

# Stochastic Adaptation in Finite Games Played by Heterogeneous Populations

Jens Josephson\*

Stockholm School of Economics

SSE/EFI Working Paper Series in Economics and Finance No. 475

November 15, 2001

## Abstract

In this paper, I analyze stochastic adaptation in finite  $n$ -player games played by heterogeneous populations of myopic best repliers, better repliers and imitators. In each period, one individual from each of  $n$  populations, one for each player role, is drawn to play and chooses a pure strategy according to her personal learning rule after observing a sample from a finite history. With a small probability individuals also make a mistake and play a pure strategy at random. I prove that, for a sufficiently low ratio between the sample and history size, only pure-strategy profiles in certain minimal closed sets under better replies will be played with positive probability in the limit, as the probability of mistakes tends to zero. If, in addition, the strategy profiles in one such set have strictly higher payoffs than all other strategy profiles and the sample size is sufficiently large, then the strategies in this set will be played with probability one in the limit. Applied to  $2 \times 2$  Coordination Games, the Pareto dominant equilibrium is selected for a sufficiently large sample size, but in all symmetric and many asymmetric games, the risk dominant equilibrium is selected for a sufficiently small sample size.

Keywords: Bounded rationality, Evolutionary game theory, Imitation, Better replies, Markov chain, Stochastic stability, Pareto dominance, Risk dominance.

JEL classification: C72, C73.

---

\*I thank Alexander Matros and Jörgen W. Weibull for helpful comments. Financial support from Bankforskningsinstitutet is gratefully acknowledged. This paper is available at <http://swopec.hhs.se/hastef/abs/hastef0475.htm>.

## 1 Introduction

There is an extensive literature on learning in games, which investigates the long-run outcomes when boundedly rational individuals use simple rules of adaptation, or *learning rules*, to play games.<sup>1</sup> A common assumption in this literature is that all individuals employ the same learning rule. In this paper, I depart from this, in many cases unrealistic, assumption and allow for several different learning rules in each population.

More specifically, I analyze stochastic learning in finite  $n$ -player games with heterogeneous populations of myopic *best repliers*, *better repliers* and *imitators*. The best repliers observe a sample from a finite history of their opponents' past play, calculate an empirical distribution, and choose a best reply to this distribution. The better repliers observe a sample from a finite history of their opponents' and their own population's past play, and choose any pure strategy which gives at least as high expected payoff against the empirical distribution of the opponents' play. This behavior can be seen as a special case of Simon's (1955) satisfying behavior. Imitators, finally, observe a sample of their own population's past play and payoffs and either choose a pure strategy with maximum average payoff, or a pure strategy with maximum payoff.<sup>2</sup>

I prove that for any finite game, if the ratio between the sample and history size is sufficiently small, the resulting unperturbed Markov chain converges with probability one to a *minimal closed set under better replies* of the game. This result is independent of population shares, as long as the share of better repliers is positive, and implies convergence to a strict Nash equilibrium in weakly acyclical games. It is also consistent with previous results in a different framework. Ritzberger and Weibull (1995) show that for a large class of

---

<sup>1</sup>See Fudenberg and Levine (1998) for an introduction.

<sup>2</sup>For an analysis of stochastic adaptation by homogeneous populations of better repliers, see Josephson (2000), and by homogeneous populations of imitators, see Josephson and Matros (2000).

deterministic selection dynamics in continuous time, including the replicator dynamics, a product set is asymptotically stable if and only if it is a minimal closed set under better replies.

I then perturb the stochastic process by assuming that with a small probability, the individuals make mistakes or experiment and play a pure strategy at random according to some fixed probability distribution with full support. A mathematically equivalent interpretation is that each population contains a small share of noise players who always choose a pure strategy at random. This assumption makes the process irreducible and aperiodic, and thus implies a unique stationary distribution. I calculate the support of this distribution as the probability of mistakes tends to zero. If the sample size, and the information incompleteness is sufficiently large, and if one minimal closed set under better replies strictly Pareto dominates all strategy-tuples outside the set, then the perturbed Markov chain puts probability one on this set as the level of noise tends to zero. A corollary of this is that the strictly Pareto-dominant equilibrium is selected in games of common interest and  $2 \times 2$  Coordination Games. However, the minimum sample size required for this result depends on the payoffs of the game. In fact, for sample sizes below a certain critical level, the *risk-dominant* equilibrium is always selected in all symmetric Coordination Games, and also in many asymmetric Coordination Games, a finding which seems to be consistent with experimental evidence (see Van Huyck, 1997).

The basic setting in this paper is similar to that in Young (1993a,1998), who analyzes the dynamics of a homogenous population of myopic best repliers. The present paper is also related to earlier papers on stochastic learning with heterogeneous populations, but it differs in that the analysis is applied to general finite games, and not only special classes of such games. Kaniovski, Kryazhimiskii, and Young (2000) study adaptive dynamics in  $2 \times 2$  Coordination Games played by heterogeneous populations of myopic best repliers, conformists (who do what the majority does), and nonconformists (who do

the opposite of what the majority does). They show that the resulting process may have limit cycles even when the proportion of non-best repliers is arbitrary small. Kaarbøe and Tieman (1999) use the Kandori, Mailath, and Rob (1993) framework to study strictly supermodular games played by best repliers and imitators. In their model, the Pareto efficient equilibrium is selected. In Weibull and Sáez Marti (1998), Young's (1993b) evolutionary version of Nash's demand game is played by one population of myopic best repliers only, and another population of best repliers and clever agents who play a best reply to the best reply. The authors show that Young's predictions are still valid in the presence of such clever agents. Matros (2000) extends this result to finite two-player games.

This paper is organized as follows. I start by describing the model in Section 2. In Section 3, I present the general results and in Section 4, the results for  $2 \times 2$  Coordination Games. Section 5 concludes. Omitted proofs can be found in the Appendix.

## 2 Model

The basic setting is similar to that of Young (1993a, 1998), although my notation is slightly different. Let  $\Gamma$  be a finite  $n$ -player game in strategic form. Let  $X_i$  be the finite set of pure strategies  $x_i$  available to player  $i \in \{1, \dots, n\} = N$  and let  $\Delta(X_i)$  be the set of probability distributions  $p_i$  over these pure strategies. Define the product sets  $X = \prod X_i$  and  $\square(X) = \prod_i \Delta(X_i)$  with typical elements  $x$  and  $p$ , respectively. Let  $p_i(x_i)$  denote the probability mass on pure strategy  $x_i$  and let  $p(x) = \prod_{i \in N} p_i(x_i)$ . I write  $x_{-i} \in \prod_{j \neq i} X_j = X_{-i}$  and  $p_{-i} \in \prod_{j \neq i} \Delta(X_j) = \square(X_{-i})$  to represent the pure strategies and distributions of the pure strategies of player  $i$ 's opponents.

Let  $C_1, \dots, C_n$  be  $n$  finite and non-empty populations, each consisting of a share  $\delta_i^\alpha > 0$  of *imitators*, a share  $\delta_i^\beta > 0$  of *best repliers*, and a share  $\delta_i^\gamma > 0$  of *better repliers* (all to be defined below). The populations need not be of equal size, nor do they necessarily have equal shares of learning rules. Each

member of population  $C_i$  is a candidate to play role  $i$  in the game  $G$  and has payoffs represented by the utility function  $\pi_i : X \rightarrow \mathbb{R}$ , and expected payoffs represented by the function  $u_i : \square(X) \rightarrow \mathbb{R}$ . In slight abuse of notation, I write  $u_i(x_i, p_{-i})$  instead of  $u_i(p_i, p_{-i})$  if  $p_i(x_i) = 1$ .

Let  $t = 1, 2, \dots$  denote successive time periods. The stage game  $\Gamma$  is played once in each period. In period  $t$ , one individual is drawn at random from each of the  $n$  populations and assigned to play the corresponding role. The individual in role  $i$  chooses a pure strategy  $x_i^t$  from her strategy space  $X_i$ , according to her personal decision rule (to be defined below). The pure-strategy profile  $x^t = (x_1^t, \dots, x_n^t)$  is recorded and referred to as *play* at time  $t$ . The *history* or *state* at time  $t \geq m$  is the sequence  $h^t = (x^{t-m+1}, \dots, x^t)$ , where  $m$  is the *memory size* of all individuals. Let  $H = X^m$  be the finite set histories of length  $m$  and let  $h$  be an arbitrary element of this set. Assume an arbitrary initial history  $h^m = (x^1, \dots, x^m)$  at time  $m$ .

Strategies are chosen as follows. In every period  $t > m$ , each individual drawn to play the game inspects a sample of size  $s$  drawn without replacement from the most recent  $m$  periods.<sup>3</sup> The draws are independent for the various individuals and across time. If an imitator is drawn, she observes a sample of size  $s$  of population-specific strategy and payoff realizations and chooses a pure strategy which is included in the sample, and has maximum average payoff or, alternatively, maximum empirical payoff. If a best replier is drawn, she inspects a sample of size  $s$ , consisting of  $n - 1$  independent draws of her opponents' strategy realizations and calculates an empirical distribution,  $\hat{p}_{-i}$ . she then chooses a pure strategy which is a best reply to this distribution. If a better replier is drawn, she inspects a sample of size  $s$ , consisting of  $n$  independent draws, of all players' strategy realizations and calculates an empirical distribution,  $\hat{p}$ . Thereafter, she chooses a pure strategy which gives at least as high expected payoff against the empirical distribution of the opponents'

---

<sup>3</sup>In fact, the sample size may differ among individuals, as long as the *minimum* sample size in each share of the population is equal to  $s$ .

strategies  $\hat{p}_{-i}$  as  $\hat{p}_i$ . Formally, she chooses a pure strategy  $x_i \in \gamma_i(\hat{p})$ , where  $\gamma_i$  is the better-reply correspondence, defined by

$$\gamma_i(p) = \{x_i \in X_i \mid u_i(x_i, p_{-i}) \geq u_i(p)\}. \quad (1)$$

For each  $x_i \in X_i$ , let  $p_i(x_i \mid h)$  be the conditional probability that the individual in role  $i$  chooses pure strategy  $x_i$ , given history  $h$ . I assume that  $p_i(x_i \mid h)$  is independent of  $t$ , and that  $p_i(x_i \mid h) > 0$  if and only if at least one of the following conditions hold:

- i) There exists a sample of size  $s$  from player  $i$ 's  $m$  most recent strategy and payoff realizations, which includes  $x_i$ , and where  $x_i$  has the maximum average realized payoff.
- ii) There exists a sample of size  $s$  from player  $i$ 's  $m$  most recent strategy and payoff realizations, which includes  $x_i$ , and where  $x_i$  has the maximum realized payoff.
- iii) There exists a sample of size  $s$  from history  $h$ , consisting of  $n - 1$  independent draws and with a sample distribution of  $\hat{p}_{-i} \in \square(X_{-i})$ , such that  $x_i \in \beta_i(\hat{p}_{-i})$ , where  $\beta_i$  is the best-reply correspondence.
- iv) There exists a sample of size  $s$  from the history  $h$ , consisting of  $n$  independent draws and with a sample distribution of  $\hat{p} \in \square(X)$ , such that  $x_i \in \gamma_i(\hat{p})$ , where  $\gamma_i$  is the better-reply correspondence defined above.

Given a history  $h^t = (x^{t-m+1}, \dots, x^t)$  at time  $t$ , the process moves to a state of the form  $h^{t+1} = (x^{t-m+2}, \dots, x^t, x^{t+1})$  in the next period. Such a state is called a *successor* of  $h^t$ . My behavioral assumptions imply that the process moves from a current state  $h$  to a successor state  $h'$  in each period, according to the following transition rule. If  $x$  is the rightmost element of  $h$ , the probability of moving from  $h$  to  $h'$  is  $P_{hh'}^{*,m,s,0} = \prod_{i=1}^n p_i(x_i \mid h)$  if  $h'$  is a successor of  $h$  and 0 if  $h'$  is not a successor of  $h$ . This defines a finite Markov chain on the finite state space of histories  $H$ . I denote the process  $P^{*,m,s,0}$  *heterogeneous adaptive play with memory  $m$  and sample size  $s$* , but often refer to it as the *unperturbed process*.

Several concepts from Markov chain theory are used below. A *recurrent*

class  $E_k$  of the process  $P^{*,m,s,0}$  is a set of states such that there is zero probability of moving from any state in the class to any state outside, and a positive probability of moving from any state in the class to any other state in the class in a finite number of periods. A state  $h$  is *absorbing* if it constitutes a singleton recurrent class. The *basin of attraction* of a state  $h'$  is the set of states  $h$  such that there is a positive probability of moving from  $h$  to  $h'$  in a finite number of periods under the unperturbed process  $P^{*,m,s,0}$ .

I also define a perturbed process in a manner completely analogous to Young (1993a, 1998). Formally, in each period there is some small probability  $\varepsilon$  that the individual in role  $i$  experiments or makes a mistake and chooses a pure strategy at random from  $X_i$ , instead of according to her decision rule. The event that the individual in role  $i$  experiments is assumed to be independent of the event that the individual in role  $j$  experiments for every  $j \neq i$ . For every role  $i$ , let  $q_i(x_i | h)$  be the conditional probability that the individual in role  $i$  chooses  $x_i \in X_i$ , given that she experiments and the process is in state  $h$ . I assume that  $q_i(x_i | h)$  is independent of  $t$  and decision rules, and that  $q_i(x_i | h) > 0$  for all  $x_i \in X_i$ . Suppose that the process is in state  $h$  at time  $t$ . Let  $J$  be a subset of  $j$  players. The probability is  $\varepsilon^j(1 - \varepsilon)^{n-j}$  that exactly the players in  $J$  experiment and the others do not. Conditional on this event, the transition probability of moving from  $h$  to  $h'$  is  $Q_{hh'}^J = \prod_{i \in J} q_i(x_i | h) \prod_{i \notin J} p_i(x_i | h)$  if  $h'$  is a successor of  $h$  and  $x$  is the rightmost element of  $h'$  and 0 if  $h'$  is not a successor of  $h$ . This gives the following transition probability of the perturbed Markov process:

$$P_{hh'}^{*,m,s,\varepsilon} = (1 - \varepsilon)^n P_{hh'}^{*,m,s,0} + \sum_{J \subset N, J \neq \emptyset} \varepsilon^{|J|} (1 - \varepsilon)^{n-|J|} Q_{hh'}^J.$$

I call the process  $P^{*,m,s,\varepsilon}$  *heterogeneous adaptive play with memory  $m$ , sample size  $s$ , experimentation probability  $\varepsilon$  and experimentation distributions  $q_i$* .

This process is irreducible and aperiodic and thus, has a unique stationary distribution  $\mu^\varepsilon$ . I study this distribution as  $\varepsilon$  tends to zero. In my analysis, I use the following standard definitions, due to Freidlin and Wentzell (1984)

and Foster and Young (1990). A state  $h$  is *stochastically stable* if  $\lim_{\varepsilon \downarrow 0} \mu^\varepsilon(h)$  exists and is positive. For each pair of distinct recurrent classes,  $E_k$  and  $E_l$ , a *kl-path* is a sequence of states  $\zeta = (h^1, h^2, \dots, h^q)$  beginning in  $E_k$  and ending in  $E_l$ . The *resistance* of this path is the sum of the resistances on the edges that compose it. Let  $r_{kl}$  be the least resistance over all *kl-paths*. Construct a complete directed graph with one vertex for each recurrent class. The weight on the directed edge  $E_k \rightarrow E_l$  is  $r_{kl}$ . A *tree* rooted at  $E_l$  is a set of directed edges such that, from every vertex different from  $E_l$ , there is a unique directed path in the tree to  $E_l$ . The *resistance* of a rooted tree is the sum of the resistances on the edges that compose it. The *stochastic potential*  $\rho(E_l)$  of a recurrent class  $E_l$  is the minimum resistance over all trees rooted at  $E_l$ .

### 3 General Results

In this section, I will present two general results on the asymptotic distribution of heterogeneous adaptive play: one theorem for all finite  $n$ -player games, and one theorem for finite  $n$ -player games that can be Pareto ranked in a special sense. In order to state these theorems, some further definitions are needed. Let  $\mathcal{X}$  be the collection of all non-empty product sets  $Y \subset X$ . Let  $\Delta(Y_i)$  be the set of probability distribution with support in  $Y_i$  and let  $\square(Y) = \prod_{i \in N} \Delta(Y_i)$  be the corresponding product set.

**Definition 1** (*Ritzberger and Weibull, 1995*) *A set  $Y \in X$  is **closed under better replies (CUBR)** if  $\gamma(\square(Y)) \subseteq Y$ . A set  $Y \in X$  is a **minimal closed set under better replies (MCUBR)** if it is closed under better replies and contains no proper subset with this property.*

Let  $H' \subseteq H$  be an arbitrary set of histories. Define the *span* of  $H'$ ,  $S(H')$ , as the product set of all pure strategies appearing in some history of  $H'$  and let  $H^Y$  be the set of all histories  $h$  such that  $S(\{h\}) \subseteq Y$ . I say that a set of histories  $H'$  is *forward invariant* if the Markov chain remains forever in  $H'$ ,



once it has reached a state in this set.

**Lemma 1** *For any MCUBR set  $Y$ ,  $H^Y$  is forward invariant under  $P^{*,m,s,0}$ .*

PROOF: It is evident that imitation of the best pure strategy in a sample containing only pure strategies in an MCUBR set cannot result in the play of a pure strategy outside the MCUBR set. Moreover, by definition, a best or better reply to a probability distribution on an MCUBR set is always contained in the MCUBR set. Hence, if the unperturbed Markov chain  $P^{*,m,s,0}$  is in a state consisting only of pure-strategy profiles in a particular MCUBR set, then a pure strategy outside the MCUBR set will thereafter never be played by any of the individuals drawn to play. *Q.E.D.*

In order to state the main theorem, one more definition is needed. I say that a set of states  $H'$  is a *minimal  $\gamma$ -configuration* if  $S(H')$  is an MCUBR set. In other words,  $H'$  is a minimal  $\gamma$ -configuration if all the strategies in some MCUBR set  $Y$ , and no other strategies, appear in the histories of  $H'$ .

**Theorem 1** *Let  $\Gamma$  be a finite  $n$ -player game.*

- i) If  $s/m$  is sufficiently small, the unperturbed process  $P^{*,m,s,0}$  converges with probability one to a minimal  $\gamma$ -configuration.*
- ii) If, in addition,  $\varepsilon$  is sufficiently small, the perturbed process  $P^{*,m,s,\varepsilon}$  puts an arbitrarily high probability on the minimal  $\gamma$ -configuration(s) with minimum stochastic potential.*

PROOF: See the Appendix.

This theorem is analogous to Theorem 1 in Josephson (2000), which states that in a homogeneous setting, where the share of better repliers in each population is one, for a sufficiently large sample size and sufficiently incomplete information ( $s/m \leq \frac{1}{|\bar{X}|}$ ), the unperturbed process converges with probability one to a minimal  $\gamma$ -configuration. The proof of Theorem 1 is based on this result combined with Lemma 1, and the observation that in each period, there

is a positive probability that only better repliers will be drawn to play from the heterogeneous populations.

Theorem 1 is also consistent with previous results for deterministic continuous time selection dynamics. Ritzberger and Weibull (1995) show that a set is MCURB if and only if it is asymptotically stable for regular selection dynamics that are *sign preserving*. This is a large class of selection dynamics, which contain several well know dynamics, such as the replicator dynamics.

It is worth noting that Theorem 1 makes no reference to population mixtures. As long as the share of better repliers is positive in all populations, the span of the recurrent sets of the unperturbed process is independent of the shares of different learning rules in the populations. However, it is clear that this does not imply that the exact shape of the asymptotic distribution is independent of the population mixtures.

One may also note that in the class of games where all minimal closed sets under best replies and MCUBR sets coincide, it is sufficient if one of the population shares of better and best repliers is positive in each of the populations. This class of games contains, for instance,  $2 \times 2$  Coordination Games.

In the special class of games where all MCUBR sets are singleton, and hence also strict Nash equilibria, Theorem 1 implies convergence to a convention, a monomorphic state which is a repetition of a strict Nash equilibrium. *Weakly acyclical* games (Young, 1993a, 1998), and games that are *weakly acyclical in  $\gamma$*  (Josephson, 2000) belong to this class.

The second result of this paper characterizes the stochastically stable states of heterogeneous play in finite  $n$ -player games with a Pareto optimal outcome.

**Definition 2** *A non-empty set of strategy-tuples  $M \subset X$  **strictly Pareto dominates** a pure-strategy profile  $y \in X \setminus M$  if, for all  $i \in N$ ,*

$$\min_{x \in M} \pi_i(x) > \pi_i(y). \quad (2)$$

**Theorem 2** *Let  $\Gamma$  be a finite  $n$ -player game with an MCUBR set  $Y$ , which strictly Pareto dominates all other pure-strategy profiles. If  $s/m \leq 1/|X|$ ,  $s$  sufficiently large, and  $\varepsilon$  is sufficiently small, the perturbed process  $P^{*,m,s,\varepsilon}$  puts arbitrarily high probability on a minimal  $\gamma$ -configuration with span  $Y$ .*

PROOF: See the Appendix.

The proof of Theorem 2 uses the following two observations. First, a state in a strictly Pareto-dominant set  $Y$  can be reached from any state outside the set if all individuals drawn to play simultaneously make a mistake and play a pure strategy in  $Y$ , and a sequence of imitators thereafter pick samples including this mistake. Second, the resistance of the reverse transition can be made arbitrarily large by choosing a sufficiently large sample and memory size.

Theorem 2 does not hold under the weaker condition that an MCUBR set strictly Pareto dominates all other MCUBR sets. Consider the game in Figure 1. In this game, there are two MCUBR sets,  $(A, a)$  and  $(C, c)$ , and  $(A, a)$  strictly Pareto dominates  $(C, c)$ . However,  $(A, a)$  does not strictly Pareto dominate  $(B, b)$ , and for a sufficiently large sample size and sufficiently incomplete information, the stochastic potential is two for both  $h_{(A,a)}$  and  $h_{(C,c)}$ . This follows since the process will make the transition from  $h_{(A,a)}$  to  $h_{(B,b)}$ , if the two players simultaneously make mistakes in period  $t$  and play  $(B, b)$ , and a sequence of  $s - 1$  imitators in population  $C_1$ , corresponding to the row player, thereafter are drawn to play and sample only  $x^t$  and earlier strategy realizations, and finally imitators in both populations sample only from plays more recent than  $x^{t-1}$  for  $m - s$  periods. Since state  $h_{(B,b)}$  is clearly in the basin of attraction of  $h_{(C,c)}$ , this implies that the stochastic potential

of  $h_{(C,c)}$  is two at most.

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

FIGURE 1.

In *games of common interest* there exists a strict Nash equilibrium – a singleton MCUBR set – which strictly Pareto dominates all other strategy-tuples. From Theorem 2, it immediately follows that for a sufficiently large memory and sample size, and sufficiently incomplete information, this equilibrium is played with probability one under the limiting distribution.

#### 4 Results for $2 \times 2$ Coordination Games

In this section, I will study the predictions of heterogeneous adaptive play in the special class of  $2 \times 2$  Coordination Games and illustrate how the stochastically stable states depend on the sample size.

Consider the game in Figure 2.

	$a$	$b$
$A$	$\pi_1(A, a), \pi_2(A, a)$	$\pi_1(A, b), \pi_2(A, b)$
$B$	$\pi_1(B, a), \pi_2(B, a)$	$\pi_1(B, b), \pi_2(B, b)$

FIGURE 2.

This game is a  $2 \times 2$  *Coordination Game* if  $(A, a)$  and  $(B, b)$  are strict Nash equilibria. It is a *symmetric*  $2 \times 2$  *Coordination Game* if, in addition, the diagonal payoffs are equal for the two players,  $\pi_1(B, a) = \pi_2(A, b)$  and  $\pi_1(A, b) = \pi_2(B, a)$ . An equilibrium  $(A, a)$  of a  $2 \times 2$  Coordination Game is *risk dominant* if its Nash product exceeds that of  $(B, b)$ :

$$\begin{aligned} & [\pi_1(A, a) - \pi_1(B, a)] [\pi_2(A, a) - \pi_2(A, b)] \\ & \geq [\pi_1(B, b) - \pi_1(A, b)] [\pi_2(B, b) - \pi_2(B, a)]. \end{aligned} \quad (3)$$

This definition (with a strict inequality) is originally due to Harsanyi and Selten (1988). Denote a monomorphic state  $h$  such that  $h = (x, \dots, x)$  by  $h_x$ .

If  $x$  is a strict Nash equilibrium, then  $h_x$  is called a *convention* (Young, 1993a, 1998). If  $x$  is risk (strictly Pareto) dominant, I say that the convention  $h_x$  is risk (strictly Pareto) dominant. From Theorem 2, the following corollary immediately follows.

**Corollary 1** *In  $2 \times 2$  Coordination Games with a strictly Pareto-dominant equilibrium, for a sufficiently large sample size, and sufficiently incomplete information, the strictly Pareto-dominant convention is a unique stochastically stable state.*

For certain payoffs, the sample size must to be very large for this result to hold. Consider the game in Figure 3, where the equilibrium  $(A, a)$  strictly Pareto dominates the equilibrium  $(B, b)$ , but where  $(B, b)$  is the unique risk-dominant equilibrium.

	$a$	$b$
$A$	2, 2	-1000, 0
$B$	0, -1000	1, 1

FIGURE 3.

If the process is in state  $h_{(B,b)}$  and the sample size  $s > 1$ , then two simultaneous mistakes, followed by a sequence of imitators, are required to reach  $h_{(A,a)}$ . One mistake is sufficient to make the reverse transition if this makes the expected payoff to playing strategy  $B$  higher than that of  $A$  for a better or best replier, or formally if

$$\frac{2(s-1) - 1000}{s} \leq \frac{1}{s} \Leftrightarrow s \leq \frac{1003}{2}. \quad (4)$$

Hence, only the risk-dominant equilibrium is selected, in the sense that the corresponding convention is stochastically stable, for  $s$  such that  $1 < s \leq 501$ . In order to ensure that only the strictly Pareto-dominant equilibrium is selected, the sample size must be so large that

$$\frac{2(s-2) - 2000}{s} > \frac{2}{s} \Leftrightarrow s > 1003. \quad (5)$$

For a sample size in between these values, i.e.  $s$  such that  $501 < s \leq 1003$ , both equilibria are selected.

More generally, define the probabilities  $q_A$ ,  $q_a$ , and  $q_{\min}$  by

$$q_A = \frac{\pi_2(B, b) - \pi_2(B, a)}{\pi_2(A, a) - \pi_2(A, b) + \pi_2(B, b) - \pi_2(B, a)}, \quad (6)$$

$$q_a = \frac{\pi_1(B, b) - \pi_1(A, b)}{\pi_1(A, a) - \pi_1(B, a) + \pi_1(B, b) - \pi_1(A, b)}, \quad (7)$$

and

$$q_{\min} = \min \{q_A, q_a, 1 - q_A, 1 - q_a\}. \quad (8)$$

In other words,  $q_A$  is the probability of strategy  $A$  which makes the expected payoff to strategies  $a$  and  $b$  identical for player 2, and  $q_a$  is the probability of strategy  $A$  which makes the expected payoff to strategies  $A$  and  $B$  identical for player 1.  $q_{\min}$  is the lowest probability required to make a player indifferent between her pure strategies. Note that in symmetric Coordination Games  $q_A = q_a$  and that (see Young, 1998) an equilibrium  $(A, a)$  is risk dominant if and only if

$$\min\{q_A, q_a\} \leq \min\{1 - q_A, 1 - q_a\}. \quad (9)$$

Let  $Z_{++}$  be the set of positive integers and let  $\lceil y \rceil$  denote the smallest integer greater than or equal to  $y$  for any real  $y$ . Further, define the function  $s_{diff} : Z_{++} \rightarrow Z_{++} \cup \{0\}$  and the set  $S^{RD} \subset Z_{++}$  as follows:

$$s_{diff}(s) = |\lceil s \min\{q_A, q_a\} \rceil - \lceil s \min\{1 - q_A, 1 - q_a\} \rceil| \quad (10)$$

$$S^{RD} = \{s \in Z_{++} : s_{diff}(s) \geq 1 \text{ and } s \leq 1/q_{\min}\}. \quad (11)$$

Note that  $s_{diff}(s)$  is increasing in  $s$ , implying that if there are integer sample sizes  $s, s'$ , and  $s''$  such that  $s < s' < s''$ , and  $s, s'' \in S^{RD}$ , then also  $s' \in S^{RD}$ .

**Theorem 3** *Let  $\Gamma$  be a  $2 \times 2$  Coordination Game and assume that information is sufficiently incomplete ( $s/m \leq 1/2$ ).*

*i) From any initial state, the unperturbed process  $P^{*,m,s,0}$  converges with probability one to a convention and locks in.*

ii) For sufficiently large  $s$  the stochastically stable states of the perturbed process  $P^{*,m,s,\varepsilon}$  correspond one to one with the conventions that are not strictly Pareto dominated.

iii) If there is a unique risk-dominant equilibrium and the set  $S^{RD}$  is non-empty, then, for sample size  $s \in S^{RD}$ , only the risk-dominant convention is stochastically stable.

iv) If both equilibria are risk dominant, then, for sample size  $s \in [1, 2/q_{\min}]$ , both conventions are stochastically stable.

PROOF: See the Appendix.

Theorem 3 gives sufficient conditions for the selection of different equilibria. The intuition behind the proof of this theorem is the same as in the above example. If the process is in state  $h_{(A,a)}$ , and  $(B, b)$  is a risk-dominant equilibrium, then, for sample sizes such that  $q_{\min} \leq 1/s$ , or equivalently  $s \leq 1/q_{\min}$ , only one mistake is necessary to make a sequence of subsequent better or best repliers switch to the risk-dominant pure strategy such that  $h_{(B,b)}$  is reached. Second, if the process is in state  $h_{(B,b)}$  and  $(A, a)$  is not strictly Pareto dominated, then two simultaneous mistakes, followed by a sequence of imitators in one of the populations, are sufficient to reach the basin of attraction of  $h_{(A,a)}$ . The condition that  $s_{diff}(s) \geq 1$  implies that the transition from  $h_{(B,b)}$  to  $h_{(A,a)}$  cannot be made with only one mistake followed by a sequence of best or better repliers.

The following corollary shows that in symmetric Coordination Games, there always exists a range of sample sizes such that the risk-dominant equilibrium is selected.

**Corollary 2** *In symmetric  $2 \times 2$  Coordination Games, if information is sufficiently incomplete ( $s/m \leq 1/2$ ) and the sample size sufficiently small ( $2 \leq s \leq 1/q_{\min}$ ), only the risk-dominant convention is stochastically stable.*

PROOF: If the symmetric game has two risk-dominant equilibria, the result follows immediately from Theorem 3 iv). If the game has a unique

risk-dominant equilibrium, then  $q_{\min} < 1/2$ . This implies that if the sample size is such that  $2 \leq s \leq 1/q_{\min}$ , then  $sq_{\min} \leq 1$ , and  $s(1 - q_{\min}) > 1$ ; hence, the requirement that  $s_{diff}(s) \geq 1$  is redundant, and the claim follows from Theorem 3 iii). *Q.E.D.*

Holding the sample and memory size fixed in symmetric games with one risk-dominant equilibrium and a different Pareto-dominant equilibrium, Corollary 2 implies that the risk-dominant equilibrium should be observed for payoffs such that  $q_{\min}$  is low, possibly both equilibria for payoffs such that  $q_{\min}$  is intermediate, and the Pareto-dominant equilibrium for payoffs such that  $q_{\min}$  is high. This appears to be in line with experimental evidence from symmetric Coordination Games (see Van Huyck (1997) for a survey), although the absolute difference between the equilibrium payoffs and the off-equilibrium payoffs also seems to matter. Given the limited number of repetitions in such experiments, one may naturally question whether the outcome observed actually corresponds to the stochastically stable state.

## 5 Conclusion

In this paper, I analyze stochastic adaptation in finite  $n$ -player games played by heterogeneous populations of myopic best repliers, better repliers and imitators. I show that for sufficiently incomplete information, and independently of the population shares, the recurrent sets of the resulting unperturbed Markov chain correspond one to one with the minimal closed sets closed under better replies of the game. Such sets are also asymptotically stable under a large class of deterministic continuous time selection dynamics, containing the replicator dynamics.

The stochastically stable states are contained in the recurrent sets with minimum stochastic potential. In games where one minimal closed set under better replies is Pareto efficient, the span of the stochastically stable states is equal to this set. This result requires a sufficiently large sample size, determined by the payoffs of the game. In all symmetric Coordination Games and



many asymmetric Coordination Games, the Pareto efficient equilibrium is selected for a sufficiently large sample size, and the risk-dominant equilibrium is selected for a sufficiently small sample size.

The main contribution of this paper is to analyze the long-run outcome when multiple rules of adaptation, previously analyzed only in isolation, are present in the populations. It is shown that the support of the limiting distribution is independent of the population shares. However, in many games, the exact shape of the limiting distribution will depend on the population shares using the different learning rules, and so will the expected payoff to the individuals employing the various rules. A next step is to study how the individuals using the different learning rules fare for different population shares, and to ask if there exists a rule, or a particular combination of rules, which is evolutionarily stable.

## Appendix

**Definition 3** (Young, 1998)  $P^\varepsilon$  is a **regular perturbed Markov process** if  $P^\varepsilon$  is irreducible for every  $\varepsilon \in (0, \varepsilon^*]$ , and for every  $h, h' \in H$ ,  $P_{hh'}^\varepsilon$  approaches  $P_{hh'}^0$  at an exponential rate, i.e.  $\lim_{\varepsilon \rightarrow 0} P_{hh'}^\varepsilon = P_{hh'}^0$  and if  $P_{hh'}^\varepsilon > 0$  for some  $\varepsilon > 0$ , then  $0 < \lim_{\varepsilon \rightarrow 0} \frac{P_{hh'}^\varepsilon}{\varepsilon^{r_{hh'}}} < \infty$  for some  $r_{hh'} \geq 0$ .

**Lemma 2**  $P^{*,m,s,\varepsilon}$  is a regular perturbed Markov process.

PROOF: The proof of Lemma 2 is completely analogous to the proof in Young (1998, p. 55) *Q.E.D.*

**Theorem 4** (Young, 1998) Let  $P^\varepsilon$  be a regular perturbed Markov process and let  $\mu^\varepsilon$  be the unique stationary distribution of  $P^\varepsilon$  for  $\varepsilon > 0$ . Then  $\lim_{\varepsilon \rightarrow 0} \mu^\varepsilon = \mu^0$  exists, and  $\mu^0$  is a stationary distribution of  $P^0$ . The stochastically stable states are precisely those states that are contained in the recurrent classes of  $P^0$  having minimum stochastic potential.

PROOF OF THEOREM 1: Part i) of Theorem 1 will be proved in two steps. In step A, I will prove that from any initial state,  $P^{*,m,s,0}$  converges to a minimal  $\gamma$ -configuration. In step B, I will prove that for any MCUBR set, there exists a minimal  $\gamma$ -configuration.

A. Let  $\gamma$ -adaptive play be the induced Markov chain when the share of better repliers in each population is one. According to Theorem 1 in Josephson (2000), for a sufficiently large sample size and sufficiently incomplete information ( $s/m \leq \frac{1}{|\mathcal{X}|}$ ), the span of each recurrent class of the  $\gamma$ -adaptive play corresponds one to one with an MCUBR set. Since, in each period, there is a positive probability that only better repliers will be selected to play, there is also a positive probability that the Markov chain, from *any* initial state, and in a finite number of periods, will end up in a state  $h$  which belongs to a recurrent class  $E_Z^\gamma$  of the better-reply dynamic, and thus only involves play of pure-strategy profiles in a corresponding MCUBR set  $Y$ . From Lemma 1, it follows that the process will never play a pure strategy outside  $Y$ , once it

has reached such a state. If  $E_Z^\gamma$  is singleton, this means that it is a recurrent class also under heterogeneous adaptive play  $P^{*,m,s,0}$ . If  $E_Z^\gamma$  is not singleton, there is a positive probability that the process makes the transition from  $h$  to any other state  $h'$  of  $E_Z^\gamma$  in a finite number of periods and, naturally, without playing any pure strategy outside  $Y$ . This follows since in each period, there is a positive probability that only better repliers will be selected to play. Hence, there exists a recurrent class  $E_Y^*$  of  $P^{*,m,s,0}$ , such that  $E_Z^\gamma \subseteq E_Y^*$  and  $S(E_Y^*) \subseteq Y = S(E_Z^\gamma)$ .

B. Conversely, I will prove that for any MCUBR set, there exists a minimal  $\gamma$ -configuration. If  $Y$  is an MCUBR set, then, by Theorem 1 in Josephson (2000), there exists a unique recurrent class  $E_Z^\gamma$  of  $\gamma$ -adaptive play such that  $S(E_Z^\gamma) = Y$ . If the Markov chain  $P^{*,m,s,0}$  is in a state involving no pure strategy outside  $Y$ , there is a positive probability that it will reach a state which belongs to  $E_Z^\gamma$  in a finite number of periods. This follows since  $P^{*,m,s,0}$ , by A., from any initial state and in a finite number of periods will end up in a state  $h$ , which belongs to a recurrent class of the better-reply dynamic, and since, by Lemma 1, the process will never play a pure strategy outside  $Y$ , once it has reached a state only involving pure-strategy profiles of  $Y$ . It thereafter follows from A. that there exists a recurrent set  $E_Y^*$  of  $P^{*,m,s,0}$ , such that  $S(E_Y^*) = Y$ .

Part ii) of Theorem 1 follows directly from Theorem 4 in the Appendix since  $P^{*,m,s,\varepsilon}$ , by Lemma 2 in the Appendix, is a regular perturbed Markov process. *Q.E.D.*

**PROOF OF THEOREM 2:** Theorem 2 will be proved in three steps. In step A, I show that the transition from any recurrent class to a Pareto-dominant minimal  $\gamma$ -configuration can always be made with at most  $n$  mistakes. In step B, I prove that for a sufficiently large sample size, the transition from a Pareto-dominant minimal  $\gamma$ -configuration to any other recurrent class requires at least  $n + 1$  mistakes. In step C, I use A and B to prove that the Pareto-dominant minimal  $\gamma$ -configuration must have minimum stochastic potential.

A. Assume that the sample size is sufficiently large and  $s/m \leq 1/|X|$  so that, by Theorem 1, the span of each recurrent class of  $P^{*,m,s,0}$  corresponds one to one with an MCUBR set. Let  $Y \subset X$  be a strictly Pareto-dominant MCUBR set, and let  $E_Y^*$  be the corresponding minimal  $\gamma$ -configuration. Assume there exists at least one other recurrent class (otherwise Theorem 2 holds trivially). The transition from such a recurrent class to  $E_Y^*$  can always be made with a probability of the order  $\varepsilon^N$  (or higher). This is, for instance, the case if the individuals in all roles experiment and play a pure-strategy profile  $y^t \in Y$  in period  $t$ , and a sequence of  $m - 1$  imitators thereafter are drawn to play in all roles, and all sample  $y^t$ .

B. Let  $E_Z^*$  be an arbitrary recurrent class, different from  $E_Y^*$ . I claim that for a sufficiently large sample size, the probability of the transition from  $E_Z^*$  to  $E_Y^*$  is at least of the order  $\varepsilon^{N+1}$ . To make a best or better replier in role  $i$ , switch to a pure strategy  $x_i \notin Y_i$  after at most  $n$  mistakes; the expected payoff to that pure strategy must be greater than for any pure strategy  $y_i \in Y_i$ ,

$$\frac{s-n}{s}u_i(y_i, p_{-i}^Y) + \frac{n}{s}u_i(y_i, p_{-i}^X) \leq \frac{s-n}{s}u_i(x_i, p_{-i}^Y) + \frac{n}{s}u_i(x_i, p_{-i}^X) \quad (12)$$

$\Leftrightarrow$

$$s \leq \frac{n [u_i(y_i, p_{-i}^Y) - u_i(y_i, p_{-i}^X) + u_i(x_i, p_{-i}^X) - u_i(x_i, p_{-i}^Y)]}{u_i(y_i, p_{-i}^Y) - u_i(x_i, p_{-i}^Y)}, \quad (13)$$

where  $p_{-i}^Y \in \square(Y_{-i})$ , and  $p_{-i}^X \in \square(X_{-i})$ . By the boundedness of payoffs and the strict Pareto dominance of  $Y$ , the right-hand side of this inequality is clearly bounded for any  $i$ ,  $y_i \in Y_i$ ,  $x_i \notin Y_i$ ,  $p_{-i}^Y \in \square(Y_{-i})$ , and  $p_{-i}^X \in \square(X_{-i})$ . Hence, there exists some finite  $\hat{s}$ , such that for  $s > \hat{s}$ , strictly more than  $n$  mistakes are necessary for a better or best replier to play a pure strategy outside  $Y$ .

Similarly, in order to make an imitator maximizing the average realized

payoff switch to a pure strategy  $x_i \notin Y_i$  after at most  $n$  mistakes,

$$\frac{(s-n)u_i(y_i, p_{-i}^Y) + (n-1)u_i(y_i, p_{-i}^X)}{s-1} \leq \pi_i(x_i, x_{-i}) \quad (14)$$

$\Leftrightarrow$

$$s \leq \frac{n[u_i(y_i, p_{-i}^Y) - u_i(y_i, p_{-i}^X)] + u_i(y_i, p_{-i}^X) - \pi_i(x_i, x_{-i})}{u_i(y_i, p_{-i}^Y) - \pi_i(x_i, x_{-i})}. \quad (15)$$

Once again, by the boundedness of payoffs and the strict Pareto dominance of  $Y$ , the right-hand side of the last inequality is bounded for any  $i$ ,  $y_i \in Y_i$ ,  $x_i \notin Y_i$ ,  $x_{-i} \in X_{-i}$ ,  $p_{-i}^Y \in \square(Y_{-i})$ , and  $p_{-i}^X \in \square(X_{-i})$ . Hence, there exists some finite  $\check{s}$ , such that if  $s > \check{s}$ , strictly more than  $n$  mistakes are necessary for a better or best replier to play a pure strategy outside  $Y$ . An imitator who simply pick the pure strategy with the maximum realized payoff can, of course, never switch to a pure strategy  $x_i \notin Y_i$  with less than  $s$  mistakes. Thus, for  $s > \max\{\hat{s}, \check{s}, N\}$ , the resistance of the transition from  $E_Y^*$  to  $E_Z^*$  must be greater than  $N$ .

C. Consider the minimum resistance tree rooted at an arbitrary recurrent class  $E_D^*$ , different from  $E_Y^*$ . In this tree, there must be a directed edge from  $E_Y^*$  to some other recurrent class  $E_Z^*$  (possibly identical to  $E_D^*$ ). Assume that the sample size is so large that the resistance of the transition from  $E_Y^*$  to  $E_Z^*$  is greater than  $N$  (this is possible by B.), and that the stochastic potential of  $E_D^*$  is smaller than or equal to that of  $E_Y^*$ . Create a new tree by deleting the edge from  $E_Y^*$  to  $E_Z^*$ , and adding an edge from  $E_D^*$  to  $E_Y^*$ . The resistance of the deleted edge is, by assumption, greater than  $N$ , and the resistance of the added edge is (by A.) smaller than or equal to  $N$ . Hence, the total resistance of the new tree is less than that of the tree rooted at  $E_D^*$ , contradicting the assumption that the stochastic potential of  $E_D^*$  is smaller than or equal to that of  $E_Y^*$ . This proves that the stochastic potential of the Pareto-dominant recurrent class  $E_Y^*$  is lower than for any other recurrent class and, by Lemma 2 and Theorem 4 in the Appendix, Theorem 2 follows. *Q.E.D.*

PROOF OF THEOREM 3: i) Convergence with probability one to a convention. By Theorem 2 in Josephson (2000), if the share of better repliers

is one in all populations and  $s/m \leq 1/2$ , then the recurrent classes of the unperturbed process are the two monomorphic states  $h_{(A,a)}$  and  $h_{(B,b)}$ . It is clear that these states are absorbing also when the shares of imitators and best repliers are positive. Moreover, since in each period, there is a positive probability that only better repliers are drawn to play, there is a positive probability of reaching one of these two states in a finite number of periods from any other state. Hence, for  $s/m \leq 1/2$ , the recurrent classes of  $P^{*,m,s,0}$  are two states  $h_{(A,a)}$  and  $h_{(B,b)}$ .

ii) Selection of equilibria that are not strictly Pareto dominated. Without loss of generality, assume that  $(A, a)$  is a strictly Pareto-dominant equilibrium. Then, the transition from  $h_{(B,b)}$  to the basin of attraction of  $h_{(A,a)}$  can always be made with two simultaneous mistakes in period  $t$  followed by a sequence of  $s - 1$  imitators in both populations, who all sample  $x_i^t$  or more recent strategy-tuples. The reverse transition requires at least three mistakes if the following two conditions are fulfilled. First, in state  $h_{(A,a)}$ , best or better repliers in one of the populations should not be able to switch to strategy  $B$  when the sample contains less than three mistakes by the other population. This is prevented if  $q_{\min} > 2/s$ , or equivalently if  $s > 2/q_{\min}$ . Second, imitators in population  $C_1$  should not be able to switch to strategy  $B$  with less than three mistakes. They can only do this if two mistakes, one after the other by different populations, make the average payoff to strategy  $B$  at least as large as that of  $A$ , or formally, if

$$\frac{(s-2)\pi_1(A,a) + \pi_1(A,b)}{s-1} \leq \pi_1(B,a) \quad (16)$$

$\Leftrightarrow$

$$s \leq \frac{2\pi_1(A,a) - \pi_1(B,a) - \pi_1(A,b)}{\pi_1(A,a) - \pi_1(B,a)} = \hat{s}_1. \quad (17)$$

It is clear that for  $s > \hat{s}_1$  this inequality does not hold, and a similar critical sample size  $\hat{s}_2$  can be computed for individuals in population  $C_2$ . Hence, for  $s > \max\{\hat{s}_1, \hat{s}_2, 2/q_{\min}\}$ , the transition to  $(B,b)$  requires at least three mistakes, whereas the reverse transition requires exactly two mistakes (if all

imitators choose the pure strategy with maximum sample payoff, i.e. nobody chooses the pure strategy with the maximum *average* payoff in the sample, then it is sufficient that  $s > 2/q_{\min}$ ).

If neither of the equilibria are Pareto dominated, then either  $\pi_1(A, a) \geq \pi_1(B, b)$  and  $\pi_2(A, a) \leq \pi_2(B, b)$ , or  $\pi_1(A, a) \leq \pi_1(B, b)$  and  $\pi_2(A, a) \geq \pi_2(B, b)$ . Without loss of generality, assume that the first two inequalities hold. Then, the transition from  $h_{(B, b)}$  to the basin of attraction of  $h_{(A, a)}$  can always be made with two simultaneous mistakes in period  $t$ , followed by a sequence of  $s - 1$  imitators in population  $C_1$ , who all sample  $x_i^t$  and earlier strategy realizations. In an analogous fashion, the transition from  $h_{(A, a)}$  to  $h_{(B, b)}$  can be made with two mistakes. If, in addition,  $s > 1/q_{\min}$ , then no better or best repliers are able to switch strategies in state  $h_{(A, a)}$  or  $h_{(B, b)}$  after only one mistake. Thus, for general Coordination Games and sample size such that  $s > \max\{\hat{s}_1, \hat{s}_2, 2/q_{\min}\}$ , the stochastic potential is lowest for the state(s) which is (are) not strictly Pareto dominated.

iii) Unique risk-dominant equilibrium. The transition from  $h_{(B, b)}$  to the basin of attraction of  $h_{(A, a)}$  can be made with  $k$  mistakes if individuals in one of the populations, say population  $C_1$ , by mistake plays  $A$   $k$  times in a row, and  $k/s \geq q_A$ . This follows since there is a positive probability that better or best repliers are drawn to play in the other population for the next  $s$  periods, and that these individuals all sample the  $k$  mistakes. Without loss of generality, assume that  $(A, a)$  is a risk-dominant equilibrium. Then, the transition from  $h_{(B, b)}$  to  $h_{(A, a)}$  requires only one mistake if  $q_{\min} \leq 1/s$ . The requirement that  $s_{diff}(s) > 1$  ensures that the reverse transition requires at least two mistakes.

iv) Two risk-dominant equilibria. If both equilibria are risk dominant, then

$$\min\{q_A, q_a\} = \min\{1 - q_A, 1 - q_a\} = q_{\min}. \quad (18)$$

This implies that for  $sq_{\min} \leq 2$ , the transition from  $h_{(A, a)}$  to  $h_{(B, b)}$ , and the

reverse transition, requires the same number of mistakes,  $k \leq 2$ . Since at least two mistakes are required to make an imitator switch strategies when the process is in a convention, the stochastic potential of the conventions will equal the number of mistakes required to make better and best repliers switch strategy for  $s \in [1, 2/q_{\min}]$ .

Theorem 3 now follows by Lemma 2 and Theorem 4. *Q.E.D.*

## References

- [1] Foster, D., and P. Young (1990): "Stochastic Evolutionary Game Dynamics," *Theoretical Population Biology* 38, 219-32.
- [2] Freidlin, M., and A. Wentzell (1984): *Random Perturbations of Dynamical Systems*. Springer-Verlag.
- [3] Fudenberg, D., and D. Levine (1998): *The Theory of Learning in Games*. The MIT Press.
- [4] Harsanyi, J., and R. Selten (1988): *A General Theory of Equilibrium Selection in Games*. The MIT Press.
- [5] Van Huyck, J. (1997): "Emergent Conventions in Evolutionary Games," to appear in C. Plott and V. Smith, eds., *Handbook of Results in Experimental Economics*, Amsterdam: North-Holland, 2001.
- [6] Josephson, J., and A. Matros (2000): "Stochastic Imitation in Finite Games," SSE/EFI Working Paper Series in Economics and Finance No. 363, Stockholm.
- [7] Josephson, J. (2000): "Stochastic Better-Reply Dynamics in Games," SSE/EFI Working Paper Series in Economics and Finance No. 400, Stockholm.
- [8] Kaarbøe, O. M., and A. F. Tieman (1999): "Equilibrium Selection under Different Learning Modes in Supermodular Games," Discussion paper nr. TI 99-061/1, Tinbergen Institute, Amsterdam.
- [9] Kaniovski, Y. M., A. V. Kryazhimiskii, and H. P. Young (2000): "Adaptive Dynamics in Games Played by Heterogeneous Populations," *Games*



and *Economic Behavior* 31, 50-96.

- [10] Matros, A. (2000): "Clever Agents in Adaptive Learning," SSE/EFI Working Paper Series in Economics and Finance No. 403, Stockholm.
- [11] Kandori, M., G. Mailath, and R. Rob (1993): "Learning, Mutation and Long-Run Equilibria in Games," *Econometrica* 61, 29-56.
- [12] Ritzberger, K., and J. W. Weibull (1995): "Evolutionary Selection in Normal-Form Games," *Econometrica* 63, 1371-1400.
- [13] Sáez Martí, M., and J. W. Weibull (1999): "Clever Agents in Young's Evolutionary Bargaining Model," *Journal of Economic Theory* 86, 268-279.
- [14] Simon, H. A. (1955): "A Behavioral Model of Rational Choice," *Quarterly Journal of Economics* 69, 99-118.
- [15] Young, P. H. (1993a): "The Evolution of Conventions," *Econometrica* 61, 57-84.
- [16] ——— (1993b): "An Evolutionary Model of Bargaining," *Journal of Economic Theory* 59, 145-68.
- [17] ——— (1998): *Individual Strategy and Social Structure*. Princeton University Press.