Stochastic imitation in finite games

Jens Josephson a,*, Alexander Matros b

a Department of Economics and Business, Universitat Pompeu Fabra, and Centre de Referencia en Economia Analítica (CREA), Ramon Trias Fargas, 25-27, 08005 Barcelona, Spain
b Department of Economics, University of Pittsburgh, 4P36 Posvar Hall, Pittsburgh, PA 15260, USA

Received 7 February 2001
Available online 22 April 2004

Abstract

In this paper we model an evolutionary process with perpetual random shocks, where individuals sample population-specific strategy-payoff pairs and imitate the most successful behavior. For finite n-player games we prove that in the limit, as the perturbations tend to zero, only strategy-tuples in minimal sets closed under single better replies will be played with positive probability. If the strategy-tuples in one such minimal set have strictly higher payoffs than all outside strategy-tuples, then the strategy-tuples in this set will be played with probability one in the limit, provided the minimal set is a product set and the sample is sufficiently large.

© 2004 Elsevier Inc. All rights reserved.

JEL classification: C72; C73

1. Introduction

In most game-theoretical models of learning, the individuals are assumed to know a great deal about the structure of the game, such as their own payoff function and all players’ available strategies. However, for many applications, this assumption is neither reasonable nor necessary; in many cases, individuals may not even be aware that they are playing a game. Moreover, equilibrium play may be achieved even with individuals who have very little knowledge of the game, an observation made already in 1950 by John F. Nash. In
his unpublished PhD thesis (1950), he referred to it as “the ‘mass-action’ interpretation of equilibrium points.” Under this interpretation:

“It is unnecessary that the participants have full knowledge of the total structure of the game, or the ability and inclination to go through any complex reasoning processes. But the participants are supposed to accumulate empirical information on the relative advantages of the various pure strategies at their disposal.”

In the present paper, we develop a model in this spirit, where individuals are only required to know their own available pure strategies and a sample of the payoffs that a subset of these strategies have earned in the past. We use an evolutionary framework with perpetual random shocks similar to Young (1993), but our assumption of individual behavior is different. Whereas the individuals in his model play a myopic best reply to a sample distribution of their opponents’ strategies, the individuals in our model imitate other individuals in their own population. Imitation is a behavior with both experimental, empirical, and theoretical support.1

More specifically, we assume that in every period, individuals are drawn at random from each of \( n \) arbitrary-size populations to play a finite \( n \)-player game. Each of these individuals observes a sample from a finite history of her population’s past strategy and payoff realizations. Thereafter, she imitates by choosing the most attractive strategy in her sample. This could, for instance, be the strategy with the highest average payoff, or that with the highest maximum payoff. In the special case when each population consists of only one individual, this behavior can be interpreted as a special kind of reinforcement learning.2 In this case, each individual plays the most successful strategy in a sample of her own previous strategy choices. With some small probability, the individuals also make errors or experiment, and instead choose any strategy at random from their set of strategies.3 Altogether, this results in an ergodic Markov process, which we denote imitation play, on the space of histories. We study the stationary distribution of this process as the experimentation probability tends to zero.

Imitation in a stochastic setting has previously been studied by Robson and Vega-Redondo (1996), who modify the framework of Kandori et al. (1993) to allow for random matching. More precisely, they assume that in each period, individuals are randomly matched for a finite number of rounds and tend to adopt the strategy with the highest average payoff across the population. Robson and Vega-Redondo (1996) assume either

1 For experimental support of imitation, see, for example, Apesteguia et al. (2003), Huck et al. (1999, 2000), and Duffy and Feltovich (1999); for empirical support, see Graham (1999), Wermers (1999), and Griffiths et al. (1998); and for theoretical support, see Björnerstedt and Weibull (1996) and Schlag (1998, 1999).

2 This behavior is related to one of the interpretations of individual behavior in Osborne and Rubinstein (1998), where each individual first samples each of her available strategies once and then chooses the strategy with the highest payoff realization.

3 An alternative interpretation, which provides a plausible rationale for experimentation and is consistent with the knowledge of individuals in the model, is the following: if and only if the sample does not contain all available strategies, then with a small probability, the individual instead picks a strategy not included in the sample at random.
Our model differs from these and other stochastic learning models, and has several advantages. First, we are able to prove general results, applicable to any finite \( n \)-player game, about the limiting distribution of imitation play. We are thus not restricted to the two classes of games in Robson and Vega-Redondo (1996), or even to a generic class of games, as in Young (1998). Second, we find that this distribution has some interesting properties. For instance, it puts probability one on an efficient set of outcomes in a large class of \( n \)-player games. Third, the speed of convergence of our process is relatively high. We show that in \( 2 \times 2 \) coordination games, the expected first passage time may be considerably shorter than in Young (1993), Kandori et al. (1993), and Robson and Vega-Redondo (1996), for small experimentation probabilities.

The perturbed version of imitation play is a regular perturbation of this Markov process. This implies that the methods employed by Young (1993) can be used to calculate the states that will be played with positive probability by the stationary distribution of the process as the experimentation probability tends to zero, i.e. the stochastically stable states.

We prove three results which facilitate this calculation and enable us to characterize the set of such states. First, we show that from any initial state, the unperturbed version of imitation play converges to a state which is a repetition of a single pure-strategy profile, a monomorphic state. Hence, the stochastically stable states of the process belong to the set of monomorphic states.

Second, we prove for the perturbed process that in the limit, as the experimentation probability tends to zero, only pure-strategy profiles in particular subsets of the strategy-space are played with positive probability. These sets, which we denote minimal sets closed under single better replies (minimal cusber sets), are minimal sets of strategy profiles such that no player can obtain a weakly better payoff by deviating unilaterally and playing a strategy outside the set.

Minimal cusber sets are similar to Sobel’s (1993) definition of non-equilibrium evolutionary stable (NES) sets for two-player games and to what Noldeke and Samuelson (1993) call locally stable components in their analysis of extensive form games. They are also closely related to minimal sets closed under better replies (Ritzberger and Weibull, 1995). Every minimal set closed under better replies contains a minimal cusber set and if a minimal cusber set is a product set, then it is also a minimal set closed under better replies. The relationship between minimal cusber set and the limiting distribution of imitation play should be contrasted with Hurkens’s (1995) and Young’s (1998) findings that adaptive best-reply processes for generic games selects pure-strategy profiles in minimal sets closed under best replies.

Finally, we show that in a certain class of games, imitation play selects efficient outcomes. If the pure-strategy profiles in a minimal cusber set have strictly higher payoffs than all other pure-strategy profiles, then the pure-strategy profiles in this set will be played with probability one in the limit, provided that the minimal cusber set is a product set. This is a generalization of previous results for games of common interest. Robson and Vega-Redondo (1996) prove that in their model a Pareto-dominant pure-strategy profile is selected in two-player games of common interest.
Applied to $2 \times 2$ games, our three results give clear predictions. In games without pure Nash equilibria, all four monomorphic states are stochastically stable and in games with a unique strict Nash equilibrium the corresponding monomorphic state is a unique stochastically stable state. In coordination games, imitation play selects the strictly Pareto-superior Nash equilibrium. This result differs sharply from the predictions in Young’s (1993) and Kandori et al. (1993) models, where the stochastically stable states correspond to the risk-dominant equilibria, but it is consistent with the predictions of Robson and Vega-Redondo’s (1996) model for symmetric coordination games. However, if neither equilibrium Pareto dominates the other, the latter model may select the risk-dominant equilibrium, whereas both equilibria are played with positive probability in our model.

The paper is organized as follows. In Section 2, we define the unperturbed and perturbed versions of imitation play. In Section 3, we derive general results for the limiting distribution of the process. In Section 4, we apply our results to $2 \times 2$ games and compare our findings to those in previous literature and in Section 5, we discuss an extension of the model. Omitted proofs can be found in Appendix A.

2. The model

The model described below is similar to Young (1993), but the sampling procedure is modified and individuals employ a different decision-rule. Let $\Gamma$ be an $n$-player game in strategic form with player roles $i \in N = \{1, \ldots, n\}$, pure-strategy sets $X_i$, and payoffs represented by the utility functions $\pi_i : X \to \mathbb{R}$, where $X = \prod_i X_i$. To each player role $i$ in the game $\Gamma$ corresponds a finite and non-empty population of individuals who all have the utility function $\pi_i(x)$. These populations need not be of the same size, nor need they be large.

Let $t = 1, 2, \ldots$ denote successive time periods. The game $\Gamma$ is played once in every period. In period $t$, one individual is drawn at random from each of the $n$ populations and assigned to play the corresponding role in the game $\Gamma$. The individual drawn to play in role $i$ chooses a pure strategy $x_t^i$ from her pure-strategy set $X_i$ according to a decision rule that is common to all individuals and that will be defined below. The pure-strategy profile $x_t = (x_1^t, \ldots, x_n^t)$ is recorded and referred to as the play at time $t$. The sequence of plays $h^t = (x_1^1, \ldots, x_t^t)$ is referred to as the history at time $t$.

The decision rule is defined as follows. Let $s$ and $m$ be integers such that $m \geq 4$ and $1 < s \leq m/2$. In period $t + 1$, the individual in player role $i$ inspects $s$ strategy-payoff pairs drawn without replacement from the most recent $m$ role-specific strategy-payoff pairs, $(x_t^{i-m+1}, \pi_i(x_t^{i-m+1})), \ldots, (x_t^i, \pi_i(x_t^i))$. We assume that any sample of length $s$ is drawn with positive probability, and that the samples are independent of time and of the samples drawn by individuals in other player roles. The individual in player role $i$ thereafter chooses the strategy with the highest average payoff in the sample. If there are more than one such strategy, we assume that each of them is chosen with equal probability.

---

4 Actually, the utility functions need not be identical within each population for any of the results in this paper. It is sufficient if each individual’s utility function is a positive affine transformation of the role-specific utility function.
We can think of the sampling process as beginning in period \( t = m + 1 \) from some arbitrary initial sequence of plays \( h^m \). In this period and every period thereafter, the individuals drawn to play in the different player roles all choose pure strategies according to the above decision rule. This defines a finite Markov process on the finite state space \( H = X^m \) of histories of length \( m \). Given a state \( h^t = (x_{t-m+1}, \ldots, x_t) \) at time \( t \), this process moves with probability one to a state of the form \( h^{t+1} = (x_{t-m+2}, \ldots, x_t, x_{t+1}) \) in the next period. Such a state is called a successor of \( h^t \). For each \( x_i \in X_i \), let \( p_i(x_i | h) \) be the probability that the individual in player role \( i \) chooses \( x_i \) in state \( h \). Our assumptions imply that \( p_i(x_i | h) \) is independent of time and that \( p_i(x_i | h) > 0 \) if and only if there exists a sample of \( s \) strategy-payoff pairs for player role \( i \) where \( x_i \) has maximum average payoff.

If \( h' \) is a successor of \( h \) and \( x \) is the right-most element of \( h' \), the transition probability becomes

\[
P^{m,s}(h | h') = \prod_{i=1}^{n} p_i(x_i | h).
\]

If \( h \) is not a successor of \( h' \), \( P^{m,s}(h | h') = 0 \). We denote this Markov process by \( P^{m,s} \) and refer to it as imitation play with memory \( m \) and sample size \( s \).

As an example, consider imitation play with memory \( m = 6 \) and sample size \( s = 3 \) in the \( 2 \times 3 \) game in Fig. 1.

Let \( h = ((A, a), (B, a), (A, b), (B, b), (A, c), (B, c)) \) be the initial state. Assume that the individual in the role of the row player (player 1) draws the most recent three strategy-payoff pairs in this state, \( ((B, 1), (A, 3), (B, 0)) \). This gives an average payoff of 3 to strategy \( A \) and 1/2 to strategy \( B \). Hence, the individual in role of the row player will choose strategy \( A \) in the next period. Further, assume that the individual in the role of the column role (player 2) draws the first three strategy-payoff pairs, \( ((a, 2), (a, 0), (b, 0)) \). This gives an average payoff of 1 to strategy \( a \) and 0 to strategy \( b \). Strategy \( c \) cannot be chosen since it is not included in the sample. Hence, the individual in the column role will choose strategy \( a \) in the next period. Altogether, this implies that the unperturbed process will move to state \( h' = ((B, a), (A, b), (B, b), (A, c), (B, c), (A, a)) \) in the next period.

Following Young (1993), we also define a perturbed version of imitation play. In each period and for each player role \( i \), there is some small probability \( \varepsilon > 0 \) that the individual drawn to play chooses a pure strategy at random from \( X_i \), instead of according to the decision rule described above. The event that \( i \) experiments is assumed to be independent of the event that \( j \) experiments for every \( j \neq i \) and across time periods. We denote the process defined in this way by \( P^{m,s,\varepsilon} \) and refer to it as imitation play with memory \( m \), sample size \( s \), and experimentation probability \( \varepsilon \).
3. Asymptotic behavior of imitation play

We start by examining the long-term properties of the unperturbed version of imitation play and thereafter proceed with the perturbed version. In what follows, we will make use of the following definitions. A recurrent class of the process \( P_{m,s} \) is a set of states such that there is zero probability of moving from any state in the class to any state outside, and there is a positive probability of moving from any state in the class to any other state in the class. A state \( h' \) is absorbing if it constitutes a singleton recurrent class. We denote a state \( h = (x, x, \ldots, x) \), where \( x \) is any pure-strategy profile from \( X \), by \( h_x \) and refer to it as a monomorphic state.

The following result shows that the unperturbed version of imitation play converges to a monomorphic state from any initial state.

**Theorem 1.** A subset of states is a recurrent class if and only if it is a singleton set containing a monomorphic state.

**Proof.** It is evident that any monomorphic state is an absorbing state, since any sample from a monomorphic state will contain one strategy only. We shall prove that singleton sets containing a monomorphic state are the only recurrent classes of the unperturbed process. Consider an arbitrary initial state \( h_t = (x_{t-1}^s + 1, \ldots, x_t^s) \). Since \( s/m \leq 1/2 \), there is a positive probability that all individuals drawn to play sample from \( (x_{t-1}^s + 1, \pi_{t-1}^s(x_{t-1}^s + 1)), \ldots, (x_t^s, \pi_t^s(x_t)) \) in every period from \( t + 1 \) to \( t + s \) inclusive. All of them play the pure strategy with the highest average payoff in their sample. Without loss of generality, assume that this is a unique pure strategy \( x^*_i \) for each of the player roles. With positive probability, all the individuals drawn to play thereafter sample only from plays more recent than \( x_i^* \) in every period from \( t + s + 1 \) to \( t + m \) inclusive. Since all of these samples have the form \( (x^*_i, \pi^*_i(x^*_i)), \ldots, (x^*_i, \pi^*_i(x^*_i)) \), the unique pure strategy with the highest payoff in the sample is \( x^*_i \). Hence, there is a positive probability of at time \( t + m \) obtaining a state \( h_{x^*} = (x^*, \ldots, x^*) \), a monomorphic state. It follows that the only recurrent classes of the unperturbed process are singletons containing a monomorphic state.

We now turn our attention to the perturbed version of imitation play. In particular, we are interested in the limiting distribution of this process as the experimentation probability tends to zero. By arguments similar to those in Young (1993), the perturbed process \( P_{m,s,\varepsilon} \) is a regular Markov chain, and therefore has a unique stationary distribution \( \mu^\varepsilon \) satisfying the equation \( \mu^\varepsilon P_{m,s,\varepsilon} = \mu^\varepsilon \). Moreover, by Theorem 4 in Young (1993), \( \lim_{\varepsilon \to 0} \mu^\varepsilon = \mu^0 \) exists, and \( \mu^0 \) is a stationary distribution of \( P_{m,s} \).

The following concepts are due to Freidlin and Wentzell (1984), Foster and Young (1990), and Young (1993). A state \( h \in H \) is stochastically stable relative to the process \( P_{m,s,\varepsilon} \) if \( \lim_{\varepsilon \to 0} \mu^\varepsilon(h) > 0 \). Let \( h' \) be a successor of \( h \) and let \( x \) be the right-most element of \( h' \). A mistake in the transition from \( h \) to \( h' \) is a component \( x_i \) of \( x \) that does not have the maximum average payoff in any sample of strategies and payoffs from \( h \). For any two states \( h, h' \), the resistance, \( r(h, h') \), is the total number of mistakes involved in the transition \( h \to h' \) if \( h' \) is a successor of \( h \), otherwise \( r(h, h') = \infty \). Identify the state space \( H \) with the vertices of a directed graph. For every pair of states \( h, h' \), insert a directed
edge \( h \rightarrow h' \) if \( r(h, h') \) is finite, and let \( r(h, h') \) be its resistance. For each pair of distinct monomorphic states, an \( x\!y \) -path is a sequence of states \( \xi = (h_x, \ldots, h_y) \) beginning in \( h_x \) and ending in \( h_y \). The resistance of this path is the sum of the resistances on the edges that compose it. Let \( r_{xy} \) be the least resistance over all \( x\!y \) -paths. Construct a complete directed graph with \(|X|\) vertices, one for each recurrent class. The weight on the directed edge \( h_x \rightarrow h_y \) is \( r_{xy} \). A tree rooted at \( h_x \) is a set of \(|X| - 1\) directed edges such that, from every vertex different from \( h_x \), there is a unique directed path in the tree to \( h_x \). The resistance of such a rooted tree \( \mathfrak{A}(x) \) is the sum of the resistances \( r_{xy} \) on the \(|X| - 1\) edges that compose it. The stochastic potential \( \rho(x) \) of a monomorphic state \( h_x \) is the minimum resistance over all trees rooted at \( h_x \).

The following theorem describes the long-run behavior of the perturbed process as the experimentation probability tends to zero.

**Theorem 2.** The stochastically stable states of \( P^{m,s} \) are the monomorphic states with minimum stochastic potential.

**Proof.** This follows from Theorem 1 above and Theorem 4 in Young (1993). \( \Box \)

In order to illustrate how to calculate the stochastic potential under imitation play, we present an example of a two-player game. In the game in Fig. 2, every player has three strategies, labeled \( A, B \) and \( C \) for the first player and \( a, b \) and \( c \) for the second player. The game has one strict Nash equilibrium \((A, a)\), where both players gain less than in a mixed equilibrium with the probability mixture \( 1/2 \) on \( B(b) \) and \( 1/2 \) on \( C(c) \) for the first (second) player.\(^5\)

Denote by \( x_1 \in \{A, B, C\} \) some strategy choice by player 1 and \( x_2 \in \{a, b, c\} \) some strategy choice by player 2. To find the stochastically stable monomorphic states, construct directed graphs with nine vertices, one for each monomorphic state. In Fig. 3, we illustrate two such trees. The numbers in the squares correspond to the resistances of the directed edges and the numbers in the circles represent the payoffs associated with the monomorphic states. It is easy to check that \( \rho(A, a) = 8 \), whereas all other monomorphic states have a stochastic potential of 9. Hence, the monomorphic state \( h(A, a) \) is stochastically stable.

In order to characterize the sets of monomorphic states with minimum stochastic potential a particular correspondence turns out to be useful. Let the better-reply correspondence, \( \gamma_i : X \rightarrow X_i \), be defined as follows:

\[
\gamma_i(x) = \{ x_i \in X_i \mid \pi_i(x_i, x_{-i}) \geq \pi_i(x) \}.
\]

\[\text{Fig. 2.}\]

\[
\begin{array}{ccc}
  & a & b \\
 A & 1,1 & 0,0 & 0,0 \\
 B & 0,0 & 3,2 & 0,3 \\
 C & 0,0 & 3,2 & 3,3 \\
\end{array}
\]

\[^{5}\text{There is also a third equilibrium, } (\left(\frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{2}, \frac{1}{2}\right)).\]
This concept (with a mixed-strategy domain) is originally due to Ritzberger and Weibull (1995). Using the better-reply correspondence, we can define the following set.

**Definition 1.** A non-empty set of strategy profiles $V$ is closed under single better replies, or a cusber set, if, for each $x \in V$ and $i \in N$, $(\gamma_i(x), x_{-i}) \subseteq V$. Such a set is called a minimal cusber set if it does not properly contain another cusber set.

From the definition it follows that every game contains a minimal cusber set and that any minimal closed set under better replies (Ritzberger and Weibull, 1995) contains a minimal cusber set. However, unlike minimal closed sets under better replies, minimal cusber sets are not necessarily product sets (the set \{(C, b), (C, c), (B, c)\} is a minimal cusber set of the game in Fig. 4).

We are now in a position to state the following main theorem.

**Theorem 3.** Let $\mathcal{V}$ be the set of stochastically stable monomorphic states. Then $V = \{x \in X: h_x \in \mathcal{V}\}$ is a minimal cusber set or a union of minimal cusber sets.

**Proof.** See Appendix A.

Theorem 3 establishes a relation between the stochastically stable states of imitation play and minimal cusber sets, which is similar to the relationship between the stochastically stable states and minimal curb sets first proved for a particular dynamics by Hurkens (1995), and later modified for a different dynamic by Young (1998, p. 111).
We say that a finite set $Y$ of pure-strategy profiles strictly Pareto dominates a pure-strategy profile $x$ if for any pure-strategy profile $y \in Y$, $\pi_i(y) > \pi_i(x)$, for all $i$. The following theorem shows that if the sample size is sufficiently large, imitation play selects sets of efficient outcomes in a large class of games.

**Theorem 4.** If a minimal cusber set $V$ is a product set that strictly Pareto dominates all pure-strategy profiles outside $V$, then there exists $s^* \geq 1$ such that for every sample size $s > s^*$, the stochastically stable states of $P^{m,1}$ are the monomorphic states in the set $\mathcal{V} = \{h_x \in H : x \in V\}$.

**Proof.** See Appendix A.

The intuition behind this result is that for a sufficiently large sample size, the transition from a state inside $\mathcal{V}$ to any state outside $\mathcal{V}$ requires more mistakes than the number of player roles, while the opposite transition requires one mistake per player role at most.

The requirement in Theorem 4 that $V$ be a product set is necessary, as shown by the game in Fig. 4.

In this game, the minimal cusber set $V = \{(C, b), (C, c), (B, c)\}$ strictly Pareto dominates all pure-strategy profiles outside $V$. It is evident that two mistakes are enough to move from the monomorphic state $h_{(A,a)}$ to any monomorphic state in $\mathcal{V} = \{h_x \in H : x \in V\}$. We will show that two mistakes are also enough to move from $\mathcal{V}$ to a monomorphic state outside of $\mathcal{V}$. Suppose the process is in the state $h_{(C,c)}$ at time $t$. Further, suppose that the individual in player role 1 plays $B$ instead of $C$ at time $t + 1$ by mistake. This results in play $(B, b)$ at time $t + 1$. Assume that the individual in player role 2 makes a mistake and plays $b$ instead of $c$, and that the individual in player role 1 plays $C$ in period $t + 2$. Hence, the play at time $t + 2$ is $(C, b)$. Assume that the individuals in both player roles sample from period $t - s + 2$ to period $t + 2$ for the next $s$ periods. This means that the individuals in player role 1 choose to play $B$ and the individuals in player role 2 choose to play $b$ from period $t + 3$ to period $t + s + 2$. There is a positive probability that from period $t + s + 3$ through period $t + m + 2$, the individuals in both player roles will sample from periods later than $t + 2$. Hence, by the end of period $t + m + 2$, there is a positive probability that the process will have reached the monomorphic state $h_{(B,b)}$ outside $\mathcal{V}$. It is now straightforward to show that all the monomorphic states $h_{(A,a)}$, $h_{(C,b)}$, $h_{(C,c)}$, and $h_{(B,c)}$ have equal stochastic potential.

### 4. Applications to $2 \times 2$ games

We now apply the above results to $2 \times 2$ games. By Theorem 3, it follows immediately that in $2 \times 2$ games with a unique strict Nash equilibrium, the corresponding monomorphic state is a unique stochastically stable state. It also follows that in $2 \times 2$ games without Nash equilibria in pure strategies the stochastically stable states correspond one to one with the four monomorphic states of the game.

By Theorem 4, it follows that in $2 \times 2$ games with two strict Nash equilibria, where one Nash equilibrium strictly Pareto dominates the other, there exists $s^* \geq 1$ such that for
every sample size $s > s^*$, the unique stochastically stable state corresponds one to one with the monomorphic state of the Pareto dominant equilibrium. This implies that, unlike Young’s (1993) process of adaptive play, imitation play does not generally converge to the risk-dominant equilibrium in coordination games. Our result is consistent with Robson and Vega-Redondo’s (1996) result for generic symmetric coordination games. However, the following proposition demonstrates that for the non-generic case when the equilibrium payoffs are equal for at least one of the player roles, the stochastically stable states in their model depend on the details of the adjustment process, whereas imitation play always selects both equilibria.

**Proposition 1.** In $2 \times 2$ games with two strict Nash equilibria, where neither strictly Pareto dominates the other, the stochastically stable states correspond one to one with the monomorphic states of the two equilibria.

**Proof.** See Appendix A.

We can also say something about the speed of convergence in $2 \times 2$ games by observing that in such games, the transition from an arbitrary state to a stochastically stable monomorphic state requires two mistakes at most. This immediately implies the following proposition.

**Proposition 2.** In $2 \times 2$ games, the maximum expected first passage time for the perturbed process $P_{m,s,\varepsilon}$ from any state to a stochastically stable state is at most $\delta\varepsilon^{-2}$ units of time, for some positive constant $\delta$.

This result should be contrasted with the speed of convergence in Young (1993), Kandori et al. (1993), and Robson and Vega-Redondo (1996). In Young’s (1993) model, the maximum expected first passage for a $2 \times 2$ coordination game is at least $\delta v\varepsilon^{-v}$ where $v$ depends on the sample size and both players’ payoffs. In Kandori et al. (1993) the maximum expected first passage time is of the order $\delta KMR\varepsilon^{-Nu}$, where $N$ is the size of the population and $u$ is determined by the game’s payoff structure. In Robson and Vega-Redondo (1996), the corresponding figure is $\delta RV\varepsilon^{-q}$, where $q$ is a positive integer independent of the payoffs and the current state. Thus, when $v$, $Nu$, and $q$ are greater than two and $\varepsilon$ is sufficiently small, then imitation play converges considerably faster than the processes in these three models.

5. Extensions

All results in this paper hold for a more general class of imitation dynamics. Suppose individuals samples strategy-payoff pairs in the same manner as in this paper, but that they thereafter choose a strategy with maximum payoff instead of a strategy with maximum average payoff. This behavior defines a new Markov process on the space of histories with the same set of absorbing states and (for a sufficiently large sample size) stochastically stable states as imitation play. Moreover, if each population consists of arbitrary shares
of individuals who maximize payoffs and average payoffs, respectively, then the results of this paper still hold. Hence, the model allows for a certain kind of population heterogeneity, where individuals make their choices based on different rules.

Acknowledgments

The authors are grateful for helpful comments and suggestions from Jörgen W. Weibull and two anonymous referees. We have also benefited from comments by Bo Becker, Ken Binmore, Eric van Damme, Maria Sáez-Marti, Philippe Solal, H. Peyton Young, and seminar participants at the Stockholm School of Economics 1999, the Young Economist Meeting 2000, and the First World Congress of the Game Theory Society 2000. We gratefully acknowledge financial support from the Jan Wallander and Tom Hedelius Foundation.

Appendix A

Proof of Theorem 3. We will prove Theorem 3 using five lemmata.

Lemma 1. The resistance from $h_x$ to $h_y$ is positive for any $y \neq x$. It is equal to one if and only if $y \in (\gamma_i(x), x_{-i})$ for some player $i$.

Proof. The first statement follows since a monomorphic state consists of a repetition of a single strategy profile and since only strategies included in the sample can be selected in the absence of mistakes. The resistance is equal to one if $y \in (\gamma_i(x), x_{-i})$ for some player $i$, since then $h_y$ can be reached if an individual in player role $i$ plays $y_i$ by mistake, and $m-1$ consecutive individuals in player role $i$ thereafter draw a sample including strategy $y_i$. Moreover, it is clear that if $y \notin (\gamma_i(x), x_{-i})$ for all $i$, then a single mistake is insufficient to make any individual change her strategy. Hence, the resistance of the transition from $h_x$ to $h_y$ is equal to one only if $y \in (\gamma_i(x), x_{-i})$ for some player $i$.

Lemma 2. If $x' \in (\gamma_i(x), x_{-i})$ for some player $i \in N$, then $\rho(x) \geq \rho(x')$.

Proof. By definition, $\rho(x)$ is the minimum resistance over all trees rooted at state $h_x$. Construct a tree rooted at $h_{x'}$ by taking one of the trees with minimum resistance rooted at $h_x$, adding the directed edge from $x$ to $x'$ and deleting the directed edge from $x'$. By Lemma 1, the resistance of the added edge is exactly one and that of the deleted edge at least one, so the total resistance of the new tree is $\rho(x)$ at most.

Let a better-reply path be a sequence of pure-strategy profiles $x^1, x^2, \ldots, x^l$ such that for every $k \in \{1, \ldots, l - 1\}$, there exists a unique player, say player $i$, such that $x^{k+1} \in (\gamma_i(x^k), x_{-i}^k)$ and $x_i^{k+1} \neq x_i^k$.

Lemma 3. For any two strategy profiles $x, x'$ in a minimal cusber set $V$, there exist better-reply paths $x, \ldots, x'$ and $x', \ldots, x$, which connect these strategy profiles.
Proof. Let \( V \) be a minimal cusber set. Suppose that the claim is false and there exist two pure strategy profiles \( x, x' \in V \) such that there is no better-reply path from \( x \) to \( x' \). Consider all better-reply paths, starting at strategy profile \( x \). There are a finite number of paths and a finite number of pure strategy profiles along all these paths. Collect all these strategy profiles in a set. By construction this set is a cusber set and by assumption it does not contain the vertex \( x' \), contradicting the assumption of \( V \) being a minimal cusber set.

**Lemma 4.** For every minimal cusber set \( V \) and any \( x, x' \in V \), \( \rho(x) = \rho(x') \).

**Proof.** On the one hand, by Lemma 3, there exists a better-reply path from an arbitrary strategy profile \( x \) in a minimal cusber set to any other strategy profile \( x' \) in the same minimal cusber set. Let the sequence \( x, \ldots, x' \) be such a path. By Lemma 2, the following inequalities hold for the stochastic potential of the corresponding monomorphic states \( h_x, \ldots, h_{x'} : \)

\[
\rho(x) \geq \cdots \geq \rho(x'). \tag{A.1}
\]

On the other hand, by applying Lemma 3 once more, there exist a better-reply path from the strategy tuple \( x' \) to the strategy tuple \( x \). Using Lemma 2, gives

\[
\rho(x') \geq \cdots \geq \rho(x). \tag{A.2}
\]

From the inequalities in (A.1) and (A.2) it follows that \( \rho(x) = \rho(x') \) for the monomorphic states \( h_x \) and \( h_{x'} \). \( \square \)

**Lemma 5.** If \( x \in X \) does not belong to a minimal cusber set, then there exists \( x' \in X \) such that \( \rho(x) > \rho(x') \).

**Proof.** For every strategy profile not included in any minimal cusber set, there exists a finite better-reply path which ends in some minimal cusber set. Let this path be \( x^1, x^2, \ldots, x^{T-1}, x^T \), where \( x^1 \) is an arbitrary strategy profile that does not belong to any minimal cusber set and \( x^T \) the first monomorphic state on the path belonging to some minimal cusber set, \( V \). By Lemma 2, it follows that the following inequalities hold for the stochastic potential of the corresponding monomorphic states:

\[
\rho(x^1) \geq \cdots \geq \rho(x^{T-1}) \geq \rho(x^T). \tag{A.3}
\]

We will show that, in fact, \( \rho(x^{T-1}) > \rho(x^T) \). Note that \( \rho(x^{T-1}) \) is the minimum resistance over all trees rooted at the state \( h_{x^{T-1}} \). Denote (one of) the tree(s) that minimizes resistance by \( \mathcal{Z}(x^{T-1}) \). Find in the tree \( \mathcal{Z}(x^{T-1}) \) a directed edge from some vertex \( h_y \) such that \( y \) is in the minimal cusber set \( V \), to some other vertex \( h_{y'} \) such that \( y' \) is outside this minimal cusber set. It will be shown later that there is only one such directed edge in the minimal resistance tree \( \mathcal{Z}(x^{T-1}) \). Delete in the tree \( \mathcal{Z}(x^{T-1}) \) the directed edge \( h_y \rightarrow h_{y'} \) and add the directed edge \( h_{y'} \rightarrow h_{y} \). As a result, we obtain a tree \( \mathcal{Z}(y) \) rooted at the state \( h_y \). By Lemma 1, the resistance of the deleted edge is greater than one, and the resistance of the added edge is one. Therefore, the total resistance of the new tree \( \mathcal{Z}(y) \) is less than the stochastic potential \( \rho(x^{T-1}) \). Moreover, by Lemma 4, the
monomorphic state $h_{xT}$ has the same stochastic potential as the monomorphic state $h_y$. Hence, $\rho(xT^{-1}) > \rho(xT)$.

We will now consider the tree $\mathcal{S}(xT^{-1})$ and show that there is only one directed edge from the monomorphic states which consists of a repetition of a strategy profile in a minimal cusber set to a state which consists of a repetition of a strategy profile outside the cusber set. Suppose there is a finite number of such directed edges $h_{y^j} \to h_{z^j}$, $j = 1, 2, \ldots, l$, where $y^1, \ldots, y^l$ are strategy profiles in the minimal cusber set and $z^1, \ldots, z^l$ strategy profiles outside the cusber set. It is clear that there cannot be an infinite number of outgoing edges since the game $\Gamma$ is finite. Recall that a tree rooted at vertex $h_{y^1}$ is a set of $|X| - 1$ directed edges such that, from every vertex different from $h_{y^1}$, there is a unique directed path in the tree to $h_{y^1}$. The resistance of any directed edge $h_{y^j} \to h_{z^j}$, $j = 1, 2, \ldots, l$, is at least two. By Lemma 3, there exists a finite better-reply path from vertex $y^1$ to vertex $y^2$ in the minimal cusber set. Let $y^1, f^1, \ldots, f^k, y^2$ be such a path.

Consider the vertex $h_{f^1}$. There are two mutually exclusive cases:

1. there exists a directed path from $h_{f^1}$ to one of the vertices $h_{y^2}, \ldots, h_{y^l}$ in the initial tree $\mathcal{S}(xT^{-1})$, or
2. there exists a directed path from $h_{f^1}$ to $h_{y^1}$.

In case (1), by deleting the directed edge $h_{y^1} \to h_{y^1}$ and adding the directed edge $h_{y^1} \to h_{f^1}$ to the tree $\mathcal{S}(xT^{-1})$, we obtain a new tree $\mathcal{S}^1(xT^{-1})$ with lower stochastic potential than $\mathcal{S}(xT^{-1})$, because the resistance of the directed edge $y^1 \to f^1$ is one. This means that we are done, since it contradicts the assumption of $\mathcal{S}(xT^{-1})$ being a minimal resistance tree.

In case (2), we will use the following procedure for vertex $h_{f^1}$: delete the initial directed edge from $h_{f^1}$ and add the directed edge $h_{f^1} \to h_{f^2}$. As above, there are two cases:

1. there exists a directed path from $h_{f^2}$ to one of the vertices $h_{y^2}, \ldots, h_{y^l}$ in the initial tree $\mathcal{S}(xT^{-1})$, or
2. there exists a directed path from $h_{f^2}$ to $h_{y^1}$.

In case (2), we obtain a new tree $\mathcal{S}^2(xT^{-1})$ with lower stochastic potential than $\mathcal{S}(xT^{-1})$, because the resistance of the directed edge $h_{f^1} \to h_{f^2}$ is one. This means that we are done, since it contradicts the assumption of $\mathcal{S}(xT^{-1})$ being a minimal resistance tree.

In case (2), we repeat the procedure for vertices $h_{f^2}, h_{f^3}, \ldots$. The better-reply path $y^1, f^1, \ldots, f^k, y^2$ from $y^1$ to $y^2$ is finite. Hence, after $k + 1$ steps at most, we have constructed a tree $\mathcal{S}^k(xT^{-1})$ rooted at the state $h_{xT^{-1}}$ with lower stochastic potential than $\mathcal{S}(xT^{-1})$. □

Theorem 3 follows immediately from Lemmas 4 and 5, and Theorem 2. □
Proof of Theorem 4. Suppose $V \subset X$ is a minimal cusber set which strictly Pareto dominates all pure-strategy profiles outside $V$. We will prove Theorem 4 using the following two lemmata.

**Lemma 6.** The transition from any monomorphic state $h_z$ to a monomorphic state $h_x$ such that $x \in V$ requires at most $n$ mistakes, independently of the sample size.

**Proof.** Assume that the process is in state $h_z$ and that the individuals in all player roles simultaneously make mistakes, so that $x$ is played instead of $z$. Since, by assumption, $\pi_i(x) > \pi_i(z)$ for all player roles $i$, if the individuals in all player roles sample the most recent plays for the next $m - 1$ periods, this will take the process to the state $h_x$. \hfill $\square$

**Lemma 7.** There exists $s^* \geq 1$ such that for sample size $s > s^*$, the transition from a monomorphic state $h_x$, such that $x \in V$, to a monomorphic state $h_z$, such that $z / \in V$, requires at least $n + 1$ mistakes.

**Proof.** The transition from $h_x$ to a monomorphic state $h_z$ such that $z / \in V$ can be made if individuals in one of the player roles, say $i$, make (at least) $s$ consecutive mistakes and play a strategy $z_i \notin V_i$ every time. If the individuals in player role $i$, thereafter sample from the most recent plays for $m - s$ periods, the Markov chain will end up in the state $h_z$ where $z = (z_i, x_{-i})$. Hence, if $s > n$, this kind of transition will require more than $n$ mistakes.

Alternatively, the transition from $h_x$ to a monomorphic state $h_z$ such that $z / \in V$ can be made if a single individual in one of the player roles, say $i$, makes a mistake and plays $z_i$ and individuals in other player roles thereafter (or before this) make sufficiently many mistakes to make the average payoff of $x_i$ lower than that of $z_i$ in a sample of the most recent plays. The number of mistakes required for this kind of transition will be as low as possible if $\pi_i(x)$ is as low as possible, $\pi_i(z)$ is as high as possible, and the minimum expected payoff is achieved when $i$ plays $z_i$, an individual in a different player role $j$ plays $q_j \notin V_j$, and all other individuals drawn to play $x_{-i}$.

Let $\bar{\pi}$ be the minimum payoff to any player role for any pure-strategy profile in $V$, let $\underline{\pi}$ be the minimum payoff to any player role for any pure-strategy profile outside $V$, and let $\bar{\pi}$ be the maximum payoff to any player role for any pure-strategy profile outside $V$. By assumption, $\bar{\pi} > \bar{\pi} \geq \underline{\pi}$. By the above logic, the transition from $h_x$ to $h_z$ will require at least $n + 1$ mistakes if

\[
\frac{(s - n)\bar{\pi} + (n - 1)\underline{\pi}}{s - 1} > \bar{\pi} \tag{A.4}
\]

\[
\Leftrightarrow s > \frac{n(\bar{\pi} - \underline{\pi}) + \underline{\pi} - \bar{\pi}}{\bar{\pi} - \underline{\pi}}. \tag{A.5}
\]

Note that the right-hand side in the last expression is greater than or equal to $n$. Hence, if $s > s^* = (n(\bar{\pi} - \underline{\pi}) + \underline{\pi} - \bar{\pi})/(\bar{\pi} - \underline{\pi})$, then the transition from a monomorphic state $h_x$, such that $x \in V$, to a monomorphic state $h_z$, such that $z / \in V$, requires more than $n$ mistakes. \hfill $\square$

Let $h_y$ be an arbitrary monomorphic state such that $y / \in V$ and consider a minimal resistance tree $\mathcal{T}(y)$ rooted at $h_y$. Let $h_z$ be a monomorphic state such that $x \in V$ and such
that there is a directed edge from $h_x$ to a state $h_z$, with $z \notin V$, in the tree $\mathcal{Z}(y)$. Create a new tree $\mathcal{Z}(x)$ rooted at $h_x$ by adding a directed edge from $h_y$ to $h_x$ and deleting the directed edge from $h_x$ to $h_z$ in the tree $\mathcal{Z}(y)$. By Lemma 7, the deleted edge has a resistance greater than $n$ provided that $s > s^*$, and by Lemma 6, the added edge has a resistance of at most $n$. Hence, for $s > s^*$ the total resistance of the new tree $\mathcal{Z}(x)$ is less than that of $\mathcal{Z}(y)$. Theorem 4 now follows by Lemma 4, according to which $\rho(x) = \rho(x')$ for all $x' \in V$.

**Proof of Proposition 1.** Denote the pure strategies of player role 1 by $A$ and $B$, and the pure strategies of player role 2 by $a$ and $b$. Without loss of generality, assume that $\{(A,a)\}$ and $\{(B,b)\}$ are the minimal cusber sets of the game. By Theorem 3, it follows that monomorphic states $h_{(A,a)}$ and $h_{(B,b)}$ are the only two candidates for the stochastically stable states. Suppose that only one of these monomorphic states is stochastically stable, say $h_{(A,a)}$. Let $\mathcal{Z}(A,a)$ be a minimum resistance tree with resistance $\rho(A,a)$ rooted at $h_{(A,a)}$. In this tree, there is an outgoing edge from the monomorphic state $h_{(B,b)}$.

First, note that the resistance of this edge is at least two, such that at least two mistakes are needed to move from the monomorphic state $h_{(B,b)}$. This follows since $\pi_1(B,b) > \pi_1(A,b)$ and $\pi_2(B,b) > \pi_2(B,a)$.

Second, note that two mistakes are sufficient to move the process from the monomorphic state $h_{(B,b)}$. This follows since $\pi_1(B,b) > \pi_1(A,b)$ and $\pi_2(B,b) > \pi_2(B,a)$.

Finally, create a new tree rooted at $h_{(B,b)}$ by deleting the outgoing edge from the monomorphic state $h_{(B,b)}$ in the tree $\mathcal{Z}(A,a)$ and adding an edge from $h_{(A,a)}$ to $h_{(B,b)}$. The resistance of the deleted edge is at least two and that of the added edge two. Hence, the total resistance of the new tree is at most $\rho(A,a)$, thereby contradicting the assumption that only $h_{(A,a)}$ is stochastically stable.

**References**


