

On economic applications of evolutionary game theory

Daniel Friedman*

Department of Economics, University of California at Santa Cruz,
212 Social Sciences I, Santa Cruz, CA 95064, USA

Abstract. Evolutionary games have considerable unrealized potential for modeling substantive economic issues. They promise richer predictions than orthodox game models but often require more extensive specifications. This paper exposts the specification of evolutionary game models and classifies the possible asymptotic behavior for one and two dimensional models.

Key words: Evolutionary games – Adjustment dynamics – ESS – Evolutionary equilibrium

JEL-classification: B25; C62; C72; C73

1 Introduction

Evolutionary games have seized a large and increasing share of the game theory literature in recent years. But economic applications of evolutionary game theory remain few and isolated, while a dominant share of the applied economics literature relies on orthodox game theory. A casual observer might surmise from these facts that the new evolutionary approach is intractable or not widely applicable, or that its implications are essentially the same as those of the orthodox approach.

* My thanks to the fellows and staff of CentER at Tilburg University for providing the stimulating environment in which this paper was originally written, and to K.C. Fung for his help and encouragement in beginning the project. Vince Crawford, Jorgen Weibull, three anonymous referees, and audiences at Tilburg, Arizona, UCSC and Florence provided useful comments on earlier versions. I retain sole responsibility for any remaining errors and idiosyncrasies.

In this paper I argue to the contrary that evolutionary game theory is quite tractable, that it has a wide range of applicability which differs somewhat from that of orthodox game theory, and that the evolutionary approach promises distinctive insights and implications in substantive economic applications. My main purpose is to present basic analytical tools of evolutionary game theory so that the general reader can begin to use them. In passing I mention some of the literature but do not survey it systematically. I present a classification of low dimensional evolutionary games, but it requires only very minor extensions of existing results. In short this paper is more of a user's manual than a survey or a research article.

The rest of this introduction addresses preliminary questions. Section 2 offers detailed but non-technical remarks on the specification of evolutionary game models. Section 3 classifies the asymptotic behavior in the one and two dimensional models which are likely to be the most useful for applications. Brief concluding remarks appear in Section 4. Appendix A collects some analytical details; it should be accessible to most recent PhD economists. Appendix B points out some of the relevant literature on evolutionary dynamics.

A. What is an evolutionary game?

By an evolutionary game, I mean any formal model of strategic interaction over time in which (a) higher payoff strategies tend over time to displace lower payoff strategies, but (b) there is some inertia, and (c) players do not systematically attempt to influence other players' future actions. Condition (a) is a version of the basic "survival of the fittest" maxim, while (b) distinguishes evolutionary change from revolutionary change and means that aggregate behavior does not change too abruptly. Underlying reasons include various sorts of adjustment costs, or informational imperfections, or perhaps bounded rationality. Condition (c), referred to below as the game against Nature (GAN) condition, distinguishes evolutionary games from repeated games with trigger strategy threats, etc. It can be justified by appeals to large numbers of players, or to players' myopia. I shall argue in section 2 that in practice these conditions are even less restrictive than they seem.

Some authors adopt a narrower definition of evolutionary games by imposing various additional assumptions such as large populations, bounded rationality, random matching, and observable strategies. I shall now argue that such restrictions are unnecessary.

B. Where might evolutionary games be applicable?

Biologists and mathematicians originally developed evolutionary games to address substantive questions in evolutionary biology (Maynard Smith and Price, 1973; Maynard Smith, 1982; Taylor and Jonker, 1978). Strategic interaction over time arises because the fitness of a biological trait or behavior often depends not just on the background environment but also on the prevalence in the current population of the behavior and the prevalence

of alternative behaviors. For example, an animal's aggressive behavior in acquiring resources has higher fitness when it is rare and hence less likely to encounter counteraggression.

Economic theorists have found evolutionary games to be very useful for investigating the foundations of game-theoretic solution concepts, especially Nash equilibrium (NE) and selection among multiple NE (e.g. Binmore, 1991; Fudenberg and Kreps, 1993; Samuelson and Zhang, 1992). Biological applications routinely and naturally assume myopia, random matching and genetic transmission of behaviors. Much of the theoretical economics literature retains these assumptions.

In economic applications, genetic transmission may sometimes underlie important constraints but generally has too much inertia to produce to interesting dynamics over relevant timescales. The human gene pool is essentially constant over the economist's long run of a decade or so. But genetic transmission is not the only evolutionary mechanism; indeed, there are several alternative mechanisms that produce behavioral change with an appropriate degree of inertia. Agents with different types of economic behavior can enter and exit; market outcomes (or outcomes from other economic institutions) can redistribute resources among different types of agents; and individual agents can change behavior as they accumulate experience. Later I will cite examples of evolutionary dynamics based on entry and exit, on resource redistribution, and on learning or imitation.

In short, evolutionary game models are worth considering whenever any of these mechanisms changes behavior over time and agents interact strategically, in the sense that the outcome for each agent depends on others' behavior as well as her own. If agents do not directly and systematically attempt to influence others' future behavior then evolutionary games provide a natural approach. It is especially valuable when the relevance of static equilibria are unclear, e.g., when there are several NE. Potential fields of application include industrial organization, law, economic development, international trade and policy analysis. The reader can probably think of other fields as well.

C. So why aren't economic applications common already?

In my opinion the reasons are primarily historical. Biologists originally borrowed the basic idea of evolution from economists and social philosophers, especially Malthus (1798). The biologists closely adapted evolutionary models, eventually including evolutionary games, to fit biological applications. Only very recently have economic theorists begun to readapt the models for economic applications. In particular,

(a) Much of the published evolutionary games literature centers on the biologists' static equilibrium concept called ESS, for evolutionary stable state (or strategy). I argue below that ESS can be viewed as a Nash equilibrium refinement, but the refinement is not central to economic applications.

(b) Replicator dynamics, defined below in section 2C, are often assumed in explicit dynamic models. I shall argue that replicator dynamics are

natural for the genetic mechanism but are less natural for economically relevant mechanisms such as exit or learning.

(c) A casual reading of most existing literature can easily give the impression that evolutionary games apply only to a single strategically symmetric population of agents, that the population must be infinitely large, and that interactions must consist of pairwise matches of agents. These conditions make far more sense in biological applications than in economic applications. I shall argue that none of these conditions are necessary for evolutionary game models.

Thus the theoretical literature does not yet provide clear guidance on the proper application of evolutionary games to substantive economic questions. But the potential is great. Evolutionary games offer a different perspective on information conditions than orthodox static complete information (NE) and incomplete information (Bayesian NE) models. Evolutionary games also offer a richer set of empirical predictions such as built-in selection criteria among multiple NE, and predictions of relatively rapid or slow convergence (or non-convergence) to NE. There is a cost, however. In addition to the orthodox specification of a payoff function (or game tree), the evolutionary game models usually require specification of initial conditions (or a slice of history) and sometimes also require detailed specification of adjustment (e.g. learning) dynamics.

In sum, economists must re-adapt evolutionary theory to economics before evolutionary game models can become routine and widespread. We must shed some of the biological adaptations and develop some new adaptations for economics.

D. Some relevant literature

The presentation here is based on Friedman (1991) and, despite the patient efforts of three generous referees, it remains somewhat opinionated. The reader will benefit from other perspectives.

Weibull (1992) is a good place to start. It is a self-contained introduction to evolutionary game theory, written at the advanced undergraduate or graduate student level; Weibull (1995) is a more complete and more advanced treatment. Other advanced texts begin with Bomze and Pötscher (1989) and Cressman (1992), and now include Vega-Redondo (1996), Samuelson (1997), and Fudenberg and Levine (1997). Chapter-length theoretical summaries include Eichberger (1992), Robson (1992) and chapter 9 of van Damme (1991). Hofbauer and Sigmund (1988) is a classic book for a general scientific audience. Several game theory textbooks have some material on evolutionary games; Binmore (1992) devotes chapter 9 to the topic. Influential early theoretical papers by economists and game theorists include Binmore (1987) and Fudenberg and Kreps (1988).

Several journals besides *JEE* carry relevant theoretical articles. Mailath (1992) introduces a special issue of the *Journal of Economic Theory* devoted to evolutionary game theory; and Selten (1991b) is the lead article in a special issue of *Games and Economic Behavior*. Crawford (1993) introduces a special double issue of *Games and Economic Behavior* on “adaptive

dynamics”, certainly included in evolutionary game theory on the current broad definition. Dosi et al. (1996) is a long non-technical survey of learning and evolution. Some readers might also wish to browse the special issue of *Journal of Mathematical Biology* introduced by Diekmann et al. (1996).

2 Model elements

An evolutionary game models strategic interaction over time in terms of one or more *populations* of players, a *state space* of strategies, a *stage game* in normal or extensive form, and a *dynamic* adjustment process. At each moment in time the stage game specifies the payoff (or expected payoff) to each strategy, and the state specifies the current distribution of strategies employed in each population. Over time, higher payoff strategies tend to displace lower payoff strategies, as specified by the dynamic. As the state changes over time, the payoffs to different strategies also change, so the time path can be rather complicated. The analysis focuses mainly on steady states, especially asymptotically stable steady states, and on the closely related Nash equilibria of the stage game.

This section will discuss each of these model elements at greater length, with particular attention to their specification in economic applications.

A. Populations and states

The number $K \geq 1$ of *populations* represents the number of economically distinct roles in the model, where each population $k = 1, \dots, K$ has $N_k \geq 2$ alternative strategies available to its members. The *state* then specifies the fractions of each population currently employing each alternative strategy. For example, consider a two country international trade model where firms in each country choose between two possible modes of internal organization. Here $K = 2$ because, even with identical technologies, firms in the two countries are economically distinct in that they face somewhat different goods and factor markets. Also $N_1 = N_2 = 2$ because each firm has two alternative strategies (organizational modes). The state of each population is summarized by a point p between 0 and 1 representing the fraction of firms currently using the first mode, with $1 - p$ of the firms using the second mode. The state space for each country’s firms is $S^1 = S^2 = [0, 1]$ and the overall state space $S = S^1 \times S^2$ is the square $[0, 1]^2$. Other natural $K = 2$ models might have co-evolving populations representing buyers and sellers, or representing gun producers and butter producers.

In any evolutionary game model, the number of populations can always be reduced to $K = 1$ by symmetrizing the game or by expanding $N_1 = N$ to cover several role-contingencies (Selten, 1980; Cripps, 1991; Bomze and Pötscher, 1989, Chapters 2.5 and 3.5). This device is useful for theoretical purposes but can short-circuit an application. In the gun-butter application, the role-contingent reduced model puts each entrepreneur sometimes in the gun sector and sometimes in the butter sector, and assumes that his effectiveness as a gun producer depends in part on his effectiveness as a butter

producer. An unreduced $K = 2$ model with separate evolutionary paths for each sector seems more sensible.¹

The simplest nontrivial state space is one dimensional, and arises only if $K = 1$ and $N = 2$ so $S = [0, 1]$, as in the isolated single country part of the international trade model. There are two types of two-dimensional state spaces: the square $S = [0, 1]^2$ arising from $K = 2$ and $N_1 = N_2 = 2$ as in the trade model, and the triangle. The triangle arises from a single population whose players have three alternative strategies ($K = 1$ and $N_1 = 3$). For example, one might model firms in an isolated industry as choosing either a high, medium or low advertising budget. The population fractions are non-negative and sum to 1 so the state space is the two dimensional triangle $\mathcal{T} = \{(p, q, 1 - p - q) : 0 \leq p, q \text{ and } p + q \leq 1\}$.

There are three types of three-dimensional state spaces: the cube $[0, 1]^3$ for $K = 3, N_1 = N_2 = N_3 = 2$ (e.g. a three country trade model); the prism $\mathcal{T} \times [0, 1]$ for $K = 2, N_1 = 3, N_2 = 2$ (e.g. a buyer-seller game where buyers have 3 alternative strategies and sellers have 2); and the tetrahedron (3-simplex) for $K = 1, N = 4$ (e.g. workers choose among four occupations). I will not systematically discuss dimension four and higher; the lower dimension state spaces are often the most tractable and useful in applications.

The formal structure presented below, like most of the theoretical literature, assumes there are infinitely many (indeed, a continuum of) individual players in each population. The assumption serves four purposes. First, the actual population size could be growing or declining over time without affecting the state space specification. Second, assuming a large population allows the analytic convenience of a continuous state space and dynamics specified by ordinary differential equations or difference equations. A finite population specification, by contrast, restricts the current state s to some finite lattice, and a large Markov transition matrix (or something even more cumbersome) is then required to specify the dynamics. Third, the law of large numbers can be invoked for an infinite population, allowing the model to ignore random fluctuations and differing perceptions of the current state. Last, and perhaps most important, each individual can be regarded as small so the GAN condition is easily justified.

In many applications, all four purposes can be adequately served by fairly small actual populations. The approximations involved in modeling the state space as a continuum and in ignoring random fluctuations, growth, etc, often are less troublesome than the numerous other simplifications required to construct a tractable model of some aspect of economic reality.²

¹ One possible candidate for reduction to $K = 1$ is a worker-consumer model in which an individual's fitness is the simple average of her fitness as a worker and her fitness as a consumer.

² Theorists can't afford to be so casual. See for example Boylan (1992), Maynard Smith (1982) and Riley (1979) for explication of various theoretical issues arising from finite or countably infinite populations. Binmore and Samuelson (1995) offer an enlightening integrated discussion of population size, mutation or error rates, period length and time horizon. See also the remarks in Section 2C below on local interaction.

The real issue is analogous to the price-taking approximation in competitive equilibrium models: is the GAN approximation reasonable, or must the model consider players' systematic attempts to influence others' future behavior? The answer is an empirical judgement. My own reading of the evidence is that the nature of players' interactions is at least as important as the number of players. For market interactions, laboratory results (Smith, 1982) and some field results (Bresnahan and Reiss, 1991) suggest that in terms of behavior, three buyers (or three sellers) can be a "large number".

Large population models have already been used in applications with only a few actual players. In numerical simulations of an international trade model, Friedman and Fung (1996) assume 10 firms in each country, but the analytic model is continuous. Indeed, some theorists use an evolutionary game model to analyze a situation in which each population consists of a *single* individual (e.g. Skyrms, 1986; see also Binmore, 1987–88). (The interpretation of the state s then is the profile of mixed strategies contemplated by the individual players. The dynamics represent an internal or virtual adjustment process for beliefs. I will discuss examples in Appendix B. Obviously the GAN assumption is more problematic in this individualistic interpretation.) My point is that an applied economist should not require thousands of actual players before considering a price-taking competitive model or an infinite population evolutionary model.

B. Stage games

Evolutionary games involve strategic interaction over time. At any point in time the strategic interaction is expressed as a stage game, an ordinary static game either in normal (strategic) form or in extensive (game tree) form. For simplicity I deal here only with stage games in normal form. The reader interested in extensive form stage games should consult Fudenberg and Kreps (1988), Fudenberg and Levine (1993, 1997), or Cressman (1997).

A normal form stage game is defined by a function $f(r, s)$, called the *fitness* or (expected) payoff function, where $r \in S$ is the (possibly mixed) strategy chosen by a specific player and $s \in S$ is the current state. When there are $K \geq 2$ populations, f has K components whose values do not depend on components of r corresponding to other populations, as illustrated below.

Most existing literature emphasizes one of two linear specifications for fitness functions. The first, dubbed the *symmetric case* by Maynard Smith (1982) involves a single population ($K = 1$) and an $N \times N$ matrix $A = ((a_{ij}))$. Players are drawn randomly in pairs and receive expected payoff³ $f(r, s) = r^T A s$ when playing mixed strategy r , a point in the N -simplex $S = \{(x_1, \dots, x_N) : x_i \geq 0 \text{ and } \sum_1^N x_i = 1\}$, against a population which in aggregate plays mixed strategy $s \in S$.

³ The T -superscript denotes the transpose. All vectors are column vectors, but (except in matrix multiplication) they are written for typographical simplicity as row vectors with the T superscript suppressed.

A simple one-dimensional example from biology is called Hawk-Dove. In a single animal population each individual is either aggressive (“hawk”) or avoids conflict (“dove”). For $r = (x, 1 - x)$ representing the mixed strategy of playing hawk with probability $x \in [0, 1]$ and for $s = (p, 1 - p)$ representing the current state of the population with fraction $p \in [0, 1]$

playing hawk, the fitness function is $f(r, s) = r^T A s$, where $A = \begin{pmatrix} -1 & 2 \\ 0 & 1 \end{pmatrix}$.

The story is that damages exceed gains on average in hawk-hawk pairings so the expected payoff is -1 , that gains are costlessly split in dove-dove pairings so the expected payoff is 1 each, and that hawks get all the gains (2) and doves none (0) when they meet. Cornell and Roll (1981) use essentially the same fitness function (payoff matrix) in their discussion of financial market participants who can either buy costly information (“hawks”) or not (“doves”).

The fitness function here suppresses the distinction between *realized* payoff in a single pairwise encounter, and the *expected* payoff averaged across all possible encounters for the given population state s . The usual justification is that in a large population the expected payoff is a sufficient statistic. Another justification, explored below, is that the payoff is non-stochastic because it arises not from random pairwise encounters but rather from general interactions such as voting or markets in a large population. In the general case, of course, the payoff can be a nonlinear function of the state s .

The second popular linear specification, often referred to in the literature as *the asymmetric case*, involves two populations (so $s = (s^1, s^2)$ where each s^i is a point in the N -simplex) and two $N \times N$ matrices, A and B .⁴ The payoff functions are $f^1(r, s) = r^T A s^2$ and $f^2(r, s) = r^T B s^1$. For example, Friedman (1991) discusses a “Buyer-Seller” game where buyers can either inspect or not and sellers can either report quality honestly or try to cheat. The payoff matrices (under current notational conventions) are

$A = \begin{pmatrix} 3 & 2 \\ 4 & 1 \end{pmatrix}$ for buyers and $B = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$ for sellers, e.g. an honest seller

gets a payoff of $b_{12} = 3$ when she meets a noninspecting buyer while he gets $a_{21} = 4$.

These linear specifications are useful for simple illustrative examples but are too restrictive for some economic applications. In the first specification there is only one economic role, and the population interacts with itself. In the second there are two distinct roles, but each player interacts only with players in the other population, not with players in his own population, e.g. $f^1(r, s)$ depends on s^2 but not on s^1 . That is, the specification rules out own population effects.

Own population effects can be important in applications with two or more populations. In the international trade example, firms in the home country compete in goods markets with foreign firms, but also with other

⁴ The reader should be warned that some of the literature specifies the transpose B^T rather than B , which can cause some confusion.

home country firms, so their fitness depends on both s^1 and s^2 . The general linear specification for $K = 2$ and $N_1 = N_2 = N$ can be written as $f^1(r, s) = r^T(As^2 + Cs^1)$ and $f^2(r, s) = r^T(Bs^1 + Ds^2)$, where the own population effects are captured in the two additional $N \times N$ matrices C and D .

Many applications require non-linear state dependence (e.g., Balkenborg and Schlag, 1995). Market interaction, whether perfectly or imperfectly competitive, produces prices and profits that typically are quite nonlinear in the population distribution of behavior. For example, in a simple one-country model of firms' choice between two modes of internal organization, Friedman and Fung (1996) use the Nash-Cournot equilibrium for given s to express $f(r, s)$ as the quotient of fourth degree polynomials in s (with coefficients depending on the demand and cost parameters and the number of firms). The work of Blume and Easley (1992) on wealth dynamics in financial markets provides another example. They express each individual's expected wealth increment as a nonlinear function of the distribution of investment strategies in a single population.

Can the fitness function $f(r, s)$ be nonlinear in r as well as in s ? It is not required by any application I have seen in economics (or in biology for that matter).⁵ An anonymous referee points out that linearity in r can be regarded as a simple notational convenience for summarizing the state dependence of each (pure) strategy, and that it underlies the compatibility specifications discussed in the next subsection. Thus I assume unless otherwise stated that fitness functions are linear in r and continuously differentiable (possibly nonlinear) in s .

C. Dynamics

Evolution of the state $s \in S$ over time can be stochastic or deterministic and in discrete or continuous time, and attainable states can be the whole continuum S or a discrete subset. How should an applied economist make the choice?

As always, the application in principle determines the specification, and in practice simplicity and convenience are highly valued. For the sake of simplicity and concreteness I will focus on deterministic, continuous-time, continuous-state dynamics and refer to other forms only occasionally. Before proceeding, however, a paragraph is in order on two important stochastic approaches.

Young (1993) and Kandori, Mailath and Rob (1993) pioneered an influential discrete stochastic approach that points to risk dominance (or more generally, to the width of the basin of attraction) to select among multiple NE. Applied modelers should be cautious about adopting their conclusions, I believe, because (a) they apply to the cosmologically long run

⁵ Nonlinearity in r conceivably could be used to represent risk preferences, but I will ignore the possibility for three orthodox reasons: analytic convenience, approximate risk-neutrality for (presumably relatively small) single-period payoffs, and (when all else fails) the game theorist's interpretation of payoffs as final utilities rather than wealth increments.

and quite different medium term considerations may dominate observed behavior (e.g., Roth and Erev, 1994), and (b) they postulate an isotropic stochastic process but actual “mutation” or error processes may have important biases (e.g., Friedman, 1996) which overturn their conclusions (Bergin and Lipman, 1996). A second stochastic approach, e.g., Ellison (1993) and Ellison and Fudenberg (1995), employs local interaction models that allow small scale fluctuations to spread even in a large population. Temzelides (1997) applies this approach to modeling bank panics. Such models are potentially valuable in applications involving interactions predominantly among nearest neighbors. The present paper, however, will focus on non-local interactions of entire populations.

Continuous deterministic dynamics are convenient for at least three reasons. First, we can get such dynamics from the other kinds by taking an appropriate limit as the time increment and the scale of random fluctuations go to zero and the number of players goes to infinity in each population (e.g., Binmore and Samuelson, 1995). Second, continuity in the state over time ensures inertia in the sense called for in the introduction. Third, such dynamics are customarily expressed in the relatively tractable form of a system of ordinary differential equations (ODEs) $\dot{s} = F(s)$, where $\dot{s} = (\dot{s}^1, \dots, \dot{s}^K)$ is the rate of change in the state and $F(s)$ is a specified vector valued function. For example, let A be a 2×2 matrix such as Hawk-Dove, and let the state $s = (p, 1 - p)$. Since S is one dimensional in this example, it suffices to express the ODE as a single equation in p , say⁶

$$\dot{p} = (1, -1)^T A(p, 1 - p) = 1 - 2p \text{ for } A = \begin{pmatrix} -1 & 2 \\ 0 & 1 \end{pmatrix}.$$

For a two dimensional state space we need to specify the time rate of change for two independent variables, and so on. As explained in the Appendix A, some mild technical restrictions must be imposed on F to ensure that s is well defined and remains in the state space S (e.g. that the fraction of hawks does not become negative).

The essence of biological evolution is that fitter genotypes increase relative to less fit genotypes. The economic analogue is that (due to imitation and learning or resource redistribution or entry and exit) fitter strategies increase relative to less fit strategies in every population. This principle, called *compatibility* in Friedman (1991), implies some relationship between a fitness function f and the corresponding dynamics $\dot{s} = F(s)$. For notational simplicity, restrict attention to a single population whose strategy choices are numbered $i = 1, \dots, N$ and let e_i denote the i^{th} pure strategy. Let $\hat{f}_i(s) = f(e_i - s, s) = f(e_i, s) - f(s, s)$ denote the relative fitness of i , that is, the absolute fitness $f(e_i, s)$ of pure strategy i less the current population-weighted average fitness $f(s, s) = \sum_{i=1}^N s_i f(e_i, s)$.

There are four popular ways to specify compatibility for the rates of change \dot{s}_i (or, alternatively, for the growth rates \dot{s}_i/s_i) given by F . From strongest to weakest, the specifications are that the s_i (or the \dot{s}_i/s_i)

(a) Are *proportional* to relative fitness; or

⁶ That is, using the payoff difference D defined in Section 3A below, we set $F(s) = (D(p), -D(p))$ for $0 < p < 1$.

- (b) Have the same *rank-order* as relative fitness; or
- (c) Have the same *signs* as relative fitness; or
- (d) Are *positively correlated* with relative fitness.

Specification (a) is the most popular in the literature. Choosing the time scale so that the constant of proportionality is 1.0 and equating each growth rate \dot{s}_i/s_i to relative fitness $\hat{f}_i(s)$, one gets the system of ordinary differential equations $\dot{s}_i = s_i f(e_i - s, s)$, $i = 1, \dots, N$, generally referred to in the literature as *replicator dynamics* (or occasionally as Malthusian dynamics).

For example, the Hawk-Dove game with matrix $A = \begin{pmatrix} -1 & 2 \\ 0 & 1 \end{pmatrix}$ and the

usual parametrization $s = (p, 1 - p)$ yields $e_i - s = (1, 0) - (p, 1 - p) = (1 - p)(1, -1)$, so here replicator dynamics is the single cubic ODE $\dot{s}_1 = \dot{p} = p(1 - p)(1, -1)^T A(p, 1 - p) = p(1 - p)(1 - 2p)$.

Replicator dynamics have a compelling biological justification – genetic fitness is *measured* as the relative growth rate – but are more difficult to justify in economic applications.⁷ An alternative implementation of compatibility version (a) equates the rate of change \dot{s}_i in the interior of S to fitness $f(e_i, s)$ relative to the simple average fitness⁸ $\bar{f}(s) = \frac{1}{N} \sum f(e_i, s)$, so we have $\dot{s}_i = f(e_i, s) - \bar{f}(s)$. In the Hawk-Dove example, we have $\dot{p} = \dot{s}_1 = [(1, 0) - (1/2, 1/2)]^T A(p, 1 - p) = (1 - 2p)/2$, and more generally in one dimensional models \dot{p} is proportional to the payoff difference.

Compatibility version (b), generally referred to as monotone dynamics (Nachbar, 1990; Samuelson, 1991), has become increasingly popular in applications because of its greater plausibility. But in some applications even (b) may be too strong. It might turn out that the population fraction using third best action sometimes grows more rapidly (or shrinks more slowly) than that for the second best action. In such cases versions (c) or (d) might be useful. Of course, (b–d) all coincide when there are only $N_k = 2$

⁷ Many theorists have accepted the challenge and sought plausible economic underpinnings for replicator dynamics. None are completely persuasive but several are interesting and instructive. Binmore and Samuelson (1995) obtain replicator dynamics as a limiting case of a complicated process that combines entry and exit, random mutation and aspiration-based imitation. Blume and Easley (1992) get replicator dynamics as a limiting case of their wealth redistribution process when they directly or indirectly assume that players try to maximize the growth rate of their wealth, as opposed to maximizing terminal wealth or present value of consumption utility. Börgers and Sarin (1993) show that transients (but not necessarily asymptotics) of an individual (but not aggregate) learning model correspond to replicator dynamics in a limiting case. Unfortunately the learning model is based on rote learning, not belief learning; for example, the marginal impact of current information is independent of accumulated information. Finally, Schlag (1994) obtains replicator dynamics as a limiting case of an imitation process in which players have a special sort of risk aversion. My point here is that replicator dynamics are a handy example of compatible dynamics but applied economists should not regard them as canonical.

⁸ As explained in Friedman (1991), when working with rates of change rather than growth rates, one should define the relative fitness in terms of the simple average of non-extinct strategies rather than the population weighted average. The distinction is vacuous for specifications (b) and (d) but in principle can affect (c) as well as (a).

alternative strategies. See Weibull (1995) for further discussion of (c) and see Friedman (1991) for further discussion of (d); both versions have plausible arguments in their favor and also have attractive asymptotic properties. Appendix B is a brief, unsystematic survey of proposed dynamical specifications, most but not all of which satisfy compatibility versions (b–d).

Two further remarks may be in order. An applied economist no doubt would seek empirical evidence in choosing among the various dynamic specifications. Laboratory data appears the obvious first place to look. The good news is that laboratory study of adjustment dynamics in games has become quite popular in the last few years; the bad news is that the work so far has focused on testing one or a few models in rather special circumstances, and a consensus has not yet formed on the most promising general approaches.

Finally, note that for most popular dynamics the information requirements are quite minimal, often just the history of own payoffs. Players do not necessarily need common (or any) knowledge of opponents' types or distribution of types in order to converge to ordinary NE. This point is emphasized in the theoretical literature but the implication for applied models may not have sunk in: *so-called complete information NE can evolve even when players initially are quite ignorant of each others' circumstances.*

D. Equilibria

The biological literature emphasizes a static equilibrium concept called ESS, for evolutionarily stable state or strategy. The original definition is for a single population linear model with fitness function $f(r, s) = r^T A s$ for some given $N \times N$ matrix A . It is as follows.

Definition. A state $s \in S$ is an ESS if for all other states $x \in S$ either

- (i) $f(s, s) > f(x, s)$, or
- (ii) $f(s, s) = f(x, s)$ and $f(s, x) > f(x, x)$.

The idea is that the ESS state s resists all possible “mutations” x either because (i) they are less fit or (ii) they are equally fit at the current state (where they are rare) but less fit when they are prevalent. Using the notation $x_\epsilon = (1 - \epsilon)s + \epsilon x$, one can replace conditions (i) and (ii) by the single condition

- (iii) $f(s, x_\epsilon) > f(x, x_\epsilon)$ for $\epsilon > 0$ sufficiently small.

This alternative definition is valid for nonlinear fitness functions and can be generalized for $K \geq 2$ models (e.g. Cressman, 1995).

Economists and orthodox game theorists for many years have emphasized a somewhat weaker static equilibrium concept called NE, for Nash or noncooperative equilibrium. For a single population model the definition is simply

Definition. A state $s \in S$ is a NE if for all $x \in S$

$$(1') f(s, s) \geq f(x, s).$$

It is clear from a comparison of the definitions that every ESS is a NE; this relationship also holds for appropriate generalizations to $K \geq 2$ models.

The intuitive appeal of ESS is that it supposedly ensures that the equilibrium is stable. A formal definition of stability must refer to the dynamic F , not simply the fitness function f for the stage game. The standard dynamic definitions are that $s \in S$ is a *fixed point* if $F(s) = 0$, and that a fixed point s is locally asymptotically *stable* if every open neighborhood $N \subset S$ of s has the property that every path starting sufficiently close to s remains in N and converges asymptotically to s . Following Hirshleifer (1982), I shall refer to such a dynamically stable equilibrium point as an EE or *evolutionary equilibrium*.⁹ The largest open set of points whose evolutionary paths converge to a given EE is called its *basin of attraction*.

What are the general relationships among these static and dynamic equilibrium concepts? It turns out that even weak compatibility (version (d)) suffices to ensure that every EE of F is a NE of the underlying fitness function f .¹⁰ Under the same conditions, every NE of f is a fixed point (but not necessarily stable) for F . There are various settings (e.g., replicator dynamics) which ensure that every ESS of f is an EE of F (e.g., Cressman, 1992, 1995), but in general with weak compatible dynamics ESS is neither necessary nor sufficient for EE (e.g. Friedman, 1991).

Which equilibrium concepts are most useful in applications? I personally favor EE and NE. It is reasonable to suppose in most applications that behavior evolves over time according to some (not necessarily deterministic) dynamic. Behavior is most easily observed when it has settled down, i.e., when the state remains near a fixed point. If the dynamic is not too sluggish or noisy, then most of the time the state should be near an attractor. Some theoretical literature emphasizes limit cycles and even more complex attractors called chaotic or strange attractors (e.g. Baumol and Benhabib, 1989) but the corresponding behavior will be difficult or impossible to observe in practice unless it is recorded very frequently and precisely. In most applications, then, the useable empirical evidence will generally come from states near an EE.

There are drawbacks to working with EE. Like ESS, they are not guaranteed to exist, although virtually every normal form applied model I know has at least one EE for some compatible dynamics. When they do exist, EE can be difficult to identify analytically; in general, one must fully specify the dynamic F , not just the stage game f , and use eigenvalue

⁹ More generally, a closed set $K \subset S$ is called an *attractor* of the dynamic F if every open neighborhood N has the property that every path starting sufficiently close to K remains in N and has limit points only in K , and every point in K is a limit point of such a path. Thus an EE is an attractor that consists of a single point.

¹⁰ Remember that we are dealing with normal form games. The relationship between evolutionary equilibrium and Nash equilibrium is more complicated for extensive form games; see for example Fudenberg and Levine (1993).

techniques on its Jacobian matrix or more advanced techniques (e.g. Lyapunov functions) to verify that a fixed point is an EE. The next section shows that in some cases ESS is a useful sufficient condition for EE. However, applied economists more often focus attention on NE, and the current discussion suggests that this is a good first step. Quite generally, NE is a necessary condition for EE, and NE are directly computable from the stage game fitness function f . The next section demonstrates that sometimes a little knowledge of the dynamics can suffice to eliminate the unstable NE and isolate the EE.

3 Local classification of low-dimensional systems

Low dimensional models are especially tractable for applications. Such models permit a rather limited repertoire of asymptotic behavior. This section classifies generic behavior¹¹ in one and two dimensions for dynamics specified as ordinary differential equations and mentions some additional considerations in passing. The one dimensional results and the two-dimensional no-own-population-effects results go back at least to Zeeman (1980) and Eshel and Akin (1982). Here ESS is a sufficient condition and NE is a necessary condition for dynamic stability.

A. Linear one-dimensional games

Recall that in the one-dimensional case there is only one population ($K = 1$) and two pure strategies ($N = 2$), so the state space $S = \{(p, 1 - p) : p \in [0, 1]\}$ can be parametrized by a point p in the unit interval $[0, 1]$ representing the fraction of players currently employing the first pure strategy. The fitness function $f(r, s)$ specifies the expected payoff to any player choosing mixed strategy $r = (x, 1 - x)$, i.e. taking the first action with probability x and the second action with probability $1 - x$. If f is linear in $s = (p, 1 - p)$ as well as in r , then for some 2×2 matrix $A = ((a_{ij}))$ it can be written as $f(r, s) = r^T A s = xpa_{11} + x(1 - p)a_{12} + (1 - x)pa_{21} + (1 - x)(1 - p)a_{22}$.

For one-dimensional evolutionary games, compatibility versions (b–d) all reduce to saying that the action with the higher current payoff will tend over time to displace the action with the lower payoff. That is, the direction of change in $s = (p, 1 - p)$ is entirely determined by the sign of the payoff difference $d(s) = f(e^1, s) - f(e^2, s) = f((1, -1), s)$ between the first pure strategy $e^1 = (1, 0)$ and the second pure strategy $e^2 = (0, 1)$: if $d(s) > 0$ then p increases and s moves towards $(1, 0)$, while if $d(s) < 0$ then p decreases and s moves towards $(0, 1)$.

The analysis is completely transparent when the payoff differences $d(s)$ is written in terms of p and the components of A , so

¹¹ That is, I usually ignore knife-edge cases because small structural shocks (e.g. small perturbations of matrix elements) will eliminate these cases. Again, extensive form stage games require a more complex analysis than I will provide here.

$$\begin{aligned} D(p) &= d(s(p)) = (1, -1)^T A(p, 1-p) = (1-p)(a_{12} - a_{22}) - p(a_{21} - a_{11}) \\ &= (1-p)a_1 - pa_2, \end{aligned}$$

where the reduced parameters are $a_1 = a_{12} - a_{22}$ and $a_2 = a_{21} - a_{11}$. Then compatible dynamics ensure that $\dot{p} \geq 0$ as $D(p) \geq 0$, so only the sign of $D(p)$ matters. The graph of $D(p)$ is a straight line with intercept a_1 at $p = 0$ and value $-a_2$ at $p = 1$. Apart from the degenerate case $a_1 = a_2 = 0$ in which a player is always indifferent between her two actions, we have only three possibilities.

Type 1: If $a_1, a_2 > 0$ then the unique root of $D(p) = 0$ is $p^* = a_1/(a_1 + a_2)$. It is easy to see that p^* is the only NE and the only ESS. More importantly for present purposes, D is downward sloping so p increases (decreases) whenever it is below (above) p^* . Hence p^* is the unique evolutionary equilibrium. That is, for any compatible continuous-time dynamic $\dot{s} = F(s)$ we have $s(t) \rightarrow (p^*, 1-p^*)$ as $t \rightarrow \infty$ from any initial state $s(0)$. The same conclusion also holds for discrete time ($t = 0, 1, 2, \dots$) dynamics $\Delta s(t) = \alpha F(s(t))$ on S if we add the proviso that the adjustment rate parameter $\alpha > 0$ is not too large. (We can get unstable oscillations if p_{t+1} jumps too far over p^* .) Examples of type 1 games include symmetric versions of Matching Pennies (e.g. $A = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$ so $a_1 = a_2 = 2$) and Hawk-Dove (e.g. $A = \begin{pmatrix} -1 & 2 \\ 0 & 1 \end{pmatrix}$ so $a_1 = a_2 = 1$).

Type 2: For $a_1, a_2 < 0$, the root $p^* = a_1/(a_1 + a_2)$ of $D(p) = 0$ is still a NE of the associated symmetric bimatrix game, but now both pure strategies $p = 0$ and $p = 1$ are also NE. The pure strategies both are ESS, but p^* is not. As Figure 1 makes clear, $D(p)$ slopes upward and is negative (positive) for $p < p^*$ ($p > p^*$), so p^* is an unstable “source” separating the basins of attraction of the two evolutionary equilibria $p = 0$ and $p = 1$. An economic interpretation is that each pure strategy has increasing returns in type 2 games and decreasing returns in type 1 games. Type 2 games are often called symmetric coordination games. An example is $A = \begin{pmatrix} 5 & -1 \\ 4 & 1 \end{pmatrix}$.

Type 3: If $D(p)$ lies on or above (on or below) the p -axis for all $p \in (0, 1)$, then the second pure strategy $p = 1$ (the first pure strategy $p = 0$) is dominant. Of course, the dominant strategy is the unique NE (and ESS) and the unique evolutionary equilibrium for any sign-preserving dynamic F . This type of game is characterized by $a_1 \cdot a_2 \leq 0$ (and $|a_1| + |a_2| > 0$). The most interesting example is Prisoner’s Dilemma, in which payoffs decrease as the dominant strategy becomes more prevalent, e.g. $A = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix}$ so $a_1 = -a_2 = -1$.

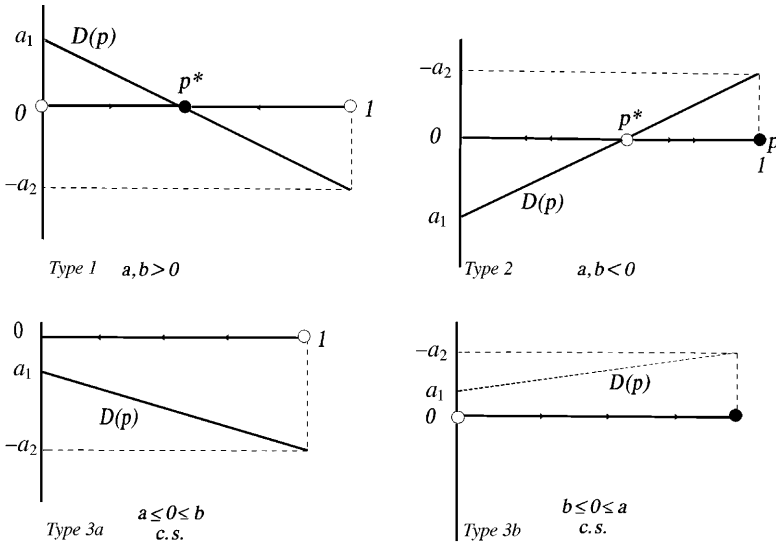


Fig. 1. Linear one-dimensional evolutionary games. *Notes:* For payoff matrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, define $a_1 = a_{12} - a_{22}$ and $a_2 = a_{21} - a_{11}$. The point $s(p) = (p, 1 - p)$ represents the current evolutionary state for $0 \leq p \leq 1$. The current payoff difference between the two pure strategies is $D(p) = (1 - p)a_1 - pa_2$

B. Two-dimensional systems: The asymmetric case

Recall the square $S = [0, 1]^2$ is a two dimensional state space that arises in two-population, two-strategy games ($K = 2, N_1 = N_2 = 2$). In the so-called asymmetric case – linear fitness functions and no own population effects – we have 2×2 matrices A and B such that $f^1(r, s) = r^T A s^2$ and $f^2(r, s) = r^T B s^1$. Define the reduced parameters $a_1 = a_{12} - a_{22}$ and $a_2 = a_{21} - a_{11}$ as before, and analogously define $b_1 = b_{12} - b_{22}$ and $b_2 = b_{21} - b_{11}$. At first it might seem that there are 9 cases of the form $x - y$ where x runs through the 3 cases in the last subsection for the a parameters and y does the same for the b parameters. However, Appendix A shows that there actually are only 3 distinct generic cases, illustrated in Figure 2.

Cases 3 – y and x – 3 (Corner): If a_1 and a_2 have opposite signs, then there is a dominant strategy for the first population. The unique EE (and NE and ESS) is the corner corresponding to the first population all adopting the dominant strategy and the second population all making the best response (see panel A of Fig. 2). The same characterization applies whenever b_1 and b_2 have opposite signs; of course, then the corner corresponds to the second population all adopting their dominant strategy and the first population best-responding. Adaptations of textbook entry deterrence models (e.g. Katz and Rosen, 1991, p. 581) are easy examples.

Cases 1 – 1 and 2 – 2 (Saddle): If all reduced parameters are positive then there is a completely mixed NE at $s^* = ((p^*, 1 - p^*), (q^*, 1 - q^*))$, where $p^* = a_1 / (a_1 + a_2)$ and $q^* = b_1 / (b_1 + b_2)$. This NE is a saddlepoint of the

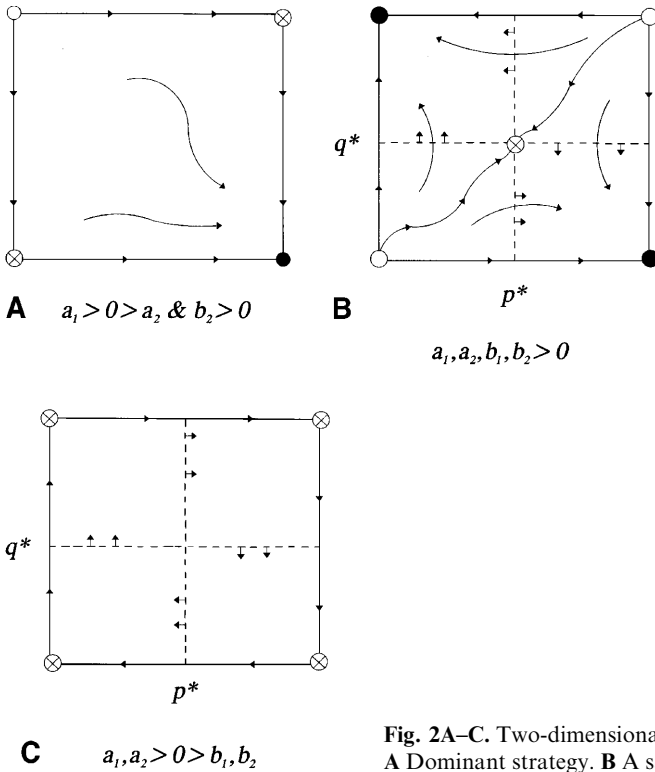


Fig. 2A–C. Two-dimensional asymmetric games. **A** Dominant strategy. **B** A saddle. **C** A center

dynamics. There are two pure strategy *NE*, both also *EE* as well as *ESS*, at the corners $(p, q) = (1, 0)$ and $(0, 1)$ of the square. Their basins of attraction are separated by the saddlepath to s^* (see panel B of Fig. 2). If all reduced parameters are negative then the result is the same except the *EE/ESS* are at the corners $(1, 1)$ and $(0, 0)$. Friedman and Fung (1996) obtain this case when there is free trade in goods but not in factors.

Cases 1 – 2 and 2 – 1 (Center): If a_i are both positive and b_i both negative, then s^* is the unique *NE*. It is not an *ESS*. The time path of every other initial state spirals around s^* clockwise (see panel C of Fig. 2). Whether s^* is asymptotically stable (an *EE*) depends on the dynamic specification. In version (a) – e.g. replicator or linear dynamics – s^* is neutrally stable and the spirals are closed orbits (cycles). Versions (b–d) allow outward spirals (unstable) or inward spirals (*EE*). The case in which the parameters b_i are both positive and the a_i are both negative is the same except that the spirals are counterclockwise. This case is illustrated in the basic Buyer-Seller game in Friedman (1991).

Remarks:

1. Zeeman (1980) and Hofbauer and Sigmund (1988) offer a fairly complete classifications for N -strategy one-population linear games under replicator dynamics.

2. For replicator dynamics there is an interesting generalization which applies to higher dimensional “asymmetric” (linear, no own-population effects) cases. It can be shown that the dynamics are volume-preserving (Hofbauer and Sigmund, 1988, Ch. 27) so internal NE (i.e. completely mixed strategy NE) can never be EE. Such NE are either saddlepoints or neutrally stable centers.

3. Selten (1991a) notes that the lack of internal EE applies to discrete-time linear dynamics as well as continuous time replicator dynamics. He shows that “anticipatory learning” can stabilize the Center case and higher dimension analogues. Roughly speaking, if the linear dynamics are $\Delta s = L(s)$ then the anticipatory dynamics are $\Delta s = L(s + L(s))$.

4. It may be worth emphasizing that the one-dimensional linear game with payoff matrix A (and payoff matrix A^T for the opponent in a pairwise random matching setting) is *not* equivalent to the two population “asymmetric” game $(A; A^T)$. Take a type 1 matrix, say $A = \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix}$. We have

seen that the one-dimensional game has a unique NE and a unique EE at $p^* = a/(a + b) = (2 - 1)/[(2 - 1) + (0 - (-1))] = \frac{1}{2}$. The two-dimensional game $(A; A^T)$ falls into case 1 - 1. Its interior NE at $(p, q) = (\frac{1}{2}, \frac{1}{2})$ is not an EE, but its two corner NE at $(p, q) = (1, 0)$ and at $(0, 1)$ are EE. This is analogous to behavior in one dimensional games of type 2, not type 1.

C. Other two-dimensional systems and nonlinear systems

Consider now the possibility of own-population effects in linear fitness functions on the square. As shown in Appendix A, it is still true that there is at most one interior NE. In addition to the saddle and center cases already encountered, two further types of behavior, called source and sink in Figure 3, are possible at an interior NE. Appendix A shows that some fitness functions have an interior NE which is an EE (a sink or stable center) for any compatible dynamic, while for other fitness functions the interior NE is a source or unstable center for every compatible dynamic and therefore never an EE. As with the center case illustrated in Figure 2C, there are also fitness functions for which some compatible dynamics make

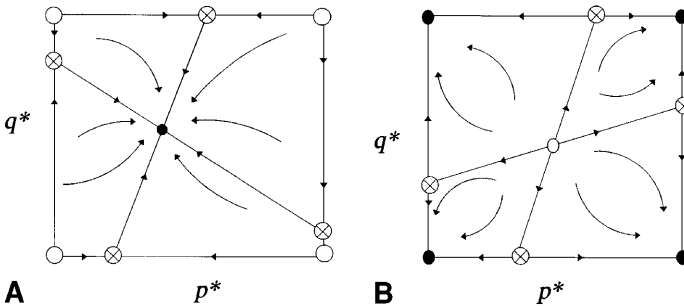


Fig. 3A,B. Other linear games on the square. **A** A sink. **B** A source

the interior NE an EE and other compatible dynamics make it unstable (not an EE). For simplicity I retain the same label, center, for this case, even though the range of compatible behavior now includes sources and saddles as well as stable and unstable spirals. Behavior along the edges of the square can exhibit all the main one dimensional cases, depending on the fitness function.

Recall that linear games on the triangle are specified by a fitness function $f(r, s) = r^T A s$, where $A = ((a_{ij}))$ is a 3×3 matrix. By varying the a_{ij} 's, one can get every type of generic behavior – a dominant strategy (so the corresponding corner is a global sink and therefore an *EE*), a dominated strategy so paths converge to the opposite edge of the triangle, or an interior *NE*. All four types of behavior illustrated in Figures 2B,C and 3 – saddle, center, source and sink – are possible at an interior *NE*. All three types of behavior illustrated in Figure 1 are possible along the edges of the triangle.

Allowing fitness functions that are non-linear in s creates fewer practical complications than one might suppose. First consider one-dimensional games. The payoff difference $D(p) = f((1, -1), s(p))$ between the two available pure strategies can cross the p -axis several times. These crossing points (sometimes augmented by one or both endpoints $p = 0, 1$) constitute the *NE* of f . Generically the crossings are transverse – the tangent to $D(p)$ is not horizontal at the roots p^* of $D(p) = 0$. Then analysis reduces locally to cases 1 and 2 in Figure 1, depending on whether D is downward or upward sloping. Specifically, an interior root p^* is always an *NE*, and is an *EE* whenever $D'(p^*) < 0$ but separates the basins of attraction of neighboring *EE* whenever $D'(p^*) > 0$, as illustrated in Figure 4.

In the absence of own-population effects, the two-population two-dimensional non-linear case similarly reduces locally to the corresponding linear (“asymmetric”) case as follows. Find the roots p^* to the equation $0 = D_2(p) \equiv f_1(e_1 - e_2, s^1(p))$ and roots q^* to the equation $0 = D_1(q) \equiv f_2(e_1 - e_2, s^2(q))$. Draw vertical lines at p^* and horizontal lines at q^* , so the square is broken down into rectangles. For generic payoff functions (e.g. $D'_i(x^*) \neq 0$ at roots x^* of $D_i(x) = 0$) and any compatible dynamics, the situation in each rectangle falls into one of the three linear cases shown in Figure 2.

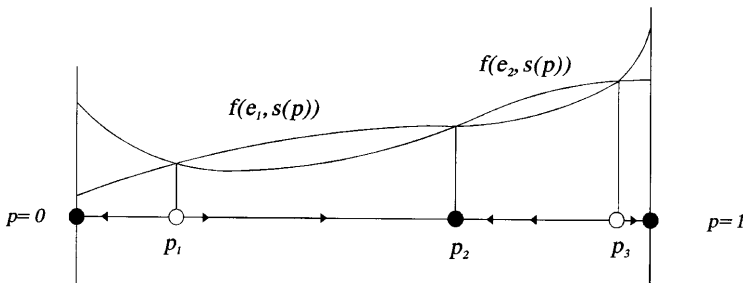


Fig. 4. A non-linear one-dimensional game. $D(p)$ here is the vertical distance between $f(e_1, s(p))$ and $f(e_2, s(p))$

General non-linear models on the square or triangle also reduce locally to one of the linear cases. That is, each interior *NE* is a local source, saddle, center or sink; it is never an *EE* in the first two cases, is always an *EE* in the last case, and its status depends on the detailed specification of dynamics in the third case. Edge and corner *NE* are either *EE* (local sinks) or not *EE* (saddles or sources) independent of the dynamic specification.

Behavior away from the *NE* can be fairly complicated in the general two-dimensional cases. For example, there can be stable or unstable limit cycles. “Strange attractors” and “chaos” are possible in three-dimensional cases, or in lower dimensions for discrete-time dynamics. In most applied work, it suffices to identify the *EE* and their basins of attraction. Empirically it will be difficult to detect systematic behavior elsewhere.

Since most economists feel more comfortable specifying stage games than specifying continuous time dynamics, it is good to know that the *EE* and their basins of attraction always are independent of the choice of compatible dynamics in dimension 1 and in three of four cases in dimension 2. In state spaces of dimension 3 and higher the dynamic specification matters in a larger fraction of cases.

4 Summary and discussion

Many interesting economic issues can be modeled in terms of interactions among a few distinct populations, each with a few alternative actions. If entry and exit or resource redistribution or learning tends to increase the prevalence of higher-payoff actions over time in each population, then an evolutionary game model may be appropriate. Time paths in such models quite generally converge to evolutionary equilibria (*EE*), a subset of the Nash equilibria for the underlying stage game. The learning (or other dynamic) process, together with the historically given initial state, determines which *EE* is attained when there is more than one.

One can be fairly specific about the *EE* and possible asymptotic behavior in one and two-dimensional continuous-time evolutionary games. As long as the dynamic specification is an ordinary differential equation compatible with the fitness function and the fitness function is qualitatively unaffected by tiny structural shocks, then there are very short lists of possibilities for local behavior. The lists apply globally when the fitness function is linear. For dimension 1, the list of three possibilities is illustrated in Figure 1. For dimension 2 and no own-population effects, there are again three possibilities, two involving interior *NE*, illustrated in Figure 2. Allowing own-population effects on the square or on the triangle introduces two additional possibilities for interior *NE*, illustrated in Figure 3. In every case except one (Fig. 2C) the status of the *NE* as an evolutionary equilibrium (*EE*) or not is determined entirely by the fitness function and is independent of the choice of compatible dynamic.

How can one use this classification in practice? It tells the applied economist directly what sorts of behavior can arise in simple evolutionary game models. For example, assumptions in the recent literature on the evolution of money (e.g. Kiyotaki and Wright, 1989; Wärneryd, 1989)

quickly boil down to saying that we have Type 2 models. Hence we can conclude from the classification scheme that the outcome will be a corner EE, which in these models means that we have a unique medium of exchange.

Perhaps less obvious but equally important, knowing the classification (and the basic economic structure) enables the applied economist to predict the qualitative effects of sizeable shocks or regime changes. For example, Friedman and Fung (1996) rely on the genericity of the classification of games whose state space is the square to conclude that small changes in trading costs or supply and demand parameters have no qualitative effect on their trade model's behavior. However, there is a large shock or regime shift when trade is opened in factors (capital and labor) as well as in goods. The classification scheme allows the authors to deduce from behavior at the four corner states that this shock causes behavior to go from case 1–1 to case 2–2. The policy implications of this transition turn out to be striking and perhaps surprising: although lowering trade barriers for goods generally has benign effects, it can happen that lowering trade barriers for capital and labor will impair global productive efficiency.

To drive the methodological point home, consider Benabou (1993), an intricate model of location choice by workers of two types. It appears that the model's qualitative features are isomorphic to those of the Friedman and Fung trade model, even though the interpretation is quite different. Hence the main results of the analysis could be obtained quite directly from the classification of games with square state spaces.

My discussion emphasizes applications in which the equilibria do not depend on the details of the dynamic specification. Economists should not conclude that dynamics are of little interest. The reason for the emphasis rather is the current lack of tractable and empirically tested models of learning and other dynamic adjustment processes. Eventually such dynamic models should become available and greatly expand the scope of evolutionary game models in economics. In the meantime I hope some readers are motivated to find new applications for the simple models presented here.

Appendix A: Some analytical details

This Appendix fills in some analytical details used in Section 3 and elsewhere. It relies on elementary methods, the most advanced being eigenvalue techniques for assessing dynamic stability (e.g. Chiang, 1984, Ch. 15).

Begin with linear games on the square, $K = 2$ and $N_1 = N_2 = 2$. Recall that the fitness function then has a representation of the form $f^1(r, s) = r^T (As^2 + Cs^1)$ and $f^2(r, s) = r^T (Bs^1 + Ds^2)$, where $Y = A, B, C, D$ are 2×2 matrices $((y_{ij}))$ and the current state $s = (s^1, s^2)$ can be written $s^1 = (p, 1 - p)$ and $s^2 = (q, 1 - q)$ for $(p, q) \in [0, 1]^2$. Define the payoff advantage of the first pure strategy in population i by $d^i(p, q) = f^i((1, -1), ((p, 1 - p), (q, 1 - q)))$, $i = 1, 2$.

The dynamics are specified as ordinary differential equations $(\dot{p} = F^1(p, q), \dot{q} = F^2(p, q))$. Trajectories $(p(t), q(t))$, also known as solution

curves or evolutionary paths, are solutions to the ODE with given initial conditions $p(0) = p_o$ and $q(0) = q_o$. I impose the substantive assumption of compatibility and two technical restrictions to ensure well-defined trajectories. Specifically I require that the dynamics (i) vary smoothly with the state, (ii) have trajectories that do not leave the state space, and (iii) are weakly compatible. For the square¹², these properties can be formalized as follows:

Definition. A dynamic $F : [0, 1]^2 \rightarrow R^2$ is *admissible* if

- (i) it is continuously differentiable,
- (ii) $F^1(0, q) \geq 0, F^1(1, q) \leq 0, F^2(p, 0) \geq 0$ and $F^2(p, 1) \leq 0$ for all $(p, q) \in [0, 1]^2$, and
- (iii) $sgnF^i(p, q) = sgnd^i(p, q)$ for all $(p, q) \in (0, 1)^2$.

Recall that the sign function $sgn(x)$ is equal to $+1, -1$ or 0 as x is positive, negative or zero.

First consider the ‘‘asymmetric’’ case $C = D = ((0))$. As noted in the text, $d^1(p, q) = (-1, 1)A(q, 1 - q)^T = (1 - q)a_1 - qa_2$, where $a_1 = a_{12} - a_{22}$ and $a_2 = a_{21} - a_{11}$, is independent of p . Similarly, $d^2(p, q) = (1 - p)b_1 - pb_2$ is independent of q because there are no own-population effects. Consequently admissible dynamics take specified signs in rectangular regions of $[0, 1]^2$ bounded by the lines $d^1(q) = 0$ and $d^2(p) = 0$.

The analysis is elementary and falls into several cases depending on the signs of the a_i and b_i . I ignore several cases where some a_i or b_i is zero because they are tedious and non-generic. Other cases are labelled $1 - x$ (respectively $x - 1$) if a_i (respectively b_i) are both positive, $2 - x$ (respectively $x - 2$) if a_i (respectively b_i) are both negative, and $3 - x$ (respectively $x - 3$) if $a_1a_2 < 0$ (respectively $b_1b_2 < 0$).

Begin with case $3 - x$. Then population 1 players have a dominant strategy because $d^1(q) > 0$ or < 0 on the interior of the square. Consequently all trajectories tend towards the edge of the square $p^* = 0$ or $p^* = 1$ corresponding to the dominant strategy. If $d^1(p^*) > 0$ (< 0) then clearly the upper end (lower end) of this edge is the limit of every trajectory as $t \rightarrow \infty$. Thus it is the unique *EE* for any admissible dynamic. The same analysis and conclusion holds for case $x - 3$: the unique *EE* is the right or left end of the edge corresponding to the dominant strategy $q^* = 0$ or $q^* = 1$ as $d^2(q^*) > 0$ or < 0 .

In the other cases the a_i 's have the same sign, so $d^1(q) = 0$ for $q^* = a_1/(a_1 + a_2) \in (0, 1)$, and likewise $d^2(p) = 0$ for $p^* = b_1/(b_1 + b_2) \in (0, 1)$. The point (p^*, q^*) obviously is a *NE* because for each population, at this point, all actions are equally fit. The vertical and horizontal lines through (p^*, q^*) divide the square into four quadrants (rectangles) in which the signs of \dot{p} and \dot{q} are constant.

Consider case $1 - 1$ and initial state in the *NE* quadrant. Since $q > q^*$ and $a_1, a_2 > 0$, we have $d^1(q) < 0$ so $\dot{p} < 0$; and since $p > p^*$ and $b_1, b_2 > 0$

¹² See Friedman (1991) for more general definitions.

we have $d^2(p) < 0$ so $\dot{q} < 0$ for any admissible dynamic. Conditions (i) and (iii) of admissibility then imply that the trajectory must either leave this quadrant in finite time or converge to its *SW* corner (p^*, q^*) . Similarly¹³, a trajectory starting in the *SW* corner must either leave the quadrant in finite time or converge to (p^*, q^*) . On the other hand, trajectories starting in the *SE* quadrant have $\dot{p} > 0$ (since $q < q^*$) and $\dot{q} < 0$ for any admissible dynamic. Hence such trajectories cannot leave the quadrant. For N any neighborhood of $(1,0)$ in the square \dot{p} and \dot{q} are bounded away from zero outside of N and therefore must enter N in finite time. Hence $(1,0)$ is an *EE* for any admissible dynamic F . The same argument establishes that the *NW* quadrant is contained in the basin of attraction for the *EE* at $(0, 1)$. The case 1 – 1 characterization in the text is completed by noting that the sets of points which leave the *NE* and *SW* quadrants are disjoint and separated by the saddle path to (p^*, q^*) .

The argument for the case 2 – 2 characterization is analogous. Straightforward sign checks on d^i show that admissible trajectories in the *SE* and *NW* quadrants relative to (p^*, q^*) consist of a saddlepath separating trajectories that leave these quadrants, and that the *NW* and *SE* quadrants lie respectively in the basins of attraction for the *EE* $(1,1)$ and $(0,0)$.

Cases 1 – 2 and 2 – 1 do not permit such definite conclusions. In case 1 – 2 we have $\dot{p} < 0$ and $\dot{q} > 0$ in the *NE* quadrant. A trajectory beginning in this quadrant must leave in finite time and enter the *NW* quadrant, since \dot{p} can be bounded away from zero along the trajectory. For analogous reasons, trajectories beginning in the *NW* quadrant must exit it in finite time for the *SW* quadrant, hence to the *SE* quadrant and back to the *NE* quadrant. A straightforward partition-of-unity construction as in Friedman (1991, p. 664) demonstrates admissible dynamics for which (p^*, q^*) is the global *EE* and other admissible dynamics for which all trajectories diverge from (p^*, q^*) . This completes the “asymmetric” case characterization.

Next consider the more general linear case with $C, D \neq ((0))$. Now $d^1(p, q)$ is a general linear function whose slope coefficient for p depends on c_{ij} in the same fashion as its q -coefficient depends on a_{ij} ; similarly for $d^2(p, q)$. *NE* (and hence *EE*) can occur only at intersections of the edges of the square and the lines $[d^i(p, q) = 0]$. Intersections at the edges and corners can be analyzed using the one-dimensional state space classification. Intersections of the lines in the interior of the square are automatically *NE* for the same reason as in the asymmetric case. Unfortunately the elementary sign-checking procedures are insufficient to characterize stability because the lines are not generally vertical or horizontal.

To characterize the stability of an interior *NE* (p^*, q^*) , use standard eigenvalue techniques. Any admissible dynamic evaluated at this point can be written $\dot{p} = \alpha d^1(p, q)$ and $\dot{q} = \beta d^2(p, q)$ for some $\alpha, \beta > 0$ by admissibility properties (i) and (iii). Then the Jacobian matrix is

¹³ A note for the technically-minded reader. Condition (i) is used to bound \dot{p} (or \dot{q}) away from zero outside of a neighborhood of the line $q = q^*$ (or $p = p^*$), to ensure exit in finite time when not on the saddle path.

$$J = \begin{pmatrix} \partial F^1 / \partial p & \partial F^1 / \partial q \\ \partial F^2 / \partial p & \partial F^2 / \partial q \end{pmatrix} = \begin{pmatrix} -\alpha(c_1 + c_2) & -\alpha(a_1 + a_2) \\ -\beta(b_1 + b_2) & -\beta(d_1 + d_2) \end{pmatrix}.$$

The characteristic roots (eigenvalues) are the solutions λ_1, λ_2 to the equation $|J - \lambda I| = 0$. By appropriate choice of the A, B, C and D matrices and the α and β coefficients one can get any desired real or complex-conjugate values for the λ_i and therefore every sort of stability behavior¹⁴. The two generic possibilities not yet encountered are λ_i both real and positive (e.g. $A = B = ((0))$ and $c_i, d_i > 0$), and λ_i both real and negative. In the first case, standard techniques show that (p^*, q^*) is a source – all trajectories beginning near it move away, and that it is a sink (EE) in the second case. See Hirsch and Smale (1974), for example, for the well-known but non-elementary analysis.

Some elaboration of this point may prevent confusion. Recall that the sum of the eigenvalues is the trace of the Jacobian matrix J and the product is the determinant of J . When $\det J$ is negative – i.e. when $cd < ab$ where $x = x_1 + x_2$ for $x = a, b, c, d$ – then all compatible dynamics produce saddles. When the determinant is positive and c and d are both positive, then $\text{tr } J = -(ac + \beta d) < 0$ for all compatible dynamics ($\alpha, \beta > 0$) so any interior NE is always an EE – it is either a sink or a stable center. Likewise, when c and d are both negative we have $\text{tr } J > 0$, so (when the determinant is positive) any interior NE must be unstable (a source or an unstable center). When c and d differ in their signs and the determinant is positive we have a generalization of the center case – the sign of $\text{tr } J$ is not determined and careful choice of compatible dynamics can produce sinks and sources as well as stable or unstable centers. As an example of this generalized center, choose matrices so that $a = -c = d = -1$ and $b = 0$. Then it is easy to verify that an interior NE is a source for $\alpha = \beta^{-1} = 0.1$, a sink for $\alpha = \beta^{-1} = 10$ and a neutral center for $\alpha = \beta = 1$.

Analysis of linear games on the triangle ($K = 1, N_1 = 3$) involves a 3×3 matrix A specifying the fitness function $f(r, s) = r^T A s$. Writing $s = (p, q, 1 - p - q)$ and $r = (x, y, 1 - x - y)$, one can tediously derive a 2×2 matrix B s.t. $f(r(x, y), s(p, q)) = (x, y)B(p, q)^T$, where the components b_{ij} are linear combinations of the a_{ij} . The three lines $d^{ij}(p, q) \equiv f(e_i - e_j, s(p, q)) = 0$ define indifference between each pair $i \neq j$ of the three strategies. Intersections of the edges ($p = 0, q = 0$ and $p + q = 1$) and these lines contain the NE . Again analysis of the stability status on edges reduces to the one-dimensional cases. Analysis of interior NE (intersections of two of the lines implies a 3-way intersection) again reduces to finding eigenvalues of the Jacobian matrix. The matrix components come directly from A (via B) when replicator or linear dynamics are assumed (see Taylor and Jonker, 1978, for example). Allowing more general compatible dynamics only increases the degrees of freedom. Therefore any interior NE

¹⁴ The reader can verify that when $C = D = ((0))$, the characteristic equation is of the form $\lambda^2 = \gamma$. Hence when $\gamma > 0$ the roots are real and of opposite sign (yielding a saddle) and when $\gamma < 0$ the roots are purely imaginary (yielding a center). Thus eigenvalue techniques lead to the same conclusions as the elementary analysis of the “asymmetric” case.

stability properties desired – center, source, sink or saddle – can arise by suitable choice of A .

This Appendix has been explicit only for linear fitness functions. Non-linearities already are allowed for in the dynamics, so not much needs to be changed for non-linear fitness functions. The main possibility is that there may be several interior intersections of the $d^i = 0$ lines, so the arguments apply locally, not necessarily globally (see Hirsch and Smale, 1974, for general techniques for local characterizations in terms of linearizations).

Appendix B: Models of adjustment dynamics

This Appendix discusses a sample of the burgeoning literature on dynamical models of learning and evolution. The sample is far from complete but it gives an impression of the variety of recent learning models (see the September 1996 issue of *Economic Journal* for more systematic surveys).

The oldest dynamic specification, going back at least to Cournot, is that players make a best response to the current state. Such dynamics can be written as $\dot{s} = F(s) \equiv \alpha(B(s) - s)$, where B is (a selection from) the best reply (correspondence) and $\alpha > 0$ is an adjustment rate parameter. The best-response assumption (here and below) is rather strong in applications where players can't see the entire distribution of play in the population (see Gilboa and Matsui, 1991, for a theoretical example of Cournot dynamics; see Fudenberg and Levine, 1997, for a thorough theoretical treatment).

Perhaps the next oldest dynamical scheme is *fictitious play* (Brown, 1951), associated with the individualistic interpretation of S . Here a state $s \in S$ represents not the fractions of populations currently choosing each action – since there is only one player in each population, the fractions are trivially 0 or 1 – but rather the subjective *beliefs* of players about other players' possible actions. Usually it is applied to 2-player games; for more players one needs further strong assumptions e.g., that each pair of players share the same beliefs regarding third parties. Players observe other players' actions each period and consider these observations as independent identically distributed samples from "Nature's" true distribution. This (mis-specified) model yields beliefs $s(t) = t^{-1} \sum_{u=1}^t a(u)$, where $a(u)$ is the discrete-time action profile at a previous dates $u < t$. Assuming players chose their expected best response, we have $a(t+1) = B(s(t))$, so the difference equation $s(t+1) = \frac{t}{t+1}s(t) + \frac{1}{t+1}B(s(t))$ characterizes discrete-time fictitious play dynamics.¹⁵ The continuous-time analogue is the ordinary differential equation $\dot{s} = \alpha_t(B(s) - s)$, where α_t is proportional to t^{-1} [see Wittman et al. (1996) for an application, Cheung and Friedman (1997) for some laboratory evidence, and again see Fudenberg and Levine (1997) for a thorough theoretical treatment].

It turns out that fictitious play dynamics converges (albeit slowly) in a wide class of simple games, and, when convergent, always converges to a

¹⁵ Note that the action profile $a(t)$ can take occasional large jumps. In fictitious play dynamics one generally gets $\dot{s} \rightarrow 0$ as $t \rightarrow \infty$ so inertia holds with respect to beliefs rather than actions.

NE. Shapley (1964) provided a first example of non-convergence. The inherent mis-specification – action profiles are neither independent nor identically distributed except (trivially) at NE – together with Shapley’s example inhibited work with fictitious play dynamics for more than 20 years. Recently, however, many variants have appeared. Milgrom and Roberts (1991) keep the same general schema (beliefs \rightarrow actions \rightarrow new beliefs) but loosen it considerably. They obtain convergence to a subset of states which contains Nash equilibria but not necessarily convergence to any particular NE. They point to applications in industrial organization and other fields. Fudenberg and Kreps (1988, 1993) also loosen the schema somewhat and introduce active learning – players may occasionally “experiment” with a strategy which is not currently an expected best response in order to learn more about their opportunities. Thus learning in part is an active investment.

Another variant on fictitious play is least-squares (LS) learning, popular in the macroeconomics literature following Marcat and Sargent (1989). The objects of LS learning are unknown parameters of a macroeconomic model, but the updating scheme, inherent misspecification, and convergence properties are essentially the same as for fictitious play.

Other individualistic discrete-time learning models include Selten (1991a), Kalai and Lehrer (1990) and Jordan (1991). Selten’s “anticipatory” dynamics are discussed at the end of section 3B. Kalai and Lehrer allow full rationality in a repeated game setting. They obtain convergence to a NE (of the repeated game) when players actively try to learn each other’s repeated game strategies and begin with sufficiently broad prior beliefs. Jordan obtains convergence to a stage game NE when myopically rational players (endowed with a common prior over possible payoff functions) learn the realized game by observing play. The convergence rate is exponential, rather than $O(t^{-1})$ as in most other individualistic learning models.

The population learning models vary considerably in the degree of rationality assumed. Crawford (1989) assumes individual-specific replicator dynamics for learning and obtains non-convergence. Foster and Young (1990) and Arthur (1989) take stochastic versions of the replicator and obtain convergence to NE. Canning (1990) allows traders Bayesian rationality (with respect to the mis-specified stationary model) or more general updating functions; a little randomization (interpreted as deaths and births) ensures convergence to NE. Matsui and Matsuyama (1991) produce a rational expectations evolutionary model by assuming large numbers of players who can switch strategies only when exogenous, randomly timed (Poisson) opportunities arrive. Players each maximize the expected present value of the payoff stream, know the model parameters and have perfect foresight. In the limit, as switching opportunities become more frequent, behavior converges globally to the risk dominant NE in 2×2 coordination games. Multiple equilibria are possible for more general stage games.

Most but not all of the evolutionary models proposed so far satisfy the weak compatibility condition (d) of section 2C. This condition (but not necessarily the monotone version (b) or the sign preserving version (c)) automatically holds for the best-reply and fictitious-play type dynamics.

Sometimes it fails for forward-looking rational dynamics, e.g. for initial segments of Matsui and Matsuyama or Jordan paths.

References

- Arthur WB (1989) A nash-discovering classifier system for finite-action games. Manuscript, Santa Fe Institute
- Balkenborg D, Schlag K (1995) On the interpretation of evolutionarily stable sets. University of Bonn, working paper
- Baumol W, Benhabib J (1989) Chaos: Significance, mechanism and economic applications. *Journal of Economic Perspectives* 3: 77–105
- Benabou R (1993) Workings of a city: Location, education, and production. *Quarterly Journal of Economics* 108: 619–652
- Bergin J, Lipman BL (1996) Evolution with state dependent mutations. *Econometrica* 64: 943–956
- Binmore K (1987–88) Modeling rational players, Parts I and II. *Economics and Philosophy* 3: 179–214, 4: 9–55
- Binmore K (1991) De-Bayesing game theory. Presentation at the International Game Theory Conference, Florence, June 1991
- Binmore K (1992) *Fun and games: A text on game theory*. Heath, Lexington, MA
- Binmore K, Samuelson L (1995) Musical chairs: Modelling noisy evolution. *Games and Economic Behavior* 11: 1–35
- Blume L, Easley D (1992) Wealth dynamics and the market selection hypothesis. *Journal of Economic Theory* 58: 9–41
- Bomze I, Pötscher B (1989) *Game theoretic foundations of evolutionary stability*. Springer, Berlin Heidelberg New York
- Börgers T, Sarin R (1993) Learning through reinforcement and replicator dynamics. Manuscript, University College, London; *Journal of Economic Theory* (forthcoming)
- Boylan R (1992) Laws of large numbers for dynamical systems with randomly matched individuals. *Journal of Economic Theory* 57: 473–504
- Bresnahan TF, Reiss PC (1991) Entry and competition in concentrated markets. *Journal of Political Economy* 99: 977–1009
- Brown G (1951) *Iterated solution of games by fictitious play. Activity analysis of production and allocation*. Wiley, New York
- Canning D (1990) Convergence to equilibrium in a sequence of games with learning. Discussion paper No. TE/89/190, London School of Economics and Political Science
- Cheung Y.-W, Friedman D (1996) Individual learning in games: Some laboratory results. *Games and Economic Behavior* 19: 46–76
- Chiang AC (1984) *Fundamental methods of mathematical economics*, 3rd edn. McGraw-Hill, New York
- Cornell B, Roll R (1981) Strategies for pairwise competitions in markets and organizations. *Bell Journal* 12: 201–213
- Crawford V (1989) Learning and mixed-strategy equilibria in evolutionary games. *Journal of Theoretical Biology* 140: 537–550
- Crawford V (1991) An “evolutionary” interpretation of Van Huyck, Battalio, and Beil’s Experimental Results on Coordination. *Games and Economic Behavior* 3: 25–59
- Crawford V (1993) Introduction: Adaptive dynamics, Parts I and II. *Games and Economic Behavior* 5: 315–319
- Cressman R (1992) *The stability concept of evolutionary game theory: a dynamic approach*. Springer, Berlin Heidelberg New York
- Cressman R (1995) Evolutionary game theory with two groups of individuals. *Games and Economic Behavior* 11: 237–253
- Cressman R (1997) Subgame monotonicity in two-person extensive form evolutionary games. Manuscript, Wilfrid Laurier University, Waterloo, Canada

- Cripps M (1991) Correlated equilibria and evolutionary stability. *Journal of Economic Theory* 55: 428–434
- Diekmann O, Christiansen F, Law R (1996) Evolutionary dynamics. *Journal of Mathematical Biology* 34: 483–483
- Dosi G, Marengo L, Fagiolo G (1996) Learning in evolutionary environments. CEEL Working Paper 1996-05, University of Trento, Italy
- Eichberger J (1992) Nash equilibrium and learning. In: Creedy et al. (eds) *Recent developments in game theory*, Edward Elgar, Aldershot, England
- Ellison G (1993) Learning, local interaction and coordination. *Econometrica* 61: 1047–1071
- Ellison G, Fudenberg D (1995) Word of mouth communication and social learning. *Quarterly Journal of Economics* 110: 93–125
- Eshel I, Akin E (1982) Coevolutionary instability of mixed nash solutions. *Journal of Mathematical Biology* 20: 123–133
- Foster D, Young H (1990) Stochastic evolutionary game dynamics. *Theoretical Population Biology* 38: 219–232
- Friedman D (1991) Evolutionary games in economics. *Econometrica* 59: 637–666
- Friedman D (1996) Equilibrium in evolutionary games: Some laboratory results. *The Economic Journal* 106: 1–25
- Friedman D, Fung KC (1996) International trade and the internal organization of firms: An evolutionary approach. *Journal of International Economics* 41: 113–137
- Fudenberg D, Kreps D (1988) Learning and equilibrium in games. Draft manuscript, Stanford University
- Fudenberg D, Kreps D (1993) Learning mixed equilibria. *Games and Economic Behavior* 5: 320–367
- Fudenberg D, Levine D (1993) Steady state learning and nash equilibrium. *Econometrica* 61: 547–574
- Fudenberg D, Levine D (1997) *Theory of learning in games*. MIT Press, Cambridge, MA
- Gilboa I, Matsui A (1991) Social stability and equilibrium. *Econometrica* 59: 859–867
- Hirshleifer J (1982) Evolutionary models in economics and law: Cooperative versus conflict strategies. *Research in Law and Economics* 4: 1–60
- Hirsch M, Smale S (1974) *Differential equations, dynamical systems, and linear algebra*. Academic Press, New York
- Hofbauer J, Sigmund K (1988) *The theory of evolution and dynamical systems*. Cambridge University Press, Cambridge
- Jordan J (1991) Bayesian learning in normal form games. *Games and Economic Behavior* 3: 60–81
- Kalai E, Lehrer E (1993) Rational learning leads to nash equilibrium. *Econometrica* 61: 29–56
- Kandori M, Mailath GJ, Rob R (1993) Learning, mutation, and long run equilibria in games. *Econometrica* 61: 1019–1045
- Katz M, Rosen H (1991) *Microeconomics*. Irwin, Boston
- Kiyotaki N, Wright R (1989) On money as a medium of exchange. *Journal of Political Economy* 97: 927–954
- Mailath GJ (1992) Introduction: Symposium on Evolutionary Game Theory. *Journal of Economic Theory* 57: 259–277
- Malthus TR (1798) *An essay on the principle population as it affects the future improvement of society with remarks on the speculations of Mr. Godwin, M. Condorcet and other writers*. Johnson, London
- Marcet A, Sargent T (1989) Convergence of least squares learning mechanisms in self-referential linear stochastic models. *Journal of Economic Theory* 48: 337–368
- Matsui A, Matsuyama K (1995) An approach to equilibrium selection. *Journal of Economic Theory* 65: 415–434
- Maynard Smith J, Price G (1973) The logic of animal conflict. *Nature* 246: 15–18

- Maynard Smith J (1982) *Evolution and the theory of games*. Cambridge University Press, Cambridge, MA
- Milgrom P, Roberts J (1991) Adaptive and sophisticated learning in repeated normal form games. *Games and Economic Behavior* 3: 82–100
- Mookherjee D, Sopher B (1994) Learning behavior in an experimental matching pennies game. *Games and Economic Behavior* 7: 62–91
- Nachbar J (1990) 'Evolutionary' selection dynamics in games: Convergence and limit properties. *International Journal of Game Theory* 19: 59–89
- Riley J (1979) Evolutionary equilibrium strategies. *Journal of Theoretical Biology* 76: 109–123
- Robson A (1992) Evolutionary game theory. In: Creedy et al. (eds) *Recent developments in game theory*. Edward Elgar, Aldershot, England
- Roth A, Erev I (1994) Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior* 8: 164–212
- Samuelson L (1991) Limit evolutionarily stable strategies in two-player, normal form games. *Games and Economic Behavior* 3: 110–128
- Samuelson L (1997) *Evolutionary games and equilibrium selection*. MIT Press, Cambridge, MA London
- Samuelson L, Zhang J (1992) Evolutionary stability in asymmetric games. *Journal of Economic Theory* 57: 363–391
- Schlag K (1994) Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits. Discussion Paper No. B-361, University of Bonn *Journal of Economic Theory* (forthcoming)
- Selten R (1980) A note on evolutionary stable strategies in asymmetric animal conflicts. *Journal of Theoretical Biology* 84: 93–101
- Selten R (1991a) Anticipatory learning in games, Chapter 3. In: Selten R (ed) *Game equilibrium models*, Vol 1. Springer, Berlin Heidelberg New York
- Selten R (1991b) Evolution, learning and economic behavior. *Games and Economic Behavior* 3: 3–24
- Shapley L (1964) Some topics in two-person games. *Advances in game theory*. *Annals of Mathematical Studies* 5: 1–28
- Skyrms B (1986) Deliberation equilibria. *Topoi* 5: 59–67
- Smith V (1982) Microeconomic systems as experimental science. *American Economic Review* 72: 923–955
- Taylor PD, Jonker LB (1978) Evolutionary stable strategies and game dynamics. *Mathematical Biosciences* 40: 145–156
- Temzelides T (1997) Evolution, coordination and banking panics. *Journal of Monetary Economics* (forthcoming)
- Van Damme E (1991) *Stability and perfection of nash equilibria*, 2nd edn. Springer, Berlin Heidelberg New York
- Vega-Redondo F (1996) *Evolution, games and economic behavior*. Oxford University Press, New York
- Wärneryd K (1989) Legal restrictions and the evolution of media of exchange. *Journal of Institutional and Theoretical Economics* 145: 613–626
- Weibull J (1992) An introduction to evolutionary games. IIESR Working Paper 347, Stockholm University
- Weibull J (1995) *Evolutionary game theory*. MIT Press, Cambridge, MA
- Wittman D, Friedman D, Crevier S, Braskin A (1996) Learning liability rules. *Journal of Legal Studies* 26: 145–164
- Young P (1993) The evolution of conventions. *Econometrica* 61: 57–84
- Zeeman E (1980) Population dynamics from game theory. In: Nitecki Z, Robinson C (eds) *Global theory of dynamical systems*. Lecture Notes in Mathematics 819. Springer, Berlin Heidelberg New York