Ph.D. Dissertation in Economics:

Evolutionary Game Theory, Markets and Conventions.

María Sáez-Martí.

University College London.

ACKNOWLEDGMENTS

I am especially grateful to Ken Binmore for his advice, criticism and encouragment. I would like to thank Tilman Börgers and Fabrizio Zilibotti for helpful comments and suggestions. Financial support from the Banco de España is gratefully acknowledged.

ABSTRACT

The Thesis applies evolutionary game theoretic ideas to the modelling of economic behaviour. The traditional approach is to assume that economic agents are perfectly rational. The rationality assumption requires that agents be very sophisticated in their decision making. This creates a gap between the behaviour postulated by economic models and real people's behaviour. Recently, theorists' attention has turned to models of bounded rationality in which agents are assumed to find their way to equilibrium by trial-and-error methods. The most widespread approach to modelling boundedly rational behaviour is the use of evolutionary game theory which also provides useful insights into the equilibrium selection problem. The first chapter is a critical survey of evolutionary game theory. In the second chapter we endogenize the learning rules in a modified version of Young's bargaining model which provides an evolutionary explanation for the asymmetric Nash bargaining solution. The Nash Demand Game is played by two different populations. Players choose their strategies in the light of some limited information about the strategies players from the other population have used in the past. An interesting result is that the better informed population has higher bargaining power. The main drawback in Young's model is that the amount of information, and therefore the bargaining powers are fixed exogenously. We endogenize players's learning rules and test for evolutionary stability. We study whether one population using a particular learning rule can be invaded by a mutant learning rules. We show that, when information is costless, the only evolutionarily stable learning rule maximizes players' information. If both populations follow the same learning rule, the equilibrium which is selected is the symmetric Nash bargaining solution. When information is costly there is a trade-off between costly learning and the rewards of being well informed. Finally we show that an economy populated by players who follow very simple imitative rules is socially more efficient than an economy of rational players. In the third and fourth chapters we introduce models which endogenize the equilibrium selection problem. We show that, in a model of a credit market an equilibrium is selected in which the market is fully developed, although the model has another equilibrium with interesting stability properties. The agents are assumed to follow very simple behavioural rules and sometimes to experiment with unplayed strategies. Similar results are obtained after the introduction of a small proportion of rational players. The model predicts interesting dynamics which are consistent with some evidence about the Great Depression. Real shocks trigger episodes of credit-crunch and temporary financial collapse which is observed in the process of adjustment towards the post shock equilibrium.

Table of Contents

List of Figures

List of Tables

Introduction

Nash equilibrium is the most widely used solution concept in economic theory. In a Nash equilibrium every player plays a strategy which maximizes his expected payoff given the strategies chosen by all other players. No player has an incentive to deviate from the strategy assigned to him, provided that all the others are playing the Nash equilibrium strategy. In other words, when a Nash equilibrium is played, no player will regret his strategy choice.

There are two main open questions concerning Nash equilibrium play,

(i) How do uncoordinated players arrive to play Nash equilibrium?

(ii) When there are multiple equilibria, which one is actually going to emerge?

In recent years, the 'evolutionary' answers to these questions have raised an increasing enthusiasm among game theorists and economists. In the evolutionary explanation the equilibrium is a long run phenomenon achieved by myopic players through iterative play. Players belong to large populations and randomly meet one another in paiwise interactions. Those strategies which perform better will be adopted by an increasing overtime proportion of people. The changes in the fractions of players employing the different strategies are caused by some type of natural selection (in a biological context) or imitation (in economic applications) of more successful strategies. In this approach, Nash equilibria are stationary points of evolutionary processes where players learn from the experience of their own population. The equilibria are not the result of complicated reasoning carried out by rational players but the result of an evolutive process.

In this Thesis we apply evolutionary game theoretic ideas to the modelling of economic behaviour. The first chapter is a critical survey of the literature mainly focussing on the problem of equilibrium selection in games. In the second chapter we propose an evolutionary model of bargaining which is close in spirit to the model of Young (1993b). In Young's model, the Nash Demand Game is played by two different populations whose members may differ in the amount of information which they use. The main result is that each population's bargaining power is determined by the least informed player in the population. We endogenize players's learning rules and characterize those which are evolutionarily stable, i.e. which cannot be invaded by a mutant learning rule under natural selection. We show that some puzzling results of Young's model are not evolutionarily stable. In the third chapter we study the implications on the equilibrium selection in asymmetric games of the introduction of a small proportion of rational players. It is known that mixed equilibria in asymmetric games cannot be evolutionarily stable. We show that the introduction of a small proportion of "rational players" makes possible to achieve mixed equilibria as long run outcomes of some darwinian processes. The fourth and last chapter is an evolutionary model of a credit market in which multiple equilibria exist. We show that, under "darwinian dynamics", the model generates episodes of credit crunch and financial collapse which are observed during the process of adjustment which follows real shocks in the economy. The predictions of the model are consistent with some empirical evidence about the Great Depression.

Chapter 1

Evolutionary Game Theory: A Critical Survey

1 Introduction

Recent research on the foundations of game theory has shifted the attention to evolutionary explanations of Nash equilibrium, since the traditional 'eductive' theories (Binmore (1987a)) have not provided a satisfactory interpretation. A Nash equilibrium is a strategy profile such that the strategy assigned to each player maximizes his expected payoff given that the other players play their strategies in the profile. In the 'eductive' interpretation, it is taken for granted that the players know the strategies that the other players will use. They are assumed to figure this out from the assumption that it is common knowledge that everybody is a rational utility-maximizing player, although there is no clear consensus among game theorist about how they do this in the general case. Even when the necessary calculation is clear, it often requires a large computational capability on the part of the players to carry through.

Equilibrium selection Traditional game theory put all its emphasis on equilibrium states, without offering a convincing general theory of how players know that a Nash equilibrium will be played, and which Nash equilibrium it is appropiate to select in case of multiplicity. All the literature on refinements of Nash equilibrium is based on the idea of requiring the equilibria to satisfy some stronger notion of rationality. A refinement is a test which allows to rule out some equilibria. Even when a refinement selects a unique equilibrium, different refinements might select different ones. The main problem with refinements of Nash equilibrium is that any of them could be defended on some grounds as the right solution concept for a particular game but none of them provides a convincing general solution.

The evolutive approach stresses the importance of dynamics in explaining the emergence of equilibrium, thus providing an alternative rationale to Nash equilibrium and an important contribution to the literature on equilibrium selection. Evolutive theories see equilibria in games as stationary points of dynamics processes representing some kind of evolutionary adaptation. The main difference between the traditional and the evolutionary approach is that observed equilibria are not the result of a rational choice but stable population distributions achieved by an evolutive process. When the dynamic processes are clearly specified it is not only stable Nash equilibria that are of interest but also other phenomena such as attractors (cycles, strange attractors, limit cycles) that, although not equilibrum points, could nevertheless serve to predict long-run behaviour (Taylor and Jonker (1978), Zeeman (1980), Schuster and Sigmund (1983)and (1986)).

Bounded rationality The evolutionary approach has also contributed to the growing litereráture on bounded rationality. The general motivation of this literature is that economics and game theory assume that players are unrealistically rational. In fact economic agents are limited in their knowledge and computational capabilities. The limitation on computational capabilities, compared with the complexity of the problems, becomes evident when we consider repeated games. For example the hundred-times repeated prisoners' dilemma has $2^{2^{100}}$ pure strategies. It is therefore natural that early attempts to model boundedly rational players occur in the literature on repeated games (Neyman (1986),Rubinstein (1986), Abreu and Rubinstein (1988)). Neyman (1986) models players as finite automata. He considers only strategies that can be programmed using automaton of fixed size. He obtains cooperation in the finitely repeated prisoners' dilemma even though all Nash equilibria result in players always defecting in the case when every strategy is available. Abreu and Rubinstein (1988) go a step further and endogenize the computational constraints of the machines. [1]

A second level of bounded rationality is related to the players' difficulty in receiving, decoding, and acting upon information they get in the course of playing games. This is the type of bounded rationality that is present in evolutionary models, where players face the same or similar situations in a repeated way. The crucial assumptions of evolutionary models are that interactions are anonymous (to rule out repeated games effects) and that players are myopic. Evolutionary models are necessarily models of bounded rationality, which is reflected in the myopic rules followed by players in deciding how to behave in future encounters.

Binmore and Samuelson (1992) consider the two different levels of bounded rationality. They consider repeated games played by automata with the election of automata driven by evolutive forces.

Biology vs. Economics. In their seminal paper 'The logic of animal conflict', Maynard Smith and Price (1972) provide an evolutionary explanation of conflicts between animals of the same species by means of individual rather than group selection. Since then, game theory has become a common tool in biology, although it was originally intended as a theory of strategic interaction between rational, utility-maximizing agents, with social science as the intended field of application. Animals are far from being the conscious, rational agents with well-defined preferences which populate the world described by economic models. In applications to biology, the assumptions of rationality and utility maximization therefore had to be abandoned and a

[1] They consider a situation in which players select a finite automata to play the infinitely repeated prisioners' dilemma. When two different automata receive the same payoffs the less complex is prefered.

new interpretation of strategies and payoffs was introduced in seeking game theoretic-explanation of animal behaviour.

> A central assumption of classical game theory is that the players will behave rationally, and according to some criterion of self-interest. Such an assumption would clearly be out of place in an evolutionary context. Instead, the criterion of rationality is replaced by that of population dynamics and the self-interest criteria by Darwinian fitness (Maynard Smith (1982)).

The payoffs are interpreted as the number of offspring and the strategies as phenotypes [2] or genetically programmed ways of behaviour.

The evolution by natural selection rests on two mechanisms. On the one hand there are random changes or mutations which introduce genetical variations. On the other hand there is need of some mechanism which selects all those mutations which happen to be useful. Natural selection works by retaining a mutations if the organism in which it has occured can be expected to leave more offspring than others in the same species which have not been subject to the mutation. The evolutionary dynamics reflect this fact. Those genes which confer higher fitness will be found in higher proportion.

Evolutionary biologists think of players in terms of populations rather than individuals. The set of pure strategies is interpreted as a set of phenotypes (the set of manifested attributes like eye colour, degree of agressiveness, blood group) which, in these simple models, are determined genetically. We may think of players as being programmed always to play the same strategy when meeting another individual. Individuals are randomly matched , reproduce asexually and breed true. The result of the interaction is a change in Darwinian fitness, i.e. the expected number of offspring. Those organism which are better than others at leaving offspring which survive to reproduce

[2] The genetic program of an organism is stored in the chromosomes in different sites (loci). The alleles are the different types of genes that may occupy a locus and the genotype is the allele that actually occurs and it determines the phenotype.

will be better represented in future generations.

The game-theoretic approach to animal behaviour has largely been critized because of the assumption of asexual reproduction. Although asexual reproduction is not a good representation of the biological world in which reproduction is mostly sexual, it might be a better approximation to the evolution of cultural traits, with culture understood as "the transmission from one generation to another, via teaching and imitation, of knowledge, values, and other factors that influence behaviour" (Boyd and Richerson (1985)). A cultural trait, such as look left and right before crossing a street, becomes part of the phenotype of an individual.

It is important to bear in mind that the "cultural parents" are not necessarily genetic ones. Cavalli-Sforza and Feldman (1981), for example, cite different circumstances in which an individual may have only one 'cultural parent' for certain traits. When the transmission takes place between parents and offspring (vertical transmission), although both parents contribute to the cultural traits inherited by the offspring, it may happen that the high specialization and division of labour among the sexes can lead to some traits being transmited in a uniparental way. Another example in which an individual has a unique cultural parent is in the relation teacher-pupil . A setting in which we think the assumption of asexual reproduction is a good approximation occurs when players are treated as roles (male/female in the battle of the sexes, seller/buyer in a market situation, borrowers/lenders in a credit market) and the strategies are possible behaviours in a role which evolve accordingly to their successes due to horizontal transmition (Cavalli-Sforza and Feldman (1981)), i.e. when the cultural trait is spread within the same generation. "The item of culture being spread horizontally acts like a microbe that reproduces and spreads rapidly because it is "infective" and has a short generation length compared to the biological generation.."(Boyd and Richerson (1985)). Clearcut examples are fashions and technical innovations.

A literal application of the evolutionary approach as it is in biology to economics is problematic. The payoffs in economic models do not represent reproductive fitness, but profits or utilities. Nor are the derivations of the dynamic processes studied by biologists adequate to model a economic environment in which people endeavor to achieve a satisfactory performance and concious imitation and learning are present. However, in an economic enviroment Darwinian dynamics can sometimes be justified by some type of imitative behaviour, with individuals revising their strategies in the light of the different strategies' relative successes. The dynamics in these cases represent the aggregate effect of the revising rules individuals employ. Nelson and Winter (1982), for instance, draw an analogy between economic competition and natural selection. In their theory firms are characterized by a set of routines (which are equivalent to genes). The profit maximizing routines happen to do better than others and spread in the population by imitation. The mutations in this framework are the techniques. Firm which have poor performances, because their profits fall bellow an aspiration level, introduce more efficient methods of production. Firm may also decide to imitate more succesful firms. At the end all firms adopt the new techniques.

Recent research justifies the use of 'darwinian' dynamics in the economic literature. Cabrales (1993) obtains the replicator dynamics in a model in which a small proportion of players are randomly paired with members from the same population and compare payoffs. The player with the lower payoff changes to the strategy of the one with higher payoff with a probability which is proportional to the payoff difference. The replicator dynamics are also obtained by Binmore and Samuelson (1993b) in a model in which, in every period, each agent has a small probability of comparing his payoff with a random aspiration level. If the payoff falls short of the aspiration level the player selects a new strategy which depends on the distribution of strategies in the population. Börgers and Sarin (1993) and Schlag (1994) justify the replicator dynamics in models with individual learning rather than in models of learning in populations. Börgers and Sarin (1993) show that the replicator dynamics are a long run approximation to a learning process due to Cross (1973). In their model, players' payoffs play a reinforcing role. Players adapt their mixed strategies according to the strategy they play and the payoffs they receive. Schlag (1994) obtain the replicator dynamics in a model where player select updating rules.

The standard assumption of random matching in large populations should be applied with caution in economics, where long term interactions and reputation are important elements. The assumption of myopia, common in all evolutionary models, which implies that players behave as if their enviroment is unchanging through time has been relaxed in the recent literature with the aim of incorporating some type of anticipatory behaviour. In models with finite number of long-lived agents it may be sub-optimal for players to choose strategies that are best replies to the current state. If all players simultaneously revise their strategies, it is possible that a good strategy today maybe bad tomorrow. It may also happen that with the choice of strategy a player influences other players' future behaviour. The same is not true when we consider continuous-time dynamics or when only a small proportion of players are allowed to revise their strategies (there is inertia in the system). In these last cases it is likely that what is optimal today will still be optimal in the near future. Banerjee and Weibull (1991) and Stahl (1992) consider models in with more sophisticaded players than those usually assumed in evolutionary models.

The rest of the chapter reviews some of the main ideas developed in evolutionary game theory. As we have mentioned above the first contributions were made by evolutionary biologist, but the increasing enthusiasm shown by game theorists and economists has contributed to the growing research in the field. The first section is devoted to the main equilibrium concept in evolutionary biology : evolutionarily stable strategies. In the second section we stress the importance of the dynamics in the equilibrium selection in games. After a general discussion about the most recent contributions we review some of the dynamics which have been most widely used in the literature. We consider deterministic and stochastic dynamics.

2 Evolutionary Stability

The main contribution of Maynard Smith and Price (1972) is the notion of evolutionarily stable strategy (ESS), which is the most widely used equilibrium concept in evolutionary game theory. An ESS is a refinement of the concept of a symmetric Nash equilibrium with the theoretical contribution consisting in an additional stability requirement. An ESS is a strategy such that, if all the members of the population adopt it, no alternative mutant strategy can invade the population. A population consisting of members adopting an ESS will cease to evolve.

The concept of ESS is originated in a setting in which all members of the population adopt the same (possibly mixed) strategy and only one type of mutant, all playing the same strategy, are allowed to enter. The ESS concept, as originally developed, applies to symmetric games in which animals from the same population are randomly chosen to occupy the two player roles. Selten (1980) extended the idea to asymmetric games by assuming that animals (who are still chosen from the same population) condition their strategy choice on the player role they are assigned. In such asymmetric games, equilibria that have alternative best replies cannot be evolutionarily stable. Selten (1980) shows that in asymmetric games, a Nash equilibrium cannot be ESS unless it is strict. All mixed equilibria are ruled out in a very wide variety of games. In particular when extensive-form games are considered, those equilibria in which some information set is not reached fail to be evolutionarily stability. As Swinkels (1992) points out ESS may fail to exist because all possible mutants have to be considered when we test for evolutionary stability. Many of the mutant strategies can be unstable themselves.

Several refinements of the concept of ESS have appeared in the literature. Examples of them include, among others, neutraly stable strategies (Maynard Smith (1982)), weak ESS (Hofbauer and Sigmund (1988)), limit ESS (Selten (1983)), equilibrium evolutionarily stable set (Swinkels (1992))and modified evolutionarily stable strategy (Binmore and Samuelson (1993b)).

In this section we introduce the concept in a model where players can use mixed strategies. We will illustrate the notion with the Hawk-Dove game. By means of an example due to Taylor and Jonker (1978) we show that ESS might not be the right equilibrium concept on which to focus.

Let us consider a large population of players. The members of the population are randomly matched to play a symmetric two-player, normal-form game. The set of pure strategies is $S = \{1, \ldots, n\}$. Players can use mixed strategies. Throughout his life, each individual plays a fixed strategy which is genetically determined. The entries in the $n \times n$ matrix $A = (a_{ij})$ represent fitnesses i.e. the expected number of offspring. Individuals reproduce asexually and breed true unless there is a mutation.

Let us assume that all the members of the population play the same strategy p and a proportion ϵ of mutants playing $q \neq p$ appear. The random matching assumption implies that a player's expected payoff in his match is equal to the expected payoff obtained if he were matched with a player using the mixed strategy $\epsilon q + (1-\epsilon)p$. This mixed strategy is exactly the population mix. The respective expected payoffs to strategies p and q in the post-mutation enviroment are given by:

$$p^T A(\epsilon q + (1-\epsilon)p) \; ;$$

$$q^T A(\epsilon q + (1-\epsilon)p) \; .$$

Definition 1 (Evolutionarily stable strategy) *A state p is an evolutionarily stable strategy if, for every state q:*

$$p^T A(\epsilon q + (1-\epsilon)p) > q^T A(\epsilon q + (1-\epsilon)p)$$

for sufficiently small ϵ.

Evolutionary forces will select against a mutant if and only if its fitness in the post-entry environment is lower than that of the incumbent strategy.

The condition above can be rewritten as the following two conditions, which are those originally proposed by Maynard Smith and Price (1972):

(a) Equilibrium condition:

$$q^T Ap \leq p^T Ap \tag{1}$$

(b) Stability condition:

$$\text{if } q \neq p \text{ and } q^T Ap = p^T Ap, \text{ then } q^T Aq < p^T Aq \tag{2}$$

The first condition is the definition of a symmetric Nash equilibrium. Evolutionary stability therefore implies Nash equilibrium play. The stability condition guarantees non-invadibility in the case when alternative best replies exist (equality in the equilibrium condition). It requires that if there is an alternative best reply, the ESS must do strictly better again the alternative best reply than the alternative best reply does against itself. The following results hold for ESS in symmetric games (see Van Damme (1987)):

(i) if (p,p) is a strict Nash equilibrium, then p is an ESS.

(ii) if p is an ESS, then (p,p) is a trembling-hand perfect equilibrium.

(iii) if p is an interior ESS, then it is unique.

(iv) if p is an ESS, then (p,p) is a proper equilibrium.

The implication that ESS in a symmetric game gives rise to a symmetric Nash equilibrium follows from the assumption that players are drawn from the same population and cannot condition their action on their role as row or column players. However, as Hofbauer and Sigmund (1988) point out, many conflicts are asymmetric. "Food is more important for a starving animal than for a replete one, while the risk of injury is smaller for a stronger contestant. In fact, asymmetries are not incidental but quite often essential features of the game: for example, in conflicts between males and females, between parents and offspring, between the owner of an habitat and intruder,or between different species".

We will consider asymmetric contests played by different populations. We consider two-player normal-form games in which $S_1 = \{1, 2..., m\}$ and $S_2 = \{1, 2, .., n\}$ are the sets of pure strategies available to the row and the column player respectively. Mixed strategies $x \in \Delta^{m-1}$ and $y \in \Delta^{n-1}$ are possible. $A = (a_{i,j})$ and $B = (b_{j,i})$ are the payoff matrices. We assume that there are two large populations $\mathcal{X}$ and $\mathcal{Y}$ and that the members of these populations are randomly matched in pairs to play the game. Player from population $\mathcal{X}$ fill in the role of row players while those from $\mathcal{Y}$ are the column players.

Definition 2 (Evolutionarily stable strategies in asymmetric games) *A pair of strategies (p, q) with $p \in \Delta^{m-1}$ and $q \in \Delta^{n-1}$ is an evolutionarily stable strategy if*

$$p^T Aq > x^T Aq \text{ for all } x \in \Delta^{m-1}, \;\; x \neq p$$

and

$$q^T Bp > y^T Bp \text{ for all } y \in \Delta^{n-1}, \;\; x \neq q$$

The definition of ESS in an asymmetric game concides with the notion of "strict" Nash equilibrium. Thus, in asymmetric games, a mixed strategy can never be evolutionarily stable.

The following example, due to Maynard Smith and Price (1972), illustrates the concept of ESS and the different results obtained when different populations are considered.

2.1 The Hawk-Dove game

Imagine two animals that are contesting a resource of value V. The resource is, for example, a territory in a favourable habitat. The animal which obtains the resource has its Darwinian fitness increased by V. Let us assume that only two pure strategies are possible.

Hawk. An animal adopting the hawk strategy always fights. It stops fighting only after being seriously injured or when the opponent retreats.

Dove. An animal adopting the dove strategy threatens in a conventional way but immediately retreats if the opponent escalates.

When two hawks meet, both will escalate and a fight will result. It is assumed that such fights have an uninjured victor and an injured loser. The victor takes the territory and increases his fitness by V. The loser's fitness is decreased by C. Each hawk in a fight has an equal chance of winning. The expected payoff to a hawk that meets another hawk is, therefore, $(V - C)/2$. When a hawk meets a dove, the dove retreats and the hawk obtains the resource with neither being injured. The hawk's fitness increases by V while the dove's fitness is unchanged. When two "doves" meet, they peacefully share the territory. Each then gains $V/2$ in fitness.

The game representation of the "hawk-dove" story is shown in Figure 1.

	dove	*hawk*
dove	$\frac{1}{2}V$ (upper right), $\frac{1}{2}V$ (lower left)	V (upper right), 0 (lower left)
hawk	0 (upper right), V (lower left)	$\frac{1}{2}$(V-C) (upper right), $\frac{1}{2}$(V-C) (lower left)

Figure 1

We assume a large population of animals which are matched in couples to play the Hawk-Dove game. Reproduction is asexual and takes place after the game is played. A new composition of the population is obtained and a new round takes place. Animals are genetically programmed to play a strategy that remains unchanged throughout thair lives. Animals breed true in the absence of mutations.

When $V > C$ hawks is the only evolutionarily stable strategy because hawk strictly dominates dove. In a population of hawks and doves, the hawks have higher Darwinian fitness. Doves will gradually disappear and in the "long run" only hawks will exist. It is a Prisoners' Dilemma situation.

When $V < C$ neither a population of all hawks nor one of all doves is evolutionarily stable. A mutant will spread through a pure population. When all are hawks, a dove has a higher fittness and will spread in the population. The opposite happens in a population of doves in which the proportion of hawks will increase.

The only evolutionary stable strategy in a monomorphic population requires that animals adopt mixed strategies. Let us assume that animals play hawk with probability x and dove with probability $(1 - x)$.[3] The only symmetric Nash equilibrim of the Hawk-Dove game when $V < C$ is $p = V/H$. Let us consider a population of p-strategists. Assume that a mutant playing $q \neq p$ appears. Since all the strategies that are played with positive probability in a mixed strategy Nash equilibrium have the same expected payoff any q will fare against p as well as p does. The *equilibrium condition* (1) is therefore satisfied with equality. We have to test for the stability of p. The *stability condition* (2) requires comparing the payoffs to the two different strategies, p and q, when matched against q The mutant strategy q will be selected against if and only if it performes against itself worse than p.It is easy to see that the *stability condition* is satisfied by p, being the only ESS of the game when $V < C$.

Now, let us assume that the column and the row players are drawn from different populations. It is then evolutionarily stable for one popualation to consist of all doves and the other of all hawks. The dove strategy is the

[3]We will identify a mixed strategy with the vector $\mathbf{x} = (x, (1 - x))$ or simply with x.

best against a population of all hawks. Against "all doves" it is better to be a hawk. A situation in which all members of both populations adopt the mixed equilibrium strategy p is not evolutionarily stable. Let us consider the pair (p, p) and a mutant q $(q \neq p)$ in population 1. The mutant will not be selected against. In the case we are considering of two separated populations, a mutant never meets itself. There is no analogue of *stability condition* (2) to restore the pre-mutation environment.

2.2 Mixed vs pure strategies

The use of mixed strategies has the nice property of convexifying the strategy space, which guarantees the existence of Nash equilibrium in finite games. In mixed strategy equilibria a player, although he is indifferent between all the pure strategies in the support of the mixed strategy, has to randomize in such a way that the opponent is indifferent between all the strategies which he is require to play with positive probability in equilibrium. A justification of mixed strategies is that people randomize to become unpredictable. This is the case in a game like "matching pennies", when the game is played repeatedly and players choices are observable. The same explanation is not valid for evolutionary models where matching is random and past behaviours of the current opponent are not observed.

Underlying Maynard Smith and Price's definition of ESS are the assumptions that populations are monomorphic, players can be genetically programmed to play mixed strategies and that they breed true. All the strategies in the support of a mixed equilibrium have the same expected payoff. For evolutionary stability, any mutant with the same support must have the same payoff against the incumbent as the incumbent itself, because all the straegies pay the same. The same is not true when we consider the payoffs a mutant receives when he meets someone like himself. Some of the pure strategies in the support of the mutant strategy may have a higher expected payoff. One could expect a change in the probabilities attached to these strategies with the population moving in the direction of increasing a pure strategy rather than towards an increase of the proportion of players employing the mixed mutant strategy.

Example 1.1 This example shows that the results may differ when we compare pure populations using mixed strategies and mixed populations using pure strategies. Equilibria in the latter case are said to be polymorphic. When there are only two pure strategies, if the mixed stategy is stable then so is a polymorphic population with the frequencies of the two strategies corresponding to the mixed equilibrium. For more general games this is not necessarily true.

Let us consider the following example by Taylor and Jonker (1978)

$$A = \begin{bmatrix} 2 & 1 & 5 \\ 5 & 1 & 0 \\ 1 & 4 & 3 \end{bmatrix}$$

The unique symmetric Nash equilibrium of this game is $p = (15, 11, 9)/35$. It is not an ESS because we can find a mutant strategy q that fares better than p in a population consisting of a large proportion of p and a small fraction of q. Consider the perturbing state $q = (18, 18, 0)/35$. The new population mix is given by $(1 - \epsilon)p + \epsilon q$. The payoffs to the different mixed strategies p and q in the post-entry enviroment are:

$$q^T A(\epsilon q + (1 - \epsilon)p) = 86/35 - 0.22\epsilon$$

$$q^T A(\epsilon q + (1 - \epsilon)p) = 86/35 - 0.19\epsilon$$

We should therefore anticipate a tendency in the population towards q.

When we test for evolutionary stability a comparison is made between the payoffs to the different mixed strategies. If we view the different groups of p-strategists and q-strategists as subpopulations made up of players using pure strategies in the proportions p and q respectively, the evolution of the system would differ.

The payoffs obtained by the three pure strategies in the post-entry enviroment are:

$$e_1^T A(\epsilon q + (1-\epsilon)p) \quad = (86-\epsilon)/35$$
$$e_2^T A(\epsilon q + (1-\epsilon)p) \quad = (86+21\epsilon)/35$$
$$e_2^T A(\epsilon q + (1-\epsilon)p) \quad = 86/35$$

where e_i is the (1×3) unit vector with a 1 in the ith row and everywhere else. Clearly, the second strategy has a higher expected payoff. The evolution of the population will drive the population towards an increase in the second strategy rather than towards q.

In the light of this example and for the reasons explained above, we therefore confine our attention to what we regard as more realistic framework that in which only pure strategies are used. In the following section we will consider the situation in which different strategies can coexist in equilibrium. In this case a mixed equilibrium is to be interpreted as a population mix with the appropiate proportions of players using the pure strategies in the support of the equilibrium. [4]

3 Dynamics and equilibrium selection

Although the definition of evolutionary stability is static it implicity relies on an unmodeled dynamic story. The main intuition behind the ESS motivation is that, if the mutant strategy performs worse than the incumbent, then selection pressures will repel the mutants. An equilibrium is an ESS if the system returns to the equilibrium once a small shock moves the system an arbitrarily small distance from the equilibrium.

In the early evolutionary game-theoretic literature, stress was put on the

[4]Hofbauer and Sigmund (1988) echo Maynard Smith (1982) in pointing out that is rare to find phenotypes that realize mixed strategies in the type of situation with which we are concerned although mixed strategies are apparently commonplace when animals are "playing the field".

stability of the different equilibria that were tested with the introduction of one-shot mutations (the concept of ESS, for instance). When the underlying dynamic forces have been explicitly modeled, it becomes evident that some other elements such as limit cycles, cycles or strange attractors are interesting phenomena to be considered. The equilibrium selected by the 'deterministic' dynamics, when they converge, depends on the initial conditions. In brief, history matters. The scenario changes when the dynamic model is continuously perturbed. Mutations, experimentation, imigration and shocks are elements which play a crucial role in the equilibrium selection. The introduction of one of these elements which continuously perturbes the system helps in the selection of a particular equilibrium. The role played by perturbations is to eliminate the path dependence. The emphasis now, is put on long run equilibra rather than on absorbing states. A long run equibrium is a probability distribution defined on the state space. It is a measure of the proportion of time spent by the system at the various states as the time horizon tends to infinity. The idea is that in the presence of continuous perturbations the system will select a set of states near which it will stay. The first work in this line is due to Foster and Young (1990). Foster and Young (1990) consider a continuous-time, continuous state-space formulation (as in Taylor and Jonker (1978)) with aggregate randomness due to several factors such as variability of payoffs (due to changes in the environment), randomness in the matching process, mutations and possible inmigrations. The accumulation of stochastic effects may change the long-run behaviour of the system. Foster and Young (1990) offer a new concept of equilibrium, i.e, stochastically stable sets, which are those states that will be observed with positive probability in the long run, when the noise of the system is small does not vanish altogether. Although the stocastically stable sets are independent of initial conditions they depend on the fine details of the dynamics.

More recently, Kandori et al. (1993), Young (1993a) and (93b) propose a discrete-time, discrete state space models which are continuously perturbed by mistakes. The novelty with respect to Foster and Young (1990) is that the randomness is introduced at the individual rather than the aggregate level. In the models of Kandori et al. (1993) and Young (1993a) the best reply dynamics defined on the discrete space state defines a Markov Chain with multiple absorbing states. The introduction of mutations eliminates the absorbing states and makes the Markov Chain ergodic.

Kandori et al. (1993) consider coordination games and show that the equilibrium which is selected is independent of the initial conditions and on the details of the Darwinian dynamics. In their model, players from a single finite population are repeatedly matched to play a coordination game. They assume Darwinian dynamics that satisfy a weak monotonicity condition reflecting the hypothesis of myopia and inertia. When there are no mutations, the dynamics are deterministic, the equilibrium that is selected is path dependent; once an equilibrium has been reached the economy stays there forever. When mistakes are allowed, they continuously perturb the system away from equilibrium making it possible to jump from the basin of attraction of one equilibrium into another. Kandori et al. (1993) focus on the long run behaviour of the system, when a non-negligible number of mutations occur, and show that such behaviour is independent of the initial conditions and the fine details of the dynamic processes considered; the weak monotonicity condition being the only restriction imposed.

Young (1993a) considers an n–person game played once every period by n players drawn at random from large populations. Each player who is called to play bases his choice of strategy on the observation of a random draw from the last m plays of the game. Young prefers to refer the rare perturbations that Kandori et al. (1993) call mutations as mistakes or experimentations. These mistakes constantly disturb the system. When the noise is small it is possible to show that the stationary distribution is concentrated around the 'stochastically stable' conventions (Foster and Young (1990)).

Most evolutionary models incorporate "selection dynamics" which describe the way players adjust their strategies according to some criteria related to current payoffs. Under the so called Darwinian dynamics the strategy distribution moves towards what is best reply to the current situation. The most commonly used are the replicator and the best-reply dynamics.

3.1 Replicator Dynamics

The most common explicit dynamic process offered as a representation of a Darwinian selection mechanism is the 'replicator dynamics'. The notion of replicator was put forward by Dawkins (1976) as the entity of the theory of natural selection. Replicators are any entity which can be copied and which determine the strategic behaviour in a game. Genes are replicators as well as rules-of-thumb, fashions and ideas. The idea behind the "replicator dynamics" is that the fitness of an organism is measured by the frequency it gets to reproduce its replicators. Those replicators which confer higher fitness to the organism carrying them will come to control a larger proportion of the organism which survive. Some replicators are more succesful ar replicating themselves than others. In the Hawk-Dove game, for instance, there are two replicators. The replicator dove induces the animal to play dove while the replicator hawk induces the hawk behaviour. The fitness of any of them depends on what replicators the other members of the populations are carrying.

3.1.1 Deterministic Replicator Dynamics

Taylor and Jonker (1978) first introduced the replicator dynamics to explain animal behaviour. We will follow them in our analysis.

Let us consider a large population of players. The members of the population are paired at random to play a symmetric two-player, normal-form game. Each individual plays a fixed strategy $i \in \{1, \ldots, n\}$ that is genetically determined, i.e. carries a replicator[5] which induces a pure strategy i. The payoffs a_{ij} in an $n \times n$ matrix A represent the expected number of offspring to a player who uses strategy i when his opponent uses strategy j. Individuals reproduce asexually and breed true provided that no mutation occurs. Let N be the total number of individuals in the population and n_i the number of i-strategists (carriers of replicator i). The state of the population at time t is given by an n-dimensional vector $x = (x_1, \ldots, x_n)$, with $x_i = n_1/N$ being defined as the proportion of players with phenotype i that is, who are programmed to use strategy i. The change in the distribution of strategies in the population is determined by the rate at which the users of each strategy reproduce. In this setting, we do not have mixed strategies. Instead, different players using different strategies may coexist in a polymorphic equilibrium. But from the point of view of a single players it is as if he were playing against an individual using a mixed strategy.

Let e_i be the unit vector with a one in the i-th row and zero elsewhere. The expected payoff (fitness) to strategy i is $e_i^T Ax$. The average fitness of the population is $x^T Ax$. Let us assume that that the rate of growth or decay of a certain strategy (replicator) is r_i and that the fitness of the strategy is an estimate of it:

$$\dot{n}_i = r_i n_i = n_i e_i^T Ax \;\; and \;\; \dot{N} = \sum N x_i r_i = N x^T Ax$$

Differentiating $x_i = n_i/N$ we get the 'replicator equations'

$$\dot{x}_i = x_i(r_i - \sum r_i x_i) = x_i(e_i^T Ax - x^T Ax) \quad \text{(i=1,...,n)} \tag{3}$$

It is easy to see that the population share that grows at the highest rate uses the strategy that pays best against the current state of the system. However, the shares of more than one strategy may be increasing.

[5]In what follows we shall use the terms strategy and replicator indifferently. We associate to a replicator the strategy it induces.

The continuous time replicator dynamic approximates the trajectories of different discrete-time replicator dynamics, either with overlapping generations (Taylor and Jonker (1978), Hofbauer and Sigmund (1988) , Binmore (1991a)) or with nonoverlapping generations (Maynard Smith (1982) and Van Damme (1987)).

Let us assume that we have discrete overlapping generations and that r_i is the number of individuals born from an i-strategist in unit time

$$n_i(t+\tau) = n_i(t)(1+\tau r_1) \quad \text{(i=1,...,n)}$$

Dividing by $N(t+\tau) = \sum n_i(1+\tau r_i)$ we get

$$x_i(t+\tau) = \frac{n_i(t)(1+\tau r_i)}{N(t)+\sum \tau r_i n_i(t)} = x_i(t)\frac{1+r_i\tau}{1+\sum r_i \tau x_i(t)} \quad \text{(i=1,...,n)}$$

Substituiting r_i by its estimate and rearranging we obtain the following discrete time 'replicator dynamics'

$$x_i(t+\tau) - x_i(t) = x_i(t)\frac{\tau(e_i^T Ax(t) - x(t)^T Ax(t))}{\tau x(t)^T Ay(t)+1} \quad \text{(i=1,...,n)} \qquad (4)$$

A problem of the process we have considered is that the population may grow without limit. We can obtain the same dynamics if we assumed a constant population. Binmore (1991a) for instance obtains the same equations in a model in which the size of the population is kept fixed. After reproduction some individuals die (food is a scarce resource). All individuals, including newborns, can die.

The continuous time replicator dynamics (3) is obtained by taking the limit as $\tau \to 0$ in (4)

$$\lim_{\tau \to 0} \frac{x_i(t+\tau) - x_i(t)}{\tau} = x_i(t)(e_i^T Ax(t) - x(t)^T Ax(t)) \quad \text{(i=1,...,n)}$$

If we consider non-overlapping generations i.e. all members from the current generations die and are substituted by newborn individuals, we have that $n_i(t+1) = r_i n_i(t)$ and $N(t+1) = \sum r_i n_i$

$$x_i(t+1) = \frac{\tau n_i(t) r_i}{\sum r_i n_i(t)} = \frac{x_i(t) r_i}{\sum r_i x_i(t)} \quad \text{(i=1,...,n)}$$

which can be rewritten after substituiting r_i by its estimate as

$$x_i(t+1) - x_i(t) = x_i(t) \frac{(e_i^T Ax(t) - x(t)^T Ax(t))}{x(t)^T Ay(t)} \quad \text{(i=1,...,n)} \qquad (5)$$

Van Damme (1987) obtains equation (3) from the straightforward approximation $x_i(t+1) - x_i(t) \sim \dot{x}_i$ and disregarding the denominator. This changes the length but not the direction of the vector field. A general expresion which encompasses (4) and (5) is

$$x_i(t+1) - x_i(t) = x_i(t) \frac{(e_i^T Ax(t) - x(t)^T Ax(t))}{x(t)^T Ay(t) + C} \quad \text{(i=1,...,n)} \qquad (6)$$

with C ($C = 0$ in the nonoverlapping case and $C = 1$ in Binmore (1991a) and Taylor and Jonker (1978)) interpreted as the 'common background fitness' (Hofbauer and Sigmund (1988). If a large enough constant C is added to all the payoffs, the strategies make a small contribution to the relative fitness and therefore the different strategies will grow slowly. The larger C the better the discrete time replicator dynamics (6) approximate the continuous one (3).

In 2×2 games, if the mixed strategy p (in a monomorphic population) is stable then so is the corresponding polymorphism (where only pure strategies are allowed). The same is not true with games with more than 2 strategies (see Taylor and Jonker (1978))

The following results hold for $n \times n$ symmetric games (see Van Damme (1987)):

(i) If p is a Nash equilibrium, then p is a fixed point of the replicator dynamics, but the converse is not true.

(ii) if p=e_i for $i = 1, 2...n$, then p is stationary in the replicator dynamics (any monomorphic population is a fixed point).

(iii) If p is an ESS, then the corresponding polymorphism is asymptotically stable in the replicator dynamics.

(iv) If p is a mixed ESS, then the corresponding polymorphism is globally stable in the replicator dynamics.

(v) If p is a local attractor the replicator dynamics, then p is a Nash equilibrium but the converse need not be true.

An interesting question is whether evolution wipes out irrational behaviour from the population. Samuelson and J.Zhang (1992) show that the population shares of non-rationalizable pure strategies converge to zero along any 'interior' dynamic path in the continuous replicator dynamics. [6] If convergence to an interior state occurs, then it is a Nash equilibrium (Nachbar (1990)). Dekel and S.Scotchmer (1992) show that if players can only inherit pure strategies, then strategies which are never best reply can survive replicator dynamics. The result is based on the fact that better than average strategies grow under replicator dynamics even though they might never be best reply strategies. They show therefore that evolution need not select the fittest strategy.

Example 2.1 Let us consider the Hawk-Dove game represented in Figure 1. Let x be the proportion of players using the hawk strategy. The replicator dynamics in this case is:

$$\dot{x} = x(1 - x)\frac{C}{2}(\frac{V}{C} - x)$$

It is easy to see that, when $C > V$ $\dot{x} > 0$ (< 0) if $x < V/C$ $(x > V/C)$ and when $C < V$ $\dot{x} > 0$ for all x.

[6] Cabrales and Sobel (1992) extend this result to discrete time replicator dynamics (with C=0) when only a small fraction of players reproduce each period.

The following figure represent the dynamics in these two cases. The filled dots are the stable points. The white dots are stationary points that are not stable (nor Nash equilibria).

Phase diagram when $V > C$

Phase diagram when $V < C$

Figure 2.

In the first case, the dynamics converge to a stable polymorphism which corresponds to the mixed equlibrium of the game. In the second case the only stable population consists of only hawks.

Example 2.2 Let us consider the replicator dynamics in the Hawk-Dove game when it is played by members of two different populations.

The pair (x, y) represents the proportion of players from each population using the hawk strategy. The replicator equations are given by:

$$\dot{x} = x(1-x)\frac{C}{2}(\frac{V}{C} - y)$$

$$\dot{y} = y(1-y)\frac{C}{2}(\frac{V}{C} - x)$$

Figure 3 represents the phase diagram of such dynamics for $C > V$.

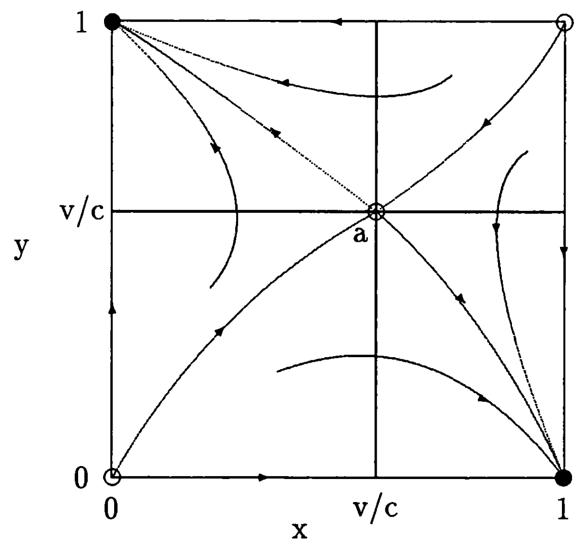

Figure 3.

The mixed Nash equilibrium a= $(V/C, V/C)$ is not a stable point of the replicator dynamics; it is saddle. Almost all orbits converge to one of the two equilibria in pure strategies, namely (0,1) and (1,0).

Example 2.3 An example of asymptotic convergence to a Nash equilibrium which is not an ESS is the example of Taylor and Jonker (1978) reported in section 1.2. The following figure is a representation of such a example. At the vertices only one pure strategy is played.

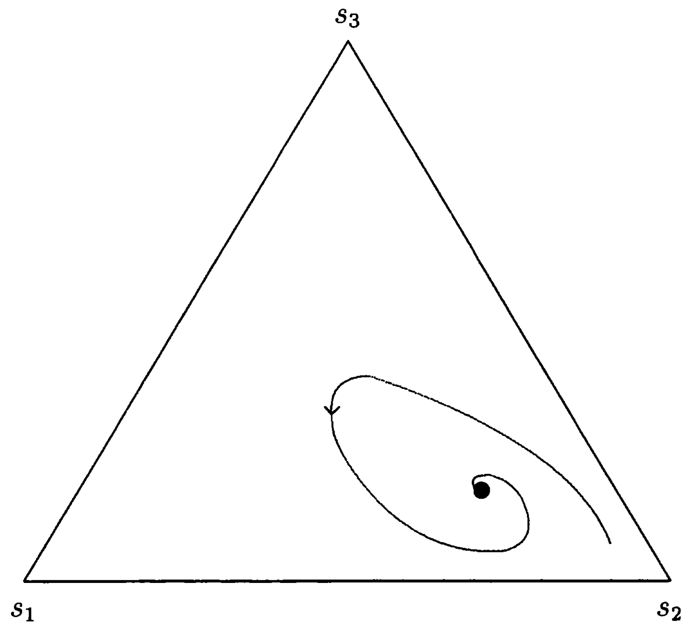

Figure 4.

A feature of this example, which is common to all interior equilibria that are asymptotically stable but not ESS, is that they are always foci. The equilibrium is reached after a period in which some cycling behaviour is observed, due to the higher growth rate of some pure strategy.

In many circumstances, the deterministic replicator dynamics does not help in selecting among different Nash equilibria.

All strict Nash equilibria are asymptotically stable points of the replicator dynamics.

Example 2.4 Let us consider the following coordination game with two strict equilibria (A,A) and (B,B):

	A	B
A	a, a	0, 0
B	0, 0	b, b

Figure 5

Let x be the proportion of player using strategy A. Under replicator dynamics,

$$\dot{x} = x(1-x)(a+b)(x - \frac{b}{a+b})$$

for which the phase diagram is shown in Figure 6.

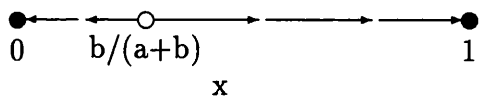

Figure 6.

The population composition that will be achieved through Darwinian selection will depend on the initial state of the population. Either (A,A) or (B,B) will be selected. The mixed equilibria (a/(a+b),a/(a+b)) is a repellor. As we will see in the next section, the results obtained from the deterministic model change when stochastic effects are taken into account.

3.1.2 Stochastic replicator dynamics

Foster and Young (1990) first introduced stochastic differential equations in evolutionary biology (in symmetric games). The main claim of Foster and Young is that neither ESS nor the more general concept of an attractor are the right concepts when stochastic effects are taken into account. Their main criticism of the previous approach is that only one-shot mutations are considered. The stability is tested with the introduction of mutants that either die out and the system goes back to the equilibrium (if it is asymptotically stable) or the population is destabilized and a new state is reached (in the case of convergence). When continuous mutations are considered it may be possible that their effects accumulate and even push the system away, from asymptotic attractor. They therefore stress *stochastically stable sets* rather than absorbing sets.

Due to the stochastic nature of the postulated dynamics, we cannot say that the system will be with certainty in an absorbing state (or set) but probabilistic statements are possible. The route followed by Foster and Young (1990) is to characterize the limiting behaviour of the stochastic system when the variance of the noise tends to zero. The idea is to find the probability for the system to be near any given state and to characterize the limiting distribution. This will put weight only on the stochastically stable equilibria. The main result in their paper is that the stochastic model has an ergodic distribution, i.e, asymptotic distributions that are independent of the initial conditions.

As pointed out by Foster and Young (1990), biological models such us that introduced by Taylor and Jonker (1978) are inherently stochastic. The stochastic forces come from different sources: variability of payoffs (fitnesses), randomness due to the matching process, background mutation and immigration. In Foster and Young's model, the aggregate noise is approximated by a continuous-time, continuous-space Wiener process **W** which enters additively into the continuous-time replicator equations (3) to give the following equation

$$dx(t)_i = x(t)_i(e_i^T Ax(t)) - x(t)Ax(t))dt + \sigma d(\mathbf{W}(t)))_i \text{ i=1,...,n} \qquad (7)$$

where **W** is a continuous, white-noise process with zero mean and unit variance covariance matrix. The first term on the right hand side represents the drift, it is the standard replicator dynamics. They study the asymptotic properties of the system as the variance σ^2 of the noise becomes arbitrarly small.

Definition 3 (Stochastically stable state) *The population vector x^* is stochastically stable if, as $\sigma \to 0$, the limiting density assigns positive probability to every small neighbourhood of x^*.*

In the coordination game of Figure 5, the stochastically stable states are either 1 or 0 depending on whether $a > b$ or $b > a$. The intuition behind this result is that when $a > b$ the basin of attraction for 1 (everybody playing A) is larger ($x > b/(a+b)$). It is possible to go from 1 to 0 and from 0 to 1, because the noise pushes the system away from the attractors, but the selection pressure towards 1 is greater. It is crucial for Foster and Young's result that the boundaries be reflecting rather than absorbing. If it were not the case, the system as the muations tends to dissappear will eventually hit the boundaries because some strategy dies off. Once this happens the process can never leave the boundaries again. This problem is solved by restricting the process to the interior of the state space.

We can understand the selection pressure as a a river with two different branches which flow into two different lakes (the asymptotically stable points). The stochastic effects are the forces that push a boat upstream, fighting against the strength of the water which follows its natural course. The longer the distance to go upstream (the length of the branch) and the greater the strength of the flow (the relative payoffs) the more difficult it is to go from one lake to the other. The state of the system, in this parable, is the place where the boat is at any time. We can therefore define a probability measure over the different locations. The weight of a particular state in the measure is the probability of finding the system at that state. As the noise vanishes, it becomes more and more difficult for the boat to swim upsteam.

The densitity function concentrates on either lake, or on both, depending on the strength of the flow and on the length of the branches. Figure 7 is a representation of the story described above. The rest points of the deterministic dynamics are the two lakes (filled dots) and the state which joints the two branches (white dot). The latter is not stable, any mutation will push the boat into any of the branches and into the corresponding lake.

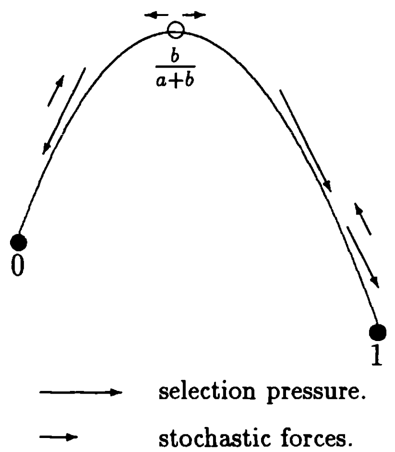

Figure 7.

Foster and Young (1990) show that the stochastically stable states are not necessarily evolutionarily stable.

Related work of Fudenberg and Harris (1992), studies the evolution of the continuous-time replicator dynamics in 2×2 symmetric games, when the payoffs are subject to aggregate shocks. The main difference with the previous paper is that Foster and Young (1990) add the stochastic term to the continuous replicator dynamics for the population shares (x_i)[7] (see equation (7)) while Fudenberg and Harris add the term to the expected payoffs in the equations reflecting the evolution of the sizes of the subpopulations using the different strategies and obtain the corresponding replicator dynamics for

[7]The share of strategy i is x_i. If the population is of size R, the members of the population playing i are of size $r_i = Rx_i$

population shares. A problem of the model is that the population may grow without limit. The shocks on each strategy i are independent of the strategy with which it is matched. Shocks on different strategies are independent. The main result is that the model does not always have an ergodic distribution, and the asymptotic behaviour can depend on the initial conditions. For example, in the coordination game (Game 2) as the variances (σ_A^2 and σ_B^2) go to zero the probability that the system converges to $x = 1$ ($x = 0$) goes to one if the process starts in $x > b/(a+b)$ ($x < b/(a+b)$. When also mutations are introduced in the model of Fudenberg and Harris, the dynamic process is ergodic for any positive rate of mutations. In the coordination game the equilibrium selected coincides with the stochastically stable equilibrium of Foster and Young (1990). Cabrales (1993) extends the model developed by Fudenberg and Harris (1992) to n-player games that are not necessarily symmetric and shows that strictly dominated strategies have little asymptotic weight.

3.2 Best-reply dynamics

In the replicator dynamics more than one strategy may be increasing their shares in the population. Those strategies which fare better than the 'average' increase their frequencies, and viceversa. In this section we consider the 'best reply dynamics'. Under this dynamics, a strategy will increase its frequency in the population only if it is a best reply to the current distribution of strategies in the population. In the first part we introduce the deterministic best reply dynamics. We will compare them with the replicator dynamics by means of few simple examples. As for the replicator dynamics we will consider both symmetric and asymmetric games. In the second section we will introduce the models of Kandori et al. (1993) and Young (1993a). The two models have a finite state space and the dynamics can be represented by a Markov chain. When ongoing mutations are introduced the right equilibrium concept to focus on is, as in Foster and Young (1990), long run equilibrium rather than absorbing state.

3.2.1 Deterministic Best Reply Dynamics

We consider two-person, symmetric games. Let us consider a situation in which, in each time interval the game is played $1/\tau$ $(0 < \tau \leq 1)$ times. Each time the game is played $\alpha\tau$ $(0 \leq \alpha \leq 1)$ players are given the chance of revising their strategies. Those who change act myopically choosing the best reply to the current state of the system. When more than one best reply exist, each is chosen with equal probability.

Let $BR(x) \subseteq S$ be the set of strategies which are best replies to the state x and let $|BR(x)|$ be its cardinality.

The dynamics are represented by the following difference equations.

$$\forall i \in BR(x(t)) \quad x_i(t+\tau) = x_i(t) + \frac{\tau\alpha}{|BR(x(t))|}(1 - \sum_{k \in BR(x(t))} x_k(t)) \quad (8)$$

$$\forall j \notin BR(x(t)) \quad x_j(t+\tau) = x_j(t) - \tau\alpha x_j(t) \quad (9)$$

where α measures the population responsiviness.

Equation (8) represents the dynamics of those strategies which are a best reply to the current state. The first term on the right hand side is the fraction of players who are already using a best reply and do not change their strategies. The second term is the proportion of players who were playing some strategy which was not a best reply to $x(t)$ and change to i. Only a proportion $\alpha\tau$ of the players who are adopting a strategy which is not a best reply to the current state of the system have the opportunity of selecting a new strategy. We have assumed that they select any of the best replies $i \in BR(x(t))$ with the same probability $1/|BR(x(t))|$. Equation (9) describes the dynamics for the strategies which are not best replies to the current state. All those player who are given the chance of revising their strategies (a proportion $\alpha\tau$) move to any of the best replies.

It is interesting to notice that the strategies do not depend on the payoffs, except insofar as these determine the best reply correspondence. When $\tau = \alpha = 1$ we have what is commonly known as 'discrete-time best-reply dynamics'.

With $\tau = 1$ and $\alpha < 1$ we have what we shall call 'slow best reply-dynamics' (in discrete time). We obtain the 'the continuous-time best-reply dynamics' by making $\tau \to 0$

$$\forall\, i \in BR(x) \qquad \dot{x}_i(t) = \frac{\alpha}{|BR(x(t))|}(1 - \sum_{k \in BR(x(t))} x_k(t)) \tag{10}$$

$$\forall j \notin BR(x(t)) \qquad \dot{x}_j(t) = -\alpha x_j \tag{11}$$

If we compare these dynamics with the replicator dynamics (3) we observe that only the best-paying strategies grow in the population. The two dynamics are the same, for interior points, in the 2×2 case, but with different speeds. The difference is that the rate of adjustment is independent of the relative payoffs for best-reply dynamics. Another difference between the two dynamics is that in the 'best reply' dynamics the stationary points are necessarily Nash equilibria. The same is not true for the replicator dynamics where any distribution in which only a pure strategy is played is stationary.

When $BR()$ is a singletone we obtain the following solution to the differential equations (10)-(11)

$$\text{for } i = BR(x(t)) \qquad x_i(t) = 1 - a_i e^{-\alpha t}$$

$$\forall j \neq BR(x(t)) \qquad x_j(t) = a_j e^{-\alpha t}$$

where a_k is a constant of integration. It follows that:

$$\frac{x_j(t)}{x_k(t)} = \text{constant} \;\; \forall j, k \neq BR(x(t)) \tag{12}$$

this implies that the relative proportions of those strategies which are not best replies remain unchanged.

Example 2.5 Let us consider for example the 'Rock-Scissors-Paper' game:

$$A = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}$$

We can represent the state space in an equilateral triangle. Each vertex corresponds to the state in which all players use the same pure strategy. The 'Rock-Scissors-Paper game' has a unique equilibrium p^*=(1/3,1/3,1/3).

Figure 8 represents the state space and the best reply regions.

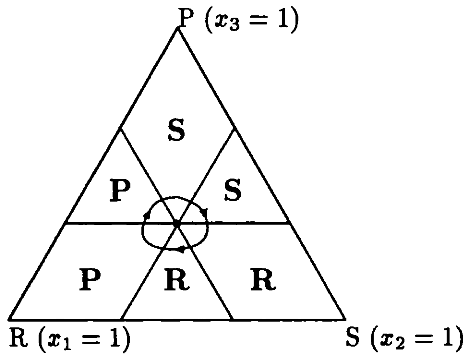

Figure 8.

This game is an example of non convergence in the 'replicator dynamics'. Under replicator dynamics, $x_1x_2x_3$ is a constant of motion (see Hofbauer and Sigmund (1988)). The dynamics are represented by closed orbits around the equilibrium point (see Figure 8).

Let us consider the continuous, best-reply dynamics (10)-(11). Condition (10) implies that the orbits are straight lines pointing at the vertex in which only the best reply to the current state is played.

Figure 9 represents the continuous time, best reply dynamics for 'Rock-Scissors-Paper'. The system converges from any point to the mixed equilibrium.

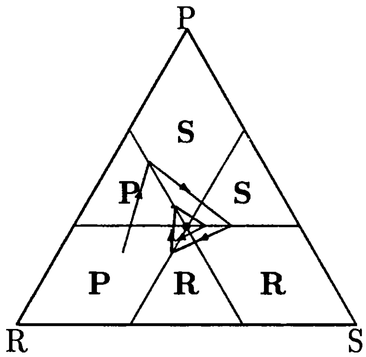

Figure 9.

If we consider the discrete time dynamics (6)-(7) with $\alpha = \tau = 1$ from any interior point the system converges to a cycle that 'jumps' (clockwise) from one vertex to another clockwise.

We now consider the best reply dynamics when a game is played by two populations. Let x(t) be the state at time t of population $\mathcal{X}$. We denote by y(t) the state of population $\mathcal{Y}$. We assume, as in the one-population case, that in time interval τ a proportion $\tau\alpha_{\mathcal{X}}$ ($\tau\alpha_{\mathcal{Y}}$) is given the oportunity of changing its strategy. Those who do change their strategies move to a best reply to the current state of the system. The discrete time 'best reply dynamics equations are given by:

$$\forall i \in BR(y(t)) \quad x_i(t+\tau) = x_i(t) + \frac{\tau\alpha_{\mathcal{X}}}{|BR(y(t))|}(1 - \sum_{k \in BR(y(t))} x_k(t)) \tag{13}$$

$$\forall j \notin BR(y(t)) \quad x_j(t+\tau) = x_j(t) - \tau\alpha_{\mathcal{X}} x_j(t) \tag{14}$$

$$\forall i \in BR(x(t)) \quad y_i(t+\tau) = y_i(t) + \frac{\tau\alpha_{\mathcal{Y}}}{|BR(x(t))|}(1 - \sum_{k \in BR(x(t))} y_k(t)) \tag{15}$$

$$\forall j \notin BR(x(t)) \quad y_j(t+\tau) = y_j(t) - \tau\alpha_{\mathcal{Y}} y_j(t) \tag{16}$$

The interpretation of equations (13) and (15) and (14) and (16) is similar to the interpretations offered for equations (6) and (7), respectively. The main difference is that each populations responds to the state of the other. The corresponding continuous time dynamics are given by:

$$\begin{aligned}
\forall i \in BR(y(t)) \quad \dot{x}_i &= \frac{\alpha_{\mathcal{X}}}{|BR(y(t))|}(1 - \sum_{k \in BR(y(t))} x_k(t)) \forall i \in BR(y(t)) \\
\forall j \notin BR(y(t)) \quad \dot{x}_j &= -\alpha_{\mathcal{X}} x_j(t) \\
\forall i \in BR(x(t)) \quad \dot{y}_i &= \frac{\alpha_{\mathcal{Y}}}{|BR(x(t))|}(1 - \sum_{k \in BR(x(t))} y_k(t)) \\
\forall \notin BR(x(t)) \quad \dot{y}_j &= -\alpha_{\mathcal{Y}} x_j(t)
\end{aligned}$$

When $BR()$ is a singletone, the solution to these differential equations is

$$\begin{aligned}
x_i(t) &= 1 - a_i^{\mathcal{X}} e^{-\alpha_{\mathcal{X}} t} & i = BR(y(t)) \\
x_j(t) &= a_j^{\mathcal{X}} e^{-\alpha_{\mathcal{X}} t} & \forall j \neq BR(y(t)) \\
y_i(t) &= 1 - a_i^{\mathcal{Y}} e^{-\alpha_{\mathcal{Y}} t} & i = BR(x(t)) \\
y_j(t) &= a_j^{\mathcal{Y}} e^{-\alpha_{\mathcal{Y}} t} & \forall j \neq BR(x(t))
\end{aligned}$$

The relative proportions of those strategies which are not best replies, and therefore are decreasing, remains constant. When the resposiveness in both populations is the same, the trajectories are straight lines pointing at the vertex where only the best reply is played. If we allow for different degrees of responsivenes in the different populations we have that:

$$\frac{z_j(t)}{z_k(t)} = \text{const} \quad \forall j, k \neq BR(w(t)) \;\; z, w \in \{x, y\} \;\; z \neq w$$

$$\frac{x_j(t)}{y_k(t)} = \text{const } \; e^{-(\alpha_{\mathcal{X}} - \alpha_{\mathcal{Y}})t} \;\; \forall j \neq BR(y(t)) \;\; and \;\; \forall k \neq BR(x(t))$$

Example 2.6 Figures 11 and 12 represent different continuous dynamics in 'matching pennies' played by different populations.

	HEADS	*TAILS*
HEADS	$1, -1$	$-1, 1$
TAILS	$-1, 1$	$1, -1$

Matching Pennies.

Figure 10.

In Figure 11 we have represented the Best-reply dynamics for Matching Pennies. The proportion of players in population $\mathcal{X}$ ($\mathcal{Y}$) using Heads is x (y). This dynamics converge to the unique (mixed) equilibrium of the game. In the long run equilibrium, half of each population will play Heads and the other half Tails.

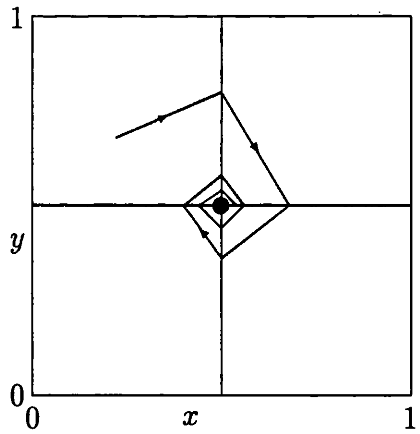

Figure 11. Best-reply dynamics

Figure 12 represents the continuous Replicator dynamics. The dynamics are characterized by closed orbits around the mixed equilibrium.

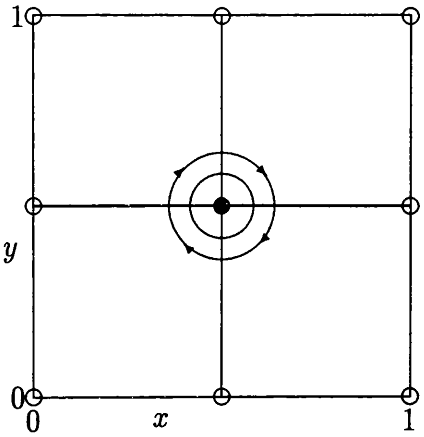

Figure 12. Replicator Dynamics

3.2.2 Stochastic Best Reply Dynamics.

Until now we have only considered dynamic systems with a continuous state space, which we have represented as a simplex. In this section we introduce two models with discrete state space. The models due to Kandori et al. (1993) and Young (1993a) are important recent contributions two the literature of equilibrium selection in games. After introducing the main motivation of the models we will, by means of very simple examples, illustrate the main idea behind the equilibrium selection which is the same in both models.

Kandori, Mailath and Rob's dynamics. Kandori et al. (1993) introduce a discrete time system with a finite population, where each individual 'mutates' from one strategy to another with positive probability. There are three hypotheses in their model: (*i*) The *inertia hyphotesis* implies that not all players simultaneously and instantaneously adjust to the environment. The assumption can be justified by the existence of adjustment costs and imperfect knowledge about the relation between payoffs and strategies. (*ii*) The *myopia hypothesis* implies that people do not take into account the long run implications of their choices. Finally, (*iii*) the *mutation* (experimentation) hypothesis, according to which players, from time to time, play an arbitrary strategy. This last hypothesis has a nice economic interpretation:

there is a small probability that a player exits the population and is replaced by a newcomer who knows nothing about the game and chooses a strategy at random.

Their main result is that the model has an ergodic distribution whose limit as the probability of mistakes goes to zero is concentrated at the risk-dominant equilibrium in 2×2 coordination games. This result is independent of the details of the underlying deterministic adjustment process provided that it satisfies a *Darwinian property* which implies that the best strategy must be better represented in the following period. The dynamic system considered by Kandori et al. (1993) can make discrete jumps. When the mutation rate is small the system spends most of its time at a strict Nash equilibria. Sometimes enough individuals mutate simultaneously to shift the state of the system to the basin of attraction of the other equilibrium. The limit distribution is shown to depend on the relative probabilities of the shifts from one equlibrium to the other, which only depends on the sizes of the basins of attraction.

In this section we use a simple example to illustrate the main ideas of Kandori et al. (1993). We will consider a simple bargaining game which has only one symmetric equilibrium (in mixed strategies) and two asymmetric equilibria in pure strategies. We will see that, in this case, the equilibrium selection depends on the speed of adjustment.

Consider a finite population of size N. We could consider either of the two following matching technologies: (i) Each player is matched,in each period $t = 0, 1, ..$, with each of the remaining players to play the bargaining game represented in Figure 13. A player will be part of $N-1$ encounters. (ii) Each player is randomly matched with another. The payoffs have the alternative interpretation of average payoffs in case (i) or expected payoffs in case (ii). We shall follow the first interpretation.

	LOW	*HIGH*
LOW	b, b	b, a
HIGH	a, b	0, 0

$$0 < b < a$$

Figure 13.

This game has two asymmetric Nash equilibria in pure strategies ,(Low,High) and (High,Low) and a symmetric equilibrium in mixed strategies. In the mixed equilibrium Low is played with probability $p = b/a$. At the beginning of each period, players choose a pure strategy which they will play in all the encounters.

We define a state of the system by z_t, which denotes the number of players adopting the strategy Low. The state space is, therefore

$$Z = \{0, 1, \ldots, N\}.$$

The average payoffs to the different strategies , in state z_t are:

$$\begin{aligned} \pi_{Low}(z_t) &= b \\ \pi_{High}(z_t) &= \frac{z_t}{N-1}a \end{aligned}$$

We assume that the players' choice of strategy moves towards the best reply strategy to the current state of the population. It follows that the strategy with the highest payoff will be more frequent in the future. We will assume the following *Darwinian property*:

$$sign(z_{t+1} - z_t) = sign(\pi_{Low}(z_t) - \pi_{High}(z_t))$$

Let $\mathcal{F}$ describe the dynamics of the system. The state tomorrow depends on the state today, $\mathcal{F}(z_t) = z_{t+1}$.

We will consider the following extreme cases:

- the *best reply dynamics* ($\tau = \alpha = 1$) given by the following rule:

$$\mathcal{F}_B(z_t) = \begin{cases} N & , \quad if\ \pi_{Low}(z_t) > \pi_{High}(z_t) \\ z_t & , \quad if\ \pi_{Low}(z_t) = \pi_{High}(z_t) \\ 0 & , \quad if\ \pi_{Low}(z_t) < \pi_{High}(z_t) \end{cases}$$

- the *Slow besr-reply dynamics* with only one player revising his strategy:

$$\mathcal{F}_S(z_t) = \begin{cases} z_t + 1 & , \quad if\ \pi_{Low}(z_t) > \pi_{High}(z_t) \\ z_t & , \quad if\ \pi_{Low}(z_t) = \pi_{High}(z_t) \\ z_t - 1 & , \quad if\ \pi_{Low}(z_t) < \pi_{High}(z_t) \end{cases}$$

Furthermore, we assume that with probability ϵ each player mutates, changing his intended strategy. The plausible story behind the mutations is that with probability 2ϵ, each player dies and is replaced by a new player who knows nothing about the game. The newcomer chooses either strategy with probability 1/2.

We have the following timing:

$$z_t \longrightarrow \boxed{\text{SELECTION}} \longrightarrow \mathcal{F}(z_t) \longrightarrow \boxed{\text{MUTATIONS}} \longrightarrow z_{t+1}$$

Figure 14.

The state z_t is observed by everybody. Players decide what strategy to use. The 'intended' strategies are described by $\mathcal{F}(z_t)$. Some players die and are replaced by newcommers who choose their strategies at random. The new state z_{t+1} s observed.

This structure defines the following stochastic difference equation:

$$z_{t+1} = \mathcal{F}(z_t) + x_t - y_t, \tag{17}$$

where x_t and y_t have binomial distributions:

$$x_t \sim Bin(N - \mathcal{F}(z_t), \epsilon)\ and\ y_t \sim Bin(\mathcal{F}(z_t), \epsilon).$$

The x_t on the right hand side of equation 11, are those players who having decided to play High, die and are replaced by a newcomer who plays Low. Those newcomers playing High and which substitute players who had decided to play Low are y_t.

The difference equation (17) defines a Markov chain on the finite state space Z, with transition probabilities,

$$p_{ij} = Prob(z_{t+1} = j | z_t = i),$$

with $P(\epsilon) = [p_{ij}]$ being the Markov matrix.

With $\epsilon > 0$, all the elements of $P(\epsilon)$ are strictly positive. In this case the Markov chain has a unique stationary distribution (see Kandori et al. (1993)) $\mu = (\mu_0, \mu_1, \ldots, \mu_N) \in \Delta_N$ with $\Delta_N \equiv \{q \in R^{N+1} | q_i \geq 0 \text{ for } i = 0, 1, \ldots, N \; and \; \Sigma_i q_i = 1\}$ satisfying

$$\mu P(\epsilon) = \mu \tag{18}$$

We are interested in the long run behavior of the system when the probability of mutation is small. The aim is to find the *limit distribution* when it exists;

Definition 4 *A limit distribution μ^* is defined by*

$$\mu^* = \lim_{\epsilon \to 0} \mu(\epsilon).$$

Those states which have positive weight in the limit distribution are said by Kandori et al. (1993) to be the long run equilibria.

When $\epsilon = 0$, the system of equations defined by (18), with P=P(0), may have multiple stationary distributions. This occurs, for example, when the game has multiple symmetric Nash equilibria in pure strategies. The distribution that puts weight 1 on the state in which all players play the same pure strategy Nash equilibrium is always a stationary distribution of $P(0)$. Unlike the case with positive mutation rate, the long run distribution depends on the initial state.

Kandori et al. (1993) offer a technique to compute the long run distribution. This technique uses trees connecting all the states in the state space Z. Let us consider a state z and a directed graph on Z such that, each state except z has a unique successor and there are no closed loops. This collection of arrows connects the elements of Z in such a way that every state in $Z \setminus \{z\}$ is the initial point of one and only one arrow, and from any state in $Z \setminus \{z\}$ there is a sequence of arrows leading to z. Such a directed graph is called a z–tree. We associate to each arrow an integer denoting the number of mutations needed to reach one state from the other in a period. The idea behind Kandori et al. (1993) characterization of the long run distribution is to compute all possible z–trees for all states z in Z.

A transition between two connected states in the tree can occur as result of two possible forces:

> the *selection mechanism* which facilitates the transitions without the need of mutations; in this case the number associated to the arrow connecting both states will be zero.
>
> the *mutations*; in this case the transition will be associated with positive integer, corresponding to the number of mutations needed to reach one state (the end of the arrow) from the other (the origin of the arrow).

Kandori et al. (1993) show that the long run distribution will put positive weight on those states whose z–tree has the minimum cost of transition, understood as the total number of mutations needed to construct the tree [8].

Example 2.7 Let us consider a population of size $N = 5$ and the bargaining game of Figure 13, with $a = 2/3$ and $b = 1/3$. The mixed equilibrium is $p = 1/2$. The state space is $Z = \{0, 1, 2, 3, 4, 5\}$.

[8]It is possible for a given state two construct several z-trees among which the one with the smallest number of mutations is considered.

Let us consider the 'slow' selection mechanism $\mathcal{F}_S(z_t)$. When nobody has played Low, $z_t = 0$, the best reply is High. As only one player at a time is allowed to revise his strategy $\mathcal{F}_S(0) = 1$. There are needed 0 mutations to go from 0 to 1. The same argument applies for the transition between 1 and 2. When in the current state of the system there are 3 or more players playing Low the best reply is High. Given the 'slow' adjustment process, from any of these states the system moves to a new state with one player less playing Low. States 2, 3 and 4 can be reached from 3, 4 and 5, respectively, without the need of any mutation.

Figure 15 represents a 2-tree. The numbers in squares represent the states. Numbers in brackets report the number of mutations needed to have the transition represented by the arrow. It is a 2-tree because state 2 can be reached from any other state and from all states except 2 there is one and only one way out (arrow). The resistance associated to this particular 2-tree is 0. All the transitions need 0 mutations.

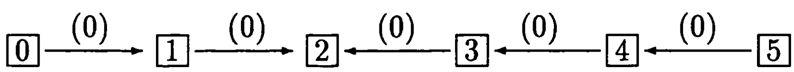

Figure 15. 2–tree

Let us consider now the 0-tree of Figure 16. The transtions out of 1, 3, 4 and 5 are as before. How can the system go from 2 to 0? If the state of the system is 2, all players are indifferent between playing High and Low, $\mathcal{F}(2) = 2$. To move from 2 to 0 we need that 2 players who intended to play Low mutate and play high. This will happen with probability ϵ^2. The easiest way to go from 2 to 0 is through 2 simultaneous mutations. All the other transitions take place under the 'selection mechanism'.

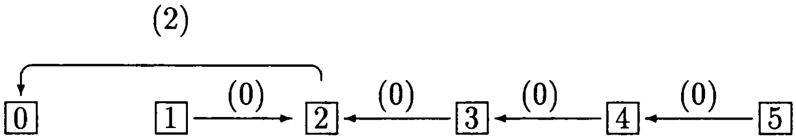

Figure 16. 0–tree

We can construct trees for all the states. The 2–tree we have built is the one that minimizes the number of mutations for all possible trees and all states and, therefore, 2 is the only long run equilibrium of the system.

Let us now consider the best reply dynamics $\mathcal{F}_B(z_t)$. Under this selection mechanism all players simultaneouly change to the best reply strategy. In this case the states with the smallest tree are 0 and 5. Consider the 5-tree in Figure 17. The transitions between 5 and 0 need 0 mutations, $\mathcal{F}(0)_B = 5$ and $\mathcal{F}_B(5) = 1$. The same is true for the arrows out of 1, 3 and 4 ($\mathcal{F}_B(1) = 5$, $\mathcal{F}(3) = \mathcal{F}(3) = 0$).

To connect state 2 with any other state we need at least one mutation ($\mathcal{F}(2) = 2$). Let us assume, for example, that one player who had decided to play High dies and is substituted by a new player who plays Low. This will happen with probability ϵ. We can draw an arrow from 2 to 3 which has one mutation associated to it. From 3 it is possible to go to 5 without the need of further mutations.

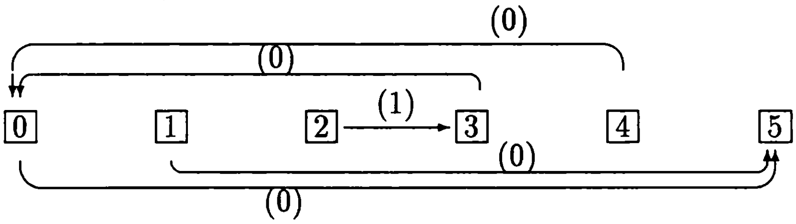

Figure 17. 5–tree

With similar argument we can build the following 0-tree.

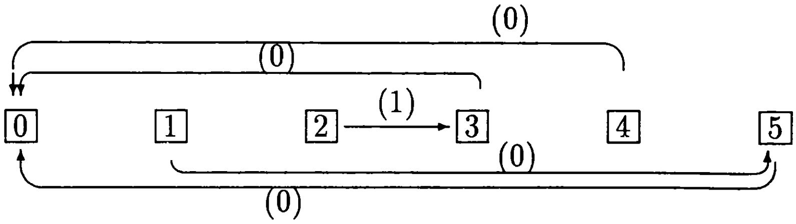

Figure 18. 0–tree

Any z-tree for $z = \{1, 2, 3, 4\}$ needs more than 1 mutation and therefore the long run distribution will put weight on 0 and 5 only.

The equilibria which appear in the long run distribution are those from which more mutations are needed to scape. Mutations are the driving force of the equilibrium selection results. The system can only move from one equilibrium to another once a sufficient number of mutations have occured. The most unsatistactory feature of the model is that the selection is driven, in many cases, by very unlikely events. This fact implies that the economy can spend very long periods at a wrong equilibrium (see Ellison (1992)).

Evolution of conventions. In recent paper Young (1993a) explains the emergence and persistence of an equilibrium in n-person weakly acyclical games. These games have the property that from any initial choice of strategies, there exist a sequence of best replies, in which only one player changes his strategy, that lead to a strict, pure strategy Nash equilibrium.

The model of Young (1993a) explains the emergence of social conventions. Money is one of the standard examples used by economist to illustrate the importance of conventions. Other examples include the use of a language, driving on the right-hand side or using compatible computers. A common feature of conventions is that everybody prefers to adhere to the conventional behaviour when everybody else does. A convention is a rule that tells you how to act in certain common situations which might otherwise be ambiguous (Wärneryd (1990)). In other words, conventions solve the problem of multiplicity in coordination games.

How do conventions emerge in a world of uncoordinated players? In the model of Young (1993a) a convention emerge because players may coordinate by chance. In Young's model a n-person game is played once each period by n players drawn at random from finite, large populations. To rule out any possible of learning at the individual level or of building up a reputation, players are assumed to be different in each encounter. Although the individual player has no memory, the society has recorded evidence about the most recent plays of the game (truncated history). The m most recent plays are recorded. Each time the game is played a new n-tuple of strategies is recorded and the oldest element in the history is lost.

Each player forms his believes about the remaining $n-1$ players by looking at what similar opponents did in the recent past. A player has access to some limited information about past plays of the game. He can observe a random sample reporting how the game was played in some previous encounters. Each element in the sample, whose size may differ among players, contains $n-1$ strategies. A player does not observe what other players in the same role have done in the past. In a 2-person game, for instance, the row player has information about previous plays of the column players but not about other row players' strategies. Players are assumed to play a best reply to the sample they have drawn. The *adaptive play* just described defines a Markov Chain whose states are the truncated histories of play. This process is similar to fictitious player with the difference that in the latter the players base their decisions on the entire history and not only on a limited sample of the recent history.

An interesting result is that, in weakly acyclical games, adaptive play converges with probability one to a strict, pure strategy Nash equilibrium provided that samples are sufficiently incomplete and that players never make mistakes. An equilibrium, in this contex, is the m-repetition of the same n-tuple Nash equilibrium strategy profile, which is called a convention. The conventions are the only absorbing states of the system. The convention which is selected cannot be determined a priori. It will depend on the initial state and on the sampling processess (of players and samples). Once the system is in a convention, all players, by following what is conventional, behave optimally. The convention is established for ever. The convergence result follows from the fact that players may coordinate 'by chance' and if they do so often enough the process will eventually lock in to a convention. The result is driven by the fact that the finite memory allows the players to forget past miscoordinations and once a certain equilibrium has been played for long enough, being what everyone remembers, it becomes the conventional way of playing the game.

Example 2.8 Consider the 2-person game in Figure 19.

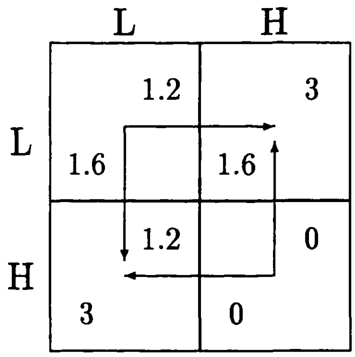

Figure 19.

The arrows show the acyclical structure of the game. The transition between (L, L) and (L, H) takes place because the column player moves to the 'best reply' to L. In any of the transitions only one player moves to the best reply.

Let us assume that both players (row and column) always draw samples of size 2. The history length is of 4.

The following sequence of four pairs is a history (state) of game.

$$\left\{ \begin{pmatrix} H \\ H \end{pmatrix}, \begin{pmatrix} H \\ H \end{pmatrix}, \begin{pmatrix} L \\ L \end{pmatrix}, \begin{pmatrix} H \\ L \end{pmatrix} \right\}$$

In each pair $\binom{i}{j}$ the upper term i represents a strategy used by a row player and the lower term j the demand of a column player. The row player will observe some of the lower terms and the column player some of the upper ones.

The following sequence of 4-period histories represents a possible evolution of the state of the system. The one-starred elements are drawn by the row player. The elements marked with the two stars are drawn by the column player. A state is obtained from the previous one by deleting the first element and adding in the last position a new pair. The entries in the new pair are best replies to the sample drawn from the previous state. For instance, the last element in the sequence corresponding to $(t+1)$ is obtained calculating the best replies to the starred elements in t.

$$\left\{\binom{L}{H}, \binom{H^{**}}{H}, \binom{L}{L^*}, \binom{H^{**}}{L^*}\right\}_t$$

$$\left\{\binom{H}{H}, \binom{L}{L}, \binom{H^{**}}{L^*}, \binom{H^{**}}{L^*}\right\}_{t+1}$$

$$\left\{\binom{L}{L}, \binom{H}{L^*}, \binom{H^{**}}{L^*}, \binom{H^{**}}{L}\right\}_{t+2}$$

$$\left\{\binom{H^{**}}{L}, \binom{H}{L^*}, \binom{H}{L^*}, \binom{H^{**}}{L}\right\}_{t+3}$$

$$\left\{\binom{H}{L}, \binom{H}{L}, \binom{L}{L}, \binom{H}{L}\right\}_{t+4}$$

The last state contains 4 repetitions of a Nash equilibrium. It is a convention. The following state will be again the same repetition of the Nash equilibrium, independently of the sample which is drawn. Such a state is an absorbing state and the system will remain there for ever. For notational convenience we will refer to the m repetition of the same pair $\binom{i}{j}$ as $\binom{i}{j}_m$

We could have selected different samples and converge to the other convention of the system

$$\left\{\binom{L}{H}, \binom{L}{H}, \binom{L}{H}, \binom{L}{H}\right\} = \binom{L}{H}_4$$

The problem of the indeterminacy of the equilibrium is solved by Young by introducing mistakes. When players have a small probability of playing an arbitrary strategy the system has no absorbing states. The mistakes constantly push the system away from the equilibrium. The system has a unique stationary distribution which, as the noise goes to zero, puts positive weight only on the stochastically stable states (see section).

Example 2.9 Let us assume that the society is in the convention $\binom{H}{L}_4$ in time t and that the column player player makes a mistake. He plays High although Low is the best reply to a sample containing all High's. The state of the system in $(t+1)$ will be

$$\left\{\begin{pmatrix}H^{**}\\L\end{pmatrix},\begin{pmatrix}H^{**}\\L^{*}\end{pmatrix},\begin{pmatrix}H\\L\end{pmatrix},\begin{pmatrix}H\\H^{*}_{m}\end{pmatrix}\right\}_{t+1}$$

Low is the best reply to a sample with 1 High and 1 Low drawn by the row player ($1.6 > 3/2$). If the mistake is drawn and no further mistakes take place then the following states will be

$$\left\{\begin{pmatrix}H\\L\end{pmatrix},\begin{pmatrix}H^{**}\\L\end{pmatrix},\begin{pmatrix}H^{**}\\H^{*}_{m}\end{pmatrix},\begin{pmatrix}L\\L^{*}\end{pmatrix}\right\}_{t+2}$$

The system, after a row player has made a mistake, has enter the basin of attraction of the convention $\binom{L}{H}_4$. Observe the following 'possible' sequence of states

$$\left\{\begin{pmatrix}H\\L\end{pmatrix},\begin{pmatrix}H\\H^{*}_{m}\end{pmatrix},\begin{pmatrix}L^{**}\\L\end{pmatrix},\begin{pmatrix}L^{**}\\L^{*}\end{pmatrix}\right\}_{t+3}$$

$$\left\{\begin{pmatrix}H\\H_{m}\end{pmatrix},\begin{pmatrix}L^{**}\\L\end{pmatrix},\begin{pmatrix}L\\L^{*}\end{pmatrix},\begin{pmatrix}L^{**}\\H^{*}\end{pmatrix}\right\}_{t+4}$$

$$\left\{\begin{pmatrix}L^{**}\\L\end{pmatrix},\begin{pmatrix}L\\L^{*}\end{pmatrix},\begin{pmatrix}L^{**}\\H^{*}\end{pmatrix},\begin{pmatrix}L\\H^{*}\end{pmatrix}\right\}_{t+5}$$

$$\left\{\begin{pmatrix}L\\L\end{pmatrix},\begin{pmatrix}L^{**}\\H^{*}\end{pmatrix},\begin{pmatrix}L^{**}\\H^{*}\end{pmatrix},\begin{pmatrix}L\\H\end{pmatrix}\right\}_{t+6}$$

$$\left\{\binom{L}{H},\binom{L}{H},\binom{L}{H},\binom{L}{H}\right\}_{t+7}$$

We can also compute the number of mistakes needed to go out from $\binom{L}{H}_c$. Let m_c and m_r be the number of mistakes made by column players and row players respectively. Assume that the state of the system is $\binom{L}{H}_4$. How many mistakes (Low's) should a sample of size 2 drawn by the row player contain for High being the best reply?

$$3m_c > 1.6 \quad \Rightarrow m_c = 2$$

How many mistakes (High's) should a sample of size 2 drawn by the column player contain for Low being the best reply?

$$1.2 > 3\frac{2 - m_r}{2} \quad \Rightarrow m_r = 2$$

It is easier to go out from $\binom{H}{L}_4$ than from $\binom{L}{H}_4$. As the mistake rate goes to zero, the latter infinetely more likely than the first and it will be the only stochastically stable convention.

Young's equilibrium selection is based on counting the number of mistakes needed to jump from one equilibrium into the basin of attraction of another one. The same criticism we made to the model of Kandori et al. (1993) applies to Young's model. The economy can spend very long periods at an equilibrium which is not stochastically stable. An advantadge of Young's dynamics with respect to those of Kandori et al. is that the number of mistakes needed to scape from one convention can be considerably reduced if some people sample only a small proportion of the memory. In this case it is easier that players start playing a non-conventional strategy. The time spent at the "wrong" equilibrium may be shortened.

4 Conclusions

Traditional game theory has focus on equilibrium states without providing satisfactory answers to the questions of how players know that an equilibrium will be played and which one in the case of multiplicity. The evolutionary approach has provided new answers to these questions stressing the importance of dynamic processes in the selection of equilibria. Other long run phenomena such as cycles and strange attractors turn out to be interesting predictions once the dynamic processes are fully specified.

The most traditional evolutionary approach to the equilibrium selection problem in games focused on evolutionarily stable states. The stability of such states was tested with the introduction of one shot perturbation. In this cases the equilibrium which was selected was dependent on the initial conditions. In this approach history was playing a very important role in the equilibrium selection. Recent work by Foster and Young (1990), Kandori et al. (1993) and Young (1993a) and (93b) has eliminated the path dependence by introducing continuous mutations. For this authors it makes only sense to study the probabilities of finding the system at a particular equilibrium. The main drawback of this last approach is that it neglects many equilibrium states that having zero mass in the long distribution maybe observed during very long periods. We think it is useful to make a further distinction between the "long run" and the "ultralong run" horizons. Binmore and Samuelson (1993a) define as "long run equilibria" all those equilibria where the system can spend very long periods and which typically depend on where the process started. We can consider the "long run equilibria" in the works of Kandori et al. (1993) and Young (1993a) as "ultralong run" phenomena which emerge when the role of history has been eliminated. It is only in this new time horizon where the concept makes sense.

Chapter 2

Are large windows efficient? Evolution of learning rules in a bargaining model

1 Introduction

A standard feature of bargaining games is their multiplicity of equilibria. In the early fifties, Nash proposed two different approaches to solve the multiplicity problem. In a first paper, Nash (1950) formulated a set of axioms (Invariance, Symmetry, Pareto Efficiency and Independence of Irrelevant Alternatives) which define properties that the outcome is required to satisfy, and which turn out to characterize a unique solution to the bargaining problem. Nash describes a 'bargaining problem' with all von Neumann and Morgenstern utility pairs representing the possible agreements available to the bargainers, and the utility pair that results in the case that no agreement is reached (the status quo). Nash shows that the unique solution satisfying the four axioms is given by the deal which maximazes the 'Nash product'. When the Symmetry axiom, which asserts that in a symmetric situation, neither player will accept an agreement giving him a lower utility than his opponent's, is abandoned, the other axioms together with the 'bargaining powers' associated to each player determine the 'asymmetric' Nash bargaining solution. In a second paper Nash (1953) obtains precisely the same bargaining outcome by analyzing a static bargaining model, the Nash Demand Game, in which the players simultaneously announce demands, which they receive if and only if the demands announced are compatible. The Nash Demand Game has many Nash equilibria (for example, any Pareto-efficient outcome is a Nash equilibrium). In order to select only one equilibrium, Nash required that an equilibrium be robust to perturbations involving some uncertainty about the location of the Pareto frontier of the negotiation set S. When the perturbed Demand Game approaches the unperturbed game (for which the Pareto boundary is known with certainty), all the Nash equilibria of the perturbed game converge on the Nash solution (see Binmore (1987a) and (1987b))

The Nash solution is supported by various strategic models. Nash (1953) himself, with the perturbed Demand Game, provides a noncooperative support to his axiomatic solution. Another noncooperative defense of the Nash solution is the model of Rubinstein (1982). In Rubinstein's model two players bargain over a pie of size 1. Each period, one of the players proposes a partition and the other player either rejects or accepts. In the latter case, the game finishes and the agreement is implemented. If the offer is rejected, the play goes to the next period, in which it is the rejecting player the one who makes the offer. The unique subgame–perfect equilibrium of the game converges to the Nash bargaining solution, when the time interval between subsequent offers approaches zero. Furthermore, the bargaining powers are determined by the players' time preferences. (See Rubinstein (1982) for the assumptions under which his result holds). In the case when the players are equipped with different discount factors it turns out that the most patient player enjoys a larger bargaining power.

More recently, Young (1993b) has provided a new interpretation of the Nash bargaining solution that still uses the Nash Demand Game, but leads to an asymmetric outcome in which the players have different bargaining powers. However the interpretation of these bargaining powers differs markedly from Rubinstein's interpretation. The approach followed by Young (1993b) is to embed the Nash Demand Game in an evolutionary framework in order to explain the emergence and persistence of one particular outcome. An interesting feature of the model is that it provides an appealing interpretation of the bargaining powers that characterize the asymmetric Nash bargaining solution. In Young's model the Nash bargaining game, over how to share a pie, is played repeatedly by members from two different large populations. Two players, one from each population, are randomly selected to play the game; players announce a share of the pie which they get if the demands are compatible, otherwise they get nothing. A crucial assumption of Young's model is that players learn how to play the game from the past behaviour of members from the other population. Players have access to a random sample, whose size may differ among players, drawn from the most recent demands which have been announced by the opponents. They take their sample as a predictor of the behaviour of the player they will face, and usually play a best reply to the empirical distribution derived from the sample. However, this behaviour is perturbed by rare 'mutations' so that sometimes the players make mistakes and announce a demand that is not a best reply to any possible sample.

An important feature of Young's dynamic process is that, in the limiting case when the mutation rate goes to zero, the system converges to a fixed Pareto-efficient division that corresponds to the asymmetric Nash bargaining solution, with the bargaining powers determined by the distribution of sample sizes. The model implies that, when all members of the same population observe a sample of the same size, it is the better informed population which gets the larger share of the cake. A less appealing result is obtained when people with different sample sizes coexist in the same population. In this case poorly informed players exert a negative externality on the better informed members of their population. The population's bargaining power is determined by the members who draw the smallest sample, even though such individuals may be present only in very small numbers.

Young obtains the Nash bargaining solution under very weak informational assumptions: Players only know their own preferences and a small sample of what happened in some recent past. This is what typically happens in many real world situations. Students seeking houses to rent know how much landlords have asked in the past for similar apartments while landlords know by experience how much students are willing to pay for a flat in some particular areas. The model has several drawbacks. The crucial elements in determining the population's bargaining powers and therefore their shares of the cake are fixed exogenously. Young leaves the most important element unexplained. The model has the unsatisfactory prediction that the share received by a population with one million types who use large samples is determined by just one further type who uses a small sample. It does not explain why different types of players, probably receiving different payoffs, may co-exist in the same population.

In this paper we endogeneize the size of the samples drawn by the players. We will present a model similar to Young's, but with the added feature that people can change their learning rule by altering the sample size or 'window' used. We will assume that players observe the payoffs received by other members from the same population and from time to time decide to imitate more successful behaviours. It is therefore as though people care about their relative performances within the social class to which they belong. If a learning rule performs better than its rivals, it is natural to expect that it will be employed by a growing proportion of people over time.

We can identify two opposite forces that affect the amount of information gathered by the players. On the one hand, players with small samples are more likely to draw a 'wrong sample' when the system is close to but not at a convention and to play a non–optimal strategy. If information is free, big samples will give a higher expected payoff. On the other hand, when the probability of mistakes is small, the system is close to a convention most of the time and those players who sample few elements will play optimally almost as often as players with big samples. If sampling costs grow with the size of the sample taken, 'small windows' will do better, and evolution will tend to reduce the amount of information gathered by the players.

In the paper we show that:

(i) When there are no sampling costs and the level of noise is arbitrarily small, people with 'larger' window sizes perform better, on average, than people with smaller sample sizes within the same population.

(ii) When there is no noise, for any positive sampling cost 'smaller' sample sizes perform better than larger samples. In this extreme case the evolutionary process converges with probability one to a convention and remains there for ever. In this case there is no need for agents to collect more than one unit of information - it is enough to see one car to realize that Londoners drive on the left. If we assume some type of imitation or Darwinian selection, we will observe, in a noiseless world, a tendency for 'well informed' people to disappear.

(iii) When there are sampling costs, one can always find small enough levels of noise such that the 'smallest' sample size will always pay best. The intuition underlying this fact is that, as the noise vanishes, so also does the advantage of sampling.

Finally, we characterize the evolutionarily stable sample sizes. The idea is to allow the entry of new people who bring with them new behaviours and to test their fitness in the enviroment. If there are samples that perform better than others, they will invade the population because everybody will adopt them. In the limiting case in which the noise tends to zero, we can characterize not only the evolutionarily stable sample size, but also the long-run convention of the system. We compare our results with a situation in which each population can decide how much to sample. In particular we consider a thought experiment in which the different populations elect a representative to play the game who is committed to choosing a certain sample size. On comparing the Nash equilibrium of this game, where sample sizes are the strategies, with the outcome of an economy populated by uncoordinated myopic players who follow very simple imitative behaviour, we find that the latter is socially more efficient. We shall show that the economy of myopic players will always converge to the symmetric bargaining solution while this is not necessarily true for the economy populated by 'rational' players.

We offer an informal discussion of the case in which the asymptotic results do not hold. The assumption of very small mistake rates is made, in the works of Kandori et al. (1993) and Young (1993a and 1993b), for reasons of tractability rather than because the noisy case is thought uninteresting. The results of some simulations of the model show that, when the noise is large, the difference in the profitability of different sample sizes depends not only on the level of noise but also on the distribution of sample sizes in the populations. We provide a very simple example in which, when sampling is costly, large sample sizes are better than smaller ones for 'intermediate' levels of noise while they are worse for both small and large noise rates.

The paper is organized as follows: In the next section we introduced the model. In the third section we characterized players' behaviour in terms of their windows sizes which is useful to compare expected payoffs to players using different learning rules. In the last two sections we present the main results of the paper.

2 The model.

Suppose that the unperturbed Nash Demand Game is played once every period by two players respectively drawn at random from two large populations which we follow Young in calling population I (landlords) and population II (tenants). Each player announces a share of the crop, and receives his demand only when the pair of demands is compatible. Each player forms his beliefs about the enviroment he is facing knowing some part of the available information about what other people have done in the past. Landlords (tenants) have access to a 'library' that contains information about m past demands of members of the other population. Players have access only to their own population's library. A landlord (tenant) decides what strategy to use by taking a random sample of size k (w) from his population's library and then playing the best reply to it. Players from the same population may use samples of different sizes.

The information stored in the landlord's (tenant's) library evolves as follows. Every time the game is played, the strategy played by the tenant (landlord) is stored. However, since the library has a limited capacity of m units of information another record of a play must leave the library.

The main difference of our model as compared with Young's lies in the definition of the state space. A state of the system in the model of Young is the 'ordered sequence' of the last m plays of the game. In our model, the demands are not ordered according to the time they entered the libraries. Every time the game is played a new element enters the library and the element it replaces is chosen at random from those previously present. Such a change in the model reduces considerably the state space. Our model has the advantadge of being much more tractable than Young's. The changes in the model simplifies the analysis without doing any violence to the essentials of the process.

We are interested in characterizing the evolution of the information stored in the libraries, since it determines the probability distribution of the future behaviour in the two populations. The evolution of the information in the libraries can be represented by a Markov chain defined on the state space Z. Let Z, the set of all possible stocks in the library, be characterized as follows:

$$Z = \{\{(z_1^I, z_2^I, \ldots, z_n^I), (z_1^{II}, z_2^{II}, \ldots, z_n^{II})\} | z_i^j \in \{0, 1, 2, \ldots, m\}, \sum z_i^j = m\}$$

where z_i^j is the number of times the strategy i is recorded in the population j's library and n is the dimension of the strategy space (all possible announcements). In order to characterize the evolution of the state of the system we make the following assumptions:

Assumption 1. Every sample is drawn with the same positive probability.

Assumption 2. Every record of a past play leaves the library with the same positive probability.

Assumption 3. With positive probability ϵ, players make mistakes[1] by playing a strategy chosen at random. When a mistake is made, all strategies are possible. We will assume that all strategies occur with the same probability ϵ/n.

Assumption 4. The probability densities for window sizes k and w are $f(k)$ and $g(w)$ in populations I and II respectively. The probability that a landlord uses sample size $\tilde{k}$ is equal to $f(\tilde{k})$. The probability that a tenant uses sample size $\tilde{w}$ is $g(\tilde{w})$.

We now define a convention. Consider the state in which $z_{1-x}^I = m$ and $z_x^{II} = m$. Whatever samples are drawn, the landlords will then demand x and

[1]For sake of simplicity we assume that the probability of mistakes is the same in both populations. The results do not change if different rates are assumed.

the tenants $(1-x)$. The states consisting of such a Pareto efficient division of the crop are the *'conventions'* of the system, i.e. the states that reproduce themselves[2]. The main feature of such conventional behaviour is that any player prefers to conform to it if everybody else does so. For notational convention we will refer to the m-repetions of the same partition $(x, 1-x)$ as c_x.

Since our model differs from Young's, it is neccesary to confirm the following result. The proof is similar, although the current model allows a much less restrictive constraint on the minimum necessary sample size. Young requires that at least some players sample at most $m/2$ records in their libraries, whereas the following proposition works with $m/2$ replaced by $m-1$. ($(m-1)$ instead of $m/2$).

Proposition 1 *If at least one agent in each population samples at most m-1 elements the system converges almost surely to a convention.*

Proof. We need to prove that it is possible that the same sample will be drawn for some time with the result that the same best reply until is made, until we have built up homogeneous library records, one for each population, that correspond to Pareto efficient divisions of the crop.

To this end, we consider the extreme case in which some players in each population sample exactly $m-1$ records, while the remainder may sample all the records. Suppose that players from population I who sample $(m-1)$ records happen to be selected to play the game $(m-1)$ times and that they happen to sample the same elements and so all play the same best reply x. This possibility occurs with positive probability, because the last element which enters the library can leave it in the following period. We can obtain a state of population I that contains $(m-1)$ copies of the same demand x. These $(m-1)$ elements can remain for some time in the library

[2]The conventions have the property that they are the only absorbing sets of the model we are considering.

of population II, and so be drawn by the players from this population, who will, therefore, demand $(1-x)$. We thereby build a state of the system that contains $(m-1)$ copies of the observation $(1-x)$ in the library of population I and $(m-1)$ copies of the observation x in the library of population II. There is a positive probability that these are the samples drawn next period and that the elements that are different leave the corresponding library. We have reached a convention (c_x) in $(2m-1)$ periods with positive probability p. The probability that a convention is not reached in $s(2m-1)$ periods is $(1-p)^s$, which goes to zero as $s \to \infty$. □

A convention can be abandoned only when some people start deviating from the behaviour prescribed by it. This is why in our model we introduce the possibility that people may make mistakes and play a strategy that is not a best reply to the sample they have drawn.

To illustrate this point consider the 2×2 bargaining game of Figure 1:

	LOW	*HIGH*
LOW	b, b	b, a
HIGH	a, b	0, 0

$$a > b > 0$$

Figure 1: Game 1.

The state space is:

$$Z = \{(z_1, z_2) | z_1, z_2 \in \{0, 1, 2, \ldots, m\}\}$$

where z_1 (z_2) denotes the number of Low's in the library to which players from I (II) have access.

This simple game has only two conventions: (m,0) and (0,m). Let us assume that the established convention is (m,0) and that some tenents start making mistakes. Sometimes they demand Low although the best reply to any sample containing all Lows is High. The mistakes will, with positive probability, remain for some time in the landlords' library. It is possible that a landlord will draw the mistakes and, if there are sufficiently many, he will, then, play a strategy that is not the conventional one. If all players have the same utility function, the probability that agents deviate from a convention in response to mistakes made in the other population will depend on the size of the sample they draw.

The introduction of mistakes keeps the system continuously in motion. Under assumptions (1)-(4) we can obtain a Markov Chain defined on the state space Z with the transition matrix:

$$M(\epsilon) = M(\epsilon; m, f(k), g(w), G) = [p_{i,j}], \quad (i, j \in Z) ,$$

where the transition probability $p_{i,j}$ is the probability of moving from state i to state j in one period [3].

Introducing mistakes makes the Markov chain irreducible, i.e; all the states intercomunicate. As the Markov chain is also aperiodic, it is therefore ergodic and has a unique stationary distribution, i.e.,there exists a unique distribution $(1 \times |Z|)$ vector μ_ϵ such that:

$$\mu_\epsilon M(\epsilon) = \mu_\epsilon \tag{1}$$

thus, system settles down in the long run to a distribution which is independent of the initial conditions. The solution to (1) is a correspondence $\Gamma : \epsilon \Longrightarrow \Delta^{|Z|-1}$ which is upper hemicontinuous. The equilibrium selection is continuous with respect to perturbations (see Kandori et al. (1992)).

[3]In what follows we will write $M(\epsilon)$ instead of $M(\epsilon; m, f(k), g(k), G)$. The game, the memory size and the distributions of window sizes are fixed.

The interpretation of the probabilities attached to each state in the long-run distribution, is the time spent by the system in the corresponding state.

It turns out that this long-run distribution often put most of its mass at just one of the possible conventions when the mutation rate is small. As the mutation rate tends to zero, all other states are assigned zero mass. We then say that the remaining convention has been selected in the long run. This conclusion is no longer valid when the mutation rate is set to zero from the outset. The convention that is then observed in the long-run depends on the initial conditions. When players do not make mistakes, the conventions are absorbing states, i.e. once the system is at a convention it is impossible to escape.

The trick in selecting a particular stationary distribution out of all possible conventions is the introduction of a small amount of noise into the system. Out of all the possible conventions we select one, by introducing noise in the system and letting it tend to zero.

We follow Young in assigning a 'resistance' to each convention. The convention that is selected is that from which it is most difficult to escape or, seen from another perspective, the one which is easiest to reach from any other convention. The computation of the 'resistance' associated to one particular convention involves counting the minimum number of mutations needed to reach such a convention from any other. As the mutation rate tends to zero only those states which are easiest to reach will be observed in the long run.

The convention which has the smallest resistance is therefore the one from which it is most difficult to escape. When we consider the possibility of going from one convention to another we have only to consider the minimum number of mistakes one of the libraries has to contain for 'the most mistake-sensitive player' to be capable of drawing a sample that prescribes a non-conventional choice. When all players from the same population have the same utility function, the most sensitive player is the person who draws the smallest sample.

Proposition 2 *(Young (1993a)) When the rate ϵ of mistakes goes to zero, the stationary distribution will put weight only on the convention or conventions with minimal resistance .*

The main result of Young's study of the Nash Demand Game is the selection of the asymmetric Nash bargaining solution as the long-run convention of the system. There are two conditions that need to be satisfied. The rate of mistakes has to be positive but vanishingly small. Also a very finely meshed stategy space has to be considered. That is to say

$$S = \{0, \delta, 2\delta, 3\delta, \ldots, 1-\delta, 1\}$$

where the mesh-size $\delta > 0$ must be taken to be sufficiently small.

Definition 1 (Asymmetric Nash Bargaining Solution) *The Asymmetric Nash Bargaining Solution is the division $(x, 1-x)$ that maximizes*

$$\{u(x)\}^a \{v(1-x)\}^b \text{ subject to } 0 \leq x \leq 1$$

where u and v are the utility functions of players 1 and 2 respectively, and a and b are their bargaining powers.

The following proposition, drawn from Young (1993b), is true in our model:

Proposition 3 *(Young (1993b)) The evolutionary process described above converges to the asymmetric Nash solution as $\delta \to 0$, with each population's bargaining power equal to the smallest sample size used by an individual in that population.*

Proof. Young's proof also applies in our model. There is a correspondence between the conventions in our model and those in Young's. The proof of the theorem depends on computing the number of mistakes needed to abandon one convention in order to enter the basin of attraction of another convention and the considering the limit as $\delta \to 0$. □

The proposition above implies that determining the convention that will be observed most of the time, requires focusing only on the evolution of $\underline{k} = \min\{supp\ f(k)\}$ and $\underline{w} = \min\{supp\ g(w)\}$, where f and g are the densities that describe the distribution of sample sizes in populations I and II (see Assumption 4). The long run convention of the system is determined by the members of the populations who draw the smallest sample. In the conventions, the payoffs obtained by members from the same population are the same independently of the amount of information gathered.

The fact that people make mistakes implies that with positive probability the system is not at a convention at any particular time. In such cases, the expected payoff to different window sizes may differ. In order to compare the profitability to different sample sizes, we need to obtained the expected payoffs in each state as well as the long-run distribution of the system.

3 Long-run payoffs

We will assume that all players have the same utility function. This assumption is necessary to rule out the influence of different levels of risk aversions in the selection of the long run convention. The difference in the expected behaviour of two players from the same population then depends only on how much information they gather from their library (sample size) and not on their attitudes towards risk.

Risk-aversion plays an important role in the models of Young (1993b) and Rubinstein (1982). In both models it is the player who is more risk averse who gets the smallest share of the cake.

Players' strategies. Players use best replies against random samples drawn from their libraries. Before the sample is chosen, we can characterize the anticipated behaviour of each player as a mixed strategy, with the probabilities attached to each strategy being determined by the current stock of information on record. Each landlord (tenant) drawn to play, faces a tenant (landlord) who behaves as if he were using a mixed strategy. Notice that each player has some limited knowledge about the past behaviour of the other population, but not about his own.

Consider Game 1 of Figure 1 and a landlord who draws a sample of size k. The state of the system is $z = (z_1, z_2)$ and the recorded history contains m past plays of the game.

Let us define $p_i^s(z_{-i}, x)$ as the probability with which player i plays strategy s when the state of the system is z and he draws a sample of size x.

When the library contains m records, there are $\binom{m}{k}$ possible samples of size k. High will be the best reply when the sample drawn contains at least $l_b(k)$ lows [4]

$$l_b(k) = \left[\frac{b}{a}k\right]^+$$

The probability with which a player from population 1, sampling k records, plays high in state z is therefore

$$p_1^H(z_2, k) = \binom{m}{k}^{-1} \sum_{l_b \leq l \leq k} \binom{z_2}{l} \binom{m - z_2}{k - l}$$

This probability is non-decreasing in z_2 and is zero for $z_2 < l_b$ and one for $z_2 > m - (k - l_b)$.

The strategy played by any member of population I (II) depends on the state of the system in the other population, the payoffs (through l_b) and on the size of the sample.

Example Choose the payoffs a and b of Game 1 to obtain:

[4] $[x]^-$ ($[x]^+$) denotes the greatest (smallest) integer smaller or equal (greater or equal) than x.

	LOW	*HIGH*
LOW	1.2, 1.2	1.2, 3
HIGH	3, 1.2	0, 0

Figure 2.

Assume that there are two types of landlord. When called upon to play, Type 1 samples only one past record; Type 2 samples three. All tenants sample 2 units of information. Tables 2, 3 and 4 report the mixed strategy (the probability of playing High) used by tenants, Type 1 and Type 2 landlords respectively. Each entry corresponds to one state of the system. The horizontal dimension is the state of the tenants, i.e. the number of times in landlords' library that a tenant played Low. The vertical dimension corresponds to the state of the landlords. Both range from 0 to 8. The entry (6,3) in Table 1 says, for instance, that a tenant will play High with probability 27/28 when the landlords' library records that tenants played Low 3 times and the tenants' library record that landlords played Low 6 times. The same entries in tables 2 and 3 represent the probabilities assigned to High by landlords who sample 1 and 2 records respectively.

	0	1	2	3	4	5	6	7	8
0	0	0	0	0	0	0	0	0	0
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
2	$\frac{13}{28}$	$\frac{13}{28}$	$\frac{13}{28}$	$\frac{13}{28}$	$\frac{13}{28}$	$\frac{13}{28}$	$\frac{13}{28}$	$\frac{13}{28}$	$\frac{13}{28}$
3	$\frac{9}{14}$	$\frac{9}{14}$	$\frac{9}{14}$	$\frac{9}{14}$	$\frac{9}{14}$	$\frac{9}{14}$	$\frac{9}{14}$	$\frac{9}{14}$	$\frac{9}{14}$
4	$\frac{11}{14}$	$\frac{11}{14}$	$\frac{11}{14}$	$\frac{11}{14}$	$\frac{11}{14}$	$\frac{11}{14}$	$\frac{11}{14}$	$\frac{11}{14}$	$\frac{11}{14}$
5	$\frac{25}{28}$	$\frac{25}{28}$	$\frac{25}{28}$	$\frac{25}{28}$	$\frac{25}{28}$	$\frac{25}{28}$	$\frac{25}{28}$	$\frac{25}{28}$	$\frac{25}{28}$
6	$\frac{27}{28}$	$\frac{27}{28}$	$\frac{27}{28}$	$\frac{27}{28}$	$\frac{27}{28}$	$\frac{27}{28}$	$\frac{27}{28}$	$\frac{27}{28}$	$\frac{27}{28}$
7	1	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1	1

Table 1

	0	1	2	3	4	5	6	7	8
0	0	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{5}{8}$	$\frac{3}{4}$	$\frac{7}{8}$	1
1	0	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{5}{8}$	$\frac{3}{4}$	$\frac{7}{8}$	1
2	0	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{5}{8}$	$\frac{3}{4}$	$\frac{7}{8}$	1
3	0	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{5}{8}$	$\frac{3}{4}$	$\frac{7}{8}$	1
4	0	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{5}{8}$	$\frac{3}{4}$	$\frac{7}{8}$	1
5	0	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{5}{8}$	$\frac{3}{4}$	$\frac{7}{8}$	1
6	0	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{5}{8}$	$\frac{3}{4}$	$\frac{7}{8}$	1
7	0	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{5}{8}$	$\frac{3}{4}$	$\frac{7}{8}$	1
8	0	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{5}{8}$	$\frac{3}{4}$	$\frac{7}{8}$	1

Table 2.

	0	1	2	3	4	5	6	7	8
0	0	0	$\frac{3}{28}$	$\frac{2}{7}$	$\frac{1}{2}$	$\frac{5}{7}$	$\frac{25}{28}$	1	1
1	0	0	$\frac{3}{28}$	$\frac{2}{7}$	$\frac{1}{2}$	$\frac{5}{7}$	$\frac{25}{28}$	1	1
2	0	0	$\frac{3}{28}$	$\frac{2}{7}$	$\frac{1}{2}$	$\frac{5}{7}$	$\frac{25}{28}$	1	1
3	0	0	$\frac{3}{28}$	$\frac{2}{7}$	$\frac{1}{2}$	$\frac{5}{7}$	$\frac{25}{28}$	1	1
4	0	0	$\frac{3}{28}$	$\frac{2}{7}$	$\frac{1}{2}$	$\frac{5}{7}$	$\frac{25}{28}$	1	1
5	0	0	$\frac{3}{28}$	$\frac{2}{7}$	$\frac{1}{2}$	$\frac{5}{7}$	$\frac{25}{28}$	1	1
6	0	0	$\frac{3}{28}$	$\frac{2}{7}$	$\frac{1}{2}$	$\frac{5}{7}$	$\frac{25}{28}$	1	1
7	0	0	$\frac{3}{28}$	$\frac{2}{7}$	$\frac{1}{2}$	$\frac{5}{7}$	$\frac{25}{28}$	1	1
8	0	0	$\frac{3}{28}$	$\frac{2}{7}$	$\frac{1}{2}$	$\frac{5}{7}$	$\frac{25}{28}$	1	1

Table 3.

Notice that the strategy used by each population depends neither on its own state (horizontal dimension for tenants and vertical for landlords), nor on the distribution of types in the two populations. Notice also, comparing tables 3 and 4, that there are two subsets of states in which a pure strategy is played and that these subsets grow with the size of the sample.

In the proposition which follows we formalize the intuition provided by the previous example. We characterize the set of states in which players use a pure strategy, when not making a mistakes.

Let us consider Game 1 and define the following subsets of Z:

$$\underline{Z}_{x,y} = \{(z_1, z_2)|z_1 \leq x, z_2 \leq y\}$$

$$\overline{Z}_{x,y} = \{(z_1, z_2)|z_1 \geq x, z_2 \geq y\}$$

Consider a landlord who samples k elements, with l which are Low's. Low will be the best reply to that sample if $l \leq kb/a$. All possible samples drawn from:

$$\underline{Z}_{m,[\frac{kb}{a}]^-} = \{(z_1, z_2)|z_1 \leq m, z_2 \leq \left[\frac{kb}{a}\right]^- \}$$

will have at most $[kb/a]^-$ lows and every landlord who samples k elements will play Low with probability one.

High will be played with probability one by any landlord with a sample of size k in all the states in $\overline{Z}_{0,[m-k(a-b)/a]^+}$. In any other state it is possible to find samples of size k to which Low is the best reply as well as samples of equal size to which High is the best reply.

We are interested in identifying the sets of states in which only pure strategies are played.

Consider the following correspondence $s^i : R^2 \Rightarrow Z$,

$$s^i(m, x) = \{(z_1, z_2)|\ p_i^s(z_{-i}, x) = 1\}$$

Note that $L^1(m, k) = \underline{Z}_{m,[\frac{kb}{a}]^-}$ and $H^1(m, k) = \overline{Z}_{0,[m-k(a-b)/a]^+}$.

Proposition 4 *$L^1(m, \lambda k) \subseteq L^1(m, k)$ and $H^1(m, \lambda k) \subseteq H^1(m, k)$ for $0 \leq \lambda \leq 1$.*

Proof. It follows from the definitions of L^1 and H^1.

Proposition 4 states that the set of states in which pure strategies are played shrinks with the sample size.

This result can be extended to the case in which there are more than two strategies. The different sample sizes need the same 'proportion of mistakes' to start playing a nonconventional strategy when the utility functions are the same.

Player's payoffs. In order to compare the profitabilities to different sample sizes we need to obtain the payoffs in each state as well as the long-run distribution μ_ϵ.

Different sample sizes will have the same expected payoffs in all those states in which the mixed strategies are the same. From the preceding proposition, we know that in some states different sample sizes prescribe the same pure strategy. Clearly, in all such states, players using different amounts of information will have the same expected payoff, independently of the opponent's strategy. In all other states different sample sizes prescribe different mixed strategies and, therefore, will have different expected payoffs.

The profitability of a learning rule (characterized by its sample size) depends not only on the rate of mistakes but also on the composition of the populations which determines the actual long-run distribution. Let $\pi^k(\epsilon; z, g(w))$ be the expected (gross) payoff in state z to a player who samples k units of information when the rate of mistakes is ϵ. We can decompose this expected payoff into two components:

$$\pi_k(\epsilon; z, g(w)) = (1-\epsilon)\pi_k^s(\epsilon; z, g(w)) + \epsilon\pi^t(\epsilon; z, g(w)) \qquad (2)$$

The first part of the payoff, $\pi_k^s(\epsilon; z, g(w))$, is received when the player uses the information provided by the sample he has drawn. Players do not tremble, and therefore use the information available to them with probability $(1-\epsilon)$.

The second component, $\pi^t(\epsilon; z, g(w))$, is the payoff obtained when the player trembles and therefore plays an arbitrary strategy. This component which does not depend on the sample size is the same for all members from the same population.

Let $c(k) \geq 0$ be the cost of a sample of size k. We obtain the net payoffs by substracting $c(k) \geq 0$ from the right-hand side of equation (2).

The (gross) expected payoff to a player who uses a sample of size k when the long-run distribution of the system is μ_ϵ is given by:

$$\pi_k(\epsilon; f(k), g(w)) = \sum_{z \in Z} \pi_k(\epsilon; z, g(w))\mu_\epsilon(z)$$

where $\mu_\epsilon(z)$ is the weight of the state z in the long-run distribution μ_ϵ.

We are interested in pairwise comparisons of sample sizes. In accounting for differences in window sizes, we shall consider first the simplest case.

Consider Game 1. Assume that all tenants are characterized by the same sample size w. Consider two different samples k and $k' < k$ in the population of landlords.

Proposition 5 *Let us consider sample sizes k and $k' < k$. For m large enough there exist integers q_1, q_2 and $q_2' \geq q_2$ such that*

$$\forall\ z \in \underline{Z}_{m,q_2} \cup \overline{Z}_{q_1,0}\ \pi_k^s(\epsilon; z, w) \geq \pi_k'^s(\epsilon; z, w)$$

$$\forall\ z \in \underline{Z}_{q_1,m} \cup \overline{Z}_{0,q_2'}\ \pi_k^s(\epsilon; z, w) \geq \pi_k'^s(\epsilon; z, w)$$

Proof. Let q_1 as the smallest z_1 such that $p_2^H(z, w) \geq b/a$, and let q_2 and q_2' be the states such that,

$$\text{for all } z_2 < q_2 \qquad p_1^H(z_2, k) \leq p_1^H(z_2, k') \ \ and$$

$$\text{for all } z_2 \geq q_2 \qquad p_1^H(z_2, k) \geq p_1^H(z_2, k')$$

The following figure is a graphical illustration of the previous proposition.

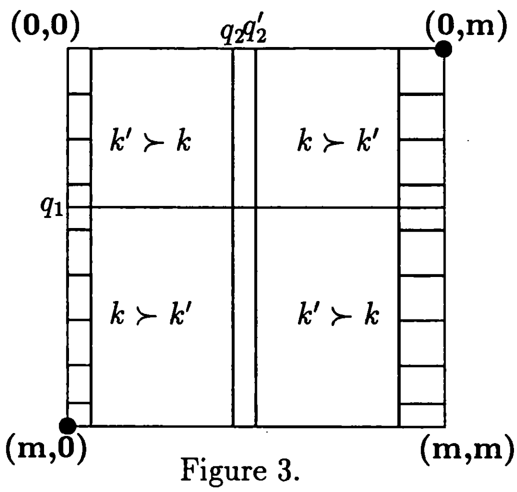

Figure 3.

The state of population II is represented in the horizontal dimension. It ranks from 0 to m and represents the number of Low's in the recorded memory. Similarly, the state of population I is represented in the vertical dimension, from up to down. The conventions are in lower-left and the upper-right corners (filled dots).

Let us consider the area labelled with $k \succ k'$ on the upper-right part of the figure, which are states in some neighbourhood of $(0, m)$. The history available to the tenants has many High's. Tenants will play Low with high probability. The best reply to $p^H < b/a$ is High. A player using a large sample will play High with higher probability than a player with a smaller sample. A similar argument applies to the states in the lower-left corner. The dashed areas correspond to the states in which landlors, either with k or k', play the same (pure) strategy and, therefore, have the same expected payoff. The smallest sample pays better in the sets labeled with $k' \succ k$.

4 The evolution of the learning rule.

The aim of this section is to endogenize the amount of information gathered by players. People can observe the payoffs of some members of their own population and imitate the most successful behaviour. Students may know how much other students are paying and how much they searched. We will assume that together with the average payoffs, players can observe the sample sizes drawn by other members from the same population.

Comparison of payoffs. We will assume that the comparison of payoffs takes place relative to the long-run distribution and so does the evolution of sample sizes. We can justify this assumption on the grounds that the adjustment periods are negligible compared with the time the system spends in the long-run distribution.

When comparing the payoffs to two different sample sizes from the same population, we have only to consider the first part in equation (2).

The difference in payoffs to window sizes k and k', evaluated relative to the long-run distribution is given by

$$D\pi(\epsilon; k, k', f(k), g(w)) = (1-\epsilon)\sum_{z\in Z}(\pi_k^s(\epsilon; z, g(w)) - \pi_{k'}^s(\epsilon; z, g(w)))\mu_\epsilon(z))$$

The function $D\pi$ is a polynomial in ϵ and it is therefore continuous.

Consider the simplest possible case. There are two types of landlords, in proportions θ and $(1-\theta)$. The first type samples k records the second type sample k'. All tenants are of the same type, and sample w past records. The rate of mistakes is the same in both populations.

Let $M(r_1, r_2)$ be the transition matrix when population I (II) follows rule r_1 (r_2) to play the game. The rule can be either to take a sample (k, k' or w) or to tremble (t). We can decompose the Markov matrix $M(\epsilon; \theta k + (1-\theta)k', w)$ as follows:

$$\begin{aligned} M(\epsilon; \theta k + (1-\theta)k', w) &= \theta((1-\epsilon)^2(M(k,w) + (1-\epsilon)\epsilon M(k,t)) \\ &\quad (1-\theta)((1-\epsilon)^2(M(k',w) + (1-\epsilon)\epsilon M(k',t)) \\ &\quad \epsilon^2 M(t,t) + (1-\epsilon)\epsilon M(t,w) \end{aligned}$$

Each time the game is played, a player sampling k is drawn with probability θ. With probabiliy $(1-\epsilon)^2$ neither he nor the opponent, who samples w with probability 1, tremble. The transition matrix is given in this case by $M(k,w)$. With probability ϵ^2 both players tremble; $M(t,t)$ describe the transition probabilities. With probability $\epsilon(1-\epsilon)$ only one player trembles. The markov matrices when only the first or only the second player tremble are $M(t,w)$ and $M(k,t)$ respectively. The terms multiplied by $(1-\theta)$ have an analogous interpretation, with k' being the sample size used by the player drawn from the first population.

When $\epsilon = 0$,

$$M(0; \theta k + (1-\theta)k', w) = \theta M(k,w) + (1-\theta)M(k',w)$$

and the system has as many absobing states as there are Pareto-efficient divisions of the cake. Once a convention has been reached, and it will happen with positive probability, the economy will remain there for ever. The set of absorbing states is independent of the composition of the populations. Independently of the value of θ, the conventions are the only states with 1 on the diagonal of $M(0,)$. Changes in θ only affect the transition probabilities but do not change the absorbing states. From any other state, it is possible to find a chain of transitions which ends up in one convention. The convention which will be selected depends on the initial conditions. In other words, history matters. In a convention, all players obtain the same payoff and, therefore $D\pi(0;k,k') = 0$. The information given by a single unit of information is as good as the whole history.

When $\epsilon = 1$,

$$M(1;\theta k + (1-\theta)k', w) = \epsilon^2 M(t,t) \tag{3}$$

The long-run distribution depends neither on the composition of the population nor on the sample sizes. Players play an arbitrary strategy. The expected payoff is the same for all players and $D\pi(1;k,k') = 0$.

When $0 < \epsilon < 1$, there are no absorbing states. The diagonal elements of $M(\epsilon;\theta k + (1-\theta)k', w)$ are all smaller than 1. The long-run distribution will depend on the specific way trembles are modeled and on the sampling process which is assumed. We shall assume that the probability of sampling an individual with sample size k is equal to its proportion in the population. The trembles have been modeled as the choice of any of the possible strategies with equal probability.

As the noise tends to zero, the long-run distribution concentrates around the convention whose basin of attraction is hardest to escape. If the noise is vanishingly small we can easily characterize the long-run distribution and the long-run payoffs. For very small noise rates, the system will be almost always in a convention, although all other states will be visited with positive probability. The closer the states are to the conventions, the higher will be their weights in the long-run distribution. For very small mistake rates we need to consider only states in neighbourhoods of the conventions to compare expected payoffs. In those states, as we have seen in the previous section, larger sample sizes have a higher expected payoff, due to the fact that, close to the conventions, it responds with smaller probability to the mistakes coming from the other population.

Proposition 6 *Let us consider two sample sizes k and $k' < k$. For $0 < \epsilon < \tilde{\epsilon}$ and large m, $D\pi(\epsilon; k, k') > 0$.*

Proof. See Appendix 1. We prove that for arbitrarily small positive ϵ, the payoffs in the states where the larger window size pays best compensate the disadvantadge in all other states.

In Appendix 2, we report the results of a simulation of Game 1 with $m = 4$ and different values of ϵ. Each entry is the probability attached to the corresponding state in the long-run distribution. We report the weight (p) of the conventions and the six neighbouring states (three for each convention) in the long-run distribution.

The larger the noise, the smaller the probability p. As the rate of noise grows p decreases and the probability mass shifts towards states in which the smallest sample performs relatively better.

We can conclude that $D\pi$ is 0 at $\epsilon = 1$ and $\epsilon = 0$ and growing at this last point.

The expected payoff to the different sample sizes depends on the mass put by the long-run distribution on all the states of the system. As the rate of mistakes ϵ grows the probability weight moves from the conventions to the other states. The set of states in which the smallest sample size has higher expected payoff depends on the payoffs, the sample sizes and the memory length. It is possible to find examples for which the largest sample is the most profitable for all rates of noise. There exists, also, the possibility that at high rates of noise, the smallest sample has a higher expected payoff. The states in which miscoordination is common are more likely when the noise is large. The smallest sample may be better as we shall see in the following numerical example.

An example Consider the game considered in the previous section. All tenants draw a sample of size 2. There are, as before, two types of tenants [5]. Type 1 tenants draw a sample of size $k = 3$. Type 2 sample $k' = 1$ units of information. The proportion of members from the landlords' population who sample $k = 3$ is given by θ. The columns report, respectively, the rate of noise of the system and the difference in payoffs ($D\pi(\epsilon; 3, 1)$ for the different rates of noise, in the long-run distribution.

ϵ	$\theta = 0$	$\theta = 0.5$	$\theta = 1$
0	0	0	0
0.05	0.024999	0.025198	0.02521
0.1	0.039749	0.04082	0.041329
0.2	0.048659	0.049013	0.050820
0.3	0.036386	0.039811	0.042752
0.4	0.023438	0.026059	0.028503
0.5	0.012609	0.014197	0.015741
0.6	0.0053594	0.0061561	0.0069475
0.7	0.0014224	0.0017458	0.00207
0.8	-0.00008059	0.00001237	0.00010585
0.9	-0.00019779	−0.00018624	-0.00017463
0.95	-0.000071602	−0.000070145	-0.000068685
1	0	0	0

Table 4 .

The last column shows the difference in expected payoffs when all members of population I sample $w = 3$, evaluated relative to the long-run distribution. For all noise rates smaller than 0.8, the larger window size pays

[5]We have selected k and k' in such a way that the structure represented in Figure 1 is preserved.

better than the smaller. Any mutant using a unit less of information will die out.

The following figures represent the two possible behaviours of $D\pi$:

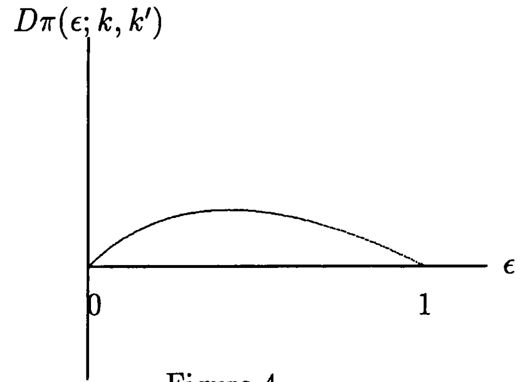

Figure 4

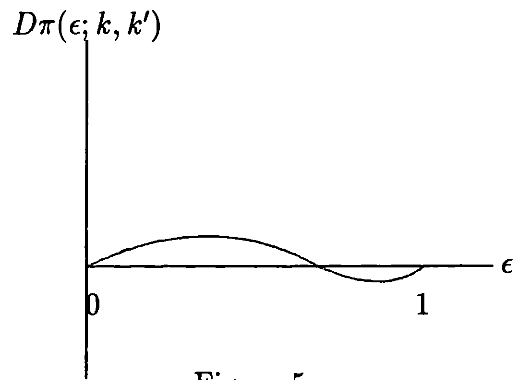

Figure 5

The function $D\pi$ is defined for fixed $f(k)$ and $g(w)$. In the following section, we introduce dynamics in the distributions of sample sizes. As the proportion of players using different sample sizes change so does $D\pi$.

The following table shows the results of running the same simulation, but keeping the noise fixed and changing θ.

θ	$D\pi(0.1;3,1)$	$D\pi(0.5;3,1)$	$D\pi(0.8;3,1)$	$D\pi(0.9;3,1)$
0	0.039749	0.012609	-0.00008059	-0.00019779
0.05	0.039887	0.01277	-0.000071324	-0.00019664
0.1	0.040017	0.01293	-0.000062045	-0.00019548
0.2	0.040257	0.013249	-0.00004347	-0.00019317
0.3	0.04047	0.013566	-0.000024876	-0.00019086
0.4	0.040657	0.013882	$-6.261\,10^{-6}$	-0.00018855
0.5	0.04082	0.0141977	0.00001237	-0.00018624
0.6	0.04096	0.014509	0.00003102	-0.00018392
0.7	0.041078	0.01482	0.00004970	-0.0001816
0.8	0.041178	0.015129	0.0000684	-0.00017928
0.9	0.04126	0.015436	0.000087115	-0.00017696
0.95	0.041296	0.015588	0.000096481	-0.0001758
1	0.041329	0.015741	0.00010585	-0.00017463

Table 5.

The simulation suggests that that advantage of big samples increases with the proportion of individuals in the population using the same sample size. The introduction of smaller samples has an effect which is similar to the increase in the noise. When there is no noise, there are some states which are ephemeral, i.e.,they are never visited unless they are the initial state of the system. The introduction of mistakes makes all states non ephemeral, opens way-outs from the conventions, and links states that before were uncomunicated. The introduction of smaller samples reduces the set of ephemeral states and opens new links between states. One can expect, that under our assumptions, an increase in the proportion of players using a small sample moves probability weight towards the states which are far from the conventions.

5 Evolutionary stability.

We can only make qualitative statements about the relation between costs, noise and evolution of learning rules in the system. It is important to realize that all the results of Young (1993a), Kandori et al. (1993) and Young (1993b) are valid when the rate of noise is close to zero. Only in this special case in which can we characterize the long-run distribution of the system.

In this section we characterize the evolutionarily stable sample sizes. The idea is to perturb the distributions of sample sizes by introducing new people with different learning rules. Stable distribution are those that survive such a disturbance. The dynamics in the compositions of the populations will be driven by some type of imitation or Darwinian selection (the survival of the fittest). It is important to notice that we have two different levels of evolution. On the one hand we have the evolution of the system, as in Young (1993b), which is driven by the adaptive play and the mistakes. On the other hand we have the evolution of the learning rules which is driven by the imitation of more profitable learning rules and by mutations which affect the sizes of the sample. We do not take the distributions of sample sizes as given. We can consider two different relevant time horizons. In the long run we take the distributions of sample sizes, f and g, as given with the system being in the long run distribution. We can think of a situation in which people adjust very slowly their learning rules compare to adjustments in the environment. In the ultra long run players have had time to adjust their learning rules. Our aim is to find two sample sizes k^* and k^* which are evolutionarily stable, i.e, cannot be invaded. In the ultra long run, the distributions f and g will put weight only on k^* and w^* respectively. As we have seen in the previous section, there is always a sample size which dominates the others, i.e has a higher expected payoff. Under darwinian dynamics the populations will be invaded by such a sample. Selection implies, in this case, homogeneous populations.

We will compare the results with an hypothetical situation in which sample sizes are selected at the population level. For this purpose we will asume that players, in each population separately, elect a representative to play the game on their behalf. The representative is characterized by the size of the sample he draws. Both populations behave this way, knowing the long-run implications of their choices.

5.1 Asymptotic results.

The asymptotic results apply in the case when the noise is very small. We can focus on the behaviour of $\underline{k}$ and $\underline{w}$, which are the sample sizes which determine the bargaining powers and the long-run distribution (Proposition 2).

The distributions of sample sizes, $f(k)$ and $g(w)$, may change over time. When different sample sizes are present in the same population, there will be a process of selection that will wipe out inefficient learning rules. Only when these changes affect either $\underline{k}$ or $\underline{w}$ will the system move to a new convention.

5.1.1 Costless window sizes

From the previous section, we know that when information is free, and the noise tends to zero, big samples have higher expected payoff than small ones, although the advantage of sampling vanishes with the noise.

Proposition 7 . *When players can change their sample sizes without cost, the only evolutionarily stable sample sizes are $\overline{k}$ and $\overline{w}$.*

Proof. Let us assume that all members in population 1 (2) are sampling $k < \overline{k}$ ($w < \overline{w}$) and that sampling is costless. These sample sizes are not evolutionarily stable. By proposition 6 any mutant who enters the population and samples more will have a higher expected payoff. □

When sampling is costless we will observe an endless process of growth in the samples. If there is a limit in peoples' capacity to retain information, there will be full employment of this capacity, which will be the only uninvadable sample size, with the population gifted with higher capacity receiving a greater share.

Nash equilibrium in sample sizes. In the analysis developed in the previous section, players do not have any concious choice of strategies. They simply apply the simple rule of playing a best reply to some observation about the past and sometimes imitate more succesful learning rules. We have assumed very little about players' information. Players only know some limited information about previous demand and payoffs and window sizes of players from the same population. In what follows we compare the results obtained above with those obtained in the extreme case of perfect forsight. We shall assume that players are committed to play as before, but they are able to computing the long run distribution and know that with their choice of sample size can affect the convention which will be selected. The situation can be modeled as a one-shot game, with sample sizes as strategies and payoffs computed in the long-run conventions.

We consider the following thought experiment: Imagine that landlords and tenants have to elect a representative (a type) to play the Nash Demand Game on behalf of the population. The rules of the game are as before, with the difference that the player is not randomly selected but chosen by the population. The representative decides how to play by sampling the number of records that characterize his type. Players are aware that their joint choices will determine the bargaining powers and their shares of the crop in the long-run. They only care about long-run payoffs. Which sample size will they choose? They will select a player with a sample size that maximizes their payoffs given the other population's choice of sample size.

The strategies spaces,

$$S^1 = \{1, 2, 3, .., \overline{k}\}$$

$$S^2 = \{1, 2, 3, .., \overline{w}\}$$

are all the possible sample sizes. The capacity limits, $\overline{k}$ and $\overline{w}$, are not necessarily the same.

Let us consider the simplest case in which the utility functions are linear.

The asymmetric Nash bargaining solution, with bargaining powers are k and w, is the partition $(x^*, 1-x^*)$ which solves,

$$\max_x x^k (1-x)^w$$

The solution is

$$x^* = \frac{k}{w+k}$$

Landlords and tenants have to elect a representative to play Game 2. Each entry correspond to the (asymmetric) Nash bargaining solution $(x^*, 1-x^*)$ for the different sample sizes.

	1	2	3	4	5
1	$\frac{1}{2}$, $\frac{1}{2}$	$\frac{1}{3}$, $\frac{2}{3}$	$\frac{1}{4}$, $\frac{3}{4}$	$\frac{1}{4}$, $\frac{4}{5}$	$\frac{1}{5}$, $\frac{5}{6}$
2	$\frac{2}{3}$, $\frac{1}{3}$	$\frac{2}{4}$, $\frac{1}{2}$	$\frac{2}{5}$, $\frac{3}{5}$	$\frac{2}{6}$, $\frac{4}{6}$	$\frac{2}{7}$, $\frac{5}{7}$
3	$\frac{3}{4}$, $\frac{1}{4}$	$\frac{3}{5}$, $\frac{2}{5}$	$\frac{1}{2}$, $\frac{1}{2}$	$\frac{3}{7}$, $\frac{4}{7}$	$\frac{3}{8}$, $\frac{5}{8}$
4	$\frac{4}{5}$, $\frac{1}{5}$	$\frac{4}{6}$, $\frac{2}{6}$	$\frac{4}{7}$, $\frac{3}{7}$	$\frac{1}{2}$, $\frac{1}{2}$	$\frac{4}{9}$, $\frac{5}{9}$

Figure 6. Game 2.

The entries on the diagonal correspond to the symmetric Nash bargaining solution, the fifty-fifty division, because both populations have the chosen the same sample size and therefore have the same bargaining powers. The optimal choice for each population, given the other's sample size, is to elect a player who samples the most he can. In the unique Nash equilibrium of this game, the two populations select a player with the largest possible sample size. If the strategy spaces, S^1 and S^2, are the same for both populations, we will observe the fifty-fifty division and people sampling at the limit of their capacity. Notice that the Nash equilibrium corresponds to the evolutionarily stable sample size. It is interesting the fact that the observed behaviour is the same with myopic imitative players and with fully rational players. Evolution leads to the same result that concious rational choice.

5.1.2 Costly sample sizes.

We now study the more realistic setting in which sampling is costly. We assume that the cost of a sample is proportional to its size; all members of the same population have the same marginal cost, c_1 for landlords and c_2 for tenants.

Proposition 8 . *When sampling is costly, the only evolutionarily stable sample sizes are* $k = w = 1$*. Furthermore, if players have the same utility fuction, the long-run convention will be the fifty-fifty division.*

Proof. When small samples are less costly a reduction in the sample size implies a saving in the cost while the worsening in the performance is negligible:

$$lim_{\epsilon \to 0} D\pi(\epsilon; k, 1) = 0 \quad \forall k > 1$$

$$\lim_{\epsilon \to 0} c(k) - c(1) > 0 \quad \forall k > 1$$

The only uninvadable sample size is 1. This result is independent of the relative costs and of the shape of the cost function. In the particular case of homogeneous utility functions the fifty-fifty division will be the rule and decision costs will be minimized.

The results differ from the situation in which representative player is chosen by the populations.

Landlords and tenants choose sample sizes k^* and w^* which maximize they long-run payoffs, taking the rival's sample sizes as given,

The Nash equilibrium of the following one shot game is selected,

	1	2	3	4	5	
1	$\frac{1}{2}-c_2$ $\frac{1}{2}-c_1$	$\frac{2}{3}-2c_2$ $\frac{1}{3}-c_1$	$\frac{3}{4}-3c_2$ $\frac{1}{4}-c_1$	$\frac{4}{5}-4c_2$ $\frac{1}{4}-c_1$	$\frac{5}{6}-5c_2$ $\frac{1}{5}-c_1$	
2	$\frac{1}{3}-c_2$ $\frac{2}{3}-2c_1$	$\frac{1}{2}-2c_2$ $\frac{2}{4}-2c_1$	$\frac{3}{5}-3c_2$ $\frac{2}{5}-2c_1$	$\frac{4}{6}-4c_2$ $\frac{2}{6}-2c_1$	$\frac{5}{7}-5c_2$ $\frac{2}{7}-2c_1$	
3	$\frac{1}{4}-c_2$ $\frac{3}{4}-3c_1$	$\frac{2}{5}-2c_2$ $\frac{3}{5}-3c_1$	$\frac{1}{2}-3c_2$ $\frac{1}{2}-3c_1$	$\frac{4}{7}-4c_2$ $\frac{3}{7}-3c_1$	$\frac{5}{8}-5c_2$ $\frac{3}{8}-3c_1$	
4	$\frac{1}{5}-c_2$ $\frac{4}{5}-4c_1$	$\frac{2}{6}-2c_2$ $\frac{4}{6}-4c_1$	$\frac{3}{7}-3c_2$ $\frac{4}{7}-4c_1$	$\frac{1}{2}-4c_2$ $\frac{1}{2}-4c_1$	$\frac{5}{9}-5c_2$ $\frac{4}{9}-4c_1$	

Figure 7. Game 3.

Game 3 is obtained from Game 2 by simply substracting the sampling costs which are proportional to the window sizes.

The unique Nash equilibrium of Game 3 is given by,

$$k^* = \frac{c_2}{(c_1+c_2)^2}$$

$$w^* = \frac{c_1}{(c_1+c_2)^2}$$

The bargaining powers are inversely related to the relative costs:

$$\frac{k^*}{w^*} = \frac{c_2}{c_1}$$

When the marginal costs are the same, $c_1 = c_2$, we will observe the fifty-fifty division. In this case there is social inefficiency because players incurr in a costs of sampling which are saved in the case in which players follow the very simple imitative behaviour we have assumed. We have an evolutionarily stable sample size ($k = w = 1$) which is not a Nash equilibrium of the game in samples sizes. The reason is that players, when changing their sample sizes, do not take into account neither the effect of their action in the long-run nor any strategic consideration. Both populations could be better-off if they agreed on sampling only one unit of information. In this case they would save the sampling costs getting half of the cake. Both parts have incentives to deviate from such an agreement. It is prissoner's dilemma situation.

An economy populated by myopic players is more efficient that one in which strategic considerations are taken into account and intra-population coordination is possible.

When the marginal costs are different the two populations get different shares, the higher one being received by those which have the smallest marginal cost.

5.2 Non asymptotic results.

In section 3 we have obtained a relation between rates of noise and differences in expected payoffs to two different sample sizes. If sample sizes are costly, the same relation defines a locus of noise rates and differential costs which makes players indifferent between two different windows.

Let $c(k)$ ($c'(k) > 0$) be the cost of keeping a window of size $k > 0$. For each ϵ and two given sample sizes k and $k' < k$ we can find a function $d(\epsilon, k, k')$, such that

$$if \;\; c(k) - c(k') = d(\epsilon; k, k') \quad then \quad \pi(\epsilon; k) - c(k) = \pi(\epsilon, z; k') - c(k')$$

Clearly, $d(\epsilon; k, k') = D\pi(\epsilon; k, k')$.

The analysis of the evolution of learning rules for non-negligible rates of noise requieres the study of the evolution of the whole $D\pi$ function. The asymptotic results do not hold. For any difference in sampling costs we can find rates of noise for which small sample sizes are more profitable, as well as other noise rates for which the largest sample is prefered. The characterization of the evolutionarily stable sample sizes requires a better understanding of how changes in the proportions of sample sizes in the populations shift the $D\pi$ function.

Consider Figure 4, the rate of noise $\epsilon = \hat{\epsilon}$ and the difference in sampling costs $c(k) - c(k') = d$

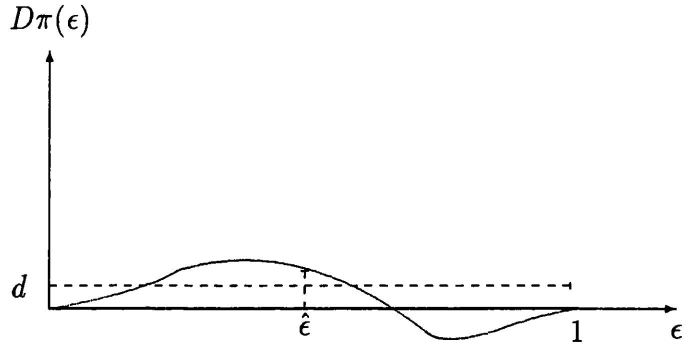

Figure 8.

the largest sample size k has, at $\hat{\epsilon}$, higher expected payoff

$$\pi(\hat{\epsilon}; k) - c(k) > \pi(\hat{\epsilon}, z; k') - c(k')$$

The proportion of player using k will grow (change in θ and f). If the change in the distribution of sample sizes f moves $D\pi$ upwards, the system will end up with an homogeneous population of k-players. Instead, if $D\pi$ moves downwards it may happen that the process of growth of k-users stops. This will occur if the new $D\pi$ falls below d at $\hat{\epsilon}$. In this last case we could have, in the ultra long-run, people in the same population using different window sizes.

As we have mentioned in section 3, the simulations of the model suggest that the advantadge of the large samples, for given noise rate, is greater the greater is its proportion in the population. This would imply that $D\pi$ moves upwards. In appendix 3 we report the results of the simulation described in section 3. The proportion of players sampling $k = 3$ and $k' = 1$ are θ

and $(1-\theta)$ respectively. All members from population II sample $w = 2$. The memory is of size $m = 4$ and $\epsilon = 0.8$. The tables 19-31 report the long run distributions for different compositions in population I. We can observe that as θ increases (the proportion of k-strategist grows) the long run distribution put more weight on those states where the larger sample size has higher expected payoff. The following table reports the difference in payoffs for $\epsilon = 0.8$ and different values of θ.

θ	$D\pi(0.8; 3, 1)$
0	-0.00008059
0.05	-0.000071324
0.1	-0.000062045
0.2	-0.00004347
0.3	-0.000024876
0.4	$-6.261\,10^{-6}$
0.5	0.00001237
0.6	0.00003102
0.7	0.00004970
0.8	0.0000684
0.9	0.000087115
0.95	0.000096481
1	0.00010585

Table 6.

The population will evolve, depending on the initial value of θ, towards $\theta = 1$ (everybody sampling k) or towards a $\theta = 0$ (everybody sampling k'). For small initial proportions of k'-players the smallest sample size performs better and will invade the population. The opposite is true for high enough $\theta's$. No general statement can be done about evolutionary stability for non-negligible rates of noise. The ultra long run distributions will depend on the initial distribution of sample sizes, the sampling costs and the rate of noise.

6 Conclusions

In this chapter we have developed an evolutionary model of bargaining with endogenous bargaining powers. In the model there are two levels of evolution and noise. On the one hand there is the evolution of the state of the system which is driven by the adaptive play and the trembles affecting players' demands. On the other hand there is the evolution of the distribution of window sizes which is continuously perturbed by mutants who employ different learning rules. When the second level of evolution is absent our model is observationally equivalent to Young's. In this case the model predicts the negative externality exerted by poorly informed players on the whole population. This result, which is obtained under the assumption of fixed samples sizes, leaves unexplained the main determinant of the bargaining powers. The model does not explain either the co-existence of different behaviours in the populations. By allowing players to imitate more succesful behaviours we endogenize the bargaining powers. We show that there will be a tendency towards homogeneous populations. All members from the same populations will, in the ultra-long run, receive the same share. It will happen not because there is a marginal player who determines the share received by everybody but because all players behave the same way. When sampling is costless both populations tend to be informed as much as they can. If there are no differences in the informational capacities of the two populations the process converges to the symmetric Nash bargaining solution. The same is true when sampling is costly, though in this case both populations sample only one unit of information. Any asymmetry in the populations' sampling costs are not reflected in the shares received. When we compare the results with those obtained with populations of rational players we observe that the economy of myopic imitative players is more efficient. The main problem with the model is that all the results are obtained in the limiting case of very small rates of mistakes. More interesting situations are those in which the rate of mistakes are not necessarily small. In this case the symptotic results do not hold and we cannot characterize the long run distributions. Some simulations seem to suggest that a closer study of the relation between the rate of mistakes and the sampling costs is needed in order to characterize the evolutionarily stable learning rules.

7 Appendix 1

Let $p_{(i,j)}^{(k,l)}$ be the transition probability between state (i,j) and (k,l). Let us consider a memory size m and a state (m,j) and consider thet all players sample 1 unit of information.

$$p_{(m,j)}^{(m,j)} = \alpha(m,j)\delta(m,j)$$

$$p_{(m,j-1)}^{(m,j)} = \beta(m,j-1)\delta(m,j-1)$$

$$p_{(m,j+1)}^{(m,j)} = \gamma(m,j+1)\delta(m,j+1)$$

$$p_{(m-1,j)}^{(m,j)} = \alpha(m-1,j)\phi(m-1,j)$$

$$p_{(m-1,j-1)}^{(m,j)} = \beta(m-1,j-1)\phi(m-1,j-1)$$

$$p_{(m-1,j+1)}^{(m,j)} = \gamma(m-1,j+1)\phi(m-1,j+1)$$

where

$$\alpha(i,j) = (1-\epsilon)(\frac{m-i}{m}\frac{j}{m} + \frac{i}{m}\frac{m-j}{m}) + \frac{\epsilon}{2}$$

$$\beta(i,j) = (1-\epsilon)\frac{m-i}{m}\frac{m-j}{m} + \frac{m-j}{m}\frac{\epsilon}{2}$$

$$\gamma(i,j) = (1-\epsilon)\frac{i}{m}\frac{j}{m} + \frac{j}{m}\frac{\epsilon}{2}$$

$$\delta(i,j) = (1-\epsilon)(\frac{m-j}{m}\frac{i}{m} + \frac{j}{m}\frac{m-i}{m}) + \frac{\epsilon}{2}$$

$$\phi(i,j) = (1-\epsilon)\frac{m-j}{m}\frac{m-i}{m} + \frac{m-i}{m}\frac{\epsilon}{2}$$

Let $\mu = \mu(\epsilon, m)$ be the long run distribution.

$$\begin{aligned}\mu_{(m,j)} = \; & \mu_{(m,j)}p_{(m,j)}^{(m,j)} + \mu_{(m,j-1)}p_{(m,j-1)}^{(m,j)} + \\ & \mu_{(m,j+1)}p_{(m,j+1)}^{(m,j)} + \mu_{(m-1,j)}p_{(m-1,j)}^{(m,j)} + \\ & \mu_{(m-1,j-1)}p_{(m-1,j-1)}^{(m,j)} + \mu_{(m-1,j+1)}p_{(m-1,j+1)}^{(m,j)}\end{aligned}$$

Let $\mu^m_{(i,j)} = \lim_{m\to\infty} \mu_{(i,j)}$

$$\mu^m_{(m,m)} = \mu^m_{(m,m)}(\frac{\epsilon}{2})^2 + \mu^m_{(m,m-1)}(\frac{\epsilon}{2})^2$$

$$lim_{\epsilon\to 0}\frac{\mu^*_{(m,m)}}{\mu^*_{(m,m-1)}} = 0$$

Solving recursively, we obtain

$$lim_{\epsilon\to 0}\frac{\mu^*_{(m,j)}}{\mu^*_{(m,j-1)}} = 0$$

We can always find an $\tilde{\epsilon}$ such that for all $\epsilon \leq \tilde{\epsilon}$, the larger window always pays best.

8 Appendix 2

	0	1	2	3	4
0	$3.8\,10^{-6}$	0.0000446	0.000187	0.000821	0.00466
1	0.000186	0.000706	0.000908	0.000697	0.000331
2	0.00532	0.00535	0.00171	0.000285	0.0000213
3	0.1	0.0247	0.00198	0.00008	$1.24\,10^{-6}$
4	0.785	0.0656	0.00171	0.0000151	$3.17\,10^{-8}$

Table 7. $\epsilon = 0.05,\ p = 0.98149$

	0	1	2	3	4
0	0.0000186	0.000184	0.000749	0.00265	0.00745
1	0.000791	0.0026	0.00344	0.00258	0.00111
2	0.0156	0.016	0.00635	0.00122	0.000105
3	0.158	0.0565	0.0073	0.000421	$8.43\,10^{-6}$
4	0.607	0.104	0.00561	0.000102	$4.28\,10^{-7}$

Table 8. $\epsilon = 0.1,\ p = 0.939487$

	0	1	2	3	4
0	0.000109	0.000856	0.00317	0.00801	0.011
1	0.00327	0.00998	0.0136	0.00987	0.00362
2	0.0371	0.0464	0.0239	0.00576	0.000599
3	0.19	0.113	0.0247	0.00233	0.0000758
4	0.35	0.128	0.0148	0.000597	$6.15\,10^{-6}$

Table 9. $\epsilon = 0.2,\ p = 0.81265$

	0	1	2	3	4
0	0.000299	0.00207	0.00684	0.0134	0.012
1	0.00654	0.0204	0.0285	0.0201	0.00654
2	0.0503	0.0768	0.0479	0.0138	0.00165
3	0.166	0.142	0.0445	0.00596	0.000284
4	0.196	0.114	0.0215	0.00146	0.0000286

Table 10. $\epsilon = 0.3,\ p = 0.671295$

	0	1	2	3	4
0	0.000583	0.0037	0.0109	0.0172	0.0114
1	0.00961	0.0313	0.0445	0.0308	0.00916
2	0.0543	0.099	0.0724	0.0242	0.00326
3	0.129	0.148	0.0614	0.011	0.000716
4	0.109	0.0909	0.0248	0.00256	0.0000835

Table 11. $\epsilon = 0.4,\ p = 0.545698$

	0	1	2	3	4
0	0.000945	0.00558	0.0147	0.0194	0.0101
1	0.012	0.0411	0.0594	0.0405	0.0113
2	0.0522	0.112	0.094	0.0358	0.00539
3	0.0948	0.139	0.074	0.017	0.00146
4	0.061	0.0686	0.0258	0.00382	0.000193

Table 12. $\epsilon = 0.5,\ p = 0.444821$

	0	1	2	3	4
0	0.00137	0.0076	0.0179	0.0201	0.00861
1	0.0137	0.0491	0.0718	0.0485	0.0128
2	0.0471	0.117	0.111	0.048	0.00799
3	0.0677	0.124	0.0826	0.024	0.00263
4	0.0345	0.0508	0.0256	0.00526	0.00039

Table 13. $\epsilon = 0.6$, $p = 0.367024$

	0	1	2	3	4
0	0.00187	0.00966	0.0203	0.0198	0.00719
1	0.0148	0.0552	0.0815	0.0547	0.014
2	0.0409	0.116	0.125	0.0601	0.0111
3	0.0476	0.107	0.0881	0.0319	0.00437
4	0.0198	0.0374	0.0249	0.00698	0.00073

Table 14. $\epsilon = 0.7$, $p = 0.307522$

	0	1	2	3	4
0	0.00245	0.0117	0.0221	0.0188	0.00592
1	0.0154	0.0594	0.0884	0.059	0.0148
2	0.0346	0.111	0.134	0.0719	0.0146
3	0.0332	0.0907	0.0914	0.0408	0.00689
4	0.0115	0.0277	0.0241	0.00912	0.0013

Table 15. $\epsilon = 0.8,\ p = 0.261521$

	0	1	2	3	4
0	0.00312	0.0137	0.0231	0.0173	0.00483
1	0.0157	0.0618	0.0924	0.0616	0.0153
2	0.0287	0.103	0.139	0.0832	0.0187
3	0.0229	0.0757	0.0932	0.0509	0.0105
4	0.00668	0.0207	0.0236	0.0119	0.00227

Table 16. $\epsilon = 0.9,\ p = 0.225114$

	0	1	2	3	4
0	0.0035	0.0147	0.0234	0.0165	0.00435
1	0.0157	0.0623	0.0934	0.0622	0.0155
2	0.026	0.0986	0.14	0.0886	0.021
3	0.0189	0.0689	0.0936	0.0565	0.0128
4	0.00511	0.018	0.0235	0.0136	0.00298

Table 17. $\epsilon = 0.95,\ p = 0.209543$

	0	1	2	3	4
0	0.00391	0.0156	0.0234	0.0156	0.00391
1	0.0156	0.0625	0.0938	0.0625	0.0156
2	0.0234	0.0938	0.141	0.0938	0.0234
3	0.0156	0.0625	0.0938	0.0625	0.0156
4	0.00391	0.0156	0.0234	0.0156	0.00391

Table 18. $\epsilon = 1,\ p = 0.195312$

9 Appendix 3

	0	1	2	3	4
0	0.0026561	0.012436	0.022212	0.017972	0.0055667
1	0.016105	0.061598	0.089242	0.058158	0.014406
2	0.03483	0.11154	0.13455	0.072529	0.014752
3	0.032162	0.088325	0.091026	0.04172	0.0071743
4	0.010718	0.025982	0.023517	0.0094132	0.0014047

Table 19. $\epsilon = 0.8$, $\theta = 0$

	0	1	2	3	4
0	0.0026349	0.012362	0.022193	0.01805	0.0056012
1	0.016037	0.061375	0.089153	0.058241	0.014444
2	0.034812	0.11146	0.13447	0.072473	0.01474
3	0.032263	0.088563	0.091067	0.041623	0.0071453
4	0.01079	0.026151	0.023575	0.0093823	0.0013944

Table 20. $\epsilon = 0.8$, $\theta = 0.05$

	0	1	2	3	4
0	0.0026139	0.012289	0.022175	0.018128	0.0056359
1	0.015968	0.061152	0.089064	0.058325	0.014483
2	0.034793	0.11138	0.13438	0.072416	0.014729
3	0.032364	0.088801	0.091107	0.041526	0.0071163
4	0.010863	0.026322	0.023633	0.0093518	0.0013843

Table 21. $\epsilon = 0.8$, $\theta = 0.1$

	0	1	2	3	4
0	0.0025721	0.012144	0.022143	0.018288	0.0057059
1	0.015831	0.060706	0.088886	0.058491	0.01456
2	0.034754	0.11121	0.1342	0.0723	0.014706
3	0.032566	0.089275	0.091188	0.041333	0.0070585
4	0.011009	0.026666	0.023752	0.0092916	0.0013641

Table 22. $\epsilon = 0.8$, $\theta = 0.2$

	0	1	2	3	4
0	0.0025308	0.012002	0.022114	0.018449	0.0057767
1	0.015693	0.06026	0.088708	0.058658	0.014638
2	0.034713	0.11103	0.13401	0.072181	0.014682
3	0.032767	0.089748	0.091269	0.04114	0.0070009
4	0.011156	0.027015	0.023876	0.0092325	0.0013441

Table 23. $\epsilon = 0.8$, $\theta = 0.3$

	0	1	2	3	4
0	0.00249	0.011861	0.022088	0.018613	0.0058482
1	0.015556	0.059815	0.088531	0.058823	0.014715
2	0.03467	0.11085	0.13382	0.072059	0.014658
3	0.032967	0.090219	0.09135	0.040948	0.0069435
4	0.011305	0.027368	0.024003	0.0091747	0.0013243

Table 24. $\epsilon = 0.8$, $\theta = 0.4$

	0	1	2	3	4
0	0.0024496	0.011722	0.022065	0.01878	0.0059206
1	0.015419	0.059371	0.088353	0.058989	0.014792
2	0.034625	0.11066	0.13362	0.071935	0.014633
3	0.033167	0.090689	0.09143	0.040757	0.0068862
4	0.011456	0.027727	0.024134	0.009118	0.0013048

Table 25. $\epsilon = 0.8$, $\theta = 0.5$

	0	1	2	3	4
0	0.0024097	0.011586	0.022046	0.01895	0.0059937
1	0.015282	0.058928	0.088176	0.059153	0.014869
2	0.034577	0.11046	0.13341	0.071808	0.014608
3	0.033367	0.091157	0.09151	0.040566	0.006829
4	0.011609	0.028091	0.024269	0.0090625	0.0012854

Table 26. $\epsilon = 0.8$, $\theta = 0.6$

	0	1	2	3	4
0	0.0023702	0.011452	0.022031	0.019122	0.0060677
1	0.015145	0.058485	0.087999	0.059318	0.014946
2	0.034527	0.11026	0.13319	0.071678	0.014582
3	0.033565	0.091623	0.091589	0.040376	0.0067721
4	0.011763	0.02846	0.024408	0.0090081	0.0012662

Table 27. $\epsilon = 0.8$, $\theta = 0.7$

	0	1	2	3	4
0	0.0023312	0.011319	0.022019	0.019297	0.0061425
1	0.015008	0.058044	0.087822	0.059481	0.015023
2	0.034475	0.11004	0.13297	0.071546	0.014556
3	0.033763	0.092086	0.091668	0.040187	0.0067154
4	0.011919	0.028833	0.024551	0.0089548	0.0012472

Table 28. $\epsilon = 0.8$, $\theta = 0.8$

	0	1	2	3	4
0	0.0022926	0.011189	0.02201	0.019474	0.0062181
1	0.014872	0.057603	0.087645	0.059644	0.0151
2	0.03442	0.10982	0.13274	0.07141	0.014529
3	0.03396	0.092548	0.091746	0.039998	0.0066588
4	0.012076	0.029212	0.024698	0.0089027	0.0012284

Table 29. $\epsilon = 0.8$, $\theta = 0.9$

	0	1	2	3	4
0	0.0022735	0.011125	0.022007	0.019564	0.0062563
1	0.014804	0.057383	0.087557	0.059726	0.015139
2	0.034392	0.10971	0.13262	0.071342	0.014516
3	0.034059	0.092778	0.091785	0.039904	0.0066306
4	0.012155	0.029404	0.024773	0.0088771	0.0012191

Table 30. $\epsilon = 0.8$, $\theta = 0.95$

	0	1	2	3	4
0	0.0022545	0.011061	0.022004	0.019655	0.0062946
1	0.014735	0.057163	0.087469	0.059807	0.015177
2	0.034364	0.10959	0.13251	0.071273	0.014502
3	0.034157	0.093007	0.091824	0.03981	0.0066024
4	0.012235	0.029596	0.024849	0.0088517	0.0012098

Table 31. $\epsilon = 0.8$, $\theta = 1$

Chapter 3

Replicator Dynamics and Rationality in Asymmetric Games.

1 Introduction.

A common concept in evolutionary game theory is that of an *evolutionarily stable strategy* (ESS) (Maynard Smith and Price (1972)) originally used in the study of animal conflicts. 'An ESS is a stategy, or phenotype, with the following property: if almost all the members of a population adopt that strategy, no alternative strategy arising by mutation can invade the population. In other words no other strategy can have as high a fitness: an ESS is a strategy that does well surrounded by copies of itself. Clearly if a population cames to consist of individuals adopting the ESS, it will cease to evolve' (Maynard Smith (1993)).

The standard definition of an ESS applies to a symmetric game. Selten (1980) shows that its natural extension to asymmetric games turns out to be very restrictive, since the concept then simply characterizes the strict Nash equilibria of the game. In order to prove his result Selten symmetrises an asymmetric game by allowing nature to make the first move by assigning a player-role to each agent at random. He shows that no mixed or polymorphic population equilibrium satisfies the ESS requirements.

Selten's result has implications for problems of equilibrium selection in an evolutionary environment. In evolutionary models, the growth rate of each strategy is positively correlated with the current relative success (fitness) of each strategy. The result suggests that mixed Nash equilibria are not in general stable long-run outcomes of an evolutionary process, when mutations are allowed. In principle, mixed equilibria could be attained by large populations in which different agents in the same population play different pure strategies. However, such outcomes are not robust in asymmetric games against the introduction of mutant strategies.

We have the following issues:

> When an asymmetric game has no Nash equilibria in pure strategies no ESS exists;
>
> When an asymmetric game has both a strict Nash equilibrium and a mixed Nash equilibrium the former is more robust even though the latter may be Pareto–superior. Socially preferred outcomes might not be selected as long-run states of some dynamic systems.

We would like to stress the fact that we can find 'darwinian' dynamic processes for which mixed Nash equilibria are asymptotic equilibria of such processes. An interesting feature of all these examples is that the mixed Nash equilibria are, under Darwinian dynamics, either centers, saddle points or foci but never nodes.

In games played by different populations, the evolution in a population depends on the strategy distribution in the other. It is a well known fact that all the strategies which are played in a mixed equilibrium equilibrium obtain the same payoff. This implies that in an evolutionary enviroment, any mutation with the same support of the equilibrium point will not be 'corrected' by the Darwinian selection in the population where the mutation has taken place. A mutation in one population will induce, however, a reaction in the other population. The return to the original equilibrium is not guaranteed under general darwinian dynamics.

In this paper we show that there exists a class of polymorphic Nash equilibria in asymmetric games that exhibits stability properties which are in some way similar to those exhibited by an ESS, provided that a small fraction of the population can anticipate others' behaviour. First, we draw from Hofbauer and Sigmund (1988) the concept of Nash-Pareto Pair in a two-population asymmetric game. Hofbauer and Sigmund prove that under the replicator dynamics, if a small deviation from a Nash–Pareto pair occurs, then evolution will not push the system further away from the equilibrium (whereas in ESS the evolution pushes the system back towards the equilibrium).

Second, we assume that each population consists of two types of agents,to be called *myopics* and *rationals.* Myopic players follow some imitative behaviour of the kind commonly found in evolutionary game theory, resulting in an increase over time of currently successful strategies. We can rationalise this behaviour in terms of bounded rationality. Myopics can be thought of as people with high costs in collecting and processing information who find it convenient to behave conservatively by sticking to their current strategy, and only periodically deciding whether or not to revise their behaviour by imitating more successful strategies. Rational players, instead, correctly anticipate changes in the environment, responding optimally to the one-stage game played each period. We assume, as always in standard discussions of the replicator dynamics, that the populations are large. The distinction between myopics and rationals can be envisaged as representing, in a highly stylised and extreme form, a world in which agents face different costs in collecting and processing information, or in the adjustment cost for changing strategies.

The idea of two classes of players along these lines was first introduced by Banerjee and Weibull (1991) in a paper intended to show the possibility of survival of non-rational agents in strategic environments. Apart from the divergent motivations, their paper differs from ours in that they consider only symmetric games with a single population and apply replicator dynamics to the evolution of rational and myopic players in the population, whereas we assume that the proportions of types in each population are fixed.

In the first part of the paper we show that when some positive fraction of agents in each population behave rationally, there exists a set of distributions of myopic agents across strategies consistent with a polymorphic Nash equilibrium. The result holds irrespective of how small or large the fraction of rational players is assumed to be. Under the same assumptions we will show that any deviation from a Nash Pareto pair never makes the deviant population (including both rational and myopic agents) better off on average. This second result relates to the main finding of the second section in which we study the effect of the introduction of rational players on the behaviour of the system under replicator dynamics. We will show that under replicator dynamics an Nash–Pareto pair is asymptotically stable in the presence of rational players.

2 An evolutionary game.

Consider a two-player normal-form game $G = \langle S_x, S_y, A, B \rangle$, where S_x is the set of n pure strategies available to player 1, S_y is the set of m pure strategies available to player 2, A and B are the respective pay-off matrices. We assume that there are two large populations $\mathcal{X}$ and $\mathcal{Y}$ and that the members of these populations are randomly matched in pairs to play the game G. Some fraction of the agents in each population behave myopically, playing a fixed strategy and only occasionally considering a revision, on the basis of the observation of the degree of fitness (relative success) of the alternative strategies. At this stage we will not commit ourselves to any particular mechanism for how such revisions are made. We will only assume that the growth of each strategy has the same rank-order as the respective fitnesses (Friedman (1992)). A certain fraction of agents in each population–no matter how small or large–behaves rationally, in the sense that they maximize their expected pay-off in each one-shot game and correctly anticipate the proportion of players who are choosing each strategy in the opponent population. No agent is able, however, to recognise the strategy that is going to be played by his opponent with whom he is individually matched at each stage.

In formal terms, we assume that a fraction σ $(0 < \sigma < 1)$ of rationals in population $\mathcal{X}$ is fixed. Similarly for the fraction τ $(0 < \tau < 1)$ of rationals in population $\mathcal{Y}$. Let Δ_n be the n-dimensional unit simplex. The vector $x \in (1-\sigma)\Delta_n$ describes the myopic agents in population $\mathcal{X}$. The coordinate x_i is the fraction of the population who are myopics currently using strategy i. The vector $X \in \sigma\Delta_n$ similarly describes the rational agents in population $\mathcal{X}$. The vectors $y \in (1-\tau)\Delta_m$ and $Y \in \tau\Delta_m$ fulfil the same roles for population $\mathcal{Y}$.

We define $u = x + X$ to be the current vector of strategy frequencies in population $\mathcal{X}$. Similarly, $v = y + Y$ is the vector of strategy frequencies in population $\mathcal{Y}$. The vector pair (u, v) will be called the current strategy distribution. We shall say that (u, v) is a (polymorphous) Nash equilibrium if it is true that (u, v) is a (mixed) Nash equilibrium when u and v are interpreted as mixed strategies. Restrictions are placed on the admissible values of X and Y in order to justify their roles in describing the current behaviour of rational agents. We require that $X_i = 0$ unless i is a best reply to v and that $Y_j = 0$ unless j is best reply to u.

The fact that a pair (X, Y) always exists for each pair (x, y) follows from Nash's theorem on the existence of Nash equilibria for finite games.

With these understandings, a state of the system is a quadruple $\langle x, X, y, Y\rangle$. We will say that a pair (x, y) is compatible with a Nash equilibrium $(u, v) = (x + X, y + Y)$ when $\langle x, X, y, Y\rangle$ is a state of the system. As previously mentioned, we draw from Hofbauer and Sigmund (1988) the concept of a Nash-Pareto Pair. This characterises a subset of Nash equilibria satisfying certain stability conditions.

Definition 1 *A strategy distribution* $(\tilde{u}, \tilde{v})$ *is said to be a Nash-Pareto Pair* *if and only if the following conditions hold:*

(i) $(\tilde{u}, \tilde{v})$ *is a Nash equilibrium, so that*

$$\begin{aligned} \tilde{u}^T A \tilde{v} &\geq u^T A \tilde{v} \text{ (for all } u\text{) and} \\ \tilde{v}^T B^T \tilde{u} &\geq v^T B^T \tilde{u} \text{ (for all } v\text{)} \end{aligned}$$

(ii) stability condition: for all states $(u, v) \in \Delta_n \times \Delta_m$ *such that equality holds in (i):*

$$\begin{aligned} u^T A v > \tilde{u}^T A v &\Rightarrow v^T B^T u < \tilde{v}^T B^T u \text{ and} \\ v^T B^T u > \tilde{v}^T B^T u &\Rightarrow u^T A v < \tilde{u}^T A v \end{aligned}$$

One might summarize the stability condition by saying that if both players switch to alternative best replies then at least one would have done better sticking to the previous strategy.

The following proposition shows that, if a deviation occurs when the system is at a Nash -Pareto pair, the deviant population is never better off than playing the equilibrium strategy. This fact has an interesting implication for the dynamics of the system because the fact that the rational players do not follow the mutants will render our system asymptotically stable.

Proposition 1 *Let* $(\tilde{u}, \tilde{v})$ *be a completely mixed Nash Pareto pair and* $\langle \tilde{x}, \tilde{X}, \tilde{y}, \tilde{Y} \rangle$ *be the state of the system. Consider a new state* $\langle x, X, \tilde{y}, Y \rangle$. *Then:*

$$(x, \tilde{y}) \text{ incompatible with } (\tilde{u}, \tilde{v}) \Rightarrow u^T A v \leq \tilde{u}^T A v,$$

where (u, v) *is the strategy distribution when the state of the system is* $\langle x, X, \tilde{y}, Y \rangle$.

Proof. The proof is by contradiction.Note first that any $u \in \Delta_n$ is an alternative best reply to $\tilde{u}$ because $\tilde{u}$ is completely mixed. Similarly any $v \in \Delta_m$ is an alternative best reply to $\tilde{v}$ because $\tilde{v}$ is completely mixed. We need only consider condition (ii) in the definition of Nash–Pareto pair . Assume that $(x, \tilde{y})$ is incompatible with $(\tilde{u}, \tilde{v})$ but $\tilde{u}^T A v < u^T A v$. Since

$(\tilde{u}, \tilde{v})$ is a Nash–Pareto pair , then $\tilde{u}^T Av < u^T Av$ implies $v^T B^T u < \tilde{v} B^T u$. But $v = (\tilde{y} + Y)$ and $\tilde{v} = (\tilde{y} + \tilde{Y})$. The rationality condition implies that $y^T Au \geq \tilde{y}^T Au$, and hence $v^T Au \geq \tilde{v}^T Au$. This says that $(\tilde{u}, \tilde{v})$ is a Nash–Pareto pair . □

A result by Schuster and Sigmund (1983) shows that a completely mixed Nash–Pareto pair in 2×2 games is a stable fixed point of the asymmetric replicator dynamics. However, such a Nash–Pareto pair turns out to be only Ljapunov stable and not asymptotically stable; geometrically, it is the centre of a system of closed orbits. We will show that the presence of a minimum degree of rationality stabilises the system and makes the Nash–Pareto pair asymptotically stable.

Let us consider the generic 2×2 game with a unique and completely mixed equilibrium (Van Damme (1983)). Matching pennies is an affine transformation of this game,

	L	R
U	a, 0	0, b
D	0, c	d, 0

$a, b, c, d > 0$

Figure 1. Game 1

The unique Nash equilibriumof this game is a Nash–Pareto pair [1]. In the absence of rationals ($\sigma = \tau = 0$), the basic replicator dynamic equations would be:

$$\dot{x} = x(1-x)[(a+d)y - d]$$

[1]The other type of 2×2 with generic pay-offs which has a completely mixed equilibrium is the Battle of the Sexes. This Nash equilibrium is not a Nash–Pareto pair and the introduction of rationals desstabilizes it.

$$\dot{y} = y(1-y)[c-(c+b)x],$$

where x and y are the fraction of players using the first strategy in populations X and Y respectively.

A modified version of the replicator dynamics which allows for a positive proportion of rational players is:

$$\dot{x} = (x+X)[1-(x+X)][(a+d)(y+Y)-d] \quad (1)$$

$$\dot{y} = (y+Y)[1-(y+Y)][c-(c+d)(x+X)] \quad (2)$$

We will refer to this system as the modified replicator dynamics .

Let us consider Figure 1; it represents all possible states of myopic players, $(x,y) \in (1-\sigma)\Delta_2 \times (1-\tau)\Delta_2$, in the game above. (Moving right increases the fraction playing U and moving up increases the fraction of L).

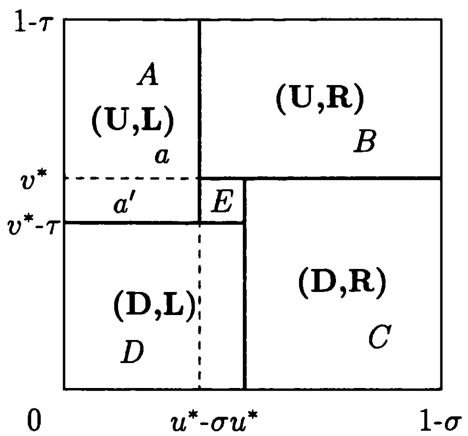

Figure 2.

Consider the behaviour of rationals when the state of the system is point $a \in A$ in Figure 2. Whatever rationals in population $\mathcal{Y}$ do, it is best reply for rationals in population $\mathcal{X}$ to play U, because the observed state $y_1 + Y_1 > v^*$. For population $\mathcal{Y}$ is optimal to play L. Let us also consider point $a' \in A$: The optimal behaviour of the rationals in $\mathcal{X}$ depends now on the choice of the rationals in $\mathcal{Y}$; if they choose L, the observed state will be $y_1 + Y_1 > v^*$ and the best reply to such state is U. But if the choice were R, then the best reply would be D because the observed state will satisfy $y_1 + Y_1 < v^*$. For the rationals in Y, L is the best reply (whatever happens in $\mathcal{X}$, the observed state will be $x_1 + X_1 > u^*$). It follows that (U,L) are the best replies in area A.

We can now write the modified replicator dynamics for any state of myopics in area A,

$$\dot{x} = (x+\sigma)[1-(x+\sigma)][(a+d)(y+\tau)-d]$$

$$\dot{y} = (y+\tau)[1-(y+\tau)][c-(c+d)(x+\sigma)]$$

The same argument applies for the other areas. In area B,

$$\dot{x} = (x+\sigma)[1-(x+\sigma)][(a+d)y-d]$$

$$\dot{y} = y(1-y)[c-(c+d)(x+\sigma)]$$

In area C,

$$\dot{x} = x(1-x)[(a+d)y-d]$$

$$\dot{y} = y(1-y)(c-(c+d)x)$$

In area D,

$$\dot{x} = x(1-x)[(a+d)(y+\tau)-d]$$

$$\dot{y} = (y+\tau)[1-(y+\tau)](c-(c+d)x)$$

This leaves area E which requires special comment. In this set of states $\dot{x} = \dot{y} = 0$, because the rationals act so as to make all strategies equally good. The observed state generates the polymorphic equilibrium in which $x + X = u^*$ and $y + Y = v^*$, with $u^* = c/(c+b)$ and $v^* = d/(d+a)$.

Proposition 2 *Let (u^*, v^*) be the completely mixed Nash–Pareto pair of Game 1. Let $\sigma > 0$ or $\tau > 0$. Then (u^*, v^*) is asymptotically stable with respect to the modified replicator dynamics .*

Proof We consider the case $\sigma > 0$, $\tau = 0$. Thus $Y_1 = 0$ in equations (1) and (2). Divide the RHS of (1) and (2) by $(y)(x+X)(1-x-X)(1-y)$. The new equations, which have the same trajectories as the old, can be rewritten as

$$\dot{x} = \frac{-d}{y} + \frac{a}{1-y} \tag{3}$$

$$\dot{y} = \frac{c}{x+X} - \frac{b}{1-(x+X)} \tag{4}$$

In a region in which X is constant, the fuction H defined by

$$H(x, y, X, 0) = c\log(x+X) + b\log(1-x-X) + d\log(y) + a\log(1-y)$$

is a constant of motion because

$$\dot{H} = \dot{x}\dot{y} - \dot{y}\dot{x} = 0$$

Thus, the system (3)-(4) is piecewise Hamiltonian.

From the expression of H, it can be verified that:

$$H(x, y, \sigma, 0) > H(x, y, 0, 0) \Leftrightarrow (x+X) < \frac{c}{b+c} = u^*$$

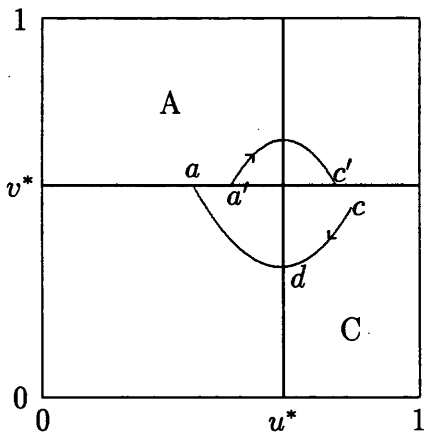

Figure 3.

Consider an initial point such that $y < v^*$ and $x + X > u^*$ (c in Figure 3). From the modified replicator dynamics equations, it follows that $X = 0$ and $dx/dt = du/dt < 0$, $dy/dt = dv/dt < 0$, while H remains constant. At some moment in time, the system will reach a state $x + X = u^*$ and $y < v^*$ (point d in Fig. 3) with $dy/dt > 0$, $dx/dt < 0$ and H constant. Once the system reaches a state in A (point a in Fig.3) with $y = v^*$ and $x + X < u^*$ the rational players, simultaneously, all change to strategy U, creating a new state of the form $\langle x, \sigma, y, 0 \rangle$ (point a') with $dy/dt > 0$, $dx/dt > 0$ and H jumps to a higher value until a new state in C (c' in Fig.3) is reached. At this point all the rationals move to play D and H jumps up again. H turns out to be a stepwise increasing function which reaches its maximum when the state of the system becomes compatible with the Nash–Pareto pair (u^*, v^*), at which $x \in [u^* - \sigma, u^*]$ and $y = v^*$). Since the system admits a nondecreasing (piecewise increasing) over time Ljapunov function H then it is an asymptotically stable fixed point of the modified replicator dynamics . The extension to the case in which $\tau > 0$ is straightforward. □

A final remark concerns the extent to which the presence of rational agents is stabilising in an evolutionary environment. We have seen that in 2×2 games with a Nash–Pareto pair , the rationals never follow the mutants and this has a stabilising effect on the equilibrium outcome. In higher-dimensional games things become more ambiguous, particularly when the Nash–Pareto pair has less than full support and other equilibria exist. In this case, if the proportion of rationals is relatively high, their presence could actually play a destabilising role. It is indeed possible to construct examples in which a mutation that would be killed off by the selection mechanism in a purely evolutionarily environment can instead invade the equilibrium because of the reaction of the rationals.

Consider the game represented in Figure 4,

	LEFT	*RIGHT*
UP	1 $2-\epsilon$	1 $2-\epsilon$
MIDDLE	2 4	4 0
DOWN	2 3	0 1

Figure 4.

For $\epsilon < 1$ this game has two Nash equilibria, one in pure strategies in which Up and Right are played with probability one and one in mixed strategies with all the strategies but Up played with probability 1/2. This last equilibrium is a Nash Pareto pair. Consider a state compatible with the mixed equilibrium and a mutation in population $\mathcal{X}$ such that the best reply for population $\mathcal{Y}$ is Right, independently of what the rationals in $\mathcal{X}$ do. If before the mutation it happened that the proportion of rationals playing Left was greater than $\epsilon/2$ then in the new population $\mathcal{Y}$'s state more than $1/2 + \epsilon/2$ will be playing Right, being Up the best strategy for $\mathcal{X}$. The system, due to the reaction of the rational players, jumps into the basin of attraction of the Pareto inferior equilibrium (Up,Right).

3 Conclusions

The presence of rational players in small proportions renders asymptotically stable some mixed equilibria in 2-player asymmetric games. The mixed equilibria are necessarily Nash Pareto pairs. This property implies that only one side can benefit from a deviation from the equilibrium strategy. Our result holds for the standard replicator dynamics but can be generalized to other dynamics. Mixed Nash Pareto pairs are characterized under darwinian dynamics by some kind of cycling behaviour around the equilibrium. By always selecting the best paying strategy the rational players move the system towards the mixed equilibrium. We only need to select the proportion of rationals large enough to amortiguate any possible divergent behaviour.

Chapter 4

An evolutionary model of development of a credit market.

1 Introduction

In this chapter we construct a simple evolutionary model of credit activity, which aims at explaining the existence of credit cycles and rationing. We model the relations between firms (borrowers) and banks (lenders) as regulated by a very simple debt contract subject to the constraint of limited liability of investors. The two parts know nothing about each others' history. For reasons which will be explained, agents cannot build up an individual reputation. Investors can decide to invest the money in a productive activity or to 'take the money and run'; lenders can detect the fraudulent behaviour and enforce its punishment, but only if they engage in a costly monitoring process. The banks can invest in a low return safe asset or lend to a borrower. They charge the same interest rate to all possible borrowers because these are ex-ante indistinguisable to them. If banks decided to monitor all borrowers, potentially bad borrowers would find it optimal to invest in the productive activity, because would certainly be penalized. If all borrowers were 'good' however, banks would maximize profits by not monitoring any project. For these reasons there exists one equilibrium in which we observe a mixture of good and bad behaviour on the borrowers' side and the banks choosing to monitor only a positive proportion of the loans granted but not all of them. This is not the only Nash equilibrium. Our model, in fact, also exhibits equilibria with 'financial collapse'. When lenders expect to find many bad borrowers they invest in the safe asset and do not grant any loan. If the proportion of borrowers who would cheat should they receive a loan is large enough the banks behave optimally an so do the borrowers.

Two are the evolutionary features which we introduce into the model. The first is that we model market interactions as taking place between members of different populations who meet and play a one-shot game representing the situation just described. Players are assumed to be of two possible types: myopic and rational (as in Banerjee and Weibull (1991) and chapter 3). Myopic players behave conservatively by sticking to a strategy, and only periodically deciding whether or not to revise their behaviour by imitating more successful strategies. Rational players always play a Nash equilibrium. The co-existence of two types of players can be thought as representing, in a highly stylised and extreme form, a world in which agents face different costs in collecting and processing information, or in the adjustment cost for changing strategies. The second feature, which is related to the existence of myopic players, is that off-equilibrium dynamics play the role of selecting one out of all possible equilibria. The fact that a large part of each population changes its strategy only from time to time introduces some inertia in the system. This sluggishness implies that the scenario faced by the players tomorrow is not too different from the one they face today. In these circumstances even rational behaviour tends to exhibit some inertia.

The first result of this chapter is that equilibria without credit activity are not stable against 'experimentation'. When some banks start granting some monitored loans good behaviour is encouraged and the credit market starts developing. So, the absence of credit market does not seem to be a robust stable outcome of the system. The second and more interesting result relates the behaviour of the system under 'darwinian dynamics' to the tendency of the economy to produce episodes of 'credit crunches' and temporary financial collapse. We show that once-over 'real shocks' which affect the equilibria can trigger episodes of credit rationing and fluctuations in the activity levels throughout the process of adjustment towards the new Nash equilibrium. Consider a situation in which the rate of fraud in the economy is small enough to allow for the existence of a fully developed credit market: all borrowers are given a loan and the behaviour of banks creates the right incentives. An external shock which affects, for example the likelihood with which borrowers repay their debts or erodes the collaterals if they exist, may make that the safe asset a more profitable investment on average. The strategy followed by banks is to start lending less. The drain of investment funds, given the banks' inability to screen the borrowers, affect also to the good borrowers and starts a process of credit crunch. This is argued to be consistent with Bernanke's (1983) description of the Great Depression.

The model has important limitations, of which we must be aware. The most evident is that the borrower-lender relationship has typically a long-term nature. This cannot be captured by the random matching, one-shot game set-up which we propose in this chapter. To rule out reputational effects in a credit relation is clearly hard to swallow. However, it can be argued that the set-up is not unrealistic for the inherently risky market for loans to new investors and small firms whose access to the credit market is sporadic and the information about whose past behaviour is little reliable. Bernanke (1983) observes that this segment of the market was significant and important during the Great Depression. Also customer relations were weakened in that period by the fact that many borrowers were separated from their banks when these were forced to close. This caused a considerable amount of borrowers to seek for credits in new banks. As we will discuss there is evidence that credit rationing issues apply more significantly to these segments of the market. So, this work may be viewed as focusing on the segment of the market which is most affected by drains in investment funds during a credit crunch: small businesses and households.

The chapter is organized as follows: In the first part we introduce the model. We will present the parable of a very economy of peasants characterized by the lack of collateral wealth. In the second section we introduce the dynamics and study the behaviour of the economy when it is populated by myopic players. In the third section rational players are introduced. We relate the findings to the empirical evidence from the Great Depression. Finally we report the results of the simulations of the model.

2 Evolution of credit activity in an economy with limited collateral

We consider a stylized primitive economy in which potential investors hold no collateral. A typical problem faced by economic systems with a low amount of collateralizable wealth is that lenders are liable to large losses in the event of bankruptcy of the debtor. Furthermore, this raises an important incentive problem for borrowers.

Assume that the probability of success of an enterprise depends on some unobservable costly effort on the part of the entrepreneur. If he borrows funds with little or no collateral, he will typically put in less effort than in a world of perfect information. This happens because he is liable, in the event of bankruptcy, only for the value of his collateral, whereas the reward to its success is limited by the payment due to the lender. This is known in the literature as a debt–overhang problem with wealth constraints (Aghion and Bolton (1991)). For expositional purposes, we start by describing a standard problem of missing market due to incentive problems in a world of rational agents.

We assume that:

- all the agents are risk-neutral;

- there exists a safe asset (outside option) in the economy that provides an exogenously given interest rate;

- potential investors needs to borrow a fixed amount of money W and have no collateral at all;

- effort is a discrete choice variable; the entrepreneur may either exert effort ($e > 0$), or do nothing. To make the argument more concrete, we can think of a peasant-entrepreneur who may either use the borrowed funds to introduce technical improvements (paying the effort cost e) and so enhance the probability of a good crop, or divert the resources, buying consumption goods, and hope that good weather makes up his lack of effort.

- there are two states of the world, which are observable by everybody at the end of each period. To go back to the previous example, if the crop is good, the borrower gets revenue $H > W$, pays back the debt with the agreed interest payment (R) and earns a net profit, whereas if it is bad the borrower gets nothing and cannot pay back the debt;

- there is no equilibrium value of R that makes the borrower's choice incentive compatible, i.e. that induces the peasant to exert effort, while giving the lender a higher expected payoff than that warranted by investing in the safe asset.

We will define the following conditional probabilities:
π=Pr (good crop | no effort);
$\pi + \alpha$=Pr (good crop | effort).

We also define r as the gross payment obtained by the lender investing W in the safe activity. Let us rule out for the moment that the agents may randomise their choices and play a mixed strategy. It is then possible to choose parameters such that a credit market would exist under perfect information, but the moral hazard problem prevents its existence in a world with imperfect information. This is the case if:

$$[\pi + \alpha][H - R] > e \ and \ (\pi + \alpha)R > r, \ for \ some \ R \tag{1}$$

but:

$$\pi R < r, \ allR \tag{2}$$

and:

$$\pi[H - R] > (\pi + \alpha)[H - R] - e \tag{3}$$

for all R's that satisfy (1).

Equation 1 guarantees that under perfect information (observable effort) there would be credit activity for some range of values of R, with borrowers exerting the effort. Equation 2 says, however, that if no effort is made by the borrower, there is no R at which lenders are willing to grant credit. Equation 3 implies that for all R's satisfying 1 there is an incentive problem (the expected payoff of the peasant is higher when he shirks). The source of this standard missing market problem is the absence of collateral to be held by investors.

The following additional assumptions are needed to fill out the model:

- there are in the market agents who act myopically, by playing an arbitrary strategy which they possibly revise periodically according to some imitative adaptive rule;

- lenders are entitled to monitor the activity of the borrowers. If cheating is detected, the lender asks for his money back (without earning interest) and the borrower is liable to legal prosecution, with a high loss in terms of utility. However, monitoring entails a cost M. We can imagine that lenders delegate and pay some specialised institution for this purpose. Monitoring reduces the risk involved in lending. In order to focus on the interesting case, we will assume that, when no cheating is detected, the payoff to the lender, net of the cost M, is still higher than if W were invested in the safe asset;

- interactions are anonymous in the sense that agents are randomly matched in a world with large populations, and so are unable to use what they know about their opponents' identity to predict their behaviour.

If monitoring the investors is costly, not all the loans will be monitored in an equilibrium with developed credit activity. Once the threat of monitoring has induced all the entrepreneurs to be 'honest', any single lender no longer has an incentive to pay the monitoring cost. Every lender would like to free-ride and exploit the fact that there is a widespread belief that any cheating will be detected and punished.

Since several strategies will coexist at any time, including strategies which are not current best–replies and agents are randomly matched, the expected payoff to each strategy will depend on the probability of matching with each of the strategies played by the opponent population. In a world with many honest borrowers, to lend without monitoring is likely to be a successful strategy. In a world with almost all dishonest borrowers the best one can do is to invest in the safe asset. It is important to remark that in our world there are agents who go to the market with some strategy in mind and only periodically revise it, imitating more successful behaviours according to some rule to be specified. For this reason, in our framework, the model is closer to the class of adverse selection problems than to that of moral hazard problems[1]. Lenders' decisions are affected by the knowledge that

[1]However, there are no trivial contractual arrangements that sustain a 'separating equilibrium', since borrowers are not good or bad as a result of some intrinsic personal characteristic, like a different disutility of exerting effort. If we like, we can think that they differ in their beliefs about the probability of being detected if they cheat. This difference cannot be exploited by the principals, by proposing distinct contractual solutions, each of which attracts a different type.

at each period there both 'good' and 'bad' types in the market that are undistinguishable ex-ante.

In the next section we formalise our model as an asymmetric normal form game with two populations and random matching.

3 Game representation and dynamic evolution of a developing credit market

Consider the following extensive form corresponding to the situation described above:

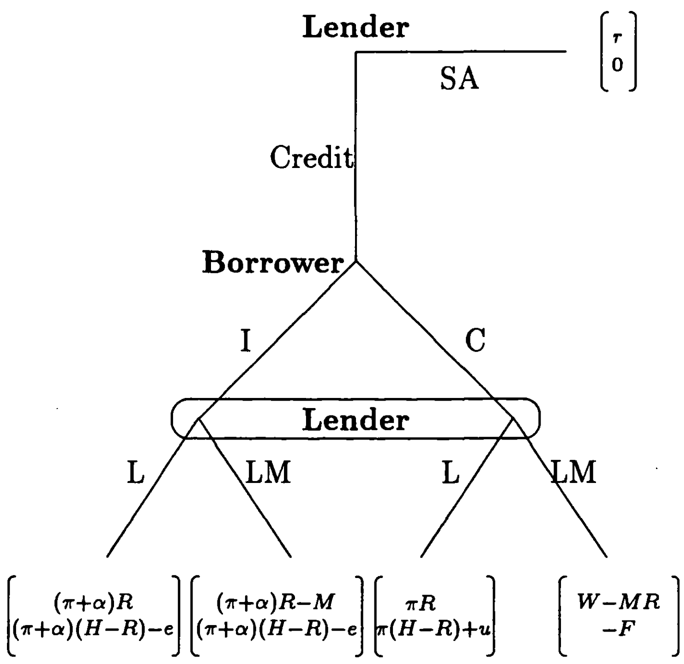

Figure 1. Extensive form

The lender decides whether to invest in a safe asset or to grant a loan. If the loan is granted the lender chooses between monitoring or not. The borrower, when he is given a loan, decides either to invest or to cheat.

The corresponding normal–form of the game is:

	I	C
SA	r ; 0	r ; 0
L	$(\pi+\alpha)R$; $(\pi+\alpha)(H-R)-e$	πR ; $\pi(H-R)+u$
LM	$(\pi+\alpha)R-M$; $(\pi+\alpha)(H-R)-e$	$W-M$; $-F$

Figure 2: Game G.

The three rows correspond to strategies SA (safe asset), L (loan) and LM (loan and monitor) respectively. The two columns correspond to strategies I (invest) and C (cheat).

The interpretation of each entry is straightforward. For example, strategy L matched with strategy I gives an expected payoff of $(\pi+\alpha)R$ to lenders (the probability of success for a honest entrepreneur times the payment agreed in the case of success) and $(\pi+\alpha)(H-R)-e$ to investors (the probability of success when effort is exerted times the net profit minus the effort cost). The quantity u represents the utility to borrowers of consuming the borrowed funds rather than investing them. The quantity F represents the disutility from the prosecution in the case of being caught when cheating.

Some notational conventions will be used in the formal discussion. Population $\mathcal{X}$ is the set of potential lenders. A vector $\mathbf{x} = (x_{SA}, x_L, x_{LM})$ records the fractions of individuals in $\mathcal{X}$ using each available strategy. We define $p = x_{SA}$, $q = x_L$ so that $x_{LM} = 1 - p - q$. A state of $\mathcal{X}$ is then identified with a pair (p,q). Population $\mathcal{Y}$ is the set of potential borrowers. A population state (y_I, y_C) represents the fractions of borrowers using each of the two available strategies. We will identify a state in population $\mathcal{Y}$ by $y = y_I$ so that $y_C = 1 - y$.

We make the following assumptions about the parameters of the payoff matrices:

$$(\pi + \alpha)R - M > r \;>\; W - M > \pi R \qquad (C1)$$

$$(W - M - \pi R)((\pi + \alpha)R - W) \;>\; (r - (W - M))(W - \pi R) \qquad (C2)$$

$$(\pi + \alpha)(H - R) - e \;>\; -F \qquad (C3)$$

$$\alpha(H - R) - e \;<\; u \qquad (C4)$$

(C1) guarantees that when there are only honest borrowers ($y = 1$), the payoffs to the different strategies for population $\mathcal{X}$ are ordered as follows: $L > LM > SA$ whereas when there are only dishonest borrowers this ordering is strictly reversed.

(C2) and (C1) together guarantee that every strategy, SA (for low levels of y), L(for intermediate values of y) and LM (for y large), is a strict best-reply for some range of y.

(C3) guarantees that, when all loans are monitored, I is the best-reply for the borrowers.

(C4) guarantees that, when no loan is monitored, C is the best-reply for the borrowers.

We will start the analysis by describing the set of Nash equilibria of game G.

Proposition 1 *If conditions (C1)-(C4) hold, then the game G has an infinite set of Nash equilibria, with two connected components: (i) a mixed equilibrium (0, q1, y1) ,where:*

$$q_1 = \frac{(\pi + \alpha)(H - R) - e + F}{\pi(H - R) + u + F}$$

$$y_1 = \frac{(W - M) - \pi R}{W - \pi R}$$

(ii) A set of Nash of equilibria N, such that p=1 and $y \in [0, y2]$*, where:*

$$y_2 = \frac{r - (W - M)}{(\pi + \alpha)R - W}$$

The mixed equilibrium corresponds to the existence of a credit market. All of the equilibria in N imply no credit market. The problem of the existence of a credit market therefore reduces to studying an equilibrium selection problem in game theory.

It is important to notice that, since no strict Nash equilibrium exists, the game has no ESS as it has been shown by Selten (1983). It is also easy to check that under our assumptions about the payoff matrices, any equilibrium belonging to N, corresponding to the absence of a credit market, is Pareto–dominated by the singleton equilibrium.

The game has only one subgame–perfect equilibrium as can be seen by solving by backward induction. Our aim here is to investigate whether the (subgame–perfect) equilibrium with a developed credit market is also supported as the long-run outcome of an evolutionary process in which (at least part of the) agents behave myopically.

4 Replicator dynamics with myopic players

Let us start with the polar case in which all agents are myopic. It is important to note that the specification of the selection mechanism, that controls how myopic agents adjust their strategies, plays an important role. We can construct examples in which the Pareto-superior equilibrium is dynamically unstable (i.e. has a repulsive nature) as well as examples in which it is asymptotically stable. The traditional replicator dynamics are a borderline case. For the game G, they take the form:

$$\begin{aligned}
\dot{p} &= p[r - \pi_{lend}] \\
\dot{q} &= q[y(\pi + \alpha)R + \pi R(1 - y) - \pi_{lend}] \\
\dot{y} &= y[(1 - p)((\pi + \alpha)(H - R) - e) - \pi_{borrow}]
\end{aligned}$$

where π_{lend} and π_{borrow} are the average payoffs for lenders and borrowers respectively.

Proposition 2 *(Dynamical properties of N.) Each state of*

$$N' = \{(p, q, y) \in \Delta_3 \times \Delta_2 | p = 1, q = 0, y \in [0, y_2[\ \}$$

is locally Ljapunov-stable. The set N' is therefore an attractor.

Proof - (see appendix 1).

Ljapunov-stability implies that if a small perturbation occurs when the state of the system belongs to N', the system will not necesarily return to the same Nash equilibrium, but will settle down in another Nash equilibrium close to it (more precisely, belonging to N). Notice, however, that this set of equilibria is less robust and stable than an ESS.

We consider now the mixed isolated equilibrium. Let us state first its dynamic properties.

Proposition 3 *Under the conditions (C1)-(C4), the isolated Nash equilibrium* $(0, q_1, y_1)$ *is a rest point belonging to the centre (linear) manifold J*

$$J \subset \Delta_3 \times \Delta_2, \ \ J = \{(p, q, y) \in \Delta_3 \times \Delta_2 | \ p = 0\}$$

Furthermore, the centre manifold J is an attractor in the neighbourhood of $(0, q_1, y_1)$ *and, more in general, is an attractor for all points in J such that:*

$$y > y^*(q) = \frac{r - (W - M) + q(W - M - \pi R)}{(\pi + \alpha)R - W + q(W - \pi R)} < y_1$$

Proof - see appendix 2.

It is easy to see that $dy^*(q)/dq > 0$ and $d^2y^*(q)/dq^2 < 0$.

Although the isolated Nash equilibrium is not an ESS it satisfies certain stability condition common to all mixed Nash equilibria which are Nash Pareto-Pairs: Let us consider a two-player normal-form game $G = \langle 2, S_x, S_y, A, B \rangle$, where S_x is the set of n pure strategies available to player 1, S_y is the set of m pure strategies available to player 2, A and B are the respective pay-off matrices. Game G is played by two populations $\mathcal{X}$ and $\mathcal{Y}$.

Definition 1 *A strategy distribution* $(\tilde{u}, \tilde{v})$ *is said to be a Nash-Pareto Pair if and only if the following conditions hold:*

(i) $(\tilde{u}, \tilde{v})$ *is a Nash equilibrium, so that*

$$\begin{aligned} \tilde{u}^T A \tilde{v} &\geq u^T A \tilde{v} \text{ (for all } u\text{) and} \\ \tilde{v}^T B^T \tilde{u} &\geq v^T B^T \tilde{u} \text{ (for all } v\text{)} \end{aligned}$$

(ii) stability condition: for all states $(u, v) \in \Delta_n \times \Delta_m$ *such that equality holds in (i):*

$$\begin{aligned} u^T A v > \tilde{u}^T A v &\Rightarrow v^T B^T u < \tilde{v}^T B^T u \text{ and} \\ v^T B^T u > \tilde{v}^T B^T u &\Rightarrow u^T A v < \tilde{u}^T A v \end{aligned}$$

The intuition behind the stability condition is that both sides cannot be better off after a deviation. If a deviation is profitable for one side, the other is penalized. Hofbauer and Sigmund (1988) show that Ljapunov-stability is a necessary condition for an equilibrium to be Nash-Pareto.

Since the isolated equilibrium is a Nash-Pareto pair, borrowers and lenders cannot simultaneously gain from a deviation from equilibrium. Small deviations that occur when the system is at the equilibrium are not corrected by the selection mechanism, and the system starts periodically oscillating about the Nash equilibrium. One could argue, however, that the equilibrium state preserves a relevant economic interpretation. It can be shown (Schuster et alia, 1981) that the time average of the orbits correspond to the equilibrium. So, if we ignore the short-run fluctuations of the economy, we could conclude that the polymorphic Nash equilibrium can be regarded as an average state of the system.

Under deterministic dynamics the basins of attraction for each absorbing set of the system are disjoint subsets of the state space. Computer simulations of the model, when the agents follow replicator dynamics, show that

a) $(0, q_2, y_2)$ is the only state in N that is adjacent to the basin of attraction of J and

b) All those states in J with $y = 0$ are adjacent to the basin of attraction of the set N.

b) All those states in J with $q = 1$ and $y \leq y_2$ are adjacent to the basin of attraction of the set N.

In order to make more precise stability statements, we need to be precise about the mutations envisaged. Like Foster and Young (1990),Kandori et al. (1993) and Young (1993a) we focus on the effect of permanent rather than transient mutations. But we are more restrictive than they and simply assume that with some positive arbitrarily small probability each agent experiments with a strategy different from that he would otherwise play. This essentially amounts to ruling out that a strategy can die, in the fashion of Foster and Young (1990). In other words, we assume that every 'wall' is a reflecting boundary, such that, when the selection mechanism pushes the system close enough to a border, ongoing experimentation generates an opposite force which maintains the state of the system in the interior of $\Delta_3 \times \Delta_2$.

Figure 3 represents the state space: Each state is given by a 3–tuple (p,q,y); the first two elements (p,q) (the horizontal dimension) representing the state of the population of lenders. The vertical coordinate, y, corresponds to the state in the population of borrowers. The thick segment is the set N of Nash equilibria and the point e represents the isolated Nash–Pareto pair.

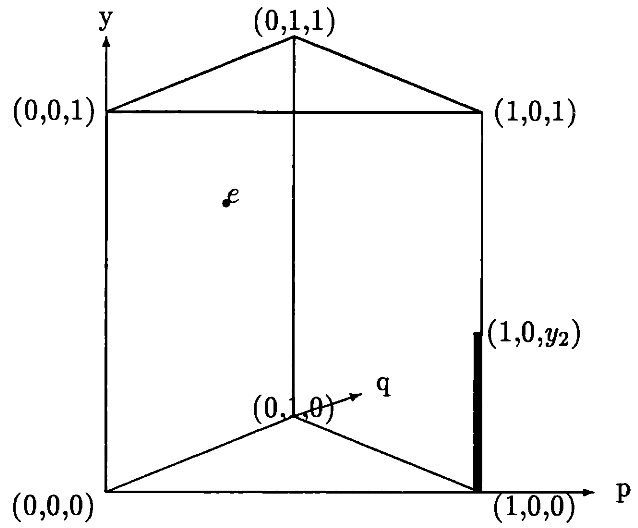

Figure 3.

Let us consider a neighbourhood of a state belonging to N. The replicator dynamics would push the system to a stationary point in N' ($p = 1, y \leq y_2$). However, given the reflecting nature of the boundaries, the system is prevented from reaching such an edge. Instead, due to ongoing mutations, a constant (arbitrary small) proportion of agents keep playing L and LM. Though proportions q and $(1 - p - q)$ do not increase in the population, because of their lower than average fitness (as far as $y < y_2$), they remain positive; in particular we assume:

$$q = n \ , \quad 1 - p - q = m$$

where n and m are the probabilities that each agents experiment with L and LM, respectively. The existence of experimentation in the population of lenders affect the payoff of borrowers. We recall that the dynamic equation of y is simply obtained as the difference between the payoff of the strategy I and the average payoff for population 2, or:

$$\dot{y} = y(\pi_I - \pi_{borrow}) = y(1 - y)(\pi(H - R) + u + F)[(m + n)q_1 - n]$$

$$\dot{y} \geq 0 \Leftrightarrow m \geq \frac{n(1 - q_1)}{q_1}$$

When the experimentation is intensive enough in monitored loans, the environment is favorable for good behaviour from borrowers and makes the strategy I increase. This ultimately drives the system away from N'. In fact, when due to the increase of y a neighbourhood of the critical state $(1, 0, y_2)$ is reached, the strategy SA starts giving a lower than average payoff (particularly, lower than LM) and credit activity starts growing with both y and (1-p-q) increasing (see Figure 4). The set of Nash equilibria N' is not robust against experimentation, when the boundaries are 'reflecting' rather than absorbing.

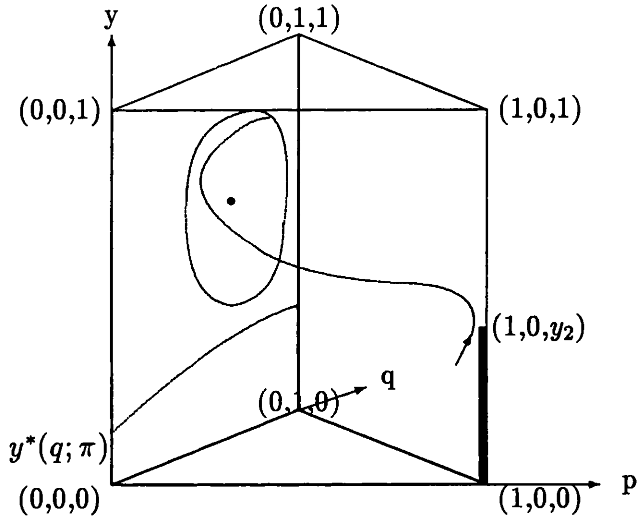

Figure 4.

Clearly, the same does not happen when experimentation is relatively intensive in unmonitored loans, against which the strategy 'cheating' exhibits the highest fitness. In this case, the system converges to the Nash equilibrium (1, 0, 0). However, when any pattern of mutation is allowed, the strategy SA, the only which is played by population $\mathcal{X}$ in the set of NE, can be 'invaded' and such equilibria are not robust to the introduction of mutants. In this case the economy leaves the trap of missing credit market.

When we allow for mutations/experimentations of the type seen before (reflecting boundaries), we can appreciate that the wall J is evolutonarily more robust than the connected set N'. Again, we assume that every agent experiments by using the non-played strategy (SA in this case) with a small probability. Let us focus first on orbits which are entirely contained in the region of the space in which $y > y^*(q)$ (see Figure 4). This implies that the outside option has a lower than average fitness and the selection mechanism never amplifies, but in fact dampers, the effects of the experimentation process.

When the selection mechanism brings the system arbitrarily close to J, experimentation keeps it at a constant distance a from the boundary. When this distance is reached the dynamic equations become:

$$\begin{aligned} \dot{p} &= 0 \\ \dot{q} &= q(\pi_L - \pi_{lend}) - \frac{qa(\pi_{lend} - \pi_{SA})}{1-a} \\ \dot{y} &= y(\pi_I - \pi_{borrow}) \end{aligned}$$

with $p = a$ (time invariant). π_i denotes, as before, the payoff of strategy i. The second term in the left hand side of $\dot{q}$ captures the process of experimentation.

In order to keep p constant at a and maintain the reflecting nature of the boundaries, it is necessary to assume that mutations into strategy SA occur at a rate $a(\pi_{borrow} - \pi_{SA})$, so as to offset the effect of the selection mechanism. It seems natural to assume that every agent, independently of the strategy which he is playing, experiments with the same positive probability. Accordingly, a proportion $q/(1-a)$ of mutants consists of people who should have played, following the replicator dynamics, the strategy L, whereas a proportion $(1-a-q)/(1-a)$ consists of people who should have played LM (the missing dimension). After substituting and rearranging, we can rewrite the previous system as follows:

$$\begin{aligned} \dot{q} &= q(1 - \frac{q}{1-a})(W - \pi R)(y - y_1) \\ \dot{y} &= y(\pi(H-R) + u + F)(1-y)(q_1(1-a) - q) \end{aligned}$$

It is easy to check that this system also admits a constant of motion, namely is represented geometrically by a set of closed loops about the equilibrium $(q_1(1-a), y_1)$.

In summary, the existence of experimentation does not destroy, contrary to the case of equilibria with missing credit market, the Ljapunov stability nature of the polymorphic equilibrium. The requirement that the system does not hit the boundary is inessential when the system converges to an orbit of a small enough amplitude. Important changes occur instead when we consider orbits with a larger amplitude. The previous proposition makes it clear, that when the proportion of cheats is relatively high ($y < y^*(q)$), J ceases to be an attractor (figure 5).

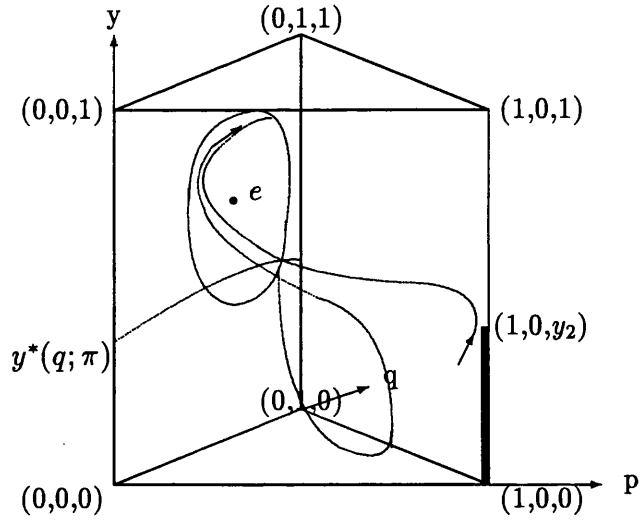

Figure 5.

When moving along orbits of high amplitude, the system periodically enters the region ($y < y^*(q)$) characterised by a high rate of cheating and bankruptcy.[2] Here, the outside option turns out to be relatively profitable and the system falls, when boundaries are reflecting, into a progressive credit crunch. Notice, however, that this process does not lead to the complete

[2]We should stress here that the model associates a positive the rate of cheating with the rate of bankruptcy, and both of them change negatively the aggregate level of output and income. This is because the expected profit is higher the higher the effort exerted.

disappearence of credit activity. It is destined to revert to a new stage of credit expansion accompanied by a reduction of the rate of bankruptcy (see figure 5). The basin of attraction of the set N, under deterministic dynamics, is adjacent to J only in the edge $(0, q, 0)$ for all q and $(0, 1, y)$ for all $y < y_2$. All other states arbitrarily close to J are destined to be mapped into long run states belonging to J itself. It is important to notice that episodes of credit crunch are observed only if the orbit's amplitude is large enough to cross the critical value y^*.

In this section we have shown that it is possible to obtain episodes of credit crunch, followed by full development of the credit activity. The less appealing feature of this result is that it relies on player's experimentation with non-played strategies. Imitative processes have the property that only existing strategies can be emulated. Once a strategy has dissappeared there is no way it appears again unless some type of experimentation is assumed. In the following section we show that it is posible to obtain the same qualitative behaviour without relying on mutations. We will assume that a tiny proportions of players behave rationally,i.e, they play Nash equilibrium. The role played by the rationals is to 'resurrect' strategies which have died or which have never before been employed. The appearence of a new stratategy,what will happen only if it is profitable, will be followed by the imitation by the myopic players.

5 Replicator dynamics with rational players

In this section we assume that the populations are made up by two types of players. A large proportion of players behave myopically, following the very simple imitative behaviour described in the previous section. A small proportion of players are rational. This last type of players always play Nash equilibria. Although very similar results could be obtained assuming that some players play 'best reply' to the current distribution of strategies, we think that our two types captures better the distintion we want to make between non rational and rational players. Best reply behaviour is not but another degree of myopia.

We assume that a certain proportion of players, σ in the population $\mathcal{X}$ (lenders) and τ in the population $\mathcal{Y}$ (borrowers), behave rationally. The rational players are characterized by their ability to compute and play optimal strategies for the one-shot game. They always played a Nash equilibrium given the distribution of myopic players.

The effect of the presence of rational players is two-fold. On one side there is a shrinkage of the absorbing sets. On the other side the rational players render the mixed equilibrium asymptotically stable under replicator dynamics (see chapter 3). We would like to emphasize that this result also holds for another specifications of the replicator dynamics, the only requirement being that the proportion of rationals is large enough.

We analyze the absorbing sets when proportions σ (in $\mathcal{X}$) and τ (in $\mathcal{Y}$) are introduced and compare them with those of the system without rationals, N and J. Let $\mathbf{x}=(p,q) \in (1-\sigma)\Delta_3$ be the vector which describes the myopic players in population $\mathcal{X}$. The vector $\mathbf{X}=(p^r,q^r) \in \sigma\Delta_3$ describes the rational agents' behaviour in population $\mathcal{X}$, it represents the rational players' optimal behaviour given the state of the system. The vectors $\mathbf{y}=y \in (1-\tau)\Delta_2$ and $\mathbf{Y}=y^r \in \tau\Delta_2$ fulfil the same roles for population $\mathcal{Y}$. A state of the system (only myopics) will be represented by a vector (p,q,y). The observed state is, therefore, $(\mathbf{x}+\mathbf{X},\mathbf{y}+\mathbf{Y})$.

We now characterize the absorbing sets of the system when the proportions of rational players in populations $\mathcal{X}$ and $\mathcal{Y}$ are σ and τ respectively.

i) $N_{\sigma,\tau} = \{(p,q,y) \in (1-\sigma)\Delta_3 \times (1-\tau)\Delta_2 \mid p = 1-\sigma, y \in [0, y_2 - \tau]\}$ and

ii) $J_{\sigma,\tau} = \{(p,q,y) \in (1-\sigma)\Delta_3 \times (1-\tau)\Delta_2 \mid p = 0, q \in [q_1 - \sigma, q_1],$
$y \in [y_1 - \tau, y_1]\}$

Assume that the state of the system is in $N_{\sigma,\tau}$. The state of the system is such that whatever the rational players from population $\mathcal{Y}$ do, the observed state will be $y + y^r < y_2$. The proportion of 'cheats' is such that the most profitable statrategy is SA. States in $N_{0,0}/N_{\sigma,\tau}$ are in the basin of attraction of the basin of attraction containing the polymorphic equilibrium. When all the rational in $\mathcal{Y}$ play I, it becomes profitable for the rationals in $\mathcal{X}$ to start granting monitored loans which incentives the good behaviour on the other side. The presence of rational players has the effect of making the system more easily invaded by mutations that are intensive enough in monitored loans.

Let us consider the set $J_{\sigma,\tau}$. Once the strategy SA has been killed off by the selection mechanism, the results for 2x2 asymmetric games (see chapter 3) apply and we know that asymptotic convergence to a Nash–Pareto pair is guaranteed under replicator dynamics. Since Ljapunov stability is preserved when the boundaries are reflecting, it should be also clear that the argument carries over to the case in which rational players are introduced. By simply applying the results of our previous chapter we can conclude that in the case of replicator dynamics the introduction of rational players turns the Ljapunov stability property of the Nash–Pareto pair into asymptotic stability. As we will see in one of the simulations reported in the last section the role of the rational players may be desestabilizing. It is possible that a 'jump', due to the joint reaction of the rational players, pushes the system into the basin of attraction of the Pareto inferior equilibrium (the dissapearance of the credit market).

The presence of a small proportion of rational players makes possible to obtain periods of recesion and credit crunch without the need of assuming any type of experimentation or mutations. Rational players will always be followed by the myopic players. The role played by the rationals will be to resurrect non played strategies whenever they are the best paying strategies. The analysis of the previous section applies here although the locus of critical values is now $y = y_2$ instead of $y^*(q) > y_2$.

6 Empirical evidence

Our dynamics admit an economic interpretation in terms of the theory of financial crises. In his analysis of the Great Depression, Bernanke (1983) observes that the fall of output and adverse development in American macroeconomy was accompanied by "exceptionally high rates of default and bankruptcy affected every class of borrower except the federal government". Bernanke argues that the financial system did not simply respond without feedback to declines in output, as it is confirmed by the fact that problems of the financial system tended to lead to output declines. In fact, according to his analysis, the initial fall in output was dramatically amplified by a subsequent period of credit contraction, in which banks switched out of loans and into more liquid investments.Bernanke observes that "Credit outstanding declined very little before October 1930, this despite a 25 percent reduction in industrial production that had occurred by that time. With the first banking crisis of November 1930, however, a long period of credit contraction was initiated. ... In October 1931 ... the net credit reduction was a record 31 percent of personal income" (p. 303). The response of the banks to the crisis and the increasing number of defaults was not to make loans to some people that they might have lent to in better times. "For example, it was reported that the extraordinary rate of default on residential mortgages forced banks and life insurance to practically stop making mortgage loans... This situation precluded many borrowers, even with good projects from getting funds ... Money (was) available in great plenty for things that are obviously safe, but not available at all for things that are in fact safe, and which under normal conditions would be entirely safe, but which are now viewed with suspicion by lenders... The idea that the low yields on treasury or blue- chip corporation liabilities during this time signaled a general state of easy money is mistaken; money was easy for a few safe borrowers but difficult for everyone else".

An apparent problem in applying our model to the analysis of developed countries issues such as the Great Crisis seems to be the assumption of absence of collateral held by borrowers. However,even in a developed country, many investors have in fact less collateral than it would be necessary to insure the lenders against the event of default. In particular, Bernanke stresses that an important effect of the crisis of the 30's was the erosion of borrowers' collateral, especially of small firms and householders. These were precisely the categories of borrowers who suffered the effect of credit rationing. It seems to be reasonable for modelling purposes to characterise the world as consisting of two types of agents: the 'ultrasafe borrowers' (large firms, government), who kept receiving funds from lenders in the form of what we call safe asset[3], and the other borrowers, who became particularly vulnerable in the 30's because of the phenomenon of the erosion of collateral. Our model seems to be applicable to this important segment of the market.

We can think of a real negative shock as the moving event, whose effect is a fall in the probability of success of each project (π in the payoff matrix) regardless of the effort taken. Let us assume that this event occurs when the system is in a long-run state described by an orbit (about the polymorphic equilibrium) of small enough amplitude to be stable against experimentation or against the presence of rational players. A fall in π has two effects: it moves the mixed Nash equilibrium to the North-West ($\delta y_1/\delta\pi < 0$, $\delta q_1/\delta\pi > 0$) and it shifts the schedule $y^*(q)$ and y_2, upwards (see figure 6). Let the orbit o

[3]We can rationalise our story more precisely by imagining that in normal conditions a certain share of total credit is granted to risky projects, which is the part of the market which we focus on in our model. When lenders choose the strategy 'safe assets' they switch resources that are normally given to risky borrowers to ultrasafe investors. This is, according to Bernanke, what happened in the 30's, as it is confirmed by the downward pressure on safe assets' interest rates.

be the long-run pre-shock equilibrium. Notice that it lies entirely above the schedule $y^*(q;\pi)$, which guarantees that Ljapunov stability is preserved under mutations. Let us suppose that the shock hits the system when it is in the downturn of the cycle, say at point a. This shifts the post- shock NE from e to e', the relevant orbit from o to o' and the locus of critical points from $y^*(q;\pi)$ to $y^{*'}(q;\pi')$. In the case represented, a becomes unstable against experimentation. As some lenders move to the outside options, others realize that this is a good choice and follow them, provoking a progressive credit crunch. Even in a case in which the productivity shock is reabsorbed soon, the credit crunch and the sluggish adjustment can make its effects long lasting before a new long run state without credit rationing emerges.

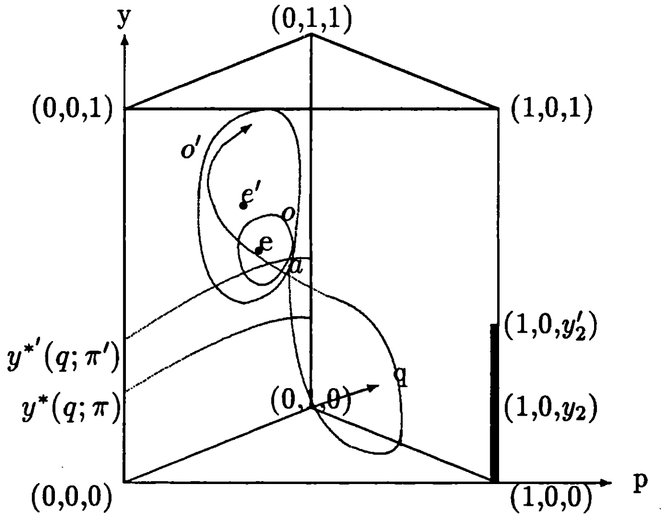

Figure 6.

The most interesting feature of the transition between Nash equilibria is that it is characterized by periods of credit crunch and recession followed by recovery and credit expansion which lead to a new long run equilibrium with the credit market fully developed.

In brief our model formalises in an evolutionary fashion Bernanke's ideas about the relation between real and financial crises as well as his interpretation of the persistence of the crisis.

7 Simulations

In this section we report the results of our simulations. The payoffs we have considered, which satisfied (C1-C4), are:

	I	*C*
SA	1 3	1 3
L	4 5	7 0
LM	4 4	0 2

Figure 7.

Figure 8 shows a case in which when all players are myopic, $\tau = \sigma = 0$. The system is initially in a state of full development of the credit market. The credit crunch starts when some people experiment with the strategy that is not played. If the experimentation occurs when the rate of cheating is high enough (under the critical relation $y(q;\pi)$ it will be followed by a period of recession after which the credit market develops again. The reduction of the loans is accompanied by a higher proportion of monitored loans which incentives the good behaviour (Invest).

Figure 8 here.

Figure 9 shows the case in which a shock moves the Nash equilibrium to the North West. The previous equilibriums is such that, under the new conditions, the best strategy is SA. The rationals start reducing the number of loans they grant. After a period of recession the new mixed equilibrium is reached.

Figure 9 here.

Figures 9.1, 9.2, 9.3, 9.4 and 9.5 show the behaviour of SA, L, LM, I and C respectively. The discrete jumps are the result of the reaction of the rationals.

At $t = 0$ there is a productivity shock which changes the fundamentals of the economy. At the pre-shock equilibrium the strategy L is no longer profitable and all the rationals 'jump' and play the strategy SA (which before was not played by anyone). Some myopic players follow the rationals and SA starts growing. The system reaches a state in which the proportion of monitored loans relative to unmonitored ones is such that the best strategy for the rational borrowers is to Invest. It arrives a moment ($t = 4$) when SA is no longer the best strategy. This will happen when the proportion of honest players is such that granting monitored loans is more profitable than the investment in the safe asset. The monitoring incentives the good behaviour. The new equilibrium is reached with more monitored loans and higher proportion of honest borrowers. The period of credit rationing emerges during the transition between equilibria.

Figure 9.1 , 9.2, 9.3, 9.4, 9.5 here.

Figure 10 reports a case where the rationals play a distabilizing role. The proportions of rational players are 'too large' and have an undesirable effect. When the rationals resurrect the startegy SA they provoke a jump into the basin of atraction of N. Initially the proportion of non-monitored loans is very high and remains high for a period due to the slow imitative process. 'Cheat' tend to grow what makes SA more and more profitable. In a certain moment the system enters a state where I is played by the rationals but the proportion of cheats is too large to revert the tendency towards the credit collapse.

Figure 10.

Figures 10.1-10.4 represent the behaviour of the safe assets, loans, monitored loans and Invest, respectively.

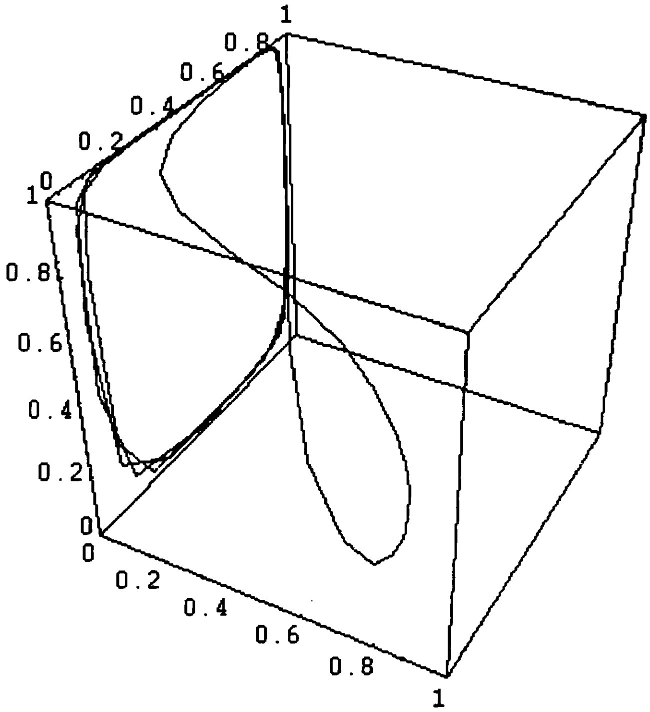

Figure 8.

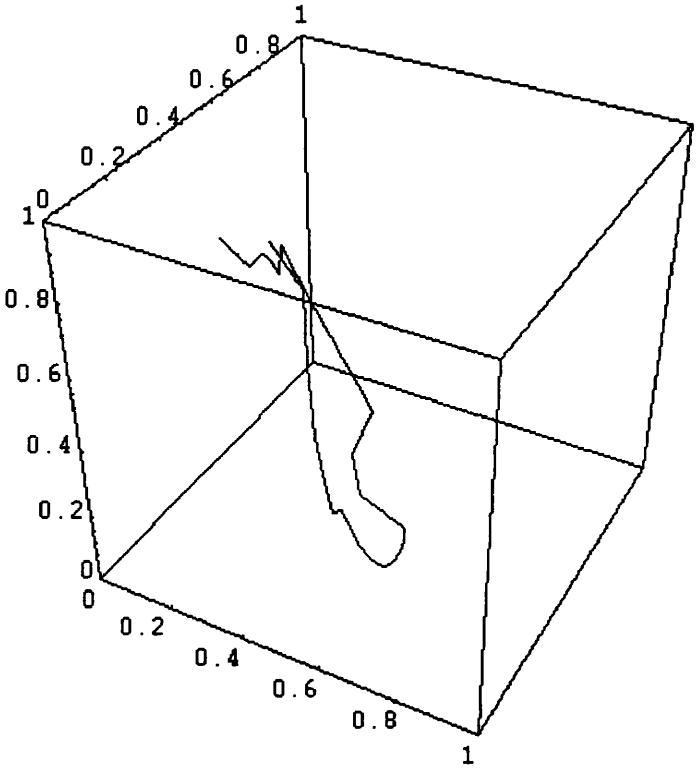

Figure 9.

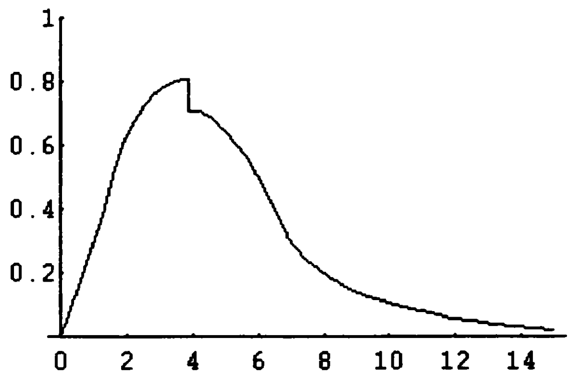

Figure 9.1. Dynamics of Safe Asset

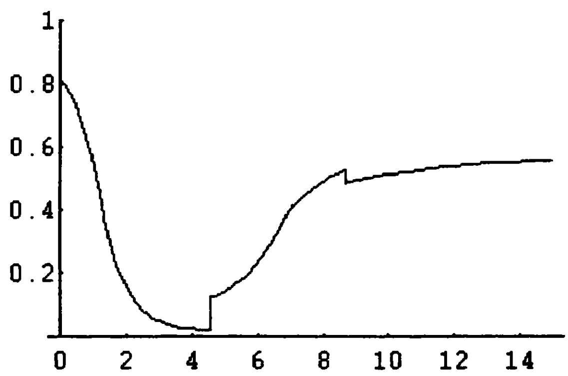

Figure 9.2. Dynamics of Loans.

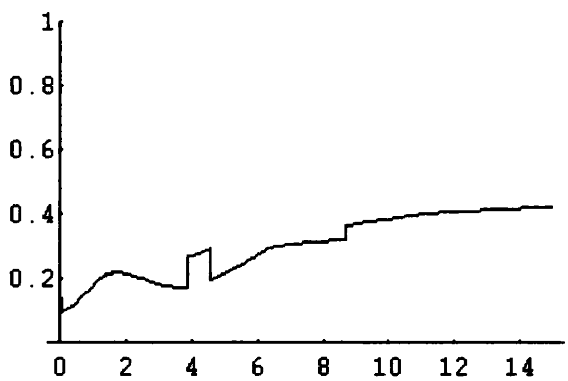

Figure 9.3. Dynamics of Monitored Loans.

Figure 9.4. Dynamics of Invest.

Figure 9.5. Dynamics of Cheat.

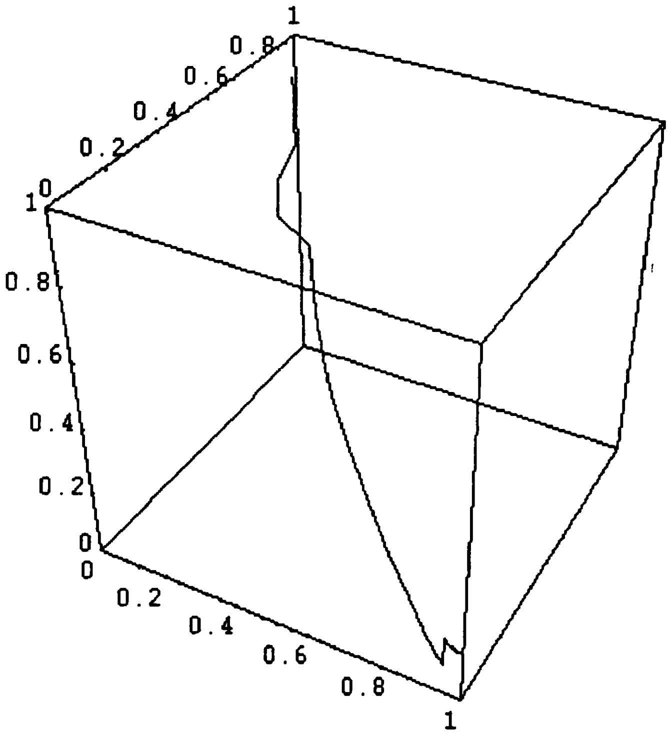

Figure 10.

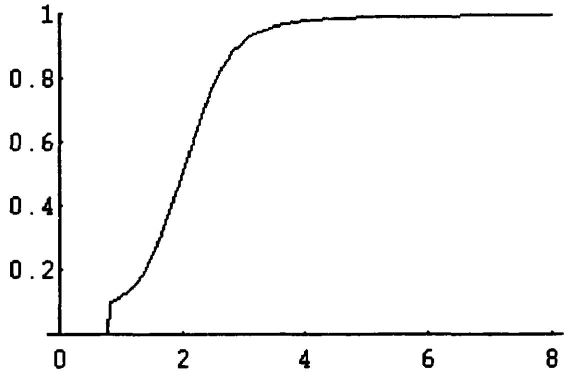

Figure 10.1. Dynamics of Safe Asset.

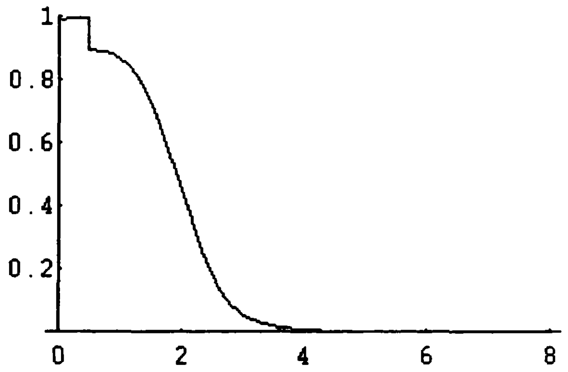

Figure 10.2. Dynamics of Loan.

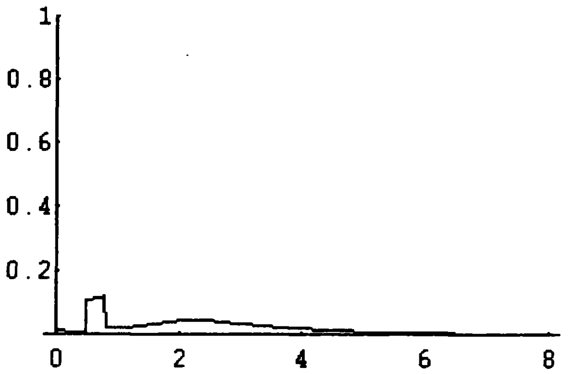

Figure 10.3. Dynamics of Monitored Loans (LM).

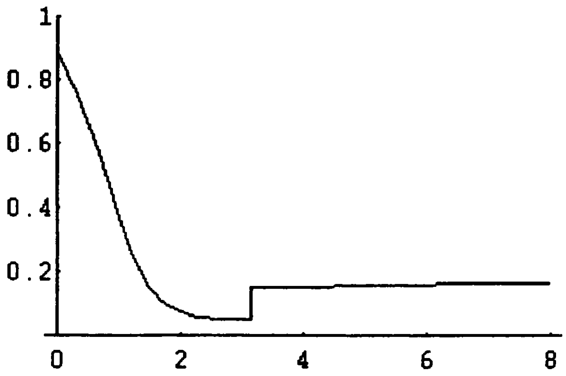

Figure 10.4. Dynamics of Invest.

Under replicator dynamics players choose their strategies based only on information about current payoffs. In the the last part of chapter 4 we study the effect on the equilibrium selection of the presence of a small proportion of more sophisticated players. We could think of the existence of some players which, due to a higher degree of sophistication or to smaller costs of gathering information, use the information regarding to the state the system and not only the information about current payoffs. Players could simply follow myopic best reply or, under stronger assumptions of rationality, some anticipatory behaviour as in Banerjee and Weibull (1991). The effect on the dynamics and on the equilibrium selection of the introduction of a small proportion of either rational players (who play a Nash equilibrium of the one shot game) or myopic best-repliers is very similar. This happens because, due to the inertia in the system, players playing best replies to the current state of the system behave 'as if' they played Nash equilibrium strategies most of the time. In both cases, the introduction of the 'sophisticated players' allows the resurrection of unplayed strategies without relying on mutations or experimentation.

In proposition 2 of chapter 3 we have shown that the presence of a small proportion of rational players renders the mixed equilibrium asymptotically stable. The convergence result rests on the collective rationality of the rational players who 'coordinate' to play the mixed equilibrium. This assumption is clearly very strong in a model which pretends to explain the emergence of Nash equilibria in a world of boundedly rational players. Yet, very similar results are obtained under the assumption that the 'sophisticated' players select best replies to the current state of the system.

The following corollary to proposition 2 proves the convergence of the system to a σ-neighbourhood of the mixed equilibrium when a σ-proportion of players use myopic best replies. All the results concerning the equilibrium selection and the dynamics with a σ-proportion of rational players hold true with a σ-proportion of players who follow best reply dynamics, provided that σ is small enough.

Corollary 1 *Let (u^*, v^*) be the completely mixed Nash Pareto Pair of Game 1 (figure 1 in chapter 3). Let σ and τ be the proportion of players who follow myopic best reply in populations $\mathcal{X}$ and $\mathcal{Y}$, respectively. All other players follow replicator dynamics (crf (1) and (2), p.125). When $\sigma > 0$ or $\sigma > 0$ there exists a $T < \infty$ such that*

$$(u_t, v_t) \in S_{\sigma,\tau}(u^*, v^*) \;\; \forall \;\; t > T$$

where

$$S_{\sigma,\tau}(u^*, v^*) \subset \Delta^2 = \{(u,v)|u \in (u^* - \sigma, u^* + \sigma) \text{ and } v \in (v^* - \tau, v^* + \tau)\}$$

Proof We consider the case $\sigma > 0$, $\tau = 0$. Thus $Y = 0$ in equations (1) and (2). Now X takes values σ and 0 because under myopic best reply all players select the same pure strategy. The behaviour of tha 'best repliers' only differs from that of the 'rationals' in $S_{\sigma,0}(u^*, v^*)$ where the latter would coordinate to play the mixed equilibrium. The argument in proposition 2 ensures that from any (u_0, v_0) the system enters $S_{\sigma,0}(u^*, v^*)$. Let us assume that the first state in $S_{\sigma,0}(u^*, v^*)$ which is reached under (1) and (2) is z. Without loss of generality let us assume that $z = (u^* + \sigma', v^*)$ where $0 \leq \sigma' \leq \sigma$. At z all best repliers, which before were playing Up, play Down and the system jumps from z into $z' = (u^* - (\sigma - \sigma'), v^*)$. At z' the best reply is Up and the system jumps again into z and so on. The extension to the case in which $\tau > 0$ is straightforward. □

The previous corollary guarantees that with a σ-proportion of best-repliers the dynamics converge to a σ-neighbourhood of the mixed equilibrium. The dynamics at any other state of the system are the same that those obtained with a σ-proportion of rational players. The fact that there is inertia in the system guarantees that myopic best repliers almost always play Nash equilibrium strategies. The only states where the behaviour of the rational players and that of the best repliers differ are those where rationals 'coordinate' to play the mixed equilibrium. The corallary applies to those states.

8 Conclusions

In this chapter we have developed an evolutionary model of a credit market. The model predicts periods of credit rationing followed by a fully active credit market. The main limitation of our model is that there is no explicit role for financial intermediation because each lenders finances, under direct lending, a whole project. In Williamson (1986) financial intermediatiors arise because with direct lending there is a duplication of monitoring costs. Each borrowers borrows from several lenders, and each of them monitors in case of default. Extensions of the model could consider more explicitely this aspect of 'delegated monitoring' and study more sophisticated micro-foundations of the simple model proposed.

However, we believe that such extensions would not alter the main idea of the paper. Credit rationing can be explained as a non-equilibrium phenomenon within an evolutionary framework with exogenous real shocks. In particular, credit crunches are observed during the process of adjustment towards the post shock equilibrium. The advantage of this approach with respect to the equilibrium analysis a la Stiglitz and Weiss (1981) is that it allows to fully characterize the dynamics of a credit cycle, as opposed to just rationalizing the existence of credit rationing.

Although the evolutionary elements of the model model are not very realistic whenever long-term relations are important, we have argued, in line with Bernanke (1983) that such long-term relations are of little importance in significant segments of the credit market where information about agents' past behaviour is either inexistent (newcomers) or unreliable. An example of such a market is the market for loans to new investors and small firms. We provide some evidence about the credit cycles observed during the Great Depression.

9 Appendix 1

Consider any state $n \in N'$. In any point contained in the neighbourhood of n the selection mechanism make SA increase and both L and LM decrease in the population. In fact, (C1) guarantees that SA is dominating strategy when $y < y_2$. That LM decreases in the population can be proved by the fact that, as $p \approx 1$ and $q \approx 0$, the payoff of LM is lower than the average pay-off if and only if:

$$y_t < \frac{r - (W - M)}{(\pi + \alpha)R - W}$$

This is always the case for points in N', since the right hand side is y_2. That L decreases in the population descends trivially from the fact that L gives the lowest pay-off amongst the strategies available to the lenders in the interval considered. □

10 Appendix 2

First, we prove that the dynamic system restricted to the absorbing set J (subset of R^2), that is invariant, describes a continuum of closed orbits surrounding the point $(0, q_1, y_1)$. This proves that J is a centre manifold of the dynamic system defined on Z. To prove this, we show that the system is 'Hamiltonian', that is there exists a function $H(q_t, y_t)$ defined in the interior of J, such that:

$$\begin{aligned} \dot{H}(q_t, y_t) &= 0 \\ H(q_t, p_t) &= const \end{aligned}$$

This function attains its maximum at the equilibrium $(0, q_1, y_1)$. The constant level sets

$$\{(q, y) \;\; \in intJ | H(q, y) = const\}$$

are closed orbits around the equilibrium (Hofbauer and Sigmund (1988)).

It is easy to verify that the function:

$$H(q,y) = [(W-M) - \pi R]\ln y + M\ln(1-y) + [(\pi+\alpha)(H-R) - e + F]\ln q$$
$$+[\pi(H-R) + u - (\pi+\alpha)(H-R) - e]\ln(1-q)$$

satisfies these conditions. Next, we show that in a subset J' including $(0, q_1, y_1)$ of the absorbing set J is attractor. We prove this by showing the existence of a local Ljapunov Function and appealing to the Ljapunov theorem that states: Let $V : Z' \rightarrow R$ (Z' is defined in the text) be continously differentiable. If for some solution $t \rightarrow \mathbf{x}(t)$, the derivative V'(t) of the map $t \rightarrow V(\mathbf{x}(t))$ satisfies the inequality $V'(t) \geq 0$, then:

$$\omega(\mathbf{x}) \cap Z' \in \{\mathbf{x} \in Z' | \dot{V}(\mathbf{x} = 0\}$$

(all orbits starting in Z' converge to J). The proof can be found in Hofbauer and Sigmund (1988), page 49. In our case, such a function exists and is of the form:

$$V(p,q,y) = q\ln q + (1-q)\ln(1-q) + k[y\ln y + (1-y)\ln(1-y)]$$

whose time derivative is:

$$\dot{V} = p[W - M - r + y((\pi+\alpha)R - W) + k(y_1 - y)[e - c - F - (\pi+\alpha)(H-R)]$$

where:

$$k \equiv \frac{W - \pi R}{\pi(H-R) + u + F}$$

It can be easily verified that: (i)$V'(t) = 0$, when (p, q, y) have the same support as the equilibrium (namely, when p = 0);. (ii)$V'(t) > 0$, when $p > 0$ and

$$y > \frac{r - (W-M) + y_1 k[c + (\pi+\alpha)(H-R) - e + F]}{(\pi+\alpha)R - W + k[c + (\pi+\alpha)(H-R) - e + F]} \equiv y^*$$

This proves that the centre manifold is attractor for all $(q,y) \in J$, such that $y > y^*$. The last thing to prove is that $y_1 > y^*$. If this is true, then 'close' to the equilibrium point the selection mechanism 'pushes' the economy towards the centre manifold. The previous inequality can be rearranged to give:

$$y_t < \frac{r - (W - M)}{(\pi + \alpha)R - W}$$

By substituting the expression of y_1, this becomes:

$$\frac{W - M - \pi R}{W - \pi R} > \frac{r - (W - M)}{(\pi + \alpha)R - W}$$

that is guaranteed by assumption (C2). □

Bibliography

Abreu, D. and Rubinstein, A. , 1988. The structure of Nash equilibrium in repeated games with finite automata. *Econometrica*, 56:1259–1282.

Aghion, P and Bolton, P . A trickle-down theory of growth and development. Mimeo, 1991.

Aumann, R. , 1985. What is game theory trying to accomplish? In Arrow, K. and Honkapohja, S. , editors, *Frontiers of Economics*, pages 28–88. Basil Blackwell, Oxford.

Banerjee, A and Weibull, J , 1991. Rationality and learning in economics. In *Evolutionary selection and rational behaviour.* Blackwell, Oxford. (forthcoming in A.Kirman,M Salmon eds).

Bardhan, P. , 1984. *Land,Labor, and Rural Poverty.* Oxford Univ.Press, Delhi.

Bernanke, B. , 1983. Nonmonetary effects of the finantial crisis in the propagation of the great depression. *American Economic Review*, 73:257–276.

Binmore, K. and Dasgupta, P. , 1987. *The Economics of Bargaining.* Basil Blackwell, Oxford.

Binmore, K. G. and Samuelson, L. , 1992. Evolutionary stability in repeated games played by finite automata. *Journal of Economic Theory*, 57:278–305.

Binmore, K. and Samuelson, L. . An economist's perspective on the evolution of norms. Lecture delivered at the International Seminar on New Institutional Economics, Wallerfagen, 1993, 1993a.

Binmore, K. and Samuelson, L. . Muddling through: Noisy equilibrium selection. University College London, mimeo, 1993b.

Binmore, K. , 1987a. Modeling rational players, I and II. *Economics and Philosophy*, 3 and 4:179–214 and 9–55.

Binmore, K. , 1987b. Nash bargaining and incomplete information. In *The Economics of Bargaining.* Basil Blackwell, Oxford.

Binmore, K. , 1987c. Perfect equilibria in bargaining models. In *The Economics of Bargaining.* Basil Blackwell, Oxford.

Binmore, K. G. . Social contract III: Evolution and utilitarianism. Technical Report CREST 89-05, University of Michigan, 1988. (forthcoming in *Constitutional Political Economy*).

Binmore, K. G. , 1991a. *Fun and Games.* D. C. Heath, Lexington, Mass.

Binmore, K. G. . Social contract IV: Evolution and convention. To appear in Harsanyi's Festschrift published by Springer-Verlag, 1991b.

Binmore, K. , 1992. Debayesing game theory. In Skyrms, B. , editor, *Studies in Logic and the Foundations of Game Theory: Proceedings of the Ninth International Congress of Logic, Methodology and the Philosophy of Science.* Kluwer, Dordrecht.

Börgers, Tilman. and Sarin, Rajiv. . Learning through reinforcement and replicator dynamics. University College London, 1993.

Boyd, R. and Richerson, P. , 1985. *Culture and the Evolutionary Process.* University of Chicago Press, Chicago.

Cabrales, Antonio and Sobel, Joel. , 1992. On the limit points of discrete selection dynamics. *Journal of Economic Theory*, 57:407–420.

Cabrales, A. . Stochastic replicator dynamics. University of California, San Diego, 1993.

Cavalli-Sforza, L.L. and Feldman, M.W. , 1981. *Cultural Transmission and Evolution.* Princeton University Press, Princeton.

Cross, J.G. , 1973. A stochastic learning model of economic behaviour. *Quaterly Journal of Economics*, 87:239–266.

Dawkins, R. , 1976. *The Selfish Gene.* Oxford University Press, Oxford.

Dekel, E and S.Scotchmer, 1992. On the evolution of optimal behaviour. *Journal of Economic Theory*, 57:392–407.

Ellison, Glenn. . Learning, local interaction and coordination. Mimeo, MIT, 1992.

Elster, J. , 1979. *Ulysses and the Syrens: Studies in Rationality and Irrationality.* Cambridge University Press, Cambridge.

Foster, D. and Young, P. , 1990. Stochastic evolutionary game dynamics. *Theoretical Population Biology*, 38:219–232.

Friedman, D. , 1992. Evolutionary games in economics. *Econometrica*, 59:637–666.

Fudenberg, D. and Harris, C. , 1992. Evolutionary dynamics with aggregate shocks. *Journal of Economic Theory*, 57:420–441.

Hofbauer, S. and Sigmund, K. , 1988. *The Theory of Evolution and Dynamical Systems.* Cambridge University Press, Cambridge.

Kandori, M. , Mailath, G. , and Rob., R. . Evolution of equilibria in the long run: A general theory and applications. Working Paper, University of Pennsylvania, 1992.

Kandori, M. , Mailath, G. , and Rob, R. , 1993. Learning, mutation and long run equilibria in games. *Econometrica*, 61 (1):29–56.

Maynard Smith, J. and Price, G. , 1972. The logic of animal conflict. *Nature*, 246:15–18.

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

Maynard Smith, J. , 1993. *The Theory of Evolution.* Canto edition, Cambridge University Press, Cambridge.

Nachbar, J. , 1990. Evolutionary selection dynamics in games: Convergence and limit properties. *International Journal of Game Theory*, 19:59–89.

Nash, J. , 1950. The bargaining problem. *Econometrica*, 18:155–162.

Nash, J. , 1953. Two-person cooperative games. *Econometrica*, 21:128–140.

Nelson, R and Winter, S.G. , 1982. *An Evolutionary Theory of Economic Change.* Belknap Press of Harvard Univ. Press, Cambridge,MA/London.

Neyman, A. , 1986. Bounded complexity justifies cooperation in the finitely repeated prisoners' dilemma. *Economic Letters*, 19:227–229.

Osborne, M. and Rubinstein, A. , 1990. *Bargaining and Markets.* Academic Press, Inc., San Diego.

Rubinstein, A. , 1982. Perfect equilibrium in a bargaining model. *Econometrica*, 50:97–109.

Rubinstein, A. , 1986. Finite automata play the repeated prisoners' dilemma. *Journal of Economic Theory*, 39:83–96.

Samuelson, L. and J.Zhang, 1992. Evolutionary stability in asymmetric games. *Journal of Economic Theory*, 57:363–391.

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

Schlag, K.H. . Why imitate, and if so, how? exploring a model of social evolution. Unpublished, University of Bonn, 1994.

Schuster, P and Sigmund, K. , 1983. Replicator dynamics. *Journal of theoretical Biology*, 100:533–538.

Selten, R. , 1980. A note on evolutionarily stable strategies in asymmetric animal conflicts. *Journal of theoretical Biology*, 84:93–101.

Selten, R. . Evolutionary stability in extensive 2-person games. Bielefeld Working Papers 121 and 122, 1983.

Selten, R. . Evolution, learning and economic behavior. Kellogg Graduate School, Northwestern University, 1989. Nancy L. Schwarz Memorial Lecture.

Stahl, D. . Evolution of smart$_n$ players. University of Texas, mimeo, 1992.

Stiglitz, J and Weiss, A , 1981. Credit rationing in markets with imperfect information. *American Economic Review*, 70:393–410.

Swinkels, J.M. , 1992. Evolutionary stability with equilibrium entrants. *Journal of Economic Theory*, 57:306–332.

Taylor, P. D. and Jonker, L.B , 1978. Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences*, 40:145–156.

Van Damme, E. , 1983. *Refinements of the Nash Equilibrium Concept.* Springer, Berlin.

Van Damme, E. , 1987. *Stability and Perfection of Nash Equilibria.* Springer, Berlin.

Wärneryd, K. , 1990. *Economic Conventions. Essays in Institutional Evolution.* Stockholm School of Economics.

Williamson, S. D , 1986. Costly monitoring, finantial intermediation, and equilibrium credit rationing. *Journal of Monetary Economics*, 18:159–179.

Young, P. , 1993a. The evolution of conventions. *Econometrica*, 61 (1):57–84.

Young, P , 1993b. An evolutionary model of bargaining. *Journal of Economic Theory*, 59:145–168.

Zeeman, E.C. , 1980. *Dynamics of the Evolution of Animal Conflicts.* Lecture Notes in Mathematics 819, Springer Verlag, Berlin.