

# Nash Equilibrium and Evolutionary Stability in Large and Finite Populations

Vincent P. CRAWFORD \*

**ABSTRACT.** — This paper studies the correspondence between Nash equilibrium and evolutionary stability in large and finite populations. Whenever the payoff function of the game that describes the simultaneous interaction of the individuals in the population, and thereby determines their fitnesses, is sufficiently continuous, an evolutionarily stable strategy (ESS) in a large population corresponds to a symmetric Nash equilibrium in that game, and a strict, symmetric Nash equilibrium in that game corresponds to a large-population ESS. This correspondence continues to hold, approximately, in finite populations; and it holds exactly for strict pure-strategy equilibria in sufficiently large finite populations.

---

## Équilibre de Nash et stabilité évolutionnaire dans des populations grandes et finies

**RÉSUMÉ.** — L'article étudie la correspondance entre équilibre de Nash et stabilité évolutionnaire dans des populations grandes et finies. Si la fonction de paiement du jeu qui décrit l'interaction des individus dans la population — en fait leur capacité d'adaptation — est suffisamment continue, une stratégie évolutionnairement stable correspond, dans une grande population, à un équilibre symétrique et un équilibre de Nash symétrique strict du jeu correspond à une stratégie évolutionnairement stable d'une grande population. Cette correspondance tient par approximation dans des populations finies. Elle tient exactement pour des équilibres stricts en stratégies pures dans des populations finies et assez grandes.

---

\* V. P. CRAWFORD: Department of Economics, University of California, San Diego, La Jolla, California 92093, USA. This paper is reprinted, with minor changes and additions, from the *Journal of Theoretical Biology*, 145, (1990), pp. 83-94, where it appeared under the title "Nash Equilibrium and Evolutionary Stability in Large- and Finite- Population 'Playing the Field' Models". I am grateful to Steven BOYCE, Yong-Gwan KIM, Dennis SMALLWOOD, Joel SOBEL, and Maxwell STINCHCOMBE for helpful advice, to Yong-Gwan KIM for research assistance, and to the National Science Foundation (USA) for research support under grant SES 8703337.

# 1 Introduction

---

The concept of an evolutionarily stable strategy (ESS) was introduced by MAYNARD SMITH and PRICE [1973] and MAYNARD SMITH [1974] to describe the long-run effects of selection for successful strategies in environments where an individual's expected rate of reproduction, or fitness, is jointly determined by its own strategy and other individuals' strategies. In the basic model, it is assumed that individuals are randomly matched, in pairs, to play a symmetric two-person game; this assumption is referred to below as "pairwise random matching". Individuals are identical except for their strategies for playing the game; these are inherited and fixed for life. Individuals reproduce asexually, with fitnesses determined by their payoffs in the game, and breed true, passing on their strategies unchanged to their offspring. Finally, the population is sufficiently large that each individual's strategy has a negligible effect on the population strategy frequencies, and the realized rate of reproduction of those using each strategy can be identified, via the law of large numbers, with the fitness that equals its mathematical expectation.

Maynard Smith defined an ESS for this model as a strategy (pure or mixed) that, if played by all members of a large monomorphic population, has strictly higher fitness than any mutant strategy that enters the population with sufficiently low frequency. (The definition is essentially the same for polymorphic populations, in which individuals play different strategies in equilibrium, with the qualification that the population must then have higher fitness, on average, than any mutant.) This definition rests on the intuition that, if the members of a population all play an ESS, then mutants that enter the population with low frequency will reproduce more slowly than individuals who play the ESS. As a result, the relative frequencies of the mutants will 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 those frequencies from the payoff matrix alone, an important simplification. This intuition has since been formalized by TAYLOR and JONKER [1978], ZEEMAN [1979], and HINES [1980 *a*, 1980 *b*] (*see also* CRESSMAN and HINES [1984]), who showed how to use the notion of evolutionary stability to characterize the locally asymptotically stable equilibrium values of the strategy frequencies in monomorphic or polymorphic populations, whether they are growing, fixed, or shrinking.

It is well known (*see for example* MAYNARD SMITH [1982, p. 14]) that the concept of evolutionary stability is closely related to the game-theoretic notion of a Nash equilibrium. (A Nash equilibrium is a combination of strategies for the players in a game such that each player's strategy maximizes its payoff, given the strategies of the other player or players. A strict Nash equilibrium is one in which each player's strategy uniquely maximizes its payoff, given the others' strategies. A symmetric Nash equilibrium is one in which each player plays the same strategy.) When individuals' fitnesses are determined by their payoffs in the game played by

matched pairs in the model just described, a large population playing a strategy that does not maximize fitness against itself is vulnerable to a low-frequency mutation that does maximize fitness against that strategy. It follows that an ESS must correspond to a symmetric Nash equilibrium in that game.<sup>1</sup> It can also be shown, conversely, that in this model any strict, symmetric Nash equilibrium corresponds to an ESS in a large population. The results of Taylor and Jonker, Zeeman, and Hines linking evolutionary and dynamic stability can therefore be viewed as an adaptive justification for the Nash equilibrium.

In recent years, the large-population pairwise random matching analysis just summarized has been extended in many ways; HINES [1987] provides a good survey. One important extension allows finite as well as large populations. Another allows individuals to interact in more general ways, under the rubric "playing the field". This paper studies the relationship between evolutionary stability and Nash equilibrium in large- and finite-population models in which individuals play the field.

The playing the field model, introduced by MAYNARD SMITH [1982, pp. 23-27] and attributed there to Hammerstein, generalizes the pairwise random matching model as follows. In the pairwise random matching model, each individual, at each instant, in effect faces a mixed strategy in a two-person game equal to the average of the other individuals' mixed strategies. The individual's fitness is therefore determined by its own strategy and the frequencies of the other individuals' strategies. Moreover, the linearity in probabilities of expected reproduction rates and the probabilistic independence of matching make the fitness function that describes this relationship linear in the individual's own mixed strategy for any given value of the others' strategy frequencies, and linear in the others' strategy frequencies for any given value of the individual's own strategy. The playing the field model maintains the assumptions that each individual's fitness is determined by its own strategy and the frequencies of the other individuals' strategies, and that fitness is (as implied by its definition as an expectation) linear in the individual's own mixed strategy, but relaxes the assumption that fitness is linear in the other individuals' strategy frequencies. This allows the model to describe either simultaneous play, by the entire population, of a single game with very general payoff function – hence the term "playing the field" – or a wide variety of more complex patterns of interaction, including pairwise random matching as a special case.

The correspondence between Nash equilibrium and evolutionary stability in large populations is readily extended from the pairwise random matching model to the playing the field model, with evolutionary stability defined as before and Nash equilibrium defined for the game that describes the

---

1. I say "correspond" rather than "is" because an ESS is defined as a single strategy, with the understanding that it is played by all members of a monomorphic population, but a Nash equilibrium is defined as a combination of strategies, one for each player.

simultaneous interaction of the individuals in the population.<sup>2</sup> Section 2 provides a formal statement and proof of this result. It is shown there that its validity depends on the (previously unstated) assumption that the fitness function satisfies a weak continuity condition, which is implied by linearity in the pairwise random matching model but need not hold in the playing the field model. The importance of continuity is illustrated by discussing some experiments with human subjects conducted by VAN HUYCK, BATTALIO, and BEIL [1990], using a game whose fitness function has a large-population limit with discontinuities that have striking consequences for observed behavior, even in finite populations.

Section 3 studies the correspondence between Nash equilibrium and evolutionary stability in a finite-population playing the field model, building on the recent contributions of SCHAFFER [1988] and MAYNARD SMITH [1988].<sup>3</sup> Schaffer and Maynard Smith argued that the original ESS definition remains adequate for finite populations (with one qualification, discussed in Section 3) provided that individuals' effects on the population strategy frequencies are properly taken into account in comparing the fitnesses of mutants and ESS-players. They showed that in the Hawk-Dove game, the mixed-strategy ESS in a finite population generally differs from the mixed-strategy ESS in a large population (and therefore from the Nash equilibrium), but converges to it as the population grows. Thus, the large-population correspondence between evolutionary stability and Nash equilibrium discussed in Section 2 breaks down in finite populations, but holds approximately for sufficiently large finite populations, at least in this example.

Section 3 extends this approximate correspondence to a wide class of playing the field models with fitness functions that satisfy the weak continuity condition used in Section 2's analysis. As Maynard Smith's and Schaffer's analyses of the Hawk-Dove example suggest, the correspondence between mixed-strategy Nash equilibrium and evolutionary stability is at best approximate in finite populations, however large. However, the correspondence is exact for strict, pure-strategy equilibria in any sufficiently large finite population. By contrast, it is shown by example in Section 3 that a sequence of mixed-strategy finite-population ESS's can converge to

- 
2. See for example MAYNARD SMITH [1982, p. 24] or SCHAFFER [1988, pp. 470-471]. The correspondence for the playing the field model is a true generalization of the one found earlier for the pairwise random matching model: To apply the playing the field model, it is necessary to view individuals' interactions in the pairwise random matching model as a game played simultaneously by the entire population, with the expected payoffs of individuals' strategies evaluated before they are matched. It is easily verified that strategies that are in symmetric Nash equilibrium in that game are also in symmetric Nash equilibrium in the two-person game played by matched pairs, and vice versa.
  3. In Schaffer's and Maynard Smith's finite-population analyses, the realized rate of reproduction of those individuals using each strategy is identified with the fitness that equals its mathematical expectation, even though the law of large numbers does not fully justify this substitution in a finite population. That assumption is also maintained here. Thus, in what follows "large" refers to populations in which each individual's strategy is assumed to have a negligible effect on the population strategy frequencies; "finite" and "large finite" refer to populations in which this is not assumed.

a limit that is not a large-population ESS, and a mixed-strategy large-population ESS need not be the limit of any sequence of finite-population ESS's.

## 2 Large Populations

---

This section studies the correspondence between Nash equilibrium and evolutionary stability in large populations.

The large-population playing the field model can be formalized as follows. Assume that the individuals in the population play symmetric roles in the game that describes their simultaneous interaction; that the game has a finite number of pure strategies; and that the influence on each individual's fitness of other individuals' strategies is completely determined by their frequency distribution, so that it does not matter "who does what" among the other individuals. In a large population, an individual's fitness can be taken to be determined by its own strategy and the population strategy frequencies including the individual's strategy, because excluding the individual's strategy has a negligible effect on the population frequencies. Let the probability vectors  $r$  and  $s$  represent an individual's mixed strategy and the population strategy frequencies, respectively, and let  $E(r|s)$  denote the fitness of an individual playing  $r$  when the population frequencies are given by  $s$ . As noted in Section 1, in the pairwise random matching model  $E(r|s)$  is linear in  $r$  for any given value of  $s$ , and vice versa. It is natural to maintain the assumption that  $E(r|s)$  is linear in  $r$  in the playing the field model, because  $r$  is a mixed strategy and fitness is defined as an expected reproduction rate. There is no reason to assume linearity in  $s$ , however. Here I require only that, if  $p$  is the ESS under consideration, then there exists a neighborhood of  $s=p$  throughout which  $E(r|s)$  is continuous in  $s$  for all  $r$ .

In terms of the notation used here, a large-population ESS can be defined as a mixed strategy,  $p$ , such that for each mixed strategy  $q \neq p$ , there exists an  $\bar{\epsilon}(q) > 0$  for which

$$(1) \quad E(p|(1-\epsilon)p + \epsilon q) > E(q|(1-\epsilon)p + \epsilon q)$$

whenever  $0 < \epsilon < \bar{\epsilon}(q)$ . Thus, as in Maynard Smith's original formulation for the pairwise random matching model, a large-population ESS is a strategy that, if initially played by all members of a monomorphic population, has strictly higher fitness than any mutant strategy that enters the population with sufficiently low frequency, after the mutants enter.

Maynard Smith's original formulation left open the possibility that  $\bar{\epsilon}(q)$ , the frequency of a mutant strategy that is deemed "sufficiently low", might vary with the mutant. It is sometimes useful in what follows to strengthen this definition to require the existence of an  $\bar{\epsilon} > 0$ , independent

of  $q$ , such that (1) holds for all  $q \neq p$  whenever  $0 < \varepsilon < \bar{\varepsilon}$ . It is plain that such an  $\bar{\varepsilon}$  would also meet the requirements on  $\bar{\varepsilon}(q)$  in the definition just stated. However, the existence of an  $\bar{\varepsilon}(q) > 0$  that meets those requirements does not immediately imply that there exists an  $\bar{\varepsilon}$  such that  $0 < \bar{\varepsilon} \leq \bar{\varepsilon}(q)$  for all  $q \neq p$ , so that  $\bar{\varepsilon}(q)$  is bounded above zero as the stronger definition requires. Nevertheless, VICKERS and CANNINGS [1987, Theorem 2] showed that such an  $\bar{\varepsilon}$  must always exist for the pairwise random matching model with a finite number of pure strategies, so that the weak and strong versions of the large-population ESS definition are equivalent for that model. CRAWFORD [1990] generalized this result to the playing the field model for fitness functions that satisfy the above continuity condition, showing that if  $p$  is a large-population ESS by the weak definition, and there exists a neighborhood of  $s=p$  throughout which  $E(r|s)$  is continuous in  $s$  for all  $r$ , then  $p$  is a large-population ESS by the strong definition as well.

The main result of this section, linking Nash equilibrium and evolutionary stability in large-population playing the field models with finite numbers of pure strategies, can now be stated:

**THEOREM 1:** If  $p$  is a large-population ESS (by either definition) and the fitness function,  $E(r|s)$ , is continuous in  $s$  at  $s=p$  for all  $r$ , then  $p$  corresponds to a symmetric Nash equilibrium in the game that describes the simultaneous interaction of the individuals in the population. Conversely, if  $p$  corresponds to a strict, symmetric Nash equilibrium in that game, and there exists a neighborhood of  $s=p$  throughout which  $E(r|s)$  is continuous in  $s$  for all  $r$ , then  $p$  is a large-population ESS (by either definition).

*Proof:* If  $p$  is a large-population ESS and the fitness function  $E(r|s)$  is continuous in  $s$  at  $s=p$  for all  $r$ , then letting  $\varepsilon \rightarrow 0$  in (1) yields

$$(2) \quad E(p|p) \geq E(q|p)$$

for all  $q \neq p$ . Because excluding an individual's strategy has a negligible effect on the strategy frequencies in a large population, (2) implies that  $p$  maximizes any individual's fitness, given that the other individuals are playing  $p$ . Thus,  $p$  corresponds to a symmetric Nash equilibrium in the game that describes the simultaneous interaction of the individuals in the population.

Conversely, if  $p$  corresponds to a strict, symmetric Nash equilibrium in that game, so that (2) holds with strict inequality for all  $q \neq p$ , then the continuity of  $E(r|s)$  in  $s$  at  $s=p$  for all  $r$  implies that, for each  $q \neq p$ , there exists an  $\bar{\varepsilon}(q) > 0$  such that (1) holds whenever  $0 < \varepsilon < \bar{\varepsilon}(q)$ . When  $E(r|s)$  is also continuous in  $s$  throughout a neighborhood of  $s=p$ , there also exists (by the result proved in CRAWFORD [1990]) an  $\bar{\varepsilon} > 0$  such that (1) holds for all  $q \neq p$  whenever  $0 < \varepsilon < \bar{\varepsilon}$ . Thus,  $p$  is a large-population ESS (by either definition).  $\square$

**Remarks:** For games with fitness functions that are continuous everywhere, Theorem 1's conclusion can be inferred directly from MAYNARD SMITH's [1982, pp. 23-24] discussion of the playing the field model. With continuous fitness functions in mind, he defined an ESS as a mixed strategy,

$p$ , such that, for each mixed strategy  $q \neq p$ , either

$$(3) \quad E(p|p) > E(q|p),$$

or both

$$(4) \quad E(p|p) = E(q|p)$$

and (1) for all sufficiently small  $\varepsilon > 0$ . For continuous fitness functions, this definition is clearly equivalent (for either translation of "sufficiently small" considered above) to the definition based on (1) used here. Without continuity, however, a strategy  $p$  might satisfy (3) but violate (1) for some  $q \neq p$  for all  $\varepsilon > 0$ , making it vulnerable to invasion by  $q$ -players at low frequencies even though it qualifies as an ESS by Maynard Smith's definition.

The formulation of the large-population ESS definition based on (1) captures the idea of evolutionary stability even for discontinuous fitness functions, thereby making it possible to study the importance of continuity for the correspondence between Nash equilibrium and evolutionary stability. The proof of Theorem 1 shows that  $E(r|s)$  need only be continuous in  $s$  at  $s=p$  for a large-population ESS to correspond to a Nash equilibrium, or for a strict, symmetric Nash equilibrium to correspond to a large-population ESS for the weak version of the ESS definition. In general, however,  $E(r|s)$  must be continuous throughout a neighborhood of  $s=p$  for the correspondence to hold for the strong version of the definition.

Large-population playing the field models with discontinuous fitness functions may appear to be no more than curiosities. However, they can arise naturally as limits of nonpathological finite-population models. My interest in such models stems from trying to understand the behavior observed in experiments conducted by VAN HUYCK, BATTALIO, and BEIL [1990] (henceforth "VHBB"). In VHBB's treatments A and A', groups of 14-16 human subjects repeatedly played a symmetric game, in each play of which they simultaneously chose an "effort level" that could take any integer value from 1 to 7. The minimum effort level chosen in the group determined the group's total output, which was then shared equally among its members. A subject's effort was costly to him, with higher levels more costly than lower levels. But effort was sufficiently productive that the resulting output shares more than repaid the cost if all subjects chose the same effort level. Thus, each subject's reward was maximized when all chose the highest possible effort level, but because the minimum effort level in the group created a bottleneck, the balance of the others' efforts was wasted if anyone shirked.

In the game just described, any effort level, when chosen by all members of the group, corresponds to a strict, symmetric Nash equilibrium, and all members prefer symmetric equilibria with higher effort levels to those with lower effort levels. As a result, standard game-theoretic equilibrium refinements either do not discriminate among these equilibria, or favor the one with the highest effort level because members unanimously prefer it to the other equilibria (see VHBB [1990], or CRAWFORD [1991]). Nevertheless, in repeated play VHBB's subjects rapidly approached the Nash equilibrium

with the lowest allowable effort level, even though they unanimously preferred any other equilibrium outcome. Thus, any theory that explains VHBB's experimental results must explain this systematic discrimination between the game's strict Nash equilibria.

A learning model that explains the behavior VHBB observed was outlined in CRAWFORD [1991]. Even though the true explanation in this case cannot be literally evolutionary, the dynamics of the learning model have a strong evolutionary flavor. The proposed explanation rests, in part, on the observation that, in the playing the field model that (except for the differences between learning and evolution) describes VHBB's experimental environments, only the Nash equilibrium in which all individuals choose the lowest possible effort level is an ESS: In any population, large or finite, lower effort levels yield the same share of the population's realized output as higher effort levels, but a lower cost. It follows that a mutation to an effort level that is the lowest in the population, after the mutation, then has higher fitness than all other effort levels (and higher than average fitness in a polymorphic population).

Evolutionary stability can discriminate between strict, symmetric Nash equilibria in this game, despite Theorem 1, because its large-population limit has a fitness function that is discontinuous in the population strategy frequencies at some of these equilibria.<sup>4</sup> The fitness function is easily seen to be continuous whenever the frequency of the lowest allowable effort level is positive, hence throughout a neighborhood of the equilibrium at which all individuals choose that effort level. But raising the frequency of the lowest allowable effort level a small amount above zero causes a discontinuous drop in output, hence in each individual's share of output. This makes the fitness function discontinuous whenever the frequency of this effort level is zero, hence at each of the six strict, symmetric equilibria with higher effort levels. That these equilibria do not correspond to ESS's shows that Theorem 1's continuity condition is necessary for this part of the correspondence.

It is important to note that, at least when subjects play only pure strategies (as they appeared to in the experiments), continuity of the fitness function in the population strategy frequencies is not a substantive restriction in the finite populations of VHBB's (or anyone else's!) experiments: the graph of the fitness function in a finite population is a discrete set, which can always be filled in continuously. As a result, the game that describes VHBB's experimental environments is not pathological in any of the senses in which this term is normally used. Nevertheless, the strong fitness interactions that appear as discontinuities in the large-population limit of that environment make evolutionary stability discriminate between equilibria in precisely the same way in finite populations. This suggests that large-population playing the field models with discontinuous fitness functions can serve

---

4. Note that for each individual's fitness in the large-population limit of this environment to be completely determined by its own strategy and the population strategy frequencies, it must depend on the lowest effort level with strictly positive frequency in the population, not on the lowest effort level in the population.



as useful idealizations of environments with large finite populations. In particular, the existence and location of discontinuities in the large-population limit of the fitness function can signal the presence of the kind of interaction that causes finite populations to discriminate between strict, symmetric Nash equilibria.

### 3 Finite Populations

---

This section studies the correspondence between Nash equilibrium and evolutionary stability in playing the field models with finite populations. It is shown, for games whose fitness functions satisfy a continuity condition like the one maintained in Section 2, that the limit of a sequence of finite-population ESS's that converges, as the population grows, must correspond to a symmetric Nash equilibrium in the game that describes the simultaneous interaction of the individuals in the population. Conversely, any strict, symmetric Nash equilibrium corresponds to a finite-population ESS in any sufficiently large finite population. Thus, the correspondence between Nash equilibrium and evolutionary stability in large populations discussed in Section 2 extends, at least approximately, to sufficiently large finite populations. It is also shown however, by example, that a sequence of mixed-strategy finite-population ESS's can converge to a limit that is not a large-population ESS, and that a mixed-strategy large-population ESS need not be the limit of any sequence of finite-population ESS's.

An individual's fitness is in general determined by its own strategy, the population frequencies of the other individuals' strategies, and the size of the population. Assuming for simplicity only that fitness does not depend directly on population size, the fitness function is again denoted  $E(r|s)$ , where  $r$  is the individual's mixed strategy and  $s$  gives the frequencies of the other individuals' strategies.<sup>5</sup> The size of the population is denoted  $N$ .

Assuming for simplicity that any mutants switch strategies from the ESS, so that the size of the population remains fixed at  $N$ , an ESS can be defined

---

5. This notation is consistent with that in Section 2, because excluding an individual's strategy has a negligible effect on the population frequencies in a large population. The assumption that population size does not directly affect fitness is implicitly maintained in most large-population analyses, and is often carried over to finite-population analyses. As noted in CRAWFORD [1991], it is satisfied for VHBB's [1990] experimental environments. However, this section's results and arguments remain valid even when population size directly affects fitness, as long as this effect is continuous in the sense that the sequence of fitness functions parameterized by population size converges, uniformly, to a limiting function as the population grows.

for this model as a mixed strategy,  $p$ , such that for each  $q \neq p$ , either

$$(5) \quad E\left(p \left| \frac{N-2}{N-1} p + \frac{1}{N-1} q \right.\right) > E(q|p),$$

or both

$$(6) \quad E\left(p \left| \frac{N-2}{N-1} p + \frac{1}{N-1} q \right.\right) = E(q|p)$$

and

$$(7) \quad E\left(p \left| \frac{N-3}{N-1} p + \frac{2}{N-1} q \right.\right) > E\left(q \left| \frac{N-2}{N-1} p + \frac{1}{N-1} q \right.\right).$$

Conditions (5) and (6) play the role of condition (1) in the large-population ESS definition, requiring that ESS-players must have fitness at least as great as any mutant that enters the population with sufficiently low frequency – in this case,  $1/N$ .<sup>6</sup> Condition (7) requires that, if the fitness of an ESS-player equals that of a given mutant after it enters the population alone, then an ESS-player must have strictly higher fitness after two such mutants enter the population. Taken together, conditions (5)-(7) ensure that ESS-players have strictly higher fitness than any mutant that enters the population with sufficiently low frequency, with “sufficiently low” taken to mean either  $1/N$  or  $2/N$  depending on whether (5) or (6) is satisfied for the mutant in question.<sup>7</sup>

This finite-population ESS definition is somewhat weaker than the standard large-population definition, in that it allows (6)-(7) instead of requiring (5) unconditionally. SCHAFFER [1988] and MAYNARD SMITH [1988] formulated the definition this way because a nondegenerate mixed strategy – that is, one that assigns positive probability to more than one pure strategy – cannot satisfy the strict inequality in (5) in a finite-population pairwise random matching model. This fact is illustrated by their analyses of the Hawk-Dove example, but it is best understood by noting, following Schaffer, that requiring that  $p$  satisfy either (5) or (6) for each  $q \neq p$  is equivalent to

6. Conditions (5) and (6) differ from conditions (3) and (4) in my restatement of MAYNARD SMITH's [1982, pp. 23-24] ESS definition for the large-population playing the field model in that (5) and (6) assume that mutants have positive frequency.

7. SCHAFFER [1988] and MAYNARD SMITH [1988] did not discuss the choice between requiring (7) if (6) holds and requiring the analog of (7) for any number of identical mutants greater than one (Schaffer's "stability" condition ( $b'$ )). The latter condition is very strong in general, but happens to be satisfied by the finite-population ESS (defined as in the text) in the Hawk-Dove example. The formulation chosen here is closer to the large-population ESS definition, and therefore better suited to studying the relationship between large- and finite-population models. The definition in the text may make it possible to extend to sufficiently large finite populations the correspondence between evolutionary and local dynamic stability established for large populations by TAYLOR and JONKER [1978], ZEEMAN [1979], and HINES [1980 *a*, 1980 *b*].

requiring that setting  $q=p$  solves the problem

$$(8) \quad \max_{q \in Q} \left[ E(q|p) - E\left(p \left| \frac{N-2}{N-1}p + \frac{1}{N-1}q \right. \right) \right],$$

where  $Q$  is the simplex of feasible mixed strategies: the objective function in (8) then takes the value zero when  $q=p$ , and is otherwise less than or equal to zero, as (5) or (6) require.

With pairwise random matching,  $E(r|s)$  is linear in  $r$  for any given value of  $s$ , and vice versa. The objective function of problem (8) is therefore linear in  $q$ ; hence its solution cannot be unique, as (5) requires, unless it is a pure strategy. This problem does not arise in large populations because mutants that enter a large population with positive frequency are sufficiently numerous to interact with each other, no matter how low their frequency.<sup>8</sup> The finite-population ESS definition proposed by Schaffer and Maynard Smith admits interactions between mutants by allowing  $p$  to solve problem (8) nonuniquely, as long as any alternative solution satisfies the two-mutant "stability" condition (7).

The main result of this section, linking Nash equilibrium and evolutionary stability in finite-population playing the field models with finite numbers of pure strategies, can now be stated:

**THEOREM 2:** If a sequence of finite-population ESS's converges, as the population grows, to a limit,  $p$ , such that the fitness function,  $E(r|s)$ , is continuous in  $s$  at  $s=p$  for all  $r$ , then  $p$  corresponds to a symmetric Nash equilibrium in the game that describes the simultaneous interaction of the individuals in the population. Conversely, if  $p$  corresponds to a strict, symmetric Nash equilibrium in that game and there exists a neighborhood of  $s=p$  throughout which  $E(r|s)$  is continuous in  $s$  for all  $r$ , then  $p$  is a finite-population ESS in any sufficiently large finite population.

*Proof:* Recall that, under the maintained assumptions, the same fitness function can be used to describe the relationship between an individual's fitness and individuals' strategies in large and finite populations, and that in either case,  $E(r|s)$  gives the fitness of  $r$  when the frequencies of the other individuals' strategies are given by  $s$ . If  $p$  satisfies the conditions in the first part of the theorem, then letting  $N \rightarrow \infty$  in (5)-(6) and using the continuity of  $E(r|s)$  at  $s=p$  for all  $r$  yields

$$(9) \quad E(p|p) \geq E(q|p)$$

8. The analogous large-population ESS characterization would require only that  $q=p$  uniquely solves  $\max_{q \in Q} [E(q|(1-\epsilon)p + \epsilon q) - E(p|(1-\epsilon)p + \epsilon q)]$  for any sufficiently small  $\epsilon > 0$ . Interactions between mutants make this problem's objective function nonlinear in  $q$  in general, even with pairwise random matching. This difference between large and finite populations is not an artifact of my notational conventions, because  $E(q|(1-\epsilon)p + \epsilon q)$  can be interpreted, in a large population, as the fitness of a  $q$ -player when the population strategy frequencies, excluding its strategy, are  $(1-\epsilon)p + \epsilon q$ .

for all  $q \neq p$ . It follows that  $p$  corresponds to a symmetric Nash equilibrium in the game that describes the simultaneous interaction of the individuals in the population.

Conversely, if  $p$  corresponds to a strict, symmetric Nash equilibrium, so that (9) holds with strict inequality for all  $q \neq p$ , then the continuity of  $E(r|s)$  in  $s$  at  $s=p$  for all  $r$  implies that, for each  $q \neq p$ , there exists an  $\underline{N}(q)$  such that (5) holds whenever  $N > \underline{N}(q)$ . If  $E(r|s)$  is also continuous in  $s$  throughout a neighborhood of  $s=p$ , then, by the argument of CRAWFORD [1990], there also exists an  $\underline{N}$ , independent of  $q$ , such that (5) holds for all  $q \neq p$  whenever  $N > \underline{N}$ . Thus,  $p$  is an ESS in any sufficiently large finite population.  $\square$

**Remarks:** Given the discussion in Section 2, it may seem puzzling that the continuity of the fitness function has implications for the correspondence between Nash equilibrium and evolutionary stability in finite populations. But even though continuity of the fitness function is not a substantive restriction in any given finite population, the existence of a continuous function that describes how fitness depends on the population strategy frequencies for arbitrarily large values of  $N$  is a substantive restriction.<sup>9</sup>

Because pure-strategy Nash equilibria are normally strict, and strict, symmetric Nash equilibria are normally large-population ESS's, Theorem 2 shows that the limit of a convergent sequence of pure-strategy finite-population ESS's is normally a pure-strategy large-population ESS, and vice versa. But the theorem is silent on whether this conclusion remains valid for mixed-strategy ESS's, which cannot correspond to strict Nash equilibria. I now show that this conclusion does not extend to mixed-strategy ESS's, by presenting two examples: one with a sequence of mixed-strategy finite-population ESS's whose limit is not a large-population ESS, and one with a large-population ESS that is not the limit of any sequence of finite-population ESS's.

The first example is a pairwise random matching model, with the payoff matrix of the game played by matched pairs given by

	H	D	M
H	0	2	1
D	1	0	1
M	0	2	2

In the matrix, only the Row player's payoffs are shown; the Column player's payoffs can be determined by symmetry. The upper left  $2 \times 2$  block of this payoff matrix is a special case of the Hawk-Dove game analyzed by SCHAFER [1988] and MAYNARD SMITH [1988]. They showed that this  $2 \times 2$  game has a unique large-population ESS at  $h=2/3$ , where  $h$  and  $1-h$  are the

9. Similarly, when fitness depends directly on population size, the existence of a fitness function that depends continuously on both the population strategy frequencies and population size (in the sense of footnote 5) is a substantive restriction.

probabilities of playing H and D, respectively, and a unique finite-population ESS, for all  $N \geq 3$ , at  $h = h_N^*$ , where  $h_N^* \equiv (2N-3)/(3N-6)$ . I now argue that  $(h_N^*, 1-h_N^*, 0)$  is a finite-population ESS in the  $3 \times 3$  game displayed above for any  $N \geq 3$ , but that its limit as  $N \rightarrow \infty$ ,  $(2/3, 1/3, 0)$ , is not a large-population ESS in that game.

To verify that  $(h_N^*, 1-h_N^*, 0)$  is a finite-population ESS in the  $3 \times 3$  game for any  $N \geq 3$ , note first that when  $p = (h_N^*, 1-h_N^*, 0)$  and  $q = (h, 1-h, 0)$ , with  $0 \leq h \leq 1$ , then

$$\begin{aligned}
 (10) \quad E(q|p) &= 2h \frac{N-3}{3N-6} + (1-h) \frac{2N-3}{3N-6} \\
 &= \frac{N-2}{N-1} \left[ 3 \frac{2N-3}{3N-6} \frac{N-3}{3N-6} \right] + \frac{1}{N-1} \left[ 2 \frac{2N-3}{3N-6} (1-h) + \frac{N-3}{3N-6} h \right] \\
 &= E \left( p \left| \frac{N-2}{N-1} p + \frac{1}{N-1} q \right. \right).
 \end{aligned}$$

When, however,  $q = (0, 0, 1)$ ,

$$\begin{aligned}
 (11) \quad E(q|p) &= 2 \frac{N-3}{3N-6} < \frac{N-2}{N-1} \left[ 3 \frac{2N-3}{3N-6} \frac{N-3}{3N-6} \right] + \frac{1}{N-1} \\
 &= E \left( p \left| \frac{N-2}{N-1} p + \frac{1}{N-1} q \right. \right).
 \end{aligned}$$

Thus, given the linearity of the objective function of problem (8), a mixed strategy  $q \in Q$  solves (8) when  $p = (h_N^*, 1-h_N^*, 0)$  if and only if it is of the form  $(h, 1-h, 0)$ , with  $0 \leq h \leq 1$ . Because  $(h_N^*, 1-h_N^*, 0)$  takes this form, it is a finite-population ESS in this game provided that each alternative solution of (8) satisfies the subsidiary stability condition (7). It therefore follows from Schaffer's and Maynard Smith's demonstration that  $(h_N^*, 1-h_N^*)$  is a finite-population ESS in the  $2 \times 2$  Hawk-Dove game that  $(h_N^*, 1-h_N^*, 0)$  is a finite-population ESS in the  $3 \times 3$  game displayed above for any  $N \geq 3$ .

It remains to show that  $(2/3, 1/3, 0)$  is not a large-population ESS in the  $3 \times 3$  game. It is instructive to take an indirect route to this result. A player's strategy (pure or mixed) is said to be dominated by another of that player's strategies if its expected payoff is never higher, and sometimes lower, than the other strategy's. For instance, M dominates H in the  $3 \times 3$  game displayed above. A dominated strategy cannot be a large-population ESS in either the pairwise random matching model or the playing the field model, because a strategy that dominates it has fitness at least as large when both face the same population strategy frequencies, contradicting the strict inequality in (1). Because M dominates H in the  $3 \times 3$  game, the mixed strategy  $(0, 1-h, h)$  dominates the mixed strategy  $(h, 1-h, 0)$  whenever  $0 < h \leq 1$ . Thus,  $(0, 1/3, 2/3)$  dominates  $(2/3, 1/3, 0)$ , and  $(2/3, 1/3, 0)$  cannot be a large-population ESS, even though it is the limit of a sequence of finite-population ESS's. By contrast, the dominated strategy  $(h_N^*, 1-h_N^*, 0)$  is an ESS in any finite population.

The second example shows that a large-population ESS need not be the limit of any sequence of finite-population ESS's. Thus, the condition in the second part of Theorem 2 that  $p$  corresponds to a strict, symmetric Nash equilibrium cannot be weakened to allow  $p$  to be a large-population ESS.

The example is also a pairwise random matching model, this time with payoff matrix

	H	D	M
H	0	2	0
D	1	0	1
M	1	0	0

The analysis of the Hawk-Dove game in the upper-left  $2 \times 2$  block is as before. It is easily verified that the large-population ESS  $(2/3, 1/3)$  in that  $2 \times 2$  game is also a large population ESS in the  $3 \times 3$  game displayed above: if a mutant strategy that assigns positive probability to M enters the population with low frequency, it must have strictly lower fitness than the large-population ESS in the Hawk-Dove game, because D dominates M and the large-population ESS equates the fitnesses of H, D, and the ESS.

The only possible candidate for a finite-population ESS that converges to  $(2/3, 1/3, 0)$  is  $(h_N^*, 1 - h_N^*, 0)$ , but this strategy is not a finite-population ESS for any  $N \geq 3$  in the  $3 \times 3$  game under consideration. To see this, note that if  $p = (h_N^*, 1 - h_N^*, 0)$  and  $q = (0, 0, 1)$ , then

$$(12) \quad E(q|p) = \frac{2N-3}{3N-6} > \frac{N-2}{N-1} \left[ 3 \frac{2N-3}{3N-6} \frac{N-3}{3N-6} \right] + \frac{1}{N-1} \frac{N-3}{3N-6}$$

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

so that  $p = (h_N^*, 1 - h_N^*, 0)$  satisfies neither (5) nor (6) for  $q = (0, 0, 1)$ .

These examples are not pathological, but they are nongeneric. I suspect that the failures of continuity of the ESS correspondence that they illustrate are, in fact, nongeneric. I am inclined to view these failures as a criticism of Maynard Smith's original large-population ESS definition, not of the finite-population definition formulated by SCHAFFER [1988] and MAYNARD SMITH [1988]. However, only a finite-population generalization of the dynamic analyses of TAYLOR and JONKER [1978], ZEEMAN [1979], and HINES [1980 a, 1980 b] will show which approach is more useful.

## • References

- CRAWFORD, V. P. (1990). — "On the Definition of an Evolutionarily Stable Strategy in the 'Playing the Field' Model", *Journal of Theoretical Biology*, 143, pp. 269-273.
- CRAWFORD, V. P. (1991). — "An 'Evolutionary' Interpretation of Van Huyck, Battalio, and Beil's Experimental Results on Coordination", *Games and Economic Behavior*, 3, pp. 25-59.

- CRESSMAN, R. and HINES, W. G. S. (1984). – “Correction to the Appendix of ‘Three Characterizations of Population Strategy Stability’”, *Journal of Applied Probability*, 21, pp. 213-214.
- HINES, W. G. S. (1980a). – “Three Characterizations of Population Strategy Stability”, *Journal of Applied Probability*, 17, pp. 333-340.
- HINES, W. G. S. (1980b). – “Strategy Stability in Complex Populations”, *Journal of Applied Probability*, 17, pp. 600-610.
- HINES, W. G. S. (1987). – “Evolutionarily Stable Strategies: A Review of Basic Theory”, *Theoretical Population Biology*, 31, pp. 195-272; “Erratum”, 33, p. 114.
- MAYNARD SMITH, J. (1974). – “The Theory of Games and the Evolution of Animal Conflicts”, *Journal of Theoretical Biology*, 47, pp. 209-221.
- 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?”, *Journal of Theoretical Biology*, 130, pp. 247-251.
- MAYNARD SMITH, J. and PRICE, G. R. (1973). – “The Logic of Animal Conflict”, *Nature*, London, 246, pp. 15-18.
- SCHAFFER, M. E. (1988). – “Evolutionarily Stable Strategies for a Finite Population and a Variable Contest Size”, *Journal of Theoretical Biology*, 132, pp. 469-478.
- TAYLOR, P. and JONKER, L. (1978). – “Evolutionarily Stable Strategies and Game Dynamics”, *Mathematical Biosciences*, 40, pp. 145-156.
- VAN HUYCK, J., BATTALIO, R. and BEIL, R. (1990). – “Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure”, *American Economic Review*, 80, pp. 234-248.
- VICKERS, G. T. and CANNINGS, C. (1987). – “On the Definition of an Evolutionarily Stable Strategy”, *Journal of Theoretical Biology*, 129, pp. 349-353.
- ZEEMAN, E. C. (1979). – “Population Dynamics from Game Theory”, In *Global Theory of Dynamical Systems*, Lecture Notes in Mathematics, vol. 819. Berlin and New York: Springer-Verlag.