EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

Learning, Evolution and Price Dispersion

Ed Hopkins

Thesis submitted for assessment with a view to obtaining the Degree of Doctor of the European University Institute

Florence, January 1996

EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

Learning, Evolution and Price Dispersion

Ed Hopkins

The Thesis Committee consists of:

Prof. Ken Binmore, University College London
" Alan Kirman, EUI, Supervisor
" Mark Salmon, EUI, Co-Supervisor
" John Sutton, London School of Economics
" Larry Samuelson, University of Wisconsin

Acknowledgements

I would like to thank the following people

for comments on my thesis or parts of it:
Tilman Börgers, Russell Davidson, Dan Friedman, Debora di Gioacchino, Esther Hauk, Josef Hofbauer, Lavan Mahadeva, Larry Samuelson and Rob Seymour;

Ken Binmore for a great deal of advice and support;

all the people at GREQAM, Marseille for their hospitality 1994-1995;

Mark Salmon for many helpful discussions and continued supervision;

Jessica Spataro and Jacqueline Bourgonje for so much help and assistance;

my parents for too many things to list here;

all my friends and colleagues at the EUI and in particular, Ciro D'Amore, for four happy years in Florence;

Alan Kirman, who introduced me to the literature on learning and evolution, and whose continuous advice, help and encouragement have made this thesis possible.

Chapter 1

Introduction

Habitually, economists attack a problem by calculating an equilibrium. Though this can often be gruelling for the economist, calculation of equilibrium strategies by the agents themselves is typically assumed to be costless and, even in dynamic models, timeless. This may be defensible given perfectly rational agents with complete information and unlimited intellectual capacity. With less heroic assumptions, the route to equilibrium is likely to be more protracted. Learning, in the economic context, is the explicit modelling of this process whereby equilibrium is achieved.

The link between economics and evolution may not at first seem obvious. In particular, the neoclassical paradigm of the maximising individual and biological evolution, where intentionality is only noticeable by its absence, seem diametrically opposed. Yet Friedman, in his famous essay *The Methodology of Positive Economics* (1953) marshalled a form of social evolution in defence of the neoclassical approach. He suggested that firms which failed to maximise their profits would not survive the rigours of the competitive market, and hence would be supplanted by others. However, if the predictions of neoclassical economics are to be empirically accurate, it is necessary to assume that such an evolutionary process has already taken place, or is in any case unobservable. To Friedman, modelling this process was not worthwhile.

The situation has changed dramatically over the last couple of years. There has been an explosion in research on learning and on evolution. While the analysis of imperfect information has been a mainstay of economic theory for many years, what marks out the recent literature is that the assumption of the maximising-agent is absent. Conventionally, the revelation that the behaviour of agents was less than fully optimal would be a sign of fundamental weakness in a model. However, models which possess agents whose behaviour is extremely myopic have become commonplace. Furthermore, there has been a widespread use of evolutionary models developed by biologists to explain animal behaviour.

This change has come about not just because of skepticism about the realism of the optimising model, but also because of doubts about its internal consistency. The economist's model of rationality may not be complete in that there may be situations where it fails to prescribe *any* action. More commonly, it fails in the sense that the action its prescribes is not unique. With the growth of game theoretic models in microeconomics, and rational expectations models in macroeconomics, multiple equilibria have become ever more frequent. It is a problem to which the literature on refinements, extensive though it is, has provided only a partial solution. Lastly, an economic equilibrium is, by definition, a state where, (or at least a state where agents possess the belief that), a departure from current behaviour would lead to a reduction in payoff or profit. In other words, an equilibrium is defined by a set of conjectures held by agents about what would happen out of equilibrium. Hence, without a serious treatment of what happens out of equilibrium it is not clear whether equilibrium is defined at all.

To this point, therefore, economists' growing interest in learning and evolutionary models has mainly been directed at two issues. First, are existing equilibrium concepts, (Nash, rational expectations), robust to the introduction of limited information or bounded rationality? That is, can a learning or evolutionary process ensure that agent behaviour converges to equilibrium strategies? Second, can this process pick out a particular equilibrium in situations of multiple equilibria? It might be the case that the path of out-of-equilibrium adjustment would only lead in one direction.

It is too early to judge the success of this research programme. It is also not clear what the criteria for success would be. Perhaps the final impact will only be a lengthening of the long list of refinements of Nash equilibrium. Alternatively even if more significant progress is made it will be on a purely abstract level. Once the present unsatisfactory foundations to maximisation and optimisation have been repaired and replaced, mainstream economics will continue largely unaffected. Or possibly, given the previously central position held by the rationality assumption in modern economics, the method and practice of economics will have been substantially changed.

The purpose of this introduction is to set the content of this thesis in the context of the existing research in the field. The first section deals in quite general terms with the possibility and the consequences of evolutionary argument in economics. The second examines whether such analysis can serve to discriminate between different equilibria. The third section is concerned with the particular problems that are posed to an evolutionary approach by mixed strategies. The fourth looks at applying learning models to economic problems. Along the way, the thesis itself is summarised.

1.1 Friedman and the "Classic Defence"

The argument that Milton Friedman put forward in his *The Methodology of Positive Economics* (1953) has since been dubbed the "Classic Defence". This is because it attempts to justify the neoclassical methodology, but also perhaps because it is still the reflexive response of the present generation of economists to criticism of the neoclassical orthodoxy. However, this "Defence" is more complex than is usually admitted and has often been misinterpreted. The major criticism, to this day, of the neoclassical approach is that agents neither possess the required information nor mental skills to calculate an optimum strategy. The rival school of behavioural economics is based on the assumption that there is a wide gap between optimal behaviour and the actual behaviour of economic agents. However, Friedman's principal argument was that the assumptions of economics, or indeed any science, are warranted not by their strict accuracy but the accuracy of the predictions they imply. In particular, the predictions generated by the assumptions of rational maximising agents will not be too far wrong.

Friedman employs two separate and quite different arguments in support of the empirical accuracy of the neoclassical model, and in particular, its hypothesis that economic agents act so as to maximise their returns. First, (an argument which I will label "Defence 1"), individuals will learn through experience and trial and error. The example given is the expert billiard player who can pot balls without being able to explain how he does it. This has been called "as if" rationality in that an agent behaves as if he knew the parameters and laws of motion of the system in which he is in. Friedman's second argument, "Defence 2", concerns not the learning of individuals but a form of social evolution. Competing firms may use different methods or claim different objectives. However, only those firms that, whether they intend to or not, maximise profits will survive. Competitive pressures will "select" for profit maximisation. What these two arguments have in common is that the adaptive process, whether learning or evolution, is assumed already to have taken place. Thus its actual form is immaterial because competitive pressures ensure that agents will be observed in (approximate) equilibrium.

Both arguments are vulnerable to the criticism that convergence to the optimum, or close to it, cannot simply be assumed. Indeed there are many counterexamples. First, there is evidence, both from experiments and from real economic situations, of systematic suboptimal behaviour. Second, the recent literature on learning and evolution has thrown up its share of examples of non-convergence. Furthermore, even when Friedman's argument seems valid, one can argue how close "as if" rationality is to full rationality. The ability to perform one action well because of repetition is not the same as the generalised capacity for optimisation that the conventional definition of rationality requires[1]. There is also a strong difference between rationality in the traditional economics of single agent decision problems and the rationality required in the strategic situations analysed by game theory. Indeed, Friedman's two arguments seem divided along these lines.

The transfer of evolutionary arguments to human society is more problematic than at first it might seem. Game theory has been successful in evolutionary biology[2] partially because in this context the idea of Darwinian fitness exists as a single unambiguous scale with which to measure payoffs. That is, the outcome of the game is assessed by its direct impact on the rate of reproduction of the participants. Hence, selection operates against strategies which have poor fitness because animals adopting these strategies have few offspring. Selten (1991) argues that such evolutionary forces are too slow to have a significant impact within a social timescale. Only the most extreme examples of suboptimal behaviour will result in the extinction of their exponents.

Nonetheless, it seems a reasonable assumption, and it is Friedman's argument, that also in human society, successful strategies spread and unsuccessful strategies die out, even though the mechanism is the death of ideas or institutions rather than of individuals. However, in an economic context, payoffs may only be observable as money payments to particular individuals. The exact relation between monetary payoffs and rate of spread of ideas in a population is less clear. To be fair to Friedman,

[1] For an extended critique see Winter (1986).

[2] Maynard Smith (1982) and Hofbauer and Sigmund (1988) are both full of examples.

he advanced Defence 2 for firms not for people. The identification of a firm's fitness with its financial position may not be too controversial. The problem is that if the unit of selection is the firm, it is unclear what the conclusions are for the behaviour of individuals.

What kind of rationality therefore do Friedman's arguments support? As Sen (1985) puts it, there are two different common conceptions of what constitutes economic rationality: consistency in choices made, and the pursuit of self-interest. Though these definitions are not mutually exclusive, they are quite distinct. Defence 1 seems to support consistency: agents will learn to make good choices. However, it is less clear where Defence 2 leads. It is not an argument for, as is Defence 1, "as if" rationality. While Friedman's billiard player certainly has the objective of potting the ball, economic agents may be unaware that they are maximising profits. Or as in the analysis of Elster (1978), this type of evolutionary explanation is not a part of, but an alternative to, explanation by the theory of rational choice. Decision makers are being pushed along by forces beyond their control. Choices made may be optimal with respect to economic survival but may be suboptimal with reference to the agents' own aims and preferences.

In fact, what is taken to be a defence of the rationality of economic agents is a hypothesis on something on which economics has traditionally been silent: the content of agents' preferences. For example, consider an agent with the intransitive preferences $A \succeq B \succeq C \succeq A$, and another completely rational (in the first sense) agent with transitive preferences $A \succeq B \succeq C$. As is well-known, the second agent could run a money pump against the first. However, there is nothing in the definition of rationality as consistency such that $A = \$20, B = \$30, C = \$40$ could not be true. While someone with the more normal preferences $C \succeq B \succeq A$ could presumably clear up at the expense of the previous two, this merely emphasises the point that consistency of choices is not enough for survival. One must also have suitable preferences.

This is illustrated by the results of Blume and Easley (1992). They find that a rational updating of beliefs is less important to market survival than the rate at which agents save, or, in other words, the exact form of their preferences over present and future consumption. One could argue that these results provide support for the idea of economic rationality as pursuit of self-interest. This is not necessarily the case. An agent can have a discount rate which is too high to permit long term economic survival, while at the same time maximising personal utility. In this case, subjective assessment of self-interest is at variance with either economic or evolutionary fitness.

In conclusion, there may be certain behaviour which maximises the chance of economic survival, there is, however, nothing in either of the two definitions of rationality, consistency or self-interest, that necessitates a rational agent to choose it. It is simply not clear, *a priori* what type of behaviour evolution selects for, and to what degree this overlaps with existing definitions of rational behaviour. Recent research has produced only very partial answers.

1.2 Lucas's Conjecture

Nonetheless, the evolutionary analysis of economic behaviour potentially offers a compromise between the neoclassical and behaviourist approaches. It can seek to explain where to expect optimal behaviour and why in some cases, behaviour, which is sub-optimal from the neoclassical perspective, can persist. Perhaps the first person to suggest the use of learning dynamics to select between different equilibria was Robert Lucas (1987). His careful analysis of the methodological issues presents a modification of the neoclassical position. For example, he notes that even laboratory pigeons settle down to consistent and stable choice patterns. "The economic theory of choice is thus interpreted as a description of a kind of stationary "point" of this dynamic, adaptive process." Economic theory has little to say about the learning process before this happens. Indeed, in the real world, unlike in a laboratory experiment, one cannot start observations from "t=0" when the agents are still unfamiliar with the problems they face. Secondly when dealing with macro data it would be impossible to untangle individual adaptive behaviour. This behaviour is likely to be idiosyncratic depending upon (unobservable) initial beliefs, which will evolve according to any number of other variables which agents think relevant. Thus follows, Lucas claims, the necessity of assuming real economies to be in equilibrium.

Lucas does make one concession that surely follows from certain problems that Friedman did not have to face. There have been a growing number of economic models that demonstrate that the usual assumptions of economics may not lead to any predictions, let alone accurate ones. Thus in the cases where economic theory does not give a clear prediction, that is, in the case, of multiple equilibria, Lucas suggests an examination of adaptive processes to help in the selection of a particular equilibrium. Lucas was prepared to conjecture that "reasonable" adaptive processes should only pick out "reasonable" equilibria.

The term "Lucas's Conjecture" was coined by Woodford (1990), who goes on to show that a learning process can lead to the kind of "sunspot" equilibria that Lucas conjectured would be avoided. In general, the evolutionary/learning literature gives the conjecture little support. In game theory, even early studies such as Nachbar (1990) found little overlap between existing equilibrium refinements and stability under adaptive processes. More recently Nöldeke and Samuelson (1993) do not find much support for subgame perfect equilibrium, which has been part of most of the successful empirical applications of game theory in recent years.

These results complement the extensive literature on experimental games. One famous example is the ultimatum bargaining game, in which experimental subjects consistently and persistently fail to play the subgame equilibrium (see for example, Thaler, 1988). Some have therefore claimed that game theory in general fails to predict and that in real life people do not fit the economist's model of the maximising individual. Of course, the aim of using evolutionary game theory is to show that these two propositions are not necessarily linked. Indeed, the results of Roth and Erev (1995), and those of Gale, Binmore and Samuelson (1995) seem to indicate that stability under a learning dynamic seems to predict the behaviour of experimental subjects rather better than subgame perfect equilibrium. Thus, Lucas was correct in at least one sense. It is possible to use learning dynamics to select between equilibria and hence to predict real world behaviour. However, these predictions are often at variance with those of neoclassical economics.

The main contribution of this thesis in this area is to be found in Chapter 2, *Learning, Matching and Aggregation.* There, it is shown that refinements to Nash equilibrium based on evolutionary considerations do have relevance to social and economic problems. This is because the aggregation of learning behaviour across a population of agents is qualitatively similar to evolutionary dynamics. This holds for learning rules such as those employed by Roth and Erev (1995), or for example, fictitious play. Given that results are similar for a wide range of adaptive processes, both learning and evolution, it is possible to make quite strong predictions about the stability properties of different equilibria. In Chapter 4, these results are applied to the problem of multiple equilibria in models of price dispersion. Even in this rather more complex environment, learning dynamics can be used to discriminate between different equilibria and different models, the stability of equilibrium depending on what search rule consumers use.

1.3 The Problem with Mixed Strategies

Sherlock Holmes (in his *Adventures*) takes a train to the Channel ports pursued by Moriarty. Holmes gets off at Canterbury, assuming Moriarty will take the fast train to catch him at Dover. But why did Holmes reason,

> *SH*: I know that Moriarty is following me, so I should alight at Canterbury,

and not have realised Moriarty could reason as follows,

> *M*: I know that SH knows I am following him. Therefore, he will get off at Canterbury, and therefore, so will I.

And,

> *SH*: I know that Moriarty knows that I know he is following me. Moriarty will go to Canterbury, so I can continue safely to Dover.

Forty years later Morgenstern[3] was able to see that there was a infinite regress involved in this situation. Within a few more years a formal solution was found to this problem. The Nash equilibrium for the game between Holmes and Moriarty is that each should make a random choice between Dover and Canterbury.

Mixed strategies continue to cause problems. A literal application of game-theoretic models with mixed strategy equilibria would have economic agents randomising over their possible choices. However, as sceptics like to point out, we do not often see governments, industrial managers or others, rolling dice, and the idea that they should remains bizarre[4]. Harsanyi (1973) introduced the idea of *purification*, whereby each player's payoffs are subject to random shocks the value of which

---

[3]Quoted in Brams (1993). This famous example is also discussed in Schelling (1960)

[4]Equally, for those not versed in game theory, the deterministic choice Sherlock Holmes made is probably more intuitively plausible than if Conan Doyle had had his hero toss a (suitably-weighted) coin.

is private information. These shocks to payoffs will cause each player positively to prefer one of the pure strategies in the support of the mixed strategy equilibrium. However, his opponents will have as little idea of what action he will take as if he were randomising. They, in turn, will choose a strategy according to their own private information and will be equally unpredictable.

Rubinstein (1990) find such arguments unsatisfactory. Either this private information is truly exogenous to the game, for example, what the manager ate for breakfast, or the game has been incompletely described. The assumption that decision-makers make use of completely irrelevant factors may be more unpalatable than having them spin roulette wheels. Put another way, agents may behave as though "substantively rational", their behaviour is optimal in that it is not predictable to their opponents, even though their actions are not "procedurally rational", they form their decisions through, for example, irrational hunches. As Elster (1989) argues, if agents are not explicitly randomising, they are not fully rational.

Purification clearly relies on some kind of evolutionary or adaptive argument. The dependence on the breakfast menu is a rule of thumb which has evolved to a close approximation of the optimal behaviour, randomisation. We might therefore look to the growing literature on learning in games for a model to provide such a result. Unfortunately, most of the perceived failures of learning and evolution are with mixed strategies. Shapley (1964) shows that the fictitious play process fails to converge for a class of games which possess unique mixed strategy equilibria. Crawford (1985) demonstrated that for another class of learning rules all mixed strategy equilibria are unstable. However, more recently, Fudenberg and Kreps (1993) have demonstrated convergence to a mixed strategy equilibrium, when, in the spirit of purification, payoffs are subject to random shocks.

The literature on learning and evolution has also thrown up two further ways of implementing mixed strategies without requiring any agent to randomise. First, if a player's opponent is randomly drawn from a large population then that player effectively faces a random distribution over the strategies present in that population. This type of model originates in the evolutionary biology literature. Second, the learning mechanism known as fictitious play can converge to a situation where opponents cycle between different pure strategies. The time average of strategies played is equal to a mixed strategy equilibrium.[5] Thus a mixed strategy can be interpreted as some kind of average, either across players or across time.

Hence, one can take a positivist stance. Some empirical studies have discovered economic agents behaving at least as if they were randomising. That is, the empirical frequency of pure strategies played approximated the predicted frequency. When a model has been confirmed in such a way, it seems superfluous then to ask the participants whether they were really randomising. Furthermore, what we are looking for is effective unpredictability. Agents might actually use some random-number generator to choose their course of action. Anybody familiar with computing will know that these algorithms are, in fact, deterministic. What we have is a case where the deterministic strategy is simply too complicated to predict. This is all we require.

However, it is easy to find examples where predicted and empirical frequencies do not coincide. Voter turn-out in elections is a prominent example (see for example, Elster, 1989). Even if we accept that agents may behave as though they randomise, there is the further question as to how they came to do so, and why they continue to do so. Why follow some complicated deterministic strategy when tossing a coin is both the rational strategy and relatively costless?

[5]One can also interpret this as a form of correlated equilibrium.

The answer to this may lie in the following. A large part of what lies behind mixed-strategy equilibria is the assumption of common-knowledge. That is, both participants know that the other knows to the infinite degree. It is the knowledge that any attempt to deduce an optimum course of action leads to an infinite regress which forces the rational agent to randomise. In practice, information may be incomplete. For example, it may have been the case that Holmes knew that (or indeed assigned any positive probability to the fact that) Moriarty did not know that Holmes knew Moriarty was following him. In which case, Holmes was right to get off at Canterbury without further thought.

However, common knowledge might fail not because of lack of information but because agents' capacity to process that information is bounded. Bacharach, Shin and Williams (1992) introduce the idea of *depth limits*. They draw on the empirical fact that human beings have limits on the number of iterations of the knowledge operator ("I know that you know that ...") that they can carry out. Clearly, if agents are bounded in this way, then the infinite regress, which randomisation is meant to resolve, does not in fact occur. The example of Conan Doyle/Holmes shows that even intelligent people can fail to use the capacity that they have.

This is borne out by the experimental results obtained by Rubinstein and Tversky (1993). The subjects were asked to hide an object from the other participants in one of three locations, one of which was in some way prominent. The choice of hiding-place showed a marked bias against the prominent location, thus suggesting that the participants had the same depth limit as Sherlock Holmes, that is, one. Perhaps, however, a better interpretation would be that of Stahl and Wilson (1994) who claim to find amongst experimental subjects a distribution of levels of sophistication. In such a case, in the presence of naive play which can be exploited, it may not be optimal even for rational agents to randomise.

In Chapter 2, *Learning, Matching and Aggregation*, there are results which have considerable relevance to the problems of mixed strategy equilibrium. It is demonstrated that when learning behaviour is aggregated across a large population there are many qualitative similarities with evolutionary dynamics. For example, learning can lead to a situation where the strategy profile within a population matches a mixed strategy equilibrium, even though each agent plays a pure strategy. Furthermore, while Crawford (1985,1989) found that for a class of learning rules, mixed strategy equilibria are unstable, I have shown that for this type of rule many mixed equilibria are stable in average strategy. That is, although no member of the population plays a mixed strategy with the correct frequency, the average frequency in the population does correspond to the mixed strategy equilibrium. However, equally importantly, it is shown in Chapter 2 that there are many mixed equilibria which are unstable for both learning and evolutionary dynamics. This potentially offers a testable prediction on the ability of learning theory to predict the behaviour of human agents, in that convergence is predicted for some games and divergence for others. Furthermore, Chapter 4 shows that this result has some economic significance.

Evolutionary game theory has also shown how randomisation can be avoided altogether. This occurs when a game is asymmetric in the evolutionary sense, that is, when one player is drawn from a different population from his opponent(s). An example would be "Battle of the Sexes" game, where males play only against females. But also consider the "intersection game" below. It is meant to represent the problem of two motorists meeting at an intersection. There are three Nash equilibria, {STOP, GO}, {GO, STOP}, and a mixed equilibrium where both players randomise over the two options. There is therefore a positive probability of {GO, GO} occurring. However, if one motorist is told that, for example, she is the "red" player, while the other is the "green" player, it is clear that they can use this information to coordinate their play to avoid this unpleasant outcome.

	STOP	GO
STOP	0,0	1,2
GO	2,1	-2,-2

(1.1)

Those that fail to coordinate will be displaced by those that can, in that the latter will on average earn a higher payoff. Consequently, the only evolutionary stable outcomes for the game (1.1) are where one player stops and the other goes. It is also true that the mixed equilibrium is not dynamically stable under a wide range of adjustment processes (see Chapter 3). This makes intuitive sense for this game. The problem is however, that no mixed strategy equilibria are asymmetrically stable in asymmetric games, even in the case where the mixed strategy is unique. In the case of the games analysed by Shapley (1964), there is no interpretation of instability in terms of coordination.

In Chapter 3, *Learning and Evolution in a Heterogeneous Population*, a partial solution to this problem is suggested. When there is both learning and evolution, convergence properties are quite different from when they are considered separately. We show that in fact in this context problem games such as the one considered by Shapley (1964) are stable.

1.4 Mixed Strategies in Market Games

One criticism that could be levelled at much of the literature on learning and evolution is that it has concentrated on normal form games, when many economic problems do not fit into this structure. Of course, when payoffs are non-linear and strategies are chosen from a continuum, as they typically are in market models, the analysis becomes more complicated. However, in such environments there may be solutions to be found to some of the problems discussed above. As the founder of evolutionary game theory puts it, "The context in which mixed ESS's are likely is that in which the individual is 'playing the field' " (Maynard Smith 1982, p76). What Maynard Smith calls "playing the field", that is, when each individual interacts simultaneously with all other members of the population, is recognisable to economists as a market. For example, in Cournot type competition, firms do not engage in pairwise interaction but affect each others profits through the price mechanism.

Maynard Smith's argument is that, as we have seen, in pairwise conflicts asymmetries can be used to coordinate on a pure strategy equilibrium. In the anonymity of what an economist would call market interaction such devices are not available. Only in this case are mixed strategies likely. There is a mathematical argument to support this. The result that mixed ESS's are generically unstable for asymmetric games (see Hofbauer and Sigmund, 1988) follows from the linearity of payoffs in normal form games. If payoffs are nonlinear, and one would expect this in most market situations, then there is the possibility of stability even in asymmetric situations.

Up to now, however, learning and evolutionary dynamics have not been frequently applied to games with nonlinear payoffs. In Chapter 4, *Price Dispersion: an Evolutionary Approach*, I investigate the stability of a mixed equilibrium for a game where agents can choose from a continuum of strategies and payoffs are indeed nonlinear. It is necessary to use mathematics which are somewhat more advanced than in the rest of the thesis. But the underlying issues are the same as they were in Chapter 2. Some mixed equilibria are stable under evolutionary and learning dynamics. Some are unstable. Into which category do dispersed price equilibria fall?

This question was first investigated by Diamond (1971), who examined how prices would be adjusted in an economy where consumers had imperfect information about prices. This is a reminder that the last few years is not the first time that disequilibrium models have been in fashion. Diamond found that prices would converge to the level that maximised joint profits. At this equilibrium no consumer would search. The result of the price adjustment process is not a dispersed price equilibrium because under Diamond's specification, no such equilibrium exists. Subsequent research revealed the existence of dispersed price equilibria, if for example, there was heterogeneity in the costs of buyers and sellers. However, all such work was in an equilibrium framework and the question of stability was not addressed.

A dispersed price equilibrium is a mixed strategy equilibrium in two senses. First, by definition, in equilibrium a number of prices are charged. This may be because different sellers are adopting different pure strategies, or because sellers are randomising over prices. Second, such behaviour can only be supported in equilibrium if consumers differ in their behaviour. For example, if all consumers had the same reservation price, a dispersed price equilibrium would not be possible. Because payoffs are non-linear, there is no fundamental reason why such an equilibrium, despite being mixed, should not be stable under evolutionary or learning dynamics. As it turns out, for reasons of economics, not mathematics, some such equilibria are unstable, others stable.

However, there is one other lesson to be learnt from the literature on price dispersion. Equilibrium in the model of Varian (1980), for example, cannot be supported in the manner of evolutionary game theory, by different sellers using different pure strategies. Sellers must randomise. The stability or otherwise of this form of equilibrium is very much more difficult to establish. This is something to be investigated in future research.

In conclusion, the idea of mixed strategy equilibrium, though it certainly has its problems, is not an empty one. It can be used to explain social, economic and biological phenomenon as long as one does not have the unrealistic expectation that all agents concerned should randomise with the exact equilibrium frequencies. Instead, it should be enough to find that the predictions hold on average either across time or across a population of players, or across both. However, there are many mixed strategy equilibria which are unstable under a wide range of adaptive dynamics. For these equilibria we should not expect even this weak form of convergence.

Chapter 2

Learning, Matching and Aggregation

Abstract

Fictitious play and "gradient" learning are examined in the context of a large population where agents are repeatedly randomly matched. We show that the aggregation of this learning behaviour can be qualitatively different from learning at the level of the individual. This aggregate dynamic belongs to the same class of simply defined dynamic as do several formulations of evolutionary dynamics. We obtain sufficient conditions for convergence and divergence which are valid for the whole class of dynamics. These results are therefore robust to most specifications of adaptive behaviour.

2.1 Introduction

There has been an increasing interest in using evolutionary models to explain social phenomena, in particular, the evolution of conventions. However, evolutionary models have not achieved universal acceptance. There has been some skepticism as to the degree to which evolutionary dynamics are relevant to economic situations. In an evolutionary system, nature chooses the individuals who embody superior strategies. In human society, individuals learn: they choose strategies that seem superior. There is no certainty that the dynamics generated by the two different processes are identical. But if one insists on basing social evolution on decisions taken by individual agents this presents its own problems. What does individual learning behaviour look like when aggregated across a population? Little research has been done on this issue and the results that do exist, as we shall see below, are not encouraging.

There are a number of potential responses. One adopted by Binmore and Samuelson (1994) is to devise a learning scheme which approximates the dynamics generated by evolution. Thus the results of evolutionary game theory could be recreated by learning. Another is to generalise the evolutionary dynamics by abandoning particular functional forms and looking at wide classes of dynamics which satisfy "monotonicity" or "order compatibility" (Nachbar, 1990; Friedman, 1991; Kandori et al., 1993). The hope is that even if learning behaviour is not identical to evolution, it is sufficiently similar to fall within these wider categories. However, in this paper, a different approach is taken. Rather than designing learning models to suit our purposes, we examine two existing models of learning behaviour current in the literature. This is done in the context of a large random-mixing population.

The question of aggregation of learning behaviour is of interest in its own right. As can be seen in for example, Crawford (1989) or Canning (1992), learning behaviour aggregated across a large population can be qualitatively different from behaviour at the level of the individual. Indeed, we show that aggregation can solve many of the problems encountered in existing learning models. Secondly, the resultant dynamics are not in general identical to evolutionary dynamics on a similarly defined population. They may not even satisfy monotonicity. However, they all belong to a class of dynamics which for reasons that will become apparent we will call "positive definite", and share much of their qualitative behaviour.

Fictitious play, our first learning model, was in fact introduced as a means of calculating Nash equilibrium, or in the terminology of the time in order to "solve" games (Brown, 1951; Robinson, 1951). Play was "fictitious" in that it was assumed to be a purely mental process by which agents would decide on a strategy. The fictitious play algorithm selects a pure strategy that is a best reply to the average past play of opponents. One can interpret this as though each player uses past play as a prediction of opponents' current actions. This is, of course, in the spirit of the adjustment process first suggested by Cournot in the 19th century. While it might not be clear *a priori* where such a naive form of behaviour might lead, in fact, it has been shown, for example, that the empirical frequencies of strategies played approaches a Nash equilibrium profile in zero-sum games (Robinson, 1951) and in all 2×2 games (Miyasawa, 1961).

More recently, fictitious play has again attracted interest, this time as a means of modelling learning[1]. This, however, is an interpretation that is problematic. The positive results noted above are qualified by the realisation that convergence of fictitious play is not necessarily consistent with the idea of players "learning" an equilibrium. Convergence to a pure strategy equilibrium is relatively straightforward: after a certain time, each player will keep to a single pure strategy. However, as Young (1993),

[1]Some of the many to have considered fictitious play or similar processes are Canning (1992), Fudenberg and Kreps (1993), Jordan (1993), Milgrom and Roberts (1991), Monderer and Shapley (1993), Young (1993).

Fudenberg and Kreps (1993), Jordan (1993) all note, convergence in empirical frequencies to a mixed Nash equilibrium may only entail that play passes through a deterministic cycle (of increasing length) through the strategies in its support. In one sense, players' "beliefs" converge, even if their actions do not, in that in the limit they will be indifferent between the different strategies in the support of the Nash equilibrium. However, if players' beliefs are predictions of their opponents' play, while correct on average, they are consistently incorrect for individual rounds of play. Implicit in fictitious play is also a strong degree of myopia. In choosing strategies, players take no account of the fact that opponents are also learning. Similarly, if as noted above, play converges to a cycle, players do not respond to the correlated nature of play. Finally, apart from the case of zero-sum games, there is no easy method of determining whether fictitious play converges.

There are other models of learning in games. We can identify a class of learning rules as being based on gradient-algorithms. The behaviour postulated is perhaps even more naive than under fictitious play[2], indeed, these models were first developed by psychologists and animal-behaviourists for non-strategic settings. More recently they have been applied to game-theory by Harley (1982), Crawford (1985; 1989), Börgers and Sarin (1993), Roth and Erev (1995). Unlike fictitious play-like processes agents do not play a single pure strategy which is a best-reply, agents play a mixed strategy. If a strategy is successful the probability assigned to it is increased, or in the terminology of psychologists, the "behaviour is reinforced". Thus such models are sometimes called "learning by reinforcement" or "stimulus learning". As these models' other name "gradient" suggests, behaviour is meant to climb toward higher payoffs. Adjustment is therefore slower and smoother than under fictitious play. However, the results obtained are not notably more positive. Crawford (1985) showing

[2]There are other models not considered here such as the more sophisticated Bayesian learning of Kalai and Lehrer (1993).

for example that all mixed strategy equilibria are unstable.

Aggregation can help with these problems. Fudenberg and Kreps (1993) in fact propose the idea of a random-mixing population of players as a justification for the myopia of fictitious play-like learning processes. If there is sufficient anonymity such that each player cannot identify his opponent and sufficient mixing, each player has a sequence of different opponents, then players may have little incentive to develop more sophisticated strategies. A population of players also offers a different interpretation of mixed-strategy equilibrium. The distribution of strategies in the population as a whole mimics a mixed-strategy profile. This is an equilibrium concept familiar from evolutionary game theory. This type of mixed equilibrium can be stable under either fictitious play or gradient learning.

The main contribution of this paper is to demonstrate that is possible to obtain precise results on the aggregation of learning behaviour and that furthermore, that the aggregate dynamics thereby obtained are qualitatively very similar to evolutionary dynamics. In fact, we show that the replicator dynamics, in both pure and mixed strategy forms, the aggregate dynamics generated by fictitious play, and also the aggregate dynamics generated by gradient learning, all belong to a simply-defined class of dynamics. We then show that for all of this class that regular Evolutionary Stable Strategies (ESSs) are asymptotically stable. Thus we show that refinements to Nash equilibrium based on evolutionary considerations are relevant also for learning models. Secondly, unlike existing models of learning in large populations, such as Canning (1992) and Fudenberg and Levine (1993), explicit results on the stability of particular equilibria are obtained. Perhaps most importantly we obtain results which are robust to different specifications of learning rules or evolutionary dynamics. Hence we can hope that these results have some predictive power.

2.2 Learning and Evolutionary Dynamics

We will examine learning in the context of two-player normal-form games, $G = (\{1,2\}, I, J, A, B)$. I is a set of n strategies available to player 1, J a set of m strategies for player 2. Payoffs are determined by A, a $n \times m$ matrix of payoffs, and B, which is $m \times n$. A has typical element a_{ij}, which is the payoff an agent receives when playing strategy i against an opponent playing strategy j. However, we will largely be dealing with games that are "symmetric" in the evolutionary sense, that is, games for which $A = B$.[3] Generalisations to the asymmetric case are briefly discussed in Section 7. We will often be dealing with a population of players, each playing a single pure strategy. In this case, the distribution of strategies within the population will be described by a vector $\mathbf{x} \in S_n = \{\mathbf{x} = (x_1, ..., x_n) \in \mathrm{R}^n : \Sigma x_i = 1, x_i \geq 0$ for $i = 1, ..., n\}$. As, in this paper, vectors will be treated ambiguously as either rows or columns, to avoid any further confusion, the inner product will be carefully distinguished by the symbol "$\cdot$", that is, the result of $\mathbf{x} \cdot \mathbf{x}$ is a scalar.

We follow Shapley (1964) and implement the fictitious play algorithm in the following way. A player places a weight on each of her strategies (we can think of these as beliefs as to the relative effectiveness of the different strategies) which we can represent as a vector $\mathbf{w} = (w_1, w_2, ..., w_n)$ and at any given time plays the strategy which is given the highest weight. Each player updates these weights after each round of play so that if her opponent played strategy j,

$$w_i(t+1) = w_i(t) + a_{ij} \text{ for } i = 1, ..., n. \tag{2.1}$$

Players can also be modelled as maintaining a vector of relative frequencies of opponents' past play (as in Fudenberg and Kreps, 1993; Young, 1993). They then choose

[3] And all players are drawn from the same population. For a fuller discussion of the difference between symmetric and asymmetric contests see van Damme (1991) or Hofbauer and Sigmund (1988).

strategies that maximise expected payoffs as though this vector represented the current (mixed) strategy of their opponents. The two methods are entirely equivalent. Note that the weights here are (less initial values) simply the relative frequencies multiplied by payoffs.

Up to now we have contrasted learning and evolution purely on the basis of their origins, one being a social, the other a natural process. However, they are also often modelled in contrasting fashion. Fictitious play and Cournotian dynamics both assume that agents play some kind of best response. This can involve discontinuous jumps in play. Taking as an example the following game which is variously known as "chicken", "hawk-dove" or "battle of the sexes",

$$A = B = \begin{array}{|c|c|} \hline 0 & a \\ \hline 1-a & 0 \\ \hline \end{array} \qquad 1 > a > 0\,, \tag{2.2}$$

Figure 2.1a gives the simple best-reply function for (2.2),where each agent in a large population plays a best-reply to the current distribution of strategies[4]. Here x represents the proportion of the population playing the first strategy. If x is greater than (respectively less than) a, then the whole population switches to strategy 2 (strategy 1). Hence, there is a discontinuity at the point ($x = a$) where the players are indifferent between their two strategies (there is no particular consensus in the literature about how players should behave when indifferent between two or more strategies).

In contrast, the evolutionary *replicator dynamics*, whether in continuous or discrete time, are derived on the basis that the proportional rate of growth of each strategy is equal to the difference between its payoff $(A\mathbf{x})_i$ (the ith element of the vector in parentheses) and the average payoff in the population[5] $\mathbf{x}\cdot A\mathbf{x}$. D is a positive constant.

$$\dot{x}_i = x_i[(A\mathbf{x})_i - \mathbf{x}\cdot A\mathbf{x}] \text{ or } x_i(t+1) = x_i(t)\frac{(A\mathbf{x})_i + D}{\mathbf{x}\cdot A\mathbf{x} + D} \tag{2.3}$$

[4]This is a dynamic as used by, for example Kandori, Mailath and Rob (1994). This is fictitious play with a one-period memory.

[5]In a biological context, this arives from relative reproductive success (see Hofbauer and Sigmund,

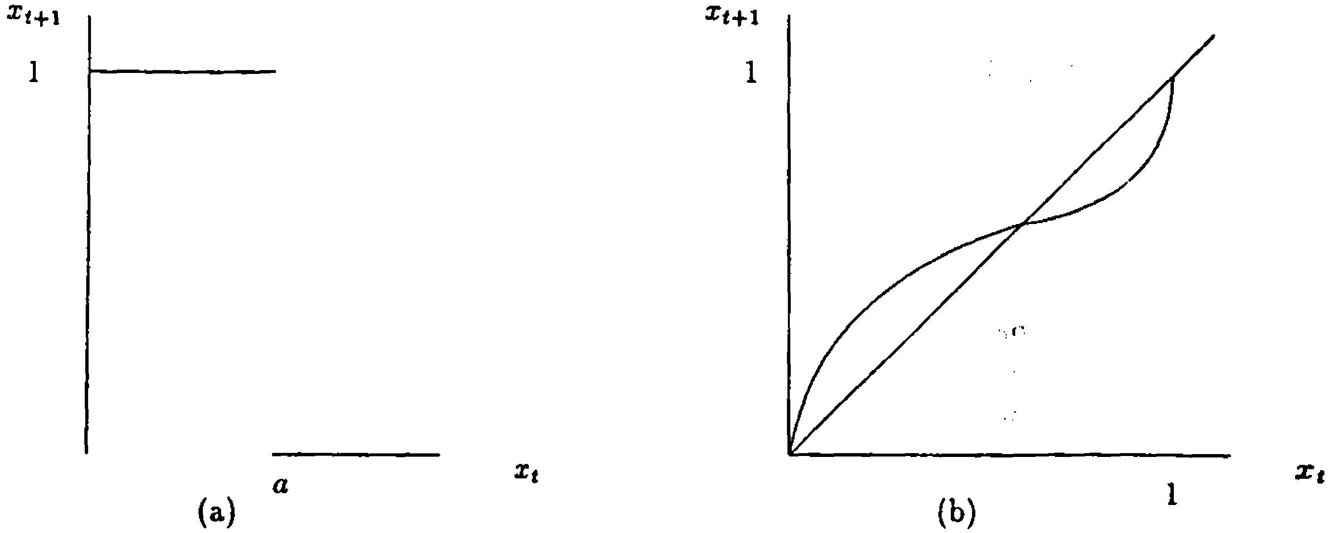

Figure 2.1: Dynamics: (a) best response (b) replicator dynamics

Clearly, both dynamics are continuous, the system moving smoothly toward the strategies earning the highest payoff. The replicator dynamic (in discrete time) for the game (2.2) is drawn in Figure 2.1b. The interior mixed equilibrium is a global attractor, the pure equilibria at $x = 0, 1$ being unstable.

Important in evolutionary theory is the idea of an Evolutionary Stable Strategy, that is, "a strategy such that, if all members of a population adopt it, then no mutant strategy could invade the population under the influence of natural selection." (Maynard Smith, 1982, p10). For a large random matching population the conditions are

Definition: An *Evolutionary Stable Strategy* (ESS) is a strategy profile $\mathbf{q}$ that satisfies the Nash equilibrium condition

$$\mathbf{q} \cdot A\mathbf{q} \geq \mathbf{x} \cdot A\mathbf{q} \tag{2.4}$$

for all $\mathbf{x} \in S_n$ and for all $\mathbf{x}$ such that equality holds in (2.4), $\mathbf{q}$ must also satisfy the

1988) but may also be an appropriate assumption in modelling learning in a human population (for example, Binmore and Samuelson, 1994).

stability condition

$$\mathbf{q} \cdot A\mathbf{x} > \mathbf{x} \cdot A\mathbf{x} \tag{2.5}$$

The first condition states that to be an ESS, a strategy must be a best-reply to itself. Were it not so, a population playing that strategy could easily be invaded by agents playing the best reply. The second condition demands that if there are a number of alternative best replies, than the ESS must be better against them than they are against themselves. Thus if a mutant strategy which was an alternative best reply were to enter the population, those agents playing it would on average have a lower payoff than those playing the ESS and therefore would not grow in number.

There is a strong connection between stability under evolutionary dynamics and the static concept of ESS.

Proposition 1 *Every ESS is an asymptotically stable equilibrium for the continuous time replicator dynamics but the converse is not true. That is, there are asymptotically stable states for the replicator dynamics which are not ESSs.*

Proof: See, for example, van Damme (1991, Theorem 9.4.8). □

Fictitious play can also converge on the mixed equilibrium of (2.2), but in a rather different manner. Setting $a = 0.5$, imagine two players both with initial weights of $(0.25, 0)$. That is, they both prefer their first strategy for the first round of play. Both consequently receive a payoff of 0. Each players observe which strategy the opponent chose. They then update the weights/beliefs according to the payoffs that they would receive against that strategy. Thus according to (2.1), weights now stand at $(0.25, 0.5)$. They now both prefer the second strategy. One can infer that player 1 believes that her opponent will continue to play her first strategy, and likewise for player 2. After the second round of play, in which again both players receive 0, the vectors stand at $(0.75, 0.5)$. It can be shown that, firstly, that the players continually miscoordinate, always receiving a payoff of 0, and that, secondly, in the limit, both play their first strategy with relative frequency 0.5, and their second with frequency 0.5. This corresponds to the mixed strategy equilibrium of (2.2). However, the players' behaviour seems to correspond only tangentially with the idea of a mixed-strategy equilibrium.

The concept of a mixed strategy equilibrium in use in evolutionary game theory seems more intuitive. It is also an average but not across time but across the differing behaviour of a large population: the aggregate strategy distribution is a mixed strategy equilibrium. One might hope that if each individual used a learning rule that like the replicator dynamics was a continuous function of payoffs, similarly well-behaved results could be obtained. However, Crawford (1985; 1989) demonstrates that in fact mixed strategy equilibria, and hence many ESSs, are not stable for a model of this kind. However, while these results are correct, they do not tell the whole story in the context of a random-mixing population. The mixed strategy of individuals will not approach the equilibrium of the two player game, nonetheless, we are able to prove convergence for the mean strategy in the population for all regular ESSs.

What we are going to show is that with a large population of players who are continually randomly matched, this type of outcome is possible even under fictitious play. This does not follow automatically from aggregation. In particular, if all players in the population have the same initial beliefs, the time path for the evolution of strategies will be the same as for fictitious play with two players[6]. Imagine in the above example, there were an entire population of players with initial weights of (0.25, 0). No matter with whom they are matched they will meet an opponent playing, strategy 1. Hence, all players will update their beliefs at the same rate, and the same cycle is reproduced. However, this is only possible given the concentration of the population

[6] A fact which Fudenberg and Kreps (1993) exploit. They do not consider the case where, within a population of players, individuals possess differing beliefs.

on a single point. If instead there is a non-degenerate distribution of weights across the population, it may be that not all the population will change strategy at once.

Imagine now that the players have initial weights or beliefs $(b, 0)$ where b is uniformly distributed on $[0, 1]$. Only those in the population with $b \leq 0.5$, that is half the population, will change strategy after the first round of play. In fact, we have arrived immediately at the population state equivalent to the mixed strategy equilibrium with half the population playing each strategy. It is easy to check that under random matching, in such a state, there is no expected change in each player's strategy. In this case, aggregation has had a smoothing effect because there was sufficient heterogeneity in the population. We will go on to make a somewhat more precise statement about convergence of fictitious play in a random matching environment. A necessary first step is to consider the modelling of random matching itself in more detail.

2.3 Matching Schemes

Any study of the recent literature on learning and evolution will reveal, firstly, that random matching within a large population of players is a common assumption, and secondly, that there are several ways of modelling such interaction. This diversity is in fact important both in terms of what it implies for theoretical results and in what cases are such results applicable. For example, there are some economic or social situations where random matching might seem a reasonable approximation of actual interaction, others where it will not. Only in some cases will agents be able to obtain information about the result of matches in which they were not involved, and so on.

Fudenberg and Kreps (1993) suggest three alternative schemes. Assuming a large population of potential players (they suggest 5000 as a reasonable number), they propose the following:

> "Story 1. At each date t, one group of players is selected to play the game...They do so and their actions are revealed to all the potential players. Those who play at date t are then returned to the pool of potential players.
>
> Story 2. At each date t there is a random matching of all the players, so that each player is assigned to a group with whom the game is played. At the end of the period, it is reported to all how the entire population played....The play of any particular player is never revealed.
>
> Story 3. At each date t there is a random matching of the players, and each group plays the game. Each player recalls at date t what happened in the previous encounters in which he was involved, without knowing anything about the identity or experiences of his current rivals."
>
> Fudenberg and Kreps (1993, p333)

It is worth drawing out the implications of these different matching schemes. Story 3 is the "classic" scheme assumed as a basis for the replicator dynamics. The population is assumed to be infinite and hence, despite random matching, the dynamics are deterministic (this has been rigorously analysed by Boylan, 1992). It is also decentralised and does not require, as do Stories 1 and 2, any public announcements of results by some auctioneer-like figure. However, there are other procedures similar to Story 2 which do not require such a mechanism. These include,

> Story 2a. In *each* round[7], the players are matched according to Story 1 or Story 3 an infinite number of times.

[7]The "round" is the time-unit of, in evolutionary models, reproduction, in learning models, decision. That is, strategy frequencies are constant within a round, even if the round contains many matches.

Story 2b. In *each* round there is a "round-robin" tournament, where each player meets each of his potential opponents exactly once.

Stories 2a and 2b have been used in the learning literature principally for reasons of tractability[8]. They ensure a deterministic result to the matching procedure even when population size is finite. The infinite number of matchings in Story 2a, by the law of large numbers, ensures that a proportion equal to the actual frequency over the whole population of opponents playing each strategy will be drawn to play. What Stories 2, 2a and 2b have in common is that all players know the exact distribution of strategies in the population when choosing their next strategy. There is little room for the diversity of beliefs one might expect in a large population.

In contrast, under Story 3, as the overall distribution of strategies is not known, it makes more sense to use past matches to estimate the current distribution. Furthermore, depending upon with which opponent they are matched, different players will receive different impressions about the frequency of strategies in the population of opponents. Under Story 3, if the population is finite, even if players use a deterministic rule to choose their strategy, such as the fictitious play algorithm, the evolution of the aggregate strategy distribution is stochastic. In this paper, however, we concentrate on the case of an infinite population, where both Story 2 and Story 3 produce deterministic results.

2.4 Population Fictitious Play

The next stage is to examine population fictitious play (PFP) where learning takes place in a large random-mixing population. We will deal both with the case where the population is large but finite, and with the case where the population is taken to be a continuum of non-atomic agents (an assumption familiar from evolutionary

[8]See for example, Kandori et al. (1993), Binmore and Samuelson (1994).

game theory). While the beliefs of a given individual can be represented by point, the beliefs of the population will be represented by a distribution over the same space. We investigate how the distribution of beliefs, and therefore how the distribution of strategies, changes over time. It will help to create some new variables. Let $p_{ij} = w_j - w_i,\ j \neq i$. Thus, $\mathbf{p}_i$ is a vector of length $n-1$. We will use this to work in $\mathbf{R}^{n-1}$ instead of $\mathbf{R}^n$. For example, if a player has to choose between two strategies, we can summarise her beliefs by the variable p_{12}. If $p_{12} < 0$ she prefers her first strategy, if $p_{12} > 0$ her second, and if $p_{12} = 0$ she is indifferent. A player's decision rule or reaction function can then be considered as a mapping from the space of beliefs to strategies, i.e. $\mathbf{R}^{n-1} \rightarrow S_n$, that is, the n-simplex. This mapping will not, in general be continuous for individual players: the fictitious play assumption limits players to pure strategies. See also Figure 2.1a.

Let F_i be the population distribution function of $\mathbf{p}_i$ over $\mathbf{R}^{n-1}$. Agents will play a strategy if it is the strategy given the highest weight in their beliefs. In other words, the beliefs of those playing strategy i must be in $E_i = \{\mathbf{p}_i \in \mathbf{R}^{n-1} : p_{ij} \leq 0, \forall\, j \neq i\}$. What if agents are indifferent between two or more strategies, that is, if their beliefs, for some j are such that $p_{ij} = 0$? One way to finesse this problem would be to assume that initial beliefs are given by irrational numbers and payoffs by rational ones or vice versa. Another method is to assume that beliefs are given by a continuous distribution on $\mathbf{R}^{n-1}$. In any of these cases then, if the proportions of the population playing each of the n strategies is given by the vector $\mathbf{x} \in S_n$, $x_i = F_i(\mathbf{0})$, where $\mathbf{0}$ is a vector of zeros of length $n-1$. For example, if all agents have the beliefs $p_{ij} < 0\ \forall\ j$ then $x_i = F_i(\mathbf{0}) = 1$.

At the basis of the deterministic model of PFP is the assumption that agents update their beliefs as if they knew $\mathbf{x} \in S_n$, the true current distribution of strategies in the population. This could be supported by Story 3 in an infinite population or by Story 2 in a finite or infinite population. We are, however, going to treat each

x_i as a continuous variable and assume that the probability of meeting an opponent playing strategy i is x_i.[9] For example, over a period of length Δt, each agent is matched within a single large population. If this matching is repeated an arbitrarily large number of times in each period (Story 2a), each agent will meet a proportion x_i of opponents playing strategy i. We assume that in a period of length Δt, players adjust their beliefs by Δt as much as they would in a period of length 1. According to (2.1), which describes the fictitious play algorithm, we have for each agent

$$\mathbf{w}(t+\Delta t) = \mathbf{w}(t) + \Delta t \ A\mathbf{x}. \tag{2.6}$$

Similarly we can derive a system of difference equations for $\mathbf{p}$, the vector of the agent's beliefs,

$$\mathbf{p}_i(t+\Delta t) = \Gamma(\mathbf{p}_i, \mathbf{x}) = \mathbf{p}_i(t) + \Delta t \ [(A\mathbf{x})_{j\neq i} - (A\mathbf{x})_i], \tag{2.7}$$

where $(A\mathbf{x})_{j\neq i}$ is a vector of length $n-1$, constructed of all the elements of $A\mathbf{x}$ except $(A\mathbf{x})_i$. We will be interested in the properties of the inverse of the function Γ with respect to $\mathbf{p}_i$ to be written $\Gamma^{-1}(\mathbf{p}_i)$. Given that $\Gamma(.)$ is a simple linear function the existence of Γ^{-1} is therefore guaranteed. In fact, we have

$$\Gamma^{-1}(\mathbf{p}_i) = \mathbf{p}_i(t) + \Delta t \ [(A\mathbf{x})_i - (A\mathbf{x})_{j\neq i}] \tag{2.8}$$

To illustrate the properties of the deterministic model with a simple example, we consider 2×2 symmetric games, that is, games where every player must choose between the same two strategies. Let $F_t(p)$ be the cumulative distribution of $p = p_{12} = -p_{21}$ on $\mathbf{R}$. This distribution of beliefs determines the distribution of strategies. As the t subscript indicates, this distribution will change endogenously over time, as the beliefs of each agent are updated according to (2.7). This is shown in Figure 2.2, (in

[9]These are both approximations if the population is finite. We treat finite populations with greater accuracy in Section 7.

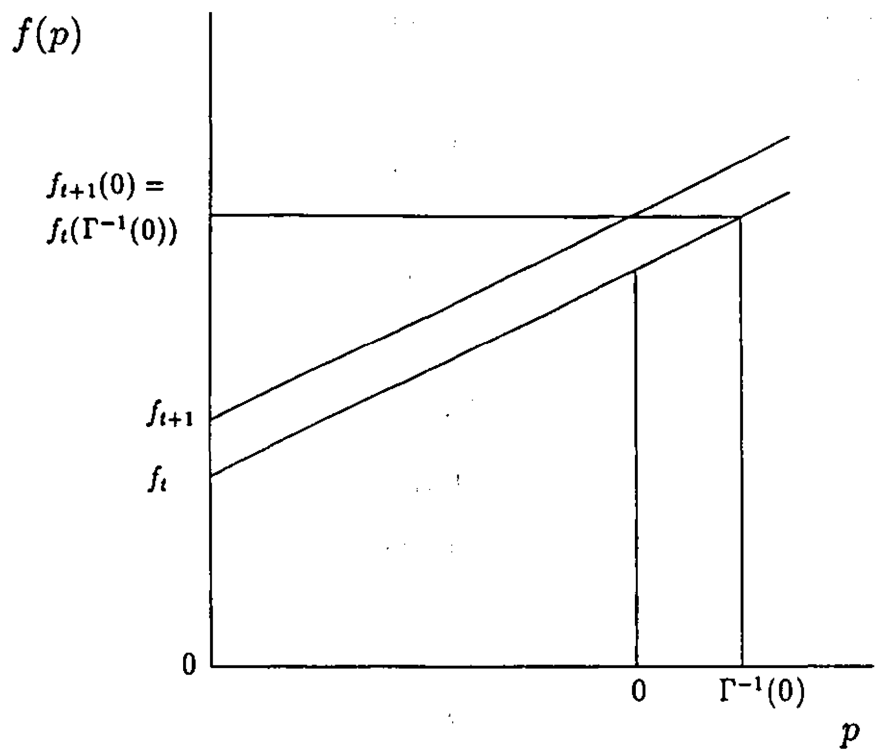

Figure 2.2: Change in the distribution of beliefs

the figure, a density function $f = dF/dp$ is assumed; its existence is not necessary to the analysis of this section). In particular,

$$\begin{array}{l} \Gamma^{-1}(p) > p : F_{t+\Delta t}(p) = F_t(p) + \int_p^{\Gamma^{-1}(p)} dF = F_t(\Gamma^{-1}(p)) \\ \Gamma^{-1}(p) < p : F_{t+\Delta t}(p) = F_t(p) - \int_{\Gamma^{-1}(p)}^{p} dF = F_t(\Gamma^{-1}(p)) \end{array} \tag{2.9}$$

Any agents possessing beliefs equal to $\Gamma^{-1}(0)$ will update their beliefs to p_0. If $\Gamma^{-1}(0) > 0$, as is the case in Figure 2.2, $F(0)$ will increase by the proportion of agents who possessed beliefs on the interval $[0, \Gamma^{-1}(0)]$. The linear nature of (2.7) implies that the whole distribution simply shifts to the left or to the right. This in turn will have an effect on the distribution of strategies. For example, an agent whose beliefs change from $p = 1$ to $p = -1$ will change from her second to her first strategy. By definition, if F is continuous at $p = 0$, that is, there is no mass of agents indifferent between strategies, $x_1 = F(0)$ and hence

$$x_1(t + \Delta t) = F_t(\Gamma^{-1}(0)) = F_t(\Delta t[(A\mathbf{x})_1 - (A\mathbf{x})_2]). \tag{2.10}$$

That is, in Figure 2.2, x_1 increases by an amount equal to the shaded area. It is not difficult to extend this analysis to games of n strategies. In a time interval of length Δt, the change in x_i is given by

$$x_i(t + \Delta t) \; = F_t(\Gamma^{-1}(\mathbf{0})) \; = F_{it}(\Delta t[(A\mathbf{x})_i - (A\mathbf{x})_{j \neq i}]) \; , \tag{2.11}$$

where F_i is the joint cumulative distribution function of $\mathbf{p}_i$ on $\mathbf{R}^{n-1}$. Clearly, if a strategy i currently has a higher expected payoff than any other strategy, then the proportion of the population playing that strategy x_i is increasing. We can state that more formally as:

Lemma 1 *If $(A\mathbf{x})_i > (A\mathbf{x})_j \; \forall \; j \neq i$ then p_{ij} is strictly decreasing at a rate bounded away from zero $\forall \; j \neq i$ and x_i is increasing. If $(A\mathbf{x})_i = (A\mathbf{x})_j \; \forall \; j$ then $p_{ij} \forall \; j \neq i$, and $x_i \; \forall \; i$, are constant.*

While the state variable of the PFP process is the distribution of agents' beliefs, our main focus of interest is the distribution of strategies. We therefore define a fixed point for the PFP process as a population strategy profile which is unchanging under the dynamic specified by (2.7), even though beliefs may continue to change. We find a one-to-one correspondence between fixed points and strategy distributions that are Nash equilibria of the game. Mixed strategies are supported by the appropriate distribution of pure strategies across the population. For the proof of the following proposition, we assume that if an agent is indifferent between two or more strategies, the choice of which of these strategies to play can be made according to any method. However, once that choice is made, no further change in strategy will be made as long as the agent remains indifferent.

Proposition 2 *A strategy profile* $\mathbf{q}$ *in the simplex* S_n *is a fixed point for the deterministic PFP dynamic if and only if it is a Nash equilibrium.*

Proof: We can start by observing that if $\mathbf{q}$ is a Nash equilibrium then from (2.4) above, if $I_0 \subseteq I$ is the set of strategies in the support of $\mathbf{q}$, then

$$\forall\, i,j \in I_0\ (A\mathbf{q})_i = (A\mathbf{q})_j \geq (A\mathbf{q})_k\ \forall\, k \notin I_0 \tag{2.12}$$

(a) **If.** If an agent plays i, she must prefer it. That is, $w_i \geq w_j\ \forall\, j$. From Lemma 1 and (2.12), no agent will change preference either between the strategies in the support of $\mathbf{q}$ or toward any other strategy.

(b) **Only if.** Let $\mathbf{q}$ now denote a rest point which is not a Nash equilibrium. Let $I_0 \subseteq I$ be the set of strategies in its support. If $\mathbf{q}$ is not a Nash equilibrium then there must be a set of strategies I_k such that $\exists\, i \in I_0\ (A\mathbf{q})_i < (A\mathbf{q})_k\ \forall\, k \in I_k$. From (2.7), for each agent playing strategy i, $w_i - w_k$ must be decreasing at a constant rate as long as the system is at $\mathbf{q}$. Within finite time, a positive measure of agents playing i must switch to a strategy in I_k. Hence the system is no longer at $\mathbf{q}$. □

The following propositions are also immediate consequents.

Proposition 3 *All pure strict Nash equilibria are asymptotically stable.*

Proof: A pure strict Nash equilibrium is a state $\mathbf{q} \in S_n$ with one strategy i in its support such that there exists an open ball B with centre $\mathbf{q}$ such that in $B \cap S_n$, $(\mathrm{Ax})_i > (\mathrm{Ax})_j \ \forall\ j \neq i$. Clearly, if the system enters B by the previous Lemma it cannot leave. While in B, for all agents, each $p_{ij}\ \forall\ j \neq i$ is decreasing at a non-vanishing rate. In finite time, all agents play i. □

Proposition 4 *All strictly dominated strategies have zero population share in finite time.*

Proof: This follows immediately from Lemma 1. □

These results are hardly surprising given that we have a population of agents that play only best replies, but they are sufficient to show convergence for many games. However, because mixed strategy equilibria are never strict, to deal with them we will need to change our approach.

2.5 Positive Definite Dynamics

We will now modify our existing model in two important ways. First, we will move from discrete to continuous time. This is not a neutral step. Our defence is that a discrete time model implies that all players are matched, and hence update their behaviour, simultaneously, a degree of coordination unlikely in a large population. Second, it is necessary to impose additional assumptions to ensure that the distribution of beliefs is continuous. For example, if there were mass points, discontinuous jumps in the value of $\mathbf{x}$ would be possible as positive measures of players switched beliefs. As we have seen the deterministic cycles of normal fictitious play are possible even in the large population model, but only with extreme restrictions on initial beliefs. Indeed, any perturbation to the distribution of beliefs will change the dynamic behaviour substantially.

Zeeman (1981) faced a similar problem in modelling mixed-strategy evolutionary dynamics. We follow the same strategy of assuming that the distributions we consider are subject to noise. For Zeeman, who was considering a biological model this was caused by mutations. Here, we can either assume that players make idiosyncratic, independently distributed mistakes in updating their beliefs, or, in the spirit of purification (see also Fudenberg and Kreps, 1993), we can imagine that each individual payoffs are subject to idiosyncratic shocks. More formally, we imagine a once-off shock of the form:

$$\mathbf{w}(t+\Delta t) = \mathbf{w}(t) + \eta, \tag{2.13}$$

where η is a vector of normally-distributed independent random variables each with zero mean and finite variance. This would rule out the possibility of mass points of agents holding exactly the same beliefs. For example, in the two strategy case, if $p = -1$ for all agents, that is, they all prefer their first strategy, with the addition of the noise, beliefs would instead be normally distributed with mean -1. We can choose the variance of η sufficiently small such that the new distribution approximates the old arbitrarily closely. Indeed, as Zeeman notes, distributions which satisfy our conditions are open dense in the set of all distributions. We state these conditions in more detail:

Assumption of Continuity: the distribution of beliefs is such that F_i is absolutely continuous with respect to $\mathbf{p}_i$. There exist continuously differentiable density functions $f_{ij} = f_{ji} = dF_i/dp_{ij}$ on $\mathbf{R}^{n-1}$ such that $f_{ij} > 0$ everywhere on $\mathbf{R}^{n-1}$.

The last inequality in turn implies that $x_i(t) > 0 \ \forall \ i, t$. However, it is possible for the system to approach the boundary of the simplex asymptotically. Consider the case where there is a single strictly dominant strategy i. In the previous section, we saw that, without noise, within a finite time only that strategy would be played. Here, the noise means that some agents will always prefer other strategies, but over time the numbers doing so will drop away to zero. The reason is that, from (2.6) and (2.13), we have $E[p_{ij}(t+\Delta t) - p_{ij}(t)] < 0 \; \forall \; j \neq i$, the strength of preference for the dominated strategies is always decreasing. The result is that, $\lim_{t\to\infty} \Pr[w_j + \eta_j > w_i + \eta_i] = 0$. Hence, $\lim_{t\to\infty} x_j = 0$ and $\lim_{t\to\infty} F(\mathbf{0}) = 1$.

We are now going to take the continuous time limit. Returning to Figure 2.2, in discrete time, all agents with beliefs in the interval $[0, \Gamma^{-1}(0)]$ changed strategy. As we will see, moving to continuous time is equivalent of taking the limit $\Gamma^{-1}(0) \to 0$. That is, the rate of change at any given point in time is going to depend on the number of agents who are, at that instant, passing from preference of one strategy to preference of another. In other words, the rate of change will be proportional to the density of agents at the point of indifference, in Figure 2.2, $f(0)$. Subtracting x_i from both sides of (2.11),

$$x_i(t+\Delta t) - x_i(t) = F_{it}(\; \Delta t[\,(A\mathbf{x})_i - (A\mathbf{x})_{j\neq i}]\;) - F_{it}(\mathbf{0}). \tag{2.14}$$

Given the presence of a random disturbance in (2.13), the reader may be surprised to see none in the above formula. The errors, however, have been subsumed in the distribution function F_i. Note that the right hand side of (2.14) can be approximated by $\Delta t \sum_{j\neq i} f_{ij}(\mathbf{0})[(A\mathbf{x})_i - (A\mathbf{x})_j]$, and that this approximation increases in accuracy as Δt and hence Γ^{-1} approach zero. Next, we divide through by Δt and take the limit $\Delta t \to 0$ to obtain

$$\dot{x}_i = \sum_{j\neq i} [(A\mathbf{x})_i - (A\mathbf{x})_j]\; f_{ij}(\mathbf{0}). \tag{2.15}$$

This also can be derived from $dF_i/dt = dF_i/d\mathbf{p}_i \cdot d\mathbf{p}_i/dt$. The last term of the chain can be obtained from (2.8) by subtracting $\mathbf{p}_i(t)$ from both sides, dividing by Δt, and taking the limit $\Delta \to 0$. It is also consistent with the theory of surface integrals which scientists and engineers use to calculate the flow of fluid (in this case, beliefs) across a surface. It will be useful to write (2.15) in matrix form,

$$\dot{\mathbf{x}} = Q(F(t))A\mathbf{x}. \tag{2.16}$$

(For the sake of simplicity, we will often suppress the extra arguments that follow Q). Clearly, (2.15) is very close to the continuous-time replicator dynamics (2.3) and the linear dynamics proposed by Friedman (1991),

$$\dot{x}_i = \frac{1}{n}\sum_{j\neq i}[(A\mathbf{x})_i - (A\mathbf{x})_j] \tag{2.17}$$

In particular, if the distribution of beliefs is symmetric, such that $f_{ij} = f_{ik}$, $\forall\, j, k$, then the continuous time PFP is identical to the linear dynamics. However, if the distribution is such that $f_{ij} = x_i x_j$, then the replicator dynamics are reproduced. In any case, without placing any restrictions on the shape of the distribution, we have the following results

Lemma 2

1. *Every element of Q is continuously differentiable in* $\mathbf{x}$,
2. $\lim_{x_i \to 0} Q_{ij} = 0\ \forall\, j$,
3. $Q\mathbf{u} = \mathbf{0}$, *where* $\mathbf{u}$ *denotes the vector* $(1, 1, ..., 1)$,
4. *Q is positive semi-definite. That is, if $A\mathbf{x}$ is not a multiple of $\mathbf{u}$, and $\mathbf{x}$ has full support, then* $A\mathbf{x} \cdot QA\mathbf{x} > 0$,
5. *Q is symmetric.*

Proof: Q has a diagonal $Q_{ii} = \sum_{j\neq i} f_{ij}$ and off-diagonal $Q_{ij} = Q_{ji} = -f_{ij}$. Satisfaction of Conditions 1 and 2 is guaranteed by the Continuity Assumption. Hence at a vertex of S_n, Q consists of zeros. Clearly $Q\mathbf{u} = \mathbf{u} \cdot Q = 0$. However, $\mathbf{x} \cdot Q\mathbf{x} = \sum_{j\neq i} f_{ij}(x_i - x_j)^2 \geq 0$. □

Geometrically, the operator Q maps the vector of payoffs $A\mathbf{x}$ from $\mathbf{R}^n$ to the subspace $\mathbf{R}^n_0 = \{\mathbf{z} \in \mathbf{R}^n : \mathbf{u} \cdot \mathbf{z} = 0\}$ (if the vector $QA\mathbf{x}$ did not add to zero then $\mathbf{x}$ would cease to add to one). It has nullspace $\mathbf{u}$. That is, at a mixed Nash equilibrium where payoffs are equal ($A\mathbf{x}$ is a multiple of $\mathbf{u}$), $\dot{\mathbf{x}} = \mathbf{0}$. For other Nash equilibria, if a strategy j is not in the support of $\mathbf{q}$, then at $\mathbf{q}$, $f_{ij} = 0$. Because Q is positive definite the angle between $A\mathbf{x}$ and $QA\mathbf{x}$ is less than 90°. This last property is what Friedman (1991) calls "weak compatibility".

Definition: Any dynamic of the form $\dot{\mathbf{x}} = QA\mathbf{x}$, where the matrix Q, satisfies the above 5 conditions, we call a *positive definite dynamic.*

We can demonstrate that evolutionary concepts are important in the context of population fictitious play. In particular, we can show that all ESSs are asymptotically stable. First we need a preliminary result,

Lemma 3 *Any ESS* $\mathbf{q}$ *is negative definite with respect to the strategies in its support. That is,* $(\mathbf{x} - \mathbf{q}) \cdot A(\mathbf{x} - \mathbf{q}) < 0$ *for all* $\mathbf{x}$ *with the same support as* $\mathbf{q}$ *(see van Damme, 1991; Theorem 9.2.7).*

The following lemma and proposition are based upon work of Hines (1980), Hofbauer and Sigmund (1988) and Zeeman (1981). However, the result obtained here generalises the above results and indeed extends beyond the continuous time PFP process to any dynamics which are symmetric positive definite transformations of the vector of payoffs $A\mathbf{x}$.

Lemma 4 *If* A *is negative definite when constrained to* $\mathbf{R}_0^n$ *(that is,* $\mathbf{z} \cdot A\mathbf{z} < 0 \; \forall \; \mathbf{z} \in \mathbf{R}_0^n$*), then* QA *is a stable matrix (i.e. all its eigenvalues have negative real parts when* QA *is constrained to* $\mathbf{R}_0^n$*).*

Proof: The eigenvalue equation is $QA\mathbf{z} = \mu\mathbf{z}$ for some $\mathbf{z} \in \mathbf{C}_0^n = \{\mathbf{z} = \mathbf{z}_1 + \mathbf{z}_2 i \in \mathbf{C}^n : \mathbf{z}_1, \mathbf{z}_2 \in \mathbf{R}_0^n\}$. We can construct a vector $\mathbf{y}$ such that $\mathbf{z} = Q\mathbf{y}$, where $\mathbf{z} \in \mathbf{C}_0^n$. By the symmetry of Q, we have $\mathbf{y}^c \cdot Q = \mathbf{z}^c$ where $\mathbf{z}^c$ is the conjugate of the complex vector $\mathbf{z}$. This gives us

$$\mathbf{y}^c QA\mathbf{z} = \mathbf{z}^c \cdot A\mathbf{z} = \mu\mathbf{y}^c \cdot \mathbf{z} = \mu\mathbf{y}^c \cdot Q\mathbf{y} \tag{2.18}$$

As Q is symmetric positive definite, $\mathbf{y}^c \cdot Q\mathbf{y}$ is real and positive. The real part of $\mathbf{z}^c \cdot A\mathbf{z}$ is negative, hence the real part of μ is negative. Since all its eigenvalues are negative or have negative real part for eigenvectors in $\mathbf{R}_0^n$, QA is a stable matrix on that space. □

A strategy profile $\mathbf{q}$ is a **regular ESS** if it is an ESS that satisfies the additional requirement that all strategies that are a best reply to $\mathbf{q}$ are in its support. We are now able to prove

Proposition 5 *All regular ESSs are asymptotically stable for any positive definite dynamic.*

Proof: Let $\mathbf{q}$ be a fully mixed ESS. Differentiating $Q(\mathbf{x})A\mathbf{x}$ with respect to $\mathbf{x}$ and evaluating at $\mathbf{q}$, we obtain $Q(\mathbf{q})A + dQ/d\mathbf{x}\, A\mathbf{q}$. At a Nash equilibrium $QA\mathbf{x} = \mathbf{0}$. It follows that for each x_i, $dQ/dx_i\ A\mathbf{q} = \mathbf{0}$. Thus the Jacobian of the system at $\mathbf{q}$ is given by $Q(\mathbf{q})A$. By Lemma 4 all its eigenvalues have real part negative.

If a regular ESS $\mathbf{q}$ is on a face $S_q \subset S_n$, that is, $q_i > 0$ if and only if $i \in I_q \subset I$, then it is also asymptotically stable under the continuous time positive definite dynamic. Because it is an ESS, A is a negative definite form on S_q, and so is QA is stable on S_q. It remains to show that the dynamic will approach S_q from the interior of S_n.

We adapt the proof of Zeeman (1981). Define $\lambda = \mathbf{u}\ \mathbf{q} \cdot A\mathbf{q} - A\mathbf{q}$. This is a vector whose ith element is zero for $i \in I_q$ and positive for $i \notin I_q$. Hence, we can define the function $\Lambda = \lambda \cdot \mathbf{x} \geq 0$, with equality on S_q, and $\dot{\Lambda} = \lambda \cdot QA\mathbf{x}$. We choose an ϵ such that for all $\mathbf{x}$ in some neighbourhood of $\mathbf{q}$, $\mathbf{x} = \mathbf{q} + \xi$ with $|\ \xi_i\ | < \epsilon$, and $|\ Q_{ij}\ | < \epsilon$ for $i \notin I_q$ by Conditions 1 and 2 of the definition of a positive definite dynamic. Then

$$\dot{x}_i = \sum_j Q_{ij}(A\mathbf{q})_j + \sum_{j,k} Q_{ij} A_{jk} \xi_k$$

Now, if $i \notin I_q$ then the first term of the above is of order ϵ, the second is of the order ϵ^2. Thus, in the neighbourhood of $\mathbf{q}$ we can approximate $\dot{\Lambda}$ by $\lambda \cdot Q(\mathbf{u}\ \mathbf{q} \cdot A\mathbf{q} - \lambda) = -\lambda \cdot Q\lambda < 0$. □

What is particularly attractive about this result is that to determine stability one no longer has to examine the potentially complicated function $Q(\mathbf{x})$. Instead, one can confine attention to the properties of A alone. For example, for the PFP dynamics it is not necessary to know the shape of the distribution of beliefs. The last two conditions on Q are the substantive ones. Positive definiteness seems a minimal condition to place upon a dynamic. Nonetheless, it becomes a sufficient condition for stability when combined with symmetry. Why this should lead to asymptotic stability for ESSs can be seen in the traditional economic terms of convexity and concavity. A "positive definite" dynamic is a gradient-climber. The negative definiteness of ESSs of course implies concavity. Any positive definite dynamic will move "uphill" toward the ESS. Condition 1 is the necessary condition for an unique solution to the differential equation (2.16). Condition 2 ensures that the dynamic remains upon the simplex. Of course, both the replicator dynamics and Friedman's linear dynamics satisfy these conditions (the latter only on the interior of the simplex).

The importance of symmetry can be illustrated by comparing positive definiteness with the Friedman's (1991) concept of order compatibility or the monotonicity of Nachbar (1990) and Samuelson and Zhang (1992). Monotonicity requires that $\dot{x}_i/x_i > \dot{x}_j/x_j$ iff $(A\mathbf{x})_i > (A\mathbf{x})_j$, and order compatibility, $\dot{x}_i > \dot{x}_j$ iff $(A\mathbf{x})_i > (A\mathbf{x})_j$. It is easy to check that if a dynamic can be written $\dot{\mathbf{x}} = Q(\mathbf{x})A\mathbf{x}$ both monotonicity and order compatibility imply the positive definiteness of Q (as Friedman points out order compatibility implies weak compatibility which is equivalent to positive definiteness). However, monotonicity and order compatibility do not imply symmetry. The existence of asymmetric order-compatible dynamics is what enables Friedman (1991) to demonstrate that ESSs may be unstable under order compatible dynamics.

Similarly, there are dynamics which are monotonic but which diverge from ESSs. Conversely, there are positive definite dynamics which are not monotone or order compatible.

In the case of only two strategies, for any such positive definite dynamic, we have

$$\dot{x}_1 = Q_{11}[(A\mathbf{x})_1 - (A\mathbf{x})_2] \tag{2.19}$$

For 2×2 games, the orbits produced by the positive definite dynamics will, after a suitable rescaling, be identical.

Proposition 6 *For 2×2 games, all positive definite dynamics generate orbits which are identical up to a change in velocity.*

Proof: Continuous dynamical systems are invariant under positive transformations, which represent a change in velocity (see for example, Hofbauer and Sigmund, 1988, p92). By positive definiteness Q_{11} is positive on the relevant state space. □

2.6 Mixed Strategy Dynamics

The replicator dynamics do not allow individuals the use of mixed strategies. As van Damme (1991) notes it would be preferable to examine mixed strategy dynamics which permit this possibility. The problem is that they are less tractable than the replicator dynamics which they generalise. In this section, we are able to show that they also fall within the class of positive definite dynamics. Furthermore, we show that the aggregation of gradient learning can be treated in a similar manner.

Zeeman (1981, Section 5) examines the properties of the mixed-strategy replicator dynamics (see also Hines, 1980). The main assumption is that there is an infinite random-mixing (Story 3) population whose individuals play mixed strategies. Thus each individual can be represented by a vector $\mathbf{y} \in S_n$. The population is summarised by a distribution F on S_n. The mean strategy in the population is given by $\mathbf{x} = \int \mathbf{y}\, dF$ and the symmetric covariance matrix $Q_m = \int(\mathbf{x}-\mathbf{y})(\mathbf{x}-\mathbf{y})\ dF$ (m is for mixed-strategy dynamic). Zeeman worked only with distributions that were *full*, that is, distributions for which Q_m has maximal rank amongst those populations having the same mean $\mathbf{x}$. As noted above, Zeeman justified this restriction by appealing to mutations. Summarising his results, we have

Lemma 5 *If* $\mathbf{x}$ *is in the interior of* S_n *then* $\mathbf{z} \cdot Q_m \mathbf{z} > 0$ *for any* $\mathbf{z}$ *which is not a multiple of* $\mathbf{u}$. *(Zeeman 1981, p265).*

Assuming as for the pure strategy replicator dynamic that the proportional growth rate of a strategy is equal to the difference between its and the average payoff gives

$$\dot{f}(\mathbf{y}) = f(\mathbf{y})[\mathbf{y} \cdot A\mathbf{x} - \mathbf{x} \cdot A\mathbf{x}]$$

and hence

Lemma 6 *The dynamic for the mean mixed strategy satisfies* $\dot{\mathbf{x}} = Q_m A\mathbf{x}$. *(Zeeman 1981, p266).*

We can find similar results for the type of learning dynamics considered by Harley (1982), Börgers and Sarin (1993), Crawford (1989) and Roth and Erev (1995). This may seem strange in that, first, Börgers and Sarin rightly point out this learning process when aggregated across a population of players is not identical to the replicator dynamics for either pure or mixed strategies, and that, second, Crawford proves that in such a large population, under such dynamics the mixed-strategy equilibrium of a simple game like (2.2) is unstable. However, Crawford's definition of a mixed-strategy equilibrium is the state where every agent plays the equilibrium mixed-strategy, that is, in game (2.2), they all play their first strategy with probability a. However, I would argue that in a random-mixing population this definition is over-strict. It is possible to have a state where the average strategy in the population, and hence, the expected strategy of an opponent, is equal to the mixed strategy equilibrium, although no agent plays the exact mixed strategy equilibrium profile. For example, the ith member of the population could play her first strategy with probability $a + \epsilon_i$ with $\sum \epsilon_i = 0$.

We assume, as for fictitious play, that each player has a vector $\mathbf{w}$, each element representing the "confidence" placed on each strategy. However, rather than choosing the strategy with the highest weight, each player plays strategy i with probability

$$y_i = \frac{w_i}{\sum_{i=1}^n w_i} = \frac{w_i}{W}.$$

Thus, here, in a similar way to the model of Zeeman, we can represent each individual as a point $\mathbf{y} \in S_n$, distributed according to a function F. However, here we have to take account of the magnitude of W, the sum of an agent's weights. We assume that they are distributed on $\mathbf{R}$ according to a function G, and let H be the joint distribution function (incorporating F and G) on $S_n \times \mathbf{R}$. And again, in a large random-mixing population, the probability of meeting an opponent playing strategy i will be x_i, where again we define the population mean as $\mathbf{x} = \int \mathbf{y}\ dF$. However, rather than strategy distributions being changed according to an evolutionary process, each individual learns by adjusting the probability that she plays each strategy in relation to the payoff that the strategy earns. If a strategy is chosen, and playing that strategy yields a positive payoff, then the probability of playing that strategy is "reinforced" by the payoff earned. In particular, if an individual plays strategy i against an opponent playing strategy j, then the ith element of $\mathbf{w}$ is increased by the resulting payoff, again scaled by the length of the period Δt,

$$w_i(t + \Delta t) = w_i(t) + \Delta t\, a_{ij}.$$

However, all other elements of $\mathbf{w}$ remain unchanged. This is the "Basic Model" of Roth and Erev (1995), who give a number of reasons why this may be a reasonable approximation of human learning. Thus the expected change is given by,

$$E\left[w_i(t+\Delta t)\right] = y_i\left(w_i(t) + \Delta t\,(A\mathbf{x})_i\right) + (1-y_i)w_i(t). \tag{2.20}$$

There are three important differences between this learning rule and fictitious play. First, it is stochastic, not deterministic. Second, while under fictitious play, agents have a limited capacity for assessing what they might have received if they had used some other strategy, here agents only consider what actions they actually play and what payoffs they actually receive (this type of learning model was developed to analyse animal behaviour). Third, for the probabilities to remain well defined, we must require all payoffs to be non-negative[10], and that all agents start with all elements of their vector $\mathbf{w}$ strictly positive. From (2.20), we can obtain

$$E\left[y_i(t+\Delta t) - y_i(t)\right] = \frac{\Delta t\, y_i\left((A\mathbf{x})_i - \mathbf{y}\cdot A\mathbf{x}\right)}{W + \Delta t\, \mathbf{y}\cdot A\mathbf{x}}. \tag{2.21}$$

This is a special case[11] of the RPS rule of Harley(1982). Crawford (1989) characterises individual behaviour in a large population of players by the deterministic continuous time equation,

$$\dot{y}_i = y_i[(A\mathbf{x})_i - \mathbf{y}\cdot A\mathbf{x}]. \tag{2.22}$$

Börgers and Sarin (1993) show that by using a slightly different specification of the updating rule one can obtain a continuous time limit similar to Crawford's equation (2.22)[12]. The advantage of the approach of Börgers and Sarin and Crawford is that learning behaviour is easier to characterise, but only at the cost of additional assumptions.

[10] Either we consider only games with positive payoffs, or we add a positive constant to all payoffs sufficiently large to make them positive. Clearly such a transformation would make no difference to a game's strategic properties, though, in a dynamic context it can change the rate of adjustment. See the discussion of discrete time processes in the next section.

[11] The equation (2.21) can be obtained by setting what Harley calls the "memory factor" to 1.

[12] It would be the same if Börgers and Sarin considered as did Crawford a single random-mixing population.

In any case, the next step is to derive an expression for the evolution of the population mean. If we think of the change made by each agent as a draw from the distribution that describes the population, $x_i(t+\Delta t) - x_i(t)$ is then the sample mean. Hence, the variance of the change in x_i is decreasing in the number of agents. Thus, if the population is infinite, then the evolution of the population mean will be deterministic (the case of a finite population will be dealt with in the next section). We have

$$\begin{aligned} x_i(t+\Delta t) - x_i(t) &= \int E[y_i(t+\Delta t) - y_i(t)]\, dH \\ &= \int \Delta t\, y_i[(A\mathbf{x})_i - \mathbf{y}\cdot A\mathbf{x}]/(W + \Delta t\, \mathbf{y}\cdot A\mathbf{x})\, dH \\ &= \int \Delta t\, y_i[\mathbf{e}_i - \mathbf{y}]/(W + \Delta t\, \mathbf{y}\cdot A\mathbf{x})\, dH \cdot A\mathbf{x}. \end{aligned}$$

where $\mathbf{e}_i$ is a vector of zeros except for a 1 in the ith position and $W + \Delta t\, \mathbf{y}\cdot A\mathbf{x} > 0$ (by the assumption of non-negative payoffs). We divide through by Δt and take the continuous time limit. This in turn gives us,

$$\dot{\mathbf{x}} = Q_g A\mathbf{x} \tag{2.23}$$

where the g-subscript is for gradient learning. The diagonal of Q_g has the form $\int y_i(1-y_i)/W \quad dH$, the off-diagonal $-\int y_i y_j/W \; dH$. Hence Q_g is symmetric and $Q_g\mathbf{u} = \mathbf{0}$. Clearly $\mathbf{z}\cdot Q_g\mathbf{z} = \sum_{i\neq j}\int y_i y_j/W \; dH \; (z_i - z_j)^2 \geq 0$. Consequently Q_g is positive semi-definite. To obtain the model of either Börgers and Sarin (1993) or Crawford(1989) it simply necessary to set $W = 1$ for all agents. Clearly this would not change the conclusion that although $Q_g \neq Q_m$,

Proposition 7 *The mean of the mixed strategy replicator dynamic and the mean of the gradient learning process are positive definite dynamics.*

This, together with Proposition 5, extends the existing results on gradient dynamics.

An Example. Take the game (2.2), assume $a = .5$, that $F(y_1) = y_1^2$, and hence $x_1 = 2/3$. Under the mixed strategy replicator dynamics, we have $\dot{f}(y_1) =$

$2y_1[1/9 - y_1/3]$. That is, those agents playing the first strategy with probability less than one third, and hence far from the equilibrium strategy, are increasing in number. For the gradient dynamics, we have $\dot{y}_1 = -y_1(1 - y_1)/(6W)$. In words, all agents are decreasing the weight they place on their first strategy. This also demonstrates the difference between the two dynamics. The evolutionary dynamic replaces badly-performing agents by better performers[13], under the gradient dynamics, all agents respond to the situation by changing strategy. As Crawford (1989) discovered, the state where all agents have $y_1 = 0.5$ is not going to be stable. In this example, the agents who are currently playing the "equilibrium" mixed strategy ($y_1 = 0.5$) are respectively dying off and moving away from it. However, for both dynamics we have $\dot{x}_1 = Q_{11}[1/2 - x_1]$, and hence the mean strategy clearly approaches the equilibrium[14].

2.7 Games without ESSs

Since the concept of an ESS is a strong refinement on Nash equilibrium and consequently there are many games which do not possess any equilibrium which satisfies its conditions, one might wonder how positive definite dynamics perform in these cases. For any constant-sum game for any $\mathbf{x} \in S_n$, $\mathbf{x} \cdot A\mathbf{x} = v$, where $v$is the value of the game. It follows, if the game has a fully mixed equilibrium $\mathbf{q}$, that $(\mathbf{x} - \mathbf{q}) \cdot A(\mathbf{x} - \mathbf{q}) = 0$. From Proposition 5 and in particular (2.18) we have that,

Corollory 1 *The eigenvalues of the linearisation of any positive definite dynamic at a fully mixed Nash equilibrium of a zero-sum game have zero real part.*

[13]Though perhaps this type of dynamic could be reproduced in a population that learns by imitation.

[14]Harley (1982, p624) reproduces two graphs of the results he obtained from simulations of a similar game using his learning model. Two things are apparent: the population mean approaches the mixed strategy equilibrium, the strategy of individual players (typically) does not.

This result unfortunately is of the "anything can happen" type. For the linear dynamics (2.17), because they are linear, the Corollary implies that such an equilibrium must be a neutrally stable centre (it is easy to check that $V = \frac{1}{2}(\mathbf{x}-\mathbf{q})\cdot(\mathbf{x}-\mathbf{q})$is a constant of motion in this case). For non-linear dynamics the fact that their linearisations have zero eigenvalues may hide asymptotic stability or instability.

Secondly, there are games which possess equilibria which are positive definite. It is an obvious corollary of Proposition 5 that positive definite dynamics diverge from such equilibria. This can prove useful in terms of equilibrium selection. Unstable positive definite equilibria can be rejected in favour of stable ESSs. This works well in games with both ESSs and positive definite equilibria.

$$A = \begin{array}{|c|c|c|} \hline 0 & a_1 & -b_1 \\ \hline -b_2 & 0 & a_2 \\ \hline a_3 & -b_3 & 0 \\ \hline \end{array} \quad a_i, b_i > 0,\ i = 1,2,3 \tag{2.24}$$

But the game (2.24) has an unique equilibrium which, for example, for $a_i = 1, b_i = 3,\ i = 1,2,3$ is positive definite. Hence, no positive definite dynamic can converge. This might seem problematic, but in fact it offers a strong empirical prediction. For rational players under the full-information assumptions of conventional game theory, for a game with an unique Nash equilibrium it should not matter whether it is positive or negative definite. However, we can conjecture that in a random-matching environment under experimental conditions, the strategy frequencies of human subjects would converge if, for example, $a_i = 3$ and $b_i = 1$ but not if $a_i = 1$ and $b_i = 3$. This conjecture we can make with a degree of confidence because so many different specifications of adaptive learning are consistent with positive definite dynamics. Such divergence is not necessarily "irrational" or "myopic". Indeed, if $a_i = 1, b_i = 3,\ i = 1,2,3$ average payoffs are at a minimum at the mixed equilibrium. Divergence increases average payoffs.

The robustness of these results, however, does depend on the property of positive or negative definiteness. For equilibria which are neither positive neither negative definite, it is possible for stability properties to vary according to the exact specification of the dynamics. Such equilibria can be attractors or repellors. Using (2.24) again as an example, the pure strategy replicator dynamics converge iff $a_1a_2a_3 > b_1b_2b_3$, the linear dynamics iff $a_1 + a_2 + a_3 > b_1 + b_2 + b_3$, while simulation suggests that the PFP dynamics will converge to any equilibrium of the game which is not positive definite.

We conclude this section with discussion of the extension of the above results to discrete time and to asymmetric games. Consider a positive definite dynamic such that

$$\mathbf{x}(t+1) = \mathbf{x}(t) + QA\mathbf{x}, \tag{2.25}$$

where Q again satisfies the five conditions outlined above. In this case, pure strategies which are regular ESSs will be asymptotically stable, the second part of the proof of Proposition 5 applying equally well in discrete time. The problem is, as always, with mixed strategies. From (2.25), the linearisation at a fully mixed fixed point $\mathbf{q}$ will be

$$I + Q(\mathbf{q})A. \tag{2.26}$$

As we have shown, the eigenvalues of QA are negative. If however, they are too "large", the absolute values of the eigenvalues of $I + QA$ will be greater than one. So it is possible for a discrete time positive definite process to diverge from a mixed ESS. This is going to depend on the magnitude of the change in strategy distribution made each period. In the case of a pure strategy equilibrium, it must be true that $\| \mathbf{x} - \mathbf{q} \| > \| QA\mathbf{x} \|$ otherwise the dynamic would jump over the fixed point and out of the simplex. In contrast, unless the rate of change is sufficiently slow, it is possible to shoot right past a mixed-strategy equilibrium. Note that, for example, for the discrete time replicator dynamics given in (2.3), the rate of adjustment is decreasing in the constant D. Hence, stability of ESSs can be assured if D is sufficiently large. In the case of gradient learning, the rate of change is decreasing over time as the size of individuals' weights (W in the notation of the last section) increases. Furthermore, in the case of positive definite equilibria, where QA has positive eigenvalues, then all the eigenvalues of the linearisation (2.26) are clearly greater than one and the equilibrium will most certainly be unstable.

In the case of asymmetric games, it is well known that no mixed strategy equilibria are ESSs. Furthermore, it is also well known that mixed strategy equilibria are either saddles or centers for the replicator dynamics (Hofbauer and Sigmund, 1988). It is easy to show that this result generalises to all positive definite dynamics. In particular, let x give the strategy frequencies in the first population and y in the second, and $\dot{\mathrm{x}} = QA\mathrm{y}$, $\dot{\mathrm{y}} = PB\mathrm{x}$, where Q and P are positive definite matrices satisfying the conditions outlined above. Then the argument outlined in Hofbauer and Sigmund (p142-3) goes through unchanged.

2.8 Conclusion

There has been some debate as to whether the replicator dynamics, in spite of their biological origins, can serve as a learning dynamic for human populations. The results obtained here on one level give some support to the skeptics. The aggregation of learning behaviour across a large population is not in general identical to the replicator dynamics, in either their pure or mixed strategy formulation. However, it is clear that all these dynamics, whether of learning or evolution, share many of the same properties.

This is valuable in that, as the literature on learning and evolution has been growing at a significant rate over the past few years, there has been a proliferation of different models and consequently different results. The hope here is that we have obtained a result that is reasonably robust: ESSs are asymptotically stable for many apparently different adaptive processes when these processes are aggregated across a large random-mixing population. An ESS is quite a strong refinement on Nash equilibrium. Furthermore, it has been discredited in the eyes of some because it does not correspond exactly to asymptotic stability under the pure strategy replicator dynamics (Proposition 1). However, these are not the only dynamics of interest, and for results on stability that are robust to different specifications, the concept of ESS is the one that is relevant. In extending existing results on fictitious play, gradient learning and mixed-strategy replicator dynamics, it has been the negative definiteness of ESSs which has been essential.

Researchers have begun to test the predictions of models of learning and evolution by carrying out experiments. The results presented in this paper may be relevant in several ways. First, they are in accordance with the results reported by Friedman (1995), who reproduced in the laboratory the anonymous random matching environment considered here. In what he terms "Type 1 Games", Friedman found convergence in average strategy to a mixed ESS although most subjects tended to stick to a single pure strategy. Second, Mookherjee and Sopher (1994), for example, attempt to determine whether fictitious play or gradient type rules best describe the learning behaviour of their subjects. As we have shown, the differences between these two types of model, in a random-matching environment at least, are smaller than previously thought. Our results would also point to a reason why Gale et al. (1995), using the replicator dynamics, and Roth and Erev (1995), using a gradient type learning process obtain similar results in trying to simulate the behaviour of experimental subjects playing the ultimatum bargaining game. Third, there has been some debate (Brown and Rosenthal, 1990; Binmore, Swierzbinski, and Proulx, 1994) about what constitutes convergence to equilibrium in experimental games. What we show here is that it may be foolish to expect more than convergence in the average strategy in a population of players. Last, we offer further predictions to be tested. Games which possess ESSs should converge. For games which possess positive definite equilibria, our predictions are equally clear. Learning processes should not converge to such equilibria.

Finally, as we noted in Section 1, under fictitious play for some mixed strategy equilibria there is convergence in beliefs without convergence in play. In the random-mixing models considered here, the opposite is possible. The distribution of strategies in the population matches exactly the equilibrium strategy profile. However, individual agents play any mix over the strategies in its support, including a single pure strategy. One might say that none has "learnt" the mixed strategy equilibrium, but equally, given the assumption of random matching none has an incentive to change strategy.

Chapter 3

Learning and Evolution in a Heterogeneous Population

Abstract

A framework is proposed for investigating the effect of evolutionary selection on a population where some agents learn. It is shown that learning behaviour when aggregated has different properties than when considered at the level of the individual and that a combination of learning and evolution has different properties in terms of stability than when considered separately. Convergence is shown for all 2×2 games and a famous 3×3 example.

3.1 Introduction

Game theorists have recently shown an increasing interest in modelling both learning and evolution. Nash equilibrium (and its refinements) place strong requirements on the rationality and the computational ability of players and on the information they must possess. In switching to models with boundedly-rational agents the hope has been not only to weaken those demands but also to select between equilibria in a manner which is more intuitive. Unfortunately, the dynamics considered do not necessarily converge and thus fail to give clear predictions. The results here indicate that in part this failure arises from too narrow a focus. Most research has concentrated on properties of individual algorithms. We examine a model where there is both learning and evolution and find quite different results from when they are considered separately. In particular, there is convergence for a wider class of games.

There are obvious similarities between the properties of adaptive learning and evolutionary dynamics. Typically, both are concerned with the development of the distribution of strategies within some large population[1]. As Cabrales and Sobel (1992) show, evolutionary dynamics under certain conditions can be "consistent with adaptive learning" in the sense of Milgrom and Roberts (1991). But this is only a condition on the asymptotic behaviour of a selection or learning process. In the short run, although "consistent", different processes may behave quite differently. In particular, while selection dynamics are typically smooth functions of current strategy distributions, under fictitious play or Cournotian dynamics, where players make best responses to previous play(s) of opponents, there can be discontinuous jumps in play. Convergence to mixed strategies is in particular troublesome (for example, see Fudenberg and Kreps 1993; Jordan, 1993). Here it is shown that if one aggregates such

[1]Some papers in the first camp include Milgrom and Roberts (1991), Kandori et al. (1993), Young (1993); in the second, Nachbar (1990) , Samuelson and Zhang (1992).

behaviour across a large population, smoothness is obtained.

The standard evolutionary dynamic framework assumes that agents compete in some game and then reproduce according to the success they obtain. Here I make the (strong) assumption that the population is randomly matched an infinite number of times in each "generation" to play the game. The population is heterogeneous in that some agents learn. At the end of their "lifespan" agents reproduce according to the success of the strategies they develop, or, to be precise, according to the limit of this learning process. Thus, there are two mechanisms that can change the mix of strategies in the population. Agents can change their own strategies, a "learning" process, and an evolutionary mechanism also chooses between different agents, the "selection" process.

The combination of the two has quite different implications for the stability of equilibrium than each considered in isolation. We show that the distribution of strategies in the population converges to Nash frequencies for all 2×2 asymmetric games and also for a famous 3×3 game first proposed by Shapley in 1964. Shapley's original pessimistic result has been confirmed and generalised by more recent research, (Jordan, 1993). It is therefore particularly striking that, even given the particular assumptions of this model, that a population can converge to the Nash equilibrium of such a game.

3.2 Learning and Selection

In this section, we first set out a standard model of evolutionary dynamics. We then explain why mixed strategies of asymmetric games are typically unstable in this setup. We go on to modify the model by the introduction of a simple learning process.

An infinite population is repeatedly, randomly matched to play a two-player normal-form game, $G = (\{1,2\}, \mathcal{I}, \mathcal{J}, A, B)$. We develop the model and notation on the basis that the game is asymmetric (in the evolutionary sense), in which case the players labelled 1 are drawn from a different "population" from the players labelled 2. For example, in the "Battle of the Sexes" game, players are matched so that a female always plays against a male. $\mathcal{I}$ is a set of n strategies, available to the first population, $\mathcal{J}$, the set of m strategies of the second population. Payoffs for the first population are determined by A, a $n \times m$ matrix of payoffs, with typical element a_{ij}, which is the payoff a member of the first population receives when playing strategy i against a member of the second population playing strategy j. B, with typical element b_{ji}, is the $m \times n$ equivalent for the second population. There are $n+m$ "types" of agent, each associated with one strategy. The state of the system can thus be summarised by the proportions of the population playing each strategy $\mathbf{x} = (x_1, ..., x_n)$, $\mathbf{y} = (y_1, ..., y_m)$. That is, the state space is the Cartesian product of the simplexes, $S_n \times S_m$ where $S_n = \{\mathbf{x} = (x_1, ..., x_n) \in \mathbf{R}^n : \Sigma x_i = 1, x_i \geq 0, \text{for } i = 1, ..., n\}$. Define the interior (or, int $S_n \times S_m$), as all states where all types have strictly positive representation, and define the boundary as all states where at least one type has zero representation. The symbol "$\cdot$" indicates multiplication by a transpose, and the notation $(A\mathbf{y})_i$ indicates the ith element of the vector in parentheses.

The problem with which we are really concerned with here is the generic instability of mixed strategy equilibria in asymmetric games under adaptive dynamics. Hofbauer and Sigmund (1988) set out the reasons for this in the case of evolutionary dynamics. In an environment where each member of the first population is randomly matched with a member of the second, the expected payoffs for the first population are $A\mathbf{y}$ and $B\mathbf{x}$ for the second. We assume that

$$\dot{\mathbf{x}} = Q(\mathbf{x})A\mathbf{y} \text{ and } \dot{\mathbf{y}} = P(\mathbf{y})B\mathbf{x} \tag{3.1}$$

where Q, P are symmetric positive definite matrices. This is a very general formulation for adaptive processes, including the evolutionary replicator dynamics and some learning processes as special cases (see Chapter 2). If we linearise the dynamics at a fully mixed fixed point ξ, we obtain

$$R = \begin{pmatrix} 0 & Q(\xi)A \\ P(\xi)B & 0 \end{pmatrix}. \tag{3.2}$$

Because the payoffs of the first population depend only on $\mathbf{y}$ and not $\mathbf{x}$ and conversely for the second population, the trace of this matrix (3.2) is zero. Consequently the eigenvalues are either a mixture of negative and positive or they have zero real part. In discrete time all such equilibria are always unstable. If we replace $\dot{\mathbf{x}}$ and $\dot{\mathbf{y}}$ in (3.1) by $\mathbf{x}(t+1) - \mathbf{x}(t)$ and $\mathbf{y}(t+1) - \mathbf{y}(t)$ respectively, then the linearisation at a mixed equilibrium is $I + R$, and has eigenvalues $1 + \mathbf{r}$, where $\mathbf{r}$ is the vector of eigenvalues of R. Given that, as we have seen, R possesses either a mixture of positive and negative eigenvalues or eigenvalues with zero real part, it is easy to show that the matrix $I + R$ always has at least one eigenvalue of absolute value greater than one. However, in continuous time, in the case of eigenvalues with zero real part, the linearisation does not determine stability, this will be determined by the equations' higher order terms. Such equilibria are not structurally stable in that small changes in the structure of the game or the dynamics will affect stability. For example, equilibria can be stable in continuous time even though unstable in discrete time.

However, most mixed equilibria in asymmetric games are saddlepoints. Saddlepoints are of course unstable and this property is structurally stable. In other words, small variations in the specifications of the dynamic cannot make the equilibrium stable. Instability of these mixed equilibria can often make intuitive sense in that there are games which also possess stable pure equilibria which seem more plausible outcomes (see the discussion of asymmetric games in Maynard Smith, 1982; or in the context of human society, Sugden, 1989). Or to put it another way, the instability allows us to select between equilibria. However, there are many games which possess a unique mixed equilibrium. A famous example is the following game first discussed by Shapley (1964). This is a saddle, with convergence only occurring if the first population starts in its equilibrium state, that is, with each of the three strategies with equal representation. From all other initial conditions, any dynamic satisfying (3.1) will diverge from equilibrium.

$$A = \begin{array}{|c|c|c|} \hline 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 1 \\ \hline \end{array} \quad B = \begin{array}{|c|c|c|} \hline 0 & 0 & 1 \\ \hline 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline \end{array} \tag{3.3}$$

Could heterogeneity help with this problem? That is, if there were a diversity in the types of behaviour present in either one or both populations, could this change the stability properties of mixed equilibria? In the framework we have sketched up to now, it will not. If we require learning rules to be expressible in the manner of (3.1) as a positive definite transformation of the payoffs, it does not matter how many such rules are present in the population. It is easy to verify that the aggregation of any such rules would itself be a positive definite function of the payoffs. To produce real qualitative change, it is necessary to consider a wider deviation.

It is tempting to look in the direction of best response dynamics or fictitious play, because they offer behaviour which is qualitatively different. Rather than offering a smooth reaction to payoffs, there can be discrete jumps in play. It is not that this in itself makes convergence properties any better. For example, Krishna and Sjöström (1995) have recently found that mixed strategies are also generically unstable for fictitious play. Rather it is the possibilities offered by the combination of different processes. Banerjee and Weibull (1995) consider the case where a proportion of the population is "rational", that is, they play a strategy which is a best reply to the current state of the population[2]. The result is quite striking. Every Nash equilibrium becomes stable. It is very interesting to see rationality only on a part of the population

[2]The rational players know that there are rational players amongst their opponents. The best reply is calculated on this basis.

is enough to give results very similar to those of standard game theory. However, given that one of the main reasons for investigating adaptive dynamics is to select between equilibria, another approach is required.

The basic model is modified here by the addition of a type capable of inheriting rules more complex than simply to play a fixed strategy. Thus, we now have $n+1$, $m+1$ types, and we work in $S_{n+1} \times S_{m+1}$. We can think of each generation being divided into an infinite number of subperiods $(0, 1, ..., s, ...)$. As a reminder, the selection process operates between generations, the learning process within generations. We assume that the $n+1$th and $m+1$th type adjust their strategies so that they play what is an optimum response to the strategy of their previous opponent: the "best-response" or Cournotian dynamic. Similar behavioural hypotheses have been employed in recent learning literature (for example, Milgrom and Roberts, 1991; Kandori et al., 1993; Young, 1993), but here the implementation is particularly simple. Agents do not need to know anything about the overall distribution of strategies in the population or to have a memory longer than one subperiod. Yet, as we will see, this is enough to ensure convergence to Nash equilibrium in a large class of games.

Thus, at any given time, different members of the additional type may be playing different strategies. Let $\mathbf{p}(s) = (p_1(s), ..., p_n(s))$ and $\mathbf{q}(s) = (q_1(s), ..., q_m(s))$ where $p_i(s)$ and $q_j(s)$ denote respectively the proportion of this $n+1$th type of population 1 playing the ith strategy, and the proportion of the $m+1$th type playing the jth strategy at a given subperiod s. As $\mathcal{I}, \mathcal{J}$ are finite, it is a standard result that for any pure strategy in $\mathcal{I}$, there exists at least one element of $\mathcal{J}$ which is a best response to that strategy. Or $\forall i \in \mathcal{I} \quad \exists\ b_{j^*i} \geq b_{ji}$. First, define $\mathcal{I}', \mathcal{J}'$ as those subsets of $\mathcal{I}$ and $\mathcal{J}$ respectively of strategies which have current positive representation in the two populations. Second, let μ_j represent the number of strategies in I which are equal best responses to strategy j. Third, let $\mathcal{J}_i = \{j \in \mathcal{J}' : i = \text{argmax}_{i \in \mathcal{I}}\ a_{ij}\}$ be the set of strategies to which i is the best reply, and, equivalently, let $\mathcal{I}_j = \{i \in \mathcal{I}' : j = \text{argmax}_{j\in\mathcal{J}}\, b_{ji}\}$.

The probability that an individual of type $n+1$, in population 1, meets an individual of type j in the second population is y_j. There is also a probability $q_j y_{m+1}$ of meeting an individual of type $m+1$ currently playing strategy j. In either case, faced with an opponent playing strategy j, the individual will play in the next subperiod a strategy which is a best reply to j. Thus, *within* each generation, each p_i evolves according to a mapping $S_m \to [0,1]$

$$p_i(s+1) = \sum_{j\in J_i} \frac{y_j}{\mu_j} + \sum_{j\in J_i} \frac{y_{m+1}q_j(s)}{\mu_j} \tag{3.4}$$

Thus, although individual choices are made according to the best-reply dynamic, the distribution of strategies in the population is a continuous function of the previous subperiod's distribution. We make the assumption that when there are alternative best responses each agent chooses independently. Then by the law of large numbers each alternative response is chosen by an equal number of agents. This explains the presence of μ_j, denoting the number of alternative best replies. Naturally if $\mathcal{J}_i = \emptyset$, $p_i = 0$, and if $\mathcal{J}_i = \mathcal{J}'$, $p_i = 1$. These represent respectively the cases where i is not a best reply to any strategy, and where it is the dominant strategy. Similarly, for the second population,

$$q_j(s+1) = \sum_{i\in I_j} \frac{x_i}{\mu_i} + \sum_{i\in I_j} \frac{x_{n+1}p_i(s)}{\mu_i}. \tag{3.5}$$

Lemma 7 *If, at time t, $1 > x_{n+1}(t), y_{m+1}(t)$, the learning process represented by equations (3.4), (3.5) has an unique fixed point $\mathbf{p}^*, \mathbf{q}^* \in S_n \times S_m$.*

Proof: Though they change between generations, within each generation the population proportions $\mathbf{x}, \mathbf{y}$ are fixed and are therefore constants for (3.4), (3.5). Consequently, the equations are simple, linear difference equations. Written in matrix form, they become

$$\mathbf{P}(s+1) = \mathbf{x}_0 + X_1\mathbf{P}(s)$$

where $\mathbf{P} = (p_1, ..., p_n, q_1 ..., q_m)$, $\mathbf{x}_0$ is the vector of terms in x_i, y_j, and X_1 is the matrix of terms in x_{n+1} or y_{m+1}. In equilibrium we have, $\mathbf{P}^* = (\mathbf{I} - X_1)^{-1}\mathbf{x}_0$. By inspection of (3.4), (3.5), it is possible to see that the coefficients on the p_i in the first n equations are all zero, as are the coefficients on the q_j in the next m equations. Hence, $(I - X_1)$ can be partitioned in the following manner:

$$I - X_1 = \begin{pmatrix} I & -X_{12} \\ -X_{21} & I \end{pmatrix}$$

Each column of X_{12} and X_{21} sums to y_{m+1} and x_{n+1} respectively. Thus $(I - X_1)$ is singular if and only if neither X_{12} and X_{21} are linearly independent of I, which can only be the case if $x_{n+1} = y_{m+1} = 1$. Otherwise, there is a unique fixed point, $\mathbf{P}^*(\mathbf{x}, \mathbf{y}) = (\mathbf{p}^*, \mathbf{q}^*)$. Because of the linearity of these equations, this solution will be a function of $(\mathbf{x}, \mathbf{y})$, continuous on the interior of $S_{n+1} \times S_{m+1}$. □

The exact value of this solution depend entirely on the value of the x_i, y_i and not on the value of $\mathbf{p}, \mathbf{q}$ at the beginning of the learning process. Furthermore, the sufficient condition for the existence of an unique fixed point is also a sufficient condition for convergence.

Lemma 8 *If, at time t, $1 > x_{n+1}(t), y_{m+1}(t)$, the learning process converges to its unique fixed point.*

Proof: (3.4), (3.5) represent a system of $n + m$ linear first order difference equations. The x_i, y_i are constant within each generation, and therefore are constants for (3.4), (3.5). In particular, the coefficients on the variables $\mathbf{p}, \mathbf{q}$ on the right hand side are the $x_{n+1}/\mu_i, y_{m+1}/\mu_j$, the sum of which in each equation have an upper bound in value of either x_{n+1} or y_{m+1}. By the elementary theory of difference equations if this sum is less than unity for all equations, so are all the roots of the dynamic system. □

It is worth remarking that here convergence is not convergence in empirical frequencies, a notion of convergence that has been forcefully criticised in the recent literature (Young, 1993; Fudenberg and Kreps, 1993; Jordan, 1993). In this case, one does not have to take a time average. As the limit approaches, strategies are actually played at limiting frequencies.

I make the assumption that payoffs during the learning process do not affect the rate of reproduction. Rather it is the limit of the learning process, denoted $(\mathbf{p}^*, \mathbf{q}^*)$ which determines reproductive fitness. This construction has some analytic convenience: if one assumes only a finite number of plays each period, the values of $\mathbf{p}, \mathbf{q}$ will be dependent on their (arbitrary) initial values. We would have to make further assumptions about how much of the behaviour learnt within a period is transmitted between the generations. For example, we could assume that each generation starts from scratch: at the beginning of each period $\mathbf{p}(0), \mathbf{q}(0)$ are randomly determined. That is, "children" learn nothing from their "parents". Or we can assume that the initial values are some function of play by the previous generation. However, using the limit, the value of $(\mathbf{p}^*, \mathbf{q}^*)$ will be the same in either case. This procedure is in any case defensible on the grounds that as (3.4), (3.5) are convergent, even after a few plays the process will be close to the limit[3].

As stated we use these limiting values to determine fitness. At the end of the learning process the total proportion of the first population adopting the ith strategy will be given by $z_i = x_i + x_{n+1}p_i^*$, and, the proportion of the second population adopting the jth strategy by $w_j = y_j + y_{m+1}q_j^*$. Given the assumption of random matching it is these overall distributions which decide fitness. For the first n, m types this will be, given the normal form game G,

$$\pi_{xi} = (A\mathbf{w})_i, \ \pi_{yj} = (B\mathbf{z})_j, \tag{3.6}$$

[3]Compare Harley's assumption (e): "The learning period is short compared to the subsequent period of stable behaviours" (1981, p613).

and for the learners,

$$\pi_{xn+1} = \mathbf{p}^* \cdot A\mathbf{w}, \ \pi_{ym+1} = \mathbf{q}^* \cdot B\mathbf{z} \tag{3.7}$$

With fitnesses defined, we can propose as a selection mechanism the following replicator dynamics:

$$x_i(t+1) = f_{xi}(\mathbf{x},\mathbf{y}) = x_i(t)\frac{\pi_{xi}+C}{\mathbf{z}\cdot A\mathbf{w}+C}, \quad y_j(t+1) = f_{yj}(\mathbf{x},\mathbf{y}) = y_j(t)\frac{\pi_{yj}+C}{\mathbf{w}\cdot B\mathbf{z}+C}, \tag{3.8}$$

where C is an arbitrary constant. Alternatively, taking the limit, as generations become arbitrarily short:

$$\dot{x}_i = F_{xi}(\mathbf{x},\mathbf{y}) = x_i(t)(\pi_{xi} - \mathbf{z}\cdot A\mathbf{w}), \quad \dot{y}_j = F_{yj}(\mathbf{x},\mathbf{y}) = y_j(t)(\pi_{yj} - \mathbf{w}\cdot B\mathbf{z}) \tag{3.9}$$

where $\mathbf{z}\cdot A\mathbf{w}, \mathbf{w}\cdot B\mathbf{z}$ are the average payoffs for the two populations. Inspection of (3.9) shows that this continuous selection mechanism has the following important property:

Invariance. As $\sum_{i=1}^{n+1} F_{xi} = \sum_{j=1}^{m+1} F_{yj} = 0$, the interior of the simplex is invariant under F. Starting from any interior point, the boundary is never reached in finite time. That is, if $(\mathbf{x}(0),\mathbf{y}(0)) \in \text{int } S_{n+1} \times S_{m+1}$, then $(\mathbf{x}(t),\mathbf{y}(t)) \in \text{int } S_{n+1} \times S_{m+1}$ for all $t \in \mathbf{R}$.

If we impose the condition that C is sufficiently large such that both denominator and numerator in (3.8) are strictly positive[4] for all i, j, invariance will also hold for the discrete dynamic f. Given that $\mathbf{p}^*, \mathbf{q}^*$ are themselves functions of the frequencies of types in the population, fitnesses will not be linear in $\mathbf{x}, \mathbf{y}$ - a usual assumption of the replicator dynamics - and perhaps not even be defined when x_{n+1} and y_{m+1} are

[4]An increase in the value of C is equivalent to the addition of an equal amount to the game matrices A, B. This will not change the best response structure or Nash equilibria but may change the qualitative behaviour of the discrete replicator dynamics. See Cabrales and Sobel (1992) for a discussion of the issues involved.

equal to one. However, by Lemmas 7 and 8, fitnesses are continuous functions of $\mathbf{x}, \mathbf{y}$ elsewhere. This, combined with invariance implies that from any fully-mixed initial conditions, (that is, $x_i > 0, i = 1....n + 1$ and $y_j > 0, j = 1...m + 1$), the learning process converges, and fitnesses are defined, for all $t \in \mathbf{R}$. Thus while both f and F are not continuous on all of $S_{n+1} \times S_{m+1}$ they are continuous on its interior. In other words, both f and F possess a limit even along a dynamic path with an accumulation point on the boundary of $S_{n+1} \times S_{m+1}$, even if that limit may be path-dependent.

What is important about this definition of fitness is that there is a fundamental difference from the standard evolutionary model. Fitnesses for the first population, for example, depend on $\mathbf{w}$ which through $\mathbf{q}^*$ depends on $\mathbf{x}$. Consequently $d\pi_{xi}/d\mathbf{x} \neq 0$ and any linearisation at a fully mixed equilibrium does not have the same structure as (3.2). That is, mixed strategy equilibria of asymmetric games are not generically unstable in this model. What however are their stability properties is as yet unknown. This we now investigate.

3.3 Equilibrium

Equilibrium in this model consists of a population distribution which is a rest point for both selection and learning processes. That is, a state of the system where the limit of the learning process is such that all types present in the population earn the same average payoff. In the standard evolutionary model, that is, in the absence of the learners, under the selection dynamics defined by (3.8) or (3.9), denote the rest points for the game G in the interior of $S_n \times S_m$, $(\mathbf{x}^*, \mathbf{y}^*)$. It is well known that such rest points are Nash equilibria (Hofbauer and Sigmund, 1988; Nachbar, 1990)[5]. For the extended game, the conditions for an interior rest point under the selection

[5]All states that consists of just one type are also rest points, but not all are Nash equilibria.

dynamics are

$$\pi_{x1} = ... = \pi_{xn+1}, \ \pi_{y1} = ...\pi_{ym+1} \tag{3.10}$$

Furthermore, as $(\mathbf{x}, \mathbf{y})$ are both constant if (3.10) holds, the limit for the learning process is also unchanging across all subsequent generations. The consequent distribution of strategies is a Nash equilibrium. Comparison of equations (3.6), (3.7), reveal that any values of $(\mathbf{x}, \mathbf{y})$ that satisfy the above condition (3.10), also satisfy $x_i + p_i^* x_{n+1} = x_i^*$, and $y_j + q_i^* y_{m+1} = y_j^*$. That is, it is a Nash equilibrium for the original game G in the sense that an outside observer would see, as the learning process reached its limit, strategies being played with the Nash equilibrium frequencies, $(\mathbf{x}^*, \mathbf{y}^*)$. Note that for each population there is now one less independent equation than there are independent variables. This means that any isolated equilibrium of the original game in the interior of $S_n \times S_m$ will be represented by a continuum of fixed points in the interior of $S_{n+1} \times S_{m+1}$.

Furthermore, we can show that for all 2×2 games the system will converge to a Nash equilibrium. (3.11) gives a generalised 2×2 game.

$$A = \begin{array}{|c|c|} \hline a_{11} & a_{12} \\ \hline a_{21} & a_{22} \\ \hline \end{array} \quad B = \begin{array}{|c|c|} \hline b_{11} & b_{12} \\ \hline b_{21} & b_{22} \\ \hline \end{array} \tag{3.11}$$

Define $a_1 = a_{12} - a_{22}, a_2 = a_{21} - a_{11}$; $b_1 = b_{12} - b_{22}, b_2 = b_{21} - b_{11}$. If $a_1 a_2 > 0$ and $b_1 b_2 > 0$ then there is a mixed Nash equilibrium where the first strategy of each population are represented with frequencies $(b, a) = (\frac{b_1}{b_1 + b_2}, \frac{a_1}{a_1 + a_2})$ respectively. The interesting case is when $a_1 b_1 < 0$, as in this case, the mixed Nash equilibrium is unique, yet the standard evolutionary dynamics do not converge. However, the addition of an arbitrarily small initial population of learners is enough to stabilise the dynamics. The problem with proving this is that there is, as noted above, a continuum of equilibrium points. The key is a transformation of the variables, which also has the advantage of allowing us to look at the problem on $S_2 \times S_2$.

Proposition 8 *If $a_1 b_1 < 0$, then the mixed equilibrium is asymptotically stable.*

Proof: Define $\alpha = x_1/(x_1 + x_2)$ and $\beta = y_1/(y_1 + y_2)$. Given the continuous time replicator dynamics (3.9), this gives us

$$\dot{\alpha} = \alpha(1-\alpha)[(A\mathbf{w})_1 - (A\mathbf{w})_2], \quad \dot{\beta} = \beta(1-\beta)[(B\mathbf{z})_1 - (B\mathbf{z})_2].$$

Clearly, these new equations have a fixed point where $(z_1, w_1) = (b, a)$. Taking the linearisation at this point we have the Jacobian

$$J = \begin{pmatrix} (a_1 + a_2)\alpha(1-\alpha)\frac{\partial w_1}{\partial \alpha} & (a_1 + a_2)\alpha(1-\alpha)\frac{\partial w_1}{\partial \beta} \\ -(b_1 + b_2)\beta(1-\beta)\frac{\partial z_1}{\partial \alpha} & -(b_1 + b_2)\beta(1-\beta)\frac{\partial z_1}{\partial \beta} \end{pmatrix}$$

Without loss of generality, take a_1 to be negative. Note that in this case (3.4), (3.5) gives us that $z_1^* = (x_1 + x_3 y_1 + x_3 y_3)/(1 + x_3 y_3)$ and $w_1^* = (y_1 + y_3 - x_1 y_3)/(1 + x_3 y_3)$. Substituting out x_3 and y_3, we can calculate $\partial w_1^*/\partial\alpha > 0$, $\partial w_1^*/\partial\beta < 0$, $\partial z_1^*/\partial\alpha > 0$, $\partial z_1^*/\partial\beta > 0$. Thus J has the sign pattern

$$\begin{pmatrix} - & + \\ - & - \end{pmatrix}$$

and therefore has eigenvalues with real part negative. Hence the equilibrium is asymptotically stable.

The equivalent discrete system has linearisation $I + J'$, where J' is identical to J except that the first row is divided by $(\alpha, 1-\alpha) \cdot A\mathbf{w} + C$, and the second by $(\beta, 1-\beta) \cdot B\mathbf{z} + C$. Hence, J' also has eigenvalues with real part negative, whose absolute value decreases to zero as $C \to \infty$. Thus, there is a C for which the eigenvalues of $I + J'$ are less than one in absolute value. □

Note that not all mixed strategies are stable. That is, the dynamics can still be used to select between equilibria. If $a_1 b_1 > 0$ then this interior equilibrium is a saddle. Similar arguments to those employed in Proposition 8, can be used to show that in this case, the system behaves in much the same way as standard evolutionary dynamics and flows toward the Nash equilibria located on the boundaries of the simplex.

The mixed equilibrium when $a_1 b_1 < 0$ is non-hyperbolic (that is, the linearisation has eigenvalues with zero real part) for the continuous time replicator dynamics and hence not structurally stable. In this sense, a modification of the dynamics would be expected to change their qualitative behaviour. However, for the discrete time dynamics, the equilibrium is hyperbolic and hence robustly unstable. We have shown that even in this case, the addition of learning can stabilise the equilibrium. We go on to show that it can drive convergence to an equilibrium which seems to be unstable under every form of adaptive dynamic.

3.4 A 3×3 Example

The famous example given by Shapley (1964) to demonstrate non-convergence of fictitious play is shown in (3.3). The only Nash equilibrium of this game is interior, where both row and column play each of their strategies with equal probability. As we have seen, interior (mixed) equilibria of asymmetric games are never asymptotically stable under the replicator dynamics. Thus this game does not converge for the replicator dynamics, just as it does not for fictitious play. Recent research on learning and evolution has only served to confirm the robustness of this result (see for example, Jordan, 1993). However, under this modified system this game converges to the unique Nash equilibrium.

Starting from a fully-mixed initial state, the proportions of type 4 playing each strategy evolve according to:

$$\mathbf{p}(s+1) = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} + y_4\mathbf{q}(s), \ \ \mathbf{q}(s+1) = \begin{pmatrix} x_3 \\ x_1 \\ x_2 \end{pmatrix} + x_4 \begin{pmatrix} p_3(s) \\ p_1(s) \\ p_2(s) \end{pmatrix} \tag{3.12}$$

This is a system of six linear difference equations. By Lemma 2 we know that the fixed point of this system is the limit of the learning process. This can be calculated using standard methods. It would be possible to eliminate $\mathbf{p}, \mathbf{q}$ by substitution using these results. However, it is easier to work in the other direction. We construct $z_i(t) = x_i(t) + x_4(t)p_i^*(t), w_j(t) = y_j + y_4(t)q_j^*(t)$; $i, j = 1, 2, 3$, where z_i is the total number of the first population playing strategy i, and w_j is the total number of the second playing strategy j. Note that (3.12) here implies that $\mathbf{z}(t) = (q_2^*(t), q_3^*(t), q_1^*(t))$, that $\mathbf{w} = \mathbf{p}$ and that $\mathbf{q}{\cdot}B = \mathbf{z}{\cdot}A$. There is an interior equilibrium for this system: the plane such that $x_1 = x_2 = x_3$, $y_1 = y_2 = y_3$, which we denote $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$. In such an equilibrium, (3.12) in turn implies that $\mathbf{p} = \mathbf{q} = (1/3, 1/3, 1/3)$. I now prove that the limit point of all solutions under f, given fully-mixed initial conditions, is on this plane (normally for the discrete dynamics the interior equilibrium is a repellor).

Proposition 9 *The plane of equilibria* $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ *under* f *attracts all other points on the interior of* $S_4 \times S_4$.

Proof: Define $V(\mathbf{x}, \mathbf{y}) = x_4 y_4$. Given that $x_4(t+1) = x_4(t)\dfrac{\mathbf{p}{\cdot}A\mathbf{w}}{\mathbf{z}{\cdot}A\mathbf{w}}$, and that $y_4(t+1) = y_4(t)\dfrac{\mathbf{q}{\cdot}B\mathbf{z}}{\mathbf{w}{\cdot}B\mathbf{z}}$, it follows that $V(t+1) - V(t) > 0$ if and only if

$$\mathbf{p}{\cdot}A\mathbf{w}\ \mathbf{q}{\cdot}B\mathbf{z} - \mathbf{z}{\cdot}A\mathbf{w}\ \mathbf{w}{\cdot}B\mathbf{z} = \mathbf{w}\cdot\mathbf{w}\ \mathbf{z}\cdot\mathbf{z} - \mathbf{z}\cdot\mathbf{w}\ \mathbf{w}\cdot\mathbf{z}^* > 0 \tag{3.13}$$

where $\mathbf{z}^* = (z_3, z_1, z_2)$. Divide through by $\mathbf{w}\cdot\mathbf{w}\ \mathbf{z}\cdot\mathbf{z}$ to obtain:

$$1 - \cos\theta_{zw}\cos\theta_{wz^*} \geq 0$$

It follows that $V(t+1) \geq V(t)$ with equality only at $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$. $V(\mathbf{x}, \mathbf{y})$ is therefore a strict Liapunov function on all of the interior of $S_4 \times S_4$ less $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$. It is therefore unclear whether the system will have its limit at $V = 1$, or whether it will come to rest at another point on $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$. However, in either case $\lim_{t\to\infty} \mathbf{p} = \lim_{t\to\infty} \mathbf{q} = (1/3, 1/3, 1/3)$.□

We consider briefly two other examples. The first (3.14) is the familiar ROCK-SCISSORS-PAPER game, the second (3.15) is a game proposed by Dekel and Scotchmer (1992). They show that the DUMB strategy survives in the limit under the discrete replicator dynamics although it is never a best response and therefore not rationalizable.

$A = B =$

ROCK	b	a	c
SCISSORS	c	b	a
PAPER	a	c	b

$a > b > c$ (3.14)

ROCK-SCISSORS-PAPER is well-known as a problem game. While it does converge for fictitious play, it does so only in empirical frequencies. It (typically) does not converge for the discrete replicator dynamics. As for the first example, the limit of these games when learners are also present is the unique Nash equilibrium. As both these two additional examples have a similar structure, it is not surprising that they elicit similar behaviour. The function, $x_{n+1}y_{m+1}$, will again work as a Liapunov function and shows that in both cases there is convergence in population frequencies to the unique Nash equilibrium.

$A = B =$

ROCK	1	2.35	0	0.1
SCISSORS	0	1	2.35	0.1
PAPER	2.35	0	1	0.1
DUMB	1.1	1.1	1.1	0

(3.15)

3.5 Discussion

Games such as (3.3) cause problems for conventional models because they possess cycles of best responses. One might think that random perturbation, for example, trembles or mutations, would also break up these deterministic cycles. However, this is not the case. For example, in Young (1993), cyclic games are excluded from the results on the convergence of a stochastic learning process. Adding noise enables the system to jump between the different possible paths of the original deterministic process. If, as is the case for (3.3), all such paths are divergent, the stochastic process must be also.

The fundamental reason that this model gives qualitatively different behaviour is that there are two distinct processes determining the change in the distribution of strategies, working at different speeds. By changing strategies, the learners anticipate the next stage of the cycle and "damp" the non-convergent tendencies of the original model. The dependence is two-way. Without the non-learners, the best-response process would not converge for this game.

One might argue that the simple learning rule considered here would be displaced by more sophisticated behaviour. For example, Harley (1981) claims that for a learning rule to be evolutionary stable it must be a "rule for ESSs". That is, it must be able to lead the population to the evolutionary stable strategy (ESS) in one generation. This would suggest that any learning rule which was able to survive evolutionary selection would have to be fairly flexible and sophisticated. However, some doubts have been cast on Harley's model and methodology (Maynard Smith et al., 1984; Houston and Sumida, 1987). The latter paper raises a further point. While Harley's claim might have some validity for single-agent optimisation problems, when there is strategic interaction with other agents things may be very different. That is an argument that finds support in more recent work (Banerjee and Weibull, 1995; Blume and Easley, 1992; Stahl, 1993), where rational agents do not necessarily displace less rational ones. There is no claim that the learning rule considered here is the "correct" one. However, there is also no strong evidence that evolution will select for more complex or sophisticated behaviour in a strategic environment.

Learning and evolution are ostensibly similar processes. However, while evolution is defined at the level of a population, learning is carried out by individuals. Crawford (1989) demonstrates that even when agents' learning is modelled in a similar manner to the replicator dynamics, an aggregation of their behaviour does not have the same properties in terms of stability as evolutionary dynamics. Similarly, in this paper even the most elementary learning behaviour gives increased stability when considered at the level of the population. This opens up the possibility of further research about the aggregate properties of populations where a number of different classes of behaviour are present.

Chapter 4

Price Dispersion: an Evolutionary Approach

Abstract

In many markets it is possible to find rival sellers charging different prices for the same good. Earlier research has explained this phenomenon by demonstrating the existence of dispersed price equilibria when consumers must make use of costly search to discover prices. This paper re-examines the question of price dispersion from an evolutionary, disequilibrium perspective. That is, firms and consumers adjust behaviour adaptively in response to current market conditions. We find that dispersed price equilibria are unstable when consumers use a fixed sample size search rule but may be stable when a reservation price rule is used.

4.1 Introduction

It is a common experience to find that prices vary between different sellers, giving consumers an incentive to search for low prices. Stigler (1961) introduced the notion of modelling consumer search as repeated random draws from the current distribution of prices. In response Rothschild (1973) set up a challenge. He argued that it was not enough to examine, as had Stigler, the search behaviour of consumers faced with an exogenous distribution of prices. Sellers, presumably, would only charge prices different from those of their competitors if they could make a profit by doing so. To explain price dispersion, economists must show that such price-setting behaviour was a rational response by traders to the search behaviour of consumers, and vice-versa. In game-theoretic terms, it was necessary to find a Nash equilibrium, where each firm and consumer adopted a strategy that was a best reply to the play of all other firms and consumers.

In fact, Diamond (1971) had introduced a model which satisfied Rothschild's conditions. However, the model's main result is usually viewed as a paradox. Diamond was able to show that for any positive search costs, in equilibrium, no consumer would search, and all firms would charge the price that maximised joint-profits. This is clearly a Nash equilibrium: when prices are identical, there is no incentive to search; when there is no search, there is no incentive to cut prices to increase sales. Note that the converse state where all consumers are fully-informed and all firms charge a competitive price cannot be a Nash equilibrium. For positive search costs and with all prices identical, active search is not optimal. Since consumers are not fully informed, firms can raise prices without losing all customers. While those economists raised on the "Law of One Price" might have expected price dispersion to be fragile it was surprising that the collapse was in this direction.

Faced with this challenge, subsequent authors, (a partial list includes Salop and

Stiglitz, 1977, 1982; Wilde and Schwartz, 1979; Varian, 1980; Burdett and Judd, 1983; Rob, 1985; Bester, 1988; Wilde, 1992; Benabou, 1993), produced models with dispersed price equilibria. However, there remains an unresolved problem with this earlier literature, that of multiple equilibria. There may be more than one equilibrium at which prices are dispersed, and typically, the joint-profit maximising outcome found by Diamond remains a Nash equilibrium even in the presence of these others.[1] Selecting between these equilibria is not straightforward. It is easily verified that for strictly positive search costs the joint-profit maximising outcome is a *strict* Nash equilibrium. That is, any deviation leads to strictly lower payoffs. It cannot therefore be easily dismissed.[2]

The other striking difference about the model of Diamond (1971) is that it is "A Model of Price Adjustment" not of equilibrium. There are several advantages to such a disequilibrium approach. First, it answers the question of how an economy arrives at equilibrium and why one equilibrium is chosen over another. Second, it allows a different approach to the modelling of consumer search. Some models assume as did Stigler that consumers use a fixed sample size search rule, that is, the consumer's problem is to choose a sample size n. The consumer then collects n prices and then takes the lowest offer. More popular has been the assumption of sequential search, that is, after each price quotation the consumer must decide whether to buy at that price or to obtain a further quotation. However, in both cases the common if implausible assumption is that the consumer knows the distribution of prices before starting searching. Here, just as did Diamond (1971), we can relax this assumption. Consumers do not have to know the distribution of prices in order to determine the optimum level of search effort; it can be learnt from experience or from the experience

[1]This problem occurs in many different models. See Wilde (1992).

[2]For a formal definition of strictness and its place in the heirarchy of equilibrium refinements, see for example van Damme (1991).

of other consumers.[3]

More recently, disequilibrium models have been back in fashion under the title of learning and evolution. While this work has up to now mostly been on a very abstract level, there may now be enough theoretical ammunition to analyse the problem of price dispersion. Here we present an dynamic, evolutionary model which is able to select between the multiple equilibria present in these models. Evolutionary models are not new to economics but have not always been well received by economists. In particular, the assumption of evolutionary game theory that agents use a single fixed strategy which determines their rate of reproduction may seem ill-suited to economic applications. However, the translation of models between disciplines is not intended to be over-literal. In human society, the births and deaths are of ideas and strategies, not people. We can assume that at each point in time a population of individuals have to choose between different strategies. The state of the system can be summarised by the proportions of the population playing each strategy. The system changes state as agents change strategies. In fact, in an earlier paper (Hopkins, 1995) I was able to show that if one aggregates such learning behaviour across a population, the resulting aggregate dynamic is qualitatively similar to the evolutionary replicator dynamics.[4] The exact dynamics do not have to be specified, rather it is possible, as shown by for example, Nachbar (1990), Friedman (1991), Samuelson and Zhang (1992) and Kandori, Mailath and Rob (1993), to work with wide classes of dynamics, which share certain qualitative properties. While this might not represent the behaviour of perfectly rational agents, it encompasses a wide range of plausible adaptive processes and learning schemes, including some quite sophisticated behaviour.

[3] Unfortunately, we were unable to obtain analytic results for simultaneous learning by sellers and consumers in the case of sequential search.

[4] Both Hofbauer and Sigmund (1988) and van Damme (1991) present excellent surveys of the field of dynamic evolutionary game theory, including description of the replicator dynamics.

A further criticism might be that much of the recent work in evolutionary games has been in the context of two-player normal-form games, a context that does not encompass most economic problems. Indeed, the game here is an asymmetric game with many players. These are divided in two distinct groups, buyers and sellers. The payoff of each player depends upon the actions of all other players. The payoffs are not linear. Firms can choose from a continuum of prices. It is a situation very far from that of a normal-form game. Nonetheless, it is possible to apply the same evolutionary techniques.

Up to now, evolutionary and learning models have been applied to similar problems by using a discrete approximation of a continuous strategy space (for example, Roth and Erev, 1995). This may be a problem in that, for the many models of dispersed price equilibria, the existence and character of equilibrium depends on the properties of the continuum. Hence we develop learning dynamics on the Hilbert space $\mathcal{L}_2$, where we look at the evolution of the functions describing the distribution of prices. While this does involve some technical difficulties, we are able to show that even when firms can choose from a continuum of prices, it is possible to obtain clear results on the stability of dispersed price equilibria under different assumptions. In particular, dispersed price equilibria may be stable when consumers use a reservation price rule when searching but unstable when they use a fixed sample size rule.

This result rests on quite a simple argument. Assume that firms and consumers are at some dispersed price equilibrium. Then assume that a positive mass of firms simultaneously raise prices, say from p_1 to p_2, while consumer behaviour remains unchanged. If consumers are using a reservation price rule then these firms will simply lose all the customers with reservation prices on the interval $[p_1, p_2)$. However, if consumers use a fixed sample size rule, some of them may be unlucky and only draw prices greater than or equal to p_2. Thus, the fall in demand is not as great as in the previous case. Furthermore, there is a positive return in raising prices. By raising prices, sellers increase the probability that a consumer will be "unlucky" and find only high prices, and therefore increase the expected sales of all high priced sellers. A consequence of this externality is that deviations from equilibrium are self-reinforcing. In this case, dispersed price equilibria are unstable.

4.2 Some Simple Examples

In classical perfect competition there can only be one price. While concerned with a market for a good which is entirely homogeneous, this result allows for a certain amount of other types of heterogeneity: consumers may have different marginal valuations of the good, firms different costs; but in the absence of any monopoly power, and, importantly, with perfect information, a single market-clearing price will be charged. This is in contrast with the existing evidence, formal and informal, that for many goods, there is a wide dispersion of prices.

There is a line of skepticism that argues that all apparent differences in prices can be explained by heterogeneity in terms of quality, distances in time or space. The convenience store sells its goods at a higher price, but only has significant custom when the neighbouring supermarket is closed. Shop A charges more than B, but offers superior service. The argument on quality can be countered by the existing empirical data where the variance in prices seems too large to support such an explanation. As to differences in time and space, these are not inconsistent with most models of price dispersion. Typically, they are concerned with consumers who differ in search costs, which of course may include such factors as impatience or willingness to travel. A more telling criticism is that, if sellers are clearly differentiated in either time or space, modelling search as purely random draws from a distribution of prices is not appropriate.

Nonetheless, it seems intuitively plausible that identical goods could be sold in the same time and place for different prices, some retailers selling a high volume at a low margin, others selling a low volume at a high margin. However, the possibility of such an equilibrium is going to depend quite delicately on the specification of, first, the sellers' costs, so that it is possible that a high volume is as profitable as a low volume, and, secondly, consumers' information and behaviour, their search "technology", so that demand is decreasing in price, but not so abruptly as under perfect competition.

We consider two cases, first, where consumers use a fixed sample size rule, and second, where search is sequential. In the first case, the consumer must decide how many quotations to obtain at a constant cost c (the convention is that the first quotation is free). Only once all the n price quotations have arrived can the consumer purchase from the firm that offers the lowest price. Such nonsequential search can be optimal (Morgan and Manning, 1985), and fits the case where a consumer must write away for quotations, or where a number of quotations can be obtained by buying a magazine or newspaper. Sequential search is where a consumer obtains one quotation and then decides whether to take another. The classic result is that when the distribution of prices is known, optimal sequential search takes the form of a reservation price rule. That is, the consumer decides on a target price and continues to search until it is found. What is common to both forms of search is the possibility of *ex post* heterogeneity of information. That is, while starting out with identical information, some consumers will find a better price than others. This is of course what allows the existence of a dispersed price equilibrium.

We now look at the simplest possible example of such equilibria and how learning dynamics can be applied to a market game of this kind. We are concerned with a market for a homogeneous good. For example, the same model of car or computer from a particular manufacturer is often sold by many different outlets, often at different prices. Consumers buy this product only infrequently. The sellers we can think of as a continuum of identical small shops, which buy the good from a wholesaler for a constant cost, which here we assume to be zero. We assume (an average of) μ customers per seller. Firms choose prices in order to maximise profits. There is an upper bound on prices p^*, which can be seen as the profit-maximising price in a monopoly situation. A continuum of consumers are uninformed about which firms charge which prices. They must engage in costly non systematic search in order to obtain price quotations.

For illustrative purposes we simplify even further assuming that firms have a choice between only two prices $\{\underline{p}, p^*\}$. This is not intended to be realistic but the same intuition drives the result in this case as when the analysis is much more complicated. We first assume that consumers search with a fixed sample size rule, and again for sake of simplicity, we assume that they must choose between sampling one and sampling two prices. Their expected costs, if a proportion x of the sellers charge $\underline{p}$ and $1-x$ charge p^*, will be $x\underline{p}+(1-x)p^*$ in the first case and $c+(x^2+2x(1-x))\underline{p}+(1-x)^2p^*$ in the second. If q consumers choose the first option, and $1-q$ the second, then sellers' profits are respectively

$$\underline{p}\mu[q+(1-q)(2-x)], \text{ or, } p^*\mu[q+(1-q)(1-x)]. \tag{4.1}$$

If search costs are not too high, then we have the structure of equilibrium illustrated in Figure 1. There are two values of x, $\{\underline{x}, \bar{x}\}$ for which consumers are indifferent between searching once and searching twice. The variance of prices, and hence the expected return to search, is at a maximum at $x = 0.5$. At $x = 1$ or 0, searching once dominates searching twice. The curve maps the equal profit line between the sellers' two possible strategies. It is upward sloping because as q rises the profits of firms charging p^* rises. Profits can only be kept equal if the number of low-priced firms x increases, which reduces the number of customers at the high-priced firms.

There are two interior equilibria, and one at the bottom right hand corner (the no-search outcome). The arrows are generated by the simple assumption that the

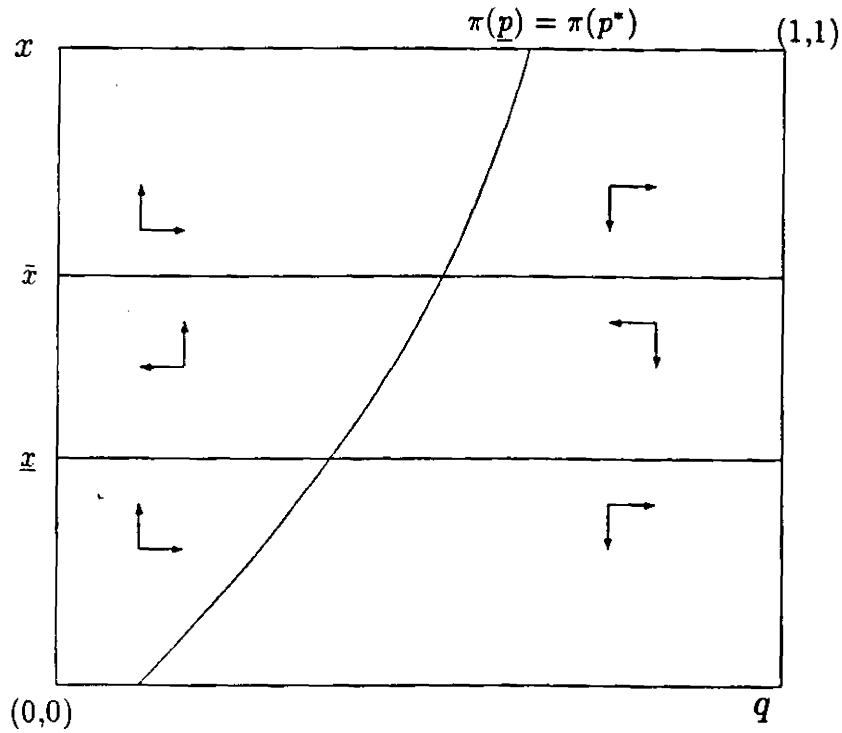

Figure 4.1: A simple example

proportion of the population playing the strategy which is earning a higher payoff will grow at the expense of the other. For example, near the no-search equilibrium, searching once is more profitable than searching twice, and p^* more profitable than $\underline{p}$. Hence the arrows point right and down. Both the interior equilibria will be unstable under any such simple dynamic. This is simply because the equal profit curve is upward-sloping. Looking at (4.1), it is clear that the profits of the firms charging p^* are increasing in their population share $1 - x$. There is a positive externality in between these sellers in that the more of them that there are, the less the probability of consumers finding a better price. Hence, a deviation from equilibrium which, for example, increases the market share of the high-priced firms, (a shift downwards in the Figure), increases their profits, leading to a deviation of increased size.

Now we consider the same simple model but with consumers adopting a reservation price rule. This time let q consumers have a reservation price greater than or equal to

p^* and $1-q$ have a reservation price on the interval $[\underline{p}, p^*)$. That is, the latter group will always refuse an offer of p^*, while all consumers will accept $\underline{p}$. Sellers' profits are therefore

$$\underline{p}\mu[q+(1-q)/x], \text{ or, } p^*\mu q,$$

while consumers must expect to spend

$$x\underline{p}+(1-x)p^*, \text{ or, } c(\frac{1}{x}-1)+\underline{p}.$$

This last result follows because a consumer has an expected $1/x$ searches before finding an offer of $\underline{p}$ (again the first quotation is free). It is easy to see that just as in the fixed sample size case, there may exist an interior equilibrium where sellers are indifferent between the two possible prices. However, the difference is that the equal profit curve for sellers is downward sloping, falling from infinity at $q=0$ to zero at $q=1$. This suggests that stability properties when consumers use a reservation price may be different from the case where they use a fixed sample size.

This is of course a very simple analysis and hence possibly misleading. It might be argued, for example, that the instability in the case of a fixed sample size rule arises solely from the inadequate nature of the discrete approximation of a continuous strategy space. The positive externality shared between the highest-priced firms depends on the existence of positive mass of firms charging that price, something that may not occur when the strategy space is a continuum. For that reason we will examine the dynamics when sellers can choose from a continuum of prices. We find that it is still the case that the dispersed price equilibrium is unstable. However, the first step is to discuss the exact structure of dispersed price equilibrium in such circumstances.

4.3 Dispersed Price Equilibrium

We consider equilibrium in both the cases where consumers use a fixed sample size rule and the case where they search sequentially. We continue to assume that there is a continuum of sellers with constant zero marginal cost. Burdett and Judd (1983) demonstrate how a dispersed price equilibrium can arise without any heterogeneity either amongst firms or consumers. The important assumption is that there is *ex post* heterogeneity in consumer information: a proportion q_1 of consumers know of one price, q_2 have two price quotations and so on. This can arise because either consumers use a fixed sample size rule or if consumers search sequentially, but the search is noisy, for each search made there is a probability q_k of finding k quotations simultaneously. If the distribution of prices is given by the cumulative distribution function $F(p)$, the probability for consumers that a given price p is the lowest that they find with two quotations is $2(1-F(p))$, after three $3(1-F(p))^2$. Hence, profits for firms are then

$$\pi(p) = p\mu \sum_{k=1}^{\infty} q_k k(1 - F(p))^{k-1} \tag{4.2}$$

where μ is the average measure of consumers per firm. Burdett and Judd show that the only possible distribution of prices in equilibrium must have continuous support on the interval $(\underline{p}, p^*)$, where $\underline{p}$ is to be determined endogenously. If there were a gap in the distribution on some interval (p_i, p_j), a firm charging p_i could raise its price to fill the gap without losing any customers. If there were a mass point in the distribution at some price p_i, a firm could cut its price from p_i by an arbitrarily small amount and gain a discrete jump in sales.

In fact, Burdett and Judd show that when consumers use a fixed sample size search rule, a dispersed price equilibrium is only possible when a proportion $1 > q > 0$ of consumers buy one quotation, and all others $1 - q$ obtain two. In equilibrium, the profit for all firms must be equal. Given (4.2), and

$$\pi(p^*) = p^*\mu q = \underline{p}\mu\left[q + 2(1-q)\right] = \pi(\underline{p}),$$

we can solve for both $\underline{p}$ and $F(p)$. Denoting the equilibrium cumulative distribution function by Φ, we have

$$\Phi = \begin{cases} 0, & p < \underline{p} \\ 1 - \dfrac{p^* - p}{p}\dfrac{q}{2(1-q)}, & \underline{p} < p \leq p^* \\ 1, & p > p^* \end{cases} \tag{4.3}$$

and

$$\underline{p} = p^* \frac{q}{2-q}$$

Note that the equilibrium distribution is continuous. There is no mass point at p^*. Thus there does not seem to be the externality present in the simple example of the last section. As we shall see this is misleading. The externality is still there, and the equilibrium is still unstable.

Is this an equilibrium for consumers? If the price distribution is given by Φ, the difference between the expected price paid by a consumer who searches once and a consumer who searches twice is given by

$$V(q) = \int p \; d\Phi(p) - 2\int p(1 - \Phi(p)) \; d\Phi(p).$$

This is a continuous function of q with a unique maximum on $(0,1)$. That is, if c is less than the maximum, there are two values of q such that $V(q) = c$, that is, such that consumers are indifferent between their two strategies. If c is higher than the maximum, so that search is not worthwhile, no dispersed price equilibrium exists. Burdett and Judd note however, that another equilibrium exists (irrespective of the value of c). It is the same outcome that Diamond (1971) found. That is, the state with $q = 1$ and, for all firms, $p = p^*$.

Diamond's result suggest that a sequential search rule is not conducive to price dispersion. In general terms, given a uniform cost to each search, a population of rational, maximising consumers will all settle on the same reservation price and will not buy for more. Sequential search is obviously therefore not consistent with a distribution of prices without a distribution of search costs as Rob (1985), Carlson and McAfee (1982) and Benabou (1993) consider.

Here, we follow Rob (1985) and consider again a continuum of sellers with constant zero marginal cost. Buyers have a (nondegenerate) distribution of search costs $v(c)$ which leads to a distribution of reservation prices $g(r)$. Obviously sellers who charge a price p will only sell to consumers with a reservation price $r \geq p$. These customers, given that they search randomly, will be equally divided amongst the $F(r)$ sellers whose prices they find acceptable. Thus profits, simply price times sales, for a firm charging a price p are given by the following expression,

$$\mu p \int_p^\infty \frac{g(r)}{F(r)} dr, \tag{4.4}$$

where μ is again the average number of consumers per firm. In equilibrium, the marginal benefit of search will equal its marginal cost, that is $\int_0^r F(s)\, ds = c$. Consequently,

$$g(r) = v\left(\int_0^r F(s)\, ds\right) F(r). \tag{4.5}$$

Thus, what really determines the existence and form of an equilibrium price dispersion is the distribution of the consumers' costs. Carlson and McAfee (1983) and Rob(1985) are able to calculate some examples for particular distributions. However, there is no simple or general theory in this case.

4.4 Evolutionary Market Dynamics

Having described some possible equilibria, we now deal with disequilibrium. We imagine the above one-shot game is repeated many times. That is, at each point in time firms must choose prices and consumers a search strategy, for example, how many price quotations to buy or a reservation price to search for. As is common in the literature on learning and evolution, agents do not play some complex intertemporal equilibrium. Instead they adjust their play of the stage game. In this context, firms change prices in the direction of increasing profits. This is not unreasonable in the context of the model. Firms are assumed "small" relative to the size of the market and have little strategic power. Consumers participate in this market only infrequently. This also means that firms have little incentive to build a reputation.

The parameter μ now represents the volume of the flow of consumers on to the market. This rate is fixed and exogenous.[5] We do not assume that consumers know the distribution of prices. Instead we assume a type of social learning of the type for example set out in Young (1993). That is, consumers have some access to information on how consumers have behaved in the past and how successful were the different strategies pursued. They obtain this information either from their own past experience of the market or from advice from more recent participants. Thus consumers act adaptively using past observations to form an estimate of the current distribution of prices. But there also needs to be some rule by which they decide on what response they make. One alternative would be for consumers to choose a search strategy which was a best response to the estimated distribution. This would be consistent with the idea of fictitious play perhaps the most popular learning model in the recent literature, (see for example, Milgrom and Roberts, 1991; Young, 1993; Fudenberg and Kreps, 1993).

However, there are other models of learning. It is possible to use the evolutionary replicator dynamics or their generalisations as a way of modelling human learning behaviour, for example, Friedman (1991), while another alternative is the learning

[5] Fershtman and Fishman (1992) consider the case where rational forward-looking consumers may decide to defer consumption if they expect prices to fall.

model considered by Roth and Erev (1995), or the learning by imitation model of Schlag (1994). In an earlier paper (Hopkins, 1995), it was demonstrated that the aggregation across a population of players of all these learning rules were qualitatively similar. For example, evolutionary stable strategies (a definition follows) are asymptotically stable for all dynamics of this class. It is possible therefore to work with the class of *positive definite* dynamics which include all these models as special cases. Within this framework, firms and consumers could have different learning rules, or behaviour could vary within the two populations; for example, some agents could play mixed strategies, some could play pure.

Much of the work on learning and evolution has been in the context of a large population of agents who are randomly matched to play a normal form game. Unfortunately this does not match a description of most markets, at least as traditionally modelled by economists. First, agents interact not by random matching but through the price mechanism. For example, in Cournot type competition, rather than being matched one-to-one, firms interact through the effect their decisions on output have on aggregate output and hence on price. Second, profits are non-linear in the firm's decision variable. Lastly, the strategy space is a continuum.

The first difficulty is easily overcome. Even in a random matching environment the aggregate is important because it determines each agent's expected payoffs. In the case of a symmetric normal form game with n strategies, let A be the $n \times n$ payoff matrix, and let each agent play a strategy (possibly mixed) $y \in S_n = \{y = (y_1, ..., y_n) \in R^n : \Sigma y_i = 1, y_i \geq 0 \text{ for } i = 1, ..., n\}$. If mixed strategies in the population are described by a distribution function F on S_n, then let $x \in S_n = \int y \, dF$ be the vector of the average propensity in the population to play each strategy (If all agents play a pure strategy, then x is simply the vector of proportions of the population following each strategy). Then, an individual playing the strategy y against a population with current state x, has expected payoff $y \cdot Ax$.

A positive definite dynamic is a dynamic which has the form,

$$\dot{x} = Q(x)Ax. \tag{4.6}$$

where Q is a symmetric semi-positive definite matrix. There are certain other conditions which are set out in Hopkins (1995). Similar conditions with modifications to allow for an infinite number of strategies appear in Section 5 of this paper. However, the most significant condition is simply that of positive definiteness. This ensures that the vector of changes in strategy frequencies $\dot{x}$ is at less than a 90° angle to the vector of payoffs Ax. This is thus a very weak formulation of the assumption that strategies with a high payoff grow at the expense of those with a lower return.

We can extend these dynamics to the case where profits are given by a nonlinear function. However, we again assume that agents choose between n strategies. The return to each strategy given the state of the population x is $\pi(x) = (\pi_1, ..., \pi_n)$. Then an agent playing any strategy $y \in S_n$ would receive a payoff $y \cdot \pi(x)$. We assume the dynamics to be given simply by

$$\dot{x} = Q(x)\pi(x). \tag{4.7}$$

The motivation for this approach is evolutionary. In evolutionary game theory an *evolutionary stable strategy* (ESS) is defined as a strategy which is "uninvadable". Agents playing some alternative strategy would not be able to supplant agents who stick to the original strategy. In the market games considered in this paper, in a similar way, we want to know whether a given equilibrium distribution of prices can resist any deviation by any firm or group of firms from their equilibrium strategy. The conditions for a state ϕ to be an ESS are firstly, that ϕ should be a best reply to itself, that is, a Nash equilibrium, or formally,

$$\phi \cdot \pi(\phi) \geq x \cdot \pi(\phi), \; \forall x \in S_n. \tag{4.8}$$

Second, if there are any alternative best replies, any strategy x for which $x \cdot \pi(\phi) = \phi \cdot \pi(\phi)$, then ϕ is better against them than they are against themselves.

$$\phi \cdot \pi(x) > x \cdot \pi(x). \tag{4.9}$$

What this last condition implies is a kind of concavity of the payoff function. If for example π were linear in x, and ϕ was a fully- mixed equilibrium ($\phi_i > 0\ \forall\ i$), (4.9) would become $(x - \phi) \cdot \pi(x - \phi) < 0$, (as $(x - \phi) \cdot \pi(\phi) = 0$). Now as both x and ϕ are vectors summing to one, $(x - \phi)$ is an element of R_0^n, that is the space $\{x \in R^n : \sum x_i = 0\}$. Thus, evolutionary stability implies (in most cases) that a linear profit function, such as for a normal form game, must be negative definite on R_0^n. Here, with nonlinear profit functions, what we require is that the linear approximation at the equilibrium point be negative definite. That is, if $\Pi_x = d\pi/dx$, then $x \cdot \Pi_x x < 0$, $\forall\ x \in R_0^n$. In fact, such negative definiteness is a sufficient condition for dynamic stability under positive definite dynamics.

Proposition 10 *An equilibrium point ϕ is asymptotically stable under positive definite dynamics if $\Pi_x(\phi)$ is negative definite on R_0^n and asymptotically unstable if $\Pi_x(\phi)$ is positive definite.*

This proposition is a special case of Proposition 12 which, together with its proof, can be found in Section 5.

4.5 Dynamics on an Infinite Space

As Burdett and Judd themselves suggested

> "Examples of further possible work include stability analysis which may give further information concerning the durability of equilibrium price dispersion and reduce the multiplicity of equilibria in the nonsequential model." (1983, p967)

In this section, we do indeed carry out a stability analysis, both for fixed sample size model of Burdett and Judd and also when consumers search sequentially. Ideally, we would want to analyse the dynamics the case where the two types of agents, firms and consumers, change their behaviour simultaneously. Unfortunately, this proves to be too complicated a model to obtain more than very partial results. For the moment, we consider only the behaviour of sellers, treating consumer behaviour as fixed. This we could justify by the likely observation that consumers adjust their behaviour much more slowly than do firms. Second, the problem for consumers is largely unstrategic, in that a consumer's payoffs are not affected by the actions of the other consumers. As we have seen in the simple examples of Section 2, the stability of equilibrium is largely determined by the adjustment process of the sellers.

While before we used a vector x to describe the state of the population of firms, now its role is taken by a density function f. To simplify things somewhat we normalise p^* to 1 and thus we consider distributions of prices on the interval [0,1]. As for a dynamic on a finite dimensional space, we take a linear approximation to the nonlinear dynamics at the equilibrium distribution and we find that, in the first case, this equilibrium is positive definite and hence unstable. In the second, that is when consumers search sequentially, the results are less clear cut. However, we are able to show that there is at least one dispersed price equilibrium which is stable.

This is not the first attempt to examine evolutionary dynamics with a continuous strategy space. Hofbauer and Sigmund (1990), for example, note that, unlike in finite dimensions, it is possible to have evolutionary stability without dynamic stability and vice versa. We do not attempt here to obtain any general results about the links between the two concepts. However, we have already seen that when the payoff function is non-linear, the condition (4.9) for evolutionary stability is not identical to the condition for dynamic stability, that is, that the linear approximation Π_x is negative definite.

Let E be a complex Hilbert space, that is a Banach space with the addition of a scalar product $\langle f, g\rangle : E \times E \to C$. An example is C^n, with the vector inner product $f \cdot \bar{f}$, where the bar indicates the complex conjugate. However, we will in particular be interested in the space $\mathcal{L}_2[0,1]$, that is, the space of (Lebesgue) measurable functions on the unit interval bounded in the norm $\|f\| = \left(\int f \ \bar{f}\, dp\right)^{1/2}$ and with inner product $\int f\bar{g}\, dp$. Let f, g denote elements of this space. The advantage of working with Hilbert space is that it is possible to recreate on function space many of the results obtained in finite dimensions using matrix algebra. Functions replace vectors, operators replace matrices, a few extra assumptions have to be made, but otherwise much is the same. We note that $\mathcal{L}_2[0,1]$ possesses the following two orthogonal subspaces. Let E_0 be a subspace such that $\int f\ dp = 0\ \forall\ f \in E_0$. Let E_1 be the space of constant functions. Note that $\langle f, g\rangle = 0,\ f \in E_0,\ g \in E_1$ and, as E_0 is closed, that $E = E_0 + E_1$.[6] Of course we will be particularly interested in density functions, that is, elements of $S_{\mathcal{L}} = \{f \in \mathcal{L}_2[0,1] : f \geq 0, \int f\, dp = 1\}$. We will, however, wish to consider dynamics on subspaces of E. If T is some closed subset of $[0,1]$, and in particular we will be looking at the interval $[\underline{p}, p^*]$, then let E_T be the elements of E with support on all of T, let E_{T1} be the elements of E constant on T, and let E_{T0} be the elements of E_0 for which $\int_T f\, dp = 0$.

We now can examine the dynamics for a continuum of prices. At any time the distribution of prices is described by $f \in S_{\mathcal{L}}$. Let $F(p) = \int_0^p f\, dr$. When not in equilibrium, firms adjust prices. In particular,

$$\dot{f} = Q(f)\pi \tag{4.10}$$

where Q is a linear operator. When f has support only on T (i.e. $f(p) > 0,\ iff\ \ p \in T$), Q possesses the following properties:

1. Q maps $E_T \to E_{T0}$.

[6] Lang(1993), Corollary V.1.8

2. $Qf = 0, \forall f \in E_{T1}$.

3. Q is positive definite elsewhere, i.e. $\langle Qf, f \rangle > 0,\ f \notin E_{T1}$.

4. Q is self-adjoint, that is, $Q = Q^*$ and hence, $\langle Qf, g \rangle = \langle f, Qg \rangle$.

5. Q is continuous.

6. At any point p, $\lim_{f(p) \to 0} \dot{f}(p) = 0$.

7. Q is compact.

The long list of conditions should not hide the generality of the dynamics specified. The substantive conditions are numbers 3 and 4. As noted in the previous section, positive definiteness ensures that population shares of the different strategies grow more or less in line with current payoffs. Property 1 ensures that $\int \dot{f}\, dp = 0$ and hence $\int f\, dp$ continues to be equal to one. Property 2 means that a mixed Nash equilibrium, that is when all strategies have the same return, is a fixed point for the dynamic. Property 6 implies that the dynamic is invariant on $S_{\mathcal{L}}$. More specifically, no strategy present in the initial distribution will disappear in finite time, nor will any new strategy be created. Thus we will want to look at cases where all prices are present in the initial distribution. This may seem somewhat restrictive, but it should be remembered that any distribution, including those where all firms charge the same price, can be approximated arbitrarily closely by a distribution with full support. Second, this formulation does not prevent the limit of the dynamic process being a state like the no-search outcome, where all sellers charge the same price.

Property 7 has two important consequences. First, unlike in finite dimensions, the spectrum of an operator on a space such as $\mathcal{L}_2$ may include elements which are not eigenvalues. However, the spectrum of a compact operator consists of its eigenvalues alone (together with zero if the space is infinite dimensional). Second, the Hilbert-Schmidt theorem (see for example, Hutson and Pym, 1980), states that the eigenfunctions of a compact, self-adjoint operator form an orthogonal basis for E. Since one eigenfunction of Q has eigenspace E_1, the others span E_0.

A concrete example of such an operator is given by the replicator dynamics, which in their $\mathcal{L}_2$ form are given by

$$\dot{f}(p) = Q(f)\pi(p) = f(p)[\pi(p) - \int_0^1 f(p)\pi(p)\,dp].$$

It is easily verified that the operator $Q(f)$ satisfies the above conditions.

We can state a preliminary result:

Lemma 9 *If Q is a compact linear operator and A is a continuous linear operator then QA is compact. (Lang, 1993; Theorem XVII.1.2).*

If A is a continuous operator, then it is positive definite when constrained to E_0 if $\langle Af, f\rangle > 0\ \forall\ f \in E_0$ and negative definite if $\langle Af, f\rangle < 0$.

Proposition 11 ***If A is positive (negative) definite on E_0 then QA has only positive (negative) eigenvalues when constrained to E_0.***

Proof: We have the eigenvalue equation $QAf = \lambda f$ where $f \in E_0$. As the eigenfunctions of Q corresponding to its nonzero eigenvalues span E_0, the image of Q is E_0. Thus, we can find a $g \in E$ such that $f = Qg$ (Q is not invertible, but this simply means that g is not unique).

$$\begin{aligned} \lambda \langle f, g\rangle &= \langle QAf, g\rangle \\ \lambda \langle Qg, g\rangle &= \langle Af, Qg\rangle \\ &= \langle Af, f\rangle \end{aligned} \tag{4.11}$$

As Q is positive self-adjoint, $\langle Qg, g\rangle$ is a positive real number. The real part of $\langle Af, f\rangle$ is positive (negative), hence all eigenvalues λ for eigenfunctions in E_0 have real part positive (negative). Furthermore, from Lemma 9, all its spectrum is positive (negative). □

We can also write the firms' profits as $\pi = \Pi f$. That is, Π is an operator mapping a distribution of prices f into a distribution of profits π. It is possible to perform calculus on function spaces such as Hilbert spaces, (see Lang, 1993; Hutson and Pym, 1980). The derivative is itself a linear operator. In particular, the operator Π is differentiable at ϕ if there exists a linear operator Π' such that

$$\lim_{\|f\|\to 0} \|\Pi(\phi + f) - \Pi(\phi) - \Pi' f\| \,/\, \|f\| = 0.$$

Looking at the profit functions, (4.2) and (4.4), we can see that they are both continuously differentiable in F. In the fixed sample size model of Burdett and Judd, firms' profits are an affine function of the distribution of prices. Differentiation simply removes the "intercept" term leaving,

$$\Pi'_F f = -2p\mu(1-q)F < 0 \quad (a.e.). \tag{4.12}$$

(The F subscript, standing for fixed sample rule is to differentiate it from the profit function in the sequential case). For the sequential search model, the marginal profits evaluated at an equilibrium Φ, are

$$\Pi'_S f = -pF\mu \int_p^\infty \frac{g}{\Phi^2} dr < 0 \quad (a.e.) \tag{4.13}$$

Profit is decreasing in the number of competitors charging a lower price.

Proposition 12 *If Π', the linearisation of the profit function taken at an equilibrium distribution ϕ with support T, is a negative definite linear operator on E_{T0} then ϕ is asymptotically stable under the positive definite dynamics (4.10) on E_T. If it is positive definite, ϕ is asymptotically unstable on E_T.*

Proof: If $\dot{f} = Q\pi$, then the linearisation at ϕ is given by

$$\frac{d\dot{f}}{df} = \frac{dQ}{df}\pi(\phi) + Q(\phi)\Pi'$$

where $\pi(\phi) = \Pi\phi$. Now, $Q\pi(\phi) = 0$, so that $(dQ/df)\pi(\phi) = 0$. The theory of Hartman and Grobman that a non-linear differential equation is locally equivalent to its linear part at a hyperbolic fixed point is valid on Banach (and hence Hilbert) spaces (Palis and de Melo, 1982). Hence, we can approximate the non-linear equation (4.10) by

$$\dot{f} = Q(\phi)\Pi' f, \quad f \in E_{T0}. \tag{4.14}$$

By Proposition 11, this linearisation has only negative (positive) eigenvalues if Π' is negative (positive) definite. In either case the linearisation is hyperbolic (there are no eigenvalues with real part zero). The solution to the linear equation (4.14) is given by $\exp(tQ(\phi)\Pi')f_0$, where f_0 is an arbitrary initial distribution of prices. (The exponent of a linear operator tL is given by the polynomial $\sum_{k=0}^{\infty}(tL)^k/k!$. Since the normed vector space of bounded linear operators on $\mathcal{L}_2$ is a Banach space, it is complete and hence the limit of this convergent series is itself an operator on $\mathcal{L}_2$.) $Q(\phi)\Pi'$ is compact because Q is compact and Π' is linear (Lemma 9). By the spectral theorem for compact operators (Lang, 1993; Theorems XVII 3.4, 3.5), we can we can make a direct sum decomposition of E into the (generalised) eigenspaces of $Q(\phi)\Pi'$. Thus, just as in the finite dimensional case, all negative eigenvalues implies asymptotic stability, all positive asymptotic instability. □

Proposition 13 *Under the positive definite dynamics (4.10), a dispersed price equilibrium is unstable when consumers use a fixed sample size rule.*

Proof: Given an equilibrium density of prices ϕ with support T, we construct an arbitrary distribution of prices $f \neq \phi$,

$$f = \phi + z, \qquad z \in E_{T0}, \int_T z^2\, dp > 0.$$

Let $Z(p) = \int_0^p z\,dr$. Of course, $Z(0) = Z(1) = 0$. As f is arbitrary, if the quadratic form $\langle \Pi'(f - \phi), (f - \phi) \rangle$ is positive then Π' is positive definite on E_{T0}.

$$\begin{aligned} \langle \Pi'(f-\phi), (f-\phi) \rangle &= \int_T z \,.\Pi' z \; dp \\ &= \int_T \Pi' z \;\; dZ \end{aligned}$$

Taking (4.12), we have

$$\int_0^1 \Pi'_F z \;\; dZ = -2\mu(1-q) \int_T pZ\, dZ,$$

which by integration by parts gives

$$\mu(1-q) \int_T Z^2\, dp > 0.$$

Hence the linearisation (4.12) is positive definite and by Proposition 12 the equilibrium is unstable on E_T. □

Of course, even if an equilibrium is unstable only on a subset of the total state space, it is still described as unstable. In contrast, in the case of consumers using a reservation price rule, the results are not so clear cut. From (4.13), we obtain, again by integration by parts

$$\int_T \Pi'_S z \;\; dZ = \frac{1}{2}\mu \int_T Z^2 \left(-p\frac{g}{\Phi^2} + \int_p^\infty \frac{g}{\Phi^2} dr \right) dp \,.$$

The sign of this expression is ambiguous. However, it is more likely to be negative than positive in the following sense. The sign depends on the relative magnitude of the two functions

$$p\frac{g}{\Phi^2} \text{ and, } \int_p^\infty \frac{g}{\Phi^2} dr. \qquad (4.15)$$

If g/Φ^2 is decreasing, as it is likely to be, then its integral (the second expression above) will be smaller than the first. In any case, we can show, using an example from Rob (1985),

Proposition 14 *There exists at least one dispersed price equilibrium which is stable under the positive definite dynamics (4.10).*

Proof: In Appendix 1.

We can place an interpretation on these two results. Firstly, it is important to realise that one consequence of (4.12) and (4.13) being negative is that it is possible to construct deviations from equilibrium which raise the profits of all sellers (less a set of measure zero). Imagine an alternative distribution f which places greater weight on high prices: $z = f - \phi$ would be negative for low prices and positive nearer 1. Assume $f < \phi$ on the interval $(\underline{p}, a)$, $f(a) = \phi(a)$, and $f > \phi$ on the interval $(a, 1)$. Because, $F < \Phi$ except at the two end points of the distribution and because profits are decreasing in F, profits are higher everywhere. For example, in the model of Burdett and Judd, profits are

$$p[q + 2(1-q)(1-F)] \geq p[q + 2(1-q)(1-\Phi)],$$

as $F \leq \Phi$ on $(\underline{p}, 1)$. The change in profits is obviously equal to $p\ 2(1-q)(\Phi - F)$, and thus is increasing in price and the difference between the two cumulative distributions. Profits are unchanged at $p = 1$ as both F and Φ must be equal to one at this point. The greatest increase in profits will occur at $p = (\Phi - F)/(2(1-q)(f-\phi)) > b$, that is, at a point where the density of firms has increased. Just as for the discrete approximation, an increase in density of firms at certain prices leads to an increase in profits for all firms charging that price. Under positive definite dynamics, this will result in the deviation from equilibrium increasing in magnitude, and is obviously destabilising. There is nothing special about this example. Since all eigenvalues of Π' (when constrained to E_0) are positive, all deviations have a similarly destructive effect.

Under sequential search a similar deviation will raise profits for all sellers where $F < \Phi$. This is because as noted above $\Pi'_S f < 0$, and hence a decrease in the number of sellers charging lower prices will raise profits. The important difference is that the maximum change in profit may occur on the interval (a, b). Whether this is the case is determined by the sign of the expression of (4.15), which of course also determines the stability of the equilibrium. If indeed $d(\Pi'_S f)/dp > 0$, the biggest increase in profits from the deviation from equilibrium falls to the firms that have kept their prices low. In this case, the number of firms charging low prices will grow and the distribution of prices will return to equilibrium.

Having considered dynamics for sellers, we now examine what happens when buyers and sellers change behaviour simultaneously. However, we were only able to obtain results for the fixed sample size case. We first obtain a preliminary result.

Proposition 15 *The no-search outcome is a strict Nash equilibrium for both buyers and sellers, when buyers use a fixed sample size search rule.*

Proof: If all firms charge their monopoly price p^* then any consumers searching more than once will undertake costly search which cannot lead to a better price. These consumers do strictly worse than consumers who sample only one seller. If all consumers make only one search then profits for a firm charging p^* are strictly higher than for any firm charging a price $p < p^*$ as all firms have the same demand. □

Proposition 16 *The no-search outcome is asymptotically stable under any positive definite dynamic.*

Proof: We consider a positive definite dynamic on $L = S_{\mathcal{L}} \times S_n$ and let $z = (f, q)$ be a typical element of L, where f is the distribution of firms' prices and q the vector of proportions of consumers choosing each of their n possible strategies (i.e. to take from 1 to n price quotations). We have $\dot{f} = Q(f)\pi_f$ and $\dot{q} = Q(q)\pi_q$. Let δ be the no-search equilibrium, that is, the state where $F(p) = 0$, $p < p^*$ and $g_1 = 1$. Let $\pi_f(p, \delta)$ be the profit of a firm charging a price p at the state δ and $\pi_f(\delta)$ is the function that describes the (hypothetical) profits for all other prices given that in fact all firms charge p^*, then define $\alpha = \pi_f(p^*, \delta) - \pi_f(\delta)$. This is a function which is zero at p^* and positive elsewhere. And in a similar way for consumers, let $\beta = \pi_q(1,\delta) - \pi_q(\delta)$, i.e. the difference at δ between the payoff for sampling once and the payoff for all other strategies. If we define $V = \langle \alpha, f \rangle + \langle \beta, q \rangle \geq 0$, then $\dot{V} = \langle Q(f)\pi_f, \alpha \rangle + \langle Q(q)\pi_q, \beta \rangle$. Since V has a unique minimum on L at δ, by the theory of Liapunov functions, δ is asymptotically stable if $\dot{V} < 0$ in the neighbourhood of δ. We choose an ϵ such that for all z in some neighbourhood of δ, $\pi_f = \pi_f(\delta) + \xi_f$ and $\pi_q = \pi_q(\delta) + \xi_q$ with $\sup \mid \xi_f \mid < \epsilon$, and $\mid \dot{f}(p) \mid < \epsilon$ for $p < p^*$ and $\sup \mid \xi_q \mid < \epsilon$, and $\mid \dot{q}_i \mid < \epsilon$ for $i > 1$ by Conditions 5 and 6 of the definition of a positive definite dynamic. Then, for example,

$$\dot{f} = Q\pi_f(\delta) + Q\xi_f.$$

Now, at any point with $p < p^*$ then the first term of the above is of order ϵ, the second is of the order ϵ^2. Thus, in the neighbourhood of δ we can approximate $\dot{V}$ by

$$\langle Q\left(\pi_f(p^*,\delta) - \alpha\right), \alpha \rangle + \langle Q\left(\pi_q(1,\delta) - \beta\right), \beta \rangle = -\langle Q\alpha, \alpha \rangle - \langle Q\beta, \beta \rangle < 0.$$

That is, δ is asymptotically stable. □

Note that in this proof neither Property 2, that Q is self-adjoint, or Property 7, that Q is compact, is used. This result therefore holds under the weakest possible conditions on dynamics.

It is important to remember that the behaviour of other buyers does not enter directly into the decision of any individual consumer. The payoff to each consumer is determined by his decision on how much and how to search and the current distribution of prices, not by the search behaviour of other consumers. Of course, there may be an indirect effect. For example, if average consumer search is very intense, then there will be a downward pressure on prices, which will in turn change consumers' expected payoffs. But it remains the case that the dynamic stability of a dispersed equilibrium is largely determined (just as in the simple examples of Section 2) by whether there is stability in the adjustment process for sellers. We use this fact to prove the following proposition, where we use a discrete approximation to the continuous distribution. Obviously though, this approximation can be arbitrarily close to the original.

Proposition 17 *When the changes of both buyer and seller behaviour are described by positive definite dynamics, the discrete approximation to the mixed equilibrium of the model of Burdett and Judd (1983) is unstable.*

Proof: In Appendix 2.

4.6 Discussion

We have shown that dispersed price equilibria are unstable when consumers use a fixed sample size rule. In contrast, when a sequential search rule is used, the adjustment process for sellers may be stable. We are not able to present results on the stability of a dispersed price equilibrium when search is sequential and buyers and sellers change their behaviour simultaneously. It is possible to conjecture that the dynamics behave in the same way in higher dimensions as they do in the simple examples of Section 2, where overall stability is determined by the stability of the sellers' adjustment process. However, there are further doubts as to whether adaptive learning by consumers would lead to such an equilibrium.

It is not known the way that real consumers search, but there are reasons to believe that this search does not take the form of a fixed reservation price rule. This is despite the fact that this has become the dominant paradigm in economic theory. Firstly, as Morgan and Manning (1985) show, the optimal search rule in many cases will take the form of a mixture between a fixed sample size and a sequential rule. That is, the searcher immediately obtains several quotations but then may take more if the offers received are unsatisfactory. Harrison and Morgan (1990) find that behaviour under experimental conditions fits this pattern. Second, as Telser (1973), Rothschild (1974) and Gastwirth (1976) all find, the optimality of a reservation price rule is not robust to the introduction of imperfect information. If one fails to calculate the reservation price correctly, it is possible to search for an arbitrarily long time without success. Or as Telser puts it "if the searcher is ignorant of the distribution, then acceptance of the first choice drawn at random from the distribution confers a lower average cost than more sophisticated procedures for a wide range of distributions" (1973, p45). Lastly, if sequential search is "noisy", in the sense of Burdett and Judd (1983), then it is similar to a fixed sample size rule.

Thus in various senses a fixed reservation price rule does not seem robust. It is therefore unlikely that it would form the endpoint of an adaptive process. A direction for further research would be to see whether the evolutionary approach can pick out simple and robust search rules.

It might be worth pointing out that the instability result arises from the fundamental weakness of dispersed price equilibria, which could also be exposed by an equilibrium approach if one desired. For example, the dispersed price equilibria in both the fixed sample size and the sequential search models are in a sense not strategically stable in that they can be undermined by deviations by coalitions of firms. As we have seen if a positive mass, no matter how small that mass, of firms raised their prices simultaneously, they would see their profits rise. In a world of perfect information, the models specified, would give rise to Bertrand competition. In such circumstances, only a coalition that comprised all sellers could be successful.

How then do we explain what seems to be an empirical fact, that prices for identical goods do vary? One possibility is "disequilibrium price dispersion". First, although a dispersed price equilibrium is unstable and the no-search outcome is stable, these are both local results. Therefore, it does not follow, although it is quite possible, that all dynamic adjustment paths that diverge from the dispersed price equilibrium must arrive at the no-search outcome That is, the learning process may never settle down to a single price or search configuration. This is something for further research.

Appendix 1

Proof of Proposition 14. Rob (1985) gives some examples of dispersed price equilibria under sequential search. We take his example 3 (1985, p501), where the equilibrium distribution of prices is uniform. The proof has two stages. First, to show that the equilibrium is negative definite on its support, T. Second, to show that, if the system is close to the equilibrium but with full support, it approaches the equilibrium. The equilibrium cumulative distribution function is $\Phi(p) = 0$, $p \leq 1$; $\Phi(p) = p - 1$, $1 \leq p \leq 2$; $\Phi(p) = 1$, $p \geq 2$. Of course, here $p^* = 2$ not 1 as we assumed above, but clearly this is not important. From the distribution of search costs that Rob gives we can use (4.5) to calculate $g(r)$:

$$g(r) = \begin{cases} \dfrac{r-1}{r^2 \log 2} & 1 \leq r \leq 2 \\ (r-1)\dfrac{9-(r-1)^2}{32 \log 2} & 2 \leq r \leq 4 \end{cases}$$

and outside these intervals $g(r) = 0$. We have

$$p\frac{g(p)}{\Phi^2(p)} = \frac{1}{p(p-1)\log 2},$$

and, for $p \leq 2$,

$$\int_p^\infty \frac{g}{\Phi^2} dr = \int_p^2 \frac{1}{r^2(r-1)\log 2} dr + \int_2^4 \frac{9-(r-1)^2}{32(r-1)\log 2} dr$$

Straightforward, if lengthy, calculation reveals that

$$\int_p^\infty \frac{g}{\Phi^2} dr - p\frac{g(p)}{\Phi^2(p)} < 0, \quad 1 \leq p < 4$$

and therefore the equilibrium is negative definite.

We now show convergence to the equilibrium from distributions not in E_T. The equilibrium profit function is sketched out in Rob (1985, p502). Of course, we have

$$\pi(p < \underline{p}) < \pi(\underline{p} \leq p \leq p^*) = \pi^* > \pi(p > p^*).$$

Define $\lambda = \pi(p^*) - \pi(\phi)$. This is a function, zero on T and positive elsewhere. We use the Liapunov function $\Lambda = \langle f, \lambda \rangle \geq 0$, with equality when f has support on T alone. This gives us $\dot{\Lambda} = \langle Q\pi, \lambda \rangle$. For $\epsilon > 0$, we can choose a neighbourhood of distributions with full support close enough to ϕ such that we have both $\pi(f) = \pi(\phi) + \xi$, with sup $|\xi| < \epsilon$, and $\left|\dot{f}(p)\right| < \epsilon$ for $p \notin T$. Then in a similar way to the proof of Proposition 16, we have

$$\dot{\Lambda} = \langle Q\pi(\phi), \lambda \rangle + \langle Q\xi, \lambda \rangle \approx \langle Q(\pi(p^*) - \lambda), \lambda \rangle = -\langle Q\lambda, \lambda \rangle < 0.$$

Thus the share in the population of prices not represented in the equilibrium distribution falls away to zero. □

Appendix 2

Proof of Proposition 17. We assume that firms can choose from only a finite number of prices. These we label $p = (p_1, p_2, ..., p_n)$, where the elements are given in increasing order $p_i > p_{i-1}$. And let the proportion of the population of firms choosing each of those prices be $x = (x_1, x_2, ..., x_n)$. Define $F_i = \sum_{j=1}^{i} x_j$. The expected payoffs (that is, the expected price plus cost of searching) for the proportion q of consumers who search once, and the $1 - q$ who search twice are respectively

$$\pi_q = -p \cdot x \text{ and } \pi_{1-q} = -c - \sum p_i(x_i^2 + 2x(1 - F_i)). \tag{4.16}$$

The probability that a consumer who makes two searches finds p_i as the lowest price is $x_i^2 + 2x_i(1 - F_i)$. These consumers we assume to be equally divided between the x_i firms charging that price. Hence if $1 - q$ consumers search twice and q once, profits are given by

$$\pi_i = p_i\mu\left[q + (1 - q)(x_i + 2(1 - F_i))\right].$$

Let

$$\dot{x} = Q(x)\pi_x(x,q) \text{ and } \dot{q} = Q(q)\pi_q(x,q)$$

where both $Q(x)$ and $Q(q)$ satisfy the Properties outlined in Section 5. Let z be the combined vector $(x,q) \in S = S_n \times S_2$. We have

$$\dot{z} = Q(z)\pi(z) = \begin{pmatrix} Q(x) & 0 \\ 0 & Q(q) \end{pmatrix} \begin{pmatrix} \pi_x \\ \pi_q \end{pmatrix}.$$

$Q(z)$ is symmetric positive definite as before. The linear approximation of this system of equations at an equilibrium point z^* is given by $Q(z^*)\Pi'$. Given the consumers' payoffs (4.16), $d\pi_q/dq$ is a matrix of zeros. Thus the matrix $Q(z^*)\Pi'$ has the form,

$$\begin{pmatrix} Q(x)\dfrac{d\pi_x}{dx} & Q(x)\dfrac{d\pi_x}{dq} \\ Q(q)\dfrac{d\pi_q}{dx} & 0 \end{pmatrix}$$

Stability in the single population case depended on whether the matrix Π' was positive or negative definite on E_0. A $n \times n$ matrix A is positive definite on E_0 iff the $(n-1)\times(n-1)$ matrix C is positive definite (van Damme, 1991), where C is defined by

$$c_{ij} = a_{ij} + a_{nn} - a_{in} - a_{nj}, \quad 1 \leq i,j \leq n-1. \tag{4.17}$$

In the two population case, if there are n strategies available to firms and m to consumers, then we want to know the sign of the eigenvalues of $Q(z^*)\Pi'$ when constrained to $E_{00} = \{x \in R^{n+m} : \sum_1^n x_i = 0, \sum_{n+1}^{n+m} x_i = 0\}$. To do this, we repeat the procedure outlined in (4.17) for each of the four submatrices of $Q(z^*)\Pi'$. We first look at $d\pi_x/dx$. In this case the matrix $(C+C^T)/2$ is

$$\mu(1-q)\begin{pmatrix} p_n - p_1 & p_n - p_2 & \cdots & p_n - p_{n-1} \\ p_n - p_2 & p_n - p_2 & \cdots & p_n - p_{n-1} \\ \vdots & \ddots & \ddots & \vdots \\ p_n - p_{n-1} & \cdots & p_n - p_{n-1} & p_n - p_{n-1} \end{pmatrix} \tag{4.18}$$

If we subtract the nth column from the $n-1$th, and the $n-1$th from the $n-2$th and so on, a upper diagonal matrix is left, with a strictly positive diagonal. Thus this matrix and hence C is positive definite. Hence by Proposition 11, $Q(x)\,d\pi_x/dx$ has only positive eigenvalues, when constrained to E_0. So necessarily it has a positive trace. Since the bottom right submatrix of $Q(z^*)\Pi'$ is all zero, its trace is zero both before and after the process (4.17). Thus, the matrix as a whole when constrained to E_{00} has a strictly positive trace and hence at least one positive eigenvalue. □

Bibliography

Learning and Evolution

Bacharach, M., Shin, H., M. Williams (1992). "Sophisticated bounded agents play the repeated dilemma," mimeo, Oxford University.

Banerjee, A., Weibull, J. (1995). "Evolutionary selection and rational behavior," in A. Kirman and M. Salmon eds. *Rationality and Learning in Economics*, Oxford: Blackwell.

Binmore, K., Samuelson, L. (1994). "Muddling through: noisy equilibrium selection," Working Paper, University College, London; University of Wisconsin.

Binmore, K., Samuelson, L., Vaughan, R. (1993). "Musical chairs: modelling noisy evolution," Working Paper, University College, London; University of Wisconsin.

Binmore, K., Swierzbinski, J., Proulx, C. (1994). "Does minimax work? An experimental study," Working Paper. University College, London.

Blume, L., Easley, D. (1992). " Evolution and market behavior," *J. Econ. Theory*, **58**, 9-40.

Börgers, T., Sarin, R. (1993). "Learning through reinforcement and replicator dynamics," Working Paper, University College, London.

Boylan, R.T. (1992). "Laws of large numbers for dynamical systems with randomly matched individuals," *J. Econ. Theory*, **57**, 473-506.

Brams, S. (1993). "Game theory and literature," *Games Econ. Behav.*, **6**, 32-54.

Brown, G.W. (1951). "Iterative solutions of fictitious play," in *Activity Analysis of Production and Allocation* (T.C. Koopmans, Ed.). New York: Wiley.

Brown, J., Rosenthal, R. (1990). "Testing the minimax hypothesis: a re-examination of O'Neill's game experiment," *Econometrica*, **58**, 1065-1081.

Cabrales, A., Sobel, J. (1992). "On the limit points of discrete selection dynamics," *J. Econ. Theory*, **57**, 407-419.

Canning, D. (1992). "Average behaviour in learning models," *J. Econ. Theory* , **57**, 442-472.

Crawford, V. (1985). "Learning behaviour and mixed-strategy Nash equilibria," *J. Econ. Behav. Organ.*, **6**, 69-78.

Crawford, V. (1989). "Learning and mixed strategy equilibria in evolutionary games," *J. Theor. Biol.*, **140**, 537-550.

van Damme, E. (1991). *The Perfection and Stability of Nash Equilibrium.* Second Edition. Berlin: Springer Verlag.

Dekel, E., S. Scotchmer (1992). "On the evolution of optimizing behavior," *J. Econ. Theory*, **57**, 392-406.

Elster, J. (1978). *Ulysses and the Sirens.* Cambridge: Cambridge University Press.

Elster, J. (1989). *Solomonic Judgements.* Cambridge: Cambridge University Press.

Friedman, D. (1991). "Evolutionary games in economics," *Econometrica*, **59**, 637-666.

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

Fudenberg, D., Kreps D. (1993). "Learning mixed equilibria," *Games Econ. Behav.*, **5**, 320-367.

Fudenberg, D., Levine D. (1993). "Steady state learning and Nash equilibrium," *Econometrica*, **61**, 547-573.

Gale, J., Binmore, K., Samuleson, L. (1995). "Learning to be imperfect: the ultimatum game," *Games Econ. Behav.*, **8**, 56-90.

Harley, C. (1981). "Learning the evolutionary stable strategy," *J. Theor. Biol.*, **89**, 611-633.

Harsanyi, J. (1973). "Games with randomly disturbed payoffs," *Int. J. Game Theory*, **2**, 1-23.

Hines, W.G.S. (1980). "Three characterisations of population strategy stability," *J. Appl. Prob.*, **17**, 333-340.

Hofbauer, J. (1994). "Stability for the best response dynamics," mimeo, Vienna.

Hofbauer, J. (1995). "Imitation dynamics for games," mimeo, Collegium, Budapest.

Hofbauer, J., Sigmund K. (1988). *The Theory of Evolution and Dynamical Systems.* Cambridge: Cambridge University Press.

Hofbauer, J., Sigmund, K. (1990). "Adaptive dynamics and evolutionary stability." *App. Math. Lett.*, **3**, 75-79.

Houston, A., Sumida B. (1987). "Learning, matching and frequency dependence," *J. Theor. Biol.*, **126**, 289-308.

Jordan, J. (1993). "Three problems in learning mixed strategy equilibria," *Games Econ. Behav.*, **5**, 368-386.

Kalai, E., Lehrer E. (1993). "Rational learning leads to Nash equilibrium," *Econometrica*, **61**, 1019-1045.

Kandori, M., Mailath G. and Rob R. (1993). "Learning, mutation and long run equilibrium in games," *Econometrica*, **61**, 29-56.

Kirman, A.P. (1993). "Ants, rationality and recruitment," *Quart. J. Econ.*, **108**, 137-156.

Lucas, R. (1987). "Adaptive behavior and economic theory," in *Rational Choice* (eds. Hogarth, R.M. and Reder, M.W.) Chicago: University of Chicago.

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

Maynard Smith, J. et al. (1984). "Symposium: Game theory and the evolution of behaviour," *Behav. Brain Sci.*, **7**, 95-125.

Milgrom, P., Roberts J. (1991). "Adaptive and sophisticated learning in normal form games," *Games Econ. Behav.*, **3**, 82-100.

Miyasawa, K. (1961). "On the convergence of the learning process in a 2×2 non-zero sum game," Mimeo, Princeton University.

Monderer, D., Shapley L. (1993). "Fictitious play property for games with identical interests," forthcoming *J. Econ. Theory.*

Mookherjee, D., Sopher, B. (1994). "An experimental study of learning and decision costs in constant-sum games," working paper.

Nachbar, J.H. (1990). "Evolutionary selection dynamics in games: convergence and limit properties," *Int. J. Game Theory*, **19**, 59-89.

Nöldeke, G., Samuelson, L. (1993). 'An evolutionary analysis of backward and forward induction.' *Games Econ. Behav.*, **5**, 425-454.

Robinson, J. (1951). "An Iterative Method of Solving a Game," *Ann. Math.*, **54**, 296-301.

Roth, A.E., Erev, I. (1995). "Learning in extensive-form games: experimental data and simple dynamic models in the intermediate term," *Games Econ. Behav.*, **8**, 164-212.

Rubinstein, A. (1991). "Comments on the interpretation of game theory," *Econometrica*, **59**, 909-924.

Rubinstein, A., Tversky, A. (1993). "Naive behavior in zero sum games," Working Paper. Tel Aviv University.

Samuelson, L. (1994). "Stochastic stability in games with alternative best replies," *J. Econ. Theory*, **64**, 35-65.

Samuelson, L., Zhang J. (1992). "Evolutionary stability in asymmetric games," *J. Econ. Theory*, **57**, 363-392.

Schlag, K. (1994). "Why imitate and if so how?" Working Paper, University of Bonn.

Sen, A. (1985). "Rationality and uncertainty," *Theory and Decision*, **18**, 109-127.

Shapley, L. (1964). "Some topics in two person games," in eds. M. Dresher et al., *Advances in Game Theory*. Princeton: Princeton University Press.

Stahl, D.O. (1993). "Evolution of smart-n players," *Games Econ. Behav.*, **5**, 604-617.

Stahl, D.O., Wilson, P.W. (1994). "Experimental evidence on players' models of other players," *Journal of Economic Behavior and Organization*, **25**, 309-327.

Sugden, R. (1989). "Spontaneous order," *J. Econ. Perspectives*, **3.4**, 85-97.

Thaler, R. (1988). "Anomalies: the ultimatum game," *J. Econ. Perspectives*, **2**, 185-206.

Winter, S.G. (1987). 'Comments on Arrow and Lucas.' in *Rational Choice* (eds. Hogarth, R.M. and Reder, M.W.) Chicago: University of Chicago.

Woodford, M. (1990). "Learning to believe in sunspots," *Econometrica*, **58**, 277-307.

Young, H.P. (1993). "Conventional equilibria," *Econometrica*, **61**, 57-84.

Zeeman, E.C. (1981). "Dynamics of the evolution of animal conflicts," *J. Theor. Biol.*, **89**, 249-270.

Price Dispersion

Benabou, R. (1993). "Search market equilibrium, bilateral heterogeneity and repeat purchases," *J. Econ. Theory*, **60**, 140-158.

Bester, H. (1988). "Bargaining, search costs and equilibrium price distributions," *Rev. Econ. Stud.*, **55**, 201-214.

Burdett, K., Judd, K. (1983). "Equilibrium price dispersion," *Econometrica*, **51**, 955-969.

Carlson, J., McAfee, P. (1983). "Discrete equilibrium price dispersion," *J. Polit. Econ.*, **91**, 480-493.

Diamond, P. (1971). "A model of price adjustment," *J. Econ. Theory*, **3**, 156-168.

Fershtman, C., Fishman, A. (1992). "Price cycles and booms: dynamic search equilibrium," *Amer. Econ. Rev*, **82**, 1221-1233.

Gastwirth, J.L. (1976). "On probabilistic models of consumer search for information," *Quart. J. Econ.*, **90**, 38-50.

Harrison, G.W., Morgan, P. (1990). "Search intensity in experiments," *Econ. J.*, **100**, 478-486.

Hutson, V., Pym, J.S. (1980). *Applications of Functional Analysis and Operator Theory.* London: Academic Press.

Lang, S. (1993). *Real and Functional Analysis.* Third Edition. New York: Springer Verlag.

Morgan, P., Manning, R. (1985). "Optimal search," *Econometrica*, **53**, 923-944.

Palis, J., de Melo, W. (1982). *The Geometrical Theory of Dynamical Systems.* New York: Springer Verlag.

Rob, R. (1985). "Equilibrium price distributions," *Rev. Econ. Stud.*, **52**, 487-504.

Rothschild, M. (1973). "Models of market organization with imperfect information: a survey," *J. Polit. Econ.*,**81**, 1283-1308.

Rothschild, M. (1974). "Searching for the lowest price when the distribution of prices is unknown," *J. Polit. Econ.*, **82**, 689-711.

Salop, S., Stiglitz, J. (1977). "Bargains and rip-offs: a model of monopolistically competitive price dispersion," *Rev. Econ. Stud.*, **44**, 493-510.

Salop, S., Stiglitz, J. (1982). "The theory of sales," *Amer. Econ. Rev.*, **72**, 1121-1130.

Stigler, G.J. (1961). "The economics of information," *J. Polit. Econ.*,**69**, 213-225.

Talmain, G. (1992). "Search fròm an unknown distribution : an explicit solution," *J. Econ. Theory*, **57**, 141-157.

Telser, L.G. (1973). "Searching for the lowest price," *Amer. Econ. Rev., Papers and Proceedings*, **63**, 41-49.

Varian, H.R. (1980). "A model of sales," *Amer. Econ. Rev*, **70**, 651-659.

Wilde, L.L. (1992). "Comparison shopping as a simultaneous move game," *Econ. J.*, **102**, 562-569.

Wilde, L.L., Schwartz, A. (1979). "Equilibrium comparison shopping," *Rev. Econ. Stud.*, **46**, 543-553.