Stochastic Imitation in Finite Games

By Jens Josephson and Alexander Matros*

March 9, 2000

Abstract. In this paper we model an evolutionary process with perpetual random shocks where individual behavior is determined by imitation. Every period an agent is randomly chosen from each of n finite populations to play a game. Each agent observes a sample of population-specific past strategy and payoff realizations. She thereafter imitates by choosing the strategy with highest average payoff in the sample. Occasionally the agents also experiment or make mistakes and choose a strategy at random. For finite n-player games we prove that in the limit, as the probability of experimentation tends to zero, only strategy-tuples in minimal sets closed under the better-reply graph 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. We also show that in 2×2 games the convergence in our model is faster than in other known models.

Keywords: Evolutionary game theory, bounded rationality, imitation, Markov chain, stochastic stability, better replies, Pareto dominance.

JEL classification: C72, C73

1. Introduction

In most game-theoretical models of learning, the agents 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 agents may not even be aware that they are

*The authors are grateful for helpful comments and suggestions from Jörgen Weibull. We have also benefited from comments by Bo Becker, Ken Binmore, Maria Sáez-Marti and Philippe Solal. We gratefully acknowledge financial support from the Jan Wallander and Tom Hedelius Foundation.

playing a game. Moreover, equilibrium play may be achieved even with agents who have very little knowledge of the game. This observation was made already in 1950 by John F. Nash. In his unpublished Ph.D. thesis (1950), he referred to it as "the 'mass-action' interpretation of equilibrium points".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 only are required to know their own available strategies and a sample of the payoffs that a subset of these strategies have earned in the past. In spite of this weak assumption, our model predicts equilibrium play in many games. We use an evolutionary framework with perpetual random shocks similar to Young (1993), but our assumption of individual behavior is different from his. 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 agents in their population.

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. She thereafter imitates by choosing the strategy with highest average payoff in her sample.[1,2] With some small probability the individuals also make errors or experiment, and instead choose any strategy at random from their set

[1]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]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.

of strategies.[3] This results in an ergodic Markov chain, 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, Mailath and Rob (1993) to allow for random matching. More precisely, they assume that agents each period 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 single- or two-population structures and obtain results for symmetric 2×2 games and 2-player games of common interest.

However, our model differs considerably from this 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 and appealing 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 2×2 Coordination Games, for small experimentation probabilities the expected first passage time may be considerably shorter than in Young (1993), Kandori, Mailath and Rob (1993) and Robson and Vega-Redondo (1996).

The perturbed version of imitation play is a regular perturbed 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

[3] If the interpretation is errors, information requirement are even weaker - individuals are only required to know a subset of their available strategies.

of the process as the experimentation probability tends to zero, the *stochastically stable states*. We prove three results which facilitate this calculation and enable us to characterize the set of such states. These results hold in finite n-player games provided the information available to the agents is sufficiently incomplete and their sample size is sufficiently large.

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 strategy-tuple. For this reason we denote such states *conventions.* Hence, the stochastically stable states of the process belong to the set of conventions. Note that our definition of a convention differs from Young's (1993) in that it does not require the repeated strategy-tuple to be a strict Nash equilibrium.

Second, we prove that in the limit, as the experimentation probability tends to zero in the perturbed process, only strategy-tuples in particular subsets of the strategy-space are played with positive probability. These sets, which we denote *minimal constructions,* are defined in the following way. By drawing directed edges from each strategy-tuple to strategy-tuples which gives exactly one of the players at least as high payoff, a *better-reply graph* is obtained. We define a *construction* as a set of strategy-tuples from which there are no outgoing such edges. A minimal construction is a minimal set with this property. Minimal constructions correspond to Sobel's (1993) definition of *non-equilibrium evolutionary stable* (*NES*) sets for 2-player games and are similar to what Nöldeke 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). We show that every minimal set closed under better replies contains a minimal construction and that if a minimal construction is a product set, then it is also a minimal set closed under better replies. The relationship between minimal constructions and the limiting distribution

of imitation play should be contrasted with Young's (1998) finding that adaptive play for generic games selects strategy-tuples in minimal sets closed under best replies.

Finally, we show that in a certain class of games, imitation play selects efficient outcomes. If the strategy-tuples in a minimal construction 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 construction is a product set. This is a generalization of Robson's and Vega-Redondo's (1996) result that a Pareto-dominant strategy-tuple under certain conditions corresponds to a unique stochastically stable state in 2-player games of common interest.

Applied to 2×2 games these three results give clear predictions. In Coordination Games where one equilibrium strictly Pareto dominates the other, imitation play selects the strictly Pareto-superior Nash-equilibrium. This result differs sharply from the predictions in Young's (1993) model, where the stochastically stable states correspond to the risk-dominant equilibria, but it is consistent with the predictions of Robson's and Vega-Redondo's (1996) model for symmetric Coordination Games. However, if neither equilibria Pareto dominates the other, the latter model may select the risk-dominant equilibrium, whereas both of the equilibria are played with positive probability on in our model. In games without pure Nash-equilibria all of the four conventions are stochastically stable.

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 investigate some of the properties of our solution concept, minimal constructions. In Section 5 we apply our results to 2×2 games and compare our findings to those in previous literature. In Section 6 we conclude.

2. Model

The model we describe below is similar to Young (1993), but the sampling procedure is modified and the agents in our model employ a different decision-rule. Let Γ be a n-player game in strategic form and let X_i be the finite set of pure strategies x_i available to player $i \in \{1, ..., n\} = N$. Define the product sets $X = \prod_i X_i$ and $\Delta = \prod_i \Delta_i$, where $\Delta_i = \{p \in R_+^{|X_i|} \mid p'\mathbf{1} = 1\}$ is the simplex of individual strategies. Let $C_1, ..., C_n$ be n finite and non-empty populations of agents. These populations need not be of the same size, nor need they be large. Each member of population C_i is a candidate to play role i in the game Γ. All agents in population C_i have payoffs represented by the utility function $\pi_i : X \to \Pi_i$, where $\Pi_i \subset R$.[4] Expected payoffs are represented by the function $u_i : \Delta \to R$. Note that we write "players" when referring to the game Γ and "agents" or "individuals" when referring to the members of the populations.

Play proceeds as follows. Let $t = 1, 2, ...$ denote successive time periods. The game Γ is played once every period. In period t, one agent is drawn at random from each of the n populations and assigned to play the corresponding role. An individual in role i chooses a pure strategy x_i^t from a subset of her strategy space X_i according to a rule that will be defined below. The strategy-tuple $x^t = (x_1^t, ..., x_n^t)$ is recorded and referred to as *play* at time t. The *history* of plays up to time t is the sequence $h^t = (x^{t-m+1}, ..., x^t)$, where m is a given positive integer, the *memory size* of all individuals.

Let $h \in X^m = H$ be an arbitrary history. Denote by $w_i = (x_i^{k_1}, ..., x_i^{k_s}) \in X_i^s$ a sample of s, $1 \leq s \leq m$, elements from the m most recent strategies choices by

[4] Actually, utility functions need not be indentical within each population for any of the results in this paper. It is sufficient if each agent's utility function is a positive affine transformation of a population-specific utility function.

individuals in population C_i and by $v_i = (\pi_i^{k_1}, ..., \pi_i^{k_s}) \in \Pi_i^s$ the corresponding payoff realizations. For any history h, the *maximum average correspondence*, $\alpha_i : X_i^s \times \Pi_i^s \twoheadrightarrow X_i$, maps each pair of strategy sample w_i and payoff sample v_i to the strategy (or the set of strategies) with the highest average payoff in the sample. Following Young (1993), we can think of the sampling process as beginning in period $t = m + 1$ from some arbitrary initial sequence of m plays h^m. In this period and every period thereafter, each agent in player position i inspects a pair (w_i, v_i) and plays a strategy $x_i \in \alpha_i(w_i, v_i)$. This defines a finite Markov chain on the finite state space H of histories. Given a history $h^t = (x^{t-m+1}, ..., x^t)$ at time t, the process moves in the next period to a state of the form $h^{t+1} = (x^{t-m+2}, ..., x^t, x^{t+1})$. Such a state is called a *successor* of h^t. The process moves from the current state h to a successor state h' in each period according to the following transition rule. For each $x_i \in X_i$, let $p_i(x_i \mid h)$ be the probability that agent i chooses strategy x_i. We assume $p_i(x_i \mid h) > 0$ if and only if there exists a sample of population-specific strategy choices w_i and payoff outcomes v_i, such that $x_i \in \alpha_i(w_i, v_i)$. If x is the rightmost element of h', the probability of moving from h to h' is $B_{hh'}^{m,s,0} = \prod_{i=1}^n p_i(x_i \mid h)$ if h' is a successor of h and $B_{hh'}^{m,s,0} = 0$ if h' is not a successor of h. We call the process $B^{m,s,0}$ *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 2×3 game in Figure 1.

	a	b	c
A	$2, 2$	$0, 0$	$3, 3$
B	$0, 0$	$1, 1$	$0, 0$

FIGURE 1

Let $h = ((A, a), (B, a), (A, b), (B, b), (A, c), (B, c))$ be the initial history. Assume that the individual in the row player position (player 1) from this history draws

the last three plays, such that $w_1 = (B, A, B)$ and $v_1 = (1, 3, 0)$. This gives an average payoff of 3 to strategy A and 1/2 to strategy B. Hence, the individual in the row player position will choose strategy A in the next period. Further assume the individual in the column player position (player 2) draws the first three plays, such that $w_2 = (a, a, b)$ and $v_2 = (2, 0, 0)$. This gives an average payoff of 1 to strategy a and 0 to strategy b. Since strategy c in not included in the sample it cannot be chosen. Hence, the column player will choose strategy a in the next period. Altogether this implies that the unperturbed process will move to the state $h' = ((B, a), (A, b), (B, b), (A, c), (B, c), (A, a))$ in the next period.

The perturbed process can be described as follows. In each period there is some small probability $\varepsilon > 0$ that each individual i drawn to play chooses a strategy at random from X_i instead of according to the imitation rule. 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. For every i let $q_i(x_i \mid h)$ be the conditional probability that i chooses $x_i \in X_i$, given that i experiments and the process is in state h. We assume that $q_i(x_i \mid h) > 0$ for all $x_i \in X_i$. Suppose that the process is in state h at time t. Let $D \subset N$ be any subset of d player positions. The probability is $\varepsilon^d(1-\varepsilon)^{n-d}$ that exactly the individuals drawn to play the positions in D experiment and the others do not. Conditional on this event the transition probability of moving from h to h' is $Q^D_{hh'} = \prod_{i \in D} q_i(x_i \mid h) \prod_{i \notin D} p_i(x_i \mid h)$ if h' is a successor of h and x is the rightmost element of h' and $Q^D_{hh'} = 0$ if h' is not a successor of h. This gives the following transition probability of the perturbed Markov process, similar to equation (2) in Young (1993, p. 67):

$$B^{m,s,\varepsilon}_{hh'} = (1-\varepsilon)^n B^{m,s,0}_{hh'} + \sum_{D \subset N, D \neq \emptyset} \varepsilon^{|D|}(1-\varepsilon)^{|N \setminus D|} Q^D_{hh'}. \tag{1}$$

The process $B^{m,s,\varepsilon}$ is denoted imitation play with memory m, sample size s, and

experimentation probabilities ε.

3. Stochastic Stability in Finite Games

In this section we turn our attention to the limiting distribution of imitation play as the experimentation probability tends to zero. We first show that we can apply some of the tools in Young (1993) to calculate this distribution. We thereafter prove that it puts positive probability only on the strategy-tuples in particular subsets of the strategy-space.

3.1. Preliminaries. In what follows we will make use of the following definitions. A *recurrent class* of the process $B^{m,s,0}$ 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. We call a state h' *absorbing* if it constitutes a singleton recurrent class. We refer to a state $h_x = (x, x, ..., x)$, where x is any strategy-tuple from X, as a *convention*.[5] In other words, a convention is a state where the individuals in each player position played the same strategy during last m periods. If each player i has $|X_i| \geq 1$ strategies in the game Γ, then there are $|X| = \prod_i |X_i|$ conventions in this game. The following result shows that conventions correspond one to one with the recurrent classes of imitation play.

Theorem 1. *All convention are absorbing states of the unperturbed process $B^{m,s,0}$. If s/m is sufficiently small, the process converges with probability one to a convention.*

Proof: It is evident that for s such that $1 \leq s \leq m$, any convention is an absorbing state, since any sample from a convention will contain only one strategy.

[5]Note that this definition differs from Young's (1993) definition of a convention, in that it does not require x to be a strict Nash equilibrium. Our definition corresponds to what Robson and Vega-Redondo (1996) denote a *monomorphic* or *pure* state.

We shall prove that if $s/m \leq 1/2$, then the conventions are the only recurrent classes of the unperturbed process. Consider an arbitrary initial state $h^t = (x^{t-m+1}, ..., x^t)$. If $s/m \leq 1/2$ there is a positive probability that all agents drawn to play sample from $x^{t-s+1}, ..., x^t$ (i.e. i, for $i \in N$, sample $x_i^{t-s+1}, ..., x_i^t$ and $\pi_i^{t-s+1}, ..., \pi_i^t$) in every period from $t+1$ to $t+s$ inclusive. All of them play the strategy with the highest average payoff in their sample. Assume without loss of generality that this is a unique strategy x_i^* for each of the player positions (if there are more than one strategy all of them have positive probability according to the assumptions). With positive probability all the agents drawn to play thereafter sample only from plays more recent than x^t in every period from $t+s+1$ to $t+m$ inclusive. Since all of these samples have the form $w_i^* = (x_i^*, ..., x_i^*)$ and $v_i^* = (\pi_i^*, ..., \pi_i^*)$, the unique strategy with the highest payoff in the sample is x_i^*. Hence, there is a positive probability of at time $t+m$ obtaining a history $h^{t+m} = (x^*, ..., x^*)$, a convention. It follows that for $s/m \leq 1/2$ the only recurrent classes of the unperturbed process are the conventions. *Q.E.D.*

Lemma 1 below implies that Theorem 3.1 in Young (1998, p. 56) applies to imitation play. This theorem, with a slightly modified notation, is referred to as Theorem 2 in this paper.

Definition 1. *(Young, 1998)* $B^{m,s,\varepsilon}$ *is a* **regular perturbed Markov process** *if* $B^{m,s,\varepsilon}$ *is irreducible for every* $\varepsilon \in (0, \varepsilon^*]$, *and for every states* h, h', $B_{hh'}^{m,s,\varepsilon}$ *approaches* $B_{hh'}^{m,s,0}$ *at an exponential rate, i.e.* $\lim_{\varepsilon \to 0} B_{hh'}^{m,s,\varepsilon} = B_{hh'}^{m,s,0}$ *and if* $B_{hh'}^{m,s,\varepsilon} > 0$ *for some* $\varepsilon > 0$, *then* $0 < \lim_{\varepsilon \to 0} \frac{B_{hh'}^{m,s,\varepsilon}}{\varepsilon^{r_{hh'}}} < \infty$ *for some* $r_{hh'} \geq 0$. *The real number* $r_{hh'}$ *is called the* **resistance** *of the transition from* h *to* h'.

Lemma 1. *Imitation play is a regular perturbed Markov process.*

PROOF: See the Appendix.

Definition 2. *(Young, 1993) Let μ^ε be the unique stationary distribution of an irreducible process $B^{m,s,\varepsilon}$. A state h is* **stochastically stable** *if* $\lim_{\varepsilon\to 0}\mu^\varepsilon(h) > 0$.

Let an unperturbed Markov process $B^{m,s,0}$ with $s/m \leq 1/2$ have recurrent classes/ conventions $h_{x^1}, ..., h_{x^{|X|}}$ (according to Theorem 1 there is a one to one correspondence between the recurrent classes and the conventions for $s/m \leq 1/2$). For each pair of distinct recurrent classes, an *xy-path* is a sequence of states $\zeta = (h_x, ..., h_y)$ that begins in h_x and ends 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 xy-paths. Construct a complete directed graph with $|X|$ vertices, one for each recurrent class. The weights on the directed edge $h_x \to h_y$ 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* $\Im(x)$ is the sum of the resistances $r_{x'x''}$ on the $|X|-1$ edges that compose it. The *stochastic potential* $\rho(x)$ of a recurrent class h_x is the minimum resistance over all trees rooted at h_x.

Theorem 2. *(Young, 1998) Let $B^{m,s,\varepsilon}$ be a regular perturbed Markov process and let μ^ε be the unique stationary distribution of $B^{m,s,\varepsilon}$ for $\varepsilon > 0$. Then $\lim_{\varepsilon\to 0}\mu^\varepsilon = \mu^0$ exists, and is a stationary distribution of $B^{m,s,0}$. The stochastically stable states are precisely those states that are contained in the recurrent classes of $B^{m,s,0}$ having minimum stochastic potential.*

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 Figure 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.

	a	b	c
A	$1,1$	$0,0$	$0,0$
B	$0,0$	$3,2$	$0,3$
C	$0,0$	$0,3$	$3,2$

FIGURE 2

The conventions $h_{(B,b)}$ and $h_{(C,c)}$ are Pareto superior to convention $h_{(A,a)}$. 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 conventions, construct directed graphs with nine vertices, one for each convention. In Figure 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 conventions. It is easy to check that $\rho(A, a) = 8$, and $\rho(B, b) = \rho(B, c) = \rho(C, b) = \rho(C, c) = 9$. Hence convention $h_{(A,a)}$ is *stochastically stable.*

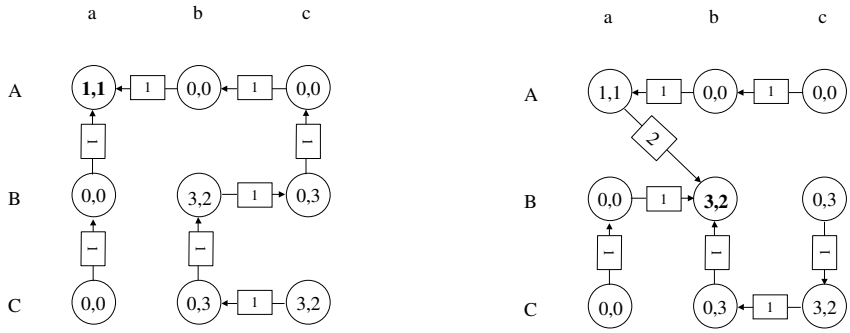

FIGURE 3–Minimum-resistance trees rooted at $h_{(A,a)}$ and $h_{(B,b)}$ respectively.

3.2. Sets of Stochastically Stable States. In this subsection we will show that the stochastically stable states of imitation play correspond to strategy-tuples in particular subsets of the strategy-space. In order to do this we need to introduce

some new concepts. Given n-player game Γ with finite strategy space $X = \prod X_i$, associate each strategy-tuple $x \in X$ with the vertex of a graph. Draw a directed edge from vertex x to vertex x' if and only if:

i) there exists exactly one player i such that $x_i \neq x_i'$, $u_i(x_i', x_{-i}) \geq u_i(x)$ and

ii) there does not exist $x_i'' \neq x_i$ and $x_i'' \neq x_i'$ such that

$u_i(x_i', x_{-i}) > u_i(x_i'', x_{-i}) > u_i(x)$.

Call the graph that is obtained in this manner the *better-reply graph* of game Γ.[6] A *better-reply path* is a sequence of the form $x^1, x^2, ..., x^l$ such that each pair (x^j, x^{j+1}) corresponds to a directed edge in the better-reply graph. A *sink* is a vertex with no outgoing edges. Clearly, x is a sink if and only if it is a strict Nash equilibrium. A better-reply graph can also contain directed cycles.

Figure 4 shows a better-reply graph for a game with two sinks - (C, a) and (B, b). The *basin* of a sink is the set of all vertices from which there exists a directed path ending at that sink. Note that a vertex may be in several basins simultaneously. For example, (B, a) is in the basin of (C, a) and also in the basin of (B, b).

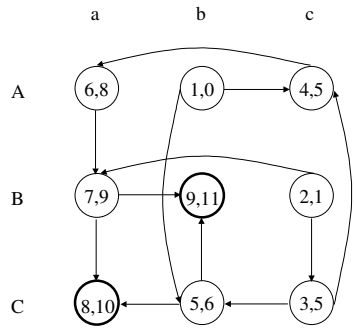

FIGURE 4–Better-reply graph.

[6]Note that condition ii) only is included for ease of exposition. The omission of this condition would complicate notation considerably without changing our results.

Lemma 2. *If x and x' are vertices of the better-reply graph of game Γ and there exists a directed edge from vertex x to vertex x', then the following inequality holds for the stochastic potentials $\rho(x)$ and $\rho(x')$ of the conventions h_x and $h_{x'}$ respectively:*

$$\rho(x) \geq \rho(x').$$

PROOF: Suppose the claim is false, such that $\rho(x) < \rho(x')$. Note that $\rho(x)$ is the minimum resistance over all trees rooted at the state h_x. Construct a new tree rooted at $h_{x'}$ by taking (one of) the tree(s) with minimum resistance rooted at h_x, adding the directed edge from vertex x to vertex x' and deleting the directed edge from x'. The resistance of the added edge is exactly one and of the deleted edge at least one, so the total resistance of the new tree is at most $\rho(x)$, which implies a contradiction. *Q.E.D.*

Definition 3. *A non-empty set of vertices V is said to be closed under the better-reply graph, or V is a* **construction,** *if there are no directed edges from any of the vertices in V to any vertex outside of V. Such a set is called a* **minimal construction** *if it does not properly contain another construction.*

From the definition follows that every game contains a minimal construction. Moreover, any sink is a minimal construction. The game in Figure 5 has two minimal constructions: a single sink (A, a) and a set $V = \{(B,b), (B,c), (C,c), (C,b)\}$, which can be considered as a single directed cycle $(B,b) \to (B,c) \to (C,c) \to (C,b) \to (B,b)$.

	a	b	c
A	$1,1$	$0,0$	$0,0$
B	$0,0$	$3,2$	$2,3$
C	$0,0$	$2,3$	$3,2$

FIGURE 5

Note that our definition of minimal constructions in normal form games corresponds to Sobel's (1993) definition of *non-equilibrium evolutionary stable (NES)* sets in 2-player games. Our concept is also similar to Nöldeke's and Samuelson's (1993) definition of *locally stable components* in extensive form games.

Lemma 3. *Let V be a minimal construction of an n-player game Γ with finite strategy space $X = \prod_{i=1}^{n} X_i$. Then, for any two vertices $x, x' \in V$, there exist better-reply paths $x, ..., x'$ and $x', ..., x$, which connect these vertices.*

PROOF: Suppose that the claim is false and there exist two vertices $x, x' \in V$ such that there is no better-reply path from x to x'. Consider all better-reply paths, which start at the vertex x. Since X is finite, there exist finitely many vertices along all these paths. Collect all these vertices. By construction this set of vertices has only ingoing edges and by assumption it does not contain the vertex x'. Hence, the constructed set of vertices is a minimal construction and a proper subset of the minimal construction V. This is a contradiction, so the claim must be true. *Q.E.D.*

To every vertex of a minimal construction corresponds a convention, which is a repetition of the associated strategy-tuple $x \in X$. Call the set of these conventions *conventions of a minimal construction.* We are now in a position to state the following main theorem.

Theorem 3. *If $s/m \leq 1/2$, s is sufficiently large and ε is sufficiently small, the perturbed process $B^{m,s,\varepsilon}$ puts arbitrarily high probability on the conventions of minimal construction(s) that minimize stochastic potential.*

PROOF: See the Appendix.

In order to prove this claim, we use Lemma 2 and Lemma 3 to show that: A) all conventions in a minimal construction have equal stochastic potential and B) for

every convention which does not belong to any minimal construction, there exists a convention with lower stochastic potential. The theorem establishes a relation between the stochastically stable states of imitation play and minimal constructions, which is similar to the relationship proved between the stochastically stable states of *adaptive play* and minimal curb sets in Theorem 7.2 of Young (1998, p. 111).

We say that a finite set Y of strategy-tuples *strictly Pareto dominates* a strategy-tuple x if for any strategy-tuple $y \in Y$ $\pi_i(y) > \pi_i(x)$, for all i. The following theorem shows that imitation play selects sets of efficient outcomes in a large class of games.

Theorem 4. *Suppose that there exists a minimal construction V, which is a product set and which strictly Pareto dominates all strategy-tuples outside of V. If $s/m \leq 1/2$, s is sufficiently large and ε is sufficiently small, the perturbed process $B^{m,s,\varepsilon}$ puts arbitrarily high probability on the conventions of V.*

PROOF: See the Appendix.

The intuition behind this result is that for sufficiently large sample size, the transition from a state inside of V to any state outside of V requires more mistakes than the number of player positions, while the opposite transition requires at most one mistake per player position. The following corollary follows immediately from Theorem 4.

Corollary 1. *If x is a strict Nash Equilibrium, which strictly Pareto dominates all other strategy-tuples, then for $s/m \leq 1/2$ and s sufficiently large, the convention h_x is a stochastically stable state.*

The requirement that V be a product set is necessary, as shown by the game in Figure 6.

	a	b	c
A	$1,1$	$0,0$	$0,0$
B	$0,0$	$0,0$	$5,3$
C	$0,0$	$2,4$	$5,4$

FIGURE 6

In this game, the minimal construction $V = \{(C,b),(C,c),(B,c)\}$ strictly Pareto dominates all strategy-tuples outside of V. It is evident that two mistakes are enough to move from convention $h_{(A,a)}$ to any convention in the minimal construction V. We will show that two mistakes are also enough to move from the minimal construction V to an outside state. Suppose the process is in the state $h_{(C,c)}$ at time t. Further suppose that the agent in player position 1 plays B instead of C at time $t+1$ by mistake. This results in play (B,c) at time $t+1$. Assume the agent in player position 2 makes a mistake and plays b instead of c, and that the agent in player position 1 plays C in period $t+2$. Hence, play at time $t+2$ is (C,b). Assume the agents in both player positions sample from period $t-s+2$ to period $t+2$ for the next s periods. This means that the agents in player position 1 chooses to play B and the agents in player position 2 chooses 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 agents in both player positions 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 convention $h_{(B,b)}$ outside of the minimal construction V. It is now straightforward to show that all of the conventions $h_{(A,a)}$, $h_{(C,b)}$, $h_{(C,c)}$ and $h_{(B,c)}$ have equal stochastic potential.

4. Properties of Minimal Constructions

In this section we investigate the relationship between minimal constructions and other set-wise concepts. We also analyze whether strictly dominated strategies can be included in minimal constructions and possible invariance under payoff transformations.

4.1. Relation to Minimal Sets Closed Under better replies. We start by the relation between minimal constructions and so called minimal sets closed under better replies. It is maybe not surprising that there is a quite strong connection between these two concepts.

Definition 4. *(Ritzberger and Weibull, 1995) Let the* **better-reply correspondence** $\gamma = \times_{i \in N} \gamma_i : \Delta \twoheadrightarrow X$ *be defined as follows*

$$\gamma_i(p) = \{x_i \in X_i \mid u_i(x_i, p_{-i}) \geq u_i(p)\}, \forall i \in N.$$

A product set of strategies is a set of form $Y = \Pi_i Y_i$, where each Y_i is a non-empty subset of X_i. Let $\Delta(Y_i)$ denote the set of probability distributions over Y_i, and let $\Delta(Y) = \Pi_i \Delta(Y_i)$ denote the product set of such distributions. Let $\gamma_i(Y)$ denote the set of strategies in X_i that are better replies by i to some distribution $p \in \Delta(Y)$ and let $\gamma(Y) = \Pi_i \gamma_i(Y)$.

Definition 5. *A non-empty product set $Y \subset X$ is said to be* **closed under better replies** *(or Y is a* **cubr** *set) if $\gamma(Y) \subset Y$. A set is called a* **minimal cubr** *set if it is cubr and contains no proper subset which is cubr.*

Proposition 1. *Every minimal cubr set contains a minimal construction. If a minimal construction is a product set, then it coincides with a minimal cubr set.*

PROOF: We start from the first claim. Let $Y = \Pi_i Y_i$ be a minimal cubr set. Associate each strategy-tuple $x \in X$ with the vertex of a graph and draw a better-reply graph for game Γ on the finite strategy space X. By the definition of a minimal cubr set, there are no directed edges from the set of vertices in Y to any vertices outside of Y in the better-reply graph. This means that the set of vertices in Y is a minimal construction or a construction containing a minimal construction.

The last claim follows immediately from the definitions of a minimal cubr set and a minimal construction. *Q.E.D.*

Combined with Theorem 3 this implies that if all the minimal constructions of a game are product sets, then imitation play converges to the conventions of minimal cubr sets. Note that every minimal construction not necessarily is included in a minimal cubr set. Consider the game in Figure 7.

	a	b	c
A	$2,2$	$1,1$	$0,0$
B	$1,1$	$0,0$	$3,3$
C	$0,0$	$3,3$	$3,3$

FIGURE 7

In this game, there is a unique minimal cubr set, $\{(A,a)\} = Y$, and two minimal constructions, $V_1 = \{(A,a)\}$ and $V_2 = \{(B,c),(C,b),(C,c)\}$. Hence V_2 is not contained in any minimal cubr set.

4.2. Relation to Curb Sets. The minimal curb set is another set-wise concept. Young (1998) show that there is a strong relation between minimal curb sets and stochastically stable states in his model. Let $Z = \Pi_{i=1}^{n} Z_i$ be a product set of strategies, and let $\beta_i(Z_{-i})$ denote the set of strategies in X_i that are best replies by i to some distribution $p_{-i} \in \Delta(Z_{-i})$. Define $\beta(Z) = \Pi_{i=1}^{n} \beta_i(Z_{-i})$.

Definition 6. *(Basu and Weibull, 1991) A non-empty product set $Z \subset X$ is said to be* **closed under rational behavior** *(or Z is a* **curb** *set) if $\beta(Z) \subset Z$. Such a set is called a* **minimal curb** *set if it does not properly contain a curb set.*

It is straightforward to show that $\beta(Z) = Z$ for any minimal curb set Z. Let the *span* of a non-empty set of vertices V, denoted by $S(V)$, be the product set of all strategies that appear in some strategy-tuple of V. By the following two examples, we will show that a minimal curb set and (the span of) a minimal construction are different set-wise concepts. In the game in Figure 8, the minimal cubr set, the minimal construction and the minimal curb set are all different.

	a	b	c
A	$1,3$	$3,2$	$3,1$
B	$3,2$	$2,3$	$2,1$
C	$2,3$	$1,1$	$1,2$

FIGURE 8

In this game, $\{A,B,C\} \times \{a,b,c\} = Y$ is a minimal cubr set. There is a unique minimal construction: $V = \{(A,a),(C,a),(B,a),(B,b),(A,b)\}$ with span $S(V) = \{A,B,C\} \times \{a,b\}$. It is obvious that $\{A,B\} \times \{a,b\} = Z$ is the minimal curb set. Hence $S(V) \cap Y = S(V)$ and $S(V) \cap Z = Z$ in this game.

One may ask whether the minimal curb set always is included in the span of a minimal construction. In other words, is the set-wise relationship $S(V) \cap Z = Z$ true in general? The answer is NO. Consider the game in Figure 9.

	a	b	c
A	$3,1$	$-10,0$	$1,3$
B	$-10,0$	$1,3$	$3,1$
C	$1,3$	$3,1$	$-10,0$
D	$0,0$	$0,0$	$0,0$

FIGURE 9

In this game, A is a best reply to a, a is a best reply to C, C is a best reply to b, b is a best reply to B, B is a best reply to c, c is a best reply to A. Hence any curb set that involves one or more of $\{A, a, C, b, B, c\}$ must involve all of them. Hence it must involve D, because D is a best reply to the probability mixture $1/3$ on a, $1/3$ on b,and $1/3$ on c. We conclude that $\{A, B, C, D\} \times \{a, b, c\} = Z$ is the unique minimal curb set. However, the unique minimal construction is $V = \{(A, a), (C, a), (C, b), (B, b), (B, c), (A, c)\}$ with span $S(V) = \{A, B, C\} \times \{a, b, c\}$, such that $S(V) \cap Z = S(V) \neq Z$.

4.3. Relation to Strictly Dominated Strategies. Our model assumes that agents are boundedly rational. In this subsection we ask whether any dominated strategies will be played in the stochastically stable states of imitation play. Consider a 2×2 game where strategy A strictly dominates strategy B for the row player. In such a game, the minimal construction will consist either of (A, a) or (A, b) or both of them. From Theorem 3 follows that in 2×2 games, a strictly dominated strategy will never be used with positive probability in the long run.

However, in games with more than two strategies, strictly dominated strategies may be played with positive probability. Consider the game in Figure 10.

	a	b
A	$0, 3$	$2, 0$
B	$2, 0$	$1, 3$
C	$1, 1$	$0, 1$

FIGURE 10

In this game, B strictly dominates C. However, the unique minimal construction is $V = \{(A, a), (C, a), (C, b), (B, b), (A, b), (B, a)\}$ and by Theorem 3, imitation play will put positive probability on all strategy-tuples in this minimal construction in the

long run.

Moreover, as illustrated by the game in Figure 11, a weakly dominated strategy may be used with positive probability in 2×2 games. In this game, the unique minimal construction is $V = \{(A, b), (B, b), (B, a)\}$.

	a	b
A	$0, 0$	$3, 3$
B	$3, 3$	$3, 3$

FIGURE 11

4.4. Payoff Transformations. Minimal constructions are invariant under affine payoff transformations and local payoff shifts. This follows since the better-reply graph is unaffected by such changes. The stochastically stable states of imitation play are also invariant under affine payoff transformations, but not under local payoff shifts. The two games in Figure 12 illustrate the latter claim.

	a	b
A	$3, 3$	$0, 0$
B	$0, 0$	$2, 2$

	a	b
A	$3, 3$	$2, 0$
B	$0, 2$	$4, 4$

FIGURE 12

In the left game, (A, a) is stochastically stable according to Theorem 4. By adding 2 units to the each of the row player's payoffs in column b and to each of the column player's payoff in row B in this game, the right game is obtained. In the right game (B, b) is stochastically stable by Theorem 4.

5. APPLICATIONS TO 2×2 GAMES

In this section we apply the results from Section 3 to 2×2 games. First, we find the stochastically stable states in three special classes of games and second, we study the

speed of convergence in general games. In all of the following games we denote player 1's strategies A and B, and player 2's strategies a and b respectively.

5.1. Stochastically Stable States in Special Classes of Games. In this subsection we analyze the stochastic stability in games with two strict Nash equilibria, games with one strict Nash equilibrium and games without Nash equilibria in pure strategies. We start with the class of games that have a unique strict Nash equilibrium. This class includes for example *Prisoners' Dilemma Games*.

Proposition 2. *In 2×2 games with a unique strict Nash equilibrium, if $s/m \leq 1/2$ and s is sufficiently large, the stochastically stable state correspond one to one with the convention of the unique strict Nash equilibrium.*

PROOF: Games in this class contain exactly one minimal construction, consisting of the unique strict Nash equilibrium. By Theorem 3 the corresponding convention must be the unique stochastically stable state. *Q.E.D.*

We now proceed with the class of games with two strict Nash equilibria. *Coordination Games* and *Hawk-Dove Games* are two examples of games in this class.

Proposition 3. *In 2×2 games with two strict Nash equilibria, where one Nash Equilibrium strictly Pareto dominates the other, if $s/m \leq 1/2$ and s is sufficiently large, then the unique stochastically stable state corresponds one to one with to the Pareto dominant equilibrium.*

PROOF: Games in this class contain two minimal constructions, either $\{(A, a)\}$ and $\{(B, b)\}$ or $\{(A, b)\}$ and $\{(B, a)\}$. Assume without loss of generality that $\{(A, a)\}$ and $\{(B, b)\}$ are the minimal constructions. If (A, a) strictly Pareto dominates (B, b), it must also strictly Pareto dominate the two other strategy-tuples (A, b) and (B, a).

Hence, by Corollary 1 the convention $h_{(A,a)}$ is the unique stochastically stable state for $s/m \leq 1/2$ and s is sufficiently large. *Q.E.D.*

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. In the right game in Figure 12 (A, a) is a risk dominant equilibrium whereas (B, b) is a strictly Pareto superior equilibrium and subsequently the unique stochastically stable state of imitation play. Our result is consistent with Robson's and Vega-Redondo's (1996) result for generic symmetric Coordination Games. However, for the non-generic case when $\pi_i(A, a) = \pi_i(B, b)$, the stochastically stable states in their model depend on the details of the adjustment process, whereas imitation play always selects both equilibria. Consider the following proposition.

Proposition 4. *In 2×2 games with two strict Nash equilibria, where neither strictly Pareto dominates the other, if $s/m \leq 1/2$ and s is sufficiently large, then the stochastically stable states correspond one to one with the two equilibria.*

PROOF: See the Appendix.

Finally, we consider games which do not have any Nash equilibrium in pure strategies. One of the games in this class is the *Matching Pennies Game*.

Proposition 5. *In 2×2 games without Nash equilibria in pure strategies, if $s/m \leq 1/2$ and s is sufficiently large, the stochastically stable states correspond one to one with the four conventions of the game.*

PROOF: Games in this class contain exactly one minimal construction $\{(A, a), (A, b), (B, a), (B, b)\}$. By Theorem 3 the four corresponding conventions must all be stochastically stable. *Q.E.D.*

5.2. Speed of Convergence. In this subsection we analyze the speed of convergence of imitation play.

Proposition 6. *In 2×2 games, if $s/m \leq 1/2$ and s is sufficiently large, the maximum expected first passage time for the perturbed process $B^{m,s,\varepsilon}$ from any state to a stochastically stable state is at most $\delta\varepsilon^{-2}$ units of time, for some positive constant δ.*

PROOF: The claim follows from the observation that in any 2×2 game, the transition from an arbitrary state to the basin of attraction of a stochastically stable convention requires at most two mistakes. *Q.E.D.*

This result should be contrasted with the speed of convergence in Young (1993), Kandori, Mailath and Rob (1993) and Robson and Vega-Redondo (1996). In Young's (1993) model the maximum expected first passage for a 2×2 Coordination Game is at most $\delta_Y \varepsilon^{-v}$ where v depends on the sample size and both players' payoffs. In Kandori, Mailath and Rob (1993) the maximum expected first passage time is of 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, if v, Nu and q all are greater than two and ε is sufficiently small, then imitation play converges considerably faster than the processes in these three models.

6. CONCLUSION

In this paper we develop an evolutionary model with perpetual random shocks where individuals every period choose the strategy with the highest average payoff in a finite sample of past play. We denote the resulting Markov chain imitation play. We prove that, provided information is sufficiently incomplete and the sample size sufficiently large, the stochastically stable states of imitation play are repetitions of

the strategy-tuples in minimal sets closed under the better-reply graph. We call such sets minimal constructions. These sets are related to minimal sets closed under better replies and other set-wise concepts. We also prove that if the strategy-tuples in a minimal construction have strictly higher payoffs than all outside strategy-tuples, then, provided the minimal construction is a product set, the strategy-tuples in this set will be played with probability one in the limit as the experimentation probability tends to zero. Our results give clear predictions in 2×2 games. In Coordination Games where one equilibrium strictly Pareto dominates the other, imitation play selects the strictly Pareto superior Nash equilibrium. If neither equilibria strictly Pareto dominates the other, then both are stochastically stable. Finally, we show that the speed of convergence for imitation play in many cases is higher than in other known models.

The objective of this paper is to derive predictions for general finite games in a world of truly boundedly rational agents. The assumption underlying the model, that individuals do not make decisions based on the predictions of their opponents' future strategies, but rather based on which strategies have been successful in the past, is maybe most appealing in the class of games where it is costly to obtain information about the opponents. A high cost may be due to the size or the complexity of the game or to institutional factors that prevent the release of information about the opponents. It would be particularly interesting to test the implications of our model against empirical or experimental evidence in this class of games.

Department of Economics, Stockholm School of Economics, P.O. Box 6501, S-113 83 Stockholm, Sweden

and

Department of Economics, Stockholm School of Economics, P.O. Box 6501,

S-113 83 Stockholm, Sweden

APPENDIX

PROOF OF LEMMA 1**:** This proof follows Young (1998). Given a history $h^t = (x^{t-m+1}, ..., x^t)$ at time t, the process moves in the next period to a state of the form $h^{t+1} = (x^{t-m+2}, ..., x^t, x^{t+1})$. Remember that such a state is called a *successor* of h^t. The strategy x_i^{t+1} is an *idiosyncratic choice* or *error* if and only if there exist no sample of strategy choices w_i and payoff outcomes v_i, such that $x_i^{t+1} \in \alpha_i(w_i, v_i)$. For each successor h^{t+1} of h^t, let $r_{h^t \to h^{t+1}}$ denote the total number of errors in the transition from h^t to h^{t+1}. Evidently $0 \leq r_{h^t \to h^{t+1}} \leq n$ and the probability of the transition $h^t \to h^{t+1}$ is on the order of $\varepsilon^{r_{h^t \to h^{t+1}}}(1-\varepsilon)^{n-r_{h^t \to h^{t+1}}}$. If h^{t+1} is not a successor of h^t, the probability of the transition is zero. Hence, the process $B^{m,s,\varepsilon}$ approaches $B^{m,s,0}$ at an exponential rate in ε. Since the process is also irreducible whenever $\varepsilon > 0$, it follows that $B^{m,s,\varepsilon}$ is a regular perturbed Markov process. *Q.E.D.*

PROOF OF THEOREM 3: In order to prove the claim, we will show that:
A) all conventions in a minimal construction have equal stochastic potential;
B) for every convention which does not belong to any minimal construction, there exists a convention with lower stochastic potential.
The theorem follows immediately from A) and B), by applying Lemma 1 and Theorem 2.

Let us start from A). On the one hand, by Lemma 3, there exist a better-reply path from an arbitrarily convention h_x in a minimal construction to any other convention $h_{x'}$ in the same minimal construction. Let the sequence $h_x, ..., h_{x'}$ be such a path. By Lemma 2, the following inequalities hold:

$$\rho(x) \geq ... \geq \rho(x'). \tag{2}$$

On the other hand, by applying Lemma 3 once again, there exist a better-reply path from the convention $h_{x'}$ to the convention h_x. Using Lemma 2, gives:

$$\rho(x') \geq ... \geq \rho(x). \tag{3}$$

From the inequalities in (2) and (3) follow that $\rho(x) = \rho(x')$ for any arbitrarily conventions h_x and $h_{x'}$ in a minimal construction.

Part B). Note that for every convention not included in any minimal construction, there exists a finite better-reply path which ends in some minimal construction. Let this path be $h_{x^1}, h_{x^2}, ..., h_{x^{T-1}}, h_{x^T}$, where h_{x^1} is an arbitrarily convention that does not belong to any minimal construction and h_{x^T} is the first convention on the path that belongs to some minimal construction V. By Lemma 2, it follows that

$$\rho\left(x^1\right) \geq \rho\left(x^2\right) \geq ... \geq \rho\left(x^{T-1}\right) \geq \rho\left(x^T\right). \tag{4}$$

We will show that in fact $\rho\left(x^{T-1}\right) > \rho\left(x^T\right)$. Note that $\rho\left(x^{T-1}\right)$ 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 $\Im\left(x^{T-1}\right)$. Find in the tree $\Im\left(x^{T-1}\right)$ a directed edge from some vertex $y \in X$ in the minimal construction V to some other vertex $y' \in X$ outside of this minimal construction. It will be shown later that there is only one such directed edge in the minimal resistance tree $\Im\left(x^{T-1}\right)$. Delete in the tree $\Im\left(x^{T-1}\right)$ the directed edge $y \to y'$ and add the directed edge $x^{T-1} \to x^T$. As a result we obtain a tree $\Im(y)$ rooted at the state h_y. By construction, the total resistance of the new tree $\Im(y)$ is less than the stochastic potential $\rho\left(x^{T-1}\right)$. Moreover, by part A), the convention h_{x^T} of the minimal construction has the same stochastic potential as the convention h_y. Hence $\rho\left(x^{T-1}\right) > \rho\left(x^T\right)$.

We will now consider the tree $\Im\left(x^{T-1}\right)$ and show that there is only one directed edge from every minimal construction to a state outside of the construction. Suppose

there are finitely many such directed edges $y^j \to z^j$, $j = 1, 2, ..., l$ from some minimal construction, where $y^1, ..., y^l$ are vertices in the minimal construction and $z^1, ..., z^l$ are vertices outside of the construction. It is clear that there cannot be an infinite number of outgoing edges since the game Γ is finite. Recall that a rooted at vertex y^j tree is a set of $|X| - 1$ directed edges such that, from every vertex different from y^j, there is a unique directed path in the tree to y^j. The resistance of any directed edge $y^j \to z^j$, $j = 1, 2, ..., l$ 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 construction. Let $y^1, f^1, ..., f^k, y^2$ be such a path.

Consider vertex f^1. There are two mutually exclusive cases:

1.a) there exists a directed path from f^1 to one of the vertices $y^2, ..., y^l$ in the initial tree $\Im\left(x^{T-1}\right)$, or

1.b) there exists a directed path from f^1 to y^1.

In case 1.a) by deleting the directed edge $y^1 \to z^1$ and adding the directed edge $y^1 \to f^1$ to the tree $\Im\left(x^{T-1}\right)$, we obtain a new tree $\Im^1\left(x^{T-1}\right)$ with lower stochastic potential than $\Im\left(x^{T-1}\right)$, 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 $\Im\left(x^{T-1}\right)$ being a minimal resistance tree.

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

2.a) there exists a directed path from f^2 to one of the vertices $y^2, ..., y^l$ in the initial tree $\Im\left(x^{T-1}\right)$, or

2.b) there exists a directed path from f^2 to y^1.

In case 2.a) we obtain a new tree $\Im^2\left(x^{T-1}\right)$ with lower stochastic potential than $\Im\left(x^{T-1}\right)$, because the resistance of the directed edge $f^1 \to f^2$ is one. This means

that we are done, since it contradicts the assumption of $\Im\left(x^{T-1}\right)$ being a minimal resistance tree.

In case 2.b) we repeat the procedure for the vertex $f^2, f^3, ...$ The better-reply path $y^1, f^1, ..., f^k, y^2$ from vertex y^1 to vertex y^2 is finite. Hence, after at most $k+1$ steps we have constructed a tree $\Im''\left(x^{T-1}\right)$ rooted at the state $h_{x^{T-1}}$ with lower stochastic potential than $\Im\left(x^{T-1}\right)$, which is impossible. *Q.E.D.*

Proof of Theorem 4: Define the *basin of attraction* of a state h' as the set of states h such that there is a positive probability of moving in a finite number of periods from h to h' under the unperturbed process $B^{m,s,0}$. Let V be a minimal construction which strictly Pareto dominates all strategy-tuples outside of V. Let h_y be a convention such that $y \notin V$. Consider a minimal resistance tree $\Im(y)$ rooted at h_y. Let $x \in V$ be a vertex such that there is a directed edge from x to a vertex $z \notin V$ in the tree $\Im(y)$. Assume that the stochastic potential of h_x is at least as high as the stochastic potential of h_y.

We claim that for sufficiently large s, the resistance of the edge from $x \in V$ to $z \notin V$ is greater than n. Assume without loss of generality that $s > n$ (otherwise s consecutive mistakes by agents in one population could be sufficient to transfer from h_x to the basin of attraction of h_z). Let $\pi_i(x)$ and $\pi_i(z)$ be payoffs for player i corresponding to the strategy-tuples x and z respectively. The individual in player position i will play z_i instead of x_i only if the average payoff for z_i in her sample is greater than or equal to the average payoff for x_i. With at most n mistakes this is only possible if the following inequality holds

$$\frac{(s-n)\pi_i(x) + (n-1)\pi_i(x_i, x'_{-i})}{s-1} \leq \pi_i(z), \tag{5}$$

where $\pi_i(x_i, x'_{-i})$ is the payoff to player i when she plays x_i and her opponents play any $x'_{-i} \neq x_{-i}$. It is clear that this inequality does not hold for any individual in

player position i if s is sufficiently large. Hence, the resistance from h_x to h_z must be larger than n for sufficiently large s.

Create a tree $\Im(x)$ rooted at h_x by adding a directed edge from h_y to h_x and deleting the directed edge from x to z in the tree $\Im(y)$. Provided that s is sufficiently large, the deleted edge has a resistance greater than n and the added edge has a resistance of at most n (n simultaneous mistakes are sufficient to move from any convention to the basin of attraction of h_x). Hence the total resistance of the new tree $\Im(x)$ is less than $\rho(y)$, contradicting the assumption of $\rho(y) \leq \rho(x)$. This establishes that h_x has minimal stochastic potential. Theorem 4 now follows from part A) in the proof of Theorem 3 where it is proved that all the conventions of a minimal construction have equal stochastic potential. *Q.E.D.*

PROOF OF PROPOSITION 3: Like in the proof of Proposition 2, assume without loss of generality that $\{(A,a)\}$ and $\{(B,b)\}$ are the minimal constructions of the game. By Theorem 3 follows that the conventions $h_{(A,a)}$ and $h_{(B,b)}$ are the only two candidates for the stochastically stable states. Suppose that only one of these conventions is stochastically stable, say $h_{(A,a)}$. Let $\Im(A,a)$ be a minimum resistance tree with resistance $\rho(A,a)$ rooted at (A,a). In this tree there is an outgoing edge from the convention $h_{(B,b)}$.

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

Second, note that provided that $s/m \leq 1/2$ and s is sufficiently large, two mistakes are sufficient to move the process from the convention $h_{(A,a)}$ to the convention $h_{(B,b)}$. Suppose the process is in state $h_{(A,a)}$. Since neither of the Nash equilibria is strictly Pareto superior, either $\pi_1(B,b) \geq \pi_1(A,a)$ and/or $\pi_2(B,b) \geq \pi_2(A,a)$. Assume without loss of generality that the first of these inequalities holds. Then there

is a positive probability that the agents in both player positions simultaneously make mistakes at time t. There is also a positive probability that the agents in player position 1 draw the fixed sample $(x_1^{t-s+1}, ..., x_1^t)$ with corresponding payoffs $(\pi_1^{t-s+1}, ..., \pi_1^t)$ and that the agents in player position 2 sample from plays earlier than x_2^t from period $t+1$ to and inclusive period $t+s-1$. With positive probability the agents in player position 1 play B and the agents in player position 2 play a in all of these periods. This implies that if the agents in both player positions sample from plays more recent than x_i^{t-1} from period $t+s$ to and inclusive period $t+s-1+m$, 1's sample will only contain strategy B and 2's sample will always contain strategy b and possibly strategy a. Furthermore, the average payoff of strategy b will be $\pi_2(B,b)$ as compared to an average payoff of $\pi_2(B,a)$ for strategy a when the latter strategy is included in the sample. Hence, with positive probability the process will be in state $h_{(B,b)}$ at time $t+s-1+m$.

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

References

[1] Basu, K., and J. W. Weibull (1991): "Strategy Subsets Closed under Rational Behavior", *Economics Letters*, 36, 141-146.

[2] Kandori, M., G. Mailath, and R. Rob (1993): "Learning, mutation and long-run equilibria in games", *Econometrica* 61, 29-56.

[3] Nash, J. (1950): "Non Cooperative Games", Ph.D. thesis, Princeton University.

[4] Nöldeke, G., and L. Samuelson (1993): "An Evolutionary Analysis of Backward and Forward Induction", *Games and Economic Behavior* 5, 425-454.

[5] Osborne, M. J., and A. Rubinstein (1998): "Procedurally rational players", *American Economic Review* 88, 834-847.

[6] Ritzberger, K., and J. Weibull (1995): "Evolutionary selection in normal-form games", *Econometrica* 63, 1371-1400.

[7] Robson, A. J., and F. Vega-Redondo (1996): "Efficient Equilibrium Selection in Evolutionary Games with Random Matching", *Journal of Economic Theory* 70, 65-92.

[8] Sobel, J. (1993): "Evolutionary stability and efficiency", *Economics Letters*, 42, 301-312.

[9] Young, P. H. (1993): "The evolution of conventions", *Econometrica* 61, 57-84.

[10] ——— (1998): *Individual Strategy and Social Structure*. Princeton University Press.