On the Evolution of Optimizing Behavior*

EDDIE DEKEL AND SUZANNE SCOTCHMER

*Economics Department and Graduate School of Public Policy,
University of California, Berkeley, California 94720*

We investigate whether replicator dynamics justify "survival of the fittest" when players inherit pure strategies, mixed strategies, or rules for choosing strategies. A strategy that never is a best reply can survive replicator dynamics if only pure strategies are heritable, but only rationalizable strategies survive if mixed strategies are heritable. Now suppose agents can inherit the rule of choosing a best reply to the last period's population, rather than inheriting a strategy itself. Such optimizers need not drive out players who inherit a dominated strategy. If we interpret replicator dynamics as a learning process, this means that non-optimizing behavior can survive. *Journal of Economic Literature* Classification Number: C73. © 1992 Academic Press, Inc.

1. INTRODUCTION

This paper investigates the evolutionary justification for "survival of the fittest." We provide an example where a strategy that is not fittest against any population, a "never-fittest strategy," survices. We discuss two interpretations of this result: the biological interpretation that evolution need not select the fittest, and the economic interpretation that evolution may fail to select for rationality.

We model evolution using replicator dynamics, which are based on the following biological model. In each generation each player is randomly matched with an opponent and they play a one-shot normal-form game where payoffs represent numbers of offspring. Each offspring inherits from its parent a rule that determines its strategy in the game, so successful rules proliferate.

Evolutionary models have typically restricted attention to very simple heritable rules, of the form "play s," where s is a pure strategy in the game. Recently some authors (Hines [14, 15], Zeeman [28], van Damme [7], and Robson [24]) have assumed that players can inherit mixed strategies as well. In addition to considering these two cases, we introduce a broader perspective: Players can inherit rules that determine strategies in the one-shot game as a function of the history of play (as in learning models).[1] In addition to broadening the scope of evolutionary models, this should allow one to extend the focus of learning models. These models typically ask what strategies result from exogenously given learning rules. Allowing for heritable rules can endogenize the determination of rules and can model which rules survive.[2]

The outcome of evolution depends on what is heritable. In Section 2 we show by example that if players inherit pure strategies, then never-fittest strategies can survive evolutionary forces. On the other hand, we show in Proposition 2 that if players inherit mixed strategies, then never-fittest strategies will become extinct. In Section 3 we assume that some players inherit a never-fittest strategy while others choose best replies to last period's population. Although naive, this rule seems more complex and rational than simply inheriting strategies. Nevertheless we show by example that if fails to drive out the never-fittest strategy.

It turns out that our first example relies on discrete time dynamics. Samuelson and Zhang [26] show that in continuous time the replicator dynamics with heritable pure strategies extinguish never-fittest strategies.[3] Cabrales and Sobel [5] provide an informative reconciliation of our results with theirs, and develop sufficient conditions (analogues to conditions provided by Samuelson and Zhang [26]) under which discrete time dynamics extinguish never-fittest strategies.

The biological interpretation of our results is clear: Evolution does not guarantee "survival of the fittest" since, when the fittest strategy depends on

---

[1] One could also consider the opposite approach. Whereas we allow players to inherit more complicated objects than just pure strategies, one might want to restrict attention to players who can inherit only a strict subset of the set of pure strategies. This seems most natural when the one-shot game is complicated, e.g., if it itself is an extensive form game. In interesting models of evolution of play in infinitely-repeated games, Binmore and Samuelson [3] and Fudenberg and Maskin [12] assume that the only heritable strategies are those that can be played by finite automata.

[2] Of course, the process selecting among rules, namely the replicator dynamics, is exogenous. Thus, we only move the exogenous assumption one step back.

[3] Our example in Section 3 does not have a continuous time analog, since some players form beliefs based on play in the immediately preceding period. However we can ask what happens in the limit, as the time period is made short (as in Cabrales and Sobel [5]). In this example, making the time period short does not effect our result: the never-fittest strategies survive regardless of the time period.

what other agents are doing, a never-fittest strategy might survive. This failure is distinct from, and possibly more severe than, a failure discussed in the biology literature, that evolution selects for local, not global, optima.

The economics interpretation of the model is less straightforward. If the payoffs in the game are interpreted as utility payoffs, then we have shown how strategies that never are best replies can survive. Thus, evolution fails to imply that people play "as if" rational since, whatever their beliefs, rational players would only choose best replies.

However, the motivation for replicator dynamics relies on the idea that payoffs represent reproductive success, not utility. There are two routes to justifying the economics interpretation. The simplest is to identify utility with fitness payoffs: assume that higher utility coincides with more children, and each chile inherits (either genetically, or by learning) its strategy or rule for choosing a strategy from its parent. Emulation of parents might be plausible in complex environments, as when children follow in their parents' footsteps. A second way to justify payoffs as utility is based on the view that replicator dynamics are a model of learning with bounded rationality. While popular, this justification has not yet been formalized, and such a formalization lies outside the scope of this paper. Nevertheless, it is useful to identify the features of replicator dynamics that the learning process must obey: the dynamic path shifts weight toward better-than-average strategies even if they are not best replies; worse-than-average strategies are not totally abandoned; and best replies might change along the dynamic path (so the path need not converge to a steady state). While we do not provide a learning model with these features, we feel that they are sufficiently plausible to justify the economics interpretation of our results: rational choice (in the sense of choosing best replies) is not easily justified as the outcome of evolutionary dynamics or as the limit of boundedly rational learning.

2. Better-than-Average Strategies versus Best Strategies

We begin by describing replicator dynamics, then present our example in which a strategy that never is a best reply survives, and conclude this section with our positive result: when mixed strategies are heritable, nonrationalizable strategies (as defined by Bernheim [2] and Pearce [23]) will vanish.

Consider a symmetric game with pure strategies described by the set $S = \{1, ..., n\}$, and payoffs given by the matrix $\Pi = (\pi_{ij})_{i=1, j=1}^{n}$. That is, if player 1 plays pure strategy i and player 2 plays pure strategy j, then the payoffs are π_{ij} to player 1, and π_{ji} to player 2. We will let Σ, with generic element σ, be the simplex that represents mixed strategies. We will let e_i

represent the ith vertex of Σ, and will refer to the distribution of strategies in period t as σ^t, so that the proportion of people playing strategy i in period t is σ_i^t.

Assume that people play pure strategies, players (asexually) reproduce (without mutations) according to π_{ij} and players are randomly paired from the population, which involves a continuum of people. These assumptions motivate the following replicator dynamics:[4]

$$\sigma_i^{t+1}/\sigma_i^t = (e_i \cdot \Pi \cdot \sigma^t)/(\sigma^t \cdot \Pi \cdot \sigma^t). \tag{1}$$

It is convenient to define the function $f: \Sigma \to \Sigma$ that describes the dynamic path determined by Eq. (1), $f(\sigma^t) = \sigma^{t+1}$. Typically the initial distribution, σ^0, is assumed to assign strictly positive weight to all heritable strategies. This captures the idea that mutations may introduce any strategy ($\sigma_i^0 > 0$ for all i), and that mutations are rare relative to the adjustment time of the dynamic process (so that mutations at any time t are not modeled).[5] In the replicator dynamics, if strategy i does better than the population average in period t, so that $(e_i \cdot \Pi \cdot \sigma^t) > (\sigma^t \cdot \Pi \cdot \sigma^t)$, then strategy i becomes more prevalent. The best reply to the population average (i.e., the fittest strategy in any given period) will have the highest rate of growth, since the numerator in (1) is largest for the strategy that is a best reply. However, other strategies may also grow, provided they are better than average.

EXAMPLE 1. Now we provide an example with four strategies A, B, C, and D. The proportions of people playing these strategies (i.e., population distributions) are denoted by $\sigma = (\alpha, \beta, \gamma, \delta)$. The payoffs are chosen so that for any population distribution σ there is a strategy (A, B, or C) that does better than D does against σ. We will show that, even though there is always a rule in the population that grows faster than the rule "play D," dynamic paths from a generic set of initial populations do not have the property that δ converges to zero.

The example augments a rock–scissors–paper game by adding a strategy D that is not a best reply to any mixed or pure strategy.

$$\Pi = \begin{array}{c|cccc} & A & B & C & D \\ \hline A & 1 & 2.35 & 0 & 0.1 \\ B & 0 & 1 & 2.35 & 0.1 \\ C & 2.35 & 0 & 1 & 0.1 \\ D & 1.1 & 1.1 & 1.1 & 0 \end{array}$$

---

[4] For more discussion of these dynamics see, for example, van Damme [7], Friedman [9], and Nachbar [22].

[5] Some recent papers assume that mutations can occur along the dynamics path as well as initially, see, for example, Foster and Young [8] and Kandori, Mailath, and Rob [20].

Strategy D survives because, under replicator dynamics, it proliferates if it does better than the current population average, even when it is not a best reply to the current population, or, in fact, to any population. An intuitive argument is as follows. A dynamic path, beginning from any point where $\delta = 0$, except where $\alpha = \beta = \gamma = 1/3$, will spiral outward, approaching the boundary. This follows from well-known properties of rock–scissors–paper games; for appropriate payoffs the discrete dynamics diverge (Hofbauer and Sigmund [17]). Intuitively, therefore, if δ' becomes small and the path approaches any point where $\delta = 0$ except $(1/3, 1/3, 1/3, 0)$, the dynamic path will spiral out. Now, as the path spirals close to the boundary the strategy that is a best reply changes, but whichever it is $(A, B,$ or $C)$, it is scarce. Hence it constitutes a small part of the population average. Thus D can be better than average even though it is not a best reply. Simple calculations show that, in fact, it is better than average. Hence δ' is growing, conflicting with the premise that it converges to zero.[6] The only other possibility is for the path to converge to $(1/3, 1/3, 1/3, 0)$. But the only way to get there is from a path along which $\alpha' = \beta' = \gamma'$.

PROPOSITION 1.   *In Example* 1, *a dynamic path beginning at* $(\alpha^0, \beta^0, \gamma^0, \delta^0)$, *with* $\delta^0 > 0$, *has* δ' *converging to zero if and only if* $\alpha^0 = \beta^0 = \gamma^0$.

Proof.   Let Σ_0 be the face of the simplex where $\delta = 0$, $\Sigma_0 = \{(\alpha, \beta, \gamma, \delta) : (\alpha, \beta, \gamma, \delta) \in \Sigma, \delta = 0\}$. Let L be the line in the simplex where the first three strategies get equal weight: $L = \{(\alpha, \beta, \gamma, \delta) \in \Sigma : \alpha = \beta = \gamma\}$. These two sets are shown in Fig. 1.
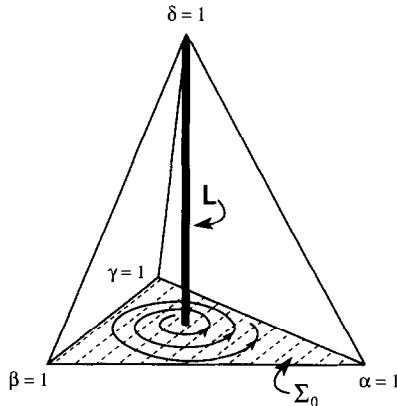

FIGURE 1

[6] In the matrix Π the payoff to playing C against A is 2.35. This payoff must be strictly greater than 2.3 for D never to be a best reply, and it must be strictly less than 2.4 for D to be better than the population average for any population near the boundary. The last condition is not necessary—for example, the result holds with the payoff of 2.4—but it simplifies the proof.

Only If. We prove the result using the following steps.

1. If δ^t converges to zero, then either the boundary of Σ_0 contains an accumulation point of the dynamic path or $(\alpha^t, \beta^t, \gamma^t)$ converges to $(1/3, 1/3, 1/3)$.

2. If the dynamic path has an accumulation point on the boundary of Σ_0, then δ^t does not converge to zero.

3a. If $(1/3, 1/3, 1/3, 0)$ is an accumulation point of a path, then the path lies in L for all t sufficiently large.

3b. If $\sigma^\tau \in L$, then for all t, $\sigma^t \in L$.

Therefore a dynamic path that starts anywhere in Σ except on the line L cannot have δ^t converging to zero. Further, if any point of a path is on L, the entire path is on L.

The first step follows from a basic result on continuous dynamic systems (see Arrowsmith and Place [1, p. 17]): Let σ'' be an accumulation point of a given dynamic path, say $\{\sigma^t\}$, and suppose σ' is an accumulation point on a dynamic path that starts at σ''. Then σ' is also an accumulation point of the dynamic path $\{\sigma^t\}$. (This result follows from continuity of the map that defines the replicator dynamics.) Consider now a dynamic path $\{\sigma^t\}$ with $\delta^t \to 0$. It follows from compactness that $\{\sigma^t\}$ has an accumulation point in Σ_0. As mentioned above (see Hofbauer and Sigmund [17]), the dynamic path starting from any point in Σ_0 other than the center has an accumulation point on the boundary. Thus, using the basic result on continuous dynamic systems, if $\{\sigma^t\}$ does not have an accumulation point only at the center, then it has an accumulation point on the boundary.

For the proof of step 2 it is helpful to think of the dynamic path given by $\rho^t = (\phi^t, \delta^t)$, where $\phi^t = \alpha^t \cdot \beta^t \cdot \gamma^t$. See Fig. 2. The proof involves three substeps.

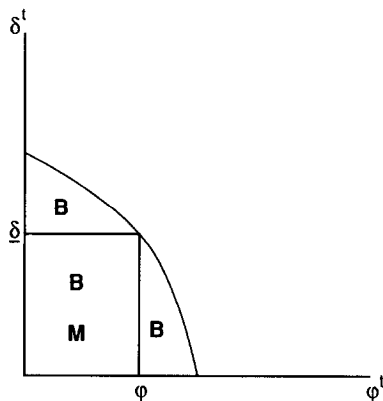

FIGURE 2

(i) There exists a closed ball with non empty interior, $B \subset [0, 1] \times [0, 1]$, that includes the origin, such that for all $\rho' \in B$, $\phi'^{+1} < \phi'$ and $\delta'^{+1} > \delta'$. This is proved as follows. For points in the interior of the simplex where $\phi > 0$ and $\delta > 0$, $\phi'^{+1}/\phi' = [\mathbf{X}_{i=\alpha, \beta, \gamma} (e_i \cdot \Pi \cdot \sigma')]/(\sigma' \cdot \Pi \cdot \sigma')^3$. This is a continuous function which is bounded below one for all points σ on the boundary of Σ_0. Therefore, in a neighborhood, G, near the boundary of Σ_0, $\phi'^{+1}/\phi' < k$ where $k < 1$. Similarly, since D is better than the population average for all populations on the boundary of Σ_0, there is a neighborhood, H, of the origin for which $\delta'^{+1}/\delta' > p$ where $p > 1$. The desired B is given by $G \cap H$.

(ii) Choose $\underline{\phi} > 0$ and $\underline{\delta} > 0$ such that: (a) the set M defined by $[0, \underline{\phi}] \times [0, \underline{\delta}]$ is a subset of B; and (b) $\phi' < \underline{\phi}$ for infinitely many t. We can satisfy (b) because of the hypothesis that the dynamic path has an accumulation point on the boundary of Σ_0.

(iii) If $\phi^T < \underline{\phi}$ and $\delta^T < \underline{\delta}$, then there exists $\delta' > 0$ and $t > T$ such that $\delta' > \delta'$. If ρ^τ is in M for all $\tau > T$, then, since $M \subset B$, δ^τ is increasing and converges to some $\delta^* > 0$. If ρ^τ leaves M, it must do so with δ' going above $\underline{\delta}$. Otherwise the path would violate the condition that, within B, ϕ is decreasing and δ is increasing. To conclude, let $\delta' = \max\{\delta^*, \underline{\delta}\}$.

Thus we have shown that, when $\{\sigma'\}$ gets sufficiently close to the boundary of Σ_0, δ' cannot remain small. Together with step 1, this implies that for δ to converge to 0, the dynamic path must be converging to the unique Nash equilibrium, $(1/3, 1/3, 1/3, 0)$. But step 3 below implies that only paths starting from the line L will converge to the Nash equilibrium.

The proof of step 3a relies on the Invariant Manifolds Theorem (Arrowsmith and Place [1, p. 68]). To apply this theorem, we recall that the space (domain and range) of a linear hyperbolic system has a direct sum decomposition into stable and unstable invariant manifolds (Arrowsmith and Place [1, p. 64]), whose dimensions sum to the dimension of the space. (A hyperbolic fixed point is one at which no eigenvalues of the Jacobian Df have modulus equal to one. The stable invariant manifold is the union of all convergent dynamic paths. The unstable invariant manifold is the union of all nonconvergent paths for which the inverse path converges.) The Invariant Manifold Theorem allows us to study the behavior of a nonlinear dynamical system in a neighborhood of a hyperbolic fixed point σ by studying the stable and unstable manifolds of its linearization, Df_σ, at that point: The locally stable and unstable invariant manifolds of the nonlinear system have the same dimensions as those of the linear system. In the nonlinear system, the line L is a stable invariant set and Σ_0 is an unstable invariant set. Since the dimensions sum to the dimension of the space, namely R^3, the stable invariant manifold has dimension one, and is therefore L: The only paths that converge lie entirely on L.

Since the replicator dynamics system turns out to be hyperbolic at $\sigma = (1/3, 1/3, 1/3, 0)$, this completes the argument for step 3a. Step 3b claims that if a path begins on L it stays on L, and points on L can only be reached from L. The first claim can be seen directly from the replicator dynamics, Eq. (1). The second claim follows from a straightforward application of the inverse function theorem: f is invertible starting at any point on L, and can therefore be reached from only one point, which must also lie on L.

If. We have already argued that a path that begins on L stays on L. It follows from the replicator dynamics in Eq. (1) that δ is strictly decreasing when the distribution of strategies starts on the line L. Therefore, such a path has a limit point, and the limit point is a Nash equilibrium (see, e.g., Nachbar [22]). Hence the limit point is $(1/3, 1/3, 1/3, 0)$. Q.E.D.

We now develop the positive result for the case of heritable mixed strategies. If a strategy is never a best reply, then there is some (possibly mixed) strategy that does better against every population (Pearce [23]). Individuals playing strategies in a neighborhood of this dominating strategy will replicate faster than, and hence drive to extinction, those individuals who play the dominated strategy. If only pure strategies can be inherited, as in the above example, then, whenever the dominating strategy is a mixed strategy, the dominating strategy will not be played by any person in the population; thus the previous argument could not be applied.

To develop the argument, we need some additional notation. Previously we used the vector $\sigma \in \Sigma$ to represent the distribution of pure strategies in the population. We now assume individuals can play mixed strategies $\sigma \in \Sigma$ themselves, and we need another notion—a distribution on Σ—to describe the distribution of mixed strategies in the population. We will let $\Theta \subset 2^\Sigma$ be the Borel field of mixed strategies in Σ, with generic elements denoted by θ. We will let $\Delta(\Sigma)$ be the set of probability measures on the Borel field of Σ, with generic element μ. The probability measure $\mu^t \in \Delta(\Sigma)$ represents the distribution of mixed strategies in the population at time t, and $\mu^t(\theta)$ is the proportion of the population playing mixed strategies in $\theta \in \Theta$.

To describe the replicator dynamics with mixed-strategy inheritance, we need notation for the average strategy played by players whose strategies lie in a particular subset θ of Σ. This average depends on μ, and is given by the function $\tau : \Delta(\Sigma) \times \Theta \to \Sigma$, where $\tau(\mu, \theta) \equiv \int_{\sigma \in \theta} \sigma \, d\mu / \mu(\theta)$. The weight assigned to pure strategy j by $\tau(\mu, \theta)$ is then $\tau_j(\mu, \theta)$. The replicator dynamics are given by the following equation, where θ is a measurable subset of Σ:

$$\mu^{t+1}(\theta) = \frac{[\tau(\mu^t, \theta) \cdot \Pi \cdot \tau(\mu^t, \Sigma)]}{[\tau(\mu^t, \Sigma) \cdot \Pi \cdot \tau(\mu^t, \Sigma)]} \cdot \mu^t(\theta). \tag{2}$$

It is easy to verify that if $\mu^t \in \Delta(\Sigma)$ then $\mu^{t+1} \in \Delta(\Sigma)$.

With this definition of the replicator dynamics, we now can state the result.[7] The next proposition states that for large t, the population will, in aggregate, play dominated strategies with arbitrarily small probability.

PROPOSITION 2. *Suppose that pure strategy j is never a best reply, so that it is dominated by a (possibly mixed) strategy $\hat{\sigma}$ that assigns zero probability to strategy j. Then, if μ_0 has full support, $\lim_{t \to \infty} \tau_j(\mu^t, \Sigma) = 0$.*

Proof. We will first show that for sufficiently large t and for $\psi > 0$, the dynamic path assigns arbitrarily small weight to the set of mixed strategies with $\sigma_j > \psi$, or almost all its weight to the subset of mixed strategies with $\sigma_j < \psi$. Since this is true for any $\psi \in (0, 1]$, the replicator dynamics remove population weight from strategies that place positive weight on strategy j. See Fig. 3.

We need to define a function r that "replaces pure strategy j with $\hat{\sigma}$ in each strategy σ." This function is given by $r(\sigma) = \sigma - \sigma_j e_j + \sigma_j \hat{\sigma}$. The proof of the following claim is straightforward, hence omitted.

Claim 1. If strategy j is strongly dominated by $\hat{\sigma}$, then for any σ such that $\sigma_j > 0$, σ is strongly dominated by $r(\sigma)$.

Choose ψ, $0 < \psi \leqslant 1$. Partition Σ into two subsets, Σ^A and Σ^B, where $\Sigma^B = \{\sigma \in \Sigma : \sigma_j < \psi\}$ and Σ^A is its complement. We will define a function $N: \Sigma \times (0, 1) \to \Sigma$ to describe neighbourhoods: $N(\sigma, \lambda) = \{\sigma' \in \Sigma : \text{for all } i \in S, |\sigma_i - \sigma'_i| < \lambda\}$, i.e., it is the λ open ball around σ under the sup norm.

Claim 2. There exists λ, $0 < \lambda < \psi$, such that for any $\sigma \in \Sigma^A$ and any $\tilde{\sigma} \in N(\sigma, \lambda) \cap \Sigma^A$, $\tilde{\sigma}$ is dominated by $r(\sigma)$. The claim follows from continuity
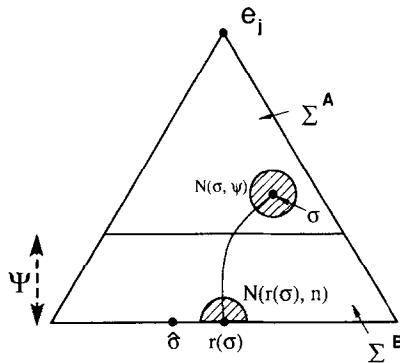

FIGURE 3

of the payoffs and of the function r, and the fact that Σ^A is compact. (Clearly λ depends on ψ, but we do not write this explicitly for notational simplicity.)

Claim 3. Using Claim 2, continuity, and compactness, there exists $\eta > 0$ such that given any $\sigma \in \Sigma^A$, for all $\sigma' \in N(\sigma, \lambda) \cap \Sigma^A$ and for all $\hat{\sigma} \in N(r(\sigma), \eta) \cap \Sigma^B$, $\hat{\sigma}$ dominates σ'.

Claim 4. Using compactness of Σ^A, there is a finite number of strategies $\{\sigma^i\}_{i=1}^{n(\lambda)}$ in Σ^A, such that the union of the neighborhoods $N(\sigma^i, \lambda)$ contains Σ^A.

Claim 5. For each $\omega > 0$, there exists T_ω such that $\mu^t(N(\sigma^i, \lambda))/\mu^t(N(r(\sigma^i), \eta)) < \omega$ for all $t > T_\omega$, and all i.

Proof of Claim 5. The replicator dynamics imply Eq. (3) below.

$$\frac{\mu^{t+1}(N(\sigma^i, \lambda))}{\mu^{t+1}(N(r(\sigma^i), \eta))} = \frac{\mu^t(N(\sigma^i, \lambda))}{\mu^t(N(r(\sigma^i), \eta))}$$

$$\cdot \frac{\tau(\mu^t, N(\sigma^i, \lambda)) \cdot \Pi \cdot \tau(\mu^t, \Sigma)}{\tau(\mu^t, N(r(\sigma^i), \eta)) \cdot \Pi \cdot \tau(\mu^t, \Sigma)}. \qquad (3)$$

We show that the rightmost fraction in (3) is bounded above by some $\varepsilon < 1$.

$$\frac{\tau(\mu^t, N(\sigma^i, \lambda)) \cdot \Pi \cdot \tau(\mu^t, \Sigma)}{\tau(\mu^t, N(r(\sigma^i), \eta)) \cdot \Pi \cdot \tau(\mu^t, \Sigma)} = \frac{\sigma \cdot \Pi \cdot \tau}{\sigma' \cdot \Pi \cdot \tau},$$

where $\sigma \in N(\sigma^i, \lambda)$, and $\sigma' \in N(r(\sigma^i), \eta)$, by the mean value theorem. By Claim 3, the right-hand side equals some $\varepsilon < 1$. Thus, the weight on the neighborhood of σ^i is decreasing at a rate bounded below one, relative to the weight on the neighborhood of $r(\sigma^i)$. This completes the proof of Claim 5,

Hence,

$$\mu^t(\Sigma^A) \leqslant \Sigma_i \, \mu^t(N(\sigma^i, \lambda)) \leqslant \Sigma_i \, \frac{\mu^t(N(\sigma^i, \lambda))}{\mu^t(N(r(\sigma^i), \eta))} \leqslant n(\lambda) \cdot \omega \qquad \text{for} \quad t > T_\omega.$$

In the argument above we can choose ψ as small as we like, and for any ψ, we can choose $n(\lambda) \cdot \omega$ as small as we like. Thus: $\tau_j(\mu^t, \Sigma) \leqslant \mu^t(\Sigma^A) + \psi \cdot \mu^t(\Sigma^B) \leqslant n(\lambda) \cdot \omega + \psi$; which we can make as small as we like.   Q.E.D.

Proposition 2 assumed that μ_0 has full support, and therefore that there is a continuum of players. A similar theorem can be proven for finite

populations which inevitably have initial distributions with finite support.[8] In particular, assume that finitely many players, say m, are drawn at random according to some distribution, $\tilde{\mu} \in \Delta(\Sigma)$, with full support. This draw determines an initial distribution, μ_0, with m elements in its support. Suppose that the population evolves according to replicator dynamics. For any $v > 0$, if m is large enough, then, if the population evolves according to replicator dynamics, with probability $1 - v$ the weight on a never-fittest strategy is eventually less than v. The proof follows the same lines as above: For any ψ, if m is large, then, with high probability, the initial distribution will include strategies belonging to all of the "dominating sets," $N(r(\sigma^1), \eta), ..., N(r(\sigma^{n(\lambda)}), \eta)$, referred to in Claims 2–4.

Thus, after a sufficiently long time T_ω, any dominated strategy is played with arbitrarily small probability. We can apply the same reasoning to assert that after a sufficiently long time, say T, the population weight on *all* dominated strategies will be less than any given $\varepsilon > 0$. Now consider a strategy k that is strictly dominated in the game that would remain if we had gone through some rounds of entirely deleting dominated strategies. Lemma 1 below asserts that there is a mixed strategy, σ', that does strictly better than k against any belief over opponents' strategies (i.e., against any population distribution) that assigns probability less than a small enough ε to those strategies that were deleted. The proof is straightforward: By hypothesis k is dominated in the game where some strategies were deleted, so the statement is true for $\varepsilon = 0$; by continuity of payoffs and finiteness of the game we can find a strictly positive ε that works for all possible strategies k, and all possible deletions.

LEMMA 1. *Given a game with strategy space S and payoffs Π, there exists $\varepsilon > 0$ such that the following is true. Suppose strategy k is strongly dominated in the game with strategy space S' and payoffs $\Pi_{S'}$, where $S' \subset S$, and $\Pi_{S'}$ is the payoff matrix Π restricted to S'. Then, there exists $\sigma' \in \Sigma$ such that $(\sigma' \cdot \Pi \cdot \sigma) > (e_k \cdot \Pi \cdot \sigma)$ for all $\sigma \in \Sigma$ such that $\Sigma_{i \in S - S'} \sigma_i < \varepsilon$.*

Therefore, by an argument similar to that yielding Proposition 2 above, there exists a $\tilde{T}$ after which k and all strategies that are strongly dominated in the smaller game, are played with arbitrarily small probability, and so on. This proves the following.

PROPOSITION 3. *If μ^0 has full support and strategy j is not rationalizable, then for any $\eta > 0$, there is a T such that $\tau_j(\mu', \Sigma) < \eta$ for $t > T$.*

3. HERITABLE RULES

The previous section illustrated that evolutionary outcomes depend on what is heritable: Never-best replies cannot survive if mixed strategies can be inherited, but might survive if only pure strategies can be inherited. Pure-strategy and mixed-strategy inheritance are both very simple rules; it is easy to imagine that organisms can inherit more complicated rules than just "play strategy s," or "play the mixed strategy σ," even though more complicated rules might be "harder" to inherit than simpler ones. Although we do not know how to rank the complexity of rules in terms of heritability, we would like to consider more complex rules than in the previous section.

An interesting class of rules, typical of learning models, is to choose a strategy optimally given beliefs, where beliefs are given by some function of the history of play.[9] Learning models typically take the distribution of rules as exogenously given and focus on the path of strategies that the rules induce.[10] We find the evolutionary perspective provocative in that it gives a framework for endogenizing the path of rules.[11] By determining which rules survive, evolution can shed light on the vexing problem of how beliefs are formed. For example, does evolution select for Bayesian updating starting from some common prior?

A general analysis of heritable rules is beyond the scope of this paper. However, we conclude with an example in which a heritable strategy that never is a best reply survives even if some players inherit the apparently more sophisticated rule "play a best reply to the distribution of strategies last period." This rule of belief formation is boundedly rational in that players do not account for the fact that if everyone obeys that rule, the distribution of strategies in the current period will differ from the distribution last period. These optimizers are a special case of the boundedly rational optimizers in the learning model of Milgrom and Roberts [21], whose result implies that a population of such optimizers will eventually play only rationalizable strategies. Since we have shown that never-best replies can

---

[9] Learning models include Canning [6], Fudenberg and Kreps [10], Fudenberg and Levine [11], Jordan [18], and Kalai and Lehrer [19].

[10] In addition to determining the strategies, the rules may also endogenously determine some aspects of behavior. For example, Fudenberg and Levine [11] adopt the exogenous assumption that players are rational Bayesians who choose strategies optimally and calculate conditional probabilities. In each period an extensive form game is played, so that these rational players will optimally decide how much to experiment in order to observe their opponents' strategies. Thus the level of experimentation is endogenous.

[11] Blume and Easley [4] provide an interesting alternative approach to endogenizing rules: they ask which investment strategies will disappear (own a vanishing share of wealth) and which will dominate (own a positive share of wealth in the limit).

survive when players inherit pure strategies, it is natural to ask whether never-best replies can survive when some players are optimizers. The following extension of Example 1 shows the answer is yes. Thus, the existence of boundedly rational optimizers (with naive beliefs as described above) will not lead to the extinction of even more naive players, and hence rationalizable strategies are not guaranteed.

EXAMPLE 2. We will again use the augmented rock–scissors–paper game described above, and will suppose that there are two heritable rules of behavior: Players can either play the dominated strategy D with probability one or they can be optimizers in the sense we have described. We claim that the replicator dynamics will allow both strategies to persist in equal numbers.

We will demonstrate that there is a stationary dynamic path in which half the population plays D every period, and the other half cycles among strategies A, B, and C. First suppose the path begins with a population in which half the population plays D and the other half, who are optimizers, play A. Strategy D yields payoff zero if the opponent is also playing D, but yields payoff 1.1 if the opponent is an optimizer, since optimizers play A, B, or C. Thus, the payoff to a player who plays D is $(.5)(1.1) = .55$. An optimizer gets .1 if the opponent is playing D, which occurs with probability .5, and gets 1 if it meets another optimizer who is also playing A. Thus, the optimizer's payoff is also .55. In the following period the payoffs will again be .55 for each type of player, but the optimizers will all play B, which is the best reply to the previous population, which was playing A and D, each with probability .5. Since the two strategies—playing D with probability one or being an optimizer—always have the same payoff, they will continue to coexist in equal numbers.

Now suppose the initial population has more than half optimizers and fewer than half D-players. Then the probability that a D-player meets an optimizer is higher than when they are equally represented. These meetings will favor D-players, who get payoff 1.1, and D-players will grow faster than optimizers until the two strategies are represented in equal proportions. A symmetric argument shows that if there are more D-players than optimizers, D-players will grow more slowly. Therefore the only outcome of the replicator dynamics is equal proportions. This completes the presentation of Example 2.

Thus, "optimizers" might be no more successful in propagating their genes than naive players who bequeath genes for playing never-best replies. This example also demonstrates the sensitivity of how one models "a little bit of irrationality": having everyone almost rational is not robust to introducing a small group of very simple-minded players. As we have shown,

introducing a few simple-minded players who copy their parents' never-best reply can destroy the conclusion that "sophisticated" players will learn to play rationalizable strategies (Gul [13], Milgrom and Roberts [21]).

REFERENCES

1. D. K. ARROWSMITH AND C. M. PLACE, "An Introduction to Dynamical Systems," Cambridge Univ. Press, New York, 1990.
2. B. D. BERNHEIM, Rationalizable strategic behavior, *Econometrica* 52 (1984), 1007-1028.
3. K. BINMORE AND L. SAMUELSON, Evolutionary stability in repeated games played by finite automata, mimeo, University of Michigan and University of Wisconsin, 1990.
4. L. BLUME AND D. EASLEY, Wealth dynamics and the market selection hypothesis, mimeo, Cornell University, 1991.
5. A. CABRALES AND J. SOBEL, On the limit points of discrete selection dynamics, mimeo, University of California, San Diego, Dept. of Economics, 1991.
6. D. CANNING, "Convergence to Equilibrium in a Sequence of Games with Learning," LSE Discussion paper TE/89/190, London School of Economics, 1989.
7. E. VAN DAMME, "Stability and Perfection of Nash Equilibria," Springer-Verlag, Berlin, 1987.
8. D. FOSTER AND P. YOUNG, Stochastic evolutionary game eynamics, *Theoretical Population Biology* 38 (1990), 219-232.
9. D. FRIEDMAN, Evolutionary games: An introduction for economists, *Econometrica* 59 (1991), 637-666.
10. D. FUDENBERG AND D. M. KREPS, A theory of learning, experimentation, and equilibrium in games, mimeo, MIT and Stanford, 1988.
11. D. FUDENBERG AND D. K. LEVINE, Steady state learning and self-confirming equilibrium, mimeo, MIT and University of California, Los Angeles, 1990.
12. D. FUDENBERG AND E. MASKIN, Evolution and cooperation in noisy repeated games, *Amer. Econ. Rev.* 80 (1990), 274-279.
13. F. GUL, Rational strategic behavior and the notion of equilibrium, mimeo, Graduate School of Business, Stanford University, 1990.
14. W. G. HINES, Three characterizations of population strategy stability, *J. Appl. Prob.* 17, 333-340.
15. W. G. HINES, Strategy stability in complex populations, *J. Appl. Prob.* 17, 600-610.
16. M. W. HIRSCH AND S. SMALE, "Differential Equations, Dynamical Systems and Linear Algebra," Academic Press, New York, 1974.
17. J. HOFBAUER AND K. SIGMUND, "Dynamical Systems and the Theory of Evolution," Cambridge Univ. Press, Cambridge, 1988.
18. J. S. JORDAN, Bayesian learning in normal form games, *Games Econ. Behavior* 3 (1991), 60-81.
19. E. KALAI AND E. LEHRER, Bayesian learning and nash equilibrium, mimeo, Kellog Graduate School of Management, Northwestern University, 1991.
20. M. KANDORI, G. MAILATH, AND R. ROB, Learning, mutation and long run equilibria in games, mimeo, University of Pennsylvania, 1990.
21. P. MILGROM AND J. ROBERTS, Adaptive and sophisticated learning in repeated normal form games, *Games and Economic Behavior* 3 (1989), 82-100.
22. J. H. NACHBAR, "Evolutionary" selection dynamics in games: Convergence and limit properties, *Int. J. Game Theory* 19 (1990), 59-89.

23. D. PEARCE, Rationalizable strategic behavior and the problem of perfection, *Econometrica* **52** (1984), 1029–1050.

24. A. J. ROBSON, An introduction to evolutionary game theory: Secret handshakes, sucker punches and efficiency, forthcoming *in* "Recent Developments in Game Theory" (J. Borland, J. Creedy and J. Eichberger, Eds.), Edward Elgar Publishing Ltd.

25. L. SAMUELSON, Evolutionary foundations of solution concepts for finite, two-player normal-form games, *in* "Theoretical Aspects of Reasoning About Knowledge" (M. Y. Vardi, Ed.), Kaufmann, Los Altos, 1988.

26. L. SAMUELSON AND J. ZHANG, Evolutionary stability in asymmetric games, mimeo, CentER for Economic Research, Tilburg University, 1991.

27. J. M. SMITH, "Evolution and the Theory of Games," Oxford Univ. Press, Oxford, 1982.

28. E. C. ZEEMAN, Dynamics of the Evolution of Animal Conflict, *J. Theoretical Biology* **89** (1981), 249–270.