EUI WORKING PAPERS IN ECONOMICS

EUI Working Paper ECO No. 94/16

Learning and Evolution in a Heterogeneous Population

ED HOPKINS

European University Institute, Florence

EUROPEAN UNIVERSITY INSTITUTE, FLORENCE

ECONOMICS DEPARTMENT

EUI Working Paper **ECO** No. 94/16

**Learning and Evolution in a
Heterogeneous Population**

ED HOPKINS

BADIA FIESOLANA, SAN DOMENICO (FI)

Learning and Evolution in a Heterogeneous Population

Ed Hopkins*

Department of Economics

European University Institute

Florence 50016, Italy

hopkins@iuecalc.iue.it

March 25, 1994

Abstract

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

1 Introduction

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

There are obvious similarities between the properties of adaptive learning and evolutionary dynamics. Typically, both are concerned with the development of the distribution of strategies within some large population[1]. As Cabrales and Sobel [3] show, evolutionary dynamics under certain conditions can be "consistent with adaptive learning" in the sense of Milgrom and Roberts [13]. But this is only a condition on the asymptotic behaviour of a selection or learning process. In the short run, although "consistent", different processes may behave quite differently. In particular, while selection dynamics are typically smooth functions of current strategy distributions, under fictitious play or Cournotian dynamics, where players make best responses to previous play(s) of opponents, there can be discontinuous jumps in play. Convergence to mixed strategies is in particular troublesome (for example, see Fudenberg and Kreps [7], Jordan [10]). Here it is shown that if one aggregates such behaviour across a large population, smoothness is obtained.

The standard evolutionary dynamic framework assumes that agents compete in some game and then reproduce according to the success they obtain. Here I make the (strong) assumption that the population is randomly matched an infinite number of times in each "generation" to play the game.

---

[1]Some papers in the first camp include Milgrom and Roberts [13], Kandori et al. [11], Young [18]; in the second, Nachbar [14] , Samuelson and Zhang [15].

1

The population is heterogeneous in that some agents learn. At the end of their "lifespan" agents reproduce according to the success of the strategies they develop, or, to be precise, according to the limit of this learning process. Thus, there are two mechanisms that can change the mix of strategies in the population. Agents can change their own strategies, a "learning" process, and an evolutionary mechanism also chooses between different agents, the "selection" process.

The combination of the two has quite different implications for the stability of equilibrium than each considered in isolation. We show that the distribution of strategies in the population converges to Nash frequencies for all 2×2 asymmetric games and also for a famous 3×3 game first proposed by Shapley [16] in 1964. Shapley's original pessimistic result has been confirmed and generalised by more recent research, (Jordan [10]). It is therefore particularly striking that, even given the particular assumptions of this model, that a population can converge to the Nash equilibrium of such a game.

2 Learning and Selection

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

tive representation, and define the boundary as all states where at least one type has zero representation. The symbol "." indicates multiplication by a transpose, and the notation $(A\mathbf{y})_i$ indicates the ith element of the vector in parentheses.

The basic model is modified here by the addition of a type capable of inheriting rules more complex than simply to play a fixed strategy[2]. Thus, we now have $n+1, m+1$ types, and we work in $S_{n+1} \times S_{m+1}$. We can think of each generation being divided into an infinite number of subperiods $(0, 1, ..., s, ...)$. As a reminder, the selection process operates between generations, the learning process within generations. We assume that the $n+1$th and $m+1$th type adjust their strategies so that they play what is an optimum response to the strategy of their previous opponent: the "best-response" or Cournotian dynamic. Similar behavioural hypotheses have been employed in recent learning literature (for example, Milgrom and Roberts [13]; Kandori et al. [11]; Young [18]), but here the implementation is particularly simple. Agents do not need to know anything about the overall distribution of strategies in the population or to have a memory longer than one subperiod. Yet, as we will see, this is enough to ensure convergence to Nash equilibrium in a large class of games.

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

The probability that an individual of type $n+1$, in population 1, meets an individual of type j in the second population is y_j. There is also a probability $q_j y_{m+1}$ of meeting an individual of type $m+1$ currently playing strategy j.

---

[2]Banerjee and Weibull [1], Stahl [17] employ a similar maneouvre.

3

In either case, faced with an opponent playing strategy j, the individual will play in the next subperiod a strategy which is a best reply to j. Thus, *within each generation*, each p_i evolves according to a mapping $S_m \to [0, 1]$

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

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

$$q_j(s+1) = \sum_{i \in I_j} \frac{x_i}{\mu_i} + \sum_{i \in I_j} \frac{x_{n+1}p_i(s)}{\mu_i} \tag{2}$$

Lemma 1 *If, at time t, $1 > x_{n+1}(t), y_{m+1}(t)$, the learning process represented by equations (1), (2) has an unique fixed point $\mathbf{p}^*, \mathbf{q}^* \in S_n \times S_m$.*

Though they change between generations, within each generation the population proportions $\mathbf{x}, \mathbf{y}$ are fixed and are therefore constants for (1), (2). Consequently, the equations are simple, linear difference equations. In equilibrium, writing them in matrix form:

$$\mathbf{P} = \mathbf{x}_0 + X_1 \mathbf{P} = (\mathbf{I} - X_1)^{-1}\mathbf{x}_0$$

where $\mathbf{P} = (p_1, ..., p_{n-1}, q_1 ..., q_{m-1})$. As $\sum p = \sum q = 1$ we can eliminate the nth and mth equations, rewriting p_n as $1 - \sum_1^{n-1} p_i$, q_m as $1 - \sum_1^{m-1} q_j$. By inspection of (1), (2), it is possible to see that $(\mathbf{I} - X_1)$ can be partitioned in the following manner:

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

Each column of X_{12} and X_{21} sums to y_{m+1} and x_{n+1} respectively. Thus $(\mathbf{I} - X_1)$ is singular if and only if neither X_{12} and X_{21} are linearly independent

4

of **I**, which can only be the case if $x_{n+1} = y_{m+1} = 1$. Otherwise, there is a unique fixed point, $\mathbf{P}^*(\mathbf{x}, \mathbf{y}) = (\mathbf{p}^*, \mathbf{q}^*)$. This solution will be a function of $(\mathbf{x}, \mathbf{y})$, continuous on the interior of $S_{n+1} \times S_{m+1}$. The exact weights depend entirely on the value of the x_i, y_i and not on the value of $\mathbf{p}, \mathbf{q}$ at the beginning of the learning process. Furthermore, the sufficient condition for the existence of an unique fixed point is also a sufficient condition for convergence.

Lemma 2 *If, at time t, $1 > x_{n+1}(t), y_{m+1}(t)$, the learning process converges to its unique fixed point.*

(1), (2) represent a system of $n + m$ linear first order difference equations. The x_i, y_i are constant within each generation, and therefore are constants for (1), (2). In particular, the coefficients on the variables $\mathbf{p}, \mathbf{q}$ on the right hand side are the $x_{n+1}/\mu_i, y_{m+1}/\mu_j$, the sum of which in each equation have an upper bound in value of either x_{n+1} or y_{m+1}. By the elementary theory of difference equations if this sum is less than unity for all equations, so are all the roots of the dynamic system.□

It is worth remarking that here convergence is not convergence in empirical frequencies, a notion of convergence that has been forcefully criticised in the recent literature (Young [18], Fudenberg and Kreps [7], Jordan [10]). In this case, one does not have to take a time average. As the limit approaches, an outside observer would see strategies actually being played at (close to) equilibrium frequencies.

I make the assumption that payoffs during the learning process do not affect the rate of reproduction. Rather it is the limit of the learning process, denoted $(\mathbf{p}^*, \mathbf{q}^*)$ which determines reproductive fitness. This construction has some analytic convenience: if one assumes only a finite number of plays each period, the values of $\mathbf{p}, \mathbf{q}$ will be dependent on their (arbitrary) initial values. We would have to make further assumptions about how much of the behaviour learnt within a period is transmitted between the generations. For example, we could assume that each generation starts from scratch: at the beginning of each period $\mathbf{p}(0), \mathbf{q}(0)$ are randomly determined. That is, "children" learn nothing from their "parents". Or we can assume that the initial values are some function of play by the previous generation. However, using the limit, the value of $(\mathbf{p}^*, \mathbf{q}^*)$ will be the same in either case. This procedure is in any case defensible on the grounds that as (1), (2) are convergent, after

a "large" number of plays the process will be arbitrarily close to the limit[3].

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

$$\pi_{xi} = (Aw)_i, \ \pi_{yj} = (Bz)_j, \tag{3}$$

and for the learners,

$$\pi_{xn+1} = \mathbf{p}^* \cdot A\mathbf{w}, \ \pi_{ym+1} = \mathbf{q}^* \cdot B\mathbf{z} \tag{4}$$

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

$$x_i(t+1) = f_{xi}(\mathbf{x}, \mathbf{y}) = x_i(t)\frac{\pi_{xi}}{\mathbf{z} \cdot A\mathbf{w}}, \quad y_j(t+1) = f_{yj}(\mathbf{x}, \mathbf{y}) = y_j(t)\frac{\pi_{yj}}{\mathbf{w} \cdot B\mathbf{z}} \tag{5}$$

or, taking the limit, as generations become arbitrarily short:

$$\dot{x}_i = F_{xi}(\mathbf{x}, \mathbf{y}) = x_i(t)(\pi_{xi} - \mathbf{z} \cdot A\mathbf{w}), \quad \dot{y}_j = F_{yj}(\mathbf{x}, \mathbf{y}) = y_j(t)(\pi_{yj} - \mathbf{w} \cdot B\mathbf{z}) \tag{6}$$

where $\mathbf{z} \cdot A\mathbf{w}, \mathbf{w} \cdot B\mathbf{z}$ are the average payoffs for the two populations. Inspection of (6) shows that this continuous selection mechanism has the following important property:

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

If we impose the condition that a_{ij}, b_{ji} and hence π_{xi}, π_{xj}, are strictly positive[4] for all i, j, invariance will also hold for the discrete dynamic f. Given that $\mathbf{p}^*, \mathbf{q}^*$ are themselves functions of the frequencies of types in the

---

[3]Compare Harley's assumption (e): "The learning period is short compared to the subsequent period of stable behaviours" [6, p613].

[4]Any game matrix with negative payoffs can be transformed by the addition of a positive constant. This will not change the best response structure or Nash equilibria but may change the qualitative behaviour of the discrete replicator dynamics. See Cabrales and Sobel [3] for a discussion of the issues involved.

population, fitnesses will not be linear in $\mathbf{x}, \mathbf{y}$ - a usual assumption of the replicator dynamics - and perhaps not even be defined when x_{n+1} and y_{m+1} are equal to one. However, by Lemmata 1 and 2, fitnesses are continuous functions of $\mathbf{x}, \mathbf{y}$ elsewhere. This, combined with invariance implies that from any fully-mixed initial conditions, (that is, $x_i > 0, i = 1....n + 1$ and $y_j > 0, j = 1...m + 1$), the learning process converges, and fitnesses are defined, for all $t \in \mathbf{R}$. Thus while both f and F are not continuous on all of $S_{n+1} \times S_{m+1}$ they are continous on its interior. In other words, both f and F possess a limit even along a dynamic path with an acculmulation point on the boundary of $S_{n+1} \times S_{m+1}$, even if that limit may be path-dependent.

3 Equilibrium

Equilibrium in this model consists of a population distribution which is a rest point for both selection and learning processes. That is, a state of the system where the limit of the learning process is such that all types present in the population earn the same average payoff. In the standard evolutionary model, that is, in the absence of the learners, under the selection dynamics defined by (5) or (6), denote the rest points for the game G in the interior of $S_n \times S_m$, $(\mathbf{x}^*, \mathbf{y}^*)$. It is well known that such rest points are Nash equilibria (Hofbauer and Sigmund [8], Nachbar [14])[5]. For the extended game, the conditions for an interior rest point under the selection dynamics are

$$\pi_{x1} = ... = \pi_{xn+1}, \ \pi_{y1} = ...\pi_{ym+1} \tag{7}$$

Furthermore, as $(\mathbf{x}, \mathbf{y})$ are both constant if (7) holds, the limit for the learning process is also unchanging across all subsequent generations. The consequent distribution of strategies is a Nash equilibrium. Comparison of equations (3), (4), reveal that any values of $(\mathbf{x}, \mathbf{y})$ that satisfy the above condition (7), also satisfy $x_i + p_i^* x_{n+1} = x_i^*$, and $y_j + q_i^* y_{m+1} = y_j^*$. That is, it is a Nash equilibrium for the original game G in the sense that an outside observer would see, as the learning process reached its limit, strategies being played with the Nash equilibrium frequencies, $(\mathbf{x}^*, \mathbf{y}^*)$. Note that for each population there is now one less independent equation than there are independent variables. This means that any isolated equilibrium of the original game in the interior

---

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

7

of $S_n \times S_m$ will be represented by a continuum of fixed points in the interior of $S_{n+1} \times S_{m+1}$.

Furthermore, we can show that for all 2×2 games the system will converge to a Nash equilibrium. This is unusual in that replicator dynamics do not converge for asymmetric games with an unique mixed strategy (Hofbauer and Sigmund [7, pp141-143]). (8) gives a generalised 2×2 game.

$$A = \begin{array}{|c|c|} \hline a_{11} & a_{12} \\ \hline a_{21} & a_{22} \\ \hline \end{array} \qquad B = \begin{array}{|c|c|} \hline b_{11} & b_{12} \\ \hline b_{21} & b_{22} \\ \hline \end{array} \qquad (8)$$

Define $a_1 = a_{12} - a_{22}, a_2 = a_{21} - a_{11}; \ b_1 = b_{12} - b_{22}, b_2 = b_{21} - b_{11}$. If $a_1 a_2 < 0$ then one of the first population's strategies dominates the other. The learners will only play the dominant strategy. Their frequency in the population will grow at the same rate as the type representing the dominant strategy, which will always be positive. Eventually only that strategy will be played in first population. If $a_1 a_2 = 0$ there is either weak dominance or complete indifference. Strategy distributions either remain unchanged, or one strategy will be played by all the population. A similar analysis can be applied for games where $b_1 b_2 \leq 0$.

If $a_1 a_2 > 0$ and $b_1 b_2 > 0$ then there is a mixed Nash equilibrium where the first strategy of each population are represented with frequencies $(b, a) = (\frac{b_1}{b_1+b_2}, \frac{a_1}{a_1+a_2})$ respectively. The interesting case is when $a_1 b_1 < 0$, as in this case the standard evolutionary dynamics do not converge. However, the addition of an arbitrarily small initial population of learners is enough to stabilise the dynamics.

Proposition 1 *If $a_1 b_1 < 0$, then the mixed equilibrium is asymptotically stable and attracts all other points on the interior of $S_3 \times S_3$.*

Proof: We give the proof for the continuous time case. Define $V_1 = x_1^{b_1} x_2^{b_2} y_1^{-a_1} y_2^{-a_2}$

$$\dot{V}_1 = V_1[(b_1 - (b_1+b_2)z_1)((A\mathbf{w})_1 - (A\mathbf{w})_2) - (a_1 - (a_1+a_2)w_1)((B\mathbf{z})_1 - (B\mathbf{z})_2)] = 0$$

or in other words V_1 is a constant of motion. All orbits on the interior of $S_{n+1} \times S_{m+1}$ will be level curves of V_1. These are convex curves which pass through the continuum of equilibria defined by $(x_1 + p_1 x_3, y_1 + q_1 y_3) = (b, a)$. This implies that all orbits either flow toward the interior equilibrium, or

8

flow toward a boundary where x_3 and/or y_3 are zero. However, without loss of generality, take a_1 to be negative, note that in this case (1), (2) imply that $\mathbf{p} = \mathbf{w}$ and that $\mathbf{q} = 1 - \mathbf{z}$ and define $V_2 = x_1^a x_2^{1-a} x_3^{-1} y_1^{1-b} y_2^b y_3^{-1}$, which has its maximum on such a boundary.

$$\dot{V}_2 = V_2[(a - w)((A\mathbf{w})_1 - (A\mathbf{w})_2) + (z - b)((B\mathbf{z})_1 - (B\mathbf{z})_2)] \leq 0$$

Evidently, orbits flow away from the boundary to the equilibrium, which must attract all the interior.□

If $a_1 b_1 > 0$ then this interior equilibrium is a saddle. Similar arguments to those employed in Proposition 1, can be used to show that in this case, the system behaves in much the same way as standard evolutionary dynamics and flows toward the Nash equilibria located on the boundaries of the simplex.

4   A 3×3 Example

The famous example given by Shapley [16] to demonstrate non-convergence of fictitious play is shown in (9). The only Nash equilibrium of this game is interior, where both row and column play each of their strategies with equal probability. Interior (mixed) equilibria of asymmetric games are never asymptotically stable under the replicator dynamics. Thus this game does not converge for the replicator dynamics, just as it does not for fictitious play. However, under this modified system this game converges to the unique Nash equilibrium.

$$A = \begin{array}{|c|c|c|} \hline 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 1 \\ \hline \end{array} \quad B = \begin{array}{|c|c|c|} \hline 0 & 0 & 1 \\ \hline 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline \end{array} \tag{9}$$

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

$$\mathbf{p}(s + 1) = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} + y_4 \mathbf{q}(s), \quad \mathbf{q}(s + 1) = \begin{pmatrix} x_3 \\ x_1 \\ x_2 \end{pmatrix} + x_4 \begin{pmatrix} p_3(s) \\ p_1(s) \\ p_2(s) \end{pmatrix} \tag{10}$$

This is a system of six linear difference equations. By Lemma 2 we know that the fixed point of this system is the limit of the learning process. This

9

can be calculated using standard methods. It would be possible to eliminate $\mathbf{p}, \mathbf{q}$ by substitution using these results. However, it is easier to work in the other direction. We construct $z_i(t) = x_i(t) + x_4(t)p_i^*(t), w_j(t) = y_j + y_4(t)q_j^*(t); \; i, j = 1, 2, 3$, where z_i is the total number of the first population playing strategy i, and w_j is the total number of the second playing strategy j. Note that (10) here implies that $\mathbf{z}(t) = (q_2^*(t), q_3^*(t), q_1^*(t))$, that $\mathbf{w} = \mathbf{p}$ and that $\mathbf{q} \cdot B = \mathbf{z} \cdot A$. There is an interior equilibrium for this system: the plane such that $x_1 = x_2 = x_3$, $y_1 = y_2 = y_3$, which we denote $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$. In such an equilibrium, (10) in turn implies that $\mathbf{p} = \mathbf{q} = (1/3, 1/3, 1/3)$. I now prove that the limit point of all solutions under f, given fully-mixed initial conditions, is on this plane (normally for the discrete dynamics the interior equilibrium is a repellor).

Proposition 2 *The plane of equilibria* $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ *under f attracts all other points on the interior of $S_4 \times S_4$.*

Proof: Define $V_3(\mathbf{x}, \mathbf{y}) = x_4 y_4$. Given that $x_4(t+1) = x_4(t)\dfrac{\mathbf{p} \cdot A\mathbf{w}}{\mathbf{z} \cdot A\mathbf{w}}$, and that $y_4(t+1) = y_4(t)\dfrac{\mathbf{q} \cdot B\mathbf{z}}{\mathbf{w} \cdot B\mathbf{z}}$, it follows that $V_3(t+1) - V_3(t) > 0$ if and only if

$$\mathbf{p} \cdot A\mathbf{w} \; \mathbf{q} \cdot B\mathbf{z} - \mathbf{z} \cdot A\mathbf{w} \; \mathbf{w} \cdot B\mathbf{z} = \mathbf{w} \cdot \mathbf{w} \; \mathbf{z} \cdot \mathbf{z} - \mathbf{z} \cdot \mathbf{w} \; \mathbf{w} \cdot \mathbf{z}^* > 0 \tag{11}$$

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

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

It follows that $V_3(t+1) \geq V_3(t)$ with equality only at $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$. $V_3(\mathbf{x}, \mathbf{y})$ is therefore a strict Liapunov function on all of the interior of $S_4 \times S_4$. $\square$

$$A = B = \begin{array}{|l|c|c|c|} \hline \text{rock} & b & a & c \\ \hline \text{scissors} & c & b & a \\ \hline \text{paper} & a & c & b \\ \hline \end{array} \quad a > b > c \tag{12}$$

We consider briefly two other examples. The first (12) is the familiar **rock-scissors-paper** game, the second (13) is a game proposed by Dekel and Scotchmer (1992). They show that the **dumb** strategy survives in the limit under the discrete replicator dynamics although it is never a best response and therefore not rationalizable. **rock-scissors-paper** is well-known as a problem game. While it does converge for fictitious play, it does so only

10

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

$$A = B = \begin{array}{|l|c|c|c|c|}
\hline
\textbf{rock} & 1 & 2.35 & 0 & 0.1 \\
\hline
\textbf{scissors} & 0 & 1 & 2.35 & 0.1 \\
\hline
\textbf{paper} & 2.35 & 0 & 1 & 0.1 \\
\hline
\textbf{dumb} & 1.1 & 1.1 & 1.1 & 0 \\
\hline
\end{array} \tag{13}$$

5 Discussion

Games such as (9) cause problems for conventional models because they possess cycles of best responses. One might think that random perturabtion, for example, trembles or mutations, would also break up these deterministic cycles. However, this is not the case (see for example, [18]). The fundamental reason that this model gives qualitatively different behaviour is that there are two distinct processes determining the change in the distribution of strategies, working at different speeds. By changing strategies, the learners anticipate the next stage of the cycle and "damp" the non-convergent tendencies of the original model. The dependence is two-way. Without the non-learners, the best-response process would not converge for this game.

One might argue that the simple learning rule considered here would be displaced by more sophisticated behaviour. For example, Harley [7] claims that for a learning rule to be evolutionary stable it must be a "rule for ESSs". That is, it must be able to lead the population to the evolutionary stable strategy (ESS) in one generation. However, some doubts have been cast on Harley's model and methodology (Maynard Smith et al., [12]; Houston and Sumida, [9]). The latter paper raises a further point. "Games against nature", or in economic terms, single-agent optimisation problems, represent a very different learning environment to a situation where there is strategic interaction with other agents. That is an argument that finds support in more recent work ([1], [2], [17]). There is no claim that the learning rule considered here is the "correct" one. However, there is also no strong evidence

11

that evolution will select for more complex or sophisticated behaviour in a strategic environment.

Learning and evolution are ostensibly similar processes. However, while evolution is defined at the level of a population, learning is carried out by individuals. Crawford [4] demonstrates that even when agents' learning is modelled in a similar manner to the replicator dynamics, an aggregation of their behaviour does not have the same properties in terms of stability as evolutionary dynamics. Similarly, in this paper even the most elementary learning behaviour gives increased stability when considered at the level of the population. This opens up the possibility of further research about the aggregate properties of populations where a number of different classes of behaviour are present.

References

1. A. Banerjee and J. Weibull, Evolutionary selection and rational behavior, forthcoming in A. Kirman and M. Salmon eds. "Rationality and Learning in Economics," Blackwell, Oxford, 1994.

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

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

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

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

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

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

8. J. Hofbauer and K. Sigmund, "The Theory of Evolution and Dynamical Systems," Cambridge University Press, Cambridge, 1988.

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

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

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

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

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

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

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

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

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

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

EUI
WORKING
PAPERS

EUI Working Papers are published and distributed by the
European University Institute, Florence

Copies can be obtained free of charge
– depending on the availability of stocks – from:

The Publications Officer
European University Institute
Badia Fiesolana
I-50016 San Domenico di Fiesole (FI)
Italy

Please use order form overleaf

Publications of the European University Institute

Economics Department Working Paper Series

To   Department of Economics **WP**
European University Institute
Badia Fiesolana
I-50016 San Domenico di Fiesole (FI)
Italy

From   Name . . . .

            Address. . . . .

             . . . .

             . . . .

             . . . .

             . . . .

(Please print)

☐ Please enter/confirm my name on EUI Economics Dept. Mailing List
☐ Please send me a complete list of EUI Working Papers
☐ Please send me a complete list of EUI book publications
☐ Please send me the EUI brochure Academic Year 1994/95

Please send me the following EUI **ECO** Working Paper(s):

No, Author   . . . .
Title: . . . .
No, Author   . . . .
Title: . . . .
No, Author   . . . .
Title: . . . .
No, Author   . . . .
Title: . . . .

Date . . . .   Signature . . . .

Working Papers of the Department of Economics
Published since 1993

ECO No. 93/1
Carlo GRILLENZONI
Forecasting Unstable and Non-Stationary
Time Series

ECO No. 93/2
Carlo GRILLENZONI
Multilinear Models for Nonlinear Time
Series

ECO No. 93/3
Ronald M. HARSTAD/Louis PHLIPS
Futures Market Contracting When You
Don't Know Who the Optimists Are

ECO No. 93/4
Alan KIRMAN/Louis PHLIPS
Empirical Studies of Product Markets

ECO No. 93/5
Grayham E. MIZON
Empirical Analysis of Time Series:
Illustrations with Simulated Data

ECO No. 93/6
Tilman EHRBECK
Optimally Combining Individual
Forecasts From Panel Data

ECO NO. 93/7
Víctor GÓMEZ/Agustín MARAVALL
Initializing the Kalman Filter with
Incompletely Specified Initial Conditions

ECO No. 93/8
Frederic PALOMINO
Informed Speculation: Small Markets
Against Large Markets

ECO NO. 93/9
Stephen MARTIN
Beyond Prices Versus Quantities

ECO No. 93/10
José María LABEAGA/Angel LÓPEZ
A Flexible Demand System and VAT
Simulations from Spanish Microdata

ECO No. 93/11
Maozu LU/Grayham E. MIZON
The Encompassing Principle and
Specification Tests

ECO No. 93/12
Louis PHLIPS/Peter MØLLGAARD
Oil Stocks as a Squeeze Preventing
Mechanism: Is Self-Regulation Possible?

ECO No. 93/13
Pieter HASEKAMP
Disinflation Policy and Credibility: The
Role of Conventions

ECO No. 93/14
Louis PHLIPS
Price Leadership and Conscious
Parallelism: A Survey

ECO No. 93/15
Agustín MARAVALL
Short-Term Analysis of Macroeconomic
Time Series

ECO No. 93/16
Philip Hans FRANSES/Niels
HALDRUP
The Effects of Additive Outliers on Tests
for Unit Roots and Cointegration

ECO No. 93/17
Fabio CANOVA/Jane MARRINAN
Predicting Excess Returns in Financial
Markets

ECO No. 93/18
Iñigo HERGUERA
Exchange Rate Fluctuations, Market
Structure and the Pass-through
Relationship

ECO No. 93/19
Agustín MARAVALL
Use and Misuse of Unobserved
Components in Economic Forecasting

ECO No. 93/20
Torben HOLVAD/Jens Leth
HOUGAARD
Measuring Technical Input Efficiency for
Similar Production Units:
A Survey of the Non-Parametric
Approach

＊＊＊