Evolutionary Games in Economics
Author(s): Daniel Friedman
Source: *Econometrica,* Vol. 59, No. 3, (May, 1991), pp. 637–666
Published by: The Econometric Society
Stable URL: http://www.jstor.org/stable/2938222
Accessed: 13/04/2008 19:25

http://www.jstor.org

EVOLUTIONARY GAMES IN ECONOMICS

By Daniel Friedman[1]

Evolutionary games are introduced as models for repeated anonymous strategic interaction. The basic idea is that actions (or behaviors) which are more "fit," given the current distribution of behaviors, tend over time to displace less fit behaviors. Simple numerical examples motivate the key concepts of fitness function and compatible dynamics, and illustrate the relation to previous biological models. Cone fields are introduced to characterize the continuous-time dynamical processes compatible with a given fitness function. The analysis focuses on dynamic steady state equilibria and their relation to the static equilibria known as NE (Nash equilibrium) and ESS (evolutionary stable state). For large classes of dynamics it is shown that all stable dynamic steady states are NE and that all NE are dynamic steady states. The biologists' ESS condition is less closely related to the dynamic equilibria. The paper concludes with a brief survey of economic applications.

KEYWORDS: Evolutionary games, evolutionary stable strategies, cone field dynamics.

1. INTRODUCTION

IN AN EVOLUTIONARY GAME each individual chooses among alternative actions or behaviors whose payoff or fitness depends on the choices of others. Over time the distribution of observed behavior in a population evolves, as fitter strategies become more prevalent. The very prevalence of a behavior can make it (and other behaviors) more fit or less fit, so dynamics can be quite complex. One can ask which behaviors become extinct and which survive over time, whether the system approaches some stable steady-state, and so forth.

The present paper has two goals. The first is to propose a tractable framework for evolutionary games which incorporates the relevant previous work by John Maynard Smith and his collaborators in biology and mathematics, but which is suitable for economic applications. The second goal is to derive substantive results on the relation between evolutionary steady states and static equilibria of the payoff or fitness function. Meeting these goals should encourage economists to employ evolutionary models in serious applied work.

The biologists emphasize pairwise interactions of individuals drawn randomly from a single population and center their analysis on a static equilibrium concept known as ESS, for "evolutionarily stable" strategy or state. The inequalities defining ESS are intended to capture the intuition of an "uninvadable" state of the population: if the population is in an ESS then a small minority employing any deviant behavior (an "invasion of mutants") will eventually disappear under natural selection. Mathematicians, beginning with Taylor and

Jonker (1978), confirmed that ESS was sufficient (although not necessary) for stability under natural selection dynamics.

Economists and other social scientists can, I believe, usefully employ many of the ideas introduced by the biologists, but the biologists' formal structure needs to be adapted and extended. A fundamental point is that biologists almost always deal with the genetic mechanism of natural selection. This mechanism admits a simple, canonical dynamical representation, which I refer to below as the Malthusian dynamic.[2] For economists the social mechanisms of learning and imitation are usually more important than the genetic mechanism. A wide variety of learning and imitation processes are conceivable and the appropriate dynamical representation seems to be highly context-dependent. To deal compactly with a diverse set of possible dynamics my formal analysis employs the device of cone fields, introduced by Smale (1976). It is not my purpose to propose any general theory of learning, so I rely on examples and references to existing literature to suggest specific dynamical processes which may be of interest.

Two other, less fundamental extensions are worth noting briefly. Most of the existing biological models consider the evolution of a single species rather than the coevolution of several distinct species. My basic model, by contrast, features interactions of several strategically distinct populations, so as to represent economic relationships between buyers and sellers, or between residents in different jurisdictions, etc. Most biological models have (bi)linear specifications and are interpreted in terms of interactions between randomly matched individuals. By contrast, my basic model allows some nonlinearities to capture possible interactions with a whole population, a possibility biologists refer to as "playing the field."

Evolutionary games as presented here are quite distinct from supergames and from differential games, even though all three examine the dynamics of repeated strategic interaction. Differential games allow payoffs to depend on the time paths of all individual players' actions, and supergames allow strategies defined on time paths even though current payoffs depend only on current actions. Such intertemporal strategies and payoffs are ruled out in evolutionary games. Thus a natural interpretation is of essentially *anonymous* interactions, in which the individuals one deals with today are not recognizably the same as those dealt with yesterday. Another distinction is that my analysis of evolutionary games focuses on the distribution of behaviors in *populations* rather than on the behavior of rational *individuals*.[3]

From the standpoint of an economist specializing in game theory, the static features of my formal model are minor variants of standard normal form games,

---

[2] These dynamics are also referred to as replicator dynamics or as (the radial projection onto the simplex of) the Lotka-Volterra equations.

[3] That is, I interpret the evolving state as a distribution of strategies across a population. One could also interpret it as the evolving (virtual or actual) mixed strategy employed by an individual, but one must then deal with some deep issues regarding rationality. In Section 4 I will cite some recent papers that employ this individualistic interpretation.

but the dynamic features and evolutionary perspective will perhaps seem novel. Even for an economist interested only in long run equilibrium outcomes, a dynamic approach offers important benefits. First, one can distinguish stable from unstable equilibria. More generally, one can find the basins of attraction, i.e., the regions of initial conditions that eventually lead to a given equilibrium. Thus in principle one can identify which equilibrium point (if any[4]) is relevant.

The next section briefly presents the formal ingredients of the model: spaces of states and strategies, fitness functions (variants of payoff functions) to specify strategic interaction, and systems of ordinary differential equations to specify dynamics. I then provide five simple examples. The first two are intended to give the reader an intuitive grasp of evolutionary games' structure and behavior and to motivate the pivotal concept of compatibility. The remaining examples make connections to previous literature from biology and elsewhere. Next I show how compatible dynamics can be summarized as cone fields, and end the section by defining static and dynamic equilibrium concepts.

The analytic results are collected in Section 3. The most important demonstrate a close general relationship between the dynamic equilibria and the Nash equilibria (NE) of a fitness function—viz: all NE are compatible dynamic equilibria (Proposition 3.2) and all compatible dynamic *stable* equilibria are NE (Proposition 3.3). Of secondary interest is the relationship between ESS and the dynamic equilibria. The ESS are always a subset of the NE, so to the extent that Maynard Smith correctly identifies the dynamically stable equilibria in terms of his ESS criteria, he provides game theorists with an appealing refinement of NE. Unfortunately in general ESS is neither necessary nor sufficient for compatible dynamic stability, although there are some interesting special cases for which it is (Proposition 3.4 and counterexamples). Section 4 briefly summarizes the argument, points to some possible generalizations of the model, and surveys some recent and prospective applications of evolutionary games. Notational conventions, technical definitions, and proofs are collected in an Appendix.

2. THE MODEL

2.1. *Basic Elements*

Consider a set of interacting *populations*, indexed $k = 1, \ldots, K$. A member of each population k has available a finite number of actions (sometimes called "behaviors" or "strategies") indexed $i = 1, \ldots, N$. (The increased generality in allowing N to depend on k is negligible and not worth the increased notational burden.) Any point r^k in the N-simplex $S^k := \{x = (x_1, \ldots, x_N): x_i \geqslant 0, \Sigma x_i = 1\}$ represents a possible *mixed strategy* for an individual member of population k. Any point s^k in the same simplex also represents the *fractions of population* k employing each available strategy. Hence the Cartesian product of K copies of

---

[4] In general dynamical systems, the state ultimately may approach a limit cycle or chaotic attractor rather than an equilibrium point. The analysis in the present paper, however, focuses on equilibrium points.

the simplex, $S := S^1 \times \cdots \times S^K$, is the *set of strategy profiles* and also is the *state space* under the maintained interpretation that interactions are anonymous. It will be used for both purposes below.[5]

Strategic interaction is summarized in a fitness function which specifies for individuals in each population the evolutionarily relevant payoff as a function of own strategy and current state. Formally a *fitness function* consists of maps f^k: $S^k \times S \to R$, $k = 1, \ldots, K$, which are assumed linear in the first (own strategy) argument $r^k \in S^k$ and continuously differentiable in the second (population state) argument $s \in S$.[6] The fitness function can be written more compactly as $f: S \times S \to R^K$, with $f(r, s) := (f^1(r^1, s), \ldots, f^K(r^K, s))$. The linearity in $x = r^k$ can be regarded as a "large numbers" assumption: for large populations the fitness of a mixed strategy is the expected fitness of its component pure strategies. This linearity permits an alternative representation in terms of the *fitness gradient vector* $\hat{f}^k(s) \in R^N$ for population k individuals given state $s \in S$, with $x \cdot \hat{f}^k(s) = f^k(x, s)$ for all $x \in S^k$.

The final basic model element is a dynamical structure specifying how a state s evolves over time. In continuous time[7] one specifies the time derivatives $\dot{s} = (\dot{s}^1, \ldots, \dot{s}^K)$, with $\dot{s}^k := (\dot{s}^k_1, \ldots, \dot{s}^k_N) := (ds^k_1/dt, \ldots, ds^k_N/dt)$, by means of some function $F: S \to R^{NK}$. Thus

$$(2.1) \qquad \dot{s} = F(s)$$

is an autonomous system of ordinary differential equations whose solution curve $s(t)$ from given initial conditions $s(0) \in S$ describes the evolution of all populations beginning at any state of interest. Some restrictions are required for legitimate dynamics. Say that $F: S \to R^{NK}$ is *admissible* if:

(a) $\qquad \sum_{i=1}^{N} F_i^k(s) = 0$ for all $s \in S$ and $k = 1, \ldots, N$;

(b) $\qquad s_i^k = 0$ implies $F_i^k(s) = 0$, and

(c) $\qquad F$ is continuous and piecewise differentiable on S.

The first two conditions ensure that s^k doesn't leave the simplex S^k: by (a) the

---

[5] The dual role of S and its relation to the anonymous population interpretation will be discussed further in the last section.

[6] Crawford (1989) shows that assuming continuity of the fitness function (in its second argument) has real bite. It may be worth mentioning that although the fitness function is closely related to the traditional payoff function of normal form games, there are formal and (in some contexts) conceptual differences. In the special case that f^k depends perhaps on the state of other populations s^{-k} but does *not* depend on own population s^k, the standard payoff function g with components $g^k(s|_k x) := f^k(x, s)$ is equivalent to the fitness function f. Own population effects and some conceptual distinctions will be illustrated in Examples 3 and 5 of the next subsection.

[7] I will not analyze discrete time dynamics (difference equations) in this paper. Apart from some minor but annoying technicalities concerning behavior at the boundary of the simplex, it appears to be a routine exercise to obtain discrete time analogs for most of my results. Indeed Nachbar (1989) and Samuelson (1987), working independently and using different techniques, obtain difference equation analogues of (the order-compatible version of) my Propositions 3.2 and 3.3 for fairly general fitness functions.

TABLE I

Payoff Bimatrix for Basic Buyer / Seller Game

Seller:	1 ("honest")	2 ("cheat")
Buyer: 1 ("inspect")	(3, 2)	(2, 1)
2 ("don't")	(4, 3)	(1, 4)

population fractions will continue to sum to unity, and (b) prevents fractions from becoming negative. Also (b) prevents the revival of extinct strategies.[8] Condition (c) is a mild technical condition to help ensure "well-behaved" solutions; actually, the weaker assumption of Lipshitz continuity guarantees unique solution curves and would suffice for most purposes. An admissible F henceforth will be referred to as a *dynamical system* (or, simply, a *dynamic*) on S.

2.2. Some Simple Examples

EXAMPLE 1: *Basic ideas*. Consider a stylized interaction between two populations, called "buyers" ($k = 1$) and "sellers" ($k = 2$). Each seller has available two possible actions or strategies, "honest" ($i = 1$) and "cheat" ($i = 2$); and each buyer also has two alternative strategies, "inspect" ($i = 1$) and "don't inspect" ($i = 2$). Since the strategy/state simplices here are 1-dimensional the state $s = ((s_1^1, s_2^1), (s_1^2, s_2^2))$ can be described by a point (p, q) in the square $[0, 1] \times [0, 1]$, with $s_1^1 = p$ as the fraction of buyers who inspect and $s_1^2 = q$ as the fraction of honest sellers, so $s_2^1 = 1 - p$ and $s_2^2 = 1 - q$.

The populations interact in an unorganized market in which individual buyers and sellers of, say, used cars meet and participation is occasional so reputational effects, etc., are negligible. The fitness of a strategy depends on the population mix of chosen strategies, e.g., cheating is less attractive to sellers when more buyers inspect, but otherwise matters are as simple as possible (e.g., random matching, no externalities, and risk neutrality). Specifically, assume that the fitness function is bilinear and given by the bimatrix of Table I.[9] Thus for a buyer the fitness of $e^1 = (1, 0) =$ inspect with probability 1, is $f^1(e^1; s) = 3q + 2(1 - q) = q + 2$ (independent of p) and $f^1(e^2; s) = 4q + (1 - q) = 3q + 1$ is the fitness of not inspecting. Likewise $f^2(e^1; s) = 3 - p$ is the fitness for a seller of being honest with probability 1 and $f^2(e^2; s) = 4 - 3p$ is the fitness of cheating.

---

[8] Presumably imitation and perhaps simple versions of Bayesian learning would not permit the revival of extinct strategies, but more flexible ideas of learning might well permit it. The direct restriction (b) has no bite in stability analysis, since all strategies, extinct or otherwise, are contemplated in a general small perturbation. See also Remark 3.2 below.

[9] The table entries are from Game 72 of Rapoport and Guyer (1966). Pitchik and Schotter (1987) analyze essentially the same strategic situation in greater depth. However, their comparative static results implicitly assume that the interior NE is stable, and the analysis to follow shows that this assumption is problematic. See also Frank (1987) for a single-population game with a similar flavor.
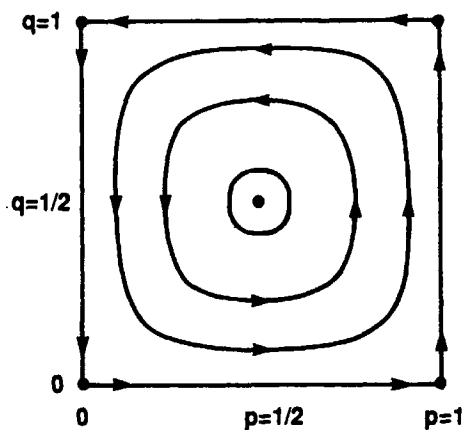
FIGURE 1

As for dynamics, assume with Malthus that the growth rate of a strategy is proportional to or (with an appropriate choice of time scale) equal to its relative fitness. That is, the growth rate of inspection $(\ln \cdot p) = \dot{p}/p$ is equal to its fitness $f^1(e^1; s)$ less the population average fitness among buyers $f^1(p; s) :=$ $pf^1(e^1; s) + (1 - p)f^1(e^2; s)$. Hence,

$$(2.2) \qquad \dot{p} = p(1 - p)(1 - 2q).$$

The growth rate $\dot{q}/q$ of honesty among sellers is its relative fitness $f^2(e^1; s) - f^2(q; s) = (1 - q)(2p - 1)$, so

$$(2.3) \qquad \dot{q} = q(1 - q)(2p - 1).$$

It is easy to check that the system (2.2)–(2.3) of coupled differential equations has five fixed points (i.e., steady-states), at the center and at the four corners of the $p - q$ square. It can be shown (see the Appendix) that all other points are on periodic trajectories circling counterclockwise, as shown in Figure 1. Thus under Malthusian dynamics the four corner fixed points are unstable and the center point is neutral, neither asymptotically stable nor unstable.

EXAMPLE 2: *Alternative dynamics*. Economically plausible alternatives to Malthusian dynamics can be constructed from specific models of imitation and learning. For example, suppose that each used car buyer transacts only once; before making his or her own choice, the buyer observes the choice of a single, randomly selected previous buyer and, with a small exogeneous probability α, also observes the predecessor's payoff. Assuming that all participants are Bayesians with diffuse priors over p and q and with common knowledge of the information structure, it evidently is rational for the new buyer to imitate the predecessor's choice when the outcome is not observed, and to make a "best reply" to the inferred seller choice when the outcome is observed. This behavior yields the discrete-time dynamic $\Delta p = \alpha(-q + [1 - q])$. Choosing the time scale so that $\Delta t = 2\alpha$ and taking the limit as the information lag $\Delta t \to 0$ one obtains the differential equation $\dot{p} = \frac{1}{2} - q$. By strict analogy the sellers' population obeys the differential equation $\dot{q} = p - \frac{1}{2}$.
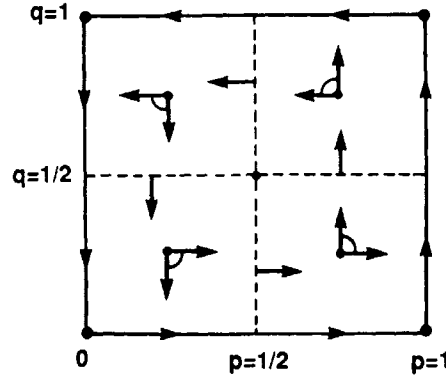
FIGURE 2

One can also directly postulate linear dynamics by setting the rate of change (not the growth rate as in Malthus) of a strategy equal to its fitness relative to the simple average (not population-weighted average) of all strategies currently employed.[10] From the Table I fitness function one calculates

$$\dot{p} = f^1(e^1;s) - \tfrac{1}{2}[f^1(e^1;s) + f^1(e^2;s)]$$

$$= \tfrac{1}{2}[f^1(e^1;s) - f^1(e^2;s)], \quad \text{i.e.,}$$

$$(2.4) \qquad \dot{p} = \tfrac{1}{2} - q$$

and similarly,

$$(2.5) \qquad \dot{q} = p - \tfrac{1}{2}$$

for (p,q) in the interior of the square. To keep p and q in $[0,1]$, override (2.4) and (2.5) by

$$(2.6) \qquad \dot{p} = 0 \text{ if } p = 0,1; \quad \dot{q} = 0 \text{ if } q = 0,1.$$

Clearly the piecewise linear system (2.4)–(2.6) has exactly the same fixed points as the system (2.2)–(2.3), and trajectories near $(\tfrac{1}{2}, \tfrac{1}{2})$ are counter-clockwise concentric circles. The main qualitative difference from the Malthusian version is that strategies can become extinct in finite time.

One can imagine many other sorts of dynamics, but it is reasonable to require *compatibility* with the fitness function in the sense that *fitter strategies should increase relative to less fit strategies*. In the present example inspection is fitter or less fit for buyers, i.e., $f^1(e^1;s) \gtrless f^1(e^2;s)$, as $q \lessgtr \tfrac{1}{2}$; and similarly $f^2(e^1;s) \gtrless f^2(e^2;s)$ as $p \gtrless \tfrac{1}{2}$. Thus the requirement is that all dynamical systems compatible with the fitness function of Table I will specify tangent vectors $(\dot{p}, \dot{q})$ such that $\dot{p} \gtrless 0$ as $q \lessgtr \tfrac{1}{2}$ and $\dot{q} \gtrless 0$ as $p \gtrless \tfrac{1}{2}$. The set of all such tangent vectors form the cones sketched at a few points in Figure 2. Compatible dynamics evidently all have the same five fixed points, the four corner points being saddles. The center point could be stable or unstable depending on which specific compatible dynamic is chosen. Using the static equilibrium concepts defined in the next

---

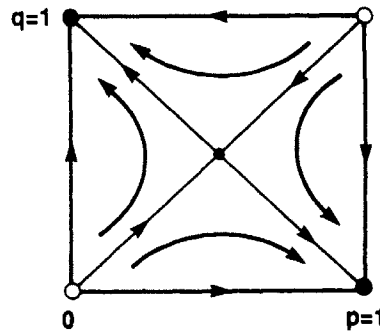[10] These linear dynamics were developed jointly with Jack Hirshleifer.

FIGURE 3

subsection it is easy to show that the center point is the unique NE and that no ESS exists for the present fitness function.

The remaining examples mainly illustrate relationships to previous literature and can be skipped with little loss of continuity.

EXAMPLE 3: *Other two-strategy games.* Any 2×2 bimatrix game can be analyzed in exactly the same way as above. For instance, a version of the Battle of the Sexes[11] has the payoff bimatrix

$$\begin{pmatrix} 2,2 & 3,4 \\ 4,3 & 1,1 \end{pmatrix}.$$

The phase portrait under Malthusian dynamics is symmetric with a saddle point at the center, "sources" at $(0,0)$ and $(1,1)$ and "sinks" at $(1,0)$ and $(0,1)$, as shown in Figure 3. The linear dynamics produce a very similar phase portrait. Indeed, it is not difficult to check that the compatibility condition defines cones in the four quadrants of the state space such that *any* compatible dynamic produces a phase portrait qualitatively similar to Figure 3. In terms of the story usually told about this game, the present analysis therefore suggests that the populations evolve so that the sex which initially was more insistent on getting its own way will ultimately always get its way and the other sex ultimately will always accommodate. Of course, to be consistent with my interpretation of evolutionary games as anonymous interactions of populations, the story should be about socially acceptable behavior on first dates and not about ongoing behavior in marriage.

Fitness functions obtained in this way from bimatrices have no own-population effects, but can readily be modified to incorporate them. For example, in the buyers and sellers story, widespread inspection (high values of p) might tie up sellers and hence impose external costs on buyers. Similarly sellers could also suffer from negative own-population effects. Thus one could justify appending a term $-ap$ to the f^1 and a term $-bq$ to the f^2 from Example 1, for small positive constants a and b. It can be shown that in this case the central fixed point becomes stable, but other fixed points (and static equilibria) do not change status.

---

[11] Biologists use the same name to refer to a quite different game.

There is no need to limit the number of interacting populations to two. For example, one could augment the basic buyer/seller game by adding a population of regulators, who can either be active (achieving the same detection rate of dishonesty, say, as buyers who inspect) or inactive. With plausible assignments of payoffs for regulators one obtains phase portraits on the unit cube in which inspection eventually disappears among buyers, and cheating by sellers and activity by regulators cycles around a neutrally stable point; this point becomes stable in the presence of appropriate external costs.

EXAMPLE 4: *A K* = 1, *N* = 3 *game from biology*. A single population can interact with itself, so that *only* own-population effects are present. The interaction can be specified as a bimatrix game with an arbitrary number of pure strategies, as long as the payoff matrix for the "row" player $A = (a_{ij})$ is the transpose of the payoff matrix for the "column" player. The transpose restriction ensures that row and column players are strategically identical, and so can be regarded as in the same population. For example, consider the standard interpretation for bimatrix games in biology, that a given individual (the "row player") playing pure strategy *i*, is paired with another individual (the "column player") drawn randomly from the same population. The row player receives a_{ij} when the column player plays *j*. If the column player is strategically the same, her payoff must be a_{ji}.

For example, Maynard Smith (1982, p. 22) specifies a territorial contest for some animal species in which three behaviors (strategies) are available: *i* = 1 or *H* ("Hawk"), *i* = 2 or *D* ("Dove"), and *i* = 3 or *B* ("Bourgeois"). The state/strategy space *S* is the 2-dimensional unit simplex in R^3 with vertices $e^1 = (1,0,0)$, $e^2 = (0,1,0)$, and $e^3 = (0,0,1)$. After rescaling, the payoff matrix (for the row player) in this game is:

$$A = \begin{pmatrix} -2 & 4 & 1 \\ 0 & 2 & 1 \\ -1 & 3 & 2 \end{pmatrix}.$$

Thus, an *H* player gets a payoff of -2 against another *H*, but $+4$ against *D* and $+1$ against *B*. The fitness function here is $f(x, s) = \sum_{i,j=1}^3 x_i s_j a_{ij} = x \cdot As$, the expected payoff to mixed strategy *x* when matched with an opponent drawn randomly from a population at state *s*. It can be shown directly from the definitions that the two NE are $\hat{s} := (\frac{1}{2}, \frac{1}{2}, 0)$ and e^3, and that e^3 is the only ESS for this fitness function.

As usual, Malthusian dynamics are defined by equating for each strategy *i* the growth rate $\dot{s}_i/s_i$ to the relative fitness $f(e^i, s) - f(s, s)$. Hence one obtains the polynomial (cubic) equations

$$(2.7) \qquad \dot{s}_i = s_i(e^i \cdot As - s \cdot As), \qquad i = 1, 2, 3.$$

As shown in Figure 4, the fixed points are e^1 and e^2 (unstable, sources), $\hat{s}$ (unstable, a saddle), and e^3 (stable, a sink). The latter is a global attractor: from
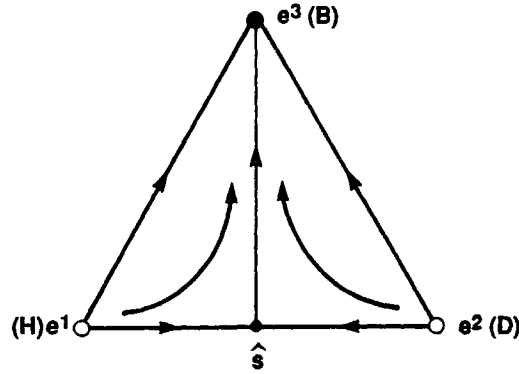
FIGURE 4

almost any initial state there is a transient tendency for the proportions of H and D to equalize, but ultimately the state evolves to 100% B.

EXAMPLE 5: *A migration process.*[12] Consider again two populations (row players and column players) each with two pure strategies, and let the matrix

$$A = \begin{pmatrix} 0 & 4 \\ 12 & 0 \end{pmatrix}$$

describe the dollar outcomes of the interaction. For $\hat{x}$ denoting the point $(x, 1 - x)$ in the 2-simplex, one sees that $h(x, q) := \hat{x} \cdot A\hat{q}$ is the expected payoff for a row player (resp. $h(x, p)$ for a column player) employing the first strategy with probability $x \in [0, 1]$ when q (resp. p) is the fraction of column (resp. row) players employing their first strategy in random pairings. One can verify that the NE for h (or A) are $(0, 1)$, $(1, 0)$ and $(0.25, 0.25)$.

Now, following James Friedman and Robert Rosenthal (1986), assume a migration dynamic for each population with the emigration rate from a strategy proportional to the difference between the overall maximum payoff and the expected payoff for the given strategy. Specifically, the maximum payoff is 12, so the row player emigration rates are $m_i(q) := (12 - a_{i1})q + (12 - a_{i2})(1 - q) = 12 - h(2 - i, q)$ for strategy $i = 1, 2$, and $m_i(p)$ gives the column player emigration rates. Consequently the population of row players is described by the differential equation

$$(2.8) \qquad \dot{p} = -p m_1(q) + (1 - p)m_2(q),$$

the two terms representing emigration from and immigration to the first strategy. The column population obeys the same equation with p and q interchanged. One can verify that the NE $(0.25, 0.25)$ is *not* a fixed point of this dynamic on the square; indeed, the only symmetric fixed point is at approximately $(0.42, 0.42)$, and it is stable.

Here the payoff function h is not a fitness function because the migration dynamic itself also affects survivability of a strategy. Specifically, one has the "decreasing returns" effect that as prevalence of a strategy increases, emigra-

---

[12] My thanks to David Kreps for suggesting this sort of example.

tion from it increases and immigration to it decreases, other things equal. The dynamics (2.8) are compatible with the fitness function $f^1(x, p, q) = \hat{x} \cdot \hat{p}(h(x, q) - 12)$.
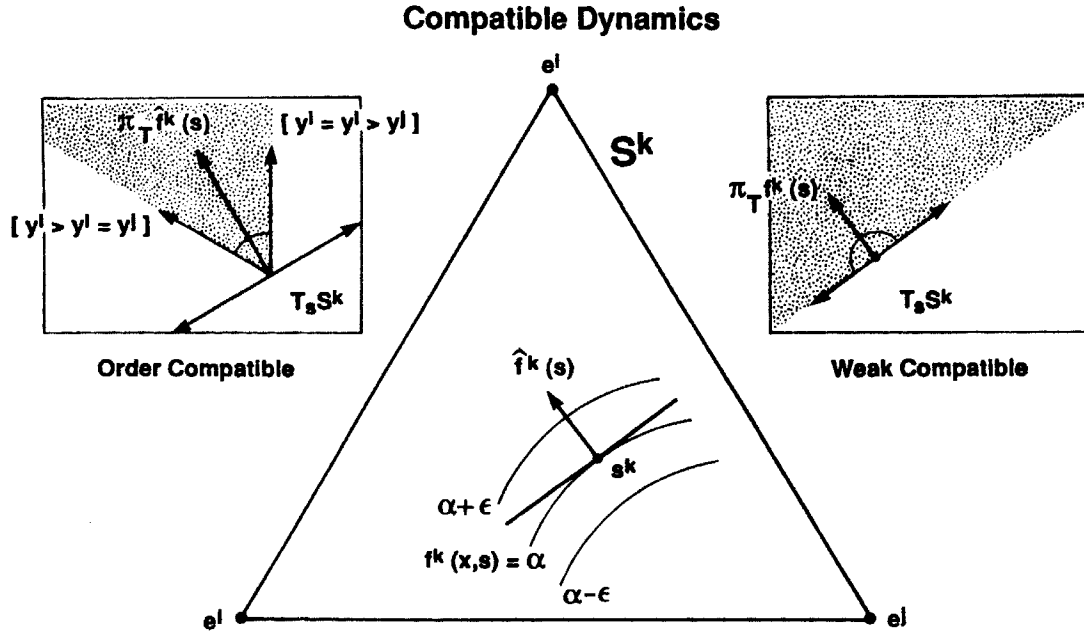
2.3. Compatibility

A given fitness function can give rise to many sorts of dynamics. Perhaps individuals change strategies when realized fitness falls below some threshold, as in Nelson and Winter (1982) or as in numerous adjustment cost models. Perhaps individuals are able to observe a sample of contemporaneous interactions and (subject to sampling error) imitate the apparently fitter strategies.[13] Perhaps rational (Bayesian) individuals gradually learn the full structure of the fitness function and the distributions of strategy choices in the populations (see Shefrin (1981) for a model with this flavor). Or perhaps, as in Crawford (1989), players employ adaptive strategies because of strategic uncertainty. As long as some sort of adjustment cost or informational imperfection or imperfect rationality precludes simultaneous jumps in behavior by a positive fraction of the population, it makes sense to describe the (co)evolution of population(s) with dynamics defined by a system of differential equations. Unfortunately in economic applications the precise specification of evolutionary dynamics seems very context-specific.

However, evolutionary dynamics are not completely arbitrary. The evolutionarily relevant aspect of a strategic interaction is not some subjective utility but rather the objective survivability or *fitness* it awards to a strategy. It is therefore axiomatic that for any dynamic compatible with a properly specified fitness function *fitter strategies should increase relative to less fit strategies*.

To formalize compatibility it is convenient to denote the set of extinct strategies in population k by the index set $I'(s^k) = \{i: s_i^k = 0\}$ and the other (nonextinct, $s_i^k > 0$) strategies by $I(s^k)$. Say the dynamic $F: S \to R^{NK}$ is *order compatible* with a fitness function f if, for each $k = 1, \ldots, K$, we have $F_i^k(s) > F_j^k(s)$ iff $f^k(e^i, s) > f^k(e^j, s)$ for all $i, j \in I(s^k)$, and $F_i^k(s) = 0$ for $i \in I'(s^k)$. The last qualification applies the ordinal relationship only to nonextinct strategies, and repeats the admissibility condition (b) that F can't revive extinct strategies.

An alternative, less stringent requirement is that there is just a positive correlation between the components of the velocity vector $\dot{s}^k = (\dot{s}_1^k, \ldots, \dot{s}_N^k)$ and those of the fitness gradient $\hat{f}^k(s) = (f^k(e^1, s), \ldots, f^k(e^N, s))$. Formally, say the dynamic $F: S \to R^{NK}$ is *weak compatible* with a fitness function f if, for all $k = 1, \ldots, K$, and $s \in S$, we have $F^k(s) = O_N$ iff $f^k(e^i, s) = \alpha$ for some $\alpha \in R^1$ and all $i \in I(s^k)$, and otherwise $F^k(s) \cdot \hat{f}^k(s) > 0$. The first part of the definition allows the tangent vector to be zero only where all nonextinct strategies are equally fit; elsewhere the tangent vector $F^k(s)$ and the fitness vector $\hat{f}^k$ make an angle of less than 90°.

---

[13] My thanks to John Riley for suggesting this interpretation.

Compatible Dynamics

FIGURE 5

Both concepts of compatibility are illustrated in Figure 5. A point $x = s^k$ in the interior of the unit 3-simplex S^k is chosen for reference, and the isoquant (α) through s^k as well as two nearby isoquants ($\alpha \pm \varepsilon$) are shown for some fitness function f^k. Order compatible dynamics will have solution curves with tangent vectors at s^k which lie inside the cone sketched on the left of Figure 5. Here $f^k(e^i, s) > f^k(e^l, s) > f^k(e^j, s)$ so the cone is open, bounded by the rays $[e^i = e^l > e^j]$ and $[e^i > e^l = e^j]$, and contains the projection π_T of $f^k(s)$ onto the tangent plane. Similarly, the diagram to the right of the simplex shows the open half-space of weak compatible admissible tangent vectors for f^k at s^k. Thus weak-compatible dynamics will have trajectories which move towards higher isoquants.[14]

In some contexts one might prefer to impose compatibility conditions on the growth rates ($\ln \cdot s_i^k$) $= \dot{s}_i^k / s_i^k$ rather than directly on the rates of change $\dot{s}_i^k$ of population fractions. Thus, let G: $S \rightarrow R^{NK}$ be differentiable and let $D(s)$ represent the $NK \times NK$ diagonal matrix with the components of the vector $s \in R^{NK}$ down the diagonal. Suppose $G^k(s)$ is perpendicular to s^k for each k, and G is compatible (order or weak) with a fitness function f. Then F: $S \rightarrow R^{NK}$, defined by $F(s) := D(s)G(s)$, will be a (n admissible) dynamic whose growth rates by construction are compatible with f. I will refer to differentiable functions G: $S \rightarrow R^{NK}$ with G^k perpendicular to s^k as *predynamics* for this reason.

A compact summary of all compatible dynamics greatly simplifies proofs of the main results. As illustrated in Figure 5, the admissible tangent vectors that are order (or weak) compatible with a given fitness function at a given state form a cone in the tangent space of S at that state. This observation suggests

[14] Of course, as $s = s(t)$ evolves, $f^k(\cdot, s)$ and its isoquants will shift.

that the technical device of cone fields, introduced in the context of exchange dynamics by Smale (1976), will be equally useful in the present context of game dynamics. In each version of compatibility (order or weak, dynamics or predynamics) one specifies for each population k and each state s in S a subset of tangent vectors (a cone Z) that are admissible and compatible in the appropriate sense. The Appendix contains formal definitions.

Two specific compatible dynamics obtained directly from the fitness function are worth mentioning. The first merely sets $F^k(s)$ equal to the orthogonal projection of $f^k(s)$ onto the tangent space of the (sub)simplex. Hence if f is bilinear this dynamic is piecewise linear, and linear on each open subsimplex. Intuitively, its trajectories head straight up the "fitness hill" on each S^k (or on the appropriate subsimplex, if some strategies are extinct in population k). We encountered this dynamic in Example 2, equations (2.4)–(2.6). I will denote it by L_f and call it the linear dynamic.

The other dynamic, denoted M_f, is analogous for growth rates. Intuitively, one renormalizes the fitness vector so it is a predynamic and then the corresponding dynamic will feature strategy growth rates equal to relative strategy fitness—that is, we obtain the Malthusian dynamic. Since M_f works with growth rates, strategies never become extinct in finite time in the trajectories it defines, while they certainly can under the L_f dynamic. The precise construction of L_f and M_f for an arbitrary fitness function f can be found in the Appendix.

2.4. Equilibrium Concepts

Two static concepts of equilibrium can be defined directly from the fitness function. Say that s is a *Noncooperative* (*or Nash*) *equilibrium point* for the fitness function f, and write $s \in NE(f)$, if $f^k(x, s) \leq f^k(s^k, s)$ for all $x \in S^k$, $k = 1, \ldots, K$. The intuition is the same as in standard game theory: the inequality says that s^k is a (weak) best reply against s for all k. Say that s is an *evolutionary stable state* for the fitness function f, and write $s \in ESS(f)$, if for each $k = 1, \ldots, K$ and $x \neq s^k \in S^k$, we have either (a) $f^k(x, s) < f^k(s^k, s)$ or (b) $f^k(x, s) = f^k(s^k, s)$ and $f^k(x, s|_k x_\varepsilon) < f^k(s^k, s|_k x_\varepsilon)$ for $\varepsilon > 0$ sufficiently small. The notation $s|_k x$ means $(s^1, \ldots, s^{k-1}, x, s^{k+1}, \ldots, s^K)$, and $s|_k x_\varepsilon$ means $(1 - \varepsilon)s + \varepsilon(s|_k x)$, i.e., a slight perturbation of s^k towards x. Of course, in the case that f is bilinear, the references to ε can be dropped from the definition.

REMARK 2.1: Obviously ESS is a refinement of NE, i.e., $ESS(f) \subset NE(f)$.

REMARK 2.2: An equivalent definition for ESS is simply that $f^k(x, s|_k x_\varepsilon) < f^k(s^k, s|_k x_\varepsilon)$ for $\varepsilon > 0$ sufficiently small. However, a two-part definition is traditional, and helps distinguish two useful cases.

REMARK 2.3: Maynard Smith's intuition, explained at length in his 1982 book, is that ESS is a stability condition. Think of x as a mutant strategy, so $s|_k x_\varepsilon$ is the state after a small invasion of population k mutants. The inequalities

defining ESS say that mutants are less fit than the dominant group (and hence presumably will disappear) either because (a) they do less well in encounters with the dominant group or (b) they do equally well in encounters with the dominant group but less well in encounters with other mutants. I will refer to an assertion that ESS is necessary or sufficient for dynamic stability as a *Maynard Smith conjecture*.

REMARK 2.4: Nash equilibrium points exist for every fitness function. The proof is a minor variant on the standard existence proof for NE in normal form games, using the Kakutani fixed point theorem.

REMARK 2.5: NE states can be characterized as those whose "extinct" strategies (i.e., i such that $s_i^k = 0$) are no more fit than nonextinct strategies ($s_i^k > 0$) and whose nonextinct strategies are equally fit. Formally, we have the following:

Let $\alpha_k := \max\{f^k(e^i, s): i = 1, \ldots, N\}$. Then s is a Nash equilibrium point iff, for each k, $f^k(e^i, s) = \alpha_k$ for all i such that $s_i^k > 0$.

The proof again is standard, with details available from the author.

REMARK 2.6: It is tempting in light of the previous remark to conjecture that ESS are the subset of NE for which extinct strategies are strictly less fit than nonextinct strategies for each population, and for which the nonextinct strategies reach peak fitness in some sense (say a negative-definiteness condition). Taylor and Jonker (1978) show that this conjecture is almost but not quite true. Following their approach, define $s^* \in S$ to be a *regular ESS* if (a) $f^k(e^i, s^*) < \alpha := f^k(s^{*k}, s^*)$ for all i such that $s_i^k = 0$, and (b) the matrix $A^k := ((\partial f^k(e^i, s^*)/\partial s_j^k))$ is negative definite on the tangent space at s^k, $\{z \in R^N: \Sigma z_i = 0, \text{ and } s_j^{*k} = 0 \Rightarrow z_j = 0\}$, for each $k = 1, \ldots, K$. Taylor and Jonker argue that every regular ESS is an ESS, and almost every ESS is regular.[15]

The dynamic equilibrium concepts are conceptually straightforward: we seek to characterize steady states, and states to which the dynamic process returns following a small perturbation. Such equilibria are defined in terms of the function F that specifies the dynamical system on S. Say that s is a *fixed point* (a.k.a. steady state or equilibrium point) for the dynamic F, and write $s \in FP(F)$, if all components of $F(s)$ are 0. Such states are steady in that $s(t) = s(0)$ for all $t \in [0, \infty)$ iff $s(0) \in FP(F)$. Say that s is an *evolutionary equilibrium* (a.k.a. stable equilibrium or persistent state) for F, and write $s \in EE(F)$, if s is a locally asymptotically stable fixed point, i.e., if s is a fixed point and it has some open neighborhood $N \subset S$ such that $s(t) \to s$ as $t \to \infty$ wherever $s(0) \in N$. Thus all states near an evolutionary equilibrium will eventually evolve towards it.

---

[15] The matrix could be only negative semi-definite, yet the inequality in part (b) of the ESS definition might still hold because of higher-order terms. The regularity condition rules out this unlikely situation. The condition will be used only in Proposition 3.4.

3. RESULTS

Given the four equilibrium concepts just defined, the theorist will instinctively ask about their relationships. The applied researcher investigating a specific fitness function f and compatible dynamic F also has a stake in the question. Evolutionary equilibria of F, as the persistent states of the model, provide the basic predictions for observables, but these equilibria may be difficult to compute directly and need not exist. The more convenient static equilibria of f take on added interest once the investigator has theoretical reassurance of their relationships to evolutionary equilibria.

For standard biological models the relationships among the equilibria are by now well known: for the Malthusian dynamic M_f of a bilinear single-population fitness function f we have

$$ESS(f) \subset EE(M_f) \subset NE(f) \subset FP(M_f);$$

see van Damme (1987) for instance. In particular, the static ESS concept is sufficient and NE is necessary for evolutionary equilibrium in standard biological models.

This section will show that the last two inclusions are quite general: for any dynamic F even weakly compatible with a given (possibly nonlinear and multi-population) fitness function f, the evolutionary equilibria of F must all be NE of f (Proposition 3.2) and all NE of f must be fixed points of the dynamic F. These results allow the researcher to confidently apply known existence results for NE and simple computational devices for FP. On the other hand, the first inclusion does not generalize as well (Proposition 3.4 and counterexamples), so the researcher can avoid possibly misleading analyses of ESS.

We begin with the preliminary result that, as Figure 5 and the choice of terminology suggest, order compatibility implies weak compatibility. The result streamlines the statement and proof of the main propositions.

PROPOSITION 3.1: *For every fitness function f, for every state s in S, and for $k = 1, \ldots, K$, the order compatible cones $Z_{oc}^k(f, s)$ for dynamics and $Z_{op}^k(f, s)$ for predynamics are subsets of the corresponding weak compatible cones $Z_{wc}^k(f, s)$ and $Z_{wp}^k(f, s)$.*

See the Appendix for a proof of this and the other propositions.

The next result confirms that a static equilibrium concept (NE) is a sufficient condition for dynamic equilibrium (FP) for any compatible dynamic. Existence of dynamic equilibrium is a corollary. The proposition also shows that the static and dynamic concepts actually coincide for interior equilibria, i.e., equilibria with no extinct strategies.

PROPOSITION 3.2: *Let the dynamic F be weak compatible with the fitness function f. Alternatively, let the predynamic G be weak compatible with f and let $F(s) = D(s)G(s)$. In either case, all Nash equilibria are fixed points and all in-*

terior fixed points are Nash equilibria, i.e., $NE(f) \subset FP(F)$ and $NE(f) \cap S^0 = FP(F) \cap S^0$.

COROLLARY 3.2.1: *Proposition 3.2 also holds with "order compatible" replacing "weak compatible."*

PROOF: Follows directly from Proposition 3.1.

COROLLARY 3.2.2: *Let the dynamic F be weak or order compatible with some fitness function f. Alternatively, let $F(s) = D(s)G(s)$ with G weak or order compatible with f. Then F has an equilibrium point, i.e., $FP(F) \neq \varnothing$.*

PROOF: Follows from Remark 2.4, Proposition 3.2, and Corollary 3.2.1.

As a complement to the previous proposition, the next result confirms that the evolutionary equilibria (EE) are always NE for a corresponding fitness function.

PROPOSITION 3.3: *Let the dynamic F be weak compatible with the fitness function f, or alternatively let $F(s) = D(s)G(s)$ and let G be a predynamic which is weak compatible with f. Then all stable equilibria of F are Nash equilibria, i.e., $EE(F) \subset NE(f)$.*

COROLLARY 3.3.1: *Proposition 3.3 also holds with "order compatible" replacing "weak compatible."*

PROOF: Follows directly again from Proposition 3.1.

REMARK 3.1: By analogy with the last part of Proposition 3.2, one migh: wonder if $EE(F)$ sometimes coincides with $NE(f)$. In particular, for $S^E = \{e^1, \ldots, e^N\}^K$ representing the extreme points of S (i.e., each population adopts a single strategy with probability 1) it is reasonable to conjecture that $NE(f) \cap S^E = EE(F) \cap S^E$ for F compatible with f, so strict NE implies ESS. It is not difficult to see that the conjecture is true if the values $f^k(e^i, s)$ are distinct (across i) for $s \in S^E$, which usually (i.e., for generic f) is the case.

REMARK 3.2: One could weaken the definition of admissibility, so that extinct strategies $i \in I'(s)$ could be revived (without perturbing the state) and only further decreases for such strategies (i.e., choice by a *negative* fraction of a population) are excluded. Weakening admissibility means reducing the number of fixed points, since $F(s) \neq 0$ is possible more often. Indeed, under the weakened definition $NE(f) = FP(F)$ and examples of unstable fixed points are harder to come by. No similar modification seems appropriate for the growth rate case, since it employs no boundary condition.

REMARK 3.3: It can be shown that Propositions 3.2 and 3.3, together with Remark 2.5 and the definitions, offer convenient computational algorithms for equilibria of evolutionary games, particularly when the fitness function is bilinear. In this case, the definition of FP provides for each subsimplex a square system of linear equations stating that the fitnesses of nonextinct strategies are equal, and that points lie in the relevant subsimplex. Solutions to the system can be tested for the additional NE, ESS, and EE properties. For details, see Friedman (1987) and the biological literature, especially Haigh (1975).

The last proposition collects positive results concerning Maynard Smith conjectures. If one is interested only in linear or Malthusian dynamics, or if one has a symmetric fitness function, then the conjectures are mostly correct. The conjectures evidently arose from consideration of a bilinear fitness function for one population and two strategies (illustrated in Figure 6 below and discussed in the last proof in the Appendix), so it is no surprise that they also hold in this case.

PROPOSITION 3.4: *Let s be a regular ESS for some fitness function f. Then:* (a) *s is an EE under Malthusian and also under linear dynamics;* (b) *if $s \in S^0$ and f is bilinear, then $ESS(f) = EE(M_f) = EE(L_f) = \{s\}$;* (c) *if $N = 2$, $K = 1$, and f is bilinear, then $ESS(f) = EE(F)$ for all weak compatible dynamics F; and* (d) *if f is symmetric, then $s \in EE(F)$ for any dynamic F which is weak or order compatible with f.*

REMARK 3.4: The symmetry condition in Proposition 3.4(d) of course is that $f(r, s) = f(s, r)$, suggesting team games. Readers acquainted with the biological literature might be confused on this point because Maynard Smith refers to the single population case $(K = 1)$ as the "the symmetric case" and to certain special cases of $K > 1$ as "the asymmetric case."[16]

REMARK 3.5: I suspect that Proposition 3.4 holds even if the ESS is not regular, but I have not been able to find a proof for part (a) when s is not interior.

I conclude this section with three counterexamples that show that ESS is neither necessary nor sufficient for the dynamic stability of a fixed point of a compatible dynamic, and that ESS and EE do not always exist. See Friedman (1987) for details.

[16] Maynard Smith applies the term "asymmetric case" to a single genetic population whose members sometimes play different "roles" (e.g., owner or intruder in a territorial contest) and have role-dependent strategy sets. The most natural way to formalize this idea in the present framework is to write an underlying general fitness function with $K =$ the number of possible roles, and then reduce it (by taking the expectation with respect to role probabilities) to a single-population fitness function defined on role-contingent strategies. See Selten (1980) for a different formal approach.
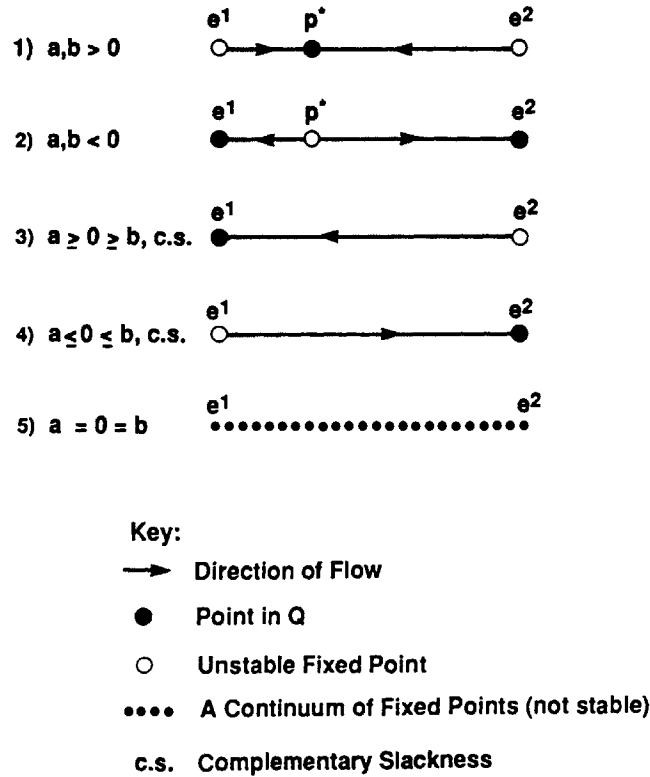
FIGURE 6

COUNTEREXAMPLE 1 (from Taylor and Jonker (1978, p. 153)): Let $K = 1$ and $N = 3$, with f bilinear and represented by the matrix

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

One readily confirms that $s = 35^{-1}$ $(15, 11, 9)$ is a NE point for A, with $\alpha = 86/35$. For $x = 35^{-1}$ $(18, 17, 0)$ we have $f(x, s) = f(s, s) = \alpha$ and $f(s, x) = 35^{-2} \cdot 2746 < f(x, x) = 35^{-2} \cdot 2773$ so $s \notin ESS(f)$. However, one can show that the eigenvalues for $DL_f(s)$ restricted to the tangent space $H(1_3)$ are $\lambda = -\frac{2}{3}(1 \pm i\sqrt{101})$ so $s \in EE(L_f)$. Taylor and Jonker show that $s \in EE(M_f)$. Hence ESS is not necessary for dynamic stability.[17]

---

[17] The biggest surprise in this counterexample is that even with Malthusian dynamics, stability is possible when certain types of mutant invasion are successful (i.e., when the ESS conditions fail). As Taylor and Jonker point out, the intuitive explanation is that ESS, which is essentially a negative-definiteness property, requires (Malthusian) trajectories to approach a fixed point directly, while the correct dynamic stability criterion allows trajectories to spiral in elliptically. An analogous confusion arose 40 years ago among economic theorists when Hicks tried to characterize the stability of multiple-market competitive equilibrium in terms of the negative definiteness of the gross substitution matrix $A = ((\partial q_i / \partial p_j))$. Paul Samuelson (1948) pointed out that the correct characterization of stability under tatonnement dynamics is that A have eigenvalues with negative real parts. In general, negative definiteness is too strong; but in the symmetric case (e.g., no income effects) the two characterizations are the same.
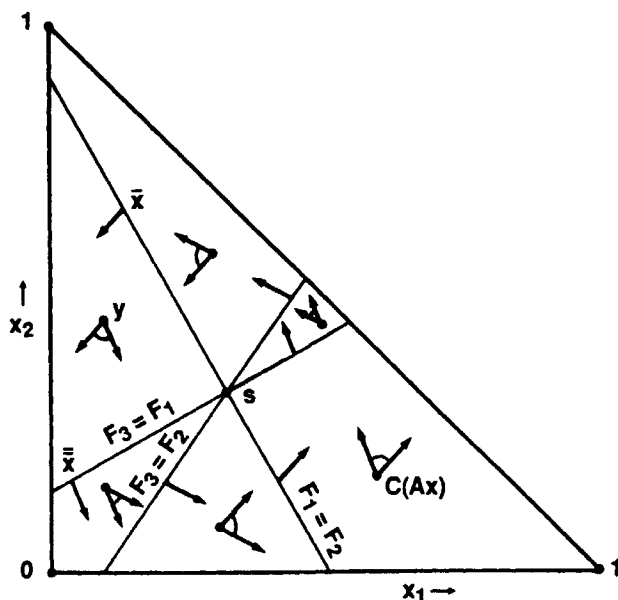
FIGURE 7

COUNTEREXAMPLE 2: Again, let f be bilinear with $K = 1$ and $N = 3$, but now set

$$A = \begin{pmatrix} -5 & 34 & -29 \\ -26 & -5 & 31 \\ 31 & -29 & -2 \end{pmatrix}.$$

The example is constructed with the convenient properties that $s = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) \in NE(f)$ and that average fitness is zero. It can be shown that A is a negative definite quadratic form on the tangent space $H(1_3)$, so s is a regular ESS. Nevertheless order compatible (not to mention weak compatible) dynamics F can be constructed for which $s \notin EE(F)$. Basically one computes the order compatible cones shown in Figure 7, and notes that by always picking tangent vectors close to the outer edge of the cones, one can intersect each of the lines at a point increasingly distant from s, i.e., one can obtain a dynamic F whose trajectories spiral rapidly outwards from s. Hence ESS is not sufficient for dynamic stability.

COUNTEREXAMPLE 3 (Sigmund, et al., reported in Zeeman (1979, pp. 483–484)): Again f is bilinear and $K = 1$ but $N = 4$ and

$$A = \begin{pmatrix} 0 & 1 & \varepsilon & 0 \\ 0 & 0 & 1 & \varepsilon \\ \varepsilon & 0 & 0 & 1 \\ 1 & \varepsilon & 0 & 0 \end{pmatrix}.$$

Consider $s = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. For $1 > \varepsilon > 0$, $\{s\} = ESS(f) = EE(L_f) = EE(M_f)$. However, for small $\varepsilon < 0$, s is a saddle point and $ESS(f) = EE(L_f) = EE(M_f) = \varnothing$. There is a stable limit cycle with a shape in S (here 3-dimensional) like the seam of a baseball.

4. DISCUSSION

4.1. *Interpretations and Extensions*

The results of the last section indicate that NE is a very useful static concept for analyzing steady states of evolutionary games. Whether one works with a specific dynamic, such as the Malthusian or the linear, or knows only that the dynamic is compatible (even just weakly compatible) with a fitness function f, the results of the last section ensure that there is always a nonempty set $NE(f)$ of points all of which are steady states and which contain all dynamically stable steady states.

An important message of the results is that ESS has no such general usefulness. There are special cases, as indicated in Proposition 3.4, in which ESS's (if they exist) will be dynamically stable steady states and therefore represent an interesting refinement of NE. However, the counterexamples show that in general the ESS property is neither necessary nor sufficient for dynamic stability of order compatible (much less weak compatible) evolutionary processes.

The evolutionary framework suggests some NE refinements with greater general appeal than ESS. Larry Samuelson (1987) shows for bilinear $K = 2$ discrete time games that convergent evolutionary processes obeying a somewhat stronger condition than order compatibility will converge to perfect NE, and that even nonconvergent processes will lead to asymptotic extinction of nonrationalizable strategies. Nachbar (1989) establishes in somewhat greater generality that the NE refinement dominance solvability (and, in some cases, weak dominance solvability) suffices to ensure the convergence of order-compatible dynamics.

Significantly sharper refinements may not be possible. Given the variety of economically plausible dynamics, it seems unreasonable to expect any static definition to pick out precisely the economically relevant equilibria. From an evolutionary perspective the quest for the ideal (static) NE refinement thus appears futile. A search for useful classes of learning and imitation dynamics seems more promising because ambiguities vanish regarding the relevant equilibrium once the dynamics (and initial conditions) are determined.

The formal framework employed here can be modified in several ways. For example, I have assumed that the evolution of the system is everywhere governed by a given compatible dynamic F, so that $\dot{s} = F(s) \in Z(f, s)$. Instead one could merely require that $\dot{s} \in Z(f, s)$, i.e., allow the learning rule to vary over time and state. Such "nondeterministic" or "differential inclusion" dynamics allow self-intersecting trajectories, but it is reasonable to conjecture that the main results will still hold, given the results of Friedman (1979) and Aubin and Cellina (1984).

A more substantial generalization of the present framework allows analysis of nonanonymous strategic interaction. The idea is that fitness could depend on *which* agents employ the various strategies, and not just on the population distributions of strategies. Hines (1980) takes a step in this direction. Building

on the biologists' distinction between monomorphic models (all individuals play the same mixed strategy except for a small subpopulation of mutants who all deviate to some other mixed strategy) and polymorphic models (all individuals play pure strategies which may differ across individuals), Hines regards states as distributions of mixed strategies (i.e., measures on S rather than points in S). He shows that the population mean evolves under Malthusian dynamics in a manner similar to that of the basic model (in essence one replaces the fitness matrix A by CA where C is the population covariance matrix), but that ESS is necessary for stability in this state space.

Schmeidler (1973) provides the static elements for general nonanonymous evolutionary games: a continuum of agents $t \in T = [0, 1]$ and Lebesgue-measurable strategy profiles x mapping T into the N-simplex. The state space is the set of all such strategy profiles and the payoff function is assumed linear in own (mixed) strategy and continuous in the state. Masso (1988) employs these static elements together with "memory strategy" dynamics adapted from Smale (1980). Both authors focus on existence of NE and emphasize the special "anonymous" case (Masso, Assumption A.2, page 61) in which the payoff depends on the state x only through its mean $\int_T x$. In the anonymous case the Schmeidler payoff function and auxiliary function reduce respectively to a fitness function f and fitness vector $\hat{f}$ as defined in Section 2.1 above. Schmeidler and Masso each show that under the anonymity assumption any NE is implementable in pure strategy choices for each player $t \in T$.

The general nonanonymous model in my view will be useful when the number of strategically distinct types of interacting agents is unbounded; otherwise the evolutionary game framework of Section 2, taking K as the bound on the number of types, seems adequate. Infinite-dimensional state spaces can also arise when there are infinitely many pure strategies, as in games of timing; for example, see Riley (1979) who also discusses the effects of finite population size in biological models.

4.2. Applications

The applications to substantial questions in economics and other social sciences will ultimately measure the value of evolutionary games. Here I will briefly and nonsystematically survey some recent applications related to evolutionary games, and then suggest some new areas of application. Several more systematic surveys are now available. Hines (1987) is a relatively recent and mathematically sophisticated summary of the biological ESS literature. Chapter 9 of van Damme (1987) provides a compact summary of game theoretic results relating to ESS and an introduction to Malthusian dynamics. Hofbauer and Sigmund (1988) is a delightful compendium of biological applications of dynamic systems theory, whose "red thread" is Malthusian dynamics. A recent brief theoretical survey can be found in Nachbar (1989).

Conlisk (1980) provides an early evolutionary approach using discrete time dynamics with lags and stochastic components. He postulates two pure strate-

gies, "optimization" and "imitation," and looks for an equilibrium population ratio of these two strategies where a quadratic loss function for nonoptimizers must balance a constant positive cost to optimization. The main result is that the fraction of optimizers in the population does not asymptotically converge to 1 as long as imitation is fitter than optimization when the population is almost all optimizers. Clearly this conclusion would also follow for *any* true order compatible dynamics.

The question of whether an investor should purchase securities information or just rely on (nearly?) efficient asset prices has a similar flavor to Conlisk's optimization vs. imitation discussion. Cornell and Roll (1981) answer this question in terms of an interior ESS of a model formally equivalent to the biologists' simplest $(K = 1, N = 2)$ static territorial conflict model. They also discuss biologists' "asymmetric" models as a device to explain the seniority system in hierarchical organizations.

James Friedman and Robert Rosenthal (1986) introduce migration dynamics for bilinear $K = 2$ games with no externalities, in order to explain deviations from NE observed in laboratory oligopolies and similar experiments. For reasons discussed in Example 2.5 above, the NE of their payoff functions are not typically dynamic equilibria.

Schofield (e.g., 1978) studies dynamic games of social choice, inspired more by nontatonnement processes than by evolutionary considerations. The idea is that at each point of policy space individual preferences and the choice rules (e.g., majority vote) define policy increments that are feasible and preferred by a relevant coalition (e.g., a majority); such increments form a cone in the tangent space at that point. Schofield shows that piecewise smooth paths (called "optimizing" curves for some reason) whose tangent vector at each point lies in the cone need not converge. Indeed, given a condition called "null dual" that is often satisfied, optimizing curves exist between most pairs of points so cycles are very common. By contrast, Smale (1976) and followers (e.g., Friedman (1979)) obtain quite strong convergence results for cone field dynamics in exchange economies.

Recently several authors have begun to model dynamic games with an evolutionary perspective on individual (not population) choice. Skyrms (1986) discusses some philosophical issues regarding rationality in games, and introduces a virtual "deliberational dynamic" (describing the adjustment of the mixed strategy contemplated by an individual) that seems formally equivalent to the linear dynamic. Binmore (1987/88) points out several other problems with rationality in the static specification of games; he emphasizes extensive-form games of imperfect information but some of the problems also arise in standard normal-form games with several NE. He suggests a dynamic resolution to the problems, and distinguishes between "eductive" processes (reminiscent of Skyrms) and "evolutive" processes in which some sort of selection mechanism operates on boundedly-rational automated players. Possibly the general non-anonymous model sketched above will be useful in pursuing the individualistic interpretation of evolutionary games.

Crawford (1989) examines laboratory games in which subjects are motivated to match the minimum or median contribution towards a public good (so the payoff is not bilinear; these are "playing the field" fitness functions). He sketches an adaptive model inspired by ESS literature (evidently a stochastic version of the linear dynamic) that seems capable of explaining the data. The paper more generally shows the promise of laboratory experiments as an empirical tool for finding specific evolutionary dynamics of economic interest.

A computerized tournament in which several game theorists submitted automated strategies for the repeated prisoners' dilemma, reported in Axelrod and Hamilton (1981), spawned an extensive and fascinating body of literature. I will not attempt a survey; see Axelrod and Dion (1988). Marks (1989) also contains a recent review that draws connections to the Holland (1980) "genetic algorithm." The tournament does not fit neatly into the present framework since it involves a supergame viz., *nonanonymous* repeated encounters in the Prisoner's Dilemma. However, Hirshleifer and Martinez-Coll (1987) re-express this supergame as an evolutionary game and show that "tit-for-tat" is not as robust a strategy as some authors have claimed.

Arthur (1989) employs a stochastic version of Malthusian dynamics arising from a learning algorithm rather than from the genetic mechanism. He shows that the limit points must be NE. Boyd and Richerson (1985) explore a variety of models that contrast and often combine genetic and social selection mechanisms. Inasmuch as economists regard a single generation of 30 years as "long-run" or "very long-run," the time scales in the latter article are exceedingly long for economic applications.

New potential economic applications of (population) evolutionary games include monetary theory, industrial organization, and international economics.[18] A recurrent question in the first field is how a medium of exchange might emerge in a laissez-faire barter economy with transactions costs—e.g., see Jones (1976) and the recent survey by Cowen and Kroszner (1987). It seems natural to address this question in terms of an evolutionary game with pure strategies representing an agent's willingness to accept specific goods he does not ultimately wish to hold; one looks for the presence of corner EE and absence of interior EE to justify a single medium of exchange.[19]

In industrial organization, one might study the "fitness" of alternative strategies such as independent entry, franchising, and merger under specific assumptions regarding product pricing.[20] One might also usefully consider $K = 2$ models for worker and firm populations, in which one could consider issues regarding unionization, apprenticeship vs. other methods of acquiring specific

[18] This list of topics is certainly not exhaustive. For example, the perspectives in Hirshleifer (1977) indicate that the economics of law is a natural area for applied models of evolutionary games.

[19] I was pleased to discover recently that Marimon, McGrattan, and Sargent (1989) take just such an approach to the issue, using dynamics based on Holland's genetic and classifier algorithms.

[20] It might be amusing to model Porter's (1980) popular "generic competitive strategies" (cost leadership, differentiation, and focus) in an evolutionary game and see whether an interior EE is possible.

skills, and compensation methods. For instance, with regard to the last issue, one wonders whether Weitzman's (1984) proposed share economy is dynamically stable, given workers' incentive not to dilute their ownership shares; one could use an evolutionary model to test the conjecture that firms which promise not to dilute might successfully invade an industry of Weitzmanian firms.

In international economics, one could consider the sectoral investment choices of entrepreneurs in autarky as a $K = 1$ model. Then, for n countries, one obtains a $K = n$ evolutionary model by coupling the autarky models under some (not necessarily free) trading regime. Such a model may be useful for discussing the dynamic impact of trade and capital restrictions on sectoral investment and output.

Another potentially fruitful avenue of investigation can be taken in conjunction with applications such as those just described or on its own: the systematic exploration of specific learning and imitation rules and their dynamic implications.

Department of Economics, University of California, Santa Cruz, CA 95064, U.S.A.

APPENDIX

A.1. *Notation and Definitions*

R^n denotes Euclidean space of finite dimension $n \geqslant 1$ with the usual basis $e^1, \ldots, e^n$, where $e^i = (0, \ldots, 1, \ldots, 0)$, the 1 being in the ith place. Vectors, denoted x, v, e^i, etc., are column vectors for purposes of matrix calculations but usually are written as row vectors as above to simplify typography. The symbol $:=$ means "is defined as." For vectors or scalars $x^1, \ldots, x^m$ and y, the notation $|_k$ refers to replacement of the kth component; e.g., if $x = (x^1, \ldots, x^m)$, then $x|_k y := (x^1, \ldots, x^{k-1}, y, x^{k+1}, \ldots, x^m)$. X^0 denotes the (relative) interior of a set $X \subset R^n$. For I a finite set, $\#I$ is the number of elements of I.

The usual *inner product* is denoted by a dot, i.e., $x \cdot y := \sum_{i=1}^n x_i y_i$ for $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n) \in R^n$. $H(x, a) := \{v \in R^n: v \cdot x = a\}$ is the $(n - 1$ dimensional) *hyperplane* with normal $x \in R^n$ translated from the origin $O_n = (0, \ldots, 0)$ along x for the distance $a \in R^1$. The corresponding *closed* and *open positive half-spaces* are $H_+(x, a) := \{v \in R^n: v \cdot x \geqslant a\}$ and $H_+^0(x, a) := \{v \in R^n: v \cdot x > a\}$. Negative half-spaces H_- are analogously defined. $H(x) := H(x, 0)$ and similarly for the half-spaces. The *positive orthant* thus is $R_+^n := \bigcap_{i=1}^n H_+(e^i)$, and the *unit simplex* is $\Delta := H(1_n, 1) \cap R_+^n$, where $1_n = (1, \ldots, 1) \in R^n$ is the *diagonal vector*.

For $s \in \Delta$, let $I(s) := \{i = 1, \ldots, N: s_i > 0\}$ and $I'(s) := \{i: s_i = 0\}$, and let $R^{I(s)}$ be the subspace of R^N spanned by $\{e^i: i \in I(s)\}$. Then s is in the subsimplex $\Delta_{I(s)} := H(1_N, 1) \cap R_+^{I(s)}$. Its *tangent space* is $H(1_N) \cap R^{I(s)}$. The *order-cone* for $x \in R^n$ is

$$C(x) := \left\{ y \in R^n: y_j < y_i \text{ iff } x_j < x_i \text{ for } i, j = 1, \ldots, n \right\}.$$

Note for later reference that $y \in C(x)$ iff $x \in C(y)$, and in particular $O_n \in C(1_n) = \{\gamma 1_n: \gamma \in R^1\}$.

A careful inspection of the definitions in the body of the paper discloses that a differentiable function $F: S \to R^{NK}$ is a dynamic which is order compatible with the fitness function f iff $F^k(s) \in C(\hat{f}^k(s)) \cap H(1_N) \cap R^{I(s^k)}$ for all $s \in S$ and all k. Hence define the order compatible cone for f at s

$$Z_{oc}^k(f, s) := C(\hat{f}^k(s)) \cap H(1_N) \cap R^{I(s^k)} \qquad\qquad (k = 1, \ldots, K).$$

Likewise, G is an order compatible predynamic if $G^k(s) \in C(\hat{f}^k(s)) \cap H(s^k) \cap R^{I(s^k)}$ so the latter expression defines the *order compatible predynamics cone for* f *at* s, $Z^k_{op}(f, s)$. For weak compatibility, one replaces $C(\hat{f}^k(s))$ by $H^0_+(\hat{f}^k(s))$ when $\hat{f}^k(s) \neq \alpha 1_{s^k}$, and by O_N otherwise. The resulting cones of weak compatible dynamics and predynamics tangent vectors are denoted Z^k_{wc} and Z^k_{wp} respectively.

Given a fitness function f, define the *linear dynamic* $L_f \colon S \to R^{NK}$ by $L_f(s) := D(1_s)(\hat{f}(s) - a(s))$, where $a(s) \in R^{KN}$ has (all N of) its k components equal to $(\#I(s^k))^{-1}\Sigma_{i \in I(s^k)} f^k(e^i, s)$, the simple average fitness of nonextinct population k strategies. Recall 1_s has zero components wherever $s^k_i = 0$ with unit components elsewhere. Obviously L_f is order compatible with f and satisfies admissibility conditions (a) and (b). However, L_f may be discontinuous on the boundaries of (sub)simplicies, so it is a slight abuse of terminology to call it a dynamic. Care is taken in the proofs involving L_f to ensure that they remain valid.

Given a fitness function f, define the *Malthusian dynamic* $M_f \colon S \to R^{NK}$ by $M_f(s) := D(s)(\hat{f}(s) - a^*(s))$, where $a^*(s) \in R^{KN}$ has its k-components (all N of them) equal to $f^k(s^k, s) = \Sigma s^k_i f^k(e^i, s)$, the population-weighted average fitness of population k strategies. It is immediate that $s \cdot (\hat{f}(s) - a^*(s)) = 0$ so $\hat{f}(s) - a^*(s)$ is indeed a predynamic which is order-compatible with f. If f is bilinear, then M_f will specify differential equations of degree 3.

A.2. Proofs

Before presenting proofs of the Propositions, I sketch the analysis behind Example 1 and Figure 1 for readers who are interested. The center point and the four corner points are clearly the only solutions to $0 = \dot{p} = \dot{q}$ in equations (2.2) and (2.3). Summarizing the equations as $(\dot{p}, \dot{q}) = F(p, q)$, one readily calculates the Jacobian

$$DF(p, q) = \begin{pmatrix} (1 - 2p)(1 - 2q) & -2p(1 - p) \\ 2q(1 - q) & (1 - 2q)(2p - 1) \end{pmatrix}.$$

Hence, at the corners ($p, q = 0$ or 1) we have

$$DF(p, q) = \pm \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix},$$

which implies that these are all saddle points. At the center point we have

$$DF(\tfrac{1}{2}, \tfrac{1}{2}) = \begin{pmatrix} 0 & -\tfrac{1}{2} \\ \tfrac{1}{2} & 0 \end{pmatrix}$$

whose eigenvalues are purely imaginary at $\pm i/2$, so up to first order the trajectories are (counterclockwise) concentric circles of period 4π. The higher order terms could stabilize or destabilize this fixed point, so further analysis is required. One employs new variables that translate the fixed point to the origin, $x = p - \tfrac{1}{2}$, $y = q - \tfrac{1}{2}$, and re-expresses equations (2.2)–(2.3) as $\dot{x} = X(x, y) := -2y(\tfrac{1}{4} - x^2)$ and $y = Y(x, y) := 2x(+y^2)$. Observe that $dx/d(-t) = -X(x, y) = X(x, -y)$ and $dy/d(-t) = -Y(x, y) = -Y(x, -y)$. Hence for initial conditions $x(0) = x_0 > 0$ and $y(0) = 0$, the solution curve satisfies $x(t) = x(-t)$ and $y(t) = -y(-t)$ for all $t \in R$. In particular, $y(t_0) = 0$ for some smallest $t_0 > 0$ (for x_0 small, $t_0 \approx 2\pi$) by the first order approximation, so the solution curve is a closed loop of period $2t_0$ which has $x = 0$ as an axis of symmetry. Similarly, $y = 0$ is also an axis of symmetry, and a simple argument shows that $x = y$ and $x = -y$ are further axes of symmetry. Alternatively, one observes that the function $H(p, q) = pq(1 - p)(1 - q)$ is constant on solution curves and has closed isoquants. Consequently the phase portrait is the same (up to time reparameterization) as the isoquant map. Re-expressing the equations in the polar coordinates $x = r\cos\theta$, $y = r\sin\theta$, one finds that $\dot{r} = 1/2 r^3 \sin 4\theta$ and $\dot{\theta} = 1/2 - r^2(1 - \cos 4\theta)$, which yield the squarish trajectories of Figure 1.

REMARK A.1: It is not always the case that the order cone of a vector lies in its own positive half-space. For instance, take $x = (1, 2, 3)$ so for $y = (-3, -2, -1)$ we have $y \in C(x)$ but $y \notin H_+(x)$. Consequently the proof of Proposition 3.1 uses the following lemma.

LEMMA A.1: *For $N \geqslant 2$, let $x, y \in H(1_N)$ and $y \in C(x)$. Then $x \cdot y > 0$ unless $x = y = O_N$.*

PROOF: Define $\|x\| := \max\{|x_i|: i = 1, \ldots, N\}$—the "sup norm," and set $U := \{x \in R^N: \|x\| = 1\}$—the unit vectors. Let $v = (1, 2, \ldots, N)$ and $Z := C(v) \cap H(1_N) \cap U$. Inspection of the extreme points of Z discloses that $\inf\{x \cdot y: x, y \in Z\}$ is achieved for x and y at the "diagonally opposite corners," viz,

$$x^* = \left(-1, \frac{1}{N-1}, \ldots, \frac{1}{N-1}\right) \quad \text{and} \quad y^* = \left(-\frac{1}{N-1}, \ldots, -\frac{1}{N-1}, 1\right).$$

But

$$x^* \cdot y^* = \frac{2}{N-1} - \frac{N-2}{(N-1)^2} = \frac{N}{(N-1)^2} > 0.$$

Note that for any $x \in R^N$, the cone $C(x)$ is isometric to $C(v)$ (possibly with N replaced by $n < N$) so the inequality is quite general. Indeed, for x, y in $H(1_N)$, if $x \neq O_N$ and $y \neq O_N$, then $x \cdot y \geqslant \|x\| \cdot \|y\| N/(N-1)^2 > 0$. If $x = O_N$ then $C(x) = \{\alpha 1_N: \alpha \in R^1\}$ so $y = \alpha 1_N$ for α such that $0 = 1_N \cdot y = \alpha 1_N \cdot 1_N = \alpha N$, i.e., $\alpha = 0$ and $y = O_N$. Since $y \in C(x)$ iff $x \in C(y)$, we also conclude that $y = O_N$ implies $x = O_N$. Hence either $x = O_N$ and $x = y = O_N$, or else $x \neq O_N$ and $x \cdot y > 0$.

REMARK A.2: The same conclusion can be established by essentially the same argument when $x, y \in H(w) \cap R^{I(w)}$ for given $w \in S^k$. One takes a w-weighted sup norm (that is, $w_i|x_i|$ replaces $|x_i|$) and replaces x^* and y^* by appropriate functions of w.

PROPOSITION 3.1: *For every fitness function f, for every state s in S, and for $k = 1, \ldots, K$, the order compatible cones $Z_{oc}^k(f, s)$ for dynamics and $Z_{op}^k(f, s)$ for predynamics are subsets of the corresponding weak compatible cones $Z_{wc}^k(f, s)$ and $Z_{wp}^k(f, s)$.*

PROOF: If $N = 1$, the state space is a singleton and the cones are all the zero vector. For given $N \geqslant 2$, fix f, s, and k. Set $y = \hat{f}^k(s)$ and let $x \in Z_{oc}^k(f, s) = C(y) \cap H(1_N) \cap R^{I(s^k)}$. If $x = O_N$ then (by definition of order-cones) it must be the case that $y = \alpha 1_{s^k}$; hence $x \in Z_{wc}^k(f, s)$. On the other hand, if $x \neq O_N$ the lemma implies that $x \in H_+^0$ so again $x \in Z_{wc}^k(f, s)$. This establishes the desired conclusion for dynamic cones. For predynamics cones one repeats the same argument using the extension of the lemma noted in Remark A.2 with $w = s^k$.   Q.E.D.

PROPOSITION 3.2: *Let the dynamic F be weak compatible with the fitness function f. Alternatively, let the predynamic G be weak compatible with f and let $F(s) = D(s)G(s)$. In either case, all Nash equilibria are fixed points and all interior fixed points are Nash equilibria, i.e., $NE(f) \subset FP(F)$ and $NE(f) \cap S^0 = FP(F) \cap S^0$.*

PROOF: Let $s \in NE(f)$ and fix k. By Remark 2.5 $\alpha := f^k(s^k, s) = f^k(e^i, s)$ for all $i \in I(s^k)$, so $\hat{f}^k(s) = \alpha 1_{s^k}$ and consequently $F^k(s) = O_N$. Alternatively, if the predynamic G is weak compatible, then the relevant cone is $Z_{wp}^k(f, s) = C(O_N) \cap H(s^k) = \{O_N\}$. In either case we conclude that $s \in FP(F)$.

For the last part of the proof, let $s \in FP(F) \cap S^0$. Then $\{O_N\} \in C(\hat{f}^k(s)) \cap H(1_N)$ for each k by the compatibility condition for F. The properties of order cones noted after their definition now imply $\hat{f}^k(s) = \gamma 1_N$ for some $\gamma \in R^1$. Alternatively if G is compatible, then $O_N = F^k(s) = D(s^k)G^k(s)$ since $s \in FP(F)$, and so $O_N = G^k(s)$ since $D(s^k)$ is invertible for $s \in S^0$. But then $\hat{f}^k(s) \in C(O_N)$ so again $\hat{f}^k(s) = \gamma 1_N$. Thus in either case we have $f^k(e^i, s) = \gamma = f^k(s^k, s)$ for all $i = 1, \ldots, N$. Since this is true for each $k = 1, \ldots, K$, we conclude from Remark 2.5 that $s \in NE(f)$.   Q.E.D.

PROPOSITION 3.3: *Let the dynamic F be weak compatible with the fitness function f, or alternatively let $F(s) = D(s)G(s)$ and let G be a predynamic which is weak compatible with f. Then all stable equilibria of F are Nash equilibria, i.e., $EE(F) \subset NE(f)$.*

PROOF: Let $s \in EE(F)$. Since $s \in FP(F)$, the last argument in Proposition 3.2 tells us that $f^k(e^i, s) = \alpha$ for all $i \in I(s^k)$ for some $\alpha \in R^1$. Suppose $\beta := \max\{f^k(e^i, s): i = 1, \ldots, N\} > \alpha$. Then set $J' := \{i: f^k(e^i, s) = \beta\}$ and $J := \{i: f^k(e^i, s) < \beta\} \supset I(s^k)$. Thus

$$C := \{r \in S: f^k(e^i, r) > f^k(e^j, r) \ \forall j \in J, \forall i \in J'\}$$

is a cone in S, which (since f is continuous) is an open neighborhood of $s \in S$. For $r \in C \cap S^0$ we have $F_i^k(r) > F_j^k(r)$ for all $i \in J'$ and $j \in J$ so $\sum_{i \in J'} \dot{r}_i^k > 0 > \sum_{j \in J} \dot{r}_j^k$ if F is compatible, or alternatively $\sum_{i \in J'} (\ln r_i^k) > 0$ if G is compatible. Hence $\sum_{i \in J'} r_i^k$ (or alternatively $\prod_{i \in J'} r_i^k$) is a positive increasing function of time for all $r(0) \in C \cap S^0$. Since this function is 0 at $r = s$, it follows that s is not stable, i.e. $s \notin EE(F)$. Consequently, the supposition that $\beta > \alpha$ is false. We conclude that

$$f(x, s) = \sum_{i=1}^{N} x_i f(e^i, s) \leqslant \alpha \sum x_i = \alpha = \sum s_i^k f(e^i, s) = f(s^k, s),$$

i.e., $s \in NE(f)$. Q.E.D.

PROPOSITION 3.4: *Let s be a regular ESS for some fitness function f. Then: (a) s is an EE under Malthusian and also under linear dynamics; (b) if $s \in S^0$ and f is bilinear then $ESS(f) = EE(M_f) = EE(L_f) = \{s\}$; (c) if $N = 2$, $K = 1$, and f is bilinear, then $ESS(f) = EE(F)$ for all weak or order compatible dynamics F; and (d) if f is symmetric, then $s \in EE(F)$ for any dynamic F which is weak or order compatible with f.*

PROOF: First suppose $s \in ESS(f) \cap S^0$, and define the function $V: S \to R^1$ by $V(r) = \frac{1}{2}(r - s) \cdot (r - s)$. Clearly V is minimized (strictly) on S at s, so it will follow that $s \in EE(L_f)$ if one can show that $\dot{V} < 0$ on S, i.e., that V is a Liapunov function for L_f (cf. Hirsch and Smale, Theorem 9.3.1, for instance). Let $A := D\hat{f}(s)$. Thus, the $NK \times NK$ matrix A has entries

$$a_{iJ}^k = \left. \frac{\partial}{\partial r_J^k} f^k(e^i, r) \right|_{r=s}.$$

For $r \in S$, let $a^k(r) := 1_N \cdot \hat{f}^k(r)/1_N \cdot 1_N = N^{-1} \sum f^k(e^i, r)$ and let $a(r) := (a^1(r)1_N, \ldots, a^k(r)1_N) \in R^{NK}$. Then up to first order for r near s we have $\dot{r} = L_f(r) \approx Ar - a(r)$, and

$$\dot{V}(r) = (r - s) \cdot \dot{r} \approx (r - s) \cdot (Ar - a(r)) = (r - s) \cdot Ar = (r - s) \cdot A(r - s).$$

The last two equalities use the facts that $r^k \cdot 1_N = 1 = s^k \cdot 1_N$ (since s and r are in the simplex) and $As = a(s)$ (since $s \in FP(L_f)$ by Proposition 3.1). Since $s \in NE(f)$ the initial hypothesis together with Remark 2.5 tells us that condition (b) of the definition of ESS holds for each k, so $rAs = sAs$ and $rAr < sAr$ for all $r \neq s \in S$. Consequently, $\dot{V}(r) \approx rAr - rAs - sAr + sAs = rAr - sAr < 0$, so $s \in EE(L_f)$. This argument can be extended as in Taylor and Jonker (1978) to establish the same conclusion if s is a regular, noninterior ESS. That is, by renumbering so that entries $i \in I(s^k)$ appear first, one can write each block A^k as an upper block triangular matrix whose lower right block is diagonal with negative entries and whose upper right block can be shown negative definite by use of the same Liapunov function employed in the proof above.

As for M_f, the proof of Zeeman (1980), or Taylor and Jonker (1978), for the case $K = 1$ can be directly extended to the case $K > 1$: one shows in essence that $V(r) = \sum_{k=1}^{K} \sum_{i=1}^{N} s_i^k \log r_i^k$ is a Liapunov function under Malthusian dynamics if s is an ESS. This establishes part (a) of the Proposition. Note that the Liapunov functions are defined on all of S^0 if $s \in S^0 \cap ESS(f)$ and the "first order approximations" in the previous paragraph are exact if f is bilinear. Hence using Theorem 9.3.2 of Hirsch and Smale one can show that there is no other point in $EE(M_f)$ or $EE(L_f)$ in this case, so (using part (a) of the Proposition) part (b) follows.

When $N = 2$ the simplexes S^k are 1-dimensional for each k, so weak and order compatibility coincide. When $K = 1$ also, the function $s(p) = pe^1 + (1 - p)e^2$ defines an isomorphism between s

in S and p in $[0, 1]$. If f is bilinear in this case then it has the matrix representation (a_{ij}), $i, j = 1, 2$. Define $a = a_{12} - a_{22}$ and $b = a_{21} - a_{11}$, and set $p^* = a/(a + b)$. It is well known that $ESS(f)$ and $EE(M_f)$ are always the same subset Q of $\{s(p^*), e^1, e^2\}$, where Q depends on the signs of a and b as indicated in Figure 6; this result can be easily verified for arbitrary compatible F by direct calculation of the fitness functions at $s(p)$.

Suppose finally that f is symmetric and s is a given regular ESS. Now f must be bilinear, so $\hat{f}(s) = As$. Since f^k is independent of r^m for $m \neq k$, it must also be independent of s^m for $m \neq k$. Thus A is a symmetric block diagonal matrix. The argument in the first paragraph of this proof establishes that A is negative definite on $H(1_N)^K$. For $V(r) := -\frac{1}{2} rAr$ we have

$$L_f(r) = L_f(r - s) = A(r - s) - a(r - s) = A(r - s) = -\text{grad}\, V(r - s),$$

i.e., $W(r) := V(r - s)$ is a potential function for L_f. Consequently, by Theorem 9.4.3 of Hirsch and Smale, the trajectories of linear dynamics cross the level surfaces of $W(r)$ orthogonally, with $\dot{W} < 0$. If F is weak compatible with f, then $F(r) \cdot L_f(r) > 0$ so W is a Liapunov function for any weak (or, by Proposition 3.1, any order) compatible dynamics F, and therefore $s \in EE(F)$.   Q.E.D.

Details for counterexample 2: Figure 7 shows the cones of (admissible and) compatible tangent vectors for f projected on the $x_1 - x_2$ plane. On the lines $F_i(x) = F_j(x)$ the cone is a single ray; e.g., at the point labelled $\bar{x}$ we have $F_3(x) > F_2(x) = F_1(x)$ since at that point $e^3 Ax > e^2 Ax = e^1 Ax$. Since $\sum F_j = 0$ by admissibility, we must have $F_1(x) = F_2(x) = -\frac{1}{2} F_3(x) < 0$, giving the indicated ray. Similarly one defines the cones at points such as $\bar{x}$ on the other lines defined by $e^i Ax = e^j Ax$, $i \neq j$. Between such lines the cone is an open 2-dimensional cone bounded by the rays defined at the adjacent lines, e.g., at the point labelled y the cone is bounded by the rays obtained from $\bar{x}$ and $\bar{\bar{x}}$. For $x \in S$, $e^1 Ax > e^2 Ax$ as

$$-5x_1 + 34x_2 - 29(1 - x_1 - x_2) > -26x_1 - 5x_2 + 31(1 - x_1 - x_2)$$

as $13x_1 + 7x_2 > 20/3$,

so the line $F_1(x) = F_2(x)$ passes through $s = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and has slope (after projection into the $x_1 - x_2$ plane) of $-13/7 \approx -1.86$. Similarly, the lines $F_1(x) = F_3(x)$ and $F_2(x) = F_3(x)$ also pass through s and have slopes $37/50 = 0.74$ and $50/43 \approx 1.16$ respectively.

Evidently, by always picking tangent vectors close to the outer edge of the cones, one can intersect each of these lines at a point increasingly distant from s, i.e., one can obtain a compatible dynamic F whose trajectories spiral rapidly outwards from s. Specifically, one can choose a partition of unity (see, e.g. Rudin (1966, pg. 40)) subordinate to the cover of S consisting of the six open regions bounded by the lines $F^i = F^j$ together with open epsilon neighborhoods of the lines. This partition of unity can be used to piece together local dynamics which take on the value of L_f at the nearest point on the *clockwise* boundary line in the regions, and take on the value of L_f at the nearest point on the line in the neighborhoods of the $F^i = F^j$ lines. For epsilon sufficiently small, such a dynamic will be arbitrarily close to the unstable dynamic whose trajectories have tangents at the outer edge of the cones.

REFERENCES

ARTHUR, W. BRIAN (1989): "A Nash-Discovering Classifier System for Finite-Action Games," Manuscript, Santa Fe Institute.

AUBIN, JEAN-PIERRE, AND ARRIGO CELLINA (1984): *Differential Inclusions*. Berlin: Springer-Verlag.

AXELROD, ROBERT, AND DOUGLAS DION (1988): "The Further Evolution of Cooperation," *Science*, 242, 1385–1390.

AXELROD, ROBERT, AND WILLIAM D. HAMILTON (1981): "The Evolution of Cooperation," *Science*, 211, 1390–1396.

BINMORE, KEN (1987 & 1988): "Modeling Rational Players," Parts I and II, *Economics and Philosophy*, 3, 179–214, and 4, 9–55.

BOYD, ROBERT, AND PETER RICHERSON (1985): *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

CORNELL, BRADFORD, AND RICHARD ROLL (1981): "Strategies for Pairwise Competitions in Markets and Organizations," *Bell Journal*, 12, 201–213.

CONLISK, JOHN (1980): "Costly Optimizers versus Cheap Imitators," *Journal of Economic Behavior and Organization*, 1, 275–293.

COWEN, TYLER, AND RANDALL KROSZNER (1987): "The Development of the New Monetary Economics," *Journal of Political Economy*, 95, 567–590.

CRAWFORD, VINCENT P. (1989): "An 'Evolutionary' Explanation of Van Huyck, Battalio, and Beil's Experimental Results on Coordination," Manuscript, Department of Economics, University of California, San Diego.

FRANK, ROBERT H. (1987): "If Homo Economicus Could Choose His Own Utility Function, Would He Want One with a Conscience?" *American Economic Review*, 77, 593–604.

FRIEDMAN, DANIEL (1987): "Evolutionary Economic Games," UCSC Working Papers in Applied Economics #156.

——— (1979): "Money-Mediated Disequilibrium Processes in a Pure Exchange Economy," *Journal of Mathematical Economics*, 6, 149–167.

FRIEDMAN, JAMES W., AND ROBERT W. ROSENTHAL (1986): "A Positive Approach to Non-Cooperative Games," *Journal of Economic Behavior and Organization*, 7, 235–251.

HAIGH, JOHN (1975): "Game Theory and Evolution," *Advances in Applied Probability*, 6, 8–11.

HINES, W. G. S. (1987): "Evolutionary Stable Strategies: A Review of Basic Theory," *Theoretical Population Biology*, 31, 195–272.

——— (1980): "Strategy Stability in Complex Populations," *Journal of Applied Probability*, 17, 600–611.

HIRSCH, MORRIS, AND STEPHEN SMALE (1974): *Differential Equations, Dynamical Systems, and Linear Algebra*. New York: Academic Press.

HIRSHLEIFER, JACK (1977): "Economics from a Biological Viewpoint," *The Journal of Law and Economics*, 20, 1–52.

HIRSHLEIFER, JACK, AND JUAN CARLOS MARTINEZ COLL (1988): "What Strategies Can Support the Evolutionary Emergence of Cooperation?" *Journal of Conflict Resolution*, 32, 367–398.

HOFBAUER, JOSEF, AND KARL SIGMUND (1988): *The Theory of Evolution and Dynamical Systems*. Cambridge: Cambridge University Press.

HOLLAND, J. H. (1980): "Adaptive Algorithms for Discovering and Using General Patterns in Growing Knowledge Bases," *International Journal of Policy Analysis and Information Systems*, 4, 245–268.

JONES, ROBERT A. (1976): "The Origin and Development of Media of Exchange," *Journal of Political Economy*, 84, 757–775.

MARKS, ROBERT E. (1989): "Niche Strategies: The Prisoner's Dilemma Computer Tournament Revisited," Working Paper 89-009, Australian Graduate School of Management, University of New South Wales.

MARIMON, RAMON, ELLEN MCGRATTAN, AND THOMAS J. SARGENT (1989): "Money as a Medium of Exchange in an Economy with Artificially Intelligent Agents," draft manuscript, Stanford University.

MAYNARD SMITH, JOHN (1982): *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.

MASSO, JORDI (1988): "Essays on Repeated Games with Continuum Player Sets," Ph.D. Dissertation, Economics Department, SUNY at Stony Brook.

NACHBAR, JOHN H. (1989): "'Evolutionary' Selection Dynamics in Games: Convergence and Limit Properties," RAND Manuscript, forthcoming in the *International Journal of Game Theory*.

NELSON, RICHARD R., AND SIDNEY G. WINTER (1982): *An Evolutionary Theory of Economic Change*. Cambridge, Massachusetts: Harvard University Press.

PITCHIK, CAROLYN, AND ANDREW SCHOTTER (1987): "Honesty in a Model of Strategic Information Transmission," *American Economic Review*, 77, 1032–1036.

PORTER, MICHAEL (1980): *Competitive Strategy*. NY: Free Press.

RAPPOPORT, ANATOL, AND MELVIN GUYER (1966): "A Taxonomy of 2×2 Games," *General Systems*, 11, 203–214.

RILEY, JOHN G. (1979): "Evolutionary Equilibrium Strategies," *Journal of Theoretical Biology*, 76, 109–123.

RUDIN, WALTER (1966): *Real and Complex Analysis*. NY: McGraw-Hill.

SAMUELSON, LARRY (1987): "Evolutionary Foundations of Solution Concepts for Finite, 2-Player, Normal-Form Games," Manuscript.

SAMUELSON, PAUL A. (1948): *Foundations of Economic Analysis*. Cambridge: Harvard U. Press.

SCHMEIDLER, DAVID (1973): "Equilibrium Points of Nonatomic Games," *Journal of Statistical Physics*, 7, 295–300.

SCHOFIELD, NORMAN (1978): "Instability of Simple Dynamic Games," *Review of Economic Studies*, 45, 575–594.

SELTEN, REINHARD (1980): "A Note on Evolutionary Stable Strategies in Asymmetric Animal Conflicts," *Journal of Theoretical Biology*, 84, 93–101.

SHEFRIN, H. M. (1981): "Games with Self-Generating Distributions," *Review of Economics and Statistics*, 48, 511–519.

SKYRMS, BRIAN (1986): "Deliberational Equilibria," *Topoi*, 5, 59–67.

SMALE, STEPHEN (1976): "Exchange Processes with Price Adjustment," *Journal of Mathematical Economics*, 3, 211–226.

——— (1980): "The Prisoner's Dilemma and Dynamical Systems Associated to Noncooperative Games," *Econometrica*, 48, 1617–1634.

TAYLOR, PETER D., AND LEO B. JONKER (1978): "Evolutionary Stable Strategies and Game Dynamics," *Mathematical Biosciences*, 40, 145–156.

VAN DAMME, ERIC (1987): *Stability and Perfection of Nash Equilibria*. Berlin: Springer-Verlag.

WEITZMAN, MARTIN (1984): *The Share Economy*. Cambridge: Harvard University Press.

ZEEMAN, E. CHRISTOPHER (1979): "Population Dynamics from Game Theory," *Proceedings of the International Conference on the Global Theory of Dynamical Systems*, Northwestern University, Evanston.