

Learning and Mixed-Strategy Equilibria in Evolutionary Games

VINCENT P. CRAWFORD

*Department of Economics, University of California, San Diego,
La Jolla, California 92093, U.S.A.*

(Received 22 November 1988, and accepted in revised form 12 June 1989)

This paper considers whether Maynard Smith's concept of an evolutionarily stable strategy, or "ESS", can be used to predict long-run strategy frequencies in large populations whose members are randomly paired to play a game, and who adjust their strategies over time according to sensible learning rules. The existing results linking the ESS to stable equilibrium population strategy frequencies when strategies are inherited do not apply to learning, even when each individual always adjusts its strategy in the direction of increased fitness, because the inherited-strategies stability results depend on aggregating across individuals, and this is not possible for learning. The stability of learning must therefore be analyzed for the entire system of individuals' strategy adjustments. The interactions between individuals' adjustments prove to be generically destabilizing at mixed-strategy equilibria, which are saddlepoints of the learning dynamics. Using the inherited-strategies dynamics to describe learning implicitly restricts the system to the stable manifold whose trajectories approach the saddlepoint, masking its instability. Thus, allowing for the interactions between individuals' strategy adjustments extends the widely recognized instability of mixed-strategy equilibria in multi-species inherited-strategies models to single-species (or multi-species) learning models.

1. Introduction

The concept of an evolutionarily stable strategy, or "ESS", was introduced by Maynard Smith & Price (1973) and Maynard Smith (1974) to describe the effects of selection for more successful strategies in environments where an individual's expected rate of reproduction, or fitness, is jointly determined by its own and other individuals' strategies. In the model originally analyzed by Maynard Smith and most often studied by subsequent writers, individuals are selected at random from a population and matched, in pairs, to play a symmetric two-person game. The individuals in the population are identical except for their strategies; these are inherited and fixed for life. The population is large enough that the differences between the expected strategy frequencies faced by different individuals are negligible, even though individuals are never paired with themselves and generally play different strategies. Individuals reproduce asexually and breed true, passing on their strategies unchanged to their offspring. Finally, an individual's current fitness is jointly determined by its strategy and the strategy of the individual with which it is currently paired, as summarized by the payoff matrix of the game†.

† In what follows, a "large" population is one in which all individuals are assumed to face the same population strategy frequencies, and a "finite" population is one in which this is not assumed. "Mixed strategy" refers in general to randomized strategies, but does not exclude "pure" (that is, unrandomized) strategies.

Maynard Smith defined an ESS for the model just described as a mixed strategy that, if played by all members of a *monomorphic* population, has strictly higher fitness than any mutant strategy that enters the population with sufficiently low frequency. (The definition is the same for *polymorphic* populations, in which individuals play different strategies in equilibrium, with the qualification that mutants must then have lower fitness than the population, on average.) The intuition for Maynard Smith's definition is that, if the members of a population all play an ESS, mutants that enter the population with low frequency will reproduce more slowly than individuals who play the ESS. The mutants' relative population frequencies will therefore approach zero over time, restoring the population strategy frequencies of the ESS. Computing an ESS should then allow the analyst to infer the possible long-run values of these frequencies from the payoff matrix alone, an important simplification.

This intuition was first formalized, for continuous-time versions of the population dynamics, by Taylor & Jonker (1978), Zeeman (1979) and Hines (1980*a,b*)[†]. Taylor & Jonker (1978) and Zeeman (1979) studied symmetric two-person finite matrix games, requiring individuals to play pure strategies, and considered large polymorphic populations, in which different strategies may persist in equilibrium. They showed that, for generic payoffs, a vector of population strategy frequencies that (when treated as a mixed strategy) satisfies the ESS condition with arbitrary mixed strategies allowed as mutations is a locally asymptotically stable (henceforth "stable") equilibrium of the population dynamics[‡]. Taylor & Jonker (1978) gave an example to show that the converse is not true in general, so that Maynard Smith's definition is overly restrictive when individuals are required to play pure strategies: Games whose players have more than two pure strategies can have equilibria whose strategy frequencies do not satisfy the ESS condition, such that any small group of pure-strategy mutants with higher average fitness than the population also has individual fitness differences that alter its strategy frequencies over time in such a way that the population strategy frequencies return to the equilibrium.

Hines (1980*a,b*) (see also Cressman & Hines, 1984; Hines, 1987, section 4; Maynard Smith, 1974, 1982, chapter 2 and appendix D; Zeeman, 1979, 1981) identified a closer link between the ESS condition and stability, showing that when individuals in a large polymorphic or monomorphic population are allowed to play mixed strategies, the ESS condition is generically necessary as well as sufficient for stability of the population strategy frequencies. A population whose strategy frequencies violate the ESS condition is vulnerable to any mixed-strategy mutant with higher fitness, because the strategy frequencies of such mutants (unlike those of groups of pure-strategy mutants) do not evolve when there are fitness differences between their constituent pure strategies. This result unifies the treatment of pure and mixed strategies on the individual level and shows how to use the concept of evolutionary stability, originally formulated for monomorphic populations, to characterize the long-run effects of strategy selection in polymorphic populations of mixed strategies.

[†] Taylor & Jonker (1978) also considered discrete-time versions of the dynamics, for which the arguments in support of Maynard Smith's definition are weaker; see also Hines (1987: 241, 245-246).

[‡] A locally asymptotically stable equilibrium is one that has a neighborhood such that any trajectory that originates in the neighborhood converges to the equilibrium.

It is important in what follows to note that, because the results just summarized relate *relative* strategy frequencies and fitnesses, they apply equally well to growing, fixed, or shrinking populations. Also, because a large population playing a strategy that does not maximize fitness against itself is clearly vulnerable to a low-frequency mutation that does maximize fitness against that strategy, an ESS must be a symmetric Nash equilibrium of the game played by matched pairs†. The stability arguments of Taylor & Jonker, Zeeman and Hines can therefore be viewed as an alternative justification for this standard game-theoretic characterization of behavior as well as for the ESS.

In recent years, the idea of evolutionary stability has been applied extensively in biology, and its usefulness has been found surprisingly robust to deviations from the original population model (see for example Hines, 1987, sections 4–6). Perhaps encouraged by this robustness, a number of biologists and social scientists have suggested using the ESS to explain behavior in human or animal populations in which inheritance of strategies is supplemented or supplanted by learning (Axelrod, 1984; Harley, 1981; Hines, 1987; Hines & Bishop, 1983; Houston & Sumida, 1987; Maynard Smith, 1982, chapters 5, 13; Sugden, 1986; Zeeman, 1979, 1981). Such applications rest implicitly on a dynamic justification like that developed for inherited strategies by Taylor & Jonker, Zeeman and Hines. Learning plainly fits the inherited-strategies model if it proceeds purely by imitation, with members of successive generations choosing strategies, once and for all, in numbers proportional to the payoffs of earlier adherents of those strategies. But this rules out individual strategy adjustment, an essential feature of learning. This paper considers whether it is possible to construct a sensible justification for using the ESS to describe the consequences of learning that involves individual strategy adjustment.

The issues raised by individual strategy adjustment stand out most clearly when it is the only source of change in population strategy frequencies. From now on, I shall use the term “learning” in this special sense, further restricting attention for simplicity to fixed populations.

There are important similarities between the inherited-strategies dynamics and sensible learning dynamics, because if each individual in a large population adjusts in the direction of increased payoffs, the population strategy frequencies also move in that direction. For this reason, it is often assumed (see, for example, Axelrod, 1984; Sugden, 1986; Taylor & Jonker, 1978: 146, 153; Zeeman, 1981: 251) that the inherited-strategies justification for the ESS extends to learning. It is shown here, however, that mixed-strategy equilibria (but not pure-strategy equilibria, in general) are generically unstable for sensible specifications of the learning dynamics. It follows that if the learning dynamics converge, they must converge to a configuration in which individuals play only pure strategies.

Learning and inherited strategies can have different implications for stability because the results of Taylor & Jonker, Zeeman and Hines depend on aggregating across individuals; aggregation is justified in large populations for inherited strategies, but not, in general, for learning. Learning must therefore be analyzed at

† An equilibrium in the game played by matched pairs will be called a “Nash equilibrium” whenever this is necessary to preserve the distinction between game-theoretic equilibria and the equilibria of the population dynamics.

the individual level, taking into account the interactions between individuals' strategy adjustments. These prove to be generically destabilizing at mixed-strategy equilibria, which are saddlepoint equilibria of the learning dynamics. Using the aggregate inherited-strategies dynamics to describe learning in effect restricts these dynamics to the stable manifold whose trajectories approach the saddlepoint, masking this instability.

The rest of the paper is organized as follows. Section 2 defines the ESS for large populations and reviews its relationship to the symmetric Nash equilibria of the game played by matched pairs, and to the stable equilibrium population strategy frequencies of the inherited-strategies dynamics. Section 3 compares the inherited-strategies dynamics with sensible learning dynamics, showing that the latter do not allow aggregation, even in large populations. Section 4 shows that mixed-strategy equilibria are generically unstable for the learning dynamics introduced in section 3. The stability analysis is carried out explicitly only for the "Hawk-Dove" example of Maynard Smith & Parker (1976) and Maynard Smith (1982); section 5 discusses the straightforward extension to more general symmetric two-person finite matrix games. Section 5 also discusses other extensions of the analysis and related work.

2. Inherited Strategies

This section defines the ESS for large populations and discusses its relationship to the symmetric Nash equilibria of the game played by matched pairs and to the stable equilibria of the inherited-strategies dynamics.

Consider a large population of identical individuals, repeatedly and anonymously paired at random to play a symmetric two-person finite matrix game. Recall that an ESS is a mixed strategy that, if initially played by all members of the population, has higher expected payoff than any mutant mixed strategy that enters the population with low frequency. Let \mathbf{q} and \mathbf{s} be vectors that give the expected frequencies with which the pure strategies are played in the population, and let $E(\mathbf{q}|\mathbf{s})$ be the expected payoff of the mixed strategy \mathbf{q} when the expected population frequencies are given by \mathbf{s} . (With random pairing in a large population, it makes no difference whether the distribution of strategies that yields these frequencies is monomorphic or polymorphic.) An ESS can then be formally defined as a mixed strategy \mathbf{s} such that for any $\mathbf{q} \neq \mathbf{s}$ and any sufficiently small $\varepsilon > 0$,

$$E[\mathbf{s}|(1-\varepsilon)\mathbf{s} + \varepsilon\mathbf{q}] > E[\mathbf{q}|(1-\varepsilon)\mathbf{s} + \varepsilon\mathbf{q}]. \quad (1)$$

Using the linearity in probabilities of expected payoffs reduces eqn (1)† to

$$(1-\varepsilon)E(\mathbf{s}|\mathbf{s}) + \varepsilon E(\mathbf{s}|\mathbf{q}) > (1-\varepsilon)E(\mathbf{q}|\mathbf{s}) + \varepsilon E(\mathbf{q}|\mathbf{q}). \quad (2)$$

This inequality holds for all small $\varepsilon > 0$ if and only if, for all \mathbf{q} ,

$$E(\mathbf{s}|\mathbf{s}) \geq E(\mathbf{q}|\mathbf{s}), \quad (3)$$

and

$$E(\mathbf{s}|\mathbf{q}) > E(\mathbf{q}|\mathbf{q}) \quad \text{whenever} \quad E(\mathbf{s}|\mathbf{s}) = E(\mathbf{q}|\mathbf{s}). \quad (4)$$

† Equations (1)-(6) are inequalities.

Inequality (3) simply requires (s, s) to be a symmetric Nash equilibrium in the game played by matched pairs; eqn (4) is a further implication of evolutionary stability, discussed in Maynard Smith (1982, chapter 2) and elsewhere.

The Hawk–Dove game can be used to illustrate this definition and its relationship to stability of the population strategy frequencies for the inherited-strategies dynamics. The Hawk–Dove game has payoff matrix

	<i>H</i>	<i>D</i>
<i>H</i>	<i>A</i>	<i>b</i>
<i>D</i>	<i>c</i>	<i>d</i>

In this matrix, *H* and *D* stand for Hawk and Dove, and only the Row player's payoffs are shown; the Column player's payoffs can be deduced by symmetry.

I begin by considering monomorphic populations. Let z denote each individual's equilibrium probability of playing *H* (so that $s_1 \equiv z$ and $s_2 \equiv 1 - z$ in terms of the more general notation used above to define the ESS). Ignoring borderline cases for simplicity, eqns (3) and (4) imply that $z = 1$ is an ESS if and only if $a > c$, $z = 0$ is an ESS if and only if $d > b$, and $z^* = (b - d)/(b - d + c - a)$ is an ESS if and only if $c > a$ and $b > d$ (in which case $0 < z^* < 1$). The first two of these conclusions follow immediately from eqn (3). To verify the third, note that playing z^* yields $(bc - ad)/(b - d + c - a)$ against the population frequency z^* , as does any alternative mixed strategy q ; thus z^* satisfies eqn (3). To satisfy eqn (4), z^* must yield a higher expected payoff against q than q does against itself, so that

$$z^*[qa + (1 - q)b] + (1 - z^*)[qc + (1 - q)d] > q[qa + (1 - q)b] + (1 - q)[qc + (1 - q)d] \quad (5)$$

or, equivalently,

$$(z^* - q)\{[qa + (1 - q)b] - [qc + (1 - q)d]\} > 0. \quad (6)$$

It is easy to verify eqn (6) from the parameter restrictions $c > a$ and $b > d$.

The inherited-strategies population dynamics for this model are easiest to describe if it is assumed, following Taylor & Jonker (1978) and Zeeman (1979), that the population is polymorphic and each individual plays a pure strategy. Then the state of a large population can be summarized by its strategy frequencies, and because all individuals who play a given strategy have the same fitness, the expected proportional rate of growth of a strategy's population frequency equals the current difference between its fitness and the population frequency-weighted average fitness of all pure strategies. Let \bar{z} denote the expected population frequency of *H*. Then $\dot{\bar{z}}$, the time rate of change of \bar{z} , equals \bar{z} times the fitness of *H* minus the population average fitness when its expected frequency is \bar{z} :

$$\begin{aligned} \dot{\bar{z}} &\equiv \bar{z}[\bar{z}a + (1 - \bar{z})b] - \{\bar{z}[\bar{z}a + (1 - \bar{z})b] + (1 - \bar{z})[\bar{z}c + (1 - \bar{z})d]\} \\ &\equiv \bar{z}(1 - \bar{z})\{[\bar{z}a + (1 - \bar{z})b] - [\bar{z}c + (1 - \bar{z})d]\}. \end{aligned} \quad (7)$$

The assumption that the population is large underlies eqn (7) in three ways: it justifies treating the population frequency as a continuous variable, it justifies using the same value of this frequency to calculate different individuals' fitnesses, and it justifies, via the law of large numbers, identifying the total realized rate of growth of the individuals playing each pure strategy and the fitness that equals its mathematical expectation. These observations can be used to show that eqn (7) holds exactly in infinite populations, and that it continues to hold, approximately, in sufficiently large finite populations.

The differential eqn (7) has three equilibria: $\bar{z} = 0$, $\bar{z} = 1$, and $\bar{z} = \bar{z}^* \equiv (b - d)/(b - d + c - a)$. Taylor & Jonker (1978) showed that, aside from borderline cases, each of these equilibria is stable if and only if the payoff parameters are such that it is an ESS. When $\bar{z} = \bar{z}^*$, for instance, the term in square brackets in the second line of eqn (7) equals zero, and

$$\partial \dot{\bar{z}} / \partial \bar{z} |_{\bar{z} = \bar{z}^*} = \bar{z}^*(1 - \bar{z}^*)(a - b - c + d) = (c - a)(d - b)/(b - d + c - a); \quad (8)$$

thus, given that $0 < \bar{z}^* < 1$, the stability condition $\partial \dot{\bar{z}} / \partial \bar{z} |_{\bar{z} = \bar{z}^*} < 0$ is equivalent to the parameter restrictions $c > a$ and $b > d$. As noted above, Taylor & Jonker, Zeeman and Hines extended this generic equivalence between the ESS condition and stability to all symmetric two-person finite matrix games in which individuals may play mixed strategies.

3. Learning and Aggregation

This section introduces sensible specifications of the learning process and shows that they do not allow the use of aggregate relationships like eqn (7) to describe the dynamics of strategy frequencies in infinite populations. My argument assumes a specific individual adjustment process for concreteness, but it will be clear that its conclusion does not depend on the details of the process.

The use of aggregate dynamics like eqn (7) to describe learning is often justified informally by noting that, if each individual in an infinite population adjusts its strategy in the direction of increased payoffs, the population strategy frequencies also adjust in that direction, as eqn (7) requires. Justifying eqn (7), however, also requires an aggregation argument. To see when aggregation is possible, imagine that each individual chooses a mixed strategy, adjusting it over time in response to the differences between the current payoffs of its pure strategies. Let s_j^i and s^i , respectively, denote individual i 's probability of playing its j th pure strategy and i 's vector of mixed-strategy probabilities. Let \bar{s} denote the vector of expected population strategy frequencies, and let e_j denote the j th pure strategy, expressed as a mixed-strategy probability vector with a one in the j th place and zeroes elsewhere. Otherwise maintaining the notation and assumptions of section 2, and ignoring boundary problems, which are not germane to the point made here, assume that, for each i and j , individual i adjusts its j th mixed-strategy probability according to

$$\dot{s}_j^i \equiv s_j^i [E(e_j | \bar{s}) - E(s^i | \bar{s})]. \quad (9)$$

The differential equation system defined by eqn (9) sets the vector of proportional rates of change of each individual's mixed-strategy probabilities equal to the vector of partial derivatives of the individual's expected payoff with respect to those probabilities, computed taking into account the linearity in probabilities of expected payoffs and that the probabilities must be rescaled proportionally so they will continue to sum to one. (Summing eqn (9) over j and recalling that $E(s^i|\bar{s}) \equiv \sum_j s_j^i E(e_j|\bar{s})$ shows that $\sum_j \dot{s}_j^i \equiv 0$ for each i .)

Thus eqn (9) defines a gradient process for each individual, adjusting its mixed strategy in the current direction of greatest payoff increase in the simplex of feasible probability vectors. These gradient processes are plausible descriptions of "atheoretical" learning, in which individuals who lack either the sophistication or the knowledge of the structure of their environment to draw complex inferences from the information they receive discover how their strategies influence their payoffs by experimenting occasionally with small adjustments, favoring those that yield the largest payoff improvements.

The gradient learning model defined by eqn (9) is also the closest individual-adjustment analog, in large populations, of the inherited-strategies dynamics studied by Taylor & Jonker, Zeeman and Hines. Summing eqn (9) over i would immediately yield those dynamics if the $E(s^i|\bar{s})$ term on its right-hand side were replaced by $E(\bar{s}|\bar{s})$. But this would make \dot{s}_j^i/s_j^i independent of s^i , and eqn (9) could then no longer ensure, even in a large population, that each individual's mixed-strategy probabilities remain in the feasible simplex. Thus, the presence of the $E(s^i|\bar{s})$ term on the right-hand side of eqn (9) reflects an inherent difference between learning and inherited strategies.

This difference has the important consequence that, even in large populations, individuals' learning processes cannot be aggregated to give $\dot{\bar{s}}$ as a function of \bar{s} alone. To see this, return to the Hawk-Dove game, letting z^i denote i 's mixed-strategy probability of playing H and again letting \bar{z} denote the expected population frequency of H . Equation (9) specializes to

$$\begin{aligned} \dot{z}^i &\equiv z^i \{ [\bar{z}a + (1 - \bar{z})b] - \{ z^i [\bar{z}a + (1 - \bar{z})b] + (1 - z^i) [\bar{z}c + (1 - \bar{z})d] \} \} \\ &\equiv z^i (1 - z^i) \{ [\bar{z}a + (1 - \bar{z})b] - [\bar{z}c + (1 - \bar{z})d] \}. \end{aligned}$$

Equation (10) does not yield an aggregate relationship between $\dot{\bar{z}}$ and \bar{z} , because its right-hand side is non-linear in z^i : In general, $\dot{\bar{z}}$ depends on the individual z^i , not just on \bar{z} . The impossibility of aggregation is easiest to see in a finite population with N members, where

$$\bar{z} \equiv \sum_{i=1}^N z^i / N \tag{11}$$

and

$$\begin{aligned} \dot{\bar{z}} &\equiv \sum_{i=1}^N \dot{z}^i / N \equiv \{ [\bar{z}a + (1 - \bar{z})b] - [\bar{z}c + (1 - \bar{z})d] \} \sum_{i=1}^N z^i (1 - z^i) / N \\ &\equiv \{ [\bar{z}a + (1 - \bar{z})b] - [\bar{z}c + (1 - \bar{z})d] \} \left[\bar{z} - \sum_{i=1}^N (z^i)^2 / N \right], \end{aligned} \tag{12}$$

which relates \bar{z} to the individual z^i as well as \bar{z} . But even in a large population, the aggregation required for a learning interpretation of eqn (7) is justified only when each individual plays the same mixed strategy throughout, so that $z^i \equiv \bar{z}$ for all i †.

4. Stability

This section studies the stability of the learning dynamics introduced in section 3, showing that individuals' strategies are generically unstable at mixed-strategy equilibria. (The implications of this result for the stability of population strategy frequencies are considered in section 5.)

Because individuals' learning processes cannot be aggregated, the analysis must be carried out for the entire system of their strategy adjustments. To give a clear account of the issues that arise in specifying this system, and to describe the interactions between individuals' adjustments precisely, I work with finite populations, otherwise maintaining the assumptions and notation of sections 2 and 3. The finite-population analysis reveals what happens in large populations by passing to the limit and shows why the normal practice of doing this at the start of the analysis is misleading in this case. It is assumed throughout that the population still has enough members to justify identifying realized population strategy frequencies and their expectations.

To adapt Maynard Smith's ESS definition to finite populations, assume a fixed population of size N and write the expected payoff of individual i playing mixed strategy \mathbf{q} as $E(\mathbf{q}|\bar{\mathbf{s}}^i)$, where

$$\bar{\mathbf{s}}^i \equiv \sum_{\substack{j=1 \\ j \neq i}}^N \mathbf{s}^j / (N-1) \quad (13)$$

gives the expected population strategy frequencies, *excluding* individual i . This generalizes the large-population definition given in section 2 to environments where excluding an individual's strategy from the population has a non-negligible effect on its strategy frequencies.

As before, an ESS, if played by all members of the population, must have an expected payoff at least as high as any mutant strategy that enters the population with sufficiently low frequency. The formal definition is derived, following Schaffer (1988) and Maynard Smith (1988), by taking "low frequency" to mean $1/N$, assuming that the single mutant switches strategies from the ESS (so that the population remains fixed at size N), and comparing the payoffs of the mutant and the individuals who continue to play the ESS. An ESS is now a mixed strategy \mathbf{s} such that, for any $\mathbf{q} \neq \mathbf{s}$,

$$E\left(\mathbf{s} \left| \frac{N-2}{N-1} \mathbf{s} + \frac{1}{N-1} \mathbf{q} \right.\right) \geq E(\mathbf{q}|\mathbf{s}); \quad (14) \ddagger$$

† Even if, by chance, $z^i = \bar{z}$ for all i initially, random deviations of realized payoffs from their expected values in finite populations would quickly throw the system off, making aggregation impossible.

‡ Equation (14) is an inequality.

inequality (14) is the finite-population analog of inequality (3). [Schaffer, 1988, supplemented inequality (14) with a "stability" condition that plays the role of inequality (4), showing that the ESS for the Hawk-Dove game implied by inequality (14) also satisfies this supplementary condition. The points made here, however, rely only on inequality (14).]

A useful alternative definition of the ESS for finite populations, also due to Schaffer, requires a mixed strategy \mathbf{s} to solve

$$\max_{\mathbf{q}} E(\mathbf{q}|\mathbf{s}) - E\left(\mathbf{s} \left| \frac{N-2}{N-1}\mathbf{s} + \frac{1}{N-1}\mathbf{q} \right.\right), \quad (15)$$

where the maximization is taken over all mixed strategies \mathbf{q} . This is plainly equivalent to requiring inequality (14), because inequality (14) holds with equality if $\mathbf{q} = \mathbf{s}$. Two important conclusions for finite populations follow from this equivalence. First, although any ESS converges to a symmetric Nash equilibrium as the population grows, an ESS is no longer a symmetric Nash equilibrium in general. Second, the implication of Nash equilibrium that the pure strategies in the support of a mixed ESS all yield equal expected payoffs against the expected population strategy frequencies no longer holds exactly. Instead, any mutant playing a mixture of strategies in the support of a mixed-strategy ESS has the same expected payoff as an individual playing the ESS, taking into account the fact that the individuals playing the ESS (unlike the mutant) are matched with the mutant with probability $1/(N-1)$. This follows from the linearity of the objective function in eqn (15) and the fact that its maximized value is zero.

I now argue that the equilibria (stable or unstable) of sensible learning dynamics in finite populations with sufficiently many members must be located at symmetric Nash equilibria. It follows that ESS's are equilibria of these dynamics only in large populations, in general, and that a learning justification for the ESS depends on establishing the stability of symmetric Nash equilibria in large populations.

Reiterating the second conclusion about the finite-population ESS from above, if \mathbf{s} is a mixed-strategy ESS, it must satisfy

$$E(\mathbf{e}_j|\mathbf{s}) = E\left(\mathbf{s} \left| \frac{N-2}{N-1}\mathbf{s} + \frac{1}{N-1}\mathbf{e}_j \right.\right) \quad (16)$$

for all pure strategies j in its support. [Requiring eqn (16) for the extreme points \mathbf{e}_j of the support of \mathbf{s} is a proxy, given the linearity in \mathbf{q} of

$$E(\mathbf{q}|\mathbf{s}) - E\left(\mathbf{s} \left| \frac{N-2}{N-1}\mathbf{s} + \frac{1}{N-1}\mathbf{q} \right.\right),$$

for requiring eqn (16) for all \mathbf{q} that have the same support as \mathbf{s} .] For individuals' adjustment processes to reach equilibrium at \mathbf{s} , they would have to respond [generalizing the expressions in eqn (16) to polymorphic populations] to differences between the $E(\mathbf{e}_j|\bar{\mathbf{s}}^i)$ and the

$$E\left(\bar{\mathbf{s}}^i \left| \frac{N-2}{N-1}\bar{\mathbf{s}}^i + \frac{1}{N-1}\mathbf{e}_j \right.\right).$$

Rescaling problems (which are non-trivial but not insurmountable) aside, there seems to be no sensible way to justify this kind of response when individuals experiment independently. If individuals' experiments are rare, simultaneous experiments by matched individuals are rare indeed. Individual i 's payoff experience, whether experimenting or not, is therefore generated against a population whose expected strategy frequencies are approximately \bar{s}^i . Thus, independent experiments cannot provide the information individuals would need to respond to differences between the $E(e_j|\bar{s}^i)$ and the

$$E\left(s^i \left| \frac{N-2}{N-1} \bar{s}^i + \frac{1}{N-1} e_j \right. \right).$$

This argument suggests that, at least when the population has many members, the finite-population analog of eqn (9) is a sensible model of the learning process. This model sets, for each i and j ,

$$\dot{s}_j^i \equiv s_j^i [E(e_j|\bar{s}^i) - E(s^i|\bar{s}^i)]. \tag{17}$$

The system defined by eqn (17) is in equilibrium at symmetric Nash equilibria, but not, in general, at ESS's in finite populations. I now consider whether eqn (17) yields a learning justification for the symmetric Nash equilibrium and the ESS in sufficiently large finite populations.

The analysis is carried out explicitly only for the Hawk-Dove game; extensions to more general games are considered in section 5. Let

$$\bar{z}^i \equiv \sum_{\substack{j=1 \\ j \neq i}}^N z^j / (N-1) \tag{18}$$

and specialize eqn (17) to

$$\begin{aligned} z^i &\equiv z^i \{ [\bar{z}^i a + (1 - \bar{z}^i) b] - \{ z^i [\bar{z}^i a + (1 - \bar{z}^i) b] + (1 - z^i) [\bar{z}^i c + (1 - \bar{z}^i) d] \} \} \\ &\equiv z^i (1 - z^i) \{ [\bar{z}^i a + (1 - \bar{z}^i) b] - [\bar{z}^i c + (1 - \bar{z}^i) d] \} \\ &\equiv z^i (1 - z^i) (\bar{z}^i \alpha + \beta), \quad i = 1, \dots, N, \end{aligned} \tag{19}$$

where $\alpha \equiv a - b - c + d$ and $\beta \equiv b - d$.

The differential equation system (19) is the finite-population analog of eqn (10). Like eqn (10), it allows aggregation only if individuals always play the same mixed strategies, so that z^i , and therefore \bar{z}^i , is independent of i . The stability analysis must therefore be carried out for the entire system.

Assume that $c > a$ and $b > d$, so that $\alpha < 0$ and $\beta > 0$, and $z^* = (b - d) / (b - d + c - a) = -\beta / \alpha$ is the only interior equilibrium of eqn (19) and the only symmetric Nash equilibrium of the game. Partially differentiating eqn (19) and evaluating the results at $z^i = -\beta / \alpha$ for all i reveals that, for all i ,

$$\partial \dot{z}^i / \partial z^i = 0 \tag{20}$$

and, for all i and all $j \neq i$,

$$\partial \dot{z}^i / \partial z^j = (-\beta / \alpha) (1 + \beta / \alpha) \alpha / (N - 1) = -\beta (1 + \beta / \alpha) / (N - 1) < 0. \tag{21}$$

The matrix of the locally linearized version of eqn (19) thus has zeroes along its main diagonal and a constant, $\kappa \equiv -\beta (1 + \beta / \alpha) (N - 1)$, everywhere else.

Because the trace of a matrix equals the sum of its eigenvalues, in this case zero, at least generically one or more of the eigenvalues must have a strictly positive real part, making the system unstable. For the Hawk-Dove game, it can be shown that the eigenvalues equal $(N - 1)\kappa \equiv -\beta(1 + \beta/\alpha) < 0$ and $-\kappa \equiv \beta(1 + \beta/\alpha)/(N - 1) > 0$, the latter having multiplicity $N - 1$ (as is necessary for the eigenvalues to sum to the trace of the matrix in this case), so that the system is always unstable. The proof, due to Dennis Smallwood (personal communication), is as follows:

Letting \mathbf{A} denote the matrix of the locally linearized system determined by eqns (20) and (21), the eigenvalues, denoted λ , and the associated eigenvectors, denoted \mathbf{x} , can be obtained by solving $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, $\mathbf{x} \neq \mathbf{0}$, where $\mathbf{0}$ denotes a vector of zeroes conformable to \mathbf{x} . Writing out $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, given that \mathbf{A} has zeroes on the main diagonal and κ 's everywhere else, yields:

$$\begin{aligned}
 &\kappa x_2 + \kappa x_3 + \dots + \kappa x_{n-1} + \kappa x_n = \lambda x_1 \\
 \kappa x_1 &\quad + \kappa x_3 + \dots + \kappa x_{n-1} + \kappa x_n = \lambda x_2 \\
 &\quad \dots \\
 &\quad \dots \\
 &\quad \dots \\
 \kappa x_1 + \kappa x_2 + \kappa x_3 + \dots &\quad + \kappa x_n = \lambda x_{n-1} \\
 \kappa x_1 + \kappa x_2 + \kappa x_3 + \dots + \kappa x_{n-1} &= \lambda x_n.
 \end{aligned}
 \tag{22}$$

Summing the equations in (22) yields

$$\kappa(N - 1) \sum_{i=1}^n x_i = \lambda \sum_{i=1}^n x_i.
 \tag{23}$$

Because κ and $(N - 1)$ differ from zero, this has only one solution for λ when $\sum_{i=1}^n x_i \neq 0$, namely $\lambda^1 \equiv \kappa(N - 1)$; the associated eigenvector can be taken to be $\mathbf{x}^1 \equiv (1, \dots, 1)$. Other solutions are possible when $\sum_{i=1}^n x_i = 0$; it is easy to verify that the $\mathbf{x}^j \equiv (1, 0, \dots, 0, -1, 0, \dots, 0)$, where -1 is the j th component, $j = 2, \dots, N$, are also eigenvectors, each with associated eigenvalue $\lambda^j \equiv -\kappa$. These N linearly independent eigenvectors and the associated eigenvalues, with multiplicity as indicated above, constitute a complete set.

Thus, all of the eigenvalues of the locally linearized system are real (as indicated by the symmetry of \mathbf{A}), and $N - 1$ of them are strictly positive. The true, non-linear system is therefore unstable, like the linear system. The system drifts away from equilibrium approximately at an exponential rate proportional to $1/N$, hence the drift is very slow when N is large. The system is unstable, however, for any finite population†.

† The instability of mixed-strategy equilibria is described as only generic because the restrictions on payoff parameters under which it was proven for the Hawk-Dove game rule out non-generic parameter configurations for which $\kappa = 0$. The zero eigenvalues that result in such cases make the locally linearized dynamics unstable, but useless in determining whether the true non-linear dynamics are stable. (The parameter configurations that cause this problem are harder to identify for more general games, but still non-generic.) Working with a large population from the start of the analysis would evidently also yield zero eigenvalues, again making the local stability analysis inconclusive.

Only if $z^i \equiv \bar{z}$ throughout for all i can eqn (19) be aggregated to obtain eqn (7), and the inherited-strategies analysis be used to show that the learning dynamics are stable. The eigenvalue of the locally linearized version of eqn (7) equals $(N-1)\kappa \equiv -\beta(1 + \beta/\alpha)$, the only stable eigenvalue of the locally linearized learning dynamics. Thus, using the inherited-strategies dynamics to analyze the stability of learning in effect restricts the system to its stable manifold, masking its instability.

5. Conclusion

This section discusses extensions of the analysis and related work. The instability result established here is mathematically closely related to the multi-species inherited-strategies results of Eshel & Akin (1983), Hines (1981) and Maynard Smith (1982, appendix J), and to the instability results established for learning processes like those considered here by Crawford (1974, 1985) in a closely related game-theoretic framework[†]. The main contribution of the present paper is to show that the already widely recognized instability of mixed-strategy equilibria for multi-species inherited-strategies dynamics extends to sensible single-species (or multi-species) learning dynamics, once the finiteness of the population and the impossibility of aggregating individuals' learning processes are recognized[‡].

The generic instability of mixed-strategy equilibria for learning dynamics extends to symmetric (or asymmetric) finite matrix games with any number of pure strategies. For any such game, the matrix of the locally linearized learning dynamics has zeroes along the main diagonal, so that, for generic payoffs, there is at least one eigenvalue with strictly positive real part. These zeroes arise because each individual's expected payoff is linear in its own mixed strategy, so that the first-order equalities required for equilibrium in individual adjustment processes like eqn (17) immediately imply that the second-order effects measured by the partial derivatives in eqn (20) equal zero. This is true of mixed-strategy equilibria in general, and their instability is therefore a robust result. In particular, individual adjustment processes with equilibria at the ESS in finite populations, as discussed in section 4, would also be unstable.

As noted in Crawford (1974, 1985) and Hines (1981), the proof given here, which ignores boundary problems, extends, generically, to equilibria in which some pure strategies are played with zero probability, as long as two or more are played with positive probability. Then, the pure strategies that have zero probability can simply

[†] There are many surface differences between evolutionary games and the environments studied by Crawford (1974, 1985); this paper follows the conventions of evolutionary game theory as closely as possible. The models of Crawford (1974, 1985) assume (by contrast) a fixed, finite population; possibly non-linear, absolute (instead of proportional) adjustment of mixed-strategy probabilities; discrete time; and repeated interaction of a single group of possibly asymmetric players, who ignore the effects of their current strategy adjustments on each other's future strategies (instead of anonymous, random pairing of identical players from a large population). The instability result survives all of these changes. It does depend on the form of the assumed learning rules, in that mixed-strategy adjustments must be differentiable functions of the payoff differences between pure strategies. Crawford (1985) gives references to the game-theoretic literature on learning.

[‡] Although the instability result established here depends on recognizing the finiteness of the population, it differs from other finite population-based criticisms of Maynard Smith's original ESS definition in the literature (see Schaffer, 1988; Maynard Smith, 1988; and the references given there) and, unlike those criticisms, remains significant in populations of indefinite size.

be eliminated from the game and the stability analysis of section 4 carried out for the reduced game. Except in borderline cases, section 4's argument is unaffected by this reduction.

Given the general instability of mixed-strategy equilibria, pure-strategy configurations are the only possible stable outcomes of the learning dynamics. My instability arguments clearly do not apply to equilibria of these dynamics at which all individuals play pure strategies. In fact, it is not hard to see that any pure-strategy Nash equilibrium that satisfies inequality (3) with strict inequality must be a locally stable equilibrium of learning dynamics like those studied here, and analysis of simple examples suggests that pure-strategy polymorphisms may also be locally stable. [The results of Hines (1980*b*) and Zeeman (1981) for the inherited-strategies dynamics are inconclusive on this point for the learning dynamics, because they depend on using the large-population assumption to justify aggregation, eliminating the interactions between individuals' adjustments studied here.]

I close by mentioning some important related work. The results of the present paper complement those of Hines & Bishop (1983), who introduced learning effects into the inherited-strategies dynamics by allowing small individual gradient adjustments in parents' strategies immediately before their transmission to offspring. They treated the differences between the expected strategy frequencies faced by different individuals in a large population as negligible, aggregating individuals' adjustment processes and thereby eliminating the effects of the interactions between them studied here. They showed that the kind of learning effects they studied can increase strategy diversity over time, and can even lead to differences between the long-run population strategy frequencies and evolutionarily stable frequencies in the analogous inherited-strategies model. By way of comparison, the analysis presented here shows that, when learning effects predominate over fitness effects, the interactions between individuals' strategy adjustments alone, in any finite population, however large, can increase strategy diversity.

Selten (1988) studied the effect of individuals' anticipations of each other's strategy adjustments in a multi-species large-population model that otherwise resembles the model of Crawford (1974, 1985), showing that such anticipations can sometimes make mixed-strategy Nash equilibria stable.

Finally, Fudenberg & Kreps (1988) developed a game-theoretic model of learning by experimentation, with individuals much more sophisticated and better informed about the structure of their environment than assumed here, and used it to evaluate the intuition that persistent experimentation eventually leads individuals to a (suitably refined) Nash equilibrium.

I am grateful to Dennis Smallwood and the referees for helpful advice, to Yong-Gwan Kim for research assistance, and to the National Science Foundation for research support under grant SES 8703337.

REFERENCES

- AXELROD, R. (1984). *The Evolution of Cooperation*. New York: Basic Books.
CRAWFORD, V. P. (1974). Learning the Optimal Strategy in a Zero-Sum Game. *Econometrica* 42, 885.

- CRAWFORD, V. P. (1985). Learning Behavior and Mixed-Strategy Nash Equilibria. *J. econ. Behav. Organization* **6**, 69.
- CRESSMAN, R. & HINES, W. G. S. (1984). Correction to the Appendix of "Three Characterizations of Population Strategy Stability". *J. appl. Probab.* **21**, 213.
- ESHEL, I. & AKIN, E. (1983). Coevolutionary Instability of Mixed Nash Solutions. *J. math. Biol.* **18**, 123.
- FUDENBERG, D. & KREPS, D. (1988). *A Theory of Learning, Experimentation, and Equilibrium in Games*. MIT and Stanford.
- HARLEY, C. B. (1981). Learning the Evolutionarily Stable Strategy. *J. theor. Biol.* **89**, 611.
- HINES, W. G. S. (1980a). Three Characterizations of Population Strategy Stability. *J. appl. Probab.* **17**, 333.
- HINES, W. G. S. (1980b). Strategy Stability in Complex Populations. *J. appl. Probab.* **17**, 600.
- HINES, W. G. S. (1981). Multi-Species Populations Models and Evolutionarily Stable Strategies. *J. appl. Probab.* **18**, 507.
- HINES, W. G. S. (1987). Evolutionarily Stable Strategies: A Review of Basic Theory. *Theor. Pop. Biol.* **31**, 195 (erratum **33**, 114).
- HINES, W. G. S. & BISHOP, D. T. (1983). On Learning and the Evolutionarily Stable Strategy. *J. appl. Probab.* **20**, 689.
- HOUSTON, A. I. & SUMIDA, B. H. (1987). Learning Rules, Matching, and Frequency Dependence. *J. theor. Biol.* **126**, 289.
- MAYNARD SMITH, J. (1974). The Theory of Games and the Evolution of Animal Conflict. *J. theor. Biol.* **47**, 209.
- MAYNARD SMITH, J. (1982). *Evolution and the Theory of Games*. Cambridge, U.K.: Cambridge University Press.
- MAYNARD SMITH, J. (1988). Can a Mixed Strategy Be Stable in a Finite Population? *J. theor. Biol.* **130**, 247.
- MAYNARD SMITH, J. & PARKER, G. A. (1976). The Logic of Asymmetric Contests. *Anim. Behav.* **24**, 159.
- MAYNARD SMITH, J. & PRICE, G. R. (1973). The Logic of Animal Conflict. *Nature, Lond.* **246**, 15.
- SCHAFFER, M. E. (1988). Evolutionarily Stable Strategies for a Finite Population and a Variable Contest Size. *J. theor. Biol.* **132**, 469.
- SELTEN, R. (1988). *Anticipatory Learning in Two-Person Games. Discussion Paper 4*. Bonn: Institute für Gesellschafts- und Wirtschaftswissenschaften, University of Bonn.
- SUGDEN, R. (1986). *The Economics of Rights, Cooperation, and Welfare*. Oxford; New York: Basil Blackwell.
- TAYLOR, P. & JONKER, L. (1978). Evolutionarily Stable Strategies and Game Dynamics. *Math. Biosci.* **40**, 145.
- ZEEMAN, E. C. (1979). Population Dynamics from Game Theory. In: *Global Theory of Dynamical Systems. Lecture Notes in Mathematics*. Vol. 819. Berlin; New York: Springer-Verlag.
- ZEEMAN, E. C. (1981). Dynamics of the Evolution of Animal Conflicts. *J. theor. Biol.* **89**, 249.