such that

(i) each $E_j \neq E_i$ is the source of exactly one edge;

(ii) from every $E_j \neq E_i$ there is a path from E_j to E_i.

The resistance of an i-tree is the sum of the resistances of its edges. The minimal resistance over all i-trees is the *stochastic potential* of the recurrent class E_i.

Lemma 3.1 (Young (1993)) *Let $\{P^\varepsilon\}$ be a regular set of perturbed processes. Then $\lim_{\varepsilon\to 0}\mu^\varepsilon$ exists and the stochastically stable states are the ones contained in the recurrent classes of P^0 with minimal stochastic potential.*

The proof of Theorem 3.2 will be based on two lemmas. Recall that the demand function is denoted by $D(p)$ and the inverse demand function is denoted by $P(Q)$. Though the state is specified by the strategies of both firms $((q_1,p_1),(q_2,p_2))$, the state where both firms play the same strategy (q,p) is denoted simply by (q,p). Recall that the grids are chosen is such a way that $\forall q$ they contain p such that $p = P(2q)$, or, equivalently, $q = \frac{D(p)}{2}$.

Lemma 3.2 *$\forall(q,p)$ with $q > \frac{D(p)}{2}$ one experimentation is needed to arrive at $(\frac{D(p)}{2},p)$.*

Proof. Since $q > \frac{D(p)}{2}$, the firms oversupply the market, that is, they both sell actually only $\frac{D(p)}{2}$. Thus, if firm i reduces q_i to $\frac{D(p)}{2}$, it has the same revenue but a lower cost, and, therefore, a higher profit. The other firm follows. ■

Lemma 3.3 *$\forall(q,p)$ with $q < \frac{D(p)}{2}$ one experimentation is needed to arrive at $(q,P(2q))$.*

Proof. Consider such (q,p) that $q < \frac{D(p)}{2}$. If firm i increases price to $P(2q)$, its demand is still q as well as the demand of the other firm. Therefore, the revenue of the deviating firm is higher, while the cost is the same. The other firm imitates. ■

We are ready for the main proof.

Proof of Theorem 3.2. Let P^0 be the imitation process without experimentation. The set of perturbed processes is $\{P^\lambda\}$ where λ is the experimentation probability. The set is regular since the transition probabilities of the perturbed processes are polynomials in λ.

By Remark 3.1 the recurrent classes of P^0 are states where both firms play the same combination of capacities and prices (p,q). Thus, the imitation process without experimentation has $(v+1)(w+1)$ recurrent classes (recall that v is the cardinality of the capacity grid and w is the cardinality of the price space grid). The resistance of a transition from one recurrent class to another is always at least 1, since at least one experimentation is needed for it.

Since there are $(v+1)(w+1)$ recurrent classes, the minimal i-tree for any of them has at least resistance $(v+1)(w+1)-1$, therefore the minimal stochastic potential is at least $(v+1)(w+1)-1$. We will construct a tree with such a resistance for the state (q_N,p_N).

First, consider arbitrary (q,p) with $q \neq \frac{D(p)}{2}$. If $q > \frac{D(p)}{2}$ by Lemma 3.2 we can arrive at the curve $q = \frac{D(p)}{2}$ by decreasing capacity with one experimentation. If $q < \frac{D(p)}{2}$ by Lemma 3.3 we can achieve the curve $q = \frac{D(p)}{2}$ by increasing price with one experimentation. Therefore, the resistance of the transitions from (q,p) with $q \neq \frac{D(p)}{2}$ to (q,p) with $q = \frac{D(p)}{2}$ is 1.

Now consider an arbitrary (q,p) with $q = \frac{D(p)}{2}$. First consider $q < q_N$. Suppose one firm deviates to (q_N, p_N). Since $p_N < p$ and $q_N < D(p_N)$ the deviating firm sells q_N and has a positive profit since it receives the profit equal to the profit in the Nash equilibrium. If $q_N \geq 2q$, the remaining firm does not sell anything, thus having zero revenue and nonpositive profit. Therefore, the remaining firm will follow the deviation. If $q_N < 2q$, the remaining firm sells the residual demand $q' = 2q - q_N < q$ at price $p = P(2q) = P(q' + q_N)$, still producing q. By the definition of the Nash equilibrium $P(2q_N)q_N - C(q_N) \geq P(q' + q_N)q' - C(q')$. The left hand side expression is the profit of the deviating firm, the right hand side is the profit of the remaining firm if it would have capacity q'; since $q > q'$ the profit of the remaining firm is even smaller. Thus, the profit of the deviating firm is higher than the remaining firm and with one experimentation we have arrived to the Nash equilibrium.

Consider now an arbitrary (q,p) such that $q \geq q_N$ and $q = \frac{D(p)}{2}$ (or, equivalently, $p = P(2q)$). Suppose one firm deviates to the next available point on the grid and on the curve $q = \frac{D(p)}{2}$, that is, the deviating firm plays $(q + \delta, P(2(q + \delta))$. Since the deviating firm has lower price, it sells all its capacity and its profit is $(q+\delta)P(2q+2\delta) - C(q+\delta)$. The remaining firm sells only $D(P(2q)) - (q + \delta) = q - \delta > 0$ since δ can be chosen to be smaller than q_N. The profit of the remaining firm is then $(q - \delta)P(2q) - C(q)$. The difference in profits between the deviating and the remaining firm is $(2P(2q) + 2qP'(2q) - C'(q))\delta + O(\delta^2)$. By choosing an appropriately small δ', $\forall \delta \leq \delta'$ the sign of this expression is determined by the sign of $2P(2q) + 2qP'(2q) - C'(q)$.

Let $F(q) := 2P(2q) + 2qP'(2q) - C'(q)$. Then $F'(q) = 6P'(2q) + 4qP''(2q) - C''(q)$. Since $P(q)$ is decreasing and concave, the first two terms are negative, and since $C(q)$ is convex, the last term is also negative. Therefore, $F'(q) < 0$. Then $F(q)$ is decreasing and there exists a unique $\hat{q}$ such that $F(\hat{q}) = 0$, $\forall q < \hat{q}$ $F(q) > 0$, and $\forall q > \hat{q}$ $F(q) < 0$. This means that along the curve $q = \frac{D(p)}{2}$, $\exists \hat{q}$ such that $\forall q < \hat{q}$ a deviation to $q + \delta$ gets imitated, and $\forall q > \hat{q}$ a deviation to $q - \delta$ gets imitated. The process can arrive to $(\hat{q}, P(2\hat{q}))$ (or to the point of the grid closest to it) by unilateral experimentations along $q = \frac{D(p)}{2}$.

We have shown above that $\hat{q}$ cannot be smaller than q_N since $\forall q < q_N$, a firm at q_N

gets higher profit. Therefore, $\hat{q} \geq q_N$. If $\hat{q} = q_N$ then we are done. Consider $\hat{q} > q_N$. Suppose the firms are at $(\hat{q}, P(2\hat{q}))$ (or at the point of the grid closest to it). Consider now that one firm experiments with $(q_N, P(\hat{q}+q_N))$. The profit of the experimenting firm is $\pi' = q_N P(\hat{q}+q_N) - C(q_N)$, the profit of the other firm is $\pi = \hat{q}P(2\hat{q}) - C(\hat{q})$. The difference $\pi' - \pi = q_N P(\hat{q}+q_N) - C(q_N) - \hat{q}P(2\hat{q}) + C(\hat{q}) = q_N P(2q_N) - C(q_N) + q_N P(\hat{q}+q_N) - q_N P(2q_N) - \hat{q}P(2\hat{q}) + C(\hat{q})$. By the definition of Nash equilibrium $\pi' - \pi \geq \hat{q}P(\hat{q}+q_N) + q_N P(\hat{q}+q_N) - q_N P(2q_N) - \hat{q}P(2\hat{q})$. From the assumptions on the inverse demand function it follows that $P''(q+q')q + 2P'(q+q') < 0$, which implies that $(P(q+q')q)'' < 0$, or that $P(q+q')q$ is concave for any q, q'. The last fact implies that $P(\alpha q_1 + (1-\alpha)q_2 + q')(\alpha q_1 + (1-\alpha)q_2) > \alpha P(q_1+q')q_1 + (1-\alpha)P(q_2+q')q_2$ $\forall q_1, q_2, q'$ $\forall \alpha \in (0,1)$. Consider $q_1 = 2q_N, q_2 = 2\hat{q}, q' = 0, \alpha = \frac{1}{2}$. Then we have $(\hat{q}+q_N)P(\hat{q}+q_N) > q_N P(2q_N) + \hat{q}P(2\hat{q})$, that is, $\pi' - \pi > 0$. Therefore, a firm experimenting with $(q_N, P(\hat{q}+q_N))$ gets a higher profit and gets imitated. If $P(\hat{q}+q_N)$ does not belong to the grid, the grid can be made finer until the point of the grid closest to $P(\hat{q}+q_N)$ brings a higher profit. Observe that making the grid finer does not change $\hat{q}$, the point to which the process can arrive along the curve $q = \frac{D(p)}{2}$ by one experimentation.

Finally, from $(q_N, P(\hat{q}+q_N))$, one arrives to $(q_N, P(2q_N))$ by one experimentation since $P(\hat{q}+q_N) < P(2q_N)$. We have constructed a i-tree for the Nash equilibrium with resistance of each edge equal to 1, thus the resistance of the tree is $(v+1)(w+1)-1$, which is the minimal possible resistance among i-trees for any stationary state. Therefore, the Nash equilibrium has non-zero weight in the limit stationary distribution, that is, the Nash equilibrium is stochastically stable. ∎

Chapter 4

Indirect Evolution in Duopoly

4.1 Introduction

A fundamental assumption in economics is that economic agents care only about their own payoff or profit and do not take into account the payoff or profit of others. This point was defended by Alchian (1950) and Friedman (1953) by stating that if it were not the case, agents not maximizing their profit would be eliminated by evolution. It has been shown recently that in a strategic interaction context with complete information this is not necessarily true (see, e.g. Bester and Güth (1998) and Koçkesen et al. (1999)) while an incomplete information setting supports the claim (see Ely and Yilankaya (1997) and Ok and Vega-Redondo (1999)).

This chapter applies the indirect evolution approach, described informally below and formally in the next section, to games which can arise from differentiated product oligopoly games. The indirect evolution approach, initiated by Güth and Yaari (1992), works on preferences rather than on strategies. The material payoff game is given, but the players may have preferences that differ from maximization of the material payoffs. They play the game with their subjective preferences. Players are either rational, or learn fast enough so that they play an equilibrium of the game with subjective preferences. Given equilibrium strategies, one can calculate the fitness of a player by substituting the equilibrium strategies into the material payoff game. Evolution selects preferences that have a higher fitness. We are interested in the stable stationary points of the evolutionary process, that is, such preferences that the population is robust to invasion of a small number of mutants with some other preferences. These points are the evolutionarily stable strategies of a game where the strategy sets are all feasible preferences and the payoffs are the material payoffs corresponding to the equilibrium strategies of the game with given subjective preferences.

Evolution is assumed to work on preferences in a large population of players who are randomly matched to play a two-player game. In the game they are either able to find an equilibrium or learn to play it sufficiently fast so that evolution uses the material payoff outcome corresponding to the equilibrium strategies. In the duopoly context evolution can be interpreted as a cultural phenomenon changing preferences (priorities) of a given firm as it observes profits and priorities of other firms in other duopoly markets.

Bester and Güth (1998) have analyzed the model with preferences restricted to a convex combination of *egoistic* (maximizing own profit) and *altruistic* (maximizing the sum of profits) preferences. In a quadratic setting they showed that (partial) altruism is evolutionarily stable with strategic complements. With strategic substitutes no altruism survives and purely egoistic preferences are evolutionarily stable. The last result appears because the preference parameter is restricted to lie between egoistic and altruistic values. We remove this restriction on the preference parameter by introducing *spiteful* preferences (maximizing the difference of profits).

Altruism resembles positive reciprocity in the sense that one may want to reward the opponent who is good to him. The spiteful preferences, in turn, resemble negative reciprocity in the sense that one may want to hurt somebody who lets him down, and therefore they are not uncommon. They are also linked to imitative behavior which was analyzed in the previous chapter, since caring about the relative payoff means that a higher profit of the opponent makes you feel worse. We show that (partially) spiteful preferences are evolutionarily stable when egoism was stable in Bester and Güth (1998).

Bester and Güth (1998) make a parallel between the strategic properties of the game (complementarity or substitutability) and the result of evolution. Strategic complementarity or substitutability are determined by the slope of the reaction function. Since the reaction function depends on (subjective) preferences strategic properties can be *subjective* or *objective* (when we consider the material payoff game). Evolutionarily stable preferences determine subjective strategic complementarity or substitutability. Therefore, subjective strategic properties are the result of evolution while objective strategic properties of the material payoff game are given data that determine how evolution works. The basic element of the model is the material payoff function and its properties, such as super- or sub-modularity can provide some light for the result. We will see, however, that they are not always enough to provide an unambiguous answer for games going beyond standard differentiated product oligopoly games but which have a similar structure.

Finally, if the preferences of the opponent are not known, we show that maximizing the material payoffs is the only evolutionarily stable preferences and, therefore, with incomplete information the claim of Alchian (1950) and Friedman (1953) is valid in the context analyzed in the chapter. As referred in the first paragraph of the chapter, this result holds for several indirect evolution models.

We proceed as follows. In Section 4.2 we introduce the indirect evolution approach and consider the Bester and Güth (1998) framework with spiteful preferences. In Section 4.3 we relax some restrictions of their model and illustrate that subjective strategic complementarity or substitutability is endogenously determined. Section 4.4 considers the model with incomplete information. Some conclusions are drawn in Section 4.5.

4.2 The Duopoly Model with Spiteful Preferences

4.2.1 The Indirect Evolution Approach

Here we formulate briefly how the indirect evolution works for general games. In the following subsection we analyze the model of Bester and Güth (1998) with spiteful preferences.

Güth and Yaari (1992) initiated the indirect evolution approach to the evolution of preferences in games. Consider a symmetric two-player game $G = (\{1,2\}, S, U)$. Denote the strategy of player 1 by s_1 and the strategy of player 2 by s_2. The payoffs $U_1(s_1, s_2), U_2(s_1, s_2)$ are the *material payoffs* or the *fitness* of players 1 and 2. Let W be a subset of the set of functions from strategy profiles (or outcomes) to real numbers, $W \subset \{V(\cdot) : S^2 \to \mathbb{R}\}$. The set W contains the possible subjective utility functions of the players and is called the set of feasible preferences.

Consider now a game $G' = (\{1,2\}, S, \{V_1, V_2\})$ with the same strategy sets but with players whose preferences are represented by the subjective utility functions $V_1(s_1, s_2)$, $V_2(s_1, s_2) \in W$. We can find an outcome of G' that is considered plausible, for example, a Nash equilibrium of G'. If G' has a unique equilibrium, it is natural to assume that the equilibrium is played. Even if G' has multiple equilibria, there are ways to select one of them (see Harsanyi and Selten (1988) or the learning process of Kandori et al. (1993)). We will address the problem of non-uniqueness of the equilibrium in our particular example later. Observe that we assume that the subjective utility functions V_1, V_2 are common knowledge.

Assume that for any given pair of preferences V_1, V_2 from W a unique outcome of the game is found, together with the corresponding strategies s_1^*, s_2^*. That is, there is a function F from $W \times W$ to the set of strategy profiles $S \times S$ that gives a "solution" for each G'. Then one can find the fitness of a player with given preferences in the particular match by substituting the equilibrium strategies into the material payoff function. Let us denote the fitness of player 1 by $U_1^*(V_1, V_2) := U_1(F(V_1, V_2))$. The material payoff for player 2 is found by symmetry, that is, $U_2^*(V_1, V_2) = U_1^*(V_2, V_1)$.

Define an *evolutionary game* $\Gamma = (\{1,2\}, W, U^*)$ as a symmetric game with strategy sets W and with payoff function $U_1^*(V_1, V_2)$. The definition of evolutionary stability for one-population symmetric games is standard (see, e.g. Weibull (1995, Ch.2))

Definition 4.1 *Strategy V^* is* ***evolutionarily stable*** *if*

(a) $U_1^*(V^*, V^*) \geq U_1^*(V, V^*)\ \forall V \in W$,

(b) if $V \neq V^*$ *and* $U_1^*(V^*, V^*) = U_1^*(V, V^*)$, *then* $U_1^*(V^*, V) > U_1^*(V, V)$.

Evolutionary stability is a refinement of Nash equilibrium since condition (a) coincides with the condition for Nash equilibrium. Moreover, if a Nash equilibrium is strict, it is evolutionarily stable since then condition (a) is satisfied with inequality and condition (b) is irrelevant.

The evolutionarily stable preferences are the ones that are evolutionarily stable in the evolutionary game Γ.

Definition 4.2 *Preferences V^* are* ***evolutionarily stable*** *if they are an evolutionarily stable strategy in Γ.*

The story behind the definition is as follows. Suppose that an arbitrarily small proportion of mutants with preferences V appears in the monomorphic population of players with preferences V^*. Since the proportion is arbitrarily small what matters more is the performance against the original preferences V^* rather than the performance against the mutant preferences V. Therefore, if $U_1^*(V^*, V^*) > U_1^*(V, V^*)$ the mutants will be driven out of the population. When $U_1^*(V^*, V^*) = U_1^*(V, V^*)$ then the matches against the mutants start to play a role and the mutant will grow if it has a higher material payoff against itself than the incumbent has against the mutant.

Note that the game Γ has as a strategy set the set of feasible preferences W. Since the set of equilibria of game Γ (in particular, the set of evolutionarily stable strategies) is affected by the changes in the strategy set, which preferences are evolutionarily stable depends on W.

Finding the evolutionarily stable preferences allows us to say which preferences are robust with respect to an invasion of a small number of mutants in a large population with random matching. Evolutionary stability does not necessarily guarantee that an evolutionary process will converge to the evolutionarily stable strategy. It can also happen that there is no evolutionarily stable strategy. However, we will see that this concept allows us to draw conclusions in certain games.

4.2.2 The Quadratic Example

Let $G = (\{1,2\}, \mathbb{R}^+, U)$ be a symmetric game with strategy sets of all nonnegative real numbers. Let $x \geq 0$ be the strategy of player 1 and $y \geq 0$ be the strategy of player 2. The material payoffs of the game are given by

$$U_1(x,y) = x(ky + m - x),\ U_2(x,y) = y(kx + m - y) \tag{4.1}$$

where $-1 < k < 1, k \neq 0, m > 0$. These restrictions, together with the assumptions on preferences introduced later, will guarantee uniqueness of the equilibrium in games with any feasible preferences .

Such a payoff function appears, for example, in differentiated product oligopoly games, where strategies are quantities produced. The demand function is linear. If the products are complements ($k > 0$), an increase in the quantity of one product leads to a reduction in its price but to an increase in the price of the other product, since the other product becomes more valuable. If the products are substitutes ($k < 0$), increasing the supply of one product reduces the price of both products. Then the above functions are the revenue functions of the firms producing the two products and, if costs are zero, they coincide with the profit functions.

The players do not necessarily maximize their material payoffs. The set of feasible preferences W consists of the following linear combinations of own and opponent's payoffs:

$$V_1(x,y) = U_1(x,y) + \alpha U_2(x,y),\ V_2(x,y) = U_2(x,y) + \beta U_1(x,y) \tag{4.2}$$

where $\alpha, \beta \in [-1,1]$ are the preference parameters. Thus, the subjective preferences are characterized by the preference parameter. The set W is then equivalent to the interval $[-1,1]$. This formulation is slightly different from the one of Bester and Güth (1998) but easier to work with. The bounds on the parameters will guarantee the uniqueness of equilibrium. We will relax them later overcoming the problem of non-uniqueness by extending the fitness function.

If $\alpha = 0$, player 1 maximizes his material payoff. A preference parameter $\alpha > 0$ means that the player is altruistic, that is, takes into account the opponent's profit with a positive weight, while $\alpha < 0$ represents spiteful preferences, that is, ones where the opponent's profit reduces one's utility. The same description is valid for player 2's parameter β. The parameters α, β are common knowledge for the players. This assumption is important and the consequences of its relaxation will be analyzed in Section 4.4. The evolutionary game is $\Gamma = (\{1,2\}, [-1,1], U^*)$, where U^* is the fitness function whose closed form in the quadratic example is found below.

When two players are matched they maximize their corresponding subjective utility functions V_1, V_2 with parameters α, β. The reaction functions derived from the first order conditions (the second order conditions are always satisfied) are

$$x = \frac{k(\alpha+1)y+m}{2}, \quad y = \frac{k(\beta+1)x+m}{2}. \tag{4.3}$$

Following Bester and Güth (1998), we say that the strategies are complements, or that the game exhibits *strategic complementarity,* when the reaction functions are upward sloping. The strategies are substitutes, or the game exhibits *strategic substitutability,* if the reaction functions are downward sloping. The sign of the slope of the reaction functions above depends only on k since $\alpha + 1$ and $\beta + 1$ are always nonnegative. If $k > 0$ then the reaction functions are upward sloping, thus strategies are complements. If $k < 0$, the reaction functions are downward sloping and strategies are substitutes.

The unique equilibrium of the game is given by

$$x^*(\alpha,\beta) = \frac{m(k(\alpha+1)+2)}{4-k^2(\alpha+1)(\beta+1)}, \quad y^*(\alpha,\beta) = \frac{m(k(\beta+1)+2)}{4-k^2(\alpha+1)(\beta+1)}. \tag{4.4}$$

Given the equilibrium strategies, the fitness of the player with preference parameter α is the material payoff she gets in the equilibrium. This defines the evolutionary game Γ on preferences.

The material payoff of player 1 as a function of preference parameters through the implied equilibrium strategies of both players is

$$U_1^*(\alpha,\beta) = -\frac{m^2(k(\alpha+1)+2)(k^2\alpha(\beta+1)+k(\alpha-1)-2)}{(4-k^2(\alpha+1)(\beta+1))^2} \tag{4.5}$$

while the material payoff function of player 2 satisfies $U_2^*(\beta,\alpha) = U_1^*(\alpha,\beta)$. The function $U_1^*(\alpha,\beta)$ is differentiable given the restrictions on α, β, and k.

An evolutionarily stable strategy for the evolutionary game Γ on preference parameters is a parameter α^* satisfying

(a) $U_1^*(\alpha^*, \alpha^*) \geq U_1^*(\alpha, \alpha^*)\ \forall \alpha \in [-1, 1]$,

(b) if $\alpha \neq \alpha^*$ and $U_1^*(\alpha^*, \alpha^*) = U_1^*(\alpha, \alpha^*)$, then $U_1^*(\alpha^*, \alpha) > U_1^*(\alpha, \alpha)$.

Condition (a) requires (α^*, α^*) to be a symmetric equilibrium of the evolutionary game Γ. To check condition (a), we fix the second argument of $U_1^*(\alpha, \beta)$, find maxima of $U_1^*(\alpha, \beta)$ with respect to the first argument and equate the arguments. The first order condition is

$$\alpha = -\frac{k(\beta+1)(k+2)}{\beta k(k-2) + k^2 - 2k - 4}. \tag{4.6}$$

After equating β to α, possible candidates for an evolutionarily stable preference parameter are

$$\alpha_1^* = -\frac{k+2}{k},\ \alpha_2^* = \frac{k}{2-k}. \tag{4.7}$$

The boundary values $\alpha^* = -1$ and $\alpha^* = 1$ are possible candidates too. These four values are the only possible candidates for evolutionary stability. Consider $\beta = -1$. Then $U_1^*(0, -1) - U_1^*(-1, -1) = \frac{m^2(k+2)^2}{16} - \frac{m^2(k+1)}{4} = \frac{m^2k^2}{4} > 0$. Thus, $\alpha^* = -1$ cannot be evolutionarily stable as it is not a best reply against itself in the evolutionary game Γ. If $\beta = 1$ then $\frac{\partial U_1^*(\alpha,1)}{\partial \alpha}|_{\alpha=1} = -\frac{m^2k^2}{8(k+1)(k-1)^2} < 0$ thus $U_1^*(\alpha, 1)$ is decreasing in α around $\alpha = 1$. It implies that there exist an $\alpha' < 1$ such that $U_1^*(\alpha', 1) > U_1^*(1, 1)$. Thus, $\alpha = 1$ is not a best reply against itself either and, therefore, is not evolutionarily stable.

Note that, for $-1 < k < 1$, α_1^* is never between -1 and 1 thus we can ignore it. Condition (a) of evolutionary stability for α_2^* says

$$\begin{aligned}
U_1^*(\frac{k}{2-k}, \frac{k}{2-k}) - U_1^*(\alpha, \frac{k}{2-k}) &\geq 0\ \forall \alpha \iff \\
\frac{m^2(k+2)(k-2)}{16(k-1)} - \frac{m^2(k+2)(k-2)(k^2(\alpha+1)^2-4)}{4(k^2(\alpha+1)+2k-4)^2} &\geq 0\ \forall \alpha \iff \\
(k^2(\alpha+1)+2k-4)^2 - (k^2(\alpha+1)^2-4)4(k-1) &\geq 0\ \forall \alpha \iff \\
(k(\alpha+1)-2\alpha)^2 &\geq 0\ \forall \alpha.
\end{aligned}$$

The last inequality is always satisfied and turns to equality only when $\alpha = \frac{k}{2-k}$. This means that α_2^* is the unique best response against itself, that is, it is a strict Nash equilibrium of the evolutionary game Γ, and, therefore, evolutionarily stable. We have

Theorem 4.1 *With preference parameters from $[-1, 1]$, in the quadratic model with $-1 < k < 1$, $k \neq 0$, the unique evolutionarily stable preference parameter is $\alpha^* = \frac{k}{2-k}$.*

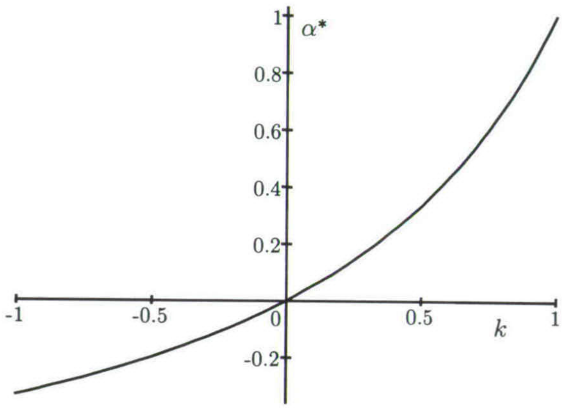

Figure 4.1: Evolutionary stable α.

This result extends the Bester and Güth (1998) example to spiteful preferences. For positive k the result is the same: some amount of altruism is evolutionarily stable. News comes when k is negative. If $k < 0$ then $\alpha^* < 0$. This means that spiteful preferences are evolutionarily stable when the strategies of the game G are substitutes. An interesting difference between positive and negative values of k is that as $k \to 1$, $\alpha^* \to 1$, that is, altruism becomes "perfect" as the degree of complementarity between strategies becomes perfect, while if $k \to -1$, $\alpha^* \to -\frac{1}{3}$, that is, spite does not gain full strength even when the degree of substitutability is perfect. These results are illustrated in Figure 4.1.

The intuition for the result is as follows. There are two material effects of a change in the preference parameter of a player. One is the change induced by the change in own equilibrium strategy and the other is the change induced by the change in the opponent's equilibrium strategy. A preference parameter can be evolutionarily stable when the two effects compensate each other, thus, no other nearby preference parameter has a higher material payoff.

Notice that an increase in a player's preference parameter always increases the equilibrium strategy of the opponent. The material payoff function is such that with strategic complements ($k > 0$), a higher strategy increases the opponent's payoff, thus a player who is more altruistic should increase her strategy that through complementarity leads to an increase in the strategy of the other player. With strategic substitutes ($k < 0$) a lower strategy increases the opponent's payoff, thus a more altruistic player chooses a lower strategy, but the strategies are substitutes so that the other player chooses a higher strategy in the equilibrium. The choice of a higher strategy by the opponent leads to an increase in the material payoff if $k > 0$, and to a decrease if $k < 0$.

With strategic complementarities ($k > 0$), an increase in a player's preference parameter means a higher strategy of the opponent and a higher own strategy. A higher opponent's strategy leads to an increase in the material payoff. A higher own strategy means an increase in the opponent's profit. If the players are spiteful, such an increase has to be compensated in the equilibrium by an increase in own profit. Then both effects point to an increase in the material payoff; therefore, a higher preference parameter would have performed better. Thus, spiteful preferences cannot be evolutionarily stable. Altruistic preferences require that in equilibrium an increase in the profit of others is offset by a decrease in the own profit. Then the effects point to different directions and when they offset each other, such altruistic preferences are possibly evolutionarily stable.

With strategic substitutes ($k < 0$), lowering one's level of altruism leads to a decrease in the opponent's strategy and to an increase in the own strategy. The decrease in the opponent's strategy leads directly to an increase in the material payoff. The increase in the own strategy means a decrease in the opponent's profit; since players are altruistic, in the equilibrium this decrease has to be compensated by an increase in own material payoff. Thus both effects work in the same direction and altruistic preferences cannot be evolutionarily stable with strategic complementarities. For spiteful preferences the equilibrium strategy is such that the decrease in the opponent's profit is compensated by a decrease in own profit; the effects work in different ways and when they are equal, such spiteful preferences are evolutionarily stable.

An implication of the above result is that in games with strategic substitutes it pays to be spiteful i.e. to care also about relative payoff (though to a certain degree) since lowering preference parameter leads to an increase in the material payoff. In a Cournot oligopoly, for example, it is evolutionarily stable to have preferences not only over own profit but also over market share since maximizing own market share implies minimizing the one of the opponent (see, e.g. Dufwenberg and Güth (1999)). In a Bertrand oligopoly it works the other way round, that is, it is evolutionary stable to care about the opponent's profit to avoid the cut-throat competition leading to low prices.

4.3 Further Extensions

A logical step further than the extension of the preferences to allow spiteful ones is to remove the restrictions on the preference parameter altogether thus allowing it to vary from $-\infty$ to $+\infty$. When α can vary from $-\infty$ to $+\infty$, we can represent preferences such as pure spite, i.e., minimizing the opponent's payoff ($\alpha \to -\infty$), relative profit maximization ($\alpha = -1$), maximizing own material payoff ($\alpha = 0$), maximizing the sum of the payoffs ($\alpha = 1$), pure altruism, i.e., maximizing the other's payoff ($\alpha \to +\infty$), and all preferences in between. This range covers a much larger span of preferences than the original model of Bester and Güth (1998).

One could argue that the extension of the preference parameter to $\alpha > 1$ is not natural and can lead to paradoxes. In my view, it is not true. First, $0 < \alpha \leq 1$ looks rather like social planner preferences and can hardly be called altruistic at all. Real altruism means that the other's payoff is more important than own, like risking own life to rescue others. Second, altruism does not lead to paradoxes more than egoism does. For example, in the Nash demand bargaining game with known preferences, the Pareto frontier contains equilibria with altruistic preferences as well as with egoistic and spiteful preferences. In complex environments, there could be no equilibria even with egoistic preferences. Thus, one should not reject altruistic preferences from the start.

There are two reasons why the above approach does not generalize automatically with this extension. The first reason is that if one keeps the restrictions $x \geq 0, y \geq 0$ then corner solutions often appear. This would require further assumptions on the cases with no or more than one equilibria than the ones below. Doing this would make the analysis more complicated without changing the qualitative results. Furthermore, the assumption of the nonnegative strategy space comes from the oligopoly interpretation of the game. However, if one sticks strictly to the standard oligopoly interpretation, one has to check also nonnegativity of the variables representing both quantities and prices. This would complicate the analysis even further and lead to cases where no obvious extension of the evolutionary game to cases with no equilibria would be possible. Therefore we remove the nonnegativity restriction altogether and consider as strategy space the whole real line $\mathbb{R}$.

Thus, the (symmetric) material payoff game is $G' = (\{1,2\}, \mathbb{R}, U)$, where $U(x,y)$ are given in (4.1) with $-1 < k < 1, k \neq 0, m > 0$. The players maximize their subjective utilities (4.2) with preference parameters $\alpha, \beta \in \mathbb{R}$.

It is possible to give an economic interpretation of the strategy space as the real line in the differentiated products oligopoly. Suppose that the market has already accumulated stocks of the two products. Then firms may still sell on the market (positive quantities supplied) or they may buy (withdraw their products) from the market (negative quantities supplied). Again, the "demand" functions are linear. A negative quantity and positive price for own product mean that a firm has to pay for withdrawing its product from the market because the market wants to keep it. A positive quantity and negative price mean that a firm has oversupplied the market and has to pay the consumers to accept its product, or, alternatively, has to pay a fine for producing too much. A negative quantity and negative price mean that the market wants to get rid of the product, so it pays the firm for withdrawing the product. For example, when the products are complements, such a situation may occur when the two products are not dangerous in combination but each of them is dangerous in itself.

Other economic interpretations of the real line strategy space and the payoff function (4.1) include bidding (or asking subsidies) for franchises when the proportions allocated to each firm depend on the difference in bids (asks), an effort game where an effort can be positive (performing a task) or negative (destructing, breaking a tool, where the unit payoff to a negative effort can be positive if the tool becomes dangerous, e.g. out of control), international duopoly where governments can impose taxes (or subsidies) on own firm, and investment games when a disinvestment (negative investment) is possible. We do not focus on any particular interpretation but on the abstract features of the game.

The second problem is that there may be no equilibrium of the game with given preference parameters, or there may be more than one (a continuum) of equilibria. In such a case the evolutionary game Γ is not well defined. A possible approach is to extend the fitness function $U_1^*(\alpha, \beta)$ to such preference parameters by continuity in the first argument.

It follows from the equilibrium strategies equation (4.4) that the game with preference parameters α', β' does not have a unique equilibrium if $4 - k^2(\alpha'+1)(\beta'+1) = 0$. For such α', β' we extend the fitness function by continuity in the first argument as $U_1^*(\alpha', \beta') = lim_{\alpha \to \alpha'} lim_{\beta \to \beta'} U_1^*(\alpha, \beta)$. This limit always exists on the extended real line $\mathbb{R} \cup \{\pm\infty\}$. We choose continuity in the first argument since then we can use first order conditions to find an evolutionarily stable strategy.

The evolutionary game is $\Gamma = (\{1,2\}, \mathbb{R}, U^*)$, where the closed form of the fitness functions U^* is given in (4.5) and in the points of discontinuity it is extended as in the previous paragraph. Given this extension of the fitness function, the evolutionary game is well defined and we can apply the concept of evolutionarily stable strategy. All the derivations of the previous section go through except that now there are no boundary candidates and that $\alpha_1^* = -\frac{k+2}{k}$ from (4.7) is a legitimate candidate for an evolutionarily stable strategy. With the above extension of the fitness function $U_1^*(\alpha, -\frac{k+2}{k}) = \frac{m^2}{4}\ \forall\alpha$. Condition (a) for evolutionary stability is satisfied with equality for any α. However, $U_1^*(-\frac{k+2}{k}, 0) = 0 < U_1^*(0,0) = \frac{m^2}{(k-2)^2}$, thus condition (b) is not satisfied and $\alpha_1^* = -\frac{k+2}{k}$ is not evolutionarily stable. The proof of the evolutionary stability of $\alpha_2^* = \frac{k}{2-k}$ goes through. We have

Theorem 4.2 *Even with preference parameters from $\mathbb{R}$, in the quadratic model with $-1 < k < 1, k \neq 0$, the unique evolutionarily stable preference parameter is $\alpha^* = \frac{k}{2-k}$.*

Notice that the signs of the slopes of the reaction functions (4.3) depend now on the preference parameters in an essential manner. Depending on whether the preference parameter is smaller or larger than -1, the game changes from one with strategic complements to one with strategic substitutes and vice versa depending on the sign of k. Thus, we need to distinguish between *objective* strategic complementarity or substitutability (when the reaction functions come from maximizing the material payoff) and *subjective* strategic complementarity or substitutability (when the reaction functions come from maximizing subjective utilities that may not coincide with the material payoffs). The subjective strategic complementarity or substitutability is determined endogenously, depending on what value of the preference parameter is evolutionarily stable.

In order to avoid confusion between objective and subjective strategic properties, it is better to use the notion of super- and sub-modularity. A twice differentiable function of two variables $U(x,y)$ is *super-modular* if $\frac{\partial^2 U}{\partial x \partial y} > 0\ \forall x, y$ and it is *sub-modular* if $\frac{\partial^2 U}{\partial x \partial y} < 0\ \forall x, y$. Bulow et al. (1985) actually use this definition to define strategic complements and substitutes. If players maximize the material payoffs then objective strategic complementarity and super-modularity are equivalent as well as objective strategic substitutability and sub-modularity.

With $-1 < k < 1$ if the function is super-modular ($\frac{\partial^2 U}{\partial x \partial y} = k > 0$) then the game with evolutionarily stable preference parameter $\alpha^* = \frac{k}{2-k}$ exhibits strategic complements, while if the function is sub-modular ($\frac{\partial^2 U}{\partial x \partial y} = k < 0$), the strategies of the resulting game are substitutes. Thus, for $-1 < k < 1$, there is still a one-to-one correspondence between subjective strategic complementarity (substitutability) and super-modularity (sub-modularity). A natural question to ask is whether this property holds for other values of k.

The parameter k measures the degree of interdependence between players' strategies. Though the differentiated product oligopoly interpretation of the game is hard to justify with $|k| > 1$ (own product usually influences own price more than the other's price), some other interpretations allow for $|k| > 1$. For example, in the effort game, the relation between tasks can be such that players perform tasks that are more important for the other player; in international duopoly it can be that the production of the firm in the other country is more important than the output of the firm in own country.

Given the extension of the fitness function, the evolutionary game on preference parameters is well defined for any k. Thus we only have to check the two possible candidates for evolutionarily stable preference parameters from (4.7). The above proof that $\alpha_1^* = -\frac{k+2}{k}$ is not evolutionarily stable works for any $k \neq 0$. Therefore, we are left with only $\alpha_2^* = \frac{k}{2-k}$ to check.

The second order condition for α_2^* to be a best reply against itself requires that

$$\frac{\partial^2 U_1^*(\alpha, \alpha_2^*)}{\partial \alpha^2} \Big|_{\alpha=\alpha_2^*} \leq 0 \Longleftrightarrow \frac{k^2 m^2 (k+2)(k-2)^5}{512(1-k)^3} \leq 0.$$

The last inequality holds when $k \in [-2, 1] \cup [2, +\infty)$. For other values of k, α_2^* is a strict local minimum, therefore it cannot be evolutionarily stable. When the second order condition is satisfied, the proof that α_2^* is the unique global maximum against itself from the previous section goes through. For boundaries one can check that α_2^* is evolutionarily stable for $k = -2$ but not for $k = 1$ or $k = 2$. This gives us the following

Theorem 4.3 *With preference parameters from* $\mathbb{R}$, *in the quadratic model with* $k \in \mathbb{R}, k \neq 0$, *there is*

(i) a unique evolutionarily stable preference parameter $\alpha^* = \frac{k}{2-k}$ *if* $k \in [-2, 1) \backslash \{0\} \cup (2, +\infty)$;

(ii) no evolutionarily stable preference parameter otherwise.

While for $k \in [-2, -1]$ the result is a natural extension of the previous result, indicating that the evolutionarily stable preference parameter has some degree of spite ($\alpha^* = -\frac{1}{2}$ when $k = -2$), a new result appears when $k > 2$. Evolutionarily stable α^* is negative and larger than one in absolute value when $k > 2$. Thus, a large degree of spite, up to minimizing opponent's payoff ($\alpha^* \to -\infty$ when $k \to 2$ from above) is evolutionarily stable when the degree of interdependence between players' strategies is high, though it is a positive interdependence!

The result also shows that the question posed above has a negative answer. For some k outside the interval $(-1, 1)$ no evolutionarily stable preference parameter exists, so we cannot say anything about the properties of the game played with the evolutionarily stable preferences. More important, when $k > 2$, the material payoff function $U(x, y)$ is supermodular while the game with the evolutionarily stable parameter α^* exhibits strategic substitutes. Thus, the property of subjective strategic complementarity or substitutability of the game played with the evolutionarily stable preferences does not have a one-to-one relation with super- or sub-modularity of the material payoff function.

4.3.1 General Payoff Function

A generalization of the material payoff function along the lines of Bester and Güth (1998) is also possible with spiteful preferences. We denote the material payoff function $U(x, y) := U_1(x, y)$ to simplify notation. Assume that the preference parameters belongs to a set $W \subset \mathbb{R}, W \ni 0$. Assume that the subjective utility functions $V(x, y) := V_1(x, y) = U(x, y) + \alpha U(y, x)$ are strictly concave for any $\alpha \in W$. Denote by U'_x the derivative of a function $U(x, y)$ with respect to the first argument and by U'_y the derivative with respect to the second argument. The notation for the second derivatives is analogous.

The necessary condition for evolutionary stability of a preference parameter α reads $\frac{\partial U(x^*(\alpha,\beta),y^*(\alpha,\beta))}{\partial \alpha}|_{\alpha=\beta} = 0$, where x^*, y^* are equilibrium strategies. By the chain rule, the necessary condition becomes

$$U'_x x^{*\prime}_\alpha + U'_y y^{*\prime}_\alpha |_{\alpha=\beta} = 0. \tag{4.8}$$

The terms in the equation are the effects referred to in the previous section: the effect of a change in own equilibrium strategy and the effect of a change in opponent's equilibrium strategy.

The equilibrium strategies are found by maximization of the subjective utility function $V(x, y) = U(x, y) + \alpha U(y, x)$ and the corresponding function for player 2. Then $x^{*\prime}_\alpha$ and $y^{*\prime}_\alpha$ can be found by the implicit function theorem. Notice that $\alpha = \beta$ implies $x^* = y^*$ because of symmetry. Then we have the following expressions for $x^{*\prime}_\alpha$ and $y^{*\prime}_\alpha$ when $\alpha = \beta$:

$$x^{*\prime}_\alpha = \frac{-U'_y(U''_{xx} + \alpha U''_{yy})}{(U''_{xx} + \alpha U''_{yy})^2 - (U''_{xy}(1+\alpha))^2}, \tag{4.9}$$

$$y^{*\prime}_\alpha = \frac{(1+\alpha)U''_{xy}U'_y}{(U''_{xx} + \alpha U''_{yy})^2 - (U''_{xy}(1+\alpha))^2}. \tag{4.10}$$

Assume that the material payoff function satisfies $(U''_{xx} + \alpha U''_{yy})^2 - (U''_{xy}(1+\alpha))^2 > 0$ $\forall \alpha \in W, \forall x, y \geq 0$. Further, assume that U'_y has the same sign for any $x, y \geq 0$. By concavity of the subjective utility functions $U''_{xx} + \alpha U''_{yy} < 0$. Then $sign(x^{*\prime}_\alpha) = sign(U'_y)$ and $sign(y^{*\prime}_\alpha) = sign(U''_{xy}U'_y)$.

Suppose that $U'_y > 0$ and $U''_{xy} > 0$, as in the quadratic example with $k > 0$. Then $x^{*\prime}_\alpha > 0$ and $y^{*\prime}_\alpha > 0$. Further, from the maximization of the subjective utility function V it follows that $U'_x = -\alpha U'_y$. Then equation (4.8) becomes $-\alpha x^{*\prime}_\alpha + y^{*\prime}_\alpha|_{\alpha=\beta} = 0$. Since both $x^{*\prime}_\alpha$ and $y^{*\prime}_\alpha$ are positive, it can hold only for $\alpha > 0$. Then $\alpha < 0$ cannot be evolutionarily stable if $U'_y > 0$ and $U''_{xy} > 0$. The same result appears when $(U''_{xx} + \alpha U''_{yy})^2 - (U''_{xy}(1+\alpha))^2 < 0$. Analogously, $\alpha < 0$ cannot be stable with $U'_y > 0$ and $U''_{xy} > 0$. The same reasoning gives that $\alpha > 0$ cannot be evolutionary stable when $U''_{xy} < 0$.

Recall that a twice differentiable function of two variables $U(x, y)$ is *super-modular* if $U''_{xy} > 0 \ \forall x, y$ and it is *sub-modular* if $U''_{xy} < 0 \ \forall x, y$. Given this definition, we can formulate the following theorem

Theorem 4.4 *Suppose it holds for the material payoff function* $U(x, y)$ *that* $sign((U''_{xx} + \alpha U''_{yy})^2 - (U''_{xy}(1+\alpha))^2) = const \neq 0$, $sign(U'_y) = const \neq 0$ *for any* $\alpha \in W$ *and any* $x, y \geq 0$. *Then, if the material payoff function is super-modular only* $\alpha > 0$ *can be evolutionarily stable and if the function is sub-modular only* $\alpha < 0$ *can be evolutionarily stable.*

The theorem extends the results to more general material payoff functions than the quadratic one. Notice that the quadratic model of section 4.2 with the set $W = [-1, 1]$ satisfy the assumptions of the theorem but the models of this section do not satisfy them since $U'_y = kx$ does not have a constant sign. Another possible generalization is the subjective utility function which in the example above takes the form $V(U_1, U_2) = U_1 + \alpha U_2$. Although it is difficult to establish analytical results in the infinite-dimensional space of functions, the intuitive reasoning in the end of the section 4.2 may go through if $V'_{U_1} > 0$ and V'_{U_2} has a constant sign and the theorem can be valid for this case too.

4.4 Incomplete Information

In the previous sections it was assumed that the players knew each other's preferences. This assumption is rather strong. In this section we relax the complete information assumption. Instead we assume that players know the distribution of preferences in the population. This assumption is used also in other works considering indirect evolution with incomplete information (Güth and Peleg (1997), Ely and Yilankaya (1997), Ok and Vega-Redondo (1999)). Thus, each encounter is a Bayesian game where the set of types is the set of feasible preferences (preference parameters) and the payoff to each type is given by its corresponding subjective utility function V.

The definition of evolutionary stability can be extended to the incomplete information setting in the following way. Suppose the population is monomorphic with preference parameter α. To analyze the evolutionary stability of this preference parameter we consider an invasion by a small number of mutants with some other preference parameter α'. The proportion ε of mutants is common knowledge and arbitrarily small. The population is large (infinite), so every member of the population faces the same distribution of types, independent of his own type. Thus, we have a Bayesian game with two types $T = \{\alpha, \alpha'\}$, with prior beliefs on the set of types $\{1-\varepsilon, \varepsilon\}$, and with the subjective utility functions for the two types

$$V_\alpha(x,y) = U_1(x,y) + \alpha U_2(x,y), \; V_{\alpha'}(x,y) = U_1(x,y) + \alpha' U_2(x,y) \tag{4.11}$$

where $U_i(x,y)$ are given in (4.1).

We are looking for an equilibrium of the Bayesian game. Given equilibrium strategies we can calculate the material payoffs of the two types. Denote the expected material payoff of type α in the equilibrium by $U^*_\alpha(\varepsilon)$ and that of mutant type α' by $U^*_{\alpha'}(\varepsilon)$.

Definition 4.3 *A preference parameter α^* is* ***evolutionarily stable with incomplete information*** *if $\exists \varepsilon^* > 0$ such that $U^*_{\alpha^*}(\varepsilon) > U^*_\alpha(\varepsilon)$ for any mutant type α for any $\varepsilon \in (0, \varepsilon^*)$.*

Note that our definition is slightly different from the specification of mutations used elsewhere in the literature. Ely and Yilankaya (1997), for example, consider non-atomic distributions of preferences and, therefore, mutations with full support on the space of preferences. Ok and Vega-Redondo (1999) have a finite population and, therefore, players of different types have a different view on the population composition. We use the above definition since it is the closest to the one for the complete information. It captures the idea that the (infinite) population is stable against invasion of an arbitrarily small proportion of mutants with particular preferences, for any mutant preferences. The only thing that is changed with respect to the complete information case is that in each encounter a player receives an uncertain payoff (depending on the type of the opponent). In the complete information case the payoff in an encounter was certain but the encounter itself was not.

We restrict ourselves to symmetric equilibria of the game. Let us denote the strategy of player 1 by $(x_\alpha, x_{\alpha'})$ where x_α is the strategy when the player is of type α and $x_{\alpha'}$ is the strategy when the player is of type α'. Let $(y_\alpha, y_{\alpha'})$ be the strategies for player 2. To find an equilibrium one has to solve

$$\max_{x_\alpha, x_{\alpha'}} (1-\varepsilon)^2 V_\alpha(x_\alpha, y_\alpha) + (1-\varepsilon)\varepsilon V_\alpha(x_\alpha, y_{\alpha'}) + \\ + \varepsilon(1-\varepsilon) V_{\alpha'}(x_{\alpha'}, y_\alpha) + \varepsilon^2 V_{\alpha'}(x_{\alpha'}, y_{\alpha'}) \tag{4.12}$$

and the corresponding problem for player 2 which we omit because of symmetry. Since we are looking for symmetric equilibria, $x_\alpha = y_\alpha$ and $x_{\alpha'} = y_{\alpha'}$ in an equilibrium.

The solution of the problems leads to

$$\begin{aligned} x_\alpha &= \frac{m(\alpha\varepsilon k - \alpha'\varepsilon k + 2)}{2(\alpha k(\varepsilon - 1) - \alpha'\varepsilon k - k + 2)}, \\ x_{\alpha'} &= \frac{m(\alpha k(\varepsilon - 1) + \alpha' k(1-\varepsilon) + 2)}{2(\alpha k(\varepsilon - 1) + \alpha'\varepsilon k - k + 2)}. \end{aligned} \tag{4.13}$$

These are the strategies of the two types in the symmetric equilibrium.

The limit of the expected material payoff of type α in the equilibrium

$$\lim_{\varepsilon \to 0} (1-\varepsilon) U_1^*(x_\alpha, x_{\alpha'}) + \varepsilon U_1^*(x_\alpha, x_\alpha) = \frac{m^2(1-\alpha k)}{(\alpha k + k - 2)^2} \tag{4.14}$$

while the limit of the expected material payoff for type α' is

$$\lim_{\varepsilon \to 0} (1-\varepsilon) U_1^*(x_{\alpha'}, x_\alpha) + \varepsilon U_1^*(x_{\alpha'}, x_{\alpha'}) = \frac{m^2(\alpha k + \alpha' k - 2)(\alpha k - \alpha' k - 2)}{4(\alpha k + k - 2)^2}. \tag{4.15}$$

The difference between the two is

$$\begin{aligned} \lim_{\varepsilon \to 0} U_\alpha^* - \lim_{\varepsilon \to 0} U_{\alpha'}^* &= \frac{k^2 m^2 (\alpha'^2 - \alpha^2)}{4(\alpha k + k - 2)^2} > 0 \Longleftrightarrow \\ \alpha'^2 - \alpha^2 &> 0 \Longleftrightarrow \mid \alpha' \mid > \mid \alpha \mid . \end{aligned} \tag{4.16}$$

Therefore, we have

Theorem 4.5 *The only evolutionarily stable preference parameter with incomplete information is* $\alpha^* = 0$.

For any other $\alpha \neq 0$ a mutant with preference parameter closer to 0 achieves a higher payoff. The intuition of the result is straightforward. Suppose the population consisted of maximizers of the material payoff when mutants invade. Since the proportion of mutants is arbitrarily small, in equilibrium the incumbents play almost the same strategy as they played before the invasion, that is, maximizing the material payoff in a game against each other. The mutants, since they face almost exclusively the incumbents, play a best response against the incumbent according to the mutants' subjective utility. This best response does not maximize the material payoff, therefore mutants have lower fitness. They have strictly lower fitness since the game is such that different types play different strategies and the material payoff function has a unique maximum, therefore the mutants cannot have the same fitness as the incumbents. Thus, with incomplete information only preferences which coincide with the material payoff maximization survive evolutionary pressure, supporting the claim of Alchian (1950) and Friedman (1953). A similar result that with incomplete information the material (or equivalent to them) preferences are evolutionary stable under certain conditions was also obtained by Ely and Yilankaya (1997) and Ok and Vega-Redondo (1999).

4.5 Conclusion

The indirect evolution approach helps to address the question which preferences will survive evolutionary pressure. However, one should not artificially restrict the set of feasible preferences. In this chapter we extend the model of Bester and Güth (1998) to a larger set of preference parameters. We have shown that when spiteful preferences are allowed they can be evolutionarily stable for a large set of values of the parameter of the material payoff function. Admittedly, this model still analyzes specific forms of the material payoff function, though some more general results are available. Nevertheless, already in this specific framework there is a variety of results showing that some basic properties of the material payoff function are not enough to draw general conclusions. The properties of the game played with the evolutionarily stable preferences may be different from the properties of the material payoff game.

Perhaps more important, the set of preferences is restricted to a one-dimensional family of utility functions. Certainly, it is possible to enlarge the set, for example, with such utilities that depend on the opponent's utility (see, e.g. Bolle (1991)), or on player's beliefs (Rabin (1993)), or preferences that take equity into account. But then the analysis for duopoly games considered here becomes rather intractable. The next chapter considers indirect evolution for quite general preferences in 2×2 games.

The persistence of spiteful preferences, found in this chapter, may be explained by the fact that a "spiteful" player gets a higher payoff than a "normal" player for a rather general class of games. An analysis of the advantage of negatively interdependent preferences can be found in Koçkesen et al. (1999) though their result should be taken with caution as it does not necessarily mean that "spiteful" players will wipe out "normal" ones. In the random matching setting there is a counter-balancing effect that "spiteful" players cooperate less with each other than "normal" players. The result of the evolutionary process depends much on the exact form of the material payoff function.

Preferences different from maximizing own material payoff are evolutionarily stable when players know each other's preferences in a match. In a sense, having certain preferences is a commitment device to use a strategy that does not maximize the material payoff but can bring a higher material payoff than that of the opponent. In an incomplete information setting only maximizing preferences survive since this commitment does not work due to unobservability. This result is of a general character and supports the claim of Alchian (1950) and Friedman (1953) that rationality will be selected by evolution. Thus, the informational issues play a large role in the determination of the outcomes of the evolutionary process.

Chapter 5

Indirect Evolution in 2×2 Symmetric Games

5.1 Introduction

The previous chapter analyzed indirect evolution of preferences for games arising from duopoly models. We argued that limitations on the set of feasible preferences may be important for the results. However, the set of possible preferences we considered in that chapter was still rather limited, and it was difficult to conduct the analysis for all preferences. In this chapter we analyze indirect evolution of preferences in simple (symmetric 2×2) games for *all* possible (von Neumann-Morgenstern) preferences.

Indirect evolution, as described in the previous chapter, works on preferences instead of on the strategies of the players. Thus, each player is supplied with genetically programmed preferences represented by a von Neumann-Morgenstern utility function. Given these preferences, players play the game. They are either rational and thus play an equilibrium from the start, or they arrive at it by a process of learning which is much faster than the process of evolution. Thus, their evolutionary success is determined by the fitness they receive from playing the equilibrium strategies. Different preferences imply different equilibria and, generally, different fitness, and the evolution selects the preferences with a higher fitness. We are interested in stationary stable points of the evolutionary process. Preferences that are observed in such points are called evolutionarily or neutrally stable preferences.

Some work using the indirect evolution approach has been done before for certain games. For example, evolution of trust (Brennan et al. (1997)), evolution of fairness

(Huck and Oechssler (1999)), evolution of preferences for sales in duopoly (Dufwenberg and Güth (1999)) have been analyzed. However, the set of feasible preferences in these papers is usually assumed to be a one-dimensional subset of all possible preferences. These restrictions may lead to results which are not robust to an enlargement of the set of feasible preferences (see Bester and Güth (1998) and the previous chapter). Therefore, it is important to consider all possible preferences.

The set of all possible preferences is huge for large games. We focus on 2×2 symmetric games since they are the simplest games and they are of importance for real life situations. The class of 2×2 symmetric games includes such extensively analyzed games as coordination problems, the prisoners' dilemma, and the "chicken" (Hawk-Dove) game. The set of all possible preferences in 2×2 symmetric games is not that large; it can be parametrized by two parameters. Thus we can consider a two-dimensional space of preferences. We are interested in the question of which preferences are stable for which class of games. We also address the related questions of whether selfish preferences, that is, the ones that coincide with fitness, survive evolution, and how the outcome of the game with evolutionarily stable preferences relates to a Nash equilibrium of the material payoffs game.

We analyze first a model of complete information and perfect coordination where players can observe each other's preferences and they can coordinate perfectly on one equilibrium, if there are multiple ones. Then selfish preferences are not necessarily stable, for example, they are not stable in the prisoners' dilemma or in the "chicken" game. Relaxing the assumption of perfect coordination does not change the result substantially though it gives some interesting insights for coordination games.

More important is the assumption of complete information. This assumption is rather strong and is relaxed to two models with incomplete information. In one model the identity of the mutant is not known but the fact that a mutant with certain preferences has appeared is known. This game then can be modeled as a Bayesian game. In the other model the existence of a mutant is not known, so the incumbent population continues to play the same strategy as before. We show that selfish preferences survive more often under incomplete than under complete information. Moreover, they (or preferences equivalent to them) are the only preferences that can be stable under incomplete information. These results support the results of the preceding chapter. The only model where selfish preferences do not survive is the Bayesian model for the "chicken" game. The reason is that players with selfish preferences play the symmetric mixed equilibrium, so any strategy is an alternative best reply to the mixed equilibrium strategy.

We formulate the model of indirect evolution in 2×2 symmetric games in Section 5.2. Section 5.3 analyzes the model. Imperfect coordination and incomplete information extensions of the model are considered in Sections 5.4 and Section 5.5. Section 5.6 discusses and concludes.

5.2 The Model

5.2.1 Games

We consider 2×2 symmetric games. The generic normal form of such games is given below, where s_1, s_2 are the strategies of the players and $\alpha, \beta, \gamma, \delta$ represent the *material payoffs* or *fitness.*

	s_1	s_2
s_1	α, α	β, γ
s_2	γ, β	δ, δ

In order to exclude uninteresting games we require that the strategies bring different payoffs at least against some strategy of the other player, that is, either $\alpha \neq \gamma$ or $\beta \neq \delta$. We assume that adding a constant to all payoffs and multiplying all payoffs by a positive constant do not change the result of evolution. This assumption is not strong since commonly used dynamic specifications of evolution, such as the replicator dynamic, and static evolutionary concepts, such as evolutionary stable strategy, are invariant with respect to these changes. Without loss of generality we can take $\alpha \geq \delta$ (otherwise rename strategies). By substracting δ from all payoffs, and multiplying them by $\frac{1}{\alpha-\delta}$ (if $\alpha > \delta$) we can consider games of the form

	s_1	s_2
s_1	$1, 1$	b, c
s_2	c, b	$0, 0$

where $b = \frac{\beta-\delta}{\alpha-\delta}, c = \frac{\gamma-\delta}{\alpha-\delta}$. If $\alpha = \delta$, we have a game of the form

	s_1	s_2
s_1	$0, 0$	b, c
s_2	c, b	$0, 0$

where $b = \beta - \delta, c = \gamma - \delta$. In what follows we use these representations of the material payoff game. Thus the parameter space is reduced to two dimensions. We focus on games with $\alpha > \delta$ since they are more interesting. When the results for games with $\alpha = \delta$ differ from the ones for games with $\alpha > \delta$ we comment on them.

For $\alpha > \delta$, depending on the values of the parameters b, c we distinguish four different types of 2×2 symmetric games:

(Eff) $1 \geq c, b \geq 0$ - "efficient dominant strategy";

(CP) $1 > c, b < 0$ - "coordination problem";

(Ch) $1 < c, b > 0$ - "chicken type game";

(PD) $1 \leq c, b \leq 0$ - "Prisoners' Dilemma type game".

As one can see, the set of symmetric 2×2 games we analyze contains interesting games which have been analyzed extensively in the economic and game-theoretic literature. To our knowledge, however, they have not been analyzed yet from the point of view of indirect evolution of preferences. We seek to close this gap. For each type of games we analyze which preferences (described in the next subsection) are stable.

5.2.2 Preferences

Let a 2×2 symmetric game be given, that is, parameters b, c are fixed. The set of strategies for each player is $S_i = \{s_1, s_2\}$. Denote the mixed extension of S_i by Σ_i. The set of mixed strategy combinations is then $\Sigma = \Sigma_1 \times \Sigma_2$. Each agent i in the population that is playing the game in randomly matched pairs is endowed with certain preferences over the set Σ of mixed strategy combinations. These *subjective preferences* are represented by a von Neumann-Morgenstern utility function $v_i(\cdot)$. The function is completely defined by its values on all four pure strategy combinations $(s_j, s_k), j, k \in \{1, 2\}$, where s_j is own strategy and s_k is the opponent's strategy. We do not make any restriction on the utility function other than being von Neumann-Morgenstern. In what follows we identify preferences with their utility function v_i.

The preferences are defined on the set of strategy combinations rather than on the set of payoff combinations since it allows to represent a larger set of preferences, and our goal is to analyze as large a set of feasible preferences as possible. If preferences are defined on the set of outcomes, the players have to be indifferent between two strategy combinations that lead to the same payoff combination. By considering preferences on the set of strategy combinations, we avoid this restriction. Our players know what the strategies are and may have biases towards one or the other strategy.

Analogously with types of games, one can define types of preferences. We are interested in finding an equilibrium of a game with given preferences. For this end we only need to know the best response correspondences for each player. Since only ordinal relationships matter for best response, for our purposes we can divide cardinal preferences according to ordinal relationships into the following subsets that we call *types*:

Type 1 (St1): $v_i((s_1, s_1)) \geq v_i((s_1, s_2)), v_i((s_2, s_1)) \geq v_i((s_2, s_2))$, at least one inequality is strict;

Type 2 (CO): $v_i((s_1, s_1)) > v_i((s_1, s_2)), v_i((s_2, s_1)) < v_i((s_2, s_2))$;

Type 3 (NC): $v_i((s_1, s_1)) < v_i((s_1, s_2)), v_i((s_2, s_1)) > v_i((s_2, s_2))$;

Type 4 (St2): $v_i((s_1, s_1)) \leq v_i((s_1, s_2)), v_i((s_2, s_1)) \leq v_i((s_2, s_2))$, at least one inequality is strict;

Type 5 (BB): $v_i((s_1, s_1)) = v_i((s_1, s_2)), v_i((s_2, s_1)) = v_i((s_2, s_2))$.

We say that preferences v_i belong to the type k if v_i satisfies the inequalities for type k. Correspondingly to the types of games above, players with type (St1) preferences perceive[1] the game as having dominant strategy s_1, while players of type (St2) think that s_2 is dominant. Type (CO) players (COoperators or COnformists) perceive that s_1 is best reply to s_1 and s_2 on s_2, while type (NC) (NonConformists) players prefer to play s_1 on s_2 and s_2 on s_1. Finally, there are preferences of type (BB) ("Big Bores") for which the strategies are equivalent. Though such preferences are rather strange, they have to be included if one is to consider all preferences. The players with such preferences are indifferent to any (mixed) strategy and therefore can be induced to play any strategy by, for example, a (private) benevolent planner[2] who maximizes their fitness (or by chance, for that matter).

An interpretation of having different preferences can be seen on the example of the prisoners' dilemma. Some agents may have selfish preferences while others might not like to let their opponents down and therefore have a higher subjective utility from mutual cooperation than from cheating against a cooperator. Yet others can be heroic unconditional cooperators who derive a higher utility even from being defected upon, that is, they prefer to sacrifice themselves in favor of the other player.

The subjective utility functions can represent many preferences. It is clear from the definition of types that such preferences as biases towards a particular strategy, the desire to conform, and the desire to differ can be represented. Preferences that depend more

[1]We often use such terms as "perceive", "think" to represent the preferences of the player. This seems to us a valid interpretation of preferences, since in game theory a player's preferences are known to the player, and usually even to all players.

[2]Maria Montero suggested to call the (private) benevolent planner "mother of the player".

on the opponent's strategy rather than on own strategy (for example, preferences for dealing with a friendly person) may be represented. Altruistic and spiteful preferences of the preceding chapter can be represented as well since one can compute the sums (or the differences) of the material payoffs for each strategy combination and compare. Furthermore, preferences represented by any well behaved function of material payoffs are feasible too. Thus, the set of feasible preferences contains many "common sense" preferences.

Though the agents know their preferences, they do not need to know what the material payoffs are. Evolution, described in the next subsection, will choose those preferences that have higher fitness.

5.2.3 Evolution and Stability

Evolution works indirectly choosing preferences through equilibrium payoffs. What is important for evolution is not the subjective utilities but the material payoffs. Roughly, the evolutionary process works as follows. The agents of the population are randomly matched to play the symmetric 2×2 game. We assume that the preferences in a match are common knowledge, an assumption which will be relaxed later. The agents are rational given their subjective utilities and therefore they play an equilibrium of the game as they perceive it. We assume that either there is a unique equilibrium of the game or there is a way to select some solution of the game, which will be specified in the following sections. The fitness of an agent is determined by the material payoffs game with the equilibrium strategies. The agents then reproduce having more offspring if they had higher fitness[3]. Offspring inherit the preferences of their parent. Thus, if an agent with preferences v_i had on average over all matches a higher fitness than an agent with preferences v_j, the proportion of agents with preferences v_i grows faster than those with preferences v_j. We consider a large population so that the Law of Large Numbers allows us to use the expected payoffs as the realized ones. Together with reproduction and selection, once in a while a mutation occurs possibly bringing new preferences into the population.

We do not model the evolutionary process more explicitly but instead turn to the analysis of which stationary states of it are stable against a mutation. We concentrate on *monomorphic* stationary states, that is, ones where all players have the same preferences.

The formal description of the indirect evolution approach is in the previous chapter.

[3]This interpretation is only suggestive. One can also think of proliferation of values through imitation, for example.

Here we repeat quickly the essential points. Denote by $u_i(v_i, v_j)$ the material payoff of an individual with preferences v_i in equilibrium (or otherwise chosen unique solution) of the game between agents with preferences v_i, v_j. The evolutionary game Γ is a symmetric two-player game with the sets of all possible preferences as the strategy sets and $u_i(v_i, v_j)$ as the payoff function (see the preceding chapter for formalities). The standard definition of evolutionarily stable strategy (e.g. Weibull (1995, Ch.2)), reformulated for preferences, is

Definition 5.1 *Preferences v_i are **evolutionarily stable** if $u_i(v_i, v_i) \geq u_j(v_j, v_i)$ $\forall v_j$ and if $u_i(v_i, v_i) = u_j(v_j, v_i)$ for $v_j \neq v_i$ implies $u_i(v_i, v_j) > u_j(v_j, v_j)$.*

Given that the space of preferences is rather large, for any preferences v_i there are always other preferences v_j which do no worse than v_i do both against v_i and against themselves. It is often so that when v_i, v_j belong to the same type of preferences they play the same strategies in equilibrium, and so are indistinguishable in terms of the material payoffs. For example, if under both v_i and v_j strategy s_1 is dominant, players with both v_i and v_j always play s_1 though it may be that v_i has a larger bias towards s_1. Therefore, we often use a weaker concept of neutral stability.

Definition 5.2 *Preferences v_i are **neutrally stable** if $u_i(v_i, v_i) \geq u_j(v_j, v_i)$ $\forall v_j$ and if $u_i(v_i, v_i) = u_j(v_j, v_i)$ implies $u_i(v_i, v_j) \geq u_j(v_j, v_j)$.*

Under neutrally stable preferences a mutant can survive but it will not grow. Neutral stability states that once perturbed, the system will stay in the neighborhood of the original state while evolutionary stability implies that the system converges back to the original state.

Given that we divided the space of preferences into types, similar definitions can be given for types of preferences.

Definition 5.3 *A type k of preferences is **evolutionarily stable** if $\forall v_i \in k$, $u_i(v_i, v_i) \geq u_j(v_j, v_i)$ $\forall v_j \notin k$ and if $u_i(v_i, v_i) = u_j(v_j, v_i)$ for $v_j \notin k$ implies $u_i(v_i, v_j) > u_j(v_j, v_j)$.*

Definition 5.4 *A type k of preferences is **neutrally stable** if $\forall v_i \in k$, $u_i(v_i, v_i) \geq u_j(v_j, v_i)$ $\forall v_j \notin k$ and if $u_i(v_i, v_i) = u_j(v_j, v_i)$ implies $u_i(v_i, v_j) \geq u_j(v_j, v_j)$.*

A type of preferences is stable if any preferences of this type cannot be invaded by mutant preferences of other types. Thus, if the monomorphic population consists of agents with preferences from a stable type, evolution cannot lead to the growth of preferences of other types.

Our goal is to find which preferences are neutrally stable and which types of preferences are evolutionarily or neutrally stable depending on the underlying game. Related questions are whether selfish preferences are stable and what outcome of the material payoffs game is observed under stable preferences.

5.3 The Analysis

5.3.1 Equilibria and Material Payoffs

As we mentioned in the previous section, the fitness of the players is defined as the material payoff with equilibrium strategies if the equilibrium is unique and by some other plausibly chosen outcome if there are multiple equilibria. In this section we will be more precise on this.

Since we divided the set of all possible preferences into types according to the best response correspondence, we can analyze games between players with preferences of given types without specifying the preferences completely. Thus, for example, type (St1) preference would mean any preferences with the property that s_1 is the (weakly) dominant strategy for them.

We focus on undominated equilibria.

Assumption 1 *Only undominated strategies are played in equilibrium.*

If a player perceives one of the strategies to be dominant (types (St1) and (St2)), her opponent can be sure that such a strategy will be played and therefore can choose a best reply to the dominant strategy. Unless the opponent in question is of type (BB), either he plays his own dominant strategy or has a unique best reply to the dominant strategy of the player. Thus, in games with one player of type (St1) or (St2) the equilibrium is unique regardless of what type of preferences different from (BB) the opponent has. The equilibria might be different, however: the game of type (St1) against type (St1) has equilibrium (s_1, s_1) while the game of type (St1) against type (NC) has equilibrium (s_1, s_2). If the opponent is of type (BB), we are free to choose any (mixed) strategy for him, as any strategy constitutes an equilibrium. We assume that any equilibrium is achievable, and that the (private) benevolent planner ("mother of the player") chooses one with the highest material payoffs for him. The idea is to give such preferences a fair chance for survival. A justification could be that there is evolution on the subpopulation of players with type (BB) preferences leading to such of them that have the highest fitness. Alternatively, it may be that just by chance the players of type (BB) hit the "right" strategy. The assumption will not influence the stability of the other preferences too often though sometimes it will, then we will comment on that. In a later section we will relax this assumption by introducing small mistakes since the players have little incentive to follow the planner's recommendations. For the time being, we make

Assumption 2 *Agents with type (BB) preferences play an equilibrium that brings them the highest material payoffs among all equilibria. Between themselves they play a symmetric equilibrium with this property.*

If a player has preferences of type (NC), that is, s_2 is best reply to s_1 and vice versa, the situation is as follows. If such a type is matched against itself there are three equilibria: two pure strategy off-diagonal ones and one mixed equilibrium. If the preferences are exactly the same, by the considerations of symmetry the mixed equilibrium is more appealing. Moreover, in one population for many models of learning, e.g. derived from the replicator dynamic, the mixed equilibrium is stable while the pure equilibria are unstable (Weibull (1995, Ch.3)). The probabilities that the mixed equilibrium puts on pure strategies depend on the exact cardinal preferences. If the preferences differ but are still both of type (NC), the symmetry argument does not apply, but, to preserve some sort of continuity of the solutions of the game with respect to the small changes in preferences, we assume that the material payoffs realized in such a game will be the ones corresponding to the mixed strategy equilibrium. Furthermore, even when an (NC) type player plays the mixed equilibrium strategy, the strategy can be made arbitrarily close to either of the pure one by choosing extreme preferences ($v_i((s_1, s_2)) \to \pm\infty$) for the opponent. Thus, the assumption is not a very strong restriction.

Assumption 3 *In a game of two agents with type (NC) preferences, the mixed equilibrium is played.*

If type (NC) agent is matched with type (CO) agent, there are no pure strategy equilibria and, therefore, the mixed strategy equilibrium is played. Again, since the game of type (BB) against type (NC) has a continuum of equilibria we allow type (BB) to choose one with the highest material payoff for him.

The problem of multiplicity of equilibria appears also with type (CO) preferences. If such an agent is matched against another agent of the same type, the perceived game is a coordination game, that is, a game with three equilibria, two pure ones on the diagonal and one mixed. With symmetric preferences, the mixed equilibrium is unstable and both pure equilibria are stable. We assume that the players of type (CO) indeed can coordinate, and so one of the pure equilibria is played, and, moreover, it is the one with a higher material payoff. A justification for this choice of the efficient equilibrium rather than of the risk-dominant equilibrium is that we want to give preferences of type (CO) a fair chance to survive, as before for preferences of type (BB). There is also a model of strategy adjustment of Robson and Vega-Redondo (1996) that leads to the efficient equilibrium in coordination games. We extend the assumption of the efficient equilibrium choice to asymmetric type (CO) preferences as well. Later in the chapter, similar to the mistakes in type (BB) planned action, we will introduce mistakes in coordination of type (CO) players. For type (BB) against type (CO) we assume the equilibrium with the highest material payoff for type (BB).

Assumption 4 *In a game of two agents with type (CO) preferences, the pure equilibrium with higher material payoffs is played.*

Summarizing the above, we can represent the evolutionary game on types by the following bimatrix (if $\alpha \neq \delta$).

	(St1)	(CO)	(NC)	(St2)	(BB)
(St1)	$1, 1$	$1, 1$	b, c	b, c	u_1^5, u_5^1
(CO)	$1, 1$	$1, 1$	u_2^m, u_3^m	$0, 0$	u_2^5, u_5^2
(NC)	c, b	u_3^m, u_2^m	u^m, u^m	b, c	u_3^5, u_5^3
(St2)	c, b	$0, 0$	c, b	$0, 0$	u_4^5, u_5^4
(BB)	u_5^1, u_1^5	u_5^2, u_2^5	u_5^3, u_3^5	u_5^4, u_4^5	u_5^5, u_5^5

The entries in the bimatrix are the material payoffs in the chosen equilibrium. By $u_\cdot^m$ we denote the fitness of the players in the mixed equilibrium. The exact value of $u_\cdot^m$ depends on the exact preferences. The entries u_i^5 represent the material payoff of type i in the equilibrium chosen by type 5 (BB) in the game against type (BB). Similarly, u_5^i represent the material payoff of type 5 (BB) in the game against type i. If $\alpha = \delta$, entries $(1, 1)$ should be substituted by $(0, 0)$ while other entries stay the same. The next subsection analyzes evolutionary stability of preferences.

5.3.2 Evolutionary Stability of Preferences

Given the evolutionary game from the previous section we will consider the types in turn for the case $\alpha > \delta$. Thus, we have that strategy s_1 is more efficient in the sense that strategy profile (s_1, s_1) is more efficient than strategy profile (s_2, s_2). This will have consequences for the stability of (St1) and (St2) preferences.

The following lemma is useful. By a symmetric strategy we mean that both players play the strategy.

Lemma 5.1 *If $b + c \leq 2$ then 1 is the highest material payoff that can be achieved by symmetric strategies. If $b + c > 2$ then there are symmetric mixed strategies that give higher material payoff.*

Proof. Let the symmetric mixed strategy put probability p on s_1. Then the material payoff players get is $p^2 + p(1-p)b + (1-p)pc = p^2(1-b-c) + p(b+c)$. If $b + c \leq 1$, the maximum with respect to p is at the borders of the interval $[0, 1]$, in our case at $p = 1$. If $b + c > 1$, the maximum is at $p^* = \frac{b+c}{2(b+c-1)} > 0$ if $p^* < 1$, and at $p = 1$ if $p^* \geq 1$. We have $1 - p^* = \frac{b+c-2}{2(b+c-1)} > 0 \iff b + c > 2$. Since the material payoff of 1 is achieved by $p = 1$, the maximum is higher than 1 in the case $b + c > 2$. ■

Now we turn to the types.

Lemma 5.2 *Type (St1) is neutrally stable and all preferences of type (St1) are neutrally stable if ($c < 1$ or ($c = 1$ and $b = 1$)) and $b + c \leq 2$.*

Proof. If $c > 1$, preferences of types (NC) or (St2) are better replies to type (St1) preferences in the evolutionary game thus type (St1) is not stable in this case. If $c < 1$, the best replies to a given type (St1) preferences are themselves, other type (St1) preferences, type (CO) preferences, and type (BB) preferences, all leading to the material payoffs $(1, 1)$. If $b + c \leq 2$, all these preferences have 1 both against themselves and against type (St1). By Lemma 5.1, type (BB) preferences can achieve a higher payoff against themselves by using a symmetric mixed strategy if $b + c > 2$, thus type (St1) is not stable in this case. For $c = 1$ all preferences are best responses to type (St1). If $b < 1$ then type (NC) preferences can have a higher payoff against themselves then what type (St1) get against them. If $b = 1$ no preferences can get more than 1, the sure payoff of type (St1). ■

Thus, the preferences that consider s_1 as the dominant strategy are neutrally stable when s_1 is really (i.e. in the material payoff game) best reply against itself and when there is no possibility of finding a symmetric mixed strategy with higher material payoffs. If we omit type (BB) preferences, the first condition (s_1 is a best reply against itself) together with $b \geq 1$ is sufficient for neutral stability of (St1) preferences.

To analyze type 2 (CO) preferences, we normalize them (if $v_2((s_1, s_1)) > v_2((s_2, s_2))$)[4] by such transformations as adding a constant and multiplying by a positive constant to $v_2((s_1, s_1)) = 1, v_2((s_2, s_2)) = 0, v_2((s_1, s_2)) = b_2 < 0, v_2((s_2, s_1)) = c_2 < 1$. The normalization is without loss of generality.

Lemma 5.3 *Type (CO) preferences that after normalization satisfy $b < \frac{c_2+b_2-1}{b_2}$ are neutrally stable if $b + c - 2 \leq 0$ and $c \leq 1$. If $b = \frac{c_2+b_2-1}{b_2}$ then such type (CO) preferences are neutrally stable if $b + c - 2 \leq 0$ and $c < 1$. Type (CO) is neutrally stable if $b + c - 2 \leq 0$, $c \leq 1$, and $b \leq 1$.*

Proof. Analogously with the type (CO) preferences normalization, we can consider any normalized type 3 (NC) preferences with $v_3((s_1, s_1)) > v_3((s_2, s_2))$ and $v_3((s_1, s_1)) = 1, v_3((s_2, s_2)) = 0, v_3((s_1, s_2)) = b_3 > 0, v_3((s_2, s_1)) = c_3 > 1$ (the proof is similar with other normalizations). When type (NC) and type (CO) players are matched against each other, they play a mixed equilibrium where type (CO) plays s_1 with probability $p_2 = \frac{b_3}{b_3+c_3-1}$, while type (NC) plays s_1 with probability $p_3 = \frac{b_2}{b_2+c_2-1}$. We should compare the expected payoff u_3 of type (NC) in equilibrium with what type (CO) gets against itself which is 1. We have $u_3 - 1 = \frac{(c_3-1)(b_2 b+1-c_2-b_2)+b_3((c_2-1)c+1-c_2)}{(c_2+b_2-1)(c_3+b_3-1)}$. The denominator is negative thus the fraction is positive if $(c_3 - 1)(b_2 b + 1 - c_2 - b_2) + b_3((c_2 - 1)c + 1 - c_2) < 0$. Since $c_3 - 1$ and b_3 can range from 0 to $+\infty$, it is possible if either $b_2 b + 1 - c_2 - b_2 < 0$ or $(c_2 - 1)c + 1 - c_2 < 0$ or both. That is, type (NC) is a better reply to type (CO) than type (CO) to itself if $b > \frac{c_2+b_2-1}{b_2}$ or $c > 1$. If $b \leq \frac{c_2+b_2-1}{b_2}$ and $c \leq 1$ and at least one of the inequalities is strict then types (St1) and (BB) are best replies to (CO), as well as type (CO) itself, all bringing the material payoff of $(1, 1)$. By the same argument as in the previous lemma, if $b + c > 2$, type (BB) can achieve a higher material payoff against itself by using a mixed strategy, and, therefore, type (CO) is not stable then. If $b + c \leq 2$, neither of the other types can achieve a material payoff higher than 1 against themselves, so type (CO) is neutrally stable. If $c = 1$ and $b = \frac{c_2+b_2-1}{b_2} = 1 + \frac{c_2-1}{b_2} > 1 \Longrightarrow b + c > 2$ and again type (BB) can do better against itself then type (CO) does. The neutral stability of type (CO) is verified only when $b \leq 1$ since only in this case the condition $b \leq \frac{c_2+b_2-1}{b_2}$ holds for any $b_2 < 0, c_2 < 1$. ■

[4] If $v_2((s_1, s_1)) \leq v_2((s_2, s_2))$ we can use another similar normalization that does not change the result since this relationship does not matter for the best response correspondence.

The conditions of the lemma can be interpreted as follows. The condition $c \leq 1$ says that s_1 is a best reply against itself, that is correctly perceived by type (CO). The condition $b \leq \frac{c_2+b_2-1}{b_2}$ indicates that not necessarily all preferences of type (CO) are stable but for any b there is a subset of non-zero measure of preferences of type (CO) that are neutrally stable. If $b \leq 1$ this subset coincides with the whole set of type (CO) preferences. Similarly to type (St1), without type (BB) preferences, type (CO) preferences would be stable even when $b + c > 2$.

Analogously with type (CO), for the analysis of type (NC) preferences consider their normalization $(v_2((s_1, s_1)) > v_2((s_2, s_2)))$ to $v_2((s_1, s_1)) = 1, v_2((s_2, s_2)) = 0, v_2((s_1, s_2)) = b_3 > 0, v_2((s_2, s_1)) = c_3 > 1$. Such preferences are never stable.

Lemma 5.4 *No type (NC) preferences are neutrally stable and type (NC) is not neutrally stable.*

Proof. Players of type (NC) with the same preferences play the symmetric mixed equilibrium with probability $p = \frac{b_3}{b_3+c_3-1}$ of playing s_1. Denote the expected material payoffs in such equilibrium by u_3.

Consider type (CO) preferences with $v_2((s_1, s_1)) > v_2((s_2, s_2))$ and $v_2((s_1, s_1)) = 1, v_2((s_2, s_2)) = 0, v_2((s_1, s_2)) = b_2 < 0, v_2((s_2, s_1)) = c_2 < 1$. The mixed strategy equilibrium played in an encounter of type (CO) and type (NC) is found in the proof of the previous lemma. Denote the expected payoff of type (CO) in equilibrium by u_2. We need to compare it with what type (NC) gets against itself. We have

$u_2 - u_3 = \frac{((c_2-1)b_3-b_2(c_3-1))(b_3b+c(1-c_3)-b_3)}{(c_2+b_2-1)(b_3+c_3-1)^2} > 0 \Longleftrightarrow$

$(1 - c_2)b_3(-(b_3b + c(1 - c_3) - b_3)) + b_2(1 - c_3)(b_3b + c(1 - c_3) - b_3) < 0.$

If $b_3b+c(1-c_3)-b_3 \neq 0$, it is always possible to find c_2 or b_2 small enough that the above inequality holds and type (NC) is not stable. Consider therefore $b_3b + c(1 - c_3) - b_3 = 0$. Since $c_3 > 1$ and $b_3 > 0$, this equality can hold only when either ($b > 1$ and $c > 0$) or ($b < 1$ and $c < 0$).

Consider the subset of type (NC) preferences satisfying $b_3b + c(1 - c_3) - b_3 = 0$. Any preferences in this subset get $u_3^* = \frac{bc}{b+c-1}$ against themselves as well as any type (CO) preferences get against them. If u_3^* is smaller then either b or c, then either type (St1) or type (St2) are a better response to type (NC). Notice that u_3^* is larger than both b and c when $b < 0$ and $c < 0$. But in this case type (NC) gets in the mixed equilibrium against type (CO) an amount smaller than type (CO) gets against itself, which is 1. ■

The intuition of the result lies in the instability of the mixed equilibrium. Since in the material payoff game not all strategies bring the same payoff, the material payoff in the mixed equilibrium is lower than in some pure strategy combination. If there is a type that can commit to use a pure strategy and achieve that pure strategy combination, such a type makes type (NC) unstable.

Type (St2) preferences are analogous to type (St1) except that type (CO) can do better against itself than type (St2) does against type (CO) thus leading to the instability of preferences (St2).

Lemma 5.5 *No type (St2) preferences are neutrally stable and type (St2) is not neutrally stable.*

Proof. Obviously, if $b > 0$, types (St1) and (NC) are a better response to (St2) than type (St2) itself. If $b \leq 0$, an alternative best response is type (CO). But since type (CO) gets 1 against itself, while type (St2) gets 0 against type (CO), mutants of type (CO) will grow. ■

The argument is similar to the one of "secret handshake" (Robson (1990)), since mutants of type (CO) can recognize each other and coordinate on a better outcome.

Since we have given the indifferent type (BB) preferences the best shot, they are stable quite often.

Lemma 5.6 *All preferences of type (BB) are neutrally stable and type (BB) is neutrally stable if $c \leq 1$ or $b \leq 0$. Moreover, if $b + c > 2$ or $c > 1$ in this region, type (BB) is evolutionarily stable.*

Proof. By assumption, type (BB) preferences play a symmetric equilibrium that maximizes the material payoffs. By Lemma 5.1 the equilibrium is (s_1, s_1), giving a payoff of 1 if $b + c - 2 \leq 0$, and s_1 is played in the equilibrium with probability $p = \frac{b+c}{2(b+c-1)}$, giving a material payoff of $u_5 = \frac{(b+c)^2}{4(b+c-1)} > 1$, if $b + c - 2 > 0$.

Type (St1) plays s_1. Then type (BB) chooses to play s_1 against type (St1) if $c \leq 1$. In this case the material payoff is 1 for both types. Type (St1) can be a best response to type (BB) if $b + c - 2 \leq 0$ but then type (BB) is neutrally stable against type (St1). If $c > 1$, type (BB) plays s_2 against type (St1). Then type (St1) has the material payoff of b. If $b \leq 1$, type (St1) has a lower payoff against type (BB) than type (BB) itself. The interesting case is when $b > 1$ and, therefore, $b + c - 2 > 0$. In this case we should compare u_5 and b. We have $b > u_5 \Longleftrightarrow c < b + 2\sqrt{b(b-1)}$, therefore type (St1) is a better response to type (BB) if $c < b + 2\sqrt{b(b-1)}$.

Type (St2) plays s_2. Then type (BB) plays s_1 if $b > 0$ and type (St2) receives c. If $c \leq 1$ type (St2) cannot be a better response to type (BB). If $c > 1$ and $b+c-2 \leq 0$, type (St2) is a better response to type (BB). Finally, if $c > 1$ and $b+c-2 > 0$, we calculate $c > u_5 \Longleftrightarrow c > \frac{2-b+\sqrt{1-b+b^2}}{3}$. Therefore, type (St2) is a better response if this inequality is satisfied. If $b \leq 0$, type (BB) chooses s_2 and type (St2) receives 0 and cannot be a best response.

Combining the inequalities, we derive that type (BB) is immune against types (St1) and (St2) when $b \leq 0$ or when $c \leq 1$. By a straightforward reasoning, given that type (BB) chooses the best equilibrium for itself, types (CO) and (NC) cannot achieve a higher payoff against type (BB) than type (BB) against itself in these regions either.

The evolutionary stability of type (BB) comes from the observation that when $a+b > 2$, the highest material payoff under symmetric strategies is achieved by a mixed strategy. When playing against other types type (BB) always chooses an equilibrium where the strategy of at least one player is pure. The payoff of the other type in such an equilibrium is strictly lower than what type (BB) gets against itself. Also, when $b \leq 0$ and $c > 1$ even with $b+c \leq 2$, by choosing the best equilibrium for himself, a player of type (BB) makes the payoff of the opponent strictly worse than 1 for any type of the opponent's preferences. ∎

Type (BB) preferences are supported by a private benevolent planner ("mother of the player") that chooses for the player the maximal material payoff equilibrium while the player himself does not care. Since the maximal material payoff is chosen, it is no surprise that such preferences are often stable. That such preferences are not always stable comes from the possibility that the opponent's material payoff will be improved even more when type (BB) chooses the best equilibrium for himself. Thus, when b, c are large, in a symmetric equilibrium a player of type (BB) cannot achieve much, but when he plays against players of types (St1) or (St2), an asymmetric equilibrium arises with higher payoffs for both players than in the symmetric one.

Combining the lemmas we get

Theorem 5.1 *We have the following pattern of neutral and evolutionary stability:*

(i) if $c \leq 1, b > 0, b+c > 2$ *then type (BB) is neutrally and evolutionarily stable;*

(ii) if $c < 1, b > 1, b+c \leq 2$ *then types (St1),(BB) and some preferences of type (CO) are neutrally stable, and no type is evolutionarily stable;*

(iii) if $c < 1, b \leq 1$ *then types (St1),(CO), and (BB) are neutrally stable and no type is evolutionarily stable;*

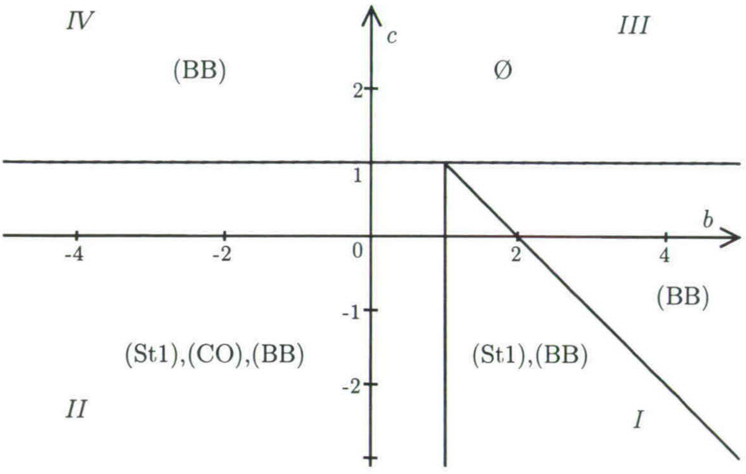

Figure 5.1: Neutrally stable types of preferences.

(iv) if $c > 1, b \leq 0$ *then type (BB) is neutrally and evolutionary stable;*
(v) if $c > 1, b > 0$ *then no type is neutrally or evolutionarily stable.*

The theorem is illustrated in Figure 5.1. The discussion of it, relating the results to the types of games discussed before, follows in the next subsection.

If preferences of type (BB) are not considered then the line $b + c = 2$ disappears from the picture and type (St1) is neutrally stable if $c < 1$ and $b > 1$. In the area $c > 1$, $b < 0$ no type is neutrally stable. If $\alpha = \delta$, the results for (St1) and (St2) should be obviously the same. The other results do not change in this case, except that c should be compared with 0 instead of with 1.

5.3.3 Discussion

In the previous section we found what preferences are neutrally stable for each particular values of b, c. Let us see what it means for the four classes of games discussed in section 5.2.

In games with an *efficient dominant strategy* ($c \leq 1, b \geq 0$, region I in the Figure) we have three subclasses. If there is a symmetric mixed strategy combination with higher material payoffs than 1 ($b + c > 2$), then only type (BB), for whom the efficient combination is chosen by a planner and who can commit to it because of his indifference, is evolutionarily stable. If there is no such mixed strategy combination, then also the "true" (that is, whose preferences are equivalent to the material payoffs) type (St1) is stable. In addition to this, some preferences of type (CO) who believe the game to be a coordination problem but still are able to coordinate on the efficient equilibrium, are stable. The smaller the off-diagonal payoff b, the larger is the set of such type (CO) preferences and when $b \leq 1$ all type (CO) preferences are stable.

In *coordination problems* ($c < 1, b < 0$, region II) the same three types as above are stable. Unconditional play of the efficient pure strategy (type (St1)) survives because it brings the highest possible material payoff and the mutants cannot achieve a higher payoff because they have to play only this strategy too. Since types (CO) and (BB) are also able to pick the efficient equilibrium, they are stable as well. Type (CO) is the type that contains the "true" preferences.

In *chicken type* games ($b > 1, a > 0$, region III) there are no neutrally stable preferences. Since we consider a one-population model, we are interested in symmetric strategies. In such games the only symmetric equilibrium is in mixed strategies. Since the equilibrium is mixed any strategy gives the same payoff against the equilibrium strategy. Whatever preferences that play a mixed equilibrium are (even the "true" preferences) there are other preferences that pick up another strategy with a higher material payoff than the "true" preferences. Thus, evolution never settles on a monomorphic outcome, either there is a polymorphic stationary population composition or the population composition always changes.

In the *prisoners' dilemma* ($b > 1, a < 0$, region IV) again only type (BB) is stable. Type (BB) is helped by a planner that picks the highest material payoff equilibrium for this type. Since the selfish Nash equilibrium is inefficient in the prisoners' dilemma, it is no surprise that the planner can find a more efficient allocation. Note that selfish preferences cannot upset the planner's ones, since the planner recognizes the mutant selfish preferences and instructs type (BB) players to use the dominant strategy s_2 against the mutants.

For type (St2) ("true") preferences, mutants of type (CO) can get the same payoff as type (St2) against (St2). If there are several mutants of type (CO), they sometimes meet each other and receive the cooperation payoff and thus have a higher payoff than type (St2) against (CO). This means that the "true" type (St2) is not neutrally stable. This is similar to the "secret handshake" argument of Robson (1990) to establish cooperation in the prisoners' dilemma. Without type (BB) cooperation cannot persist, however: once there is a majority of players who coordinate, another mutant can invade the population defecting against the coordinators.

The "true" preferences do not always survive evolution: in chicken-type games and in the prisoner's dilemma they are not stable. Notice that in these games the equilibrium with "true" preferences is not efficient, thus evolution promotes efficiency, at least in the prisoners' dilemma. The results also illustrate that a material payoff Nash equilibrium is not necessarily played with the stable preferences.

The players could use the "secret handshake" or, alternatively, they could commit to a strategy because their preferences were known to the opponent. This is a strong assumption. Another rather strong assumption we made was the perfect coordination of the players on one of the equilibria when there are multiple ones. In the next sections we relax these assumptions one by one.

5.4 Imperfect Coordination

We have assumed in the previous section that the players can perfectly coordinate on one of the multiple equilibria. Moreover, players with preferences of type (BB) can be induced to play any strategy without mistakes. These assumptions are, in our view, too strong. Coordination problems may require more effort and attention than games with dominant strategies. Coordination on an equilibrium takes time, and evolution may be working when perfect coordination is not yet achieved. Similarly, since agents with type (BB) preferences have no strong incentives to play any particular strategy, the planner may not be able to induce this strategy with certainty.

To formalize the arguments above, we assume the following. In games with a dominant strategy for either player, that is, games involving types (St1) or (St2), the rationality requirements are not so strong and therefore players play the unique undominated pure equilibrium without mistakes. In a coordination problem, that is, when type (CO) plays against itself, we assume that players *almost* coordinate on the efficient equilibrium. That is, they play s_1 with probability $1-\varepsilon_2$, where ε_2 is arbitrarily small (but larger than the proportion of mutants, so we can use the definition of evolutionary and neutral stability). With the remaining probability ε_2 the players play s_2.

Assumption 4' *In a game of two players with type (CO) preferences, they play the efficient strategy with probability* $1-\varepsilon_2$.

A similar model is applied to players of type (BB), that is, such players play the intended (by the planner) possibly mixed strategy with probability $1-\varepsilon_5$ and some other strategy with probability ε_5, for arbitrarily small ε_5.

Assumption 2' *Players with type (BB) preferences play the strategy that corresponds to the highest for them material payoff equilibrium with probability* $1 - \varepsilon_5$.

For players of type (NC), who plays the mixed equilibrium, we do not need to assume small probability of mistakes, since the symmetric equilibrium for them is unique. Moreover, even if there were mistakes in the equilibrium strategy, leading to another mixed strategy, we can consider players with other preferences of type (NC) for whom the latter strategy is the equilibrium strategy.

We have the following evolutionary game on preferences

	(St1)	(CO)	(NC)	(St2)	(BB)
(St1)	$1, 1$	$1, 1$	b, c	b, c	$\tilde{u}_1^5, \tilde{u}_5^1$
(CO)	$1, 1$	$\tilde{u}_2, \tilde{u}_2$	u_2^m, u_3^m	$0, 0$	$\tilde{u}_2^5, \tilde{u}_5^2$
(NC)	c, b	u_3^m, u_2^m	u^m, u^m	b, c	$\tilde{u}_3^5, \tilde{u}_5^3$
(St2)	c, b	$0, 0$	c, b	$0, 0$	$\tilde{u}_4^5, \tilde{u}_5^4$
(BB)	$\tilde{u}_5^1, \tilde{u}_1^5$	$\tilde{u}_5^2, \tilde{u}_2^5$	$\tilde{u}_5^3, \tilde{u}_3^5$	$\tilde{u}_5^4, \tilde{u}_4^5$	$\tilde{u}_5^5, \tilde{u}_5^5$

where the differences from the previous section are in $\tilde{u}_2$ that depends on ε_2 and in the row and the column of type (BB), where $\tilde{u}_\cdot^\cdot$ incorporates ε_5. The following result is useful.

Lemma 5.7 *If* $b + c > 2$, *then the symmetric strategy* $(1 - \varepsilon)s_1 + \varepsilon s_2$ *gives a material payoff higher than 1 for small* ε. *If* $b + c \leq 2$, *the material payoff from* $(1 - \varepsilon)s_1 + \varepsilon s_2$ *is lower than 1.*

Proof. By straightforward calculation, the material payoff when both players use $(1 - \varepsilon)s_1 + \varepsilon s_2$ is $u = 1 + (b + c - 2)\varepsilon + (1 - b - c)\varepsilon^2$. If $b + c - 2 = 0$, $1 - b - c < 0$, therefore $u < 1$. Since ε is arbitrarily small, only the second term matters if $b + c - 2 \neq 0$, which proves the lemma. ■

The result suggests that the main difference from the previous section lies in that now types (CO) and (BB) cannot achieve perfect coordination and a payoff of 1 if $b + c \leq 2$, while if $b + c > 2$, the mistakes will not hurt but rather help them. Therefore, the main difference in the results is with respect to these types that are less stable if $b + c \leq 2$ and more stable if $b + c > 2$.

Consider the case when $b + c > 2$. If $c < 1$, even without mistakes preferences of type (BB) are stable, so they are stable with mistakes that favor them. Preferences of type (CO) can be stable if $\varepsilon_2 < \varepsilon_5$, since a player is favored more by mistakes of the opponent than by own mistakes in this area. For stability of preference of type (CO), the condition $b \leq \frac{c_2+b_2-1}{b_2}$ from Lemma 5.3 still has to hold. Since it cannot be satisfied for all preferences of type (CO) if $b+c>2$ and $c<1$, some preferences of type (CO) are then stable but not all. If $\varepsilon_2 \geq \varepsilon_5$, only type (BB) preferences are stable.

Now turn to the case $b+c \leq 2$. Without mistakes, type (St1) was neutrally stable when $c<1$. Types (CO) and (BB) are best responses to type (St1) but now they cannot get a higher payoff against themselves. Thus, type (St1) is still neutrally stable and even evolutionarily stable since types (CO) and (BB) have a payoff strictly smaller than 1 against themselves. Moreover, type (St1) is a better response to types (CO) and (BB) than these types themselves since it can secure a payoff of 1 while types (CO) and (BB) get a lower payoff. Thus, preferences of types (CO) and (BB) are not stable if $b+c \leq 2$ and $c<1$.

Other reasonings used in the lemmas in the previous section go through. Therefore, we have

Theorem 5.2 *Under imperfect coordination, the following pattern of neutral and evolutionary stability emerges:*

(i) if $c \leq 1, b>0, b+c>2$ then type (BB) is neutrally and evolutionarily stable and some preferences of type (CO) are neutrally stable if $\varepsilon_2 < \varepsilon_5$;

(ii) if $c<1, b+c \leq 2$ then type (St1) is neutrally and evolutionarily stable;

(iii) if $c>1, b \leq 0$ then type (BB) is neutrally and evolutionarily stable;

(iv) if $c>1, b>0$ then no type is neutrally or evolutionary stable;

The theorem is illustrated in Figure 5.2.

As observed above, the difference from the perfect coordination case lies with respect to the preferences of types (CO) and (BB) who are directly affected by imperfect coordination. When $b+c>2$, type (CO) preferences can be stable now, if their mistakes are smaller than the mistakes of type (BB). Both types (CO) and (BB) profit from their failure to coordinate perfectly since mistakes rather help than hurt in that area. When $b+c \leq 2$, however, only type (St1), that is, unconditional preference for the more efficient strategy s_1, survives. Mistakes are costly in this region; types (CO) and (BB) are prone to them while type (St1) is not. An interesting illustration is a coordination problem. If a player happens to play the efficient strategy without giving it a thought, as if the strategy is the dominant one, he fares better than a player choosing one of the strategies consciously, which is prone to occasional mistakes. Obviously, if strategies are equally efficient ($\alpha = \delta$), any of them can be considered as dominant. Thus, driving on the

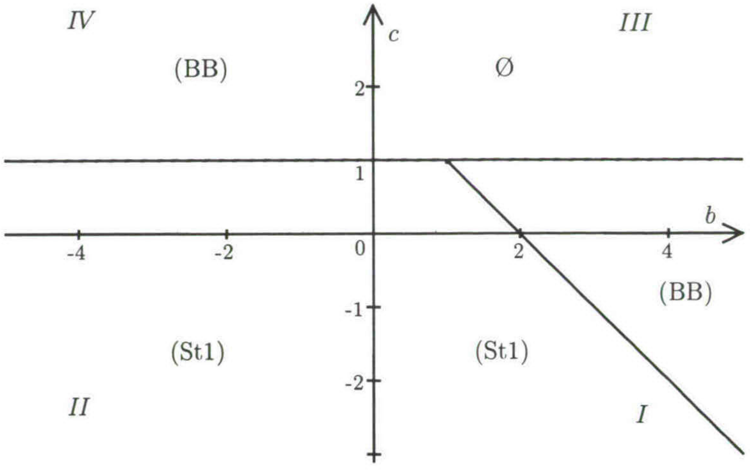

Figure 5.2: Stable types of preferences with imperfect coordination.

right became a habit rather than the result of deliberation (if one does not go out of the country; but beware of England if you think it is the dominant strategy!).

5.5 Incomplete Information

The models above considered the case where players could observe each others' preferences. The assumption is rather strong. In this section we relax this assumption and assume instead that players cannot observe each others' preferences at all. We consider two setups. The first one is when the players know that a mutant has appeared and his type but not his identity. The second setup is of ignorance when the players do not know that a mutant has appeared. We again concentrate on the analysis of monomorphic populations. We return to the case with perfect coordination, thus players of type (CO) coordinate on the efficient equilibrium (s_1, s_1) and players of type (BB) play an equilibrium with the highest material payoff for themselves. Other assumptions are also kept, so only undominated equilibria are played and players of type (NC) play the mixed strategy equilibrium.

5.5.1 Bayesian Approach

Consider a monomorphic population with preferences v_1. Similarly to the complete information case, assume that an arbitrarily small proportion ε of mutants with other preferences v_2 appeared in the population and this is known to all players. But the identity of the mutants is not known, so the players cannot recognize a mutant in a match. Since the population is infinite, each player faces the same distribution of types. Then each match is a Bayesian game with types $\{v_1, v_2\}$, priors $\{1-\varepsilon, \varepsilon\}$ and preferences specified by the utility functions v_1, v_2. One can calculate a symmetric equilibrium of the Bayesian game, where symmetric means that both players use the same strategy. Let us denote the material payoffs of preferences v_1 in equilibrium by $U_1(\varepsilon)$ and of preferences v_2 by $U_2(\varepsilon)$. Similar to the definition of evolutionarily stable with incomplete information preferences in the previous chapter, and the definition of stable types in 5.2, define

Definition 5.5 *A type of preferences k is* ***evolutionary stable with incomplete information*** *if $\exists\varepsilon^* > 0$ such that $\forall v_1 \in k$ it holds that $U_1(\varepsilon) > U_2(\varepsilon)$ $\forall v_2 \notin k$, $\forall\varepsilon \in (0, \varepsilon^*)$.*

Definition 5.6 *A type of preferences k is* ***neutrally stable with incomplete information*** *if $\exists\varepsilon^* > 0$ such that $\forall v_1 \in k$ it holds that $U_1(\varepsilon) \geq U_2(\varepsilon)$ $\forall v_2 \notin k$, $\forall\varepsilon \in (0, \varepsilon^*)$.*

The discussion of the definition and comparison of it with definitions used elsewhere can be found in the previous chapter. Again, we analyze the types of preferences in turn.

Lemma 5.8 *Types (St1) and (CO) are neutrally stable with incomplete information if $c < 1$ or $c = 1, b > 0$.*

Proof. Type (St1) always plays its dominant strategy s_1 in equilibrium. For type (CO), since the proportion of the mutant preferences is arbitrarily small, it is also optimal to play a strategy arbitrarily close to s_1 in an equilibrium of the Bayesian game. Therefore, for a mutant type to get a higher payoff the mutant has to play a better reply to s_1. If $c > 1$, s_2 is a better reply, and, for example, a mutant of type (St2) plays it. If $c = 1$, the fitness of the players depends on the performance against s_2, which is reflected in $b > 0$. Types (St1) and (CO) are not evolutionarily stable since both of them play s_1 in equilibrium of the game with only them present, achieving the same payoff. ■

Thus, type (St1) is stable with incomplete information in the same regions as with complete information, and also when $b + c > 2$ since now players with preferences of type (BB) cannot recognize each other to profit from the possibility of mixed strategies.

Type (NC) is now no problem for type (CO) since there is always an equilibrium of the Bayesian game where type (CO) plays a strategy arbitrarily close to s_1 and receives a payoff of almost 1 and type (NC) plays a strategy arbitrarily close to s_2 and can get a higher payoff only when $c > 1$.

For type (St2) the reasoning is similar, only mirrored with respect to b, c.

Lemma 5.9 *Type (St2) is neutrally stable with incomplete information if $b < 0$ or $b = 0, c > 1$.*

The reason that type (St2) is now stable is that types (CO) or (BB) cannot use the "secret handshake" anymore and have to play strategy s_2 as well. Note that type (CO) preferences can be made stable in this region, if they could coordinate on the other, inferior equilibrium.

Type (NC) that was never stable under complete information, is not stable under incomplete information either.

Lemma 5.10 *Type (NC) is not neutrally stable with incomplete information.*

Proof. Consider type 3 (NC) preferences with $v_3((s_1, s_1)) > v_3((s_2, s_2))$ (the proof is similar for the other case). Without loss of generality, we can normalize them to $v_3((s_1, s_1)) = 1, v_3((s_2, s_2)) = 0, v_3((s_1, s_2)) = b_3 > 0, v_3((s_2, s_1)) = c_3 > 1$. Consider a Bayesian game with proportions $1 - \varepsilon$ of players of type (NC) and ε of players of type (St1). Denote by $U_3(\varepsilon)$ the material payoff in an equilibrium of the game for type (NC) and by $U_1(\varepsilon)$ the material payoff of type 1 (St1). It follows from calculations that $\lim_{\varepsilon\to 0}(U_3(\varepsilon) - U_1(\varepsilon)) = \frac{(1-c_3)(b(c_3-1)-b_3(c-1))}{(b_3+c_3-1)^2}$. Since $1-c_3 < 0$, $\lim_{\varepsilon\to 0} U_3(\varepsilon) \geq \lim_{\varepsilon\to 0} U_1(\varepsilon)$ if $b(c_3 - 1) - b_3(c - 1) \leq 0$. Analogously, in the Bayesian game between players of types 3 (NC) and 4 (St2), $\lim_{\varepsilon\to 0}(U_3(\varepsilon) - U_4(\varepsilon)) = \frac{b_3(b(c_3-1)-b_3(c-1))}{(b_3+c_3-1)^2}$, thus $\lim_{\varepsilon\to 0} U_3(\varepsilon) \geq \lim_{\varepsilon\to 0} U_4(\varepsilon)$ if $b(c_3 - 1) - b_3(c - 1) \geq 0$. Thus, if $b(c_3 - 1) - b_3(c - 1) \neq 0$ either type (St1) or type (St2) have a higher material payoff in the equilibrium of the Bayesian game for some small ε.

Consider from now on preferences of type (NC) satisfying $b(c_3 - 1) - b_3(c - 1) = 0$. The equality can hold only when either ($b > 0$ and $c > 1$) or ($b < 0$ and $c < 1$). Given that $\lim_{\varepsilon\to 0}(U_3(\varepsilon) - U_1(\varepsilon)) = 0$ in this case, consider $\lim_{\varepsilon\to 0}(U_3'(\varepsilon) - U_1'(\varepsilon)) = \frac{(c_3+b_1)(c-1)^3}{(c_3-1)(b+c-1)^2}$, where $b_1 = v_1((s_1, s_2)) \geq 0$ since type (St1) preferences regard s_1 as the dominant strategy. Thus, $\lim_{\varepsilon\to 0}(U_3'(\varepsilon) - U_1'(\varepsilon)) > 0$ only if $c > 1$.

Consider now the Bayesian game between players of types (NC) and (CO). Take type (CO) preferences normalized to $v_2((s_1, s_1)) = 1, v_2((s_2, s_2)) = 0, v_2((s_1, s_2)) = b_2 <$

$0, v_2((s_2, s_1)) = c_2 < 1$. Calculations reveal that in equilibrium of this game $\lim_{\varepsilon\to 0}(U_3(\varepsilon) - U_2(\varepsilon)) = 0$ and $\lim_{\varepsilon\to 0}(U_3'(\varepsilon) - U_2'(\varepsilon)) = \frac{(b-c+1)(c-1)(b(c_2-1)-b_2(c-1))}{(1-c_3)(b+c-1)^2}$. Since $b_2, c_2 - 1 \in (-\infty, 0)$ are arbitrary, the expression can be made negative if $b - c + 1 \neq 0$. Then type (CO) preferences have a higher material payoff in equilibrium than type (NC) preferences have for small ε.

Consider therefore $b - c + 1 = 0$. Then $\lim_{\varepsilon\to 0}(U_3''(\varepsilon) - U_2''(\varepsilon)) = \frac{(c_2-b_2-1)^2(1-c)}{2(c_3-1)^2} < 0$ if $c > 1$ and $c_2 - b_2 - 1 \neq 0$. Then type (CO) preferences have a higher material payoff. For any b, c we found another type of preferences that has a higher payoff in equilibrium of the Bayesian game. Therefore, type (NC) is not neutrally stable. ■

For type (NC) that plays the mixed equilibrium with complete information, there are always some preferences that move the equilibrium of the incomplete information game towards an outcome that is better for them than for type (NC). Even in the chicken type games, when type (NC) contains the "true" preferences, there are preferences that move the mixed equilibrium of the Bayesian game towards a pure equilibrium that is more favorable for mutant preferences. Then these mutant preferences grow.

For players with preferences of type (BB), observe from the previous sections that they plays a symmetric mixed strategy combination when $b + c - 2 > 0$. However, such a strategy combination is never an equilibrium of the material payoffs game. Therefore, a mutant can appear so that it plays a best response to this strategy combination. This mutant is not recognized by the players and, therefore, it receives a higher payoff. When $b + c - 2 \leq 0$, type (BB) plays strategy s_1 against itself and by a reasoning similar to the reasoning for types (St1) and (CO) is stable when these types are stable, that is, when $c < 1$ or $c = 1, b > 0$.

Lemma 5.11 *Type (BB) is not neutrally stable if $b + c - 2 > 0$. If $b + c - 2 \leq 0$ it is neutrally stable if $c < 1$ or $c = 1, b > 0$.*

Summarizing the results, we can conclude

Theorem 5.3 *With incomplete information the following types are neutrally stable:*
(i) if $c < 1$ or $c = 1, b > 0$ then types (St1), (CO) and (BB) (if $b + c - 2 \leq 0$);
(ii) if $b < 0$ or $b = 0, c > 1$ then type (St2);
(iii) if $c > 1, b > 0$ then no type.

Compared with the results of the previous sections, now types that play a (strict) Nash equilibrium of the material payoff game are stable. For example, in coordination problems ($b < 0, c < 1$) types (St1), (CO), (St2), and (BB) are all neutrally stable, though types (St1), (CO), and (BB) play the efficient equilibrium and type (St2) plays the inefficient equilibrium. In the prisoners' dilemma ($b < 0, c > 1$) only type (St2) that plays the equilibrium of it is stable. Thus, preferences that are "true", or preferences that are equivalent to "true" preferences (in the sense that they play a Nash equilibrium of the game with "true" preferences) are stable if they play a strict or undominated Nash equilibrium. However, if players with the "true" preferences play a mixed equilibrium, such preferences need not be stable. Ely and Yilankaya (1997) and Ok and Vega-Redondo (1999) show that the "true" preferences (or preferences equivalent to them) are stable in their models, but they did not allow for mixed strategies. In this respect our model and the results of these subsection differ from theirs.

The following subsection relaxes the informational assumptions even further.

5.5.2 Ignorance

Suppose now that players in a (originally monomorphic) population do not suspect that a mutant with other preferences has appeared, and therefore they still continue to play what they used to play. The mutant, however, plays a best reply to the strategy of the original players. Clearly, if the original players played a (material payoff) best reply against themselves, the mutant can be no worse only if he plays a best reply as well. Analogously, if the original players are not playing an equilibrium, a mutant could appear that plays a best reply to the strategies of the original population and has a higher material payoff.

Denote the strategy of the players in the original population by s_i and the strategy of the mutant by s_m. The expected material payoff of the players with the original preferences is $U_i(\varepsilon) = (1-\varepsilon)u_i(s_i, s_i) + \varepsilon u_i(s_i, s_m)$ and of the players of the mutant type is $U_m(\varepsilon) = (1-\varepsilon)u_m(s_m, s_i) + \varepsilon u_m(s_m, s_m)$. In the spirit of the previous subsection, taking arbitrarily small ε,

Definition 5.7 *A type of preferences k is* ***evolutionary stable under ignorance*** *if $\forall v_i \in k$ $\lim_{\varepsilon\to 0} U_i(\varepsilon) > \lim_{\varepsilon\to 0} U_m(\varepsilon)$ $\forall v_m \notin k$.*

Definition 5.8 *A type of preferences k is* ***neutrally stable under ignorance*** *if $\forall v_i \in k$ $\lim_{\varepsilon\to 0} U_i(\varepsilon) \geq \lim_{\varepsilon\to 0} U_m(\varepsilon)$ $\forall v_m \notin k$.*

Comparing the expressions for the expected material payoffs, one can observe that $\lim_{\varepsilon\to 0} U_i(\varepsilon) > \lim_{\varepsilon\to 0} U_m(\varepsilon)$ when $u_i(s_i, s_i) > u_m(s_m, s_i)$ and if $u_i(s_i, s_i) = u_m(s_m, s_i)$ then when $u_i(s_i, s_m) > u_m(s_m, s_m)$, which is precisely the standard definition of evolutionary stability for the game with the material payoffs. Thus, if a type plays the strategy that is evolutionary stable in the game with material payoffs, then this type is at least neutrally stable under ignorance.

It is easy to see that if $c < 1$ or $c = 1, b > 0$, strategy s_1 is evolutionarily stable. Therefore, types (St1), (CO), and (BB) (if $b + c - 2 \leq 0$) are neutrally stable under ignorance if $c < 1$ or $c = 1, b > 0$. Analogously, if $b < 0$ or $b = 0, c > 1$, type (St2) is neutrally stable. A difference from the previous subsection appears in type (NC). If the material payoff game is of chicken type, the unique mixed strategy equilibrium is evolutionarily stable. Therefore, if players with preferences of type (NC) play the strategy corresponding to the equilibrium, such preferences are neutrally stable. To play the material payoff game equilibrium strategy, preferences of type (NC) with $v_3((s_1, s_1)) > v_3((s_2, s_2))$ and $v_3((s_1, s_1)) = 1, v_3((s_2, s_2)) = 0, v_3((s_1, s_2)) = b_3 > 0, v_3((s_2, s_1)) = c_3 > 1$ should satisfy $b_3(c - 1) = b(c_3 - 1)$. This holds not for all preferences of type (NC) (but it holds for the "true" preferences), therefore type (NC) is not neutrally stable.

Theorem 5.4 *With ignorance the following types are neutrally stable*

(i) if $c < 1$ or $c = 1, b > 0$ then types (St1),(CO) and (BB) (if $b + c - 2 \leq 0$);

(ii) if $b < 0$ or $b = 0, c > 1$, then type (St2);

(iii) if $c > 1$ and $b > 0$ then no type is neutrally stable but some preferences of type (NC) are neutrally stable.

The conclusions of the two last theorems are illustrated in Figure 5.3. They do not change if preferences of type (BB) are not considered and when $\alpha = \delta$ (then c should be compared with 0).

The two last theorems show that the preferences that play a Nash equilibrium of the game with material payoffs, are more often stable under incomplete information. Moreover, only preferences that are "true", or equivalent to "true" ones, are stable. Preferences of types (St1) and (St2) are stable for a larger area than in the complete information case, since they play a Nash equilibrium also in coordination games (region II). However, with the Bayesian approach, there is still a possibility for the "true" preferences not to be stable. It happens for preferences of type (NC), that play a mixed strategy equilibrium against themselves. With ignorance only the "true" preferences, or preferences equivalent to them, survive the evolutionary pressure, a result that is in line with other recent results in indirect evolution under incomplete information (Ely and Yilankaya (1997) and Ok and Vega-Redondo (1999)).

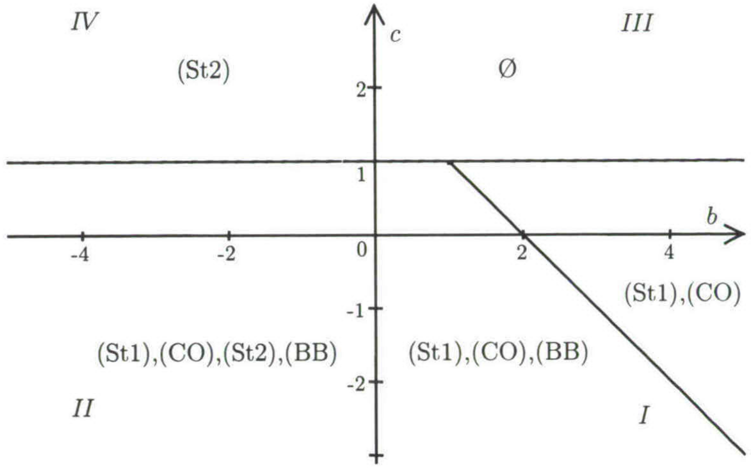

Figure 5.3: Neutrally stable types with imperfect information.

5.6 Conclusion

Of course, the results of the previous sections should be taken with caution. Many of them rely on the particular choice of equilibria (or approximate equilibria) for certain preferences. Other choices may lead to other results but the analysis indicates that the problem of coordinating is of importance. Another limitation of the analysis is that only monomorphic states were considered while non-monomorphic or type-monomorphic states can be stable as well. However, the analysis of monomorphic states allows to draw some conclusions. With complete information selfish preferences may be unstable and evolution improves efficiency, while with incomplete information, for any (strict) equilibrium in pure strategies of the material payoff game, one can find preferences that play the equilibrium strategies and such preferences are stable even if the equilibrium is inefficient.

The consideration of incomplete information supports the claim that incomplete information promotes selfishness (Ely and Yilankaya (1997), Ok and Vega-Redondo (1999)). The preferences that are equivalent to material payoffs are more often stable under incomplete information. However, they are not necessarily stable. If the material payoff game has only a mixed symmetric equilibrium, such preferences would not be stable in the Bayesian game we considered. The main difference between our model and the model of papers mentioned above lies in the possibility of mixed strategies.

Consideration of all possible preferences reduces chances for any particular preferences to be stable. Since the space of preferences is large, there is a chance that for any given preferences there are other preferences that beat them.

Inclusion of type (BB) preferences in the analysis, represented by a planner that maximizes their material payoff is, in our view, an interesting approach. They can be stable, but not always. Following advice, even for own good, is not necessarily stable, since the advice may benefit the opponent even more. In the prisoners' dilemma such preferences are the only stable ones under complete information, since there are gains from cooperation. In coordination problem, however, they are not stable (assuming a small possibility of mistakes in implementation of the planner's advice) since preferences that regard one of the strategies as dominant choose one of the equilibria more efficiently.

Chapter 6

Merger Games and Coordination

6.1 Introduction

This chapter analyses a non-cooperative game of coalition formation inspired by mergers. Mergers are an important concern of policy makers and industry regulators. Though regulations usually deal with a merger that has been proposed, it is interesting and also important to know which merger in an heterogeneous industry is more likely to be proposed, since it allows the regulators to predict the development of the industry. The proposed merger is determined endogenously; this chapter attempts to find out which of the possible mergers is most likely.

It is well known that a merger can be unprofitable for its participants, for example, it is so in the simple setting of symmetric Cournot oligopoly (see Salant et al. (1983)). Since the firm that results from a merger reduces aggregate production in equilibrium, it pushes price up thus making other firms better off, while firms participating in the merger have smaller aggregate profit. That a merger is unprofitable does not rule it out when firms are rational as firms may merge to prevent an even more disadvantageous merger, as shown in Fridolfsson and Stennek (1999).

If firms are asymmetric, a merger can be profitable in the Cournot setting since it improves efficiency. Profitable mergers with asymmetries in costs are considered in Perry and Porter (1985). A fairly general analysis of mergers in Cournot oligopoly, including a welfare analysis and policy implications is found in Farrell and Shapiro (1990). In these papers, however, a merger is given exogenously and the question is whether there are incentives to merge, that is, whether the merger is profitable. In contrast, we consider the question of which one of several profitable mergers is more likely to occur.

There have been attempts to address this question from the point of view of cooperative game theory. Rajan (1989) considers several games in characteristic function form corresponding to several conjectures about how the firms outside a merger will act; no consistency requirement is imposed on these conjectures. Horn and Persson (1996) introduce a new concept of domination between coalition structures and predict that a merger that yields maximum industry profit will arise. The dominance concept does not take into account the problem of free riding, since most of the combined profit in the dominant structure may accrue to the firm outside of a merger. Since transfers between coalitions are not allowed, it is counterintuitive that a dominant structure with most gains going to an outsider of the merger will arise.

We model the merger game as a non-cooperative game of coalition formation instead. The idea of such games is that before the actual play takes place, there is a stage of coalition formation, that is, players can make agreements that will be binding in the actual play. The stage of coalition formation has strict rules and is also modeled as a non-cooperative game. The two-stage nature of the resulting game may lead to too many equilibria; the standard refinement in this case is subgame perfection where incredible threats are eliminated.

More precisely, the merger game is modeled as a simultaneous moves, exclusive membership game (Bloch (1997)) with a restriction on the size of coalitions. That is, players simultaneously propose certain coalitions, and a coalition is formed if (and only if) all players in the coalition propose the same coalition. This way of modelling is chosen because of several reasons. Exclusive membership refers to the impossibility to force a player into a coalition; it is, in our view, a better way to model a merger, since a firm in reality can always refuse to merge. The simultaneous moves part has two advantages. First, sequential moves models often have results that depend crucially on the order of moves. Even when the results do not depend on the order of moves, predictions are not clear-cut because of multiple equilibria. Though the simultaneous moves model also has multiple equilibria, it allows to apply equilibrium selection techniques more easily. Then in the asymmetric setting a clear-cut result can be achieved. Kamien and Zang (1990,1991) have followed the way of noncooperative modeling of the merger game but the firms there were symmetric and the question addressed was how many firms would merge rather that which firms would merge, like in this chapter.

We show that if there are several profitable mergers, there are correspondingly several Nash equilibria in the game. The problem then is on which equilibrium to coordinate. We address the question of equilibrium selection in this coordination problem. With certain assumptions on the payoff division rule in the merger, the structure of the merger game is such that many refinements that use notions of small perturbations of strategies and best response dynamic lead to the same equilibrium.

An example in the chapter analyzes this coordination problem in an asymmetric Cournot triopoly. Whereas all firms may benefit from a merger, a firm may prefer most to stay out of the merger since the merger creates positive externalities. Therefore, there is a potential free rider problem. Nevertheless, we will show that a merger whose participants prefer this merger to other mergers is in a sense "focal". With a particular specification of the division of gains from merger, we show that this merger is the one with the highest internal gains. Barros (1998) uses this criterion (without justification) to chose among mergers in a model similar to ours but where cost parameters are more restricted. We provide a justification for the highest internal gains criterion.

Welfare considerations are an important issue in the analysis of mergers. Antitrust policies aim at identifying and preventing a merger that decreases welfare. In the asymmetric firm case, a merger may improve efficiency (as the less efficient firm ceases production) while still having a negative effect on welfare due to the reduced number of competitors. We analyze when the "focal" merger improves and when it hampers efficiency. We consider also the implications of the "focal" equilibrium for the producer surplus and find that it is not necessary maximized there, in contrast with the results of Horn and Persson (1996).

The structure of the chapter is as follows. Section 6.2 describes the general model. The rest of the paper focuses on the three-player case. The analysis of the game is in Section 6.3, while Section 6.4 considers the Cournot example. Section 6.5 concludes.

6.2 The Merger Game

We model the merger game as a simultaneous moves, exclusive membership game (Bloch (1997)) with partition function and restriction on the size of coalitions. Let $N = \{1, ..., n\}$ be the set of players. A *coalition* $S = \{i_1 ... i_m\}$ is a subset of N. A partition of set N into coalitions is a *coalition structure*. That is, a coalition structure is a collection of sets $\pi = \{S_1, ..., S_k\}$ such that $\cup_{i=1}^k S_i = N$ and $S_i \cap S_j = \emptyset \; \forall i \neq j$. A pair $(S; \pi)$ of a coalition S and a partition π that contains S is called an *embedded coalition*. Denote the set of all embedded coalition by Ω. A *partition function* $v : \Omega \to \mathbb{R}$ is a function from all possible embedded coalitions to the real line. For each embedded coalition $(S; \pi)$ it gives the worth of coalition S in partition π, denoted by $v(S; \pi)$.

In a general simultaneous moves, exclusive membership game players announce a coalition to which they want to belong. Depending on the announcements, a coalition structure is formed. Hart and Kurz (1983) introduced two models, differing in the mapping from the announcements to the coalition structures. In model γ a coalition S is formed if and only if all its members announced S. In model δ a coalition S is formed if and only if all its members announced the same coalition S' that may be larger than S. After the coalition structure is known, the worth of each coalition in it can be calculated by the partition function.

We restrict announcements to one-player and two-player coalitions only. A justification is that a merger usually involves only two firms at once. The restriction eliminates the need to distinguish between γ and δ models since the only possibility to form a two-player coalition is that both players in the coalition announce this coalition. Alternatively, in the merger game players simultaneously announce their intended partner. If two players have announced each other, they merge. There may be no intended partner (player i announces the singleton coalition $\{i\}$), meaning that a player does not want to form a coalition with anybody but prefers to be a singleton.

Formally, the strategy set X_i of player i is the set of all one and two-player coalitions containing i: $X_i = \{\{i\}, \{ij\}_{j=1,j\neq i}^n\}$. Denote by $x_i \in X_i$ the strategy (announcement) of player i and by $x_j \in X_j$ the strategy (announcement) of player j. Coalition $\{ij\}$ is formed if and only if $x_i = \{ij\} = x_j$.

Observe that $|X_i| = n$. Thus, an alternative representation of X_i is N. A strategy (announcement) $x_i \in N$ is interpreted as the intended partner for player i where $x_i = i$ means that player i prefers to be a singleton. Coalition $\{ij\}$ is formed if and only if $x_i = j$ and $x_j = i$. This representation of strategies is simpler but we will use both representations in what follows.

For the full description of the non-cooperative merger game a division of the coalitional worth to the two players forming a two-player coalition is needed. However, even without specification of the division rule one can state the following result. There is always a Nash equilibrium of the merger game when all players do not name any intended partners (that is $x_i = i \ \forall i$). Since to form a coalition the consent of both players is needed, a unilateral proposal to another player does not change the resulting coalition structure consisting of singletons. Payoffs remain unchanged too.

Remark 6.1 *There is always a Nash equilibrium where all players remain singletons.*

To specify how individual players value different coalition structures, we need a division of the coalitional worth for the players. We assume that there is an exogenously given division rule $\varphi_i^\pi(ij)$ that specifies what player i gets in coalition $\{ij\}$ in coalition structure π. We require the following properties from the rule.

Definition 6.1 *A division rule φ is **efficient** if $\varphi_i^\pi(ij) + \varphi_j^\pi(ij) = v(\{ij\};\pi)$ $\forall i,j$ and any coalition structure $\pi \ni \{ij\}$.*

Consider any coalition structure $\pi \ni \{ij\}$. Denote by π_{-ij} the coalition structure resulting from breaking up coalition $\{ij\}$ while other coalitions stay intact, $\pi_{-ij} = \pi \backslash \{ij\} \cup \{i\} \cup \{j\}$.

Definition 6.2 *A division rule φ is **individually rational** if $\forall i,j$ and any coalition structure $\pi \ni \{ij\}$, whenever $v(\{ij\};\pi) \geq v(\{i\};\pi_{-ij}) + v(\{j\};\pi_{-ij})$ then $\varphi_i^\pi(ij) \geq v(\{i\};\pi_{-ij})$.*

Definition 6.3 *A division rule φ is **strictly individually rational** if $\forall i,j$ and any coalition structure $\pi \ni \{ij\}$, whenever $v(\{ij\};\pi) > v(\{i\};\pi_{-ij}) + v(\{j\};\pi_{-ij})$ then $\varphi_i^\pi(ij) > v(\{i\};\pi_{-ij})$.*

That is, a rule is efficient if the worth of a coalition is fully divided. For a rule to be strictly individually rational, it should be the case that if a merged worth is higher than the sum of the separate worths, keeping other coalitions fixed, then the division rule should give payoffs higher than the individual payoffs of each of the participants in the coalition. We consider only efficient and strictly individually rational division rules.

6.3 The Three Player Case

6.3.1 Equilibria

The restriction of the model to the three-player case simplifies the analysis considerably. The main feature of the three-player case is that the value of a two-player coalition $v(\{ij\})$ is determined, since the only coalition structure compatible with a two-player coalition is $(\{ij\},\{k\})$. This will allow us to drop superscript π in the specification of a division rule. The value of a one-player coalition, though, may depend on whether the other two players formed the coalition or not. However, since a single player cannot unilaterally influence the formation of the other players' coalition, this has limited relevance. In particular, for the set of pure Nash equilibria it is not important whether the formation of a coalition hurts or helps outsiders. Spillovers, that is, how a merger affects the outsider, do not matter for the set of pure Nash equilibria of the game.

The set of pure strategy Nash equilibria in the game with three players depends only on the profitability of forming two-player coalitions. It is easy to show that if a two-player coalition has a higher worth than the sum of the individual worths, then there exists a pure Nash equilibrium of the merger game in which this coalition forms. Formally, denote the coalition structure $(\{i\}, \{j\}, \{k\})$ by π_{ijk} and $v(\{i\}; \pi_{ijk})$ by $v(i)$. Then we have

Theorem 6.1 *If $v(\{ij\}) \geq v(i) + v(j)$ then there exists a pure Nash equilibrium of the merger game in which $\{ij\}$ forms.*

Proof. Suppose player i proposes to j and player j proposes to i. Suppose also that player k wants to remain a singleton. We show that this is an equilibrium. First, player k cannot change the coalition structure by unilateral deviations and therefore receives the same payoff regardless of her strategy. If players i or j change their strategy, the resulting coalition structure would be all singletons since no coalition offers would coincide. By the individual rationality of the division rule, both players i and j receive higher (or equal) payoff when they merge, therefore they would not deviate. ■

It is also true that if a merger is unprofitable ($v(\{ij\}) < v(i)+v(j)$), at least one of the players will have incentives to remain a singleton instead. Therefore, the possible outcomes of pure Nash equilibria are characterized by the profitability of two-player coalitions: those coalitions that are profitable will arise in a pure Nash equilibrium, and those that are not profitable will not.

Since there is also the equilibrium where no merger forms, if there are profitable two-player coalitions, the game exhibits multiplicity of (pure) equilibria. Some of them, like the no-merger equilibrium, seem to be rather unstable. We want to select one equilibrium in order to be able to say more about possible outcomes of the game.

6.3.2 Equilibrium Selection

We continue to focus on pure equilibria. First, we consider static approaches to equilibrium selection like perfection and persistency (see, e.g. van Damme (1991, Ch.2)). Then we also discuss set-valued equilibrium concepts and dynamic approaches to equilibrium selection. All approaches select the same equilibrium in the merger game. One can observe that the notions of perfection and persistency make use of small perturbations, like dynamic approaches do. Thus, they are similar, and in the merger game lead to the same equilibrium.

The Static Approach

We start with the notion of (pure) perfect equilibrium for our game.

Definition 6.4 *A Nash equilibrium $x = (x_1, x_2, x_3)$ is* ***perfect*** *if there exists a sequence of completely mixed strategy profiles $\{\sigma^i\}_{i=1}^{\infty} \to x$ such that $x_j \in BR(\sigma^i_{-j})$ $\forall j = 1, 2, 3$.*

Observe that since x_j has to be a best response to a completely mixed strategy, no weakly dominated strategy can be part of a perfect Nash equilibrium. Thus, like in the previous chapter, we consider equilibria in undominated strategies.

Assume that there exists a strictly profitable merger for players i, j, that is, $v(\{ij\}) > v(i)+v(j)$. Then the strategy "stay alone" is weakly dominated for players i, j by strategy $\{ij\}$ because the only difference a unilateral switch of player i from "stay alone" to $\{ij\}$ can make is to change the payoff of player i from $v(i)$ to $\varphi_i(ij)$ if player j was proposing to i. Recall that we work with strictly individually rational division rules, thus $\varphi_i(i) > v(i)$. Therefore, any equilibrium in which any of the players i, j uses strategy "stay alone", in particular the equilibrium from Remark 6.1 where all players played "stay alone", is not perfect. Even when there are large positive externalities from a merger, that is, a player benefits from the merger of other players very much, possibly more than the participants of the merger themselves, trying to free ride on the merger of others is a dominated strategy. Thus, the (possibly inefficient) equilibrium of no merger is not selected if there is a (strictly) profitable merger.

To save space we will call strictly profitable merger simply profitable. Clearly, if there is only one profitable merger, it is selected by the perfection criterion. The situation is more interesting if there are two or three profitable mergers. Then the perfection criterion is not enough to select one merger as the following example shows.

Consider a situation when only mergers $\{12\}$ and $\{13\}$ are profitable and the division rule is such that $\varphi_1(12) > \varphi_1(13)$ and $\varphi_2(12) > \varphi_2(23) > v(2)$. Then equilibrium $(\{12\}, \{12\}, \{13\})$ with merger $\{12\}$ and equilibrium $(\{13\}, \{23\}, \{13\})$ with merger $\{13\}$ are both perfect. The first equilibrium is perfect since the strategies of all three players are best replies to any mixed strategy close enough to the equilibrium except when player 2 offers the merger with player 3 much more often than player 1 does. The second equilibrium is perfect since if the elements of the sequence $\{\sigma^i\}$ have much smaller probability on player 1 playing $\{12\}$ than on player 3 playing $\{23\}$, $\{23\}$ is a best reply for player 2. For other sequences of completely mixed strategies converging to that equilibrium it may be not true that the equilibrium strategies are best responses. These observations lead to the feeling that the second equilibrium is less attractive than the first one.

To distinguish the two equilibria we employ the concept of *persistent* equilibrium (Kalai and Samet (1984)). A *retract* is a non-empty closed convex subset of the set of mixed strategies of all players. A retract Θ is *absorbing* if for sufficiently small ε, for every strategy profile σ in the ε-neighborhood of Θ, there exists strategy profile $\rho \in \Theta$ such that ρ is a best reply to σ for all players. A *persistent* retract is a minimal absorbing retract, that is, an absorbing retract that does not contain any smaller absorbing retracts. Now we are ready for

Definition 6.5 *A* ***persistent*** *equilibrium is any equilibrium that belongs to a persistent retract.*

Intuitively, an absorbing retract attracts (under best response) the play back to itself after sufficiently small mistakes by the players. Observe that it attracts the play back for every strategy in the ε-neighborhood, which distinguishes it from perfect and proper equilibrium.

We make also an additional assumption on the division rule.

Definition 6.6 *A division rule φ is* ***acyclic*** *if $\exists i, j$ such that $\varphi_i(ij) > \varphi_i(ik)$ and $\varphi_j(ij) > \varphi_j(jk)$.*

That is, there are players i, j such that i prefers merger with j over merger with k and j prefers merger with i over merger with k under φ. Acyclicity rules out cycles in preferences over mergers. That is, the situation when i prefers merger with j, j prefers merger with k, and k prefers merger with i is ruled out. Thus, one of the three possible mergers is preferred by its participants over the merger with the third player. We call such a merger *the most preferred partners* merger. We consider from now on only acyclic division rules. In the example in the next section we present an acyclic division rule arising from quite reasonable assumptions.

Without loss of generality, suppose that the merger referred to in the above paragraph is between players 1 and 2. Then for any (mixed) strategy x_3 of player 3 the strategy profile $(\{12\}, \{12\}, x_3)$ is an equilibrium. We show that at least some of such equilibria are persistent. Consider a retract Θ consisting of $(\{12\}, \{12\}, x_3)$, where $x_3 \in S_3$. For players 1 and 2, strategy $\{12\}$ is a strict best reply to any strategy profile that belongs to Θ since $\varphi_1(12), \varphi_2(12)$ are the maximal values players 1 and 2 can get by participating in a merger, and therefore, it is a best reply to any strategy in a small enough neighborhood of Θ. Since Θ contains all strategies of player 3, it contains best replies to any strategy in the neighborhood of Θ as well. Thus, Θ is absorbing. The whole retract Θ may be not minimal; a subset of it where player 3 does not use (weakly) dominated strategies is then minimal. Any equilibrium of the form $(\{12\}, \{12\}, x_3)$ that belongs to such a subset is persistent.

Suppose now that there is another profitable merger, say, between players 1 and 3. Then $(\{13\}, x_2, \{13\})$, where x_2 puts low enough probability on $\{12\}$ so that $\{13\}$ is a best response for player 1 and low enough probability (which may be 1) on $\{23\}$ so that $\{13\}$ is a best response for player 3, is an equilibrium. Suppose there is a persistent retract Θ' that contains $(\{13\}, x_2, \{13\})$ for some x_2 as described. Any neighborhood of Θ' would contain a mixed strategy of player 1 that puts a non-zero probability on $\{12\}$ higher than the probability of player 3 playing $\{23\}$. Then $\{12\}$ would be the unique best reply for player 2. Thus, Θ' has to contain $(\{13\}, \{12\}, \{13\})$. But then $\{12\}$ is the best reply for player 1; Θ' has to contain $(\{12\}, \{12\}, \{13\})$ as well. In the previous paragraph we showed that $(\{12\}, \{12\}, \{13\})$ belongs to a persistent retract that does not contain $(\{13\}, x_2, \{13\})$. By Corollary 3 of Kalai and Samet (1984) that any two persistent retracts have empty intersection, we arrive to a contradiction. Therefore, there is no persistent retract containing $(\{13\}, x_2, \{13\})$ and such equilibrium is not persistent. A similar reasoning shows that the equilibrium when everybody stays alone is not persistent either.

Combining the reasonings above we state

Theorem 6.2 *If there are profitable mergers and the division rule φ is acyclic, in the pure strategy persistent equilibrium of the merger game the most preferred partners merger forms.*

Thus, using the static refinements of the Nash equilibrium, we can select one merger. Observe that the notion of strong Nash equilibrium does not necessarily help in this situation. A Nash equilibrium is *strong* if there is no group of players such that they can get a better payoff than in the equilibrium by a joint deviation. If a firm has large enough positive externalities from the merger of other firms, the equilibrium with such a merger is strong since this firm would not want to make a joint deviation. Thus, it may be the case that any equilibrium with a profitable merger is strong.

In the next subsection we consider set-valued and dynamic approaches for equilibrium selection.

Set-Valued Concepts and Dynamic Approach

We used the notion of persistent equilibrium to select one of the pure strategy equilibria of the merger game. This notion is formulated in such a way that though it is a refinement of Nash equilibrium, it uses sets of strategies: the equilibrium is persistent if it belongs to a persistent retract, which is a set of strategies. There are other, less demanding, set-valued notions of equilibrium that lead to the same result for the merger game if one adjusts their definitions in a way similar to the definition of persistent equilibrium.

A retract Θ is *closed under rational behavior (curb)* if $\forall i$, $\forall \sigma \in \Theta$, $\sigma' \in BR(\sigma_{-i}) \Longrightarrow (\sigma', \sigma_{-i}) \in \Theta$ (Basu and Weibull (1991)). A curb retract is minimal if it does not contain any smaller curb retract. Then we can define

Definition 6.7 *An equilibrium is* ***curb*** *if it belongs to a minimal curb retract.*

The definition of a curb retract requires that all best replies to strategies in it belong to it. Analogously with the reasoning in the case of persistent equilibrium one can see that in the case of retract $(\{12\}, \{12\}, x_3)$ with arbitrary x_3 all best replies to strategies in it belong to it. Thus this retract is curb, and, moreover, it is minimal curb since player 3 is indifferent among all his strategies. On the other hand, by definition the set of strategies $(\{13\}, x_2, \{13\})$ is contained in a curb retract only if the retract contains all strategies of player 2. But then it contains $\{12\}$ to which the unique best reply of player 1 is $\{12\}$. Thus, the retract has to contain $(\{12\}, \{12\}, x_3)$ together with $(\{13\}, x_2, \{13\})$. Then the retract is not minimal curb. A similar reasoning applies to other equilibria of the game. Thus, only equilibrium $(\{12\}, \{12\}, x_3)$ belongs to the minimal curb retract.

Remark 6.2 *A pure equilibrium of the merger game is curb only if in it the most preferred partners merger forms.*

Another set valued concept that also incorporates dynamic considerations, is *equilibrium evolutionarily stable (EES)* set, introduced by Swinkels (1992). Consider a non-empty closed subset Θ of Nash equilibria of a game. The set Θ is equilibrium evolutionarily stable if it is minimal with respect to the following condition: $\exists \varepsilon' > 0$ such that $\forall \varepsilon \in (0, \varepsilon')$, $\forall \sigma \in \Theta$, $\forall \sigma'$ it holds that $\sigma' \in BR((1-\varepsilon)\sigma + \varepsilon\sigma') \Longrightarrow (1-\varepsilon)\sigma + \varepsilon\sigma' \in \Theta$. Analogously with the above definitions

Definition 6.8 *An equilibrium is* ***EES*** *if it belongs to an EES set.*

The idea of equilibrium evolutionarily stable set is that it is robust against an invasion of small number of mutants that play a best response σ' against the population after the invasion that is represented by $BR((1-\varepsilon)\sigma + \varepsilon\sigma')$.

Consider again the set of Nash equilibria $\Theta = (\{12\}, \{12\}, x_3)$. If ε is small, $\{12\}$ continues to be the unique best reply for players 1 and 2. Since x_3 is arbitrary, $\forall \sigma \in \Theta$, Θ contains all small perturbations of σ that differ from σ only in the strategy of player 3. Thus, Θ satisfies condition for equilibrium evolutionary stability and it is also minimal with respect to it.

Consider now the set $\Theta' = (\{13\}, x_2, \{13\})$ with appropriate x_2 so that it is a subset of Nash equilibria. If it is equilibrium evolutionarily stable, by definition $\forall \sigma \in \Theta'$ it has to contain perturbations of σ that differ from σ in the strategy of player 2. But then it has to contain a strategy of player 2 that puts high probability on $\{12\}$. Then $\{12\}$ becomes a best reply for player 1. Thus, an equilibrium evolutionarily stable set containing Θ' has to contain also $\{12\}$ as strategies of players 1 and 2. But then it is not minimal since in the previous paragraph we found a minimal equilibrium evolutionarily stable set containing such combinations of strategies. Given that the reasoning applies to other equilibria of the merger game, once again we have

Remark 6.3 *A pure equilibrium of the merger game is EES only if in it the most preferred partners merger forms.*

All three concepts (persistency, curb, EES) we used select the equilibrium of the merger game with the same outcome. There are also other similar concepts, like cyclically stable set of Matsui (1992) that lead to the same results. Furthermore, Hurkens (1995) presents a learning process that converges to a minimal curb set, so in our game it converges to the equilibrium with the merger of the most preferred partners. Also, the best response dynamic with mutations (Kandori et al. (1993), Young (1993)) leads to the same result. All this comes as no surprise: all concepts use small perturbations in players' strategies to check stability of equilibria. Notice that in the selected equilibrium a merger always forms even when all firms prefer that their rivals merge. The potential free rider problem is overcome by equilibrium selection. The problem on which merger to coordinate is also overcome by possibility of small mistakes in players' strategies. The results are restricted to acyclic division rules; if a division rule is not acyclic, the refinements above do not have much bite and any of the profitable mergers can occur.

6.4 An Illustrative Example: Asymmetric Linear Cournot Triopoly

In this section we consider an example of the merger game in a linear Cournot oligopoly with asymmetric firms. Using the results of the previous section, we are able to answer the question of how asymmetries in efficiency influence the likelihood of each merger. Thus, we are able to say which of the three mergers is more likely.

There is an industry with three firms that differ in efficiency. That is, they have different marginal cost of production c_1, c_2, c_3. In what follows we take, without loss of generality, $c_1 = 0 < c_2 < c_3$. The demand side of the market is represented by the inverse demand function $P = 1 - Q$.

There are two stages in the game. The second stage is the usual Cournot oligopoly game while the first stage is the merger game. Since the second stage has a unique equilibrium, using subgame perfection enables us to reduce the game to the first stage by substituting the profits realized in the Cournot game into the first stage valuations of coalitions.

Formally, in the first stage each firm can propose, simultaneously with the other firms, a merger with another firm but not with both other firms. Monopoly is impossible either because of regulations or of high costs of negotiations with more than one partner. Of course, a firm cannot be driven into a merger, so it also has an opportunity to stay alone. Thus, each firm has three possibilities in the first stage: stay alone, propose to the more efficient firm, and propose to the less efficient firm.

If two firms propose a merger to each other, they merge; the efficiency level of the resulting firm is the maximum of the two, that is, the resulting firm has marginal cost equal to the minimum of the marginal costs of the merging firms. The firms in a merger retain their identities and each will get a certain share of the joint profit, in a way that is specified below. If no proposals coincide, no merger takes place and all three firms proceed to the second stage as separate entities.

A merger is understood as an arrangement where the identities of the firms are retained. One interpretation is that it is a temporary arrangement, say the less efficient firm places an order for production to the more efficient firm. Another interpretation is that shareholders of both merging firms should have a part of the profit of the resulting firm; one then can find out how much each shareholder gets compared with the situation before the merger.

After the first stage, there are either three (if no merger took place) or two firms in the industry. They play the usual Cournot game. With the restriction on the parameters (marginal costs) specified below, there is a unique equilibrium of the second stage game. We consider subgame perfect equilibria of the two stage game. Thus, the profits of the firms in the Cournot equilibrium of the second stage game will be understood as the payoffs after the first stage and the whole game.

If no merger has occurred, the profit of each firm in the second stage is well defined. Denote the profits in this case by π_1, π_2, π_3. If there was a merger, the merged firms i, j get some profit as one entity, which is denoted by π_{ij}. The remaining firm k gets profit π_k^{ij}, which in general is not equal to π_k. As the merged firms retain their identities they have to divide the profit in some way. We assume that there is an efficient and strictly individually rational rule φ_{ij} for the division of the payoff in this case that specifies the share $\varphi_{ij}(i)$ of player i.

6.4.1 Second Stage Payoffs

Let us analyze the second stage of the game. Three distinct situations are possible. If there was no merger, there are three firms in the industry with marginal costs $c_1 = 0 < c_2 < c_3$. If firms 1 and 2 have merged, there are two firms with marginal costs $c_1 = 0 < c_3$. If firms 1 and 3, or firms 2 and 3 have merged there are two firms with marginal costs $c_1 = 0 < c_2$. We analyze these three situations in turn.

Case $c_1 = 0 < c_2 < c_3$

Following the standard oligopoly theory, the equilibrium quantities produced by the firms are

$$q_1 = \frac{1+c_2+c_3}{4}, q_2 = \frac{1-3c_2+c_3}{4}, q_3 = \frac{1+c_2-3c_3}{4}.$$

Thus, the total quantity and price are

$$Q = \frac{3-c_2-c_3}{4}, P = \frac{1+c_2+c_3}{4}$$

which bring the profits

$$\pi_1 = (\frac{1+c_2+c_3}{4})^2, \pi_2 = (\frac{1-3c_2+c_3}{4})^2, \pi_3 = (\frac{1+c_2-3c_3}{4})^2.$$

The quantities and the price above are required to be positive to avoid corner solutions. Since $c_2 < c_3$, for positivity of all quantities and prices it is sufficient to have also $c_2 > 3c_3 - 1$. The region where both conditions hold is shown in Figure 6.1.

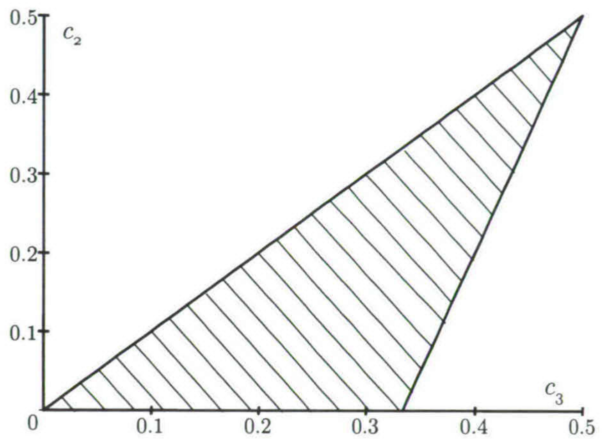

Figure 6.1: Region where quantities and prices are positive.

Case $c_1 = 0 < c_3$

This case occurs when firms 1 and 2 merge. There are two firms in the industry whose equilibrium quantities are

$q_{12} = \frac{1+c_3}{3}, q_3 = \frac{1-2c_3}{3},$

implying the total quantity and price

$Q = \frac{2-c_3}{3}, P = \frac{1+c_3}{3}.$

The profits are

$\pi_{12} = (\frac{1+c_3}{3})^2, \pi_3^{12} = (\frac{1-2c_3}{3})^2.$

Note that in the region depicted in Figure 6.1 the quantities and prices for this case are positive.

Case $c_1 = 0 < c_2$

This case arises when firms 1 and 3, or firms 2 and 3 merge. The case is completely equivalent to the case $c_1 = 0 < c_3$, where we substitute c_2 for c_3. Thus the profits of the firms are

$\pi_{13} = (\frac{1+c_2}{3})^2, \pi_2^{13} = (\frac{1-2c_2}{3})^2$, or

$\pi_1^{23} = (\frac{1+c_2}{3})^2, \pi_{23} = (\frac{1-2c_2}{3})^2.$

The region in Figure 6.1 is again sufficient for positive prices and quantities.

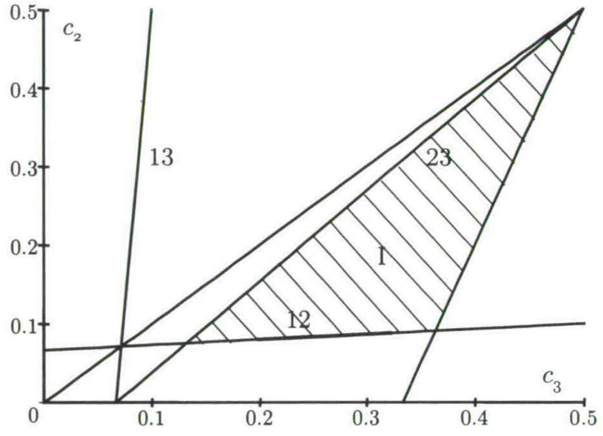

Figure 6.2: Regions where mergers are profitable.

6.4.2 Equilibria of the First Stage

Having calculated the payoffs of the second stage we can proceed to the first stage. Notice, first, that mergers are not always profitable. A merger can have as a consequence that the profit of the resulting firm is lower than the sum of the profits of the two merging firms, if there would have been no merger. Below we calculate the values of the parameters c_2 and c_3 for which mergers are profitable.

For example, the firm obtained as merger of firms 1 and 2 in the equilibrium of the second stage has profit $\pi_{12} = (\frac{1+c_3}{3})^2$. Separately, firms 1 and 2 get respectively $\pi_1 = (\frac{1+c_2+c_3}{4})^2, \pi_2 = (\frac{1-3c_2+c_3}{4})^2$. Therefore, the firms find merger profitable if $\pi_{12} > \pi_1 + \pi_2$. This inequality holds if $\frac{c_3+1}{15} < c_2 < \frac{c_3+1}{3}$ for $c_3 \in (0, 0.5)$. This region is marked 12 in Figure 6.2.

Analogously, the firm combining firms 1 and 3 gets in the second stage $\pi_{13} = (\frac{1+c_2}{3})^2$, while separately firms 1 and 3 get $\pi_1 = (\frac{1+c_2+c_3}{4})^2, \pi_3 = (\frac{1+c_2-3c_3}{4})^2$ respectively. $\pi_{13} > \pi_1 + \pi_3$ holds when $3c_3 - 1 < c_2 < 15c_3 - 1$ for $c_3 \in (0, 0.5)$. The region is marked 13 in Figure 6.2.

Finally, firms 2 and 3 find merger profitable if $\pi_{23} = (\frac{1-2c_2}{3})^2 > \pi_2 + \pi_3 = (\frac{1-3c_2+c_3}{4})^2 + (\frac{1+c_2-3c_3}{4})^2$. The inequality holds if $3c_3 - 1 < c_2 < \frac{15c_3-1}{13}$ for $c_3 \in (0, 0.5)$. This region is marked 23 in Figure 6.2.

Not every merger is profitable. If c_2 is very small, firms 1 and 2 would not like to merge. If both c_2 and c_3 are small, firms 1 and 3 would not like to merge. Finally, if c_2 is not very different from c_3, firms 2 and 3 will stay apart. The observations are in line with the result of Salant et al. (1983) that in a symmetric Cournot oligopoly a merger is unprofitable if less than 80% of the firms merge. The result also holds for small asymmetries; however, as the asymmetry between participants of the merger grows, the efficiency gain from the merger overcomes the negative effect of the merger.

Remark 6.4 *A merger is profitable only if the cost difference between its participants is large enough.*

We will focus our attention on the case when c_2 is not very small and c_3 is sufficiently different from c_2, since this is the case when all three mergers are profitable and the coordination problem is most acute. That is, we consider the intersection of regions 12, 13, and 23 in Figure 6.2, which is denoted by I. In this region all three firms prefer a merger to the situation when there is no merger. Following Theorem 6.1, all mergers are possible in equilibrium. The question is which one is more likely.

Though all firms prefer a merger to occur, they might have different preferences among possible mergers. For example, due to the positive externality of a merger in Cournot oligopoly it can happen that a firm prefers most that the other two firms merge rather than to participate itself in a merger. It is also possible that all three firms have such preferences. There is a potential free rider problem: if each firm prefers most that the other two merge, they all can try to stay alone. It is an equilibrium as Remark 6.1 indicates. But this equilibrium brings the worst possible outcome as all firms stay separate. For a strategic choice, as we have seen in the previous section, the preferences over the merger of the two other firms are irrelevant, as it cannot be influenced. The only preference relevant for a strategic choice is between two possible mergers is which the firm itself participates.

The firm evaluates these two mergers according to the profit division rule φ, which determines what merger the firm prefers more. Firm i prefers merger with j rather than with k, if $\varphi_i(ij) > \varphi_i(ik)$. In distinction from Barros (1998), who did not give justification for choosing one of the mergers, and Horn and Persson (1996), who used a cooperative concept, we apply the equilibrium selection results from the previous section to select an equilibrium.

As an example of a division rule, which seems one of the most natural ones, we consider equal division of gains from merger. Thus, $\varphi_i(ij) = \pi_i + \frac{\pi_{ij} - (\pi_i + \pi_j)}{2}$. This division arises, for example, from the Nash bargaining solution. The Nash solution can also be implemented

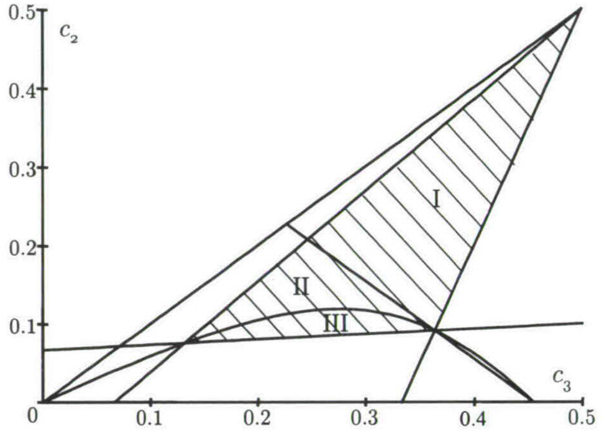

Figure 6.3: Pattern of firms' preferences over mergers.

via a non-cooperative game. It can be justified by a game where the merging firms bargain over the joint profit after they had committed not to merge with the remaining firm.

With such division rule, in the Cournot model above, one can find what firm prefers what merger. The rule is acyclic; the pattern of firms' preferences is depicted in Figure 6.3.

Firm 3 always prefers to merge with firm 1 rather than with firm 2. In region I firm 1 prefers firm 2 and firm 2 prefers firm 1; in region II firm 1 prefers firm 3 and firm 2 prefers firm 1; in region III both firms 1 and 2 prefer firm 3.

Applying Theorem 6.2, we immediately arrive at the following result.

Theorem 6.3 *In region I merger {1,2} forms in the selected equilibrium; in regions II and III merger {1,3} forms in the selected equilibrium.*

Loosely speaking, the merger that is more likely to form depends on the difference in cost parameters. If c_2 and c_3 are high enough, the most efficient firm 1 merges with firm 2, while if the cost differences are not that high, it merges with firm 3. A merger of the two least efficient firms is never selected: the efficient firm has always more to offer to attract a partner.

Since we have assumed that the division rule always gives each player the same proportion (half) of the gains from merging, it is no surprise that the merger in the selected equilibrium is the one with the highest internal gain, where internal gain is the difference between the profit of the merged firm and the sum of the profits of the separate firms. Barros (1998), though using the highest internal gains criterion, does not give any justification of it. Horn and Persson (1996) select a merger which maximizes industry profit since they ignored the free rider problem of the outsider in a merger. Thus, the result of merger selection corresponds to Barros (1998) and differs from Horn and Persson (1996). The next subsection shows that in our model the selected merger does not always maximize industry profit.

Note that the result depends much on the division rule. The division rule we consider divides the gains equally. If the division rule gives all the bargaining power in a merger to the buyer (more efficient firm) then merger of firms 2 and 3 is selected since firm 1 prefers to merge with firm 2, firm 2 now prefers to have some profit from merging with less efficient firm 3 and firm 3 is indifferent between mergers.

6.4.3 Efficiency of the Selected Equilibrium

We analyze efficiency from two points of view: from the point of view of the firms and from the point of view of a regulator whose objective is social welfare.

We call an equilibrium *profit efficient* if the sum of the profits of the firms in the equilibrium is highest among all equilibria. Since the equilibrium selected maximizes internal gains, one might think that it is also profit efficient. This is not true since the merging firms do not take into account the effect of the merger on the outsider. We show that there is a region where a profit inefficient merger forms.

The sum of the firms' profits when firms 1 and 2 merge is equal to $\frac{5c_3^2-2c_3+2}{9}$, while when firms 1 and 3 merge the aggregate profit is $\frac{5c_2^2-2c_2+2}{9}$. One can check that the profit in the first case is larger than in the second case when $c_3 < c_2 < \frac{2}{5} - c_3$ for $c_3 < \frac{1}{5}$ and $\frac{2}{5} - c_3 < c_2 < c_3$ for $c_3 > \frac{1}{5}$. The line $c_2 = \frac{2}{5} - c_3$ does not coincide with the line separating regions where the corresponding mergers are selected. There is a region where the selected merger is between firms 1 and 3 while a merger between firms 1 and 2 would have given a higher aggregate profit. The region where a profit inefficient merger forms is shown on Figure 6.4.

In region I a profit inefficient merger takes place. It happens because firm 3 would have profited a lot from merger $\{12\}$ driving the sum of profits up; given the division rule, however, firm 1 prefers to merge with firm 3 and firm 2 does not benefit that much from this merger.

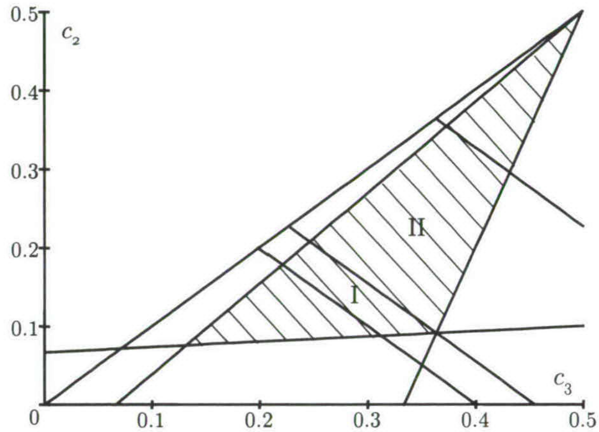

Figure 6.4: Regions of merger (in)efficiency.

If redistribution of profits was allowed, the firms could achieve a better result. This example shows that the equilibrium selected needs not be the one where the producer surplus is maximized.

Welfare consists of the sum of consumer surplus and profits. We call an equilibrium *welfare efficient* if the sum of consumer surplus and profits in the equilibrium is highest among all equilibria. When the sum of profits is higher one can expect that consumer surplus will be lower since higher profit means higher price. Therefore, the effect on welfare is a priori ambiguous.

When firms 1 and 2 merge the price is $\frac{1+c_3}{3}$, thus the consumer surplus is $\frac{c_3^2-4c_3+4}{18}$. When firms 1 and 3 merge the price is $\frac{1+c_2}{3}$, resulting in a consumer surplus of $\frac{c_2^2-4c_2+4}{18}$. Thus welfare in the first case is $\frac{11c_3^2-8c_3+8}{18}$ while in the second $\frac{11c_2^2-8c_2+8}{18}$. Comparing the two expressions we have that welfare in the first case will be higher only when $c_2 > \frac{8}{11} - c_3$. Figure 6.4 shows the region (denoted by II) where the selected merger is not welfare efficient. Note, however, that in the other two regions the equilibrium merger is welfare efficient and the coordination problem helps to improve efficiency in region I compared with the merger that maximizes producer surplus.

If the cost difference between the most efficient firm and the other two is large (c_2 and c_3 large) then it is better for welfare that firms 1 and 2 merge, thus leaving the least efficient firm on the market. Notice, however, that in this case the difference between

c_2 and c_3 is not large, thus eliminating firm 3 instead of firm 2 does not bring much in terms of efficiency. As the difference between the most efficient firm 1 and the other two firms decreases, it becomes better for welfare to eliminate the least efficient firm 3 but the firms prefer to eliminate firm 2 instead. After some point, the selected equilibrium of the merger game indeed eliminates firm 3. Eventually, if the difference decreases further, mergers become unprofitable, which is also good for welfare, since then a merger would hamper competition without bringing much in terms of efficiency.

6.5 Conclusion

We have considered a model of the merger game. The attention was restricted to three-player games; they give enough insight into the problem while avoiding the difficulties of a game with more players. The merger game we constructed has multiple equilibria; the coordination problem of which equilibrium to select was resolved by refinements that have a dynamic flavor and all led to the same result. The resulting merger is the one where both participants prefer this merger to other mergers. In the Cournot oligopoly example with a particular division rule we found that the resulting merger maximizes internal gains. There are cases when the selected equilibrium is both profit and welfare efficient. There are also cases when the equilibrium is not welfare efficient but profit efficient. Finally, it can happen that the equilibrium is welfare efficient but not profit efficient. Thus, coordination problem in our merger game gives a better outcome for welfare than in the case where firms can collude to get maximal profit.

Certainly, there are ways to extend the model to endogenous division of gains, more players and sequential merger decisions, though there appear to be many difficulties (for an attempt, see Gowrisankaran (1999)). However, it is possible to address the question of how asymmetries affect mergers with just three players. Thus, the simple merger game and its illustration with a linear triopoly model considered in this chapter provides already an insight of what may happen in an endogenous merger game and how asymmetries affect possible mergers.

Bibliography

Alchian A. (1950) "Uncertainty, Evolution, and Economic Theory", *Journal of Political Economy,* 58, 211-221.

Arthur W.B. (1993) "On Designing Economic Agents that Behave Like Human Agents", *Journal of Evolutionary Economics,* 3, 1-22.

Banerjee A. (1992) "A Simple Model of Herd Behavior", *Quarterly Journal of Economics,* 107(3), 797-817.

Barros P. (1998) "Endogenous Mergers and Size Asymmetry of Merger Participants", *Economics Letters* 60,113-119.

Basu K.,Weibull J. (1991) "Strategy Subsets Closed under Rational Behavior", *Economics Letters,* 36, 141-146.

Bester H.,Güth W. (1998) "Is Altruism Evolutionarily Stable?", *Journal of Economic Behavior and Organization*, 34(2), 193-209.

Bhattacharya R.,Waymire E. (1990) *Stochastic Processes with Applications,* Wiley.

Binmore K.,Swierzbinski J.,Proulx C. (1996) "Does Minimax Work? An Experimental Study", mimeo, University College London.

Bloch F. (1997) "Noncooperative Models of Coalition Formation in Games with Spillovers", in: Carraro C., Siniscalco D.(eds) *New directions in the economic theory of the environment*, Cambridge University Press.

Bolle F. (1991) "On Love and Altruism", *Rationality and Society, 3(2), 197-214.*

Börgers T.,Sarin R. (1996) "Naive Reinforcement Learning with Endogenous Aspiration", mimeo, University College London.

Brennan G.,Güth W.,Kliemt H. (1997) "Trust in the Shadow of the Courts", CentER DP 9789, Tilburg University.

Bulow J.,Geanakoplos J.,Klemperer P. (1985) "Multimarket Oligopoly: Strategic Substitutes and Complements", *Journal of Political Economy,* 93(5), 488-511.

Bush R.,Mosteller F. (1955) *Stochastic Models for Learning,* Wiley.

Cooper B. (1996) "Copying Fidelity in the Evolution of Finite Automata that Play the Repeated Prisoners' Dilemma", mimeo, Oxford University.

Damme E.,van (1991) *Stability and Perfection of Nash Equilibria,* Springer-Verlag.

Dufwenberg M.,Güth W. (1999) "Indirect Evolution versus Strategic Delegation: A Comparison of two Approaches to Explaining Economic Institutions", *European Journal of Political Economy,* 15, 281-295.

Ellison G. (1993) "Learning, Local Interaction, and Coordination", *Econometrica,* 61(5), 1047-1071.

Ellison G.,Fudenberg D. (1993) "Rules of Thumb for Social Learning", *Journal of Political Economy,* 101(4), 612-643.

Ellison G.,Fudenberg D. (1995) "Word-of-Mouth Communication and Social Learning", *Quarterly Journal of Economics,* 110(1), 93-126.

Ely J.,Yilankaya O. (1997) "Nash Equilibrium and the Evolution of Preferences", DP No. 1191, Northwestern University.

Farrell J.,Shapiro C. (1990) "Horizontal Mergers: An Equilibrium Analysis", *American Economic Review,* 80, 107-126.

Fridolfsson S.-O.,Stennek J. (1999) "Why Mergers Reduce Profits, and Raise Share Prices", mimeo, Stockholm University.

Friedman M. (1953) "The Methodology of Positive Economics", in: *Essays in Positive Economics,* University of Chicago Press.

Fudenberg D.,Levine D. (1998) *The Theory of Learning in Games,* MIT Press.

Gale D.,Rosenthal R. (1999) "Experimentation, Imitation, and Stochastic Stability", *Journal of Economic Theory,* 84(1), 1-40.

Gale J.,Binmore K.,Samuelson L. (1995) "Learning to Be Imperfect: The Ultimatum Game", *Games and Economic Behavior,* 8, 56-90.

Gowrisankaran G. (1999) "A Dynamic Model of Endogenous Horizontal Mergers", *RAND Journal of Economics,* 30(1), 56-83.

Güth W.,Peleg B. (1997) "When Will the Fittest Survive? - An Indirect Evolutionary Analysis", CentER DP No. 9768, Tilburg University.

Güth W.,Yaari M. (1992) "An Evolutionary Approach to Explain Reciprocal Behavior in a Simple Strategic Game", in: Witt U. (ed): *Explaining Process and Change - Approaches to Evolutionary Economics*, The University of Michigan Press.

Harsanyi J.,Selten R. (1988) *A General Theory of Equilibrium Selection in Games,* MIT Press.

Hart S.,Kurz M. (1983) "Endogenous Formation of Coalitions", *Econometrica,* 51, 1047-1064.

Horn H.,Persson L. (1996) "Endogenous Mergers in Concentrated Markets", CEPR discussion paper.

Huck S.,Normann H.-T.,Oechssler J. (1999) "Learning in Cournot Oligopoly - An Experiment", *Economic Journal,* 109, C80-C95.

Huck S.,Oechssler J. (1999) "The Indirect Evolutionary Approach to Explaining Fair Allocation", *Games and Economic Behavior,* 28, 13-24.

Hurkens S. (1995) "Learning by Forgetful Players", *Games and Economic Behavior,* 11, 304-329.

Kalai E.,Samet D. (1984) "Persistent Equilibria in Strategic Games", *International Journal of Game Theory,* 13, 129-144.

Kamien M.,Zang I. (1990) "The Limits of Monopolization through Acquisition", *Quarterly Journal of Economics,* 105, 465-499.

Kamien M.,Zang I. (1991) "Competitively Cost Advantageous Mergers and Monopolization", *Games and Economic Behavior,* 3, 323-338.

Kandori M.,Mailath G.,Rob R. (1993) "Learning, Mutation, and Long Run Equilibria in Games", *Econometrica,* 61, 29-56.

Karandikar R.,Mookherjee D.,Ray D.,Vega-Redondo F. (1998) "Evolving Aspirations and Cooperation", *Journal of Economic Theory,* 80(2), 292-331.

Koçkesen L.,Ok E.,Sethi R. (1999) "The Strategic Advantage of Negatively Interdependent Preferences", *Journal of Economic Theory,* forthcoming.

Kreps D.,Scheinkman J. (1983) "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes", *Bell Journal of Economics,* 14, 326-337.

Laslier J.F.,Topol R.,Walliser B. (1999) "A Behavioral Learning Process in Games", University of Cerge-Pontoise, Working Paper 99-03.

Mas-Colell A.,Whinston M.,Green J. (1995) *Microeconomic Theory,* Oxford University Press.

Matsui A. (1992) "Best Response Dynamics and Socially Stable Strategies", *Journal of Economic Theory,* 57, 343-362.

Narendra K.,Thathachar M.A.L. (1989) *Learning Automata: An Introduction,* Prentice Hall International.

Ok E.,Vega-Redondo F. (1999) "On the Evolution of Individualistic Preferences: Complete versus Incomplete Information Scenarios", mimeo, New York University and University of Alicante.

Palomino F. (1996) "On the Survival of Strictly Dominated Strategies in Large Populations", mimeo, Tilburg University.

Perry M.,Porter R. (1985) "Oligopoly and the Incentive for Horizontal Merger", *American Economic Review,* 75, 219-227.

Posch M. (1997) "Cycling in a Stochastic Learning Algorithm for Normal Form Games", *Journal of Evolutionary Economics,* 7, 193-207.

Possajennikov A. (1997) "An Analysis of a Simple Reinforcement Dynamics: Learning to Play an 'Egalitarian' Equilibrium", CentER Discussion Paper 9719.

Possajennikov A. (1999a) "Optimality of Imitative Behavior in Cournot Oligopoly", CentER DP 9902, Tilburg University.

Possajennikov A. (1999b) "On Evolutionary Stability of Spiteful Preferences", CentER DP 9956. Tilburg University, forthcoming in *Journal of Economic Behavior and Organization.*

Rabin M. (1993) "Incorporating Fairness into Game Theory and Economics", *American Economic Review,* 83(5), 1281-1302.

Rajan R. (1989) "Endogenous Coalition Formation in Cooperative Oligopolies", *International Economic Review,* 30(4), 863-876.

Rhode P.,Stegeman M. (1997) "Non-Nash Equilibria of Darwinian Dynamics with Applications to Duopoly", mimeo, University of North Carolina and Virginia Tech.

Robson A. (1990) "Efficiency in Evolutionary Games: Darwin, Nash and the Secret Handshake", *Journal of Theoretical Biology,* 144, 379-396.

Robson A.,Vega-Redondo F. (1996) "Efficient Equilibrium Selection in Evolutionary Games with Random Matching", *Journal of Economic Theory,* 70(1), 65-92.

Roth A.,Erev I. (1995) "Learning in Extensive-Form Games: Experimental Data and Simple Dynamics Models in the Intermediate Run", *Games and Economic Behavior,* 8, 164-212.

Ruebeck C. (1999) "Imitation Dynamics in the Repeated Prisoners' Dilemma: An Exploratory Example", *Journal of Economic Behavior and Organization,* 40(1), 81-104.

Salant S.,Switzer S.,Reynolds R. (1983) "Losses from Horizontal Merger: the Effect of an Exogenous Change in Industry Structure on Cournot-Nash Equilibrium", *Quarterly Journal of Economics*, 98, 185-199.

Samuelson L. (1997) *Evolutionary Games and Equilibrium Selection,* MIT Press.

Schaffer M. (1989) "Are Profit-maximizers the Best Survivors?", *Journal of Economic Behavior and Organization,* 12, 29-45.

Schlag K. (1998) "Why Imitate and if so, How?", *Journal of Economic Theory,* 78(1), 130-156.

Swinkels J. (1992) "Evolutionary Stability with Equilibrium Entrants", *Journal of Economic Theory,* 52, 306-332.

Vega-Redondo F. (1996) *Evolution, Games, and Economic Behavior,* Oxford University Press.

Vega-Redondo F. (1997) "The Evolution of Walrasian Behavior", *Econometrica,* 65(2), 375-384.

Weibull J. (1995) *Evolutionary Game Theory,* MIT Press.

Young H.P. (1993) "The Evolution of Conventions", *Econometrica,* 61, 57-84.

Young H.P. (1998) *Individual Strategy and Social Structure,* Princeton University Press.

Samenvatting

Stel je voor dat je een keuze tussen vele acties moet maken. De acties zouden een uitbetaling brengen, maar je weet niet precies hoeveel elke actie brengt. Wat zou je doen?

Natuurlijk, is de situatie erboven niet compleet. Er kan veel meer informatie zijn om de keuze te maken. Bijvoorbeeld, je weet dat er ook anderen zijn die in dezelfde situatie zich bevinden. Dan kan je van ze leren. Of je weet dat de acties van deze anderen invloed op je uitbetaling hebben. Er zijn veel mogelijkheden om de situatie te modelleren en sommige worden in dit proefschrift geanalyseerd. De structuur uit de eerste alinea vormt de basis voor de modellen.

Er zijn veel voorbeelden van de bovenbeschrevene situaties. Denk aan het kopen van een yoghurt: zou je hem lekker vinden? zou hij goed voor je maag zijn? Voor een interactieve situatie, stel je voor dat je een kruising op de fiets nadert (typisch voor Nederland): zou de auto die uit de andere richting komt, zelf vóór gaan of zou hij je vóór laten gaan? Deze zijn voorbeelden van (heel simpele) problemen die erboven zijn omgeschreven. Dit proefschrift is een verzameling van een aantal modellen die dergelijke situaties van een bepaald uitzichtpunt bekijken: met beperkte informatie en groeiende ervaring, wat kan en wat moet men doen?

De klassieke speltheorie geeft een antwoord op deze vraag, maar deze theorie veronderstelt ook dat men alles weet wat er te weten is over de situatie: de preferenties van de andere spelers, de relevante waarschijnlijkheden enz. Dat is niet wat in de eerste alinea werd bedoeld: daar weet men slechts welke acties er zijn en dat ze een uitbetaling opleveren. Dus moet een andere methode worden gebruikt. Toch blijft de oplossing uit de klassieke speltheorie (het evenwicht van Nash) een nuttige referentiepunt waartegen we onze modellen kunnen meten.

De tweede alinea boven beschrijft hoe men informatie over de situatie kan krijgen: meestal door ervaring, eigen of die van anderen. Om het te modelleren, moet een dynamische methode worden gebruikt. De essentie van deze methode is dat de ervaring voor toekomstige beslissingen kan worden gebruikt. De klassieke speltheorie heeft ook de herhaalde situaties geanalyseerd maar, als het boven al gezegd is, moet men dan te veel weten om op een beslissing te komen. In dit proefschrift worden modellen met beperkte rationaliteit beschouwd, waarin niet alle relevante informatie wordt gebruikt.

Om het model compleet te maken, moet een regel worden gespecificeerd hoe ervaring wordt gebruikt voor toekomstige beslissingen. Er zijn veel mogelijkheden; sommige worden in dit proefschrift geanalyseerd. Als een regel is gekozen, verandert het gebruik van acties met de tijd volgens deze regel. Welke acties worden dan uitendelijk genomen? Wat is de relatie van deze acties met de oplossing die klassieke speltheorie biedt? Zulke vragen probeert dit proefschrift te beantwoorden.

De dynamische modellen uit de bovenstaande alinea's kunnen in twee grote groepen worden verdeeld. De eerste groep bevat modellen die gebaseerd zijn op het biologische idee van evolutie, van "*survival of the fittest*". In de tweede groep bevinden zich modellen uit de psychologische wetenschap, die gebaseerd zijn op het idee van leren. Dit proefschrift is getiteld "Leren en Evolutie in Spellen en Oligopolistische Modellen". Zoals uit de titel blijkt, beschouwt het proefschrift beide groepen. Hoofdstuk 1 van het proefschrift geeft een inleiding tot dit type van modellen en definieert ook noodzakelijke begrippen.

De modellen die uit leren komen worden beschouwd in hoofdstukken 2 en 3. Hoofdstuk 2 beschouwt zogenaamde "versterking leren" (*reinforcement learning*). In dit model krijgt de actie die wordt genomen een versterking als er een positieve uitbetaling wordt gebracht door de actie. De versterking betekent dat deze actie in volgende perioden met hogere waarschijnlijkheid wordt genomen. De vraag is, zou men de beste (in de zin van verwachte uitbetaling) actie vinden wanneer elke actie versterking krijgt? In eenspelers-besluitvorming bestaat een *trade-off* tussen het vinden van (convergentie tot) de beste actie en de snelheid van leren. Door simulaties met de computer wordt een modificatie van het grondmodel gekozen, die deze trade-off het best aanpakt. Het model met deze modificatie wordt dan toegepast voor meerpersoonspellen. Er wordt getoond dat het evenwicht dat meer centraal is, dat is intussen andere evenwichten ligt, meer kansen heeft te worden gekozen door het dynamische proces van versterking leren. In de beschouwde spellen is dit evenwicht ook egalitair. Er is een verband tussen het egalitaire evenwicht en het begrip van risico-dominatie; dit verband wordt door een paar voorbeelden geïllustreerd.

Hoofdstuk 3 beschouwt een ander proces, namelijk imitatie, in een spel van Cournot-oligopolie type. Wanneer een speler de uitbetaling van een andere speler ziet, en deze uitbetaling hoger is dat de uitbetaling van de eerste speler, heeft de eerste speler de verleiding de tweede speler te navolgen. Maar als de spelers elkaars strategieën imiteren in een kleine populatie die een Cournot-oligopolie speelt, dan landen ze in een heel inefficiënte (voor ze) staat. Het hoofdstuk beschouwt twee modificaties van de simpele "imiteer de beste" (*imitate the best*) regel. De eerste modificatie is onvolmaakte imitatie waarin de spelers slechts een deel van de strategieën van de anderen kunnen zien en dus slechts dit deel kunnen navolgen. De andere modificatie veronderstelt dat interactie en imitatie in verschillende populaties plaatsvinden. Beide modificaties leiden tot een betere uitkomst voor de spelers; de combinatie van de strategieën die uitendelijk wordt gespeeld oplevert een hogere uitbetaling. Het is mogelijk dat het evenwicht (in de zin van de klassieke speltheorie) van het spel uitendelijk wordt gespeeld.

Hoofdstukken 2 en 3 van het proefschrift beschouwen lerenmodellen; hoofdstukken 4 en 5 nemen evolutie als de basis. Deze hoofdstukken proberen de evolutie van preferenties te modelleren. De evolutie werkt niet direct maar heeft invloed op de compositie van de populatie door uitbetalingen in evenwicht. Spelers ontmoeten elkaar in paren en spelen een tweepersoonspel; ze hebben bepaalde preferenties en met deze preferenties zijn ze in staat een evenwicht te spelen. De evolutie is minder rationeel en kiest kortzichtig voor vermenigvuldiging van de spelers (preferenties) die gemiddeld de hoogste uitbetalingen hebben tegen de huidige populatie. Alleen de situaties waarin oorspronkelijk een type van preferenties in de populatie aanwezig is worden geanalyseerd. Er wordt verondersteld dat een kleine hoeveelheid mutanten met een andere type van preferenties verschijnt. Als de mutanten verdwijnen of niet meer groeien, is de originele situatie stabiel tegen zulke mutaties. Het model kan niet alleen voor biologische evolutie worden toegepast maar ook voor culturele evolutie van waarden.

Hoofdstuk 4 analyseert spellen dat uit oligopolistische modellen komen. Het hoofdstuk beschouwt wanneer altruïstische preferenties (je voelt je goed voor de hoge uitbetaling van de ander) en wanneer nijdige preferenties (je voelt je slecht voor de hoge uitbetaling van de ander) stabiel zijn. In een Bertrand-type oligopolie zijn de altruïstische preferenties stabiel, terwijl in een Cournot-type oligopolie zijn de nijdige preferenties stabiel. Wanneer informatie niet volledig is (de preferenties van de anderen zijn onbekend), dan worden de "normale" egoïstische preferenties stabiel.

Hoofdstuk 5 zet deze analyse voort voor 2×2 tweepersoonspellen. Het hoofdstuk beschouwt alle mogelijke (von Neumann-Morgenstern) preferenties. In sommige spelen, namelijk in de gevangenes dilemma en in lafaard(*chicken*)-type spellen (ook soms valkduif (*hawk-dove*) genoemd), zijn de egoïstische preferenties niet stabiel, terwijl in andere spellen, zoals coördinatieproblemen, zijn ze stabiel. Het model wordt uitgebreid tot onvolmaakte coördinatie, waarin coördinatie kostbaar is, en tot onvolledige informatie over preferenties van de anderen. Met onvolmaakte coördinatie verdwijnt de stabiliteit van egoïstische preferenties en worden in coördinatieproblemen simpele preferenties, die één strategie als dominante beschouwen, stabiel. Met onvolledige informatie worden de egoïstische preferenties meer stabiel in bijna alle types van spellen en wordt uitendelijk het evenwicht van het spel gespeeld.

Hoofdstuk 6 is wat verschillend van de andere hoofdstukken. Daarin in plaats van een expliciet dynamische proces worden de technieken voor de selectie van evenwichten gebruikt die op impliciete dynamische processen baseren. Zo wordt bijvoorbeeld het beste-antwoord leren gebruikt. Als het beste-antwoord proces het spel terug naar het evenwicht (of naar de verzameling van evenwichten) trekt na een kleine verstoring, dan wordt dit evenwicht geselecteerd. Deze technieken worden toegepast voor een fusiespel. Het fusiespel wordt gemodelleerd als een niet-coöperative spel van formatie van coalities. De selectiestechnieken zijn in principe voor iedere driespelersfusiespel van toepassing. Het werk van de technieken wordt door een voorbeeld van Cournot-oligopolie gedemonstreerd. In het evenwicht dat wordt geselecteerd, vormt zich de fusie, die beide deelnemers beter dan de andere fusies vinden. De asymmetrische Cournot-triopolie wordt geanalyseerd met de vraag welke invloed het verschil tussen de bedrijven op de geselecteerde fusie heeft. Wanneer het verschil tussen de bedrijven groot is, vindt de fusie van het meest efficiënt en het minst efficiënt bedrijven plaats, en wanneer het verschil niet groot is, dan vormen de twee meest efficiënte bedrijven de fusie.

Center for Economic Research, Tilburg University, The Netherlands
Dissertation Series

No.	Author	Title	Published
1	P.J.J. Herings	Static and Dynamic Aspects of General Disequilibrium Theory; ISBN 90 5668 001 3	June 1995
2*	Erwin van der Krabben	Urban Dynamics: A Real Estate Perspective - An institutional analysis of the production of the built environment; ISBN 90 5170 390 2	August 1995
3	Arjan Lejour	Integrating or Desintegrating Welfare States? - a qualitative study to the consequences of economic integration on social insurance; ISBN 90 5668 003 X	Sept. 1995
4	Bas J.M. Werker	Statistical Methods in Financial Econometrics; ISBN 90 5668 002 1	Sept. 1995
5	Rudy Douven	Policy Coordination and Convergence in the EU; ISBN 90 5668 004 8	Sept. 1995
6	Arie J.T.M. Weeren	Coordination in Hierarchical Control; ISBN 90 5668 006 4	Sept. 1995
7	Herbert Hamers	Sequencing and Delivery Situations: a Game Theoretic Approach; ISBN 90 5668 005 6	Sept. 1995
8	Annemarie ter Veer	Strategic Decision Making in Politics; ISBN 90 5668 007 2	October 1995
9	Zaifu Yang	Simplicial Fixed Point Algorithms and Applications; ISBN 90 5668 008 0	January 1996
10	William Verkooijen	Neural Networks in Economic Modelling - An Empirical Study; ISBN 90 5668 010 2	February 1996
11	Henny Romijn	Acquisition of Technological Capability in Small Firms in Developing Countries; ISBN 90 5668 009 9	March 1996
12	W.B. van den Hout	The Power-Series Algorithm - A Numerical Approach to Markov Processes; ISBN 90 5668 011 0	March 1996
13	Paul W.J. de Bijl	Essays in Industrial Organization and Management Strategy; ISBN 90 5668 012 9	April 1996

No.	Author	Title	Published
14	Martijn van de Ven	Intergenerational Redistribution in Representative Democracies; ISBN 90 5668 013 7	May 1996
15	Eline van der Heijden	Altruism, Fairness and Public Pensions: An Investigation of Survey and Experimental Data; ISBN 90 5668 014 5	May 1996
16	H.M. Webers	Competition in Spatial Location Models; ISBN 90 5668 015 3	June 1996
17	Jan Bouckaert	Essays in Competition with Product Differentiation and Bargaining in Markets; ISBN 90 5668 016 1	June 1996
18	Zafar Iqbal	Three-Gap Analysis of Structural Adjustment in Pakistan; ISBN 90 5668 017 X	Sept. 1996
19	Jimmy Miller	A Treatise on Labour: A Matching-Model Analysis of Labour-Market Programmes; ISBN 90 5668 018 8	Sept. 1996
20	Edwin van Dam	Graphs with Few Eigenvalues - An interplay between combinatorics and algebra; ISBN 90 5668 019 6	October 1996
21	Henk Oosterhout	Takeover Barriers: the good, the bad, and the ugly; ISBN 90 5668 020 X	Nov. 1996
22	Jan Lemmen	Financial Integration in the European Union: Measurement and Determination; ISBN 90 5668 021 8	December 1996
23	Chris van Raalte	Market Formation and Market Selection; ISBN 90 5668 022 6	December 1996
24	Bas van Aarle	Essays on Monetary and Fiscal Policy Interaction: Applications to EMU and Eastern Europe; ISBN 90 5668 023 4	December 1996
25	Francis Y. Kumah	Common Stochastic Trends and Policy Shocks in the Open Economy: Empirical Essays in International Finance and Monetary Policy; ISBN 90 5668 024 2	May 1997
26	Erik Canton	Economic Growth and Business Cycles; ISBN 90 5668 025 0	Sept. 1997

No.	Author	Title	Published
27	Jeroen Hoppenbrouwers	Conceptual Modeling and the Lexicon; ISBN 90 5668 027 7	October 1997
28	Paul Smit	Numerical Analysis of Eigenvalue Algorithms Based on Subspace Iterations; ISBN 90 5668 026 9	October 1997
29	Uri Gneezy	Essays in Behavioral Economics; ISBN 90 5668 028 5	October 1997
30	Erwin Charlier	Limited Dependent Variable Models for Panel Data; ISBN 90 5668 029 3	Nov. 1997
31	Rob Euwals	Empirical Studies on Individual Labour Market Behaviour; ISBN 90 5668 030 7	December 1997
32	Anurag N. Banerjee	The Sensitivity of Estimates, Inferences, and Forecasts of Linear Models; ISBN 90 5668 031 5	December 1997
33	Frans A. de Roon	Essays on Testing for Spanning and on Modeling Futures Risk Premia; ISBN 90 5668 032 3	December 1997
34	Xiangzhu Han	Product Differentiation, Collusion and Standardization; ISBN 90 5668 033 1	January 1998
35	Marcel Das	On Income Expectations and Other Subjective Data: A Micro-Econometric Analysis; ISBN 90 5668 034 X	January 1998
36	Jeroen Suijs	Cooperative Decision Making in a Stochastic Environment; ISBN 90 5668 035 8	March 1998
37	Talitha Feenstra	Environmental Policy Instruments and International Rivalry: A Dynamic Analysis; ISBN 90 5668 036 6	May 1998
38	Jan Bouwens	The Use of Management Accounting Systems in Functionally Differentiated Organizations; ISBN 90 5668 037 4	June 1998
39	Stefan Hochguertel	Households' Portfolio Choices; ISBN 90 5668 038 2	June 1998
40	Henk van Houtum	The Development of Cross-Border Economic Relations; ISBN 90 5668 039 0	July 1998
41	Jorg Jansen	Service and Inventory Models subject to a Delay-Limit; ISBN 90 5668 040 4	Sept. 1998

No.	Author	Title	Published
42	F.B.S.L.P. Janssen	Inventory Management Systems: control and information issues; ISBN 90 5668 041 2	Sept.1998
43	Henri L.F. de Groot	Economic Growth, Sectoral Structure and Unemployment; ISBN 90 5668 042 0	October 1998
44	Jenke R. ter Horst	Longitudinal Analysis of Mutual Fund Performance; ISBN 90 5668 043 9	Nov. 1998
45	Marco Hoeberichts	The Design of Monetary Institutions; ISBN 90 5668 044 7	December 1998
46	Adriaan Kalwij	Household Consumption, Female Employment and Fertility Decisions: A microeconometric analysis; ISBN 90 5668 045 5	February 1999
47	Ursula Glunk	Realizing High Performance on Multiple Stakeholder Domains: A Resource-based Analysis of Professional Service Firms in the Netherlands and Germany; ISBN 90 5668 046 3	April 1999

No.	Author	Title	Published
48	Freek Vermeulen	Shifting Ground: Studies on the Intersection of Organizational Expansion, Internationalization, and Learning; ISBN 90 5668 047 1	February 1999
49	Haoran Pan	Competitive Pressures on Income Distribution in China; ISBN 90 5668 048 X	March 1999
50	Laurence van Lent	Incomplete Contracting Theory in Empirical Accounting Research; ISBN 90 5668 049 8	April 1999
51	Rob Aalbers	On the Implications of Thresholds for Economic Science and Environmental Policy; ISBN 90 5668 050 1	May 1999
52	Abe de Jong	An Empirical Analysis of Capital Structure Decisions in Dutch Firms; ISBN 90 5668 051 X	June 1999
53	Trea Aldershof	Female Labor Supply and Housing Decisions; ISBN 90 5668 052 8	June 1999
54	Jan Fidrmuc	The Political Economy of Reforms in Central and Eastern Europe; ISBN 90 5668 053 6	June 1999
55	Michael Kosfeld	Individual Decision-Making and Social Interaction; ISBN 90 5668 054 4	June 1999

No.	Author	Title	Published
56	Aldo de Moor	Empowering Communities - A method for the legitimate user-driven specification of network information systems; ISBN 90 5668 055 2	October 1999
57	Yohane A. Khamfula	Essays on Exchange Rate Policy in Developing Countries; ISBN 90 5668 056 0	Sept. 1999
58	Maurice Koster	Cost Sharing in Production Situations and Network Exploitation; ISBN 90 5668 057 9	December 1999
59	Miguel Rosellon Cifuentes	Essays on Financial Policy, Liquidation Values and Product Markets; ISBN 90 5668 058 7	December 1999
60	Sharon Schalk	Equilibrium Theory: A salient approach; ISBN 90 5668 059 5	December 1999
61	Mark Voorneveld	Potential Games and Interactive Decisions with Multiple Criteria; ISBN 90 5668 060 9	December 1999
62	Edward Droste	Adaptive Behavior in Economic and Social Environments; ISBN 90 5668 061 7	December 1999
63	Jos Jansen	Essays on Incentives in Regulation and Innovation; ISBN 90 5668 062 5	January 2000
64	Franc J.G.M. Klaassen	Exchange Rates and their Effects on International Trade; ISBN 90 5668 063 3	January 2000
65	Radislav Semenov	Cross-country Differences in Economic Governance: Culture as a major explanatory factor; ISBN 90 5668 065 X	February 2000
66	Alexandre Possajennikov	Learning and Evolution in Games and Oligopoly Models; ISBN 90 5668 064 1	March 2000

* Copies can be ordered from ThelaThesis,Prinseneiland 305, 1013 LP Amsterdam, The Netherlands, phone + 31 20 6255429; fax: +31 20 6203395; e-mail: office@thelathesis.nl

Ingevoerd: 1017:02-03-00 Gewijzigd: 1017:09-03-00 13:18:20 Status: 1017:02-03-00

0500 Aaa
1100 2000
1110 Aa019
1500 / 1eng/4ned
1700 / 1nl
2000 9056680641
3000 Alexandre@Possajennikov!157413462!*Alexandre Possajennikov*
4000 @Learning and evolution in games and oligopoly models / door Alexandre Possajennikov
4030 Tilburg : CentER for Economic Research, Tilburg University
4060 VII, 142 p
4061 ill
4062 24 cm
4170 @CentER dissertation series ; 66
4180 #660#!140541977!@*CentER dissertation* ; 66
4204 Met lit. opg., en een samenvatting in het Nederlands
4209 Proefschrift Katholieke Universiteit Brabant, Tilburg
7001 09-03-00 : 6ECOgraa
7100 ECO advies
8200 1593425
7800 317085484
7002 09-03-00 : 6CBMgrac
7100 CBM advies
8200 1593426
7800 317085492

Invitation

to the public defense of my thesis

Learning and Evolution in Games and Oligopoly Models

on Wednesday, March 8, 2000 at 16:15
in the Aula of Tilburg University.

There will be a reception after the defense.

Alexandre Possajennikov
Adelaarshorst 68
5042 XJ Tilburg
013-4679029

CentER Diss

Alexandre Possajennikov graduated in Mathematics from Moscow State University in 1993 and then obtained his Master's degree in Economics from the New Economic School, Moscow, in 1995. After that, he joined the CentER for Economic Research graduate program. He is working in game theory, particularly on dynamic models of learning and evolution. He currently holds a post-doc position at Dortmund University.

Dynamic models of adjustment, as well as static models of equilibrium, are important to understand economic reality. This thesis considers such dynamic models applied to economic games. The models can broadly be divided into two categories: learning and evolution. This thesis analyzes reinforcement learning and imitation dynamic on the learning side and the indirect evolution approach on the evolution side. It demonstrates the relation between the concept of Nash equilibrium and the long run outcome of the dynamic processes. The applications of the dynamic models to economic games include, among others, Cournot oligopoly and merger games.

ISBN: 90 5668 064 1