Title	Evolution with state-dependent mutations
Authors(s)	Bergin, James, Lipman, B. L.
Publication date	1994-08
Publication information	Bergin, James, and B. L. Lipman. "Evolution with State-Dependent Mutations." University College Dublin. School of Economics, August 1994.
Series	UCD Centre for Economic Research Working Paper Series, WP94/11
Publisher	University College Dublin. School of Economics
Item record/more information	http://hdl.handle.net/10197/1755
Notes	A hard copy is available in UCD Library at GEN 330.08 IR/UNI

Evolution with State - Dependent Mutations

James Bergin
Department of Economics
University College Dublin
(Visiting Professor 1993/'94)

and
B L Lipman
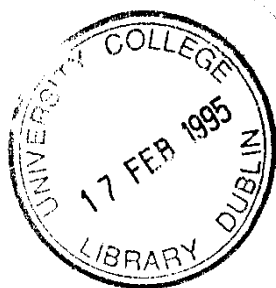Department of Economics
Queen's Univeristy
Ontario, Canada

Evolution with State–Dependent Mutations*

by

J. Bergin and B. L. Lipman

Department of Economics, Queen's University

May 1994

Abstract

Recent evolutionary models have introduced "small mutation rates" as a way of refining predictions of long-run behavior. We show that if mutation rates are allowed to vary across states, then mutations no longer narrow the set of possible predictions. In particular, given any model of the effect of mutations, any invariant distribution of the "mutationless" process is close to an invariant distribution of the process with appropriately chosen small mutation rates.

1 Introduction

A recent reformulation of some simple evolutionary dynamics has led to a surprising result: the addition of small mutation rates leads to precise long–run predictions. Here we re-examine the robustness of this result with respect to the specification of the mutation process.

To motivate the discussion, consider the evolution of strategic behavior in a population of individuals who are repeatedly matched to play some stage game. A variety of evolutionary processes have been analyzed, but to fix ideas, suppose that these individuals only change their actions occasionally, always changing myopically to a best response to the current distribution of strategies in the population. Clearly, if the initial distribution of strategies in the population is sufficiently close to a strict Nash equilibrium,[1] the distribution of strategies will converge to this equilibrium and stay there forever after. Hence any strict Nash equilibrium is the limit point of such a process. Furthermore, there are many other long–run possibilities — for example, the process may cycle forever — so that typically such dynamics have at least as many possibilities for long-run behavior as there are strict Nash equilibria for the game.

Kandori, Mailath, and Rob [1993] (henceforth KMR) and Young [1993][2] have reanalyzed such processes, showing that the addition of small probabilities of mutations change the picture significantly. Suppose that there is a small probability $\varepsilon > 0$ that a given agent changes his action for some unmodeled reason — that is, with probability ε, he deviates in some, perhaps random, fashion from the dynamic described above. This could be thought of as experimentation or mutation in the biological sense. (We comment further on this below.) If every action can be "mutated to," then we have a Markov process which has a strictly positive probability of moving from any one state to any other state. It is well-known that every such process has a unique invariant distribution and that the system converges to this distribution from *any* starting point. Thus the addition of noise in the form of mutations makes the limit of the process unique (as a function of ε)! KMR and Young consider a sequence of ε's converging to zero and analyze the limit point of the associated sequence of invariant distributions, which they call a *long–run equilibrium*. They and many others have provided characterizations of this unique long–run equilibrium for various classes of games.

Since the impact of adding mutations is so dramatic and the cause of mutations is

---

[1] That is, a Nash equilibrium where each player's strategy is the unique best reply to his opponents.

[2] See also Blume [1993], Canning [1992], Ellison [1993], Kandori and Rob [1992], Lagunoff and Matsui [1994], Nöldeke and Samuelson [1993], Robson [1994], Robson and Vega-Redondo [1994], Samuelson [forthcoming], and Young [1993].

unmodeled, one would naturally like to know how robust the analysis is to changes in the mutation process. This paper explores the implications of allowing mutation rates to vary with the state of the system. We show that the KMR/Young results are dishearteningly nonrobust in the following sense. Suppose that x is a limit point (more precisely, an invariant distribution) of the evolutionary process *without* mutations. Given *any* (continuous) model of the way mutations affect evolution, it is always possible to introduce small mutations in such a way that the unique invariant distribution with mutations converges to x as the mutation rates go to zero. Thus the uniqueness created by adding mutations completely vanishes if we relax the state independence assumption of KMR and Young. (We will make this more precise in the following section.)

We make three assumptions on the way mutation affects evolution. First, we assume that mutation rates affect the process continuously. That is, as mutation rates go to zero, the transition probabilities converge "smoothly" to those of the process without mutations. Second, when mutation rates are strictly positive, there is a unique invariant distribution which puts positive probability on every state. This is implied by but weaker than the assumption that the transition matrix with positive mutation rates is strictly positive. Finally, mutation rates are allowed to vary across states.

The first two assumptions are standard.[3] There are several reasons why mutation rates may vary with the state of the system. For example, suppose mutation is intended to represent experimentation. It is natural to allow experimentation rates to vary with the payoffs being earned by and the experience of the players, both of which vary across states. (This point is illustrated more concretely by an example in the next section.) Alternatively, suppose mutations are viewed as mistakes — either computational errors or "trembles". In the case of computational error it is reasonable to suppose that players are more likely to make computational errors in more complex situations so that mistake rates would vary with the state. Similarly, traditional formulations of trembles (Selten [1975] and Myerson [1978]) allow tremble rates to vary across information sets, naturally leading the "aggregate" tremble rate to vary with the state.

There are at least two ways to interpret our results. A pessimistic view is that it is fruitless to use small mutation rates to refine the set of long-run predictions. A more positive assessment is that the results indicate the need for further study of the mutation process and where these mutation rates come from. As we show in the conclusion, simple and

---

[3] Uniqueness of the invariant distribution at positive mutation rates is the key property exploited in this literature. Since this property arises naturally from weak and plausible assumptions on the effect of mutations, it seems quite reasonable to maintain it. Continuity also seems reasonable and is convenient.

plausible restrictions on the mutation rates can, in some examples, restore the uniqueness result of KMR and Young. We emphasize, however, that in general uniqueness can only be maintained by imposing very strong restrictions, such as state independence, on the mutation process.

In Section 2.1, we give some motivating examples in the KMR framework which demonstrate our results. In Section 2.2, we present the general model. Our main result, Theorem 1, is stated and proved in Section 3. In Section 4, we discuss some possible further restrictions on the mutation process.

2 The Model and Motivating Examples

2.1 Examples

The insight of KMR and Young, following on Foster and Young [1990], is that adding small probabilities of mutation to an otherwise standard evolutionary process can yield a unique long–run prediction. To see this more concretely, consider the following game:

$$
\begin{array}{cc}
 & \begin{array}{cc} 1 & \quad 2 \end{array} \\
\begin{array}{c} 1 \\ 2 \end{array} & \begin{pmatrix} 8,8 & 0,4 \\ 4,0 & 6,6 \end{pmatrix}
\end{array}
$$

Suppose there are three agents playing this game. In each period, each player must choose a single action which he uses when playing against each of the other two players. Suppose the action chosen is a best reply given the actions chosen by the other two players in the previous period.

It is not hard to see that the number of agents playing each strategy in a given period completely determines the future evolution of the strategies chosen. Hence we can represent this evolutionary dynamic as a Markov process with a state space given by the number of agents playing, say, strategy 2. Since there are three agents, we let $S = \{0, 1, 2, 3\}$. Let $\pi_{ij}(s)$ denote the payoff in state s to a player currently playing i who switches to strategy j, given other agents do not change their strategies. Thus, $\pi_{11}(0) = 8 + 8 = 16$ since all players expect to be matched with a player choosing 1. Similarly, $\pi_{12}(0) = 4 + 4 = 8$, so, if we are in state 0, no agent would change strategies and we would remain in state 0 forever after (giving a transition probability from state 0 to state 0 of 1). Similar calculations apply to the other states. In state 1, all agents change strategies, moving us to state 2.[4] In state

---

[4] In state 1, $\pi_{11}(1) = 8 + 0 = 8$ and $\pi_{12}(1) = 4 + 6 = 10$. Hence both agents playing action 1 will switch to action 2. Also, $\pi_{21}(1) = 8 + 8 = 16$ and $\pi_{22}(1) = 4 + 4 = 8$. Hence the agent playing action 2 also changes action.

3

2, only the single agent playing action 1 switches, moving the system to state 3.[5] In state
3, every agent is choosing a best reply so the system remains there.

This gives the following transition matrix $\mathbf{P}$ describing how the system moves between
states:

$$\mathbf{P} = \begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} \\ p_{10} & p_{11} & p_{12} & p_{13} \\ p_{20} & p_{21} & p_{22} & p_{23} \\ p_{30} & p_{31} & p_{32} & p_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Here, p_{ij} is the probability of going from state i in one period to state j in the next.

An invariant distribution is a probability distribution on S, say q, which satisfies $q = q\mathbf{P}$.
Such a distribution may be viewed as a "steady state" for the population. (See further
comments on this interpretation below.) For this transition matrix, the set of invariant
distributions is

$$\{q \mid q = \theta e_1 + (1 - \theta)e_4\},$$

where e_i is a vector in $\mathcal{R}^4$ with 1 in the ith position and 0's elsewhere. Note that the invariant
distributions correspond to probability distributions over the two strict Nash equilibria.

Now let us introduce mutations. KMR, Young, and others assume that the probability
of a mutation is some fixed ε, independent of time, the current state, or the agent. If
an agent does not mutate, he changes strategy or not according to the dynamic described
above. If he does mutate, he changes strategies with some fixed probability. For simplicity,
we assume in this example that a mutating agent chooses the opposite strategy from what
the dynamic without mutations would specify for him. Mutations are independent events
across agents and over time. Under these assumptions, the transition matrix with mutations
is:

$$\begin{pmatrix} (1 - \varepsilon)^3 & 3\varepsilon(1 - \varepsilon)^2 & 3\varepsilon^2(1 - \varepsilon) & \varepsilon^3 \\ \varepsilon^2(1 - \varepsilon) & \varepsilon^3 + 2\varepsilon(1 - \varepsilon)^2 & 2\varepsilon^2(1 - \varepsilon) + (1 - \varepsilon)^3 & \varepsilon(1 - \varepsilon)^2 \\ \varepsilon^3 & 3\varepsilon^2(1 - \varepsilon) & 3\varepsilon(1 - \varepsilon)^2 & (1 - \varepsilon)^3 \\ \varepsilon^3 & 3\varepsilon^2(1 - \varepsilon) & 3\varepsilon(1 - \varepsilon)^2 & (1 - \varepsilon)^3 \end{pmatrix}$$

For example, the mutationless process has $p_{12} = 1$, so that if one player is choosing 2, we
move to a state where two players choose this action. With mutations, this transition occurs
if either (a) no mutation takes place (probability $(1 - \varepsilon)^3$) or (b) the player who is supposed
to switch to action 1 mutates and plays action 2 instead and one of the two who is supposed
to choose action 2 mutates to action 1 (probability $2\varepsilon^2(1 - \varepsilon)$).

It is not difficult to verify that the unique invariant distribution is $q^* = (q_0^*, q_1^*, q_2^*, q_3^*)$

---

[5]   In this case, $\pi_{11}(2) = 0 + 0 = 0$ and $\pi_{12}(2) = 6 + 6 = 12$, so the agent playing 1 switches while
$\pi_{21}(2) = 8 + 0 = 8$ and $\pi_{22}(2) = 4 + 6 = 10$ so the other agents do not.

given by

$$q_0^* = \frac{\varepsilon^3 + \varepsilon^4 - 2\varepsilon^5}{Z_1}$$

$$q_1^* = \frac{9\varepsilon^3 - 15\varepsilon^4 + 6\varepsilon^5}{Z_1}$$

$$q_2^* = \frac{9\varepsilon^2 - 36\varepsilon^3 + 69\varepsilon^4 - 66\varepsilon^5 + 24\varepsilon^6}{Z_1}$$

$$q_3^* = \frac{3\varepsilon - 18\varepsilon^2 + 52\varepsilon^3 - 83\varepsilon^4 + 70\varepsilon^5 - 24\varepsilon^6}{Z_1}$$

where Z_1 is a normalization factor which makes the probabilities sum to 1. Specifically,

$$Z_1 = 3\varepsilon - 9\varepsilon^2 + 26\varepsilon^3 - 28\varepsilon^4 + 8\varepsilon^5.$$

Therefore,

$$\lim_{\varepsilon \to 0} q^* = (0, 0, 0, 1).$$

Suppose we relax these assumptions by letting the mutation rate vary with the state. If mutation is intended to model experimentation by the players, then it is difficult to see why the mutation rate in state 0 would be as large as the mutation rate in other states. In state 0, all agents play strategy 1 and so all always earn the highest possible payoff in the game. So why would players experiment in this state? By contrast, in state 3, players may experiment in hopes of reaching the other (Pareto preferred) Nash equilibrium. While we do not wish to claim that this is a necessary property of mutation rates, a reasonable model of mutations would surely allow this possibility, and hence it is important to explore its implications.

In line with this intuition, suppose that the mutation rate in states 1, 2, and 3 is ε, while the mutation rate in state 0 is ξ. The interesting case is where $\xi < \varepsilon$, so that the mutation rate at state 0 is much smaller than the mutation rate at other states, although we do not impose this yet.

Except for this change, we maintain all the assumptions from above. Specifically, we suppose that the probability that any given player mutates in a given period is dependent on the state of the system in that period, but is independent across agents and over time. If an agent does not mutate, his action choice is determined in the way the original evolutionary process specified; otherwise, he chooses the opposite of the action called for by the original process.

In this case, the transition matrix for the process with mutations is

$$\begin{pmatrix} (1-\xi)^3 & 3\xi(1-\xi)^2 & 3\xi^2(1-\xi) & \xi^3 \\ \varepsilon^2(1-\varepsilon) & \varepsilon^3 + 2\varepsilon(1-\varepsilon)^2 & 2\varepsilon^2(1-\varepsilon)+(1-\varepsilon)^3 & \varepsilon(1-\varepsilon)^2 \\ \varepsilon^3 & 3\varepsilon^2(1-\varepsilon) & 3\varepsilon(1-\varepsilon)^2 & (1-\varepsilon)^3 \\ \varepsilon^3 & 3\varepsilon^2(1-\varepsilon) & 3\varepsilon(1-\varepsilon)^2 & (1-\varepsilon)^3 \end{pmatrix}$$

5

The unique invariant distribution is $q' = (q_0', q_1', q_2', q_3')$, where

$$q_0' = \frac{\varepsilon^3[1 + f_1(\varepsilon, \xi)]}{Z_2}$$

$$q_1' = \frac{9\xi\varepsilon^2[1 + f_2(\varepsilon, \xi)]}{Z_2}$$

$$q_2' = \frac{9\xi\varepsilon[1 + f_3(\varepsilon, \xi)]}{Z_2}$$

$$q_3' = \frac{3\xi[1 + f_4(\varepsilon, \xi)]}{Z_2}$$

where

$$Z_2 = \varepsilon^3[1 + f_5(\varepsilon, \xi)] + 3\xi[1 + f_6(\varepsilon, \xi)]$$

and

$$\lim_{\xi \to 0, \varepsilon \to 0} f_i(\varepsilon, \xi) = 0, \quad \forall i.$$

In this case, the limiting distribution depends on whether ξ or ε^3 goes to zero faster. If $\xi = \varepsilon$, for example, we are back to the previous case and we know that $q_3' \to 1$ as ε and ξ go to zero. If $\xi = \alpha\varepsilon^3$, it is not hard to show that $q_0' \to 1/(1 + 3\alpha)$ and $q_3' \to 3\alpha/(1 + 3\alpha)$ as ε and ξ go to zero. Thus any q_0' and q_3' both strictly positive and summing to 1 can be generated by an appropriate choice of $\alpha \in (0, \infty)$. Finally, if $\xi = \varepsilon^4$, then $q_0' \to 1$. In short, by choosing the ε and ξ sequences appropriately, we can generate *any* of the original invariant distributions. Thus, if the mutation rate at this state goes to zero quickly enough relative to the other mutation rates, then it will have positive probability in the limit, in contrast to the KMR/Young results.

2.2 The Model

The details of the evolutionary process without mutations are irrelevant for our purposes. We assume that it is a finite state Markov process.

For intuition, imagine a large but finite population of agents who are matched over time, perhaps randomly, to play some fixed game. These agents have certain observations regarding the play at each date. Suppose that memory is finite, so that no more than a certain number of past observations can be remembered. In this case, as in the example above, the entire future of the system can be summarized by specifying the strategy of each player in each of these finitely many past periods. Call each such specification a *state*. If the set of strategies is finite, the set of states of the system will be finite as well and the stochastic process governing evolution will be a finite state Markov process.

Formally, the evolutionary process without mutations is a Markov process on a finite state space, $S = \{1, \ldots, s\}$. This process is completely described by a transition matrix $\mathbf{P}$,

where p_{ij} is the probability of a transition from state i to state j, $i, j \in S$. We will refer to $\mathbf{P}$ as the *mutationless process*.

Let Δ denote the set of probability distributions on S. A long-run prediction about the mutationless process is an element of Δ with the property that it is "unaffected by the passage of time." More formally, we focus on the invariant distributions of the mutationless process.

Definition 1 *An* **invariant distribution** *of* $\mathbf{P}$ *is a vector* $q \in \Delta$ *such that* $q = q\mathbf{P}$.

Note that q is a *row* vector. Denote by $\mathcal{I}(\mathbf{P})$ the set of invariant distributions of $\mathbf{P}$.

Intuitively, if q is our belief about the state of the system and we learn that a period of time has passed (without learning anything directly about the state), then our updated beliefs $q\mathbf{P}$ will be unchanged if q is invariant.

Remark 1 Duffie *et al.* [1989], among others, have pointed out that invariant distributions may not be an appropriate way of describing long-run behavior. For example, suppose that $\mathbf{P}$ is an identity matrix. Then any distribution on S is invariant. However, whatever state the system starts in, it will remain in that state forever. Hence a nondegenerate invariant distribution describes an outside observer's uncertainty about the initial state of the system, but is not a "prediction" in the sense that it *cannot* equal the frequency distribution of outcomes over time. As we will see, however, our results give a new interpretation of such invariant distributions. See Remark 3.

The set of invariant distributions is convex. Therefore, if there is more than one invariant distribution, there are infinitely many. As noted above, the mutationless evolutionary dynamics generally considered in the literature have the property that every strict Nash equilibrium is associated with an invariant distribution. Hence for any game with more than one strict Nash equilibrium, these dynamics have infinitely many invariant distributions. The following definition allows us to characterize these invariant distributions.

Definition 2 *A subset C of S is* **absorbing in** $\mathbf{P}$ *if for all* $i \in C$, $j \notin C$, $p_{ij} = 0$. C *is a* **minimal absorbing set of** $\mathbf{P}$ *if it is absorbing in* $\mathbf{P}$ *and no subset of it is.*

Definition 3 *A state $i \in S$ is* **transient in** $\mathbf{P}$ *if it is not contained in any minimal absorbing set of* $\mathbf{P}$.

Equivalently, a state $i \in S$ is transient if there is a j such that $p_{ij} > 0$ but j is either transient or in an absorbing set which does not include i.

A well-known fact about Markov processes[6] is

---

[6] See, *e.g.*, Iosifescu [1980] for proof.

Fact 1 *If C is a minimal absorbing set of $\mathbf{P}$, then for every $i, j \in C$, there is a μ_{ij} such that for every invariant distribution $q \in \mathcal{I}(\mathbf{P})$, either $q_i = q_j = 0$ or $q_i/q_j = \mu_{ij}$. If i is a transient state of $\mathbf{P}$, then for every invariant distribution $q \in \mathcal{I}(\mathbf{P})$, $q_i = 0$.*

In other words, given any invariant distribution, the restriction of the distribution to a minimal absorbing set, if well-defined, is unique. Furthermore, all probability is concentrated on the nontransient states — that is, on the collection of minimal absorbing sets. Hence all differences between invariant distributions correspond to differences in the relative probabilities of different minimal absorbing sets.

An immediate implication is that if the only minimal absorbing set is S itself, then there is a unique invariant distribution. In this case, $\mathbf{P}$ is said to be *irreducible*.

While the mutationless processes considered in the literature are typically not irreducible, the processes generated by adding mutations are. In KMR, a state of the system is the number of agents using a particular strategy. When they add mutations, they assume that the probability of any given agent "mutating to" any given strategy is strictly positive, giving a strictly positive probability of transiting from any one state to any other state in a single period. Hence the resulting transition matrix is irreducible. In Young, a state also includes a specification of strategies played in some finite number of past periods. Since a mutation only changes the current period strategies in his model, mutations do not create a positive probability of moving between states with radically different histories in a single period. However, the process with mutations is irreducible since there is a positive probability of *eventually* moving between any pair of states.

We call a procedure for associating transition probabilities to mutation rates a "model of mutations." The properties we require the model of mutations to satisfy are the following. First, the process with mutations must be irreducible. Second, the process with mutations converges smoothly to the mutationless process as mutation rates go to zero. Third, the mutation rates may be state dependent. More specifically, only the probability of a mutation in state i is relevant for determining movements out of state i via mutation.

To understand the motivation for the third restriction, consider the process with mutation and suppose the current state is i. How are the probabilities of transition to other states determined? By assumption, the state summarizes all relevant aspects of the system — the information the players have, the payoffs they receive, etc. For this reason, if a player does not mutate, his behavior is completely determined by the state. Similarly, if a player does mutate, his behavior conditional on this event is completely determined by the state. Finally, it is the state i mutation rate which determines the relative probabilities of these two events. Hence no mutation rate other than the state i mutation rate is relevant

8

for transitions from state i. Technically, this is not necessary for the result — transition probabilities out of one state can depend on the mutation rates for other states. However, the interpretation of mutation rates is less clear in this case. See Remark 4.

Let $\mathcal{M}$ denote the set of Markov matrices on S. We will call a vector $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_s)$ with $\varepsilon_i \in [0, 1]$ for all i a vector of *mutation rates*.

Definition 4 *A* **model of mutations** *for* $\mathbf{P}$ *is a continuous function* $\mathbf{M} : [0, 1]^S \to \mathcal{M}$ *such that (a)* $\mathbf{M}(0) = \mathbf{P}$, *(b)* $\mathbf{M}(\varepsilon)$ *is irreducible for all* $\varepsilon \gg 0$, *and (c) the elements of the* i^{th} *row of* $\mathbf{M}(\varepsilon)$ *depend only on* ε_i.

These three assumptions allow for a very wide variety of interpretations of the nature of mutations. In particular, the KMR/Young assumptions are special cases. On the other hand, we do not need this much structure for our result. These assumptions are more intuitive that the weaker statements we require. See Remark 4 for more detail.

Rather than adopt a specific model of how the addition of mutations affects evolution, we prove our theorem for *any* model of mutations. We emphasize that the model of mutations is held fixed — only mutation rates vary. The main result of the paper is easier to prove if the model of mutations is treated as a variable as well — that is, if to approach a specific invariant distribution, we vary both the vector ε *and* the function $\mathbf{M}(\cdot)$. However, in many applications, the model of mutations arises from the specification of the model, so the mutation model is exogenous and independent of the mutation rates.

We wish to characterize the long-run behavior of the system for "small" mutation rates given some fixed model of mutations. This motivates the following definition.

Definition 5 *A probability distribution* $q \in \Delta$ *is* **achievable** *with mutation model* $\mathbf{M}$ *if there exists a sequence of strictly positive mutation rate vectors* $\varepsilon^n \to 0$ *such that* $q^n \to q$ *where* $\{q^n\} = \mathcal{I}(\mathbf{M}(\varepsilon^n))$. *Let* $\mathcal{A}(\mathbf{M})$ *denote the set of achievable distributions with mutation model* $\mathbf{M}$.

KMR and Young require ε_i to be independent of i. Hence they, in effect, consider only a single sequence of ε vectors as they require all mutation rates to go to zero at the same rate.[7] Instead, we allow different mutation probabilities to go to zero at different rates.

Remark 2 The reader may find an analogy useful at this point. When defining trembling-hand perfect equilibria, one first calculates ε-perfect equilibria where agents may "tremble" and make mistakes. One then considers the limit as these mistake probabilities go to zero. However, the probabilities of different mistakes may go to zero at different rates. Similarly,

---

[7] It is conceivable that different sequences of ε's going to zero could yield different limits even when all the ε_i's are required to be equal. However, this multiplicity does not arise in the KMR and Young models.

we are allowing the probabilities of mutations in different states to go to zero at different rates.

3 The Main Result

Our main result is that with state–dependent mutation rates, any invariant distribution of the mutationless process is achievable with any mutation model.

Theorem 1   *If* **M** *is any mutation model for* **P**, *then* $\mathcal{A}(\mathbf{M}) = \mathcal{I}(\mathbf{P})$.

Proof. Let $N \subset 2^S$ denote the collection of minimal absorbing sets of **P** and let T be the set of transient states of **P**. Let $S^* = N \cup T$, let n be the cardinality of N, and let t be the cardinality of T. With some abuse of terminology, we will sometimes refer to elements of S^* as states, even though some elements are actually sets of states. Letting $\Delta(S^*)$ denote the set of probability distributions on S^*, we can define a function from Δ (the set of probability distributions on S) to this set as follows:

$$\beta_i(q) = \begin{cases} \sum_{j \in i} q_j, & \text{if } i \in N; \\ q_i, & \text{if } i \in T. \end{cases}$$

By Fact 1, if we restrict β to $\mathcal{I}(\mathbf{P})$, it is one–to–one.

Throughout, we only consider mutation rates with the property that for any two states i and j in the same minimal absorbing set of **P**, $\varepsilon_i = \varepsilon_j$. Given this, we abuse notation slightly and refer to vectors of mutation rates as $(n+t)$–tuples rather than s–tuples — that is, we will write $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n, \varepsilon_{n+1}, \ldots, \varepsilon_{n+t})$ where for $i \in N$, ε_i is the common mutation rate for the states in set i. For convenience, given any $V \subseteq N$, we write $\varepsilon = (\varepsilon_V, \varepsilon_{\sim V})$. Also, let $\Delta(V)$ denote the set of probability distributions on V — that is,

$$\Delta(V) = \{x \in \mathcal{R}^V \mid x_i \geq 0, \forall i \in V \text{ and } \sum_{i \in V} x_i = 1\}.$$

Also, let $\Delta^\circ(V)$ denote the interior of $\Delta(V)$ and $\partial\Delta(V)$ its boundary. That is,

$$\Delta^\circ(V) = \{x \in \Delta(V) \mid x \gg 0\}$$

and

$$\partial\Delta(V) = \{x \in \Delta(V) \mid x_i = 0 \text{ for some } i \in V\}.$$

Fix any mutation model **M**. We construct a family of correspondences mapping the simplex into itself as follows. First, given any $\varepsilon \gg 0$, let $I(\varepsilon) \in \Delta(S)$ denote the unique invariant distribution for the process $\mathbf{M}(\varepsilon)$.

Lemma 1   *If* $\varepsilon^k \to \varepsilon^*$, $\varepsilon^k \gg 0$ *for all* k, *then*

$$\lim_{k \to \infty} I(\varepsilon^k) \in \mathcal{I}(\mathbf{M}(\varepsilon^*)).$$

10

Hence I is continuous on $(0,1]^S$.

Proof of Lemma. Because $\mathbf{M}$ is continuous, $\mathbf{M}(\varepsilon^k) \to \mathbf{M}(\varepsilon^*)$ (entry by entry). By definition,

$$I(\varepsilon^k)\mathbf{M}(\varepsilon^k) = I(\varepsilon^k)$$

for each k. Hence for any $\delta > 0$, there is a K such that for all $k \geq K$, the ith coordinate of $I(\varepsilon^k)\mathbf{M}(\varepsilon^*)$ is within δ of $I_i(\varepsilon^k)$ for all i. Hence, letting q denote the limit of $I(\varepsilon^k)$ as $k \to \infty$ (taking a subsequence if necessary), we must have

$$q\mathbf{M}(\varepsilon^*) = q,$$

so $q \in \mathcal{I}(\mathbf{M}(\varepsilon^*))$. Clearly, if $\varepsilon^* \gg 0$, $\mathcal{I}(\mathbf{M}(\varepsilon^*))$ is a singleton so $q = I(\varepsilon^*)$. Hence $I(\varepsilon)$ is continuous on $(0,1]^S$. ∎

Given any nonempty $V \subseteq N$ with at least two elements, let $\#V$ denote the cardinality of V and define $\alpha^V : \Delta(S^*) \to \Delta(V)$ by

$$\alpha_i^V(q) = q_i + \frac{1}{\#V} \sum_{j \notin V} q_j.$$

In other words, the probability that q puts on states outside of V is "spread" over the states in V to construct a probability distribution on V only.

Next, for any nonempty $V \subseteq N$ with $\#V \geq 2$, $\eta > 0$ and $\varepsilon_{\sim V} \gg 0$, define $g_{\eta,\varepsilon_{\sim V}}^V :$ $\Delta^\circ(V) \to \Delta^\circ(V)$ by

$$g_{\eta,\varepsilon_{\sim V}}^V(x) = \alpha^V\left(\beta\left(I(\eta x, \varepsilon_{\sim V})\right)\right).$$

For notational simplicity, we often omit the $\eta, \varepsilon_{\sim V}$ subscripts.

To understand the construction, imagine that we fix the mutation rates for all states not in V at $\varepsilon_{\sim V}$ and fix the *total* of the mutation rates for states in V at η. Given a vector in $\Delta^\circ(V)$, we can translate it to a vector of mutation rates by "rescaling" by η so that these numbers correspond to the mutation rates for the states in V and then adding to the vector the mutation rates for the other states. We then use this to calculate the invariant distribution associated with these mutation rates. Finally, we use β to convert this to a distribution on S^* and then α^V to move the probability on states outside of V onto the states in V.

Obviously, α^V and β are continuous. Hence Lemma 1 implies that $g_{\eta,\varepsilon_{\sim V}}^V$ is continuous on $\Delta^\circ(V)$ for all $V \subseteq N$ with $\#V \geq 2$, all $\eta > 0$, and all $\varepsilon_{\sim V} \gg 0$.

To construct the correspondence $G_{\eta,\varepsilon_{\sim V}}^V : \Delta(V) \to \Delta(V)$, for any $x \in \Delta(V)$, let

$$G_{\eta,\varepsilon_{\sim V}}^V(x) = \{q \in \Delta(V) \mid \exists x^k \to x \text{ with } x^k \in \Delta^\circ(V) \text{ and } g_{\eta,\varepsilon_{\sim V}}^V(x^k) \to q\}.$$

Again, we will often omit the $\eta, \varepsilon_{\sim V}$ subscripts. The continuity of g^V implies that

$$G^V(x) = \{g^V(x)\}$$

for all $x \gg 0$. Hence G^V maps the interior of the simplex to itself and is identical to g^V on this subspace. By construction, G^V is upper semicontinuous. By Lemma 1 and the continuity of α^V and β,

$$G^V_{\eta, \varepsilon_{\sim V}}(x) \subseteq \alpha^V \left(\beta \left(\mathcal{I} \left(\mathbf{M}(\eta x, \varepsilon_{\sim V}) \right) \right) \right).$$

Lemma 2 *For any $x \in \partial\Delta(V)$, $G^V(x) \subset \partial\Delta(V)$. Furthermore, if $x_k > 0$, then for all $x' \in G^V(x)$, $x'_k = 0$ so $x \notin G^V(x)$.*

Proof of Lemma. Fix any nonempty $V \subseteq N$ with $\#V \geq 2$, $x \in \partial\Delta(V)$, and η and $\varepsilon_{\sim V}$ strictly positive. Consider the process $\mathbf{M}(\eta x, \varepsilon_{\sim V})$. Because $x \in \partial\Delta(V)$, there is an $i \in V$ with $x_i = 0$. Hence the set of states i is an absorbing set of $\mathbf{M}(\eta x, \varepsilon_{\sim V})$. Consider any k such that either $k \notin V$ (and $\varepsilon_k > 0$) or $x_k > 0$. This state (or collection of states) is associated with a strictly positive mutation rate. If all mutation rates were strictly positive, the probability of transiting from set k to set i in a finite number of periods would be strictly positive. Consider the sequence of states in this path. If every state in this sequence has a strictly positive mutation rate, then $\mathbf{M}(\eta x, \varepsilon_{\sim V})$ still gives this sequence of states strictly positive probability. In this event, state k must be transient in $\mathbf{M}(\eta x, \varepsilon_{\sim V})$. If some state in this sequence has a zero mutation rate, then it, too, is absorbing, again establishing that state k is transient in $\mathbf{M}(\eta x, \varepsilon_{\sim V})$. Either way, for any $q \in \mathcal{I}(\mathbf{M}(\eta x, \varepsilon_{\sim V}))$, we must have $q_k = 0$ by Fact 1. Hence for such q, $\alpha^V(\beta(q)) = \beta(q)$. So

$$G^V(x) \subseteq \beta(\mathcal{I}(\mathbf{M}(\eta x, \varepsilon_{\sim V}))).$$

If $x_k > 0$, then, for all $x' \in G^V(x)$, we must have $x'_k = 0$. Hence $x \notin G^V(x)$. Also, since there must be some $k \in V$ with $x_k > 0$, $G^V(x) \subseteq \partial\Delta(V)$. $\blacksquare$

Lemma 3 *For any nonempty $V \subseteq N$ with $\#V \geq 2$, if G^V is convex-valued, then it is surjective.*

Proof of Lemma. Suppose G^V is convex-valued but not surjective, so $G^V(\Delta(V)) \neq \Delta(V)$. Clearly, $G^V(\Delta(V))$ is simply the closure of $g^V(\Delta^\circ(V))$, so it must be closed. Hence $\Delta(V) \setminus G^V(\Delta(V))$ is open. Choose any element, say b, from the interior of this set.

Define a function $\rho_b : G^V(\Delta(V)) \to \partial\Delta(V)$ as follows. Given any $x \in G^V(\Delta(V))$, draw a straight line from b to x and let $\rho_b(x)$ be the point where this line intersects the boundary. Clearly, ρ_b is a continuous function on $G^V(\Delta(V))$. Note also that for any $x \in \partial\Delta(V)$, $\rho_b(x) = x$.

12

Define a correspondence $G_b^V : \Delta(V) \to \Delta(V)$ by

$$G_b^V(x) = \{\bar{q} \mid \bar{q} = \rho_b(q), \text{ for some } q \in G^V(x)\} = \rho_b(G^V(x)).$$

Since G^V is upper semicontinuous and ρ_b is continuous, G_b^V is upper semicontinuous as well. Furthermore, since G^V is convex-valued by hypothesis, G_b^V is convex-valued. By the Kakutani fixed point theorem, then, there is some x^* such that $x^* \in G_b^V(x^*)$. Since ρ_b maps to the boundary of the simplex, we must have $G_b^V(x^*) \subseteq \partial\Delta(V)$, so $x^* \in \partial\Delta(V)$. By Lemma 2, then, $G^V(x^*) \subset \partial\Delta(V)$. Because ρ_b maps any boundary point to itself, $\rho_b(G^V(x^*)) = G^V(x^*)$. Hence $x^* \in \partial\Delta(V)$ and $x^* \in G^V(x^*)$. But by Lemma 2, this is impossible. Hence G^V must be surjective. ∎

Lemma 4 *For any nonempty $V \subseteq N$ with $\#V \geq 2$, G^V is convex-valued.*

Proof of Lemma. The proof is by induction on $\#V$. First, consider any $V \subseteq N$ with $\#V = 2$. Clearly, Lemma 2 implies $G^V(1,0) = \{(0,1)\}$ and $G^V(0,1) = \{(1,0)\}$, while $G^V(x_1, x_2) = \{g^V(x_1, x_2)\}$ for any $(x_1, x_2) \gg 0$. Obviously, G^V is convex-valued.

So suppose the lemma is established for all $V \subseteq N$ with $2 \leq \#V \leq L - 1$ for $L \leq \#N$. We now show that it holds for all V with $\#V = L$, completing the proof. So fix any $V \subseteq N$ with $\#V = L$. Clearly, for any $x \in \Delta^\circ(V)$, $G^V(x)$ is a singleton and so is convex. Consider any $x \in \partial\Delta(V)$. Let

$$\bar{V} = \{i \in V \mid x_i = 0\}.$$

Clearly, $\bar{V}$ is nonempty and $\#\bar{V} \leq L - 1$. If $\#\bar{V} = 1$, then Lemma 2 implies that $G^V(x)$ is a singleton since it must consist only of the degenerate distribution with probability 1 on the single element of $\bar{V}$. In this case, again, $G^V(x)$ is convex. So we may as well assume that $\#\bar{V} \geq 2$.

Let

$$F^*(x) = \{q \in \Delta(V) \mid q_i = 0 \ \forall i \notin \bar{V}\}.$$

By Lemma 2, $G^V(x) \subseteq F^*(x)$. We now show that, in fact, for all $\eta > 0$ and $\varepsilon_{\sim V} \gg 0$, we have $G_{\eta, \varepsilon_{\sim V}}^V(x) = F^*(x)$.

To see this, fix any $\bar{\eta} > 0$, any $\bar{\varepsilon}_{\sim V} \gg 0$, and any $q \in F^*(x)$. By definition, the support of q must be contained in $\bar{V}$. Hence we can view q as an element of $\Delta(\bar{V})$. By the induction hypothesis, $G^{\bar{V}}$ is convex-valued and so, by Lemma 3, is surjective. Hence for every $\eta > 0$ and every $\varepsilon_{\sim\bar{V}} \gg 0$, there exists x^* such that $q \in G_{\eta, \varepsilon_{\sim\bar{V}}}^{\bar{V}}(x^*)$. Define $\varepsilon_{\sim\bar{V}}^*$ by

$$\varepsilon_j^* = \begin{cases} \bar{\varepsilon}_j, & \text{if } j \notin V; \\ \bar{\eta} x_j, & \text{if } j \in V \setminus \bar{V}. \end{cases}$$

13

Since $\varepsilon_{\sim V}$ will be fixed at $\varepsilon^*_{\sim V}$ through the rest of the argument, we will omit it from the notation when convenient.

Consider a sequence $\eta^k \to 0$ and the associated sequence $x^*(\eta^k)$ chosen so that $q \in G^V_{\eta^k}(x^*(\eta^k))$. By definition, then, for each k, there is a sequence, say $(x^k)^m \to x^*(\eta^k)$, such that

$$g^{\bar{V}}_{\eta^k}((x^k)^m) \to q.$$

Fix any sequence $\delta^k \to 0$ with $\delta^k > 0$ for all k. For each k, choose an m large enough that $g^V_{\eta^k}((x^k)^m)$ is within δ^k of q (pointwise). Let $\hat{x}^k = (x^k)^m$. Clearly, $g^V_{\eta^k}(\hat{x}^k) \to q$. By definition, the ith coordinate of $g^V_{\eta^k}(\hat{x}^k)$ is

$$\beta_i(I(\eta^k \hat{x}^k, \varepsilon^*_{\sim V})) + \frac{1}{\#\bar{V}} \sum_{\ell \notin \bar{V}} \beta_\ell(I(\eta^k \hat{x}^k, \varepsilon^*_{\sim V})).$$

Since $\eta^k \hat{x}^k \to 0$ as $k \to \infty$,

$$\sum_{\ell \notin \bar{V}} \beta_\ell(I(\eta^k \hat{x}^k, \varepsilon^*_{\sim V})) \to 0.$$

So

(1) $$\beta_i(I(\eta^k \hat{x}^k, \varepsilon^*_{\sim V})) \to q_i, \quad \forall i \in \bar{V}.$$

Define a sequence x^k in $\Delta(V)$ by

$$x^k_i = \begin{cases} \eta^k \hat{x}^k_i, & \text{if } i \in \bar{V}; \\ x_i, & \text{if } i \in V \setminus \bar{V}. \end{cases}$$

Clearly, $x^k \to x$ since $x_i = 0$ for all $i \in \bar{V}$. Also, (1) implies that $g^V_{\hat{\eta}, \ell_{\sim V}}(x^k) \to q$ (where we now view q as an element of $\Delta(V)$). Hence $q \in G^V_{\hat{\eta}, \ell_{\sim V}}(x)$.

Summarizing, we see that $F^*(x) \subseteq G^V(x)$. Hence $F^*(x) = G^V(x)$. Since $F^*(x)$ is obviously convex, this establishes that $G^V(x)$ is convex, so that G^V is convex-valued. ∎

To complete the proof of the Theorem, it suffices to show that for all $q^* \in \mathcal{I}(\mathbf{P})$, there is a sequence $\varepsilon^k \to 0$ with $\varepsilon^k \gg 0$ for all k such that $\beta(I(\varepsilon^k)) \to \beta(q^*)$. To see this, recall from Lemma 1 that

$$\lim_{k \to \infty} I(\varepsilon^k) \in \mathcal{I}(\mathbf{M}(0)) = \mathcal{I}(\mathbf{P}).$$

Also, β is one-to-one on $\mathcal{I}(\mathbf{P})$ and continuous. Hence $\beta(I(\varepsilon^k)) \to \beta(q^*)$ implies $I(\varepsilon^k) \to q^*$.

Let

$$Q = \{q \in \Delta(S^*) \mid q_i = 0, \ \forall i \in T\}.$$

Since $\beta(\mathcal{I}(\mathbf{P})) = Q$, we wish to show that for any $q \in Q$, there is a sequence $\varepsilon^k \to 0$ with $\varepsilon^k \gg 0$ such that $\beta(I(\varepsilon^k)) \to q$. Clearly, though, any $q \in Q$ can be equivalently viewed as an element of $\Delta(N)$. So fix any $q \in \Delta(N)$.

14

For ease of notation, let $G = G^N$ and $g = g^N$. By Lemma 4, G is surjective. Hence for any $\eta > 0$ and $\varepsilon_{\sim N} \gg 0$, there is a $\bar{x}$ with $q \in G_{\eta, \varepsilon_{\sim N}}(\bar{x})$. Fix a sequence $\eta^k \to 0$ and a sequence $\varepsilon^k_{\sim N} \to 0$ with $\eta^k > 0$ and $\varepsilon^k_{\sim N} \gg 0$ for all k. Let $\bar{x}^k$ be chosen so that $q \in G_{\eta^k, \varepsilon^k_{\sim N}}(\bar{x}^k)$ for all k. That is, for each k, there exists $(\bar{x}^k)^m \to \bar{x}^k$ as $m \to \infty$ such that $g_{\eta^k, \varepsilon^k_{\sim N}}((\bar{x}^k)^m) \to q$. Fix any sequence $\delta^k \to 0$ with $\delta^k > 0$ for all k. For each k, choose m such that $g_{\eta^k, \varepsilon^k_{\sim N}}((\bar{x}^k)^m)$ is within δ^k of q (pointwise). Let $\hat{x}^k = (\bar{x}^k)^m$. Clearly, $g_{\eta^k, \varepsilon^k_{\sim N}}(\hat{x}^k) \to q$.

By Lemma 1,

$$\lim_{k \to \infty} I(\eta^k \hat{x}^k, \varepsilon^k_{\sim N}) \in \mathcal{I}(\mathbf{M}(0)) = \mathcal{I}(\mathbf{P}).$$

Hence

$$\lim_{k \to \infty} I_j(\eta^k \hat{x}^k, \varepsilon^k_{\sim N}) = 0, \quad \forall j \in T.$$

Therefore,

$$
\begin{aligned}
q &= \lim_{k \to \infty} g_{\eta^k, \varepsilon^k_{\sim N}}(\hat{x}^k) \\
&= \lim_{k \to \infty} \alpha^N(\beta(I(\eta^k \hat{x}^k, \varepsilon^k_{\sim N}))) \\
&= \lim_{k \to \infty} \beta(I(\eta^k \hat{x}^k, \varepsilon^k_{\sim N})).
\end{aligned}
$$

That is, letting $\varepsilon^k_i = \eta^k x^k_i$ for each $i \in N$, we have $\beta(\varepsilon^k) \to q$, completing the proof. ∎

Remark 3 Recall from Remark 1 that some processes have invariant distributions which can never be equal to the long–run frequency distribution over states. For example, if $\mathbf{P}$ is an identity matrix, then any probability distribution is invariant. However, the only possible long–run frequency distributions are degenerate since whatever state the system begins in, it will remain in that state forever after. Theorem 1 gives a new way of viewing invariant distributions which do not equal long–run frequencies: they are approximately equal to long–run frequencies for nearby processes. In this sense, even these "odd" invariant distributions can be viewed as reasonable ways to predict the long–run behavior of the system if we believe that we may make small mistakes in our specification of the system.[8]

Remark 4 Our assumptions on mutation models can be substantially relaxed without requiring significant changes in the proof. First, the assumption that $\mathbf{M}(\varepsilon)$ is irreducible for all $\varepsilon \gg 0$ can be weakened to requiring only that for every $\varepsilon \gg 0$, $\mathbf{M}(\varepsilon)$ has a unique minimal absorbing set containing all nontransient states. Second and more importantly, our state dependence assumption that the ith row of $\mathbf{M}(\varepsilon)$ only depends on ε_i can be relaxed to the following less intuitive requirement: If $A \subseteq S$ is absorbing in $\mathbf{P}$, then

(i) $\varepsilon_i = 0$, $\forall i \in A \implies A$ is absorbing in $\mathbf{M}(\varepsilon)$, and

---

15

(ii) $\quad \varepsilon_i > 0, \quad \forall i \in A$ and $\exists B \subset S \setminus A$ absorbing in $M(\varepsilon)$

$$\Longrightarrow \quad i \text{ is transient in } \mathbf{M}(\varepsilon), \; \forall i \in A.$$

Remark 5   Theorem 1 implies that any invariant distribution of $\mathbf{P}$ is near the unique invariant distribution for the process with appropriately chosen small mutation rates. Furthermore, it is not hard to use the continuity of $\mathbf{M}$ to strengthen this as follows: For any invariant distribution $q \in \mathcal{I}(\mathbf{P})$ and any $\beta > 0$, there is an $\varepsilon^* \gg 0$ and $\delta > 0$ such that for all ε within δ of ε^*, the unique invariant distribution of $\mathbf{M}(\varepsilon)$ is within β of q. In this sense, our use of small mutation rates to approximate q is robust in that nearby small mutation rates always lead to nearby long-run predictions.

4   Conclusion

As we have shown, if mutation rates can vary in any fashion with the state, then small mutation probabilities do not refine the set of long-run predictions at all. More specifically, any long-run prediction about the mutationless process (any invariant distribution) is close to the unique invariant distribution of a nearby system with small mutation probabilities. One could interpret this result as saying that it is futile to try to refine the set of long-run predictions by adding small mutation rates to the model.

An alternative interpretation is that the source of mutation rates should be analyzed more carefully. As we show by example in this section, for some games, simple and reasonable requirements on the way mutation probabilities vary across states can give significant restrictions on long-run predictions.

For example, consider the following game:

$$
\begin{array}{c c}
 & \begin{array}{c c} 1 & 2 \end{array} \\
\begin{array}{c} 1 \\ 2 \end{array} &
\left(\begin{array}{c c} 8,8 & 0,0 \\ 0,0 & 6,6 \end{array} \right)
\end{array}
$$

As in the example of Section 2.1, suppose there are three agents playing this game where, in each period, each player must choose a single strategy which he uses against each of the other two players. Again, assume that each player chooses the strategy which is a best reply given the strategies chosen by the other two players in the previous period.

As before, we can represent this evolutionary dynamic as a Markov process with a state space given by the number of agents playing strategy 2. So let $S = \{0, 1, 2, 3\}$. Constructing

the transition matrix as in the previous example, it is easy to show that

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The set of invariant distributions is

$$\{q \mid q = \theta e_1 + (1 - \theta)e_4\},$$

where e_i is a vector in $\mathcal{R}^4$ with 1 in the ith position and 0's elsewhere.

Now suppose we add mutations. Analogously to our earlier example, suppose that the mutation rate in state i is ε_i, that mutations are independent across agents and over time, that if an agent does not mutate, he changes strategy as the process above prescribes, and that if he does mutate, he goes to the opposite strategy. The implied transition matrix is:

$$\begin{pmatrix} (1-\varepsilon_0)^3 & 3\varepsilon_0(1-\varepsilon_0)^2 & 3\varepsilon_0^2(1-\varepsilon_0) & \varepsilon_0^3 \\ (1-\varepsilon_1)^3 & 3\varepsilon_1(1-\varepsilon_1)^2 & 3\varepsilon_1^2(1-\varepsilon_1) & \varepsilon_1^3 \\ \varepsilon_2(1-\varepsilon_2)^2 & (1-\varepsilon_2)^3 + 2\varepsilon_2^2(1-\varepsilon_2) & \varepsilon_2^3 + 2\varepsilon_2(1-\varepsilon_2)^2 & \varepsilon_2^2(1-\varepsilon_2) \\ \varepsilon_3^3 & 3\varepsilon_3^2(1-\varepsilon_3) & 3\varepsilon_3(1-\varepsilon_3)^2 & (1-\varepsilon_3)^3 \end{pmatrix}$$

It is tedious but not difficult to verify that the unique invariant distribution is $q'' = (q_0'', q_1'', q_2'', q_3'')$ given by

$$q_0'' = \frac{3\varepsilon_3[1 + g_1(\varepsilon)]}{Z_3}$$

$$q_1'' = \frac{9\varepsilon_0\varepsilon_3[1 + g_2(\varepsilon)]}{Z_3}$$

$$q_2'' = \frac{\varepsilon_0\varepsilon_3 g_3(\varepsilon)}{Z_3}$$

$$q_3'' = \frac{\varepsilon_0 g_4(\varepsilon)}{Z_3}$$

where Z_3 is a normalization factor making the probabilities sum to 1 and

$$\lim_{\varepsilon \to 0} g_i(\varepsilon) = 0, \quad \forall i$$

for any sequence of ε's converging to 0. Hence as $\varepsilon \to 0$,

$$\frac{q_1''}{q_0''} \to 0$$

and

$$\frac{q_2''}{q_0''} \to 0.$$

Hence we must have $q_1'', q_2'' \to 0$.

17

In our analysis of the example in Section 2.1, we argued that it seemed most reasonable to suppose that mutation rates are smallest in the state where agents get the highest possible payoffs in the game. This suggests that we should require $\varepsilon_0 \le \varepsilon_i$ for $i = 1, 2, 3$. If we impose this restriction, then

$$\frac{q_0''}{q_3''} \ge \frac{3[1 + g_1(\varepsilon)]}{g_4(\varepsilon)}$$

for all "allowed" ε. Since this ratio goes to infinity as $\varepsilon \to 0$, this restriction implies that $q_0'' \to 1$. That is, requiring mutation rates to be lowest in the state where payoffs are the highest possible uniquely selects the Pareto optimal Nash equilibrium.

Intuitively, if we impose this kind of property on the mutation rates, the state–independent mutation rates considered by KMR give the least possible weight to the Pareto efficient outcome. Since KMR's approach selects the risk dominant equilibrium, this uniqueness result will persist whenever Pareto efficiency and risk dominance coincide.

In the example of Section 2.1, these two notions differ and so even this kind of restriction on mutation rates is insufficient to generate uniqueness. In fact, the arguments given in Section 2.1 are entirely unaffected by this restriction. Hence in that game, even with this restriction, we could still obtain any invariant distribution of the mutationless process as mutation rates go to zero. More generally, we expect that more detailed modeling of the mutation process may allow mutations to refine long–run predictions but that, absent some unreasonably strong and *ad hoc* restrictions like state independence, unique predictions will typically not be possible.

REFERENCES

Blume, L., "The Statistical Mechanics of Strategic Interaction," Cornell University working paper, 1993.

Canning, D., "Learning Language Conventions in Common Interest Signaling Games," Columbia University working paper, 1992.

Duffie, D., J. Geanakoplos, A. Mas–Colell, and A. McLennan, "Stationary Markov Equilibria," working paper, 1989.

Ellison, G., "Learning, Local Interaction, and Coordination," *Econometrica*, 61, September 1993, pp. 1047–1071.

Foster, D., and P. Young, "Stochastic Evolutionary Game Dynamics," *Theoretical Population Biology*, 38, 1990, pp. 219–232.

Fudenberg, D., D. Kreps, and D. Levine, "On the Robustness of Equilibrium Refinements," *Journal of Economic Theory*, 44, April 1988, pp. 354–380.

Iosifescu, M., *Finite Markov Processes and Their Applications*, New York: John Wiley and Sons, 1980.

Kandori, M., G. Mailath, and R. Rob, "Learning, Mutation, and Long–Run Equilibria in Games," *Econometrica*, 61, January 1993, pp. 29–56.

Kandori, M., and R. Rob, "Evolution of Equilibria in the Long Run: A General Theory and Applications," University of Pennsylvania working paper, 1992.

Lagunoff, R., and A. Matsui, "The Evolution of Cooperation in Mechanisms for Public Projects," University of Pennsylvania, working paper, 1994.

Myerson, R., "Refinements of the Nash Equilibrium Concept," *International Journal of Game Theory*, 7, 1978, pp. 73–80.

Nöldeke, G., and L. Samuelson, "An Evolutionary Analysis of Backward and Forward Induction," *Games and Economic Behavior*, 5, July 1993, pp. 425–454.

Robson, A., "The Adam and Eve Effect and Fast Evolution of Efficient Equilibria in Coordination Games," University of Western Ontario working paper, 1994.

Robson, A., and F. Vega–Redondo, "Random Pairing and Long–Run Equilibria in Games," University of Western Ontario working paper, 1994.

Samuelson, L., "Stochastic Stability in Games with Alternative Best Replies," *Journal of Economic Theory*, forthcoming.

Selten, R., "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory*, 4, 1975, pp. 25–55.

Young, P., "The Evolution of Conventions," *Econometrica*, 61, January 1993, pp. 57–84.

Young, P., "An Evolutionary Model of Bargaining," *Journal of Economic Theory*, 59, Febru-

ary 1993, pp. 145–168.

20