

Learning, Matching, and Aggregation

Ed Hopkins*

*Department of Economics, University of Edinburgh, Edinburgh EH8 9JY,
United Kingdom*

Received July 18, 1995

Fictitious play and stimulus–response/reinforcement learning are examined in the context of a large population where agents are repeatedly randomly matched. We show that the aggregation of this learning behavior can be qualitatively different from learning at the level of the individual. This aggregate dynamic belongs to the same class of simply defined dynamic as do several formulations of evolutionary dynamics. We obtain sufficient conditions for convergence and divergence which are valid for the whole class of dynamics. These results are therefore robust to most specifications of adaptive behavior. *Journal of Economic Literature* Classification Numbers: C72, D83. © 1999 Academic Press

Key Words: games; fictitious play; reinforcement learning; evolution

1. INTRODUCTION

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

*I would like to thank all the people at GREQAM, Marseille, for their hospitality while writing this paper. Alan Kirman, Ken Binmore, Tilman Börgers, Larry Samuelson, Dan Friedman, an anonymous referee, and, in particular, Josef Hofbauer, have given many helpful comments. E-mail: e.hopkins@ed.ac.uk.



There are a number of potential responses. One adopted by Binmore and Samuelson (1997) is to devise a learning scheme which approximates the dynamics generated by evolution. Thus the results of evolutionary game theory could be recreated by learning. Another is to generalize the evolutionary dynamics by abandoning particular functional forms and looking at wide classes of dynamics which satisfy “monotonicity” or “order compatibility” (Nachbar, 1990; Friedman, 1991; Kandori *et al.*, 1993). The hope is that even if learning behavior is not identical to evolution, it is sufficiently similar to fall within these wider categories. However, in this paper, a different approach is taken. Rather than designing learning models to suit our purposes, we examine two existing models of learning behavior current in the literature. This is done in the context of a large random-mixing population.

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

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

More recently, fictitious play has again attracted interest, this time as a means of modeling learning.¹ This, however, is an interpretation that is problematic. The positive results noted above are qualified by the re-

¹Some of the many to have considered fictitious play or similar processes are Canning (1992), Fudenberg and Kreps (1993), Jordan (1993), Milgrom and Roberts (1991), Monderer and Shapley (1996), and Young (1993).

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

There are other models of learning in games. We can identify a class of learning rules as being based on gradient algorithms. The behavior postulated is perhaps even more naive than under fictitious play.² Indeed, these models were first developed by psychologists and animal behaviorists for nonstrategic settings. More recently, they have been applied to game theory by Harley (1981), Crawford (1985, 1989), Börgers and Sarin (1997), and Roth and Erev (1995). Unlike fictitious playlike processes, agents do not play a single pure strategy which is a best reply; agents play a mixed strategy. If a strategy is successful the probability assigned to it is increased, or, in the terminology of psychologists, the “behavior is reinforced.” Thus such models are sometimes called “learning by reinforcement” or “stimulus–response learning.” As these models’ other name “gradient” suggests, behavior is meant to climb toward higher payoffs. Adjustment is therefore slower and smoother than under fictitious play. However, the results obtained are not notably more positive. Crawford (1985) shows, for example, that all mixed-strategy equilibria are unstable. Crawford (1989) finds a similar result in a model where gradient learning is aggregated across a large population.

As we will see, however, Crawford’s result does not tell the whole story and aggregation does help with many of the problems outlined above. Fudenberg and Kreps (1993), in fact, propose the idea of a random-mixing population of players as a justification for the myopia of fictitious play-like

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

learning processes. If there is sufficient anonymity such that each player cannot identify his opponent and sufficient mixing, each player has a sequence of different opponents, then players may have little incentive to develop more sophisticated strategies. A population of players also offers a different interpretation of mixed-strategy equilibrium. The distribution of strategies in the population as a whole mimics a mixed-strategy profile. This is an equilibrium concept familiar from evolutionary game theory. This type of mixed equilibrium can be stable under either fictitious play or gradient learning.

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

2. LEARNING AND EVOLUTIONARY DYNAMICS

We will examine learning in the context of two-player normal-form games, $G = (\{1, 2\}, I, J, A, B)$. I is a set of n strategies available to player 1, J a set of m strategies for player 2. Payoffs are determined by A , an $n \times m$ matrix of payoffs, and B , which is $m \times n$. A has typical element a_{ij} , which is the payoff an agent receives when playing strategy i against an opponent playing strategy j . However, we will largely be dealing with games that are “symmetric” in the evolutionary sense, that is, games for which $A = B$.³ Generalizations to the asymmetric case are briefly discussed in Section 7. We will often be dealing with a population of players, each playing a single pure strategy. In this case, the

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

distribution of strategies within the population will be described by a vector $\mathbf{x} \in S_n = \{\mathbf{x} = (x_1, \dots, x_n) \in \mathbf{R}^n: \sum x_i = 1, x_i \geq 0 \text{ for } i = 1, \dots, n\}$. As, in this paper, vectors will be treated ambiguously as either rows or columns, to avoid any further confusion, the inner product will be carefully distinguished by the symbol “.”; that is, the result of $\mathbf{x} \cdot \mathbf{x}$ is a scalar.

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

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

Players can also be modeled as maintaining a vector of relative frequencies of opponents’ past play (as in Fudenberg and Kreps, 1993; Young, 1993). They then choose strategies that maximize expected payoffs as though this vector represented the current (mixed) strategy of their opponents. The two methods are entirely equivalent. Note that the weights here are (less initial values) simply the relative frequencies multiplied by payoffs.

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

$$A = B = \begin{array}{|c|c|} \hline 0 & a \\ \hline 1 - a & 0 \\ \hline \end{array} \quad 1 > a > 0. \tag{2}$$

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

⁴This is a dynamic as used by, for example, Kandori *et al.* (1993). This is fictitious play with a one-period memory.

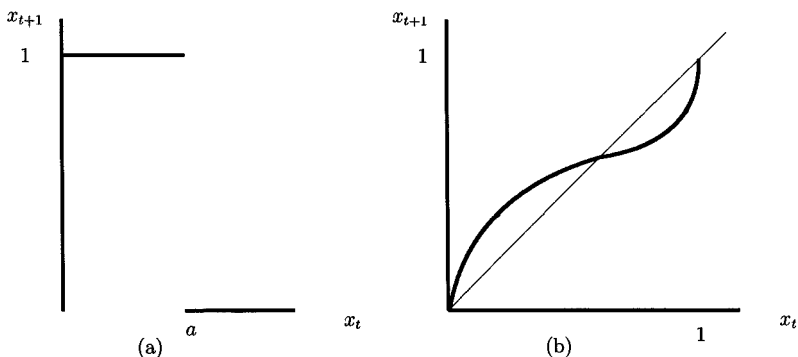


FIG. 1. Dynamics: (a) best response, (b) replicator dynamics.

In contrast, the evolutionary *replicator dynamics*, whether in continuous or discrete time, are derived on the basis that the proportional rate of growth of each strategy is equal to the difference between its payoff $(A\mathbf{x})_i$ (the i th element of the vector in parentheses) and the average payoff in the population⁵ $\mathbf{x} \cdot A\mathbf{x}$. D is a positive constant.

$$\dot{x}_i = x_i[(A\mathbf{x})_i - \mathbf{x} \cdot A\mathbf{x}] \quad \text{or} \quad x_i(t+1) = x_i(t) \frac{(A\mathbf{x})_i + D}{\mathbf{x} \cdot A\mathbf{x} + D}. \quad (3)$$

Clearly, both dynamics are continuous, the system moving smoothly toward the strategies earning the highest payoff. The replicator dynamic (in discrete time) for the game (2) is drawn in Fig. 1b. The interior mixed equilibrium is a global attractor, the pure equilibria at $x = 0, 1$ being unstable.

Important in evolutionary theory is the idea of an evolutionarily stable strategy, that is, “a strategy such that, if all members of a population adopt it, then no mutant strategy could invade the population under the influence of natural selection” (Maynard Smith, 1982, p. 10). For a large random-matching population the conditions are

DEFINITION. An *evolutionarily stable strategy* (ESS) is a strategy profile \mathbf{q} that satisfies the Nash equilibrium condition

$$\mathbf{q} \cdot A\mathbf{q} \geq \mathbf{x} \cdot A\mathbf{q} \quad (4)$$

for all $\mathbf{x} \in S_n$ and for all \mathbf{x} such that equality holds in (4), \mathbf{q} must also satisfy the stability condition

$$\mathbf{q} \cdot A\mathbf{x} > \mathbf{x} \cdot A\mathbf{x}. \quad (5)$$

⁵In a biological context, this arrives from relative reproductive success (see Hofbauer and Sigmund, 1988) but may also be an appropriate assumption in modeling learning in a human population (for example, Binmore and Samuelson, 1997).

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

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

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

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

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

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

we are able to prove convergence for the mean strategy in the population for all regular ESSs.

What we are going to show is that with a large population of players who are continually randomly matched, this type of outcome is possible even under fictitious play. Clearly, for such states to be stable, we will need a dynamic that is smooth like Fig. 1b and not discontinuous like Fig. 1a. Such smoothness does not follow automatically from aggregation. In particular, if all players in the population have the same initial beliefs, the time path for the evolution of strategies will be the same as for fictitious play with two players.⁶ Imagine in the above example, there is an entire population of players with initial weights of $(0.25, 0)$. No matter with whom they are matched they will meet an opponent playing strategy 1. Hence all players will update their beliefs at the same rate, and the same cycle is reproduced. However, this is only possible given the concentration of the population on a single point. If instead there is a nondegenerate distribution of weights across the population, it may be that not all the population will change strategy at once.

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

3. MATCHING SCHEMES

Any study of the recent literature on learning and evolution will reveal, first, that random matching within a large population of players is a common assumption and, second, that there are several ways of modeling such interaction. This diversity is, in fact, important both in terms of what it implies for theoretical results and in what cases such results are applicable. For example, there are some economic or social situations where random matching might seem a reasonable approximation of actual interaction,

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

others where it will not. Only in some cases will agents be able to obtain information about the result of matches in which they were not involved, and so on.

Fudenberg and Kreps (1993, p. 333) suggest three alternative schemes. Assuming a large population of potential players (they suggest 5000 as a reasonable number), they propose the following:

Story 1. At each date t , one group of players is selected to play the game They do so and their actions are revealed to all the potential players. Those who play at date t are then returned to the pool of potential players.

Story 2. At each date t there is a random matching of all the players, so that each player is assigned to a group with whom the game is played. At the end of the period, it is reported to all how the entire population played The play of any particular player is never revealed.

Story 3. At each date t there is a random matching of the players, and each group plays the game. Each player recalls at date t what happened in the previous encounters in which he was involved, without knowing anything about the identity or experiences of his current rivals.

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

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

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

Stories 2a and 2b have been used in the learning literature principally for reasons of tractability.⁸ They ensure a deterministic result to the matching procedure even when population size is finite. The infinite number of matchings in Story 2a, by the law of large numbers, ensures that a proportion equal to the actual frequency over the whole population of opponents

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

⁸See, for example, Kandori *et al.* (1993) and Binmore and Samuelson (1997).

playing each strategy will be drawn to play. What Stories 2, 2a, and 2b have in common is that all players know the exact distribution of strategies in the population when choosing their next strategy. There is little room for the diversity of beliefs one might expect in a large population.

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

4. POPULATION FICTITIOUS PLAY

The next stage is to examine population fictitious play (PFP) where learning takes place in a large random-mixing population. We concentrate on the case where this population is taken to be a continuum of nonatomic agents (an assumption familiar from evolutionary game theory). We develop the theory on the assumption that agents are matched according to Story 2 (or one of its variants). Story 3-type matching will be considered later.

As shown in Section 2, for the purposes of fictitious play the beliefs or weights of a given individual can be represented by point $\mathbf{w} \in \mathbf{R}^n$. The beliefs of the population will be represented by a distribution over the same space. We investigate how the distribution of beliefs, and therefore how the distribution of strategies, changes over time. It will help to create some new variables. Let $p_{ij} = w_j - w_i$, $j \neq i$. Thus \mathbf{p}_i is a vector of length $n - 1$. We will use this to work in \mathbf{R}^{n-1} instead of \mathbf{R}^n . For example, if a player has to choose between two strategies, we can summarize her beliefs by the variable p_{12} . If $p_{12} < 0$ she prefers her first strategy, if $p_{12} > 0$ her second, and if $p_{12} = 0$ she is indifferent. A player's decision rule or reaction function can then be considered as a mapping from her beliefs, a grand name for the vector \mathbf{p}_i , to strategies (i.e., $\mathbf{R}^{n-1} \rightarrow S_n$), that is, the n -simplex. This mapping will not, in general, be continuous for individual players: the fictitious play assumption limits players to pure strategies. See also Fig. 1a.

Let F_i be the population distribution function of \mathbf{p}_i over \mathbf{R}^{n-1} . Agents will play a strategy if it is the strategy given the highest weight in their beliefs. In other words, the beliefs of those playing strategy i must be in $\mathbf{R}_-^{n-1} = \{\mathbf{p}_i \in \mathbf{R}^{n-1} : p_{ij} \leq 0 \forall j \neq i\}$. What if agents are indifferent between two or more

strategies, that is, if their beliefs for some j are such that $p_{ij} = 0$? At this stage we assume that agents choose arbitrarily between the strategies over which they are indifferent. In the next section, however, we will need the stronger assumption that beliefs are given by a continuous distribution on \mathbf{R}^{n-1} . In the second case, F_i will be continuous at the origin $\mathbf{0}$, and therefore the proportion of agents playing strategy i is given unambiguously by $F_i(\mathbf{0})$. But at present, given the possibility of a mass point of indifferent agents, if the proportions of the population playing each of the n strategies is given by the vector $\mathbf{x} \in S_n$, then $x_i \leq F_i(\mathbf{0})$. However, for example, if all agents have the beliefs $p_{ij} < 0 \forall j$ then $x_i = F_i(\mathbf{0}) = 1$.

At the basis of this model of PFP is the assumption that agents update their beliefs as if they knew $\mathbf{x} \in S_n$, the true current distribution of strategies in the population. This can be supported by Story 2 in a finite or infinite population. We are, however, going to treat each x_i as a continuous variable and assume that the probability of meeting an opponent playing strategy i is x_i . For example, over a period of length Δt , each agent is matched within a single large population. If this matching is repeated an arbitrarily large number of times in each period (Story 2a), each agent will meet a proportion x_i of opponents playing strategy i . We assume that in a period of length Δt , players adjust their beliefs by Δt as much as they would in a period of length 1. According to (1), which describes the fictitious play algorithm, we have for each agent

$$\mathbf{w}(t + \Delta t) = \mathbf{w}(t) + \Delta t \mathbf{A}\mathbf{x}. \tag{6}$$

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

$$\mathbf{p}_i(t + \Delta t) = \Gamma(\mathbf{p}_i, \mathbf{x}) = \mathbf{p}_i(t) + \Delta t[(\mathbf{A}\mathbf{x})_{j \neq i} - (\mathbf{A}\mathbf{x})_i], \tag{7}$$

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

$$\Gamma^{-1}(\mathbf{p}_i) = \mathbf{p}_i(t) + \Delta t[(\mathbf{A}\mathbf{x})_i - (\mathbf{A}\mathbf{x})_{j \neq i}]. \tag{8}$$

To illustrate the properties of the model with a simple example, we consider 2×2 symmetric games, that is, games where every player must choose between the same two strategies. Let $F_i(p)$ be the cumulative distribution of $p = p_{12} = -p_{21}$ on \mathbf{R} . This distribution of beliefs determines the distribution of strategies. As the t subscript indicates, this distribution will change endogenously over time, as the beliefs of each agent are updated according to (7). This is shown in Fig. 2 (in the figure, a density function

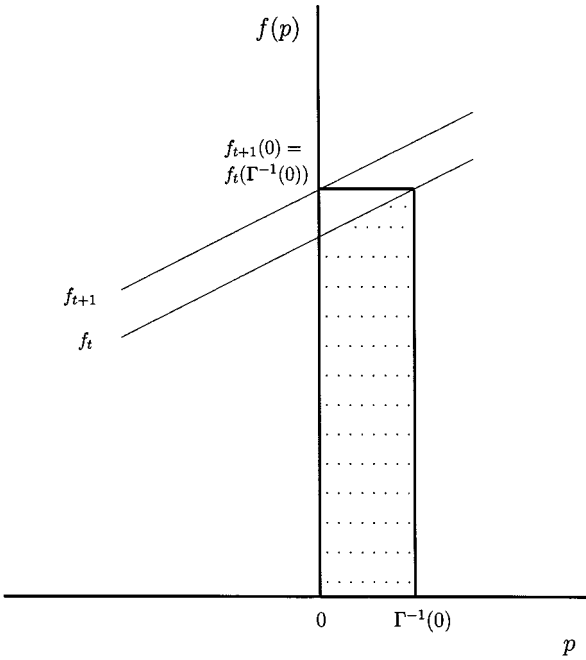


FIG. 2. Change in the distribution of beliefs.

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

$$\Gamma^{-1}(p) > p: F_{t+\Delta t}(p) = F_t(p) + \int_p^{\Gamma^{-1}(p)} dF = F_t(\Gamma^{-1}(p)), \quad (9)$$

$$\Gamma^{-1}(p) < p: F_{t+\Delta t}(p) = F_t(p) - \int_{\Gamma^{-1}(p)}^p dF = F_t(\Gamma^{-1}(p)).$$

Any agents possessing beliefs equal to $\Gamma^{-1}(0)$ will update their beliefs to $p = 0$. If $\Gamma^{-1}(0) > 0$, as is the case in Fig. 2, $F(0)$ will increase by the proportion of agents who possessed beliefs on the interval $(0, \Gamma^{-1}(0)]$. The linear nature of (7) implies that the whole distribution simply shifts to the left or to the right. This in turn will have an effect on the distribution of strategies. For example, an agent whose beliefs change from $p = 1$ to $p = -1$ will change from her second to her first strategy. By definition, $x_1 \leq F(0)$ and hence

$$x_1(t + \Delta t) \leq F_t(\Gamma^{-1}(0)) = F_t(\Delta t[(A\mathbf{x})_1 - (A\mathbf{x})_2]). \quad (10)$$

That is, in Fig. 2, x_1 increases by an amount equal to the shaded area. It is not difficult to extend this analysis to games of n strategies. In a time

interval of length Δt , the change in x_i is given by

$$x_i(t + \Delta t) \leq F_{it}(\Gamma^{-1}(\mathbf{0})) = F_{it}(\Delta t[(A\mathbf{x})_i - (A\mathbf{x})_{j \neq i}]), \quad (11)$$

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

While the state variable of the PFP process is the distribution of agents' beliefs, our main focus of interest is the distribution of strategies. We therefore define a fixed point for the PFP process as a population strategy profile which is unchanging under the dynamic specified by (7), even though beliefs may continue to change. We find a one-to-one correspondence between fixed points and strategy distributions that are Nash equilibria of the game. Mixed strategies are supported by the appropriate distribution of pure strategies across the population. For the proof of the following proposition, we assume that if an agent is indifferent between two or more strategies the choice of which of these strategies to play can be made according to any method. However, once that choice is made, no further change in strategy will be made as long as the agent remains indifferent.

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

Proof. We can start by observing that if \mathbf{q} is a Nash equilibrium then from (4) above, if $I_0 \subseteq I$ is the set of strategies in the support of \mathbf{q} , then

$$\forall i, j \in I_0 \quad (A\mathbf{q})_i = (A\mathbf{q})_j \geq (A\mathbf{q})_k \quad \forall k \notin I_0. \quad (12)$$

(a) *If.* If an agent plays i , she must prefer it. That is, $w_i \geq w_j \forall j$. From (7) and (12), no agent will change preference either between the strategies in the support of \mathbf{q} or toward any other strategy.

(b) *Only if.* Let \mathbf{q} now denote a rest point which is not a Nash equilibrium. Let $I_0 \subseteq I$ be the set of strategies in its support. If \mathbf{q} is not a Nash equilibrium then there must be a set of strategies I_k such that $\exists i \in I_0 (A\mathbf{q})_i < (A\mathbf{q})_k \forall k \in I_k$. From (7), for each agent playing strategy i , $w_i - w_k$ must be decreasing at a constant rate as long as the system is at \mathbf{q} . Within finite time, a positive measure of agents playing i must switch to a strategy in I_k . Hence the system is no longer at \mathbf{q} . ■

The following proposition is also an immediate consequent.

PROPOSITION 3. *All strict Nash equilibria are asymptotically stable.*

Proof. A strict Nash equilibrium is a state $\mathbf{q} \in S_n$ with one strategy i in its support such that there exists an $\alpha < 1$ such that for all \mathbf{x} satisfying $x_i > \alpha$, $(A\mathbf{x})_i > (A\mathbf{x})_j \forall j \neq i$. Define the set $B = \{\mathbf{x} \in S_n: x_i > \alpha\}$. Clearly, if the system enters B , it cannot leave. While in B , for all agents, each

$p_{ij} \forall j \neq i$ is decreasing at a nonvanishing rate. Each agent plays i after a finite time. ■

PROPOSITION 4. *If a strategy is strictly dominated, then it is eliminated in finite time.*

Proof. If a strategy i is strictly dominated by another strategy j , then for all agents p_{ij} is increasing at a nonvanishing rate. Any agent playing i will cease to do so after a finite time. ■

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

5. POSITIVE DEFINITE DYNAMICS

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

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

$$\mathbf{w}(t + \Delta t) = \mathbf{w}(t) + \eta, \quad (13)$$

where η is a vector of normally distributed independent random variables each with zero mean and finite variance. This would rule out the possibility of mass points of agents holding exactly the same beliefs. For example,

in the two-strategy case, if $p = -1$ for all agents, that is, they all prefer their first strategy, with the addition of the noise, beliefs would instead be normally distributed with mean -1 . We can choose the variance of η sufficiently small such that the new distribution approximates the old arbitrarily closely. Indeed, as Zeeman notes, distributions which satisfy our conditions are open dense in the set of all distributions. We state these conditions in more detail:

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

It is important to emphasize that this assumption implies significant changes from the PFP model considered in the previous section. The last inequality, for example, implies that $x_i(t) > 0 \forall i, t$. That is, only dynamics on the interior of the simplex are considered. However, it is possible for the system to approach the boundary of the simplex asymptotically. Consider the case where there is a single strictly dominant strategy i . In the previous section, we saw that, without noise, within a finite time only that strategy would be played. Here, the noise means that some agents will always prefer other strategies, but over time the numbers doing so will drop away to 0. The reason is that, from (6), we have $p_{ij}(t + \Delta t) - p_{ij}(t) < 0 \forall j \neq i$, the strength of preference for the dominated strategies is always decreasing. The result is that $\lim_{t \rightarrow \infty} \Pr[w_j + \eta_j > w_i + \eta_i] = 0$. Hence $\lim_{t \rightarrow \infty} x_j(t) = 0$ and $\lim_{t \rightarrow \infty} F_{ii}(\mathbf{0}) = 1$.

We are now going to take the continuous time limit. Returning to Fig. 2, in discrete time, all agents with beliefs in the interval $(0, \Gamma^{-1}(\mathbf{0}))$ changed strategy. As we will see, moving to continuous time is equivalent of taking the limit $\Gamma^{-1}(\mathbf{0}) \rightarrow 0$. That is, the rate of change at any given point in time is going to depend on the number of agents who are, at that instant, passing from preference of one strategy to preference of another. In other words, the rate of change will be proportional to the density of agents at the point of indifference, in Fig. 2, $f(\mathbf{0})$. Subtracting x_i from both sides of (11), we obtain

$$x_i(t + \Delta t) - x_i(t) = F_{ii}(\mathbf{p}_i^*(\Delta t)) - F_{ii}(\mathbf{0}), \tag{14}$$

where $\mathbf{p}_i^*(\Delta t) = \Delta t[(A\mathbf{x})_i - (A\mathbf{x})_{j \neq i}] = \Gamma^{-1}(\mathbf{0})$. Note that $\mathbf{p}_i^*(0) = \mathbf{0}$. Given the presence of a random disturbance in (13), the reader may be surprised to see none in the above formula. The errors, however, have been subsumed in the distribution function F_i .

PROPOSITION 5. *The continuous time limit of (14) is given by*

$$\dot{x}_i = \sum_{j \neq i} f_{ij}(\mathbf{0}) [(A\mathbf{x})_i - (A\mathbf{x})_j]. \tag{15}$$

Proof. We divide both sides of (14) by Δt . By definition, $\lim_{\Delta t \rightarrow 0} (x_i(t + \Delta t) - x_i(t))/(\Delta t) = \dot{x}_i$. We now find the limit of the right-hand side as $\Delta t \rightarrow 0$. The assumption of continuity above enables us to differentiate F_i with respect to \mathbf{p}_i . In particular, if we differentiate at $\mathbf{p}_i^*(\Delta t)$, we have $dF_i/dp_{ij} = f_{ij}(\mathbf{p}_i^*(\Delta t))$. One way to express the chain rule is, for any two differentiable functions g, h , that $\lim_{z \rightarrow 0} [g(h(z)) - g(h(0))]/z = g'(h(0)) \cdot h'(0)$. Thus

$$\lim_{\Delta t \rightarrow 0} \frac{[F_{it}(\mathbf{p}_i^*(\Delta t)) - F_{it}(\mathbf{p}_i^*(0))]}{\Delta t} = \sum_{j \neq i} f_{ij}(\mathbf{p}_i^*(0)) \cdot \left. \frac{dp_{ij}^*}{d\Delta t} \right|_{\Delta t=0}.$$

We need to evaluate the last term, but this is simple as $dp_{ij}^*/d\Delta t = (\mathbf{A}\mathbf{x})_i - (\mathbf{A}\mathbf{x})_j$. The result follows. ■

This is also a well-known problem in the physical sciences where the object is to calculate the flow of fluid (in this case, beliefs) with a known density (here f_i) subject to a vector field (here $\mathbf{A}\mathbf{x}$) across a surface, which, by careful choice of parameters, is here simply the origin. The textbook treatment of this topic assumes that the density of the liquid is fixed and exogenous. The disappearance of the time subscript on $f_{ij}(\mathbf{0})$ in (15) might suggest that this is what assumed here. However, the distribution of beliefs is endogenous and changing. Given that our main interest is in what strategies agents play, rather than the beliefs that support this behavior, we capture the time dependency of f_{ij} simply by treating it as a function of $\mathbf{x}(t)$, which, in fact, it is (we make this explicit in Proposition 7 below). As \mathbf{x} changes, payoffs change, beliefs are updated, and the distribution of beliefs shifts. For clarity, here is one striking example where we can find the exact relation between f_{ij} and \mathbf{x} .

EXAMPLE. Consider the game (2) and assume $a > 0.5$ and that beliefs at time $t = 0$ are given by a distribution $F(p) = 1/(1 + e^{\mu(0)-p})$. If the mean at time 0, $\mu(0)$, is 0, then $x_1(0) = 0.5$. Note that for this game $(\mathbf{A}\mathbf{x})_1 - (\mathbf{A}\mathbf{x})_2 = a - x_1$ and so with $x_1 = 0.5$ the return to the first strategy is higher and beliefs move uniformly to the left. In fact, considering (7) in continuous time, we have, at any point p , $\dot{p} = x_1 - a$. So as time progresses the distribution retains its shape, the only change being that its mean μ will also move such that the behavior of the system is determined by the following two equations:

$$\dot{\mu} = x_1 - a \quad \dot{x}_1 = f(0)[a - x_1] = \frac{e^\mu}{(1 + e^\mu)^2} [a - x_1]. \quad (16)$$

But we can reduce these two equations to one by realizing that $F_i(\mathbf{0}) = x_1(t) = 1/(1 + e^{\mu(t)})$. Solving for μ , we find that $\mu(t) = \log(1 - x_1)/x_1$.

Substituting this back into the second equation in (16), we obtain

$$\dot{x}_1 = x_1(1 - x_1)[a - x_1], \quad (17)$$

which is identical to (3), the continuous time replicator dynamics!

This is a remarkable result. Indeed, although we have a population of “fictitious players,” agents who base their actions on all past play, the dynamics are entirely driven by the current population state \mathbf{x} . However, the exact equivalence to the replicator dynamics is quite specific to the exact distribution of beliefs specified. What can we say in other cases? First, for this game the distribution of strategies in the population would converge smoothly to the mixed equilibrium as it would do under the replicator dynamics as long as the distribution of beliefs satisfied our continuity assumption. Of course, the value of $f(0)$ would change over time in response to changes in x_1 . But whatever the timepath of $f(0)$, as long as $f(0)$ continues to exist and remain positive, x_1 will reach a . We will go on to prove this formally, and extend our analysis to games with n strategies, using only the assumption that the density of beliefs f is a differentiable function of \mathbf{x} , the distribution of strategies (in this example, that is, with two strategies, we have $df/dx_1 = (df/dp) \cdot (dp/dF) = f'/f$).

We have developed the theory up to now using the convenient if unrealistic assumption that agents respond to the population state \mathbf{x} . However, it is possible to show that the same results can be obtained when agents are only aware of the results of the matchings in which they have been personally involved. That is, it is possible to decentralize the population learning process. Under Story 2, all agents' beliefs responded to the population state \mathbf{x} and moved in an identical way. There was a single updating function Γ . Under Story 3, how an individual updates her beliefs depends on the strategy of the opponent met in her particular matching. There are therefore n different ways of updating beliefs, one for each strategy in the population, and we need n different updating functions to take account of this.

It is easier to consider an example. Take again the game (2). If an opponent plays strategy 1, an agent will increase p_{12} by $\Delta t(1 - a)$, but, in response to strategy 2, will decrease p_{12} by $\Delta t a$. There are therefore two ways in which x_1 can change. Agents with beliefs on $(-\Delta t(1 - a), 0]$ who meet an opponent playing strategy 1 will themselves switch to strategy 2, and agents with beliefs on $[0, \Delta t a)$ will switch to strategy 1 if they encounter an opponent playing strategy 2. Note that the latter event has probability $1 - x_1$, and, given the infinite population, we assume exactly $1 - x_1$ of the agents on the interval $[0, \Delta t a)$ will change strategy. This gives us

$$x_1(t + \Delta t) - x_1(t) = (1 - x_1)F_{1t}(\Delta t a) + x_1F_{1t}(-\Delta t(1 - a)) - F_{1t}(0).$$

Taking the limit $\Delta t \rightarrow 0$ as before, we obtain from the right-hand side of the above equation $f_{12}(\mathbf{0})[(1-x_1)a-x_1(1-a)]$. Stated more formally, and also covering the case of n strategies, we have the following proposition:

PROPOSITION 6. *The continuous time limit of population fictitious play with an infinite population under matching Story 3 is also given by (15).*

Proof. Each individual meets an opponent playing strategy j with probability x_j . In this case, each w_i will change by $\Delta t a_{ij}$. More generally, the individual's beliefs change in the following way:

$$\mathbf{p}_i(t + \Delta t) = \Gamma_j(\mathbf{p}_i, \mathbf{x}) = \mathbf{p}_i(t) - \Delta t [a_{(k \neq i)j} - a_{ij}],$$

where $[a_{(k \neq i)j} - a_{ij}]$ is the vector of length $n-1$, of form $(a_{1j} - a_{ij}, \dots, a_{nj} - a_{ij})$. Define $p_{ijk}^*(\Delta t) = \Gamma_{jk}^{-1}(\mathbf{0}) = \Delta t [a_{ij} - a_{kj}]$. We assume a proportion x_1 of the population change beliefs according to Γ_1 , a proportion x_2 according to Γ_2 , etc. Therefore, the change in x_i is given by

$$x_i(t + \Delta t) - x_i(t) = \sum_{j=1}^n x_j F_{ii}((\Gamma_j^{-1}(\mathbf{0}))) - F_{ii}(\mathbf{0}).$$

Again we divide both sides by Δt and take the limit $\Delta t \rightarrow 0$. From the right-hand side, again by use of the chain rule, we obtain

$$\sum_{j=1}^n x_j \left(\sum_{k \neq i} f_{ik}(p_{ijk}^*(\mathbf{0})) [a_{ij} - a_{kj}] \right),$$

which we can reassemble to form (15). ■

Whatever the derivation of (15), clearly it is very close to the continuous time replicator dynamics (3) and the linear dynamics proposed by Friedman (1991):

$$\dot{x}_i = \frac{1}{n} \sum_{j \neq i} [(A\mathbf{x})_i - (A\mathbf{x})_j]. \quad (18)$$

We now define a class of dynamics. We then go on to show that the continuous time PFP dynamic is a member of this class.

DEFINITION. Any dynamic of the form $\dot{\mathbf{x}} = Q(\mathbf{x})A\mathbf{x}$, where the matrix $Q(\mathbf{x})$ for any \mathbf{x} on the interior of S_n satisfies the following five conditions, we call a *positive definite dynamic*.⁹

1. Every element of Q is continuously differentiable in \mathbf{x} .
2. $\lim_{x_i \rightarrow 0} Q_{ij} = 0 \forall j$.

⁹Josef Hofbauer has brought to my attention Hofbauer and Sigmund (1990) in which there is a similar formulation under the name "adaptive dynamics."

3. $Q\mathbf{u} = \mathbf{0}$, where \mathbf{u} denotes the vector $(1, 1, \dots, 1)$.
4. Q is otherwise positive definite. That is, for any $\mathbf{z} \in \mathbf{R}^n$ which is not a multiple of \mathbf{u} or $\mathbf{0}$, $\mathbf{z} \cdot Q\mathbf{z} > 0$.
5. Q is symmetric.

Condition 1 is a sufficient condition for a unique solution to the differential equation $\dot{\mathbf{x}} = Q\mathbf{A}\mathbf{x}$. Condition 2 ensures that the dynamic remains upon the simplex. Geometrically, the operator Q maps the vector of payoffs $\mathbf{A}\mathbf{x}$ from \mathbf{R}^n to the subspace $\mathbf{R}_0^n = \{\mathbf{z} \in \mathbf{R}^n: \mathbf{u} \cdot \mathbf{z} = 0\}$ (if the vector $Q\mathbf{A}\mathbf{x}$ did not add to 0 then \mathbf{x} would cease to add to 1). It has nullspace \mathbf{u} . That is, at a mixed Nash equilibrium where payoffs are equal ($\mathbf{A}\mathbf{x}$ is a multiple of \mathbf{u}), $\dot{\mathbf{x}} = \mathbf{0}$. Because Q is positive definite the angle between $\mathbf{A}\mathbf{x}$ and $Q\mathbf{A}\mathbf{x}$ is less than 90° . This last property is what Friedman (1991) calls “weak compatibility.” It is easy to see that both the replicator dynamics and Friedman’s linear dynamics satisfy the above conditions,¹⁰ but we can also establish a further result.

PROPOSITION 7. *The continuous time PFP dynamic (15) is a positive definite dynamic.*

Proof. If we write (15) in matrix form then $\dot{\mathbf{x}} = Q_f \mathbf{A}\mathbf{x}$, where the f subscript is for “fictitious play.” Q_f has a diagonal $\sum_{j \neq i} f_{ij}$ and off-diagonal $-f_{ij} = -f_{ji}$. Satisfaction of Conditions 1 and 2 is guaranteed by the Continuity Assumption. In particular, $df_{ij}/dx_k = (df_{ij}/d\mathbf{p}_i) \cdot (d\mathbf{p}_i/dx_k) = (df_{ij}/d\mathbf{p}_i) \cdot (d\mathbf{p}_i/dF_k)$. Clearly, $Q_f \mathbf{u} = \mathbf{u} \cdot Q_f = \mathbf{0}$. However, $\mathbf{x} \cdot Q_f \mathbf{x} = \sum_{j \neq i} f_{ij}(x_i - x_j)^2 \geq 0$. ■

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

LEMMA 1. *Any ESS \mathbf{q} is negative definite with respect to the strategies in its support. That is, $(\mathbf{x} - \mathbf{q}) \cdot \mathbf{A}(\mathbf{x} - \mathbf{q}) < 0$ for all \mathbf{x} with the same support as \mathbf{q} (see van Damme, 1991, Theorem 9.2.7).*

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

¹⁰The linear dynamics do not satisfy condition 2 and it is necessary to impose some other kind of boundary condition.

LEMMA 2. If A is negative definite when constrained to \mathbf{R}_0^n (that is, $\mathbf{z} \cdot A\mathbf{z} < 0 \forall \mathbf{z} \in \mathbf{R}_0^n$), then QA is a stable matrix (i.e., all its eigenvalues have negative real parts when QA is constrained to \mathbf{R}_0^n).

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

$$\mathbf{y}^c \cdot QAz = \mathbf{z}^c \cdot Az = \mu \mathbf{y}^c \cdot \mathbf{z} = \mu \mathbf{y}^c \cdot Q\mathbf{y}. \quad (19)$$

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

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

PROPOSITION 8. All regular ESSs are asymptotically stable for any positive definite dynamic.

Proof. Let \mathbf{q} be a fully mixed ESS. Differentiating $Q(\mathbf{x})A\mathbf{x}$ with respect to \mathbf{x} and evaluating at \mathbf{q} , we obtain $Q(\mathbf{q})A + dQ/dx_i A\mathbf{q}$. At a Nash equilibrium $QA\mathbf{x} = \mathbf{0}$. It follows that, for each x_i , $dQ/dx_i A\mathbf{q} = \mathbf{0}$. Thus the Jacobian of the system at \mathbf{q} is given by $Q(\mathbf{q})A$. By Lemma 2 all its eigenvalues have real part negative.

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

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

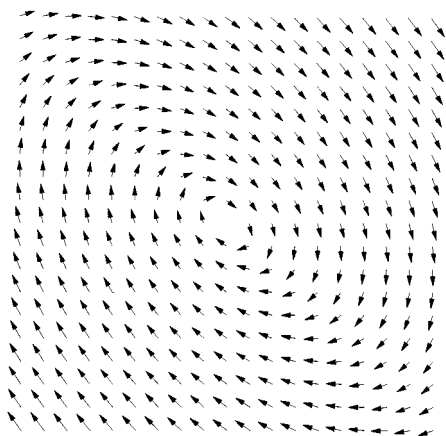
$$\dot{x}_i = \sum_j Q_{ij}(A\mathbf{q})_j + \sum_{j,k} Q_{ij}A_{jk}\xi_k.$$

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

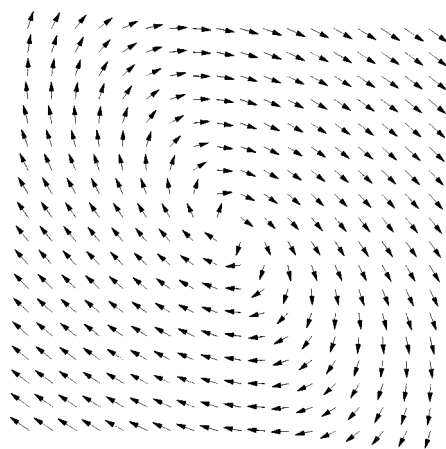
What is particularly attractive about this result is that to determine stability one no longer has to examine the potentially complicated function $Q(\mathbf{x})$. Instead, one can confine attention to the properties of A alone. For example, for the PFP dynamics it is not necessary to know the shape of the distribution of beliefs. The last two conditions on Q are the substantive ones. Positive definiteness seems a minimal condition to place upon a dynamic. Nonetheless, it becomes a sufficient condition for stability when combined with symmetry. Why this should lead to asymptotic stability for ESSs can be seen in the traditional economic terms of convexity and concavity. A “positive definite” dynamic is a gradient climber. The negative definiteness of ESSs, of course, implies concavity. This is illustrated in Fig. 3a. This represents part of the vector field $A\mathbf{x}$ (projected onto \mathbf{R}^2) from a game with $n = 3$ and which possesses a fully mixed ESS at the center of the vector field depicted. The arrows of vector field points upward and inward. Any positive definite dynamic will move “uphill” toward the ESS.

The role of symmetry in the results is more complex. First, note that the second half of the proof of Proposition 8 does not require symmetry and indeed it is easy to show that positive definiteness alone (or, equivalently, Friedman’s (1991) “weak compatibility”) is sufficient for the asymptotic stability of pure-strategy ESSs. Symmetry is, however, essential in the first part of the proof and in ensuring the stability of mixed ESSs. The symmetry requirement works by ruling out matrix functions $Q(\mathbf{x})$ that represent rotations. The importance of this is illustrated by the example given in Friedman (1991, p. 655) where he constructs a dynamic which diverges from a mixed ESS. The dynamic constructed is effectively an anticlockwise rotation of the vector field $A\mathbf{x}$. This is illustrated in Fig. 3b. This represents a positive definite but asymmetric transformation of the vector field shown in Fig. 3a. The arrows have been rotated (by less than 90°) so that they now point outward. This illustrates the fact that if Q is positive definite but not symmetric, even if A is negative definite, the matrix QA can be positive definite.

Symmetry is also what differentiates positive definite dynamics from Friedman’s (1991) concept of order compatibility or the monotonicity of Nachbar (1990) and Samuelson and Zhang (1992). Monotonicity requires that $\dot{x}_i/x_i > \dot{x}_j/x_j$ iff $(A\mathbf{x})_i > (A\mathbf{x})_j$, and order compatibility, $\dot{x}_i > \dot{x}_j$ iff $(A\mathbf{x})_i > (A\mathbf{x})_j$. One other important difference is that the usual definition of monotonicity requires only Lipschitz continuity rather than the differentiability we have specified for positive definite dynamics. It is easy to check that both monotonicity and order compatibility imply the positive definiteness of Q (as Friedman points out, order compatibility implies weak compatibility which is equivalent to positive definiteness). However, monotonicity and order compatibility do not imply symmetry. Conversely, there



(a)



(b)

FIG. 3. Vector fields: (a) negative definite, (b) positive definite.

are positive definite dynamics which are not monotone or order compatible. This is easy to show by example. Take $n = 3$ and $(A\mathbf{x})_1 = (A\mathbf{x})_2 > (A\mathbf{x})_3 = 0$. Order compatibility requires that $Q_{11} + Q_{12} = Q_{21} + Q_{22} > 0$, monotonicity $(Q_{11} + Q_{12})/x_1 = (Q_{21} + Q_{22})/x_2 > 0$, whereas the constraints on a positive definite dynamic in this context are $Q_{11} > 0$, $Q_{11}Q_{22} > Q_{12}Q_{21}$, and $Q_{12} = Q_{21}$. Clearly, they are not equivalent.

6. MIXED-STRATEGY DYNAMICS

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

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

LEMMA 3. *If \mathbf{x} is in the interior of S_n then $\mathbf{z} \cdot Q_m \mathbf{z} > 0$ for any \mathbf{z} which is not a multiple of \mathbf{u} (Zeeman, 1981, p. 265).*

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

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

and hence

LEMMA 4. *The dynamic for the mean mixed strategy satisfies $\dot{\mathbf{x}} = Q_m A\mathbf{x}$ (Zeeman, 1981, p. 266).*

We can find similar results for the dynamics, variously called gradient, stimulus-response, or reinforcement learning, considered by Harley (1981), Börgers and Sarin (1997), Crawford (1989), and Roth and Erev (1995). This may seem strange in that, first, Börgers and Sarin rightly point out this learning process when aggregated across a population of players is not identical to the replicator dynamics for either pure or mixed strategies, and that, second, Crawford proves that in such a large population, under such dynamics the mixed-strategy equilibrium of a simple game like (2) is unstable. However, Crawford's definition of a mixed-strategy equilibrium is the state where every agent plays the equilibrium mixed strategy; that

is, in game (2), they all play their first strategy with probability a . However, I would argue that in a random-mixing population this definition is overstrict. It is possible to have a state where the average strategy in the population, and hence the expected strategy of an opponent, is equal to the mixed-strategy equilibrium, although no agent plays the exact mixed-strategy equilibrium profile. For example, the i th member of the population could play her first strategy with probability $a + \varepsilon_i$ with $\sum \varepsilon_i = 0$.

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

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

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

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

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

$$E[w_i(t + \Delta t)] = y_i(w_i(t) + \Delta t (A\mathbf{x})_i) + (1 - y_i)w_i(t). \quad (20)$$

There are three important differences between this learning rule and fictitious play. First, it is stochastic, not deterministic. Second, while, under fictitious play, agents have a limited capacity for assessing what they might have received if they had used some other strategy, here agents only consider what actions they actually play and what payoffs they actually receive

(this type of learning model was developed to analyze animal behavior). Third, for the probabilities to remain well defined, we must require all pay-offs to be nonnegative,¹¹ and that all agents start with all elements of their vector \mathbf{w} strictly positive. From (20), we can obtain

$$E[y_i(t + \Delta t) - y_i(t)] = \frac{\Delta t y_i((\mathbf{A}\mathbf{x})_i - \mathbf{y} \cdot \mathbf{A}\mathbf{x})}{W + \Delta t \mathbf{y} \cdot \mathbf{A}\mathbf{x}} + O((\Delta t)^2). \quad (21)$$

This is a special case¹² of the RPS rule of Harley (1981). Crawford (1989) characterizes individual behavior in a large population of players by the deterministic continuous time equation

$$\dot{y}_i = y_i[(\mathbf{A}\mathbf{x})_i - \mathbf{y} \cdot \mathbf{A}\mathbf{x}]. \quad (22)$$

To obtain a deterministic result, again some form of Story 2 matching is assumed. Börgers and Sarin (1997) show that by using a slightly different specification of the updating rule one can obtain a continuous time limit similar to Crawford's equation (22).¹³ The advantage of the approach of Börgers and Sarin and Crawford is that learning behavior is easier to characterize, but only at the cost of additional assumptions.

In any case, the next step is to derive an expression for the evolution of the population mean. Under Story 3-type matching, the changes made in each agent's beliefs will depend on the strategy chosen by that agent and by his opponent (both random variables). Nonetheless, if we think of the change made by each agent as a draw from the distribution that describes the population, $x_i(t + \Delta t) - x_i(t)$ is then the sample mean. Hence the variance of the change in x_i is decreasing in the number of agents. Thus, if the population is infinite, then the evolution of the population mean will be deterministic. We have from (21), leaving out the higher-order terms (which would disappear anyway when we take the limit $\Delta t \rightarrow 0$),

$$\begin{aligned} x_i(t + \Delta t) - x_i(t) &= \int E[y_i(t + \Delta t) - y_i(t)] dH \\ &= \int \Delta t y_i[(\mathbf{A}\mathbf{x})_i - \mathbf{y} \cdot \mathbf{A}\mathbf{x}]/(W + \Delta t \mathbf{y} \cdot \mathbf{A}\mathbf{x}) dH \\ &= \int \Delta t y_i[\mathbf{e}_i - \mathbf{y}]/(W + \Delta t \mathbf{y} \cdot \mathbf{A}\mathbf{x}) dH \cdot \mathbf{A}\mathbf{x}, \end{aligned} \quad (23)$$

¹¹Either we consider only games with positive payoffs or we add a positive constant to all payoffs sufficiently large to make them positive. Clearly, such a transformation would make no difference to a game's strategic properties, though, in a dynamic context, it can change the rate of adjustment. See the discussion of discrete time processes in the next section.

¹²Equation (21) can be obtained by setting what Harley calls the "memory factor" to 1.

¹³It would be the same if Börgers and Sarin considered as did Crawford a single random-mixing population.

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

$$\dot{\mathbf{x}} = Q_r A\mathbf{x}, \quad (24)$$

where the r subscript is for reinforcement learning. The diagonal of Q_r has the form $\int y_i(1 - y_i)/W dH$, the off-diagonal, $-\int y_i y_j/W dH$. Hence Q_r is symmetric and $Q_r \mathbf{u} = \mathbf{0}$. Clearly, $\mathbf{z} \cdot Q_r \mathbf{z} = \sum_{i \neq j} \int y_i y_j/W dH (z_i - z_j)^2 \geq 0$. Consequently, Q_r is positive semidefinite. To obtain the model of either Börgers and Sarin (1997) or Crawford (1989), it is simply necessary to set $W = 1$ for all agents. Clearly, this would not change the conclusion that, although $Q_r \neq Q_m$,

PROPOSITION 9. *The mean of the mixed-strategy replicator dynamic and the mean of the stimulus–response/reinforcement learning process are positive definite dynamics.*

This, together with Proposition 8, extends the existing results on reinforcement learning.

EXAMPLE. Take the game (2), assume $a = 0.5$, $F(y_1) = y_1^2$, and hence $x_1 = 2/3$. Under the mixed-strategy replicator dynamics, we have $\dot{f}(y_1) = 2y_1[1/9 - y_1/3]$. That is, those agents playing the first strategy with probability less than $1/3$, and hence far from the equilibrium strategy, are increasing in number. For the gradient dynamics, under, for example, the version suggested by Crawford (22), we have for each agent $\dot{y}_1 = -y_1(1 - y_1)/6$. In words, all agents are decreasing the weight they place on their first strategy. This also demonstrates the difference between the two mixed-strategy dynamics. The evolutionary dynamic replaces badly performing agents by better performers¹⁴; under the gradient dynamics, all agents respond to the situation by changing strategy. As Crawford (1989) discovered, the state where all agents have $y_1 = 0.5$ is not going to be stable. In this example, the agents who are currently playing the “equilibrium” mixed strategy ($y_1 = 0.5$) are respectively dying off and moving away from it. However, for both dynamics we have $\dot{x}_1 = Q_{11}[1/2 - x_1]$, and hence the mean strategy clearly approaches the equilibrium.¹⁵

¹⁴Though perhaps this type of dynamic could be reproduced in a population that learns by imitation.

¹⁵Harley (1981, p. 624) reproduces two graphs of the results he obtained from simulations of a similar game using his learning model. Two things are apparent: the population mean approaches the mixed-strategy equilibrium; the strategy of individual players (typically) does not.

7. EXTENSIONS

Since the concept of an ESS is a strong refinement on Nash equilibrium and, consequently, there are many games which do not possess any equilibrium which satisfies its conditions, one might wonder how positive definite dynamics perform in these cases. For any constant-sum game for any $\mathbf{x} \in S_n$, $\mathbf{x} \cdot A\mathbf{x} = v$, where v is the value of the game. It follows, if the game has a fully mixed equilibrium \mathbf{q} , that $(\mathbf{x} - \mathbf{q}) \cdot A(\mathbf{x} - \mathbf{q}) = 0$. From Lemma 2 and in particular (19) we have that

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

This result unfortunately is of the “anything can happen” type. For the linear dynamics (18), because they are linear, the corollary implies that such an equilibrium must be a neutrally stable center (it is easy to check that $V = \frac{1}{2}(\mathbf{x} - \mathbf{q}) \cdot (\mathbf{x} - \mathbf{q})$ is a constant of motion in this case). For nonlinear dynamics the fact that their linearizations have zero eigenvalues may hide asymptotic stability or instability.

Second, there are games which possess equilibria which are positive definite. It is an obvious corollary of Proposition 8 that positive definite dynamics diverge from such equilibria. This can prove useful in terms of equilibrium selection. Unstable positive definite equilibria can be rejected in favor of stable ESSs. This works well in games with both ESSs and positive definite equilibria:

$$A = \begin{array}{|c|c|c|} \hline 0 & a_1 & -b_1 \\ \hline -b_2 & 0 & a_2 \\ \hline a_3 & -b_3 & 0 \\ \hline \end{array} \quad a_i, b_i > 0 \quad i = 1, 2, 3. \quad (25)$$

But the game (25) has a unique equilibrium which, for example, for $a_i = 1$, $b_i = 3$, $i = 1, 2, 3$, is positive definite. In this case, the vector field $A\mathbf{x}$ will be similar to that illustrated in Fig. 3b, where the payoff gradient points away from equilibrium. Gradient-climbing behavior will thus lead away from equilibrium and no positive definite dynamic can converge.

This might seem problematic, but, in fact, it offers a strong empirical prediction. For rational players under the full-information assumptions of conventional game theory, for a game with a unique Nash equilibrium it should not matter whether it is positive or negative definite. However, we can conjecture that in a random-matching environment under experimental conditions, the strategy frequencies of human subjects would converge

if, for example, $a_i = 3$ and $b_i = 1$ but not if $a_i = 1$ and $b_i = 3$. We can make this conjecture with a degree of confidence because so many different specifications of adaptive learning are consistent with positive definite dynamics. Such divergence is not necessarily “irrational” or “myopic.” Indeed, if $a_i = 1$, $b_i = 3$, $i = 1, 2, 3$, average payoffs are at a minimum at the mixed equilibrium. Divergence increases average payoffs.

The robustness of these results, however, does depend on the property of positive or negative definiteness. For equilibria which are neither positive nor negative definite, it is possible for stability properties to vary according to the exact specification of the dynamics. Such equilibria can be attractors or repellers. Using (25) again as an example, the pure-strategy replicator dynamics converge iff $a_1 a_2 a_3 > b_1 b_2 b_3$, the linear dynamics iff $a_1 + a_2 + a_3 > b_1 + b_2 + b_3$, while simulation suggests that the PFP dynamics will converge to any equilibrium of the game which is not positive definite.

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

$$\mathbf{x}(t+1) = \mathbf{x}(t) + QA\mathbf{x}, \quad (26)$$

where Q again satisfies the five conditions outlined above (examples would be the discrete time replicator dynamics (3) or the discrete dynamic specified in (23)). In this case, pure strategies which are regular ESSs will be asymptotically stable, the second part of the proof of Proposition 8 applying equally well in discrete time. The problem is, as always, with mixed strategies. From (26), the linearization at a fully mixed fixed point \mathbf{q} will be

$$I + Q(\mathbf{q})A. \quad (27)$$

As we have shown, the eigenvalues of QA are negative. If, however, they are too “large,” the absolute values of the eigenvalues of $I + QA$ will be greater than 1. So it is possible for a discrete time positive definite process to diverge from a mixed ESS. This is going to depend on the magnitude of the change in strategy distribution made each period. In the case of a pure-strategy equilibrium, the dynamic cannot jump over the fixed point and out of the simplex. In contrast, unless the rate of change is sufficiently slow, it is possible to shoot right past a mixed-strategy equilibrium. Note that, for example, for the discrete time replicator dynamics given in (3), the rate of adjustment is decreasing in the constant D . Hence the stability of ESSs can be assured if D is sufficiently large. In the case of gradient learning, the rate of change is decreasing over time as the size of individuals’ weights (W in the notation of the last section) increases. Furthermore, in the case of positive definite equilibria, where QA has positive eigenvalues, then all

the eigenvalues of the linearization (27) are clearly greater than 1 and the equilibrium will most certainly be unstable.

In the case of asymmetric games, it is well known that no mixed-strategy equilibria are ESSs. Furthermore, it is also well known that mixed-strategy equilibria are either saddles or centers for the replicator dynamics (Hofbauer and Sigmund, 1988). It is easy to show that this result generalizes to all positive definite dynamics. In particular, let \mathbf{x} give the strategy frequencies in the first population and \mathbf{y} in the second, and $\dot{\mathbf{x}} = Q\mathbf{A}\mathbf{y}$, $\dot{\mathbf{y}} = P\mathbf{B}\mathbf{x}$, where Q and P are positive definite matrices satisfying the conditions outlined above. Then the argument outlined in Hofbauer and Sigmund (1988, pp. 142–143) goes through unchanged in that the linearization around a mixed equilibrium \mathbf{q} will be given by

$$\begin{pmatrix} 0 & Q(\mathbf{q})A \\ P(\mathbf{q})B & 0 \end{pmatrix}.$$

The zero trace implies eigenvalues that are either a mix of positive and negative or have real part zero. In the latter case, we would call the equilibrium a center, but, depending on higher-order terms, it may nonetheless be stable or unstable (again see Hofbauer and Sigmund, 1988, pp. 273–283).

8. CONCLUSION

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

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

on fictitious play, stimulus–response learning, and mixed-strategy replicator dynamics, it has been the negative definiteness of ESSs which has been essential.

Researchers have begun to test the predictions of models of learning and evolution by carrying out experiments. The results presented in this paper may be relevant in several ways. First, they are broadly in accordance with the results reported by Friedman (1996), who reproduced in the laboratory the anonymous random matching environment considered here. In what he terms “Type 1 Games,” Friedman found convergence in average strategy to a mixed ESS although most subjects tended to stick to a single pure strategy. Second, Mookherjee and Sopher (1994), for example, attempt to determine whether fictitious play or gradient-type rules best describe the learning behavior of their subjects. As we have shown, the differences between these two types of models, in a random-matching environment at least, are smaller than previously thought. Our results would also point to a reason why Gale *et al.* (1995), using replicator dynamics, and Roth and Erev (1995), using a stimulus–response learning process, obtain similar results in trying to simulate the behavior of experimental subjects playing the ultimatum bargaining game. Third, there has been some debate (Brown and Rosenthal, 1990; Binmore *et al.* 1994) about what constitutes convergence to equilibrium in experimental games. What we show here is that it may be foolish to expect more than convergence in the average strategy in a population of players. Last, we offer further predictions to be tested. ESSs should be locally stable. For games which possess positive definite equilibria, our predictions are equally clear. Learning processes should not converge to such equilibria.

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

REFERENCES

- Binmore, K., and Samuelson, L. (1997). “Muddling Through: Noisy Equilibrium Selection,” *J. Econom. Theory* **74**, 235–265.
- Binmore, K., Swierzbinski, J., and Proulx, C. (1994). “Does Minimax Work? An Experimental Study,” Working Paper. London: University College.

- Börgers, T., and Sarin, R. (1997). "Learning through Reinforcement and Replicator Dynamics," *J. Econom. Theory* **77**, 1–14.
- Boylan, R. T. (1992). "Laws of Large Numbers for Dynamical Systems with Randomly Matched Individuals," *J. Econom. Theory* **57**, 473–506.
- Brown, G. W. (1951). "Iterative Solutions of Fictitious Play," in *Activity Analysis of Production and Allocation* (T. C. Koopmans, Ed.). New York: Wiley.
- Brown, J., and Rosenthal, R. (1990). "Testing the Minimax Hypothesis: A Re-Examination of O'Neill's Game Experiment," *Econometrica* **58**, 1065–1081.
- Canning, D. (1992). "Average Behaviour in Learning Models," *J. Econom. Theory* **57**, 442–472.
- Crawford, V. (1985). "Learning Behaviour and Mixed-Strategy Nash Equilibria," *J. Econom. Behav. Organ.* **6**, 69–78.
- Crawford, V. (1989). "Learning and Mixed Strategy Equilibria in Evolutionary Games," *J. Theoret. Biol.* **140**, 537–550.
- Friedman, D. (1991). "Evolutionary Games in Economics," *Econometrica* **59**, 637–666.
- Friedman, D. (1996). "Equilibrium in Evolutionary Games: Some Experimental Results," *Econom. J.* **106**, 1–25.
- Fudenberg, D., and Kreps, D. (1993). "Learning Mixed Equilibria," *Games Econom. Behav.* **5**, 320–367.
- Fudenberg, D., and Levine, D. (1993). "Steady State Learning and Nash Equilibrium," *Econometrica* **61**, 547–573.
- Gale, J., Binmore, K., and Samuelson, L. (1995). "Learning to Be Imperfect: The Ultimatum Game," *Games Econom. Behav.* **8**, 56–90.
- Harley, C. (1981). "Learning the Evolutionarily Stable Strategy," *J. Theoret. Biol.* **89**, 611–633.
- Hines, W. G. S. (1980). "Three Characterisations of Population Strategy Stability," *J. Appl. Probab.* **17**, 333–340.
- Hofbauer, J., and Sigmund, K. (1988). *The Theory of Evolution and Dynamical Systems*. Cambridge: Cambridge Univ. Press.
- Hofbauer, J., and Sigmund, K. (1990). "Adaptive Dynamics and Evolutionary Stability," *Appl. Math. Lett.* **3**, 75–79.
- Jordan, J. (1993). "Three Problems in Learning Mixed Strategy Equilibria," *Games Econom. Behav.* **5**, 368–386.
- Kalai, E., and Lehrer, E. (1993). "Rational Learning Leads to Nash Equilibrium," *Econometrica* **61**, 1019–1045.
- Kandori, M., Mailath, G., and Rob, R. (1993). "Learning, Mutation and Long Run Equilibrium in Games," *Econometrica* **61**, 29–56.
- Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge: Cambridge Univ. Press.
- Milgrom, P., and Roberts, J. (1991). "Adaptive and Sophisticated Learning in Normal Form Games," *Games Econom. Behav.* **3**, 82–100.
- Miyasawa, K. (1961). "On the Convergence of the Learning Process in a 2×2 Non-Zero Sum Game," mimeo. Princeton, NJ: Princeton Univ.
- Monderer, D., and Shapley, L. (1996). "Fictitious Play Property for Games with Identical Interests," *J. Econom. Theory*, **68**, 258–265.
- Mookherjee, D., and Sopher, B. (1994). "An Experimental Study of Learning and Decision Costs in Constant-Sum Games," Working Paper.
- Nachbar, J. H. (1990). "Evolutionary Selection Dynamics in Games: Convergence and Limit Properties," *Internat. J. Game Theory* **19**, 59–89.

- Robinson, J. (1951). "An Iterative Method of Solving a Game," *Ann. Math.* **54**, 296–301.
- Roth, A. E., and Erev, I. (1995). "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games Econom. Behav.* **8**, 164–212.
- Samuelson, L., and Zhang, J. (1992). "Evolutionary Stability in Asymmetric Games," *J. Econom. Theory* **57**, 363–392.
- Shapley, L. (1964). "Some Topics in Two Person Games," in *Advances in Game Theory* (M. Dresher *et al.*, Eds.). Princeton, NJ: Princeton Univ. Press.
- van Damme, E. (1991). *The Stability and Perfection of Nash Equilibrium*, 2nd ed. Berlin: Springer-Verlag.
- Young, H. P. (1993). "The Evolution of Conventions," *Econometrica* **61**, 57–84.
- Zeeman, E. C. (1981). "Dynamics of the Evolution of Animal Conflicts," *J. Theoret. Biol.* **89**, 249–270.