Comments welcome.

Projectbereich B

Discussion Paper No. B-296

Why Imitate, and if so, How?
Exploring a Model of Social Evolution*

Karl H. Schlag

November 1994

Wirtschaftstheoretische Abteilung III

Friedrich-Wilhelms-Universität Bonn

Adenauerallee 24-26

D-53113 Bonn

Germany

e-mail: schlag@glider.econ3.uni-bonn.de

---

* I especially wish to thank Avner Shaked for many useful comments. I also wish to thank Dirk Bergemann, Jonas Björnerstedt, Georg Nöldeke and Larry Samuelson for helpful comments and discussions. Financial support from the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 303 at the University of Bonn is gratefully acknowledged.

Abstract

In consecutive rounds, each agent in a finite population chooses an action, is randomly matched, obtains a payoff and then observes the performance of another agent. An agent determines future behavior based on the information she receives from the present round. She chooses among the behavioral rules that increase expected payoffs in any specifications of the matching scenario. The rule that outperforms all other such rules specifies to imitate the action of an agent that performed better with probability proportional to how much better she performed. The evolution of a large population in which each agent uses this rule can be approximated in the short run by the replicator dynamics.

JEL classification number: C79.
Keywords: random matching, learning, imitation, replicator dynamics.

0. Introduction

Imitation is the act of copying or mimicking the actions of others, a behavior that is often observable in the context of human decision making. Several reasons for why people might choose to imitate can be given. We will provide a theoretical foundation for imitation based on performance in an environment in which an agent's information is extremely limited[1].

Consider an individual that is repeatedly engaged in a contest (or game) against a random opponent. Before each contest the individual must choose some action (from a fixed set of actions). This action together with the action chosen by her opponent determines the success (formally, the payoff) of the individual in the contest. Between rounds the individual is able to obtain some information for selecting future actions by observing the performance of another individual that is confronted with the same setup.

Assume now that the individual has extremely limited knowledge and information about the further specifications of the model. All that the she knows about the contest is an interval that will contain her payoffs. She has no prior belief as to which actions are likely to be played by other agents. The only prior the individual assesses to unknown events is that in any given contest and specification of the model the individual conceives to be equally likely in her own position as to be in the position of any of the individuals that she observes the performance of. We assume that the individual chooses a behavioral rule that increases expected payoffs in any given round for any given specification of the parameters of the model. Thereby the individual ignores the fact that other agents in the population might also change their actions. Such rules will be called improving. Thus, the focus of our analysis will be on whether under the above assumptions an individual can extract information from her observations for future behavior in a way such that her choice of an action does not depend on a prior belief over the unobservable parameters.

Notice that the contest might be such that the payoffs of the individual are independent of the opponent's action. In the following these contests will be called degenerate. As a reference point assume for a moment an alternate (admittedly uninteresting) model in which all games are

[1] For a justification of imitation in a different context see Banerjee [1992].

degenerate and assume that the individual is aware of this fact. In this alternate model there is a very intuitive improving rule referred to in the following as "imitate if better": do not change actions unless you observe an agent who performed better in which case imitate the action of the observed agent. It turns out that this rule is no longer improving in our model in which contests generally are not degenerate.

We give a complete characterization of the improving rules. Especially it follows that an individual using an improving rule will display an imitating behavior, i.e., she will never switch to an action not observed. We derive a unique most preferable improving rule that is described in the following brief summary of the analysis.

Q: "Why imitate?"

A: "In order to increase expected payoffs under any circumstances in a model of limited information."

Q: "And if so, how?"

A: "Imitate actions that perform better with probability proportional to how much better they perform."

The new aspect of this model is the search for a behavioral rule that has certain properties in each possible configuration under limited information. Previously, individual behavior in random matching models has only been justified in some examples of environments with rich information (see Kandori, Mailath and Rob [1993], more about this in section 8).

We now give a more detailed description of the analysis, beginning with an introduction of the matching and sampling scenario.

Consider a finite population that consists of an equal number of two different types of agents, referred to in the following as type A and as type B agents (e.g., sellers and buyers). Each agent is provided with a set of actions (or strategies) S^A or S^B according to her type.

These agents interact in each of a sequence of rounds according to the following scenario. In each round each agent selects an action, is then randomly matched with an agent of the opposite type and obtains thereafter a payoff according to her action and the action of her opponent. During the matching, an agent only learns her own payoff, neither the action nor the payoff of her opponent.

Between matching rounds each agent receives information about the performance of another agent of the same type. This information consists of a payoff and of the action with which this payoff was achieved without exposing the identity of the associated agent. This so-called sampling is independent of the matching and sampling in the previous rounds and occurs according to an exogenous sampling procedure. We consider only sampling procedures where the probability of 'a' sampling 'b' is the same as that of 'b' sampling 'a'.

An example of a symmetric sampling scenario is the situation in which each agent randomly selects an agent among the agents of her type with equal probability. These specific sampling procedures will be referred to as random sampling.

We consider an individual that is about to enter the above matching and sampling scenario and must determine before her entry how to update her actions from round to round. The following information will be available to the individual when she selects her action.

The individual knows her type and her set of actions. Additionally she knows that choosing an action and being matched against an agent of the opposite type will lead to some payoff in a closed interval (denoted by $I^A \subseteq \Re$ or $I^B \subseteq \Re$, depending on her type). The individual knows that in the matching scenario she can only remember her observations from the previous round. Therefore the action the individual selects in a given round of the matching scenario can be characterized as a function of her previous action, the payoff she achieved in the previous round and of the action and the payoff of the agent she sampled in the previous round. Such a function will be called an updating rule.

Given the above considerations, the objective of the individual is reduced to selecting an updating rule before she enters the matching and sampling scenario.[2]

Of course there is a large variety of updating rules. An updating rule may specify to switch to an action that was not observed. Even among the updating rules that have the imitation property, i.e., switch only to actions whose performance was observed (in the previous round),

[2] Apart from section 7 we will ignore the fact that an action must also be specified when there is no previous experience, i.e., in the first round.

there are many different ways of reacting to the observed payoffs.

For example: i) The simplest such rule is the one that specifies to choose the same action as in the previous round regardless of the observations. ii) A very plausible rule seems to be for an agent to adapt the action of the sampled agent if it performed better than hers did in the previous round. This updating rule will be called "imitate if better". iii) An example for an updating rule that incorporates the relative performance of actions is as follows. After observing an agent of the same type that received a higher payoff than your own, adapt (imitate) the sampled agent's action with probability proportional to the difference in the achieved payoffs. In all other cases, play the same action again. We will refer to this updating rule as the proportional imitation rule. Notice that this rule is uniquely determined up to a positive (proportionality) constant. This constant is bounded by the range of payoffs that are attainable.

In the following we will specify the criteria for selecting an updating rule.

We assume that an individual does not know the size of the population. However she knows the matching and sampling scenario, especially she knows which sampling procedure will be used in the various population sizes. However she is restricted in the way she perceives the situation.

(Ignorance) We assume that the individual ignores the fact that other agents might adapt. The individual also ignores the effect her own rule has on future population distributions.

An environment will be the collection of the following specifications: the payoff function for each type of agent where payoffs to type A (B) are in I^A (I^B), the population size and the sampling procedure that is symmetric. For a fixed environment and round, a state will be the specification of the action each agent is playing.

(Uniform Prior) The only prior that the individual has is that she expects in any environment and any state to be equally likely in the position of any of the agents of her own type.[3]

[3] This assumption is easily motivated by an entry and exit scenario (see section 7).

How do the above assumptions affect the individual's calculation of the expected change in payoffs (denoted by EIP) when using a given updating rule in a given state and environment? For each agent 'a' of the same type as the individual calculate the expected increase in payoffs of this agent when using the given updating rule and when no other agent changes actions. Denote this expression by EIP(a). Here the expectation is taken with respect to the uncertainty in the matching and sampling only. Under the uniform prior assumption it then follows that EIP is the average of EIP(a) over all agents in the population that have the same type as the individual that is about to enter.

We assume that the individual does not want to choose an updating rule that performs worse than the rule of never switching. Moreover the individual conceives each environment and state possible and has no priors over these events. Therefore we choose a "distribution free approach" and assume that the individual chooses a rule that does not perform worse than not switching in any state and any environment. This leads to the following assumption.

(Improving) The individual chooses an updating rule that increases (improves) expected payoffs (i.e., $EIP \geq 0$) from one round to the next in any environment, any state and any round. Such rules are called improving. Especially the rule "never switch" is improving.

(Dominance) Moreover, the individual does not choose an improving rule that is pareto dominated by some other improving rule. One updating rule pareto dominates another updating rule if in any environment, any state and any round the former rule achieves a higher expected payoff in the next round than the later.

When indifferent between updating rules the individual evaluates their posterior "improving" abilities when the contest is degenerate, i.e., in the situation where payoffs are independent of the actions of the opposite type. Clearly, a rule that never imitates actions that achieved lower payoffs never decreases payoffs in a degenerate contest.

(Degenerate Improving) Given a set of improving rules that yield the same expected improvement

in any environment, any state and any round the individual will choose one (if it exists) that never imitates actions that achieved a lower payoff.

Given the above assumptions we proceed to search for a most preferable rule for the individual. The main theorem of the paper contains two conditions that completely characterizes the behavior of an agent using an improving updating rule:

i) The agent follows an imitative behavior: either she does not change actions or she adapts the action of the agent she sampled. Especially, an agent that samples another agent using the same action will not change her action.

The second part of the characterization is easiest formulated for the case in which two agents sample each other:

ii) An agent using an improving rule is more likely to switch if she had played the action with the lower payoff than if she had played the one with the higher payoff. Moreover the difference of these switching probabilities is proportional to the absolute difference in the payoffs the actions achieved. The implicit proportionality factor is non negative and may depend on the pair of actions.

What is the intuition behind the above result? Due to the random matching setup an agent can never rule out that a given alternative action has a strictly lower expected payoff. Switching to an action not observed might result in switching to an action not that is not used in the present population. Moreover if each action used in the population achieves the same expected payoff then this switch to an unobserved action may lead to EIP<0. Therefore condition i) must hold.

Intuition for part ii) in the above characterization is easiest to provide using the following lemma. It can be shown that the condition of improving is equivalent to requiring the imitation property (condition i)) together with the following: the individual switches more likely to an action with a higher expected payoff than vice versa. Especially the improving condition does not preclude switching to an action with a lower expected payoff. Notice that even if all the contests were degenerate a switch to a lower payoff may occur when using an improving rule, as long as the agent is more likely to switch to the higher payoff than if the roles of updating agent and sampled agent are reversed. The contests in which an increase in expected payoffs is more difficult to enforce are those in which the payoffs are a priori random because of the possibility of being matched against different actions. The main part of the proof is to show that the linear structure of taking expectations implies that the difference in the switching probabilities (condition ii)) must be linear in order to ensure the improving condition in any environment and any state. Especially this implies that the possible randomness in the payoffs associated to an action causes "imitate if better" to fail to be improving. On the otherhand we obtain that "never switch", "always switch" and that all proportional imitation rules are improving.

The proof of the main theorem reveals that the expected improvement of an improving rule only depends on the proportionality factors implicit in part ii) of the characterization of improving rules. Using the dominance assumption, the individual will therefore choose an improving rule which yields the maximal proportional factor for each pair of actions. Such an updating rule will be referred to as a "best" rule.

There are many "best" rules since only the difference of two switching probabilities (see condition ii)) is uniquely determined. Moreover, in each environment, each state and each round any two "best" rules achieve the same expected improvement. Hence the individual is a priori indifferent between these rules and selects among them using the degenerate improving assumption. The conclusion is that there is a unique "best" improving rule that never imitates actions that achieved a lower payoff. This rule is a specific proportional imitation rule.

An alternative justification of this proportional imitation rule is that it minimizes the probability of switching actions among the "best" improving rules.

Thus we have singled out a unique updating rule for the individual to use in our scenario. Especially, the analysis of our model reveals imitative behavior as an efficient method to utilize limited information.

It should be noted that the information about the payoff of the sampled agent is not needed in order to construct a "best" rule. Consider the updating rule under which the individual reviews her present action with a probability proportional to the difference between the maximal she may obtain (upper end point of I^A (I^B)) and her present payoff. Once she reviews her choice, the individual adapts the action of the agent she sampled. This rule will be referred to as the proportional reviewing rule.??quote whom?? With an appropriate proportionality constant we show that this rule is the unique "best" improving rule that does not depend on the sampled agent's payoff. Hence this rule provides the answer to the question "how to imitate?" in situations in which payoffs of sampled agents can not be observed (see also section 8).

Once we have determined what kind of updating rule each agent might choose, we are interested in the dynamics of the population as a whole. The above analysis singles out a unique rule for each agent (which is the same for agents of the same type). This endogenously justifies the analysis of a population in which agents of the same type use the same rule.

Consider a population in which agents of the same type use the same "best" improving rule, e.g., the "best" proportional imitation rule or the "best" proportional reviewing rule. Assume that the sampling rule is random sampling. Then it is shown that in the short run with high probability the frequencies of the various actions played in large populations evolve approximately according to a discrete version of the continuous replicator dynamics for two type populations (see Taylor [1979]). Especially our process and the continuous replicator dynamics move with high probability approximately in the same direction.

The continuous replicator dynamics are derived from an evolutionary (large population) model of reproduction based on fitness. An analysis of the (gradients of the) continuous replicator dynamics in view of the above result becomes relevant to understanding learning behavior in our model of adaptive agents. Moreover, instead of being a special case, the replicator dynamics are the only relevant short run dynamic adjustment process for our model with a large population.

The rest of the paper is organized as follows. Section 1 contains the basic model. In section 2 the condition of improving is introduced. Section 3 contains an example. In section 4 the main characterization theorem is established. In section 5 the "best" improving rules are characterized. In section 6 the short run population dynamics are analyzed. Section 7 contains an entry and exit scenario to motivate the uniform prior assumption. In section 8 the assumptions of the model together with the related literature are discussed.

1. The Matching and Sampling Scenario

For a finite set R let ΔR be the set of probability distributions on R. Consider the following two person game in which the two players are denoted by A and B. Let $S^A=\{A^1,..,A^{n_A}\}$ be the set of strategies (or actions) of player A and $S^B=\{B^1,..,B^{n_B}\}$ be those of player B. Let $\pi^C(x^A,x^B)$ be the payoff that player C receives when player A uses $x^A\in\Delta S^A$ and player B uses $x^B\in\Delta S^B$, $C\in\{A,B\}$ (so $\pi^C:\Delta S^A\times\Delta S^B\rightarrow\Re$). Let $\Gamma(S^A,S^B,\pi^A,\pi^B)$ denote the underlying two person normal form game. Let $I^C=[\kappa^C,\zeta^C]\subseteq\Re$ such that $\kappa^C<\zeta^C$, C=A,B.

Consider a finite population W consisting of an equal number of two different types of agents A and B. Let W^A and W^B denote the set of all agents of type A and type B respectively in the population and let $N=|W^A|=|W^B|$. We will identify the population shares in W^A and W^B with the probability distributions in ΔW^A and ΔW^B that are associated with randomly selecting an agent from W^A and W^B respectively and observing her strategy.

We will consider the following dynamic process of matching, sampling and updating in the population $W=W^A\cup W^B$.

In each round 1, 2, etc. each agent of type A is randomly matched with exactly one agent of type B. Each agent of type C is endowed with a strategy in S^C she uses when she is matched, C=A,B. When matched an agent of type C receives a payoff according to π^C but neither observes the strategy nor the payoff of her opponent.

A state of the population in a given round is the description of the strategy that each agent is using. Formally, $s:W^A\cup W^B\rightarrow S^A\cup S^B$ with $s(c)\in S^C$ when $c\in W^C$, $C\in\{A,B\}$ is a state of the population where s(c) is the strategy agent 'c' plays in round t. The event that s is the state in round t will be denoted by $\{s^t=s\}$ ($t\in\mathbb{N}$). Given such a state s, k^C_i will denote the number of agents in W^C playing the strategy C^i and $x^C\in\Re^{n_C}$ will denote the population shares of the population W^C, i.e., $x^C_i=\frac{k^C_i}{N}$, $1\leq i\leq n_C$. Given this notation, the expected payoff of an agent 'a' of type A using A^i

(i.e., $s(a)=A^i$) in state s is $\pi^A(A^i,x^B)=\sum_{j=1}^{n_B} \frac{k^B_j}{N} \pi^A(A^i,B^j)$.

Between matching rounds each agent samples an agent of her own type and receives the following information. When agent 'c' samples agent 'd' ($c,d \in W^C$, $c \neq d$) then agent 'c' observes the strategy 'd' used and the payoff 'd' achieved in the last round without observing the identity of 'd'. For each agent 'c' of type C this sampling occurs according to some exogenously given probability distribution $z_c \in \Delta\{W^C \backslash \{c\}\}$ called a sampling rule for agent 'c'. Thereby, $z_c(d)$ is the probability that agent 'c' samples agent 'd'.

The assignment of a sampling rule to each agent in the population will be called a sampling procedure. Formally, $z=(z_c)_{c \in W^A \cup W^B}$ is called a sampling procedure if $c \in W^C$ implies that z_c is a sampling rule for agent 'c' of type C ($C \in \{A,B\}$). We will call the sampling procedure z symmetric if for any $c,d \in W^C$ the probability of 'c' sampling 'd' (this event denoted by c*s*d) is the same as vice versa, i.e., P(c*s*d)=P(d*s*c).

It will be assumed that in each round both matching and sampling are independent of all previous events.

The above conditions restrict the variety of individual sampling procedures without specifying explicitly how the sampling rules of different agents relate to each other. A model in which each agent is sampled at most once (due to time constraints) is equally feasible as one in which an agent can be sampled multiple times. The sampling could be such that agents sample independently, i.e., P(c*s*d|d*s*c)=P(c*s*d) (referred to later as one-sided sampling). Similarly we allow for a model in which agents sample from each other (referred to as two-sided sampling). In this case c*s*d is the same event as d*s*c for each $c,d \in W^C$ which means P(c*s*d|d*s*c)=1. Notice that two-sided sampling implies that the sampling rule is symmetric.

We will now present some examples of symmetric sampling procedures. The situation in which each agent randomly samples an agent (with equal probability) among the agents her type (except for herself) is a symmetric sampling procedure and will be referred to as random sampling

Here we have $P(csd)=\frac{1}{N-1}$ for $c,d\in W^C$ and $c\neq d$.

Another example of a symmetric sampling procedure is the following. Imagine that agents of type A are located on a circle. Assume that agent 'a' of type A randomly samples (one or two-sided) with equal probability among his 2m closest neighbors (m to the left, m to the right, $m<N/2$). This is a sampling rule for agent 'a'. Moreover if each agent of type A uses such a rule and the agents of type B employ a similar rule then we obtain a symmetric sampling procedure.

2. Updating Rules

We now consider an individual that wants to determine a rule for updating her play between rounds in the matching and sampling scenario introduced in the previous section. The individual has a randomizing device available to her that generates independent events. Following the assumptions made in the introduction such an updating rule is characterized as follows. An updating rule for an individual (agent) of type C is formally a (random) function $F^C: S^C \times I^C \times S^C \times I^C \to \Delta S^C$ where $F^C(C^i,\chi,C^j,\psi)_r$ is the probability of playing C^r in the next round after previously playing C^i, receiving the payoff χ and sampling an agent who used C^j and received ψ, $i,j,r \in \{1,..,n_C\}$, $\chi,\psi \in I^C$.

One of the simplest updating rules is the rule "never switch", formally defined by $F^C(C^i,\chi,C^j,\psi)_i=1$ for $C^i,C^j \in S^C$ and $\chi,\psi \in I^C$. A seemingly opposite updating rule is the self explanatory rule "always switch". A more plausible rule seems to be the following rule we will refer to as "imitate if better", i.e., $F^C(C^i,\chi,C^j,\psi)_j=1$ if $\psi>\chi$ and $F^C(C^i,\chi,C^j,\psi)_i=1$ if $\psi \le \chi$. The above rules belong to a class of updating rules that are based on imitation, i.e., either the individual does not switch strategies or she switches to the strategy of the sampled agent. Updating rules with this property will be called imitating. Formally, an updating rule F^C for type C is called imitating if $F^C(C^i,\chi,C^j,\psi)_j+F^C(C^i,\chi,C^j,\psi)_i=1$ for all $C^i,C^j \in S^C$ and $\chi,\psi \in I^C$.

The following class of imitating rules, referred to as proportional imitation rules, will play an important role in our analysis. Using such a rule, the agent never imitates the strategy sampled if it achieved a lower payoff. Moreover, the imitation of a strategy that achieved a higher payoff occurs with probability proportional to the difference in the payoffs. Formally, the updating rule F^C is called a proportional imitation rule (with rate σ) if F^C is imitating and $F^C(C^i,\chi,C^j,\psi)_j=\sigma[\psi-\chi]_+$ for some $0<\sigma \le 1/(\zeta^C-\kappa^C)$, where $[\chi]_+=\chi$ when $\chi>0$ and $[\chi]_+=0$ otherwise.[4]

It is assumed that the individual chooses an updating rule that is expected to increase her

[4] A reference to agents using such a rule is given in Cabrales [1993].

payoffs from one round to the next in any round of the matching and sampling scenario. This should hold for any normal form game $\Gamma(S^A,S^B,\pi^A,\pi^B)$ with payoffs π^A in I^A and π^B in I^B, for any population size and for any state. Moreover the expected payoffs in the next round are calculated under the ignorance assumption and the uniform prior assumption from the introduction. Such an updating rule with be called improving. In the following we will formalize this condition.

Fix an updating rule F^A for type A. Let $s:W^A\cup W^B\rightarrow S^A\cup S^B$ be a state of the population. Then the probability that agent 'a' of type A using F^A plays A^j in round t+1 given that the state at time t is s, is determined by

$$P(s^{t+1}(a)=A^j|s^t=s)=\sum_i \sum_{\{c\neq a,\ s(c)=A^i\}} P(asc)\sum_{q,r=1}^{n_B}\frac{k^B_q}{N}\frac{k^B_{rq}}{N-1}F^A(s(a),\pi^A(s(a),B^q),A^i,\pi^A(A^i,B^r))_j$$

$$\text{where } k^B_{rq}=k^B_r \text{ if } r\neq q \text{ and } k^B_{qq}=k^B_q-1. \qquad (1)$$

It follows that the expected increase of the payoffs of agent 'a' between rounds t and t+1, assuming that no agent of type B changes her strategy is

$$EIP(a)=\Big[\sum_{j=1}^{n_A} P(s^{t+1}(a)=A^j|s^t=s)\pi^A(A^j,x^B)\Big]-\pi^A(s(a),x^B).$$

We are now able to evaluate the performance of the rule F^A for an individual of type A in state s prior to her knowing which agent of type A in W^A she is associated to in s. Under the uniform prior and ignorance assumptions her expected increase in payoffs (denoted by EIP^A) is $EIP^A=\frac{1}{N}\sum_{a\in W^A} EIP(a)$. EIP^A will be called the <u>expected improvement of type A using F^A in state s</u>.

We are now able to formalize the improving condition. Let $\{W_N\}_N$ be a sequence of populations and $\{z_N\}_N$ a sequence of symmetric sampling procedures such that $|W_N^A|=|W_N^B|=N$ and z_N is the sampling procedure associated with $W^A_N\cup W^B_N$. Let G be the set of normal form games $\Gamma(S^A,S^B,\pi^A,\pi^B)$ with payoffs $\pi^C(A^i,B^j)\in I^C$ for all $A^i\in S^A$, $B^j\in S^B$ and $C\in\{A,B\}$. We will call the updating rule $F^C:S^C\times I^C\times S^C\times I^C\rightarrow\Delta S^C$ <u>improving (for type C)</u> ($C\in\{A,B\}$) if for any N, any game $\Gamma(S^A,S^B,\pi^A,\pi^B)\in G$ and any state s ($s:W^A_N\cup W^B_N\rightarrow S^A\cup S^B$) the expected improvement of type C using F^C in state s under the sampling procedure z_N is non negative.

Clearly the rule "never switch" is improving for either type. Our aim is to investigate whether the individual can select among these rules given her preferences we characterized in the introduction. In order to simplify notation we will focus whenever possible on individuals of type A.

3. Examples

In order to clarify the setup of the two previous sections we now consider the performance of two seemingly plausible updating rules by means of an example. Both of these rules will fail to be improving because they can not cope with the incomplete information implicit in the model. For demonstration purposes we will consider the case where populations W^A and W^B are infinite. Consider the matching and sampling scenario of section one with random sampling.

Let $\Gamma(S^A,S^B,\pi^A,\pi^B)$ be the symmetric game with $|S^A|=|S^B|=2$ and payoffs represented in table I where $\psi>3$ and $\alpha,\beta,\delta,\gamma\in\Re$.

Table I: An asymmetric game with two strategies for each type.

	B^1	B^2
A^1	ψ, α	0, β
A^2	2, γ	1, δ

For a given state s of the population let $x^A\in\Delta S^A$ ($x^B\in\Delta S^B$) be the population shares in population W^A (W^B).

3.1 "Imitate if Better"

In the following we will calculate the expected improvement of type A associated with the updating rule "imitate if better". Let s be a state and $a,c\in W^A$ be agents such that $s(a)=A^1$, $s(c)=A^2$. The individual - if in the position of agent 'a' - will only switch from A^1 to A^2 if she is matched against B^2 and then samples an agent using A^2. Consequently, $P(s^{t+1}(a)=A^2|s^t=s)=x^A{}_2x^B{}_2$ and

$EIP(a)=x^A{}_2 x^B{}_2(1+(1-\psi)x^B{}_1)$. Similarly, were the individual in the position of agent 'c' then she would switch to A^1 if she samples an agent using A^1 that was matched against an agent of type B using B^1. Hence, $P(s^{t+1}(c)=A^1|s^t=s)=x^A{}_1 x^B{}_1$, $EIP(c)=-x^A{}_1 x^B{}_1(1+(1-\psi)x^B{}_1)$ and we obtain $EIP^A=x^A{}_1 x^A{}_2(1+(1-\psi)x^B{}_1)(1-2x^B{}_1)$.

It follows that the expected improvement when using the rule "imitate if better" is negative if $1/(\psi-1)<x^B{}_1<1/2$. Especially "imitate if better" is not improving. Notice that this is of course due to the fact that the outcome of a future contest generated by a given action is not necessarily deterministic.

3.2 The "Best Response" Rule

In order to get a feeling for how easy or difficult it is to satisfy the improving condition we will assume in this example that the individual (of type A) knows that the game in table I is being played. It follows that after sampling the individual implicitly has observed a random sample of two strategies (drawn at random without replacement) played among the agents of type B: the strategy of her own and that of the sampled agent's matched opponent.

Consider the updating rule that emerges when playing a best response to this implicit sample in the game in table I, referred to in the following as the "best response" rule. Since $\psi>3$ an agent will play A^2 in the next round if and only if she implicitly observes B^2 being played twice. It follows that $P(s^{t+1}(a)=A^2|s^t=s)=(x^B{}_2)^2$, $P(s^{t+1}(a')=A^1|s^t=s)=1-(x^B{}_2)^2$, $EIP(a)=(x^B{}_2)^2(1+(1-\psi)x^B{}_1)$, $EIP(a')=-[1-(x^B{}_2)^2](1+(1-\psi)x^B{}_1)$ and $EIP^A=[(x^B{}_2)^2-x^A{}_2](1+(1-\psi)x^B{}_1)$.

Despite the additional information of the individual, her expected improvement can be negative although she uses the best response rule. This happens when there is a sufficiently large proportion of type A agents using the best response to the mean strategy of the type B agents. Of course if the individual could observe the exact frequencies of each strategy played among the type B agents then playing a best response would have been improving.

4. Improving Rules

In the following we will show that an updating rule is improving if and only if it is imitating and it satisfies the following condition. Consider an individual using an imitating rule that samples an agent using a strategy that achieves a higher expected payoff in the present round. Then the individual must be more likely to switch strategies than if the roles of the individual and the sampled agent were reversed.

In order to be able to formalize the above statement we will need the following notation. Given a state $s:W^A \cup W^B \rightarrow S^A \cup S^B$ and $i,j,q \in \{1,..,n_A\}$, if A^i and A^j are being played among the type A agents (i.e., $k^A_i k^A_j > 0$) then let $r(A^i,A^j)_q$ denote the probability of playing A^q in the next round when playing A^i and sampling an agent using A^j. In all other cases(i.e., if $k^A_i = 0$ or $k^A_j = 0$) then let $r(A^i,A^j)_q = 0$.

LEMMA 1:

Let $|S^B| \geq 2$, let $\{W_N\}_N$ be a sequence of populations and let $\{z_N\}_N$ be a sequence of symmetric sampling procedures such that $|W_N{}^A| = |W_N{}^B| = N$ and z_N is used in $W_N = W_N{}^A \cup W_N{}^B$. Then the updating rule F^A for type A is improving if and only if i) F^A is imitating and ii) for any $1 \leq i,j \leq n_A$ and any state s with $k^A_i k^A_j > 0$, $r(A^i,A^j)_j \geq r(A^j,A^i)_i$ if and only if $\pi^A(A^j,x^B) \geq \pi^A(A^i,x^B)$.

The proof of the imitation property is quite intuitive. An agent will avoid to play a strategy that she did not observe since it might be that the strategy observed is a duplification of her own strategy whereas all strategies not observed lead necessarily to the worst outcome. Notice that imitation remains necessary to ensure the improving condition even after the event of receiving the lowest possible payoff κ^A and sampling an agent who used the same action and also obtained κ^A. This is because it may be that obtaining κ^A is an unlucky event for the own strategy in the current population but is the only outcome for any other strategy.

PROOF:

We will first show that an improving rule is imitating. Assume that F^A is improving.

Let $\chi,\psi \in I^A=[\kappa^A,\zeta^A]$ such that $\chi>\kappa^A$ or $\psi>\kappa^A$ and let N be even. Consider a population in which $k^A_1=N$ and $k^B_1=N/2$. Assume that Γ is such that $\pi^A(A^1,B^1)=\chi$, $\pi^A(A^i,B^j)=\kappa^A$, $\pi^A(A^1,B^r)=\psi$ for all $1<r\leq n_B$, $1\leq j\leq n_B$, and $1<i<n_A$. Since $\pi^A(A^1,x^B)>\pi^A(A^i,x^B)$ for i>1, all agents in W^A are playing the unique best response. Therefore any change of strategy will decrease expected payoffs. Since F^A is improving, it follows that $F^A(A^1,\chi,A^1,\psi)_1=1$.

Consider now a population in which $k^A_1=k^A_2=k^B_1=N/2$. Assume that Γ is such that $\pi^A(A^1,B^1)=\pi^A(A^2,B^1)=\chi$, $\pi^A(A^i,B^j)=\kappa^A$, $\pi^A(A^1,B^r)=\pi^A(A^2,B^r)=\psi$ for all $1<r\leq n_B$, $1\leq j\leq n_B$, and $2<i\leq n_A$. It follows $\pi^A(A^1,x^B)=\pi^A(A^2,x^B)>\pi^A(A^i,x^B)$ for i>2. Therefore any change to a strategy A^i with i>2 will decrease expected payoffs. Since F^A is improving, it follows that $F^A(A^1,\chi,A^2,\psi)_1+F^A(A^1,\chi,A^2,\psi)_2=1$.

The proof for $\chi=\psi=\kappa^A$ follows just like above when replacing ψ by ζ^A. Since the above arguments hold for any A^i and A^j it follows that F^A is imitating.

Let F^A be an imitating updating rule for type A. Then for an agent 'a' of type A,

$$EIP(a)=\sum_i \sum_{\{c\neq a, s(c)=A^i\}} P(asc)r(s(a),A^i)_i[\pi^A(A^i,x^B)-\pi^A(s(a),x^B)].$$

The expected improvement hence is

$$EIP^A=\frac{1}{N}\sum_{i,j}\sum_{\substack{\{a:s(a)=A^i\}\\ \{c:s(c)=A^j\}}} P(asc)r(A^i,A^j)_j[\pi^A(A^j,x^B)-\pi^A(A^i,x^B)].$$

Using the fact that the sampling procedure is symmetric we obtain

$$EIP^A=\Big[\frac{1}{N}\sum_{i,j}\sum_{\substack{\{a:s(a)=A^i\}\\ \{c:s(c)=A^j\}}} P(asc)\Big][r(A^i,A^j)_j-r(A^j,A^i)_i][\pi^A(A^j,x^B)-\pi^A(A^i,x^B)]. \quad (2)$$

Especially the above must hold for $1\leq i,j\leq n_A$ and a state such that $k^A_i+k^A_j=1$. Therefore condition ii) follows.

Moreover (2) shows that i) and ii) are sufficient for F^A to be improving. $\square$

The following theorem constitutes the central result of this paper and gives a complete characterization of the set of all updating rules that are improving.

THEOREM 2:

Let the assumptions of lemma 1 hold. Then the updating rule F^A is improving for type A if and only if i) F^A is imitating and ii) for all $A^i, A^j \in S^A$, $i \neq j$ there exists $0 \leq \sigma_{ij} \leq 1/(\zeta^A - \kappa^A)$ such that $F^A(A^i,\chi,A^j,\psi)_j - F^A(A^j,\psi,A^i,\chi)_i = \sigma_{ij}(\psi - \chi)$ for all $\chi, \psi \in I^A$.

The intuition behind the proof of condition ii) is as follows. The individual will switch from A^i to A^j only when playing A^i and sampling A^j. From lemma 1 it follows that only the difference between the probability of switching and this probability when the roles are reversed influences the expected improvement. Due to the linear structure of taking expectations, it turns out that only linear terms can be factored out to ensure the improving condition under any circumstances.

PROOF:

We will first show that conditions i) and ii) are sufficient. Let F^A be an updating rule for type A satisfying conditions i) and ii).

For $1 \leq i,j \leq n_A$ such that $k^A_i k^A_j > 0$ we obtain

$$r(A^i,A^j)_j = \sum_{q,r=1}^{n_B} \frac{k^B_q}{N} \frac{k^B_{rq}}{N-1} F^A(A^i,\pi^A(A^i,B^q),A^j,\pi^A(A^j,B^r))_j, \tag{3}$$

where $k^B_{rq} = k^B_r$ if $r \neq q$ and $k^B_{qq} = k^B_q - 1$.

It easily follows with (3) and condition ii) that

$$r(A^i,A^j)_j - r(A^j,A^i)_i = \sigma_{ij}[\pi^A(A^j,x^B) - \pi^A(A^i,x^B)]. \tag{4}$$

From (2) and (4) we obtain

$$EIP^A = \sum_{i<j} \Big[\frac{1}{N} \sum_{\substack{a:s(a)=A^i \\ c:s(c)=A^j}} P(asc) \Big] \sigma_{ij}[\pi^A(A^j,x^B) - \pi^A(A^i,x^B)]^2. \tag{5}$$

Therefore $EIP^A \geq 0$ and hence F^A is improving.

Let F^A be improving. Then lemma 1 implies that F^A is imitating. Therefore it is enough to show the necessity of ii) in populations in which type A uses either A^1 or A^2 and type B uses either B^1 or B^2, i.e., $k^A{}_1+k^A{}_2=N$ and $k^B{}_1+k^B{}_2=N$. To simplify notation, let $\alpha=\pi^A(A^1,B^1)$, $\beta=\pi^A(A^2,B^1)$, $\gamma=\pi^A(A^1,B^2)$, $\delta=\pi^A(A^2,B^2)$, $k^A=k^A{}_1$ and $k^B=k^B{}_1$.

For $\chi,\psi\in I^A$ let $h(\chi,\psi)=F^A(A^2,\chi,A^1,\psi)_1-F^A(A^1,\psi,A^2,\chi)_2$. Consider $1\le k^A\le N-1$.Using

$$V(k^B,N)=\frac{k^B}{N}\frac{k^B-1}{N-1}h(\beta,\alpha)+\frac{k^B}{N}\frac{N-k^B}{N-1}[h(\beta,\gamma)+h(\delta,\alpha)]+\frac{N-k^B}{N}\frac{N-k^B-1}{N-1}h(\delta,\gamma), \tag{6}$$

it is easily verified that

$$V(k^B,N)=r(A^2,A^1)_1-r(A^1,A^2)_2. \tag{7}$$

Let $\alpha>\beta$ and $\delta>\gamma$. With $\mu^*=\frac{\delta-\gamma}{\alpha-\beta+\delta-\gamma}$ it follows (with a slight abuse of notation) that

$\pi^A(A^1,\mu B^1+(1-\mu)B^2)>\pi^A(A^2,\mu B^1+(1-\mu)B^2)$ if and only if $\mu>\mu^*$. Therefore, if $\frac{k^B}{N}>\mu^*$ then from

lemma 1 and (7) it follows that $V(k^B,N)\ge 0$. This must hold for all N and $0\le k^B\le N$. Consider a sequence $\{k^B{}_N\}_N$ such that $0\le k^B\le N$ and $\frac{k^B}{N}>\mu^*$ for all N and $\frac{k^B}{N}\to\mu^*$ as $N\to\infty$. It follows that

$$\mu^{*2}h(\beta,\alpha)+\mu^*(1-\mu^*)[h(\beta,\gamma)+h(\delta,\alpha)]+(1-\mu^*)^2h(\delta,\gamma)\ge 0. \tag{8}$$

Similarly when $\frac{k^B}{N}<\mu^*$ then $V(k^B,N)\le 0$. Consider now a sequence $\{k^B{}_N\}_N$ such that $0\le k^B\le N$ and

$\frac{k^B}{N}<\mu^*$ for all N and $\frac{k^B}{N}\to\mu^*$ as $N\to\infty$. Then

$$\mu^{*2}h(\beta,\alpha)+\mu^*(1-\mu^*)[h(\beta,\gamma)+h(\delta,\alpha)]+(1-\mu^*)^2h(\delta,\gamma)\le 0. \tag{9}$$

Using (8), (9) and the definition of μ^* we obtain

$$(\delta-\gamma)^2h(\beta,\alpha)+(\alpha-\beta)(\delta-\gamma)[h(\beta,\gamma)+h(\delta,\alpha)]+(\alpha-\beta)^2h(\delta,\gamma)=0. \tag{10}$$

When $\beta=\delta$, (10) implies

$$(\alpha-\beta)h(\beta,\gamma)=(\gamma-\beta)h(\beta,\alpha). \tag{11}$$

The right hand side in (11) is a polynome in γ, therefore the left hand side is too, i.e., there exist $u(\beta)$ and $v(\beta)$ such that $h(\beta,\gamma)=u(\beta)\gamma+v(\beta)$. Moreover, comparing the coefficients of γ it follows that $h(\beta,\alpha)=u(\beta)(\alpha-\beta)$. Similarly replacing $u(\beta)\alpha+v(\beta)$ for $h(\beta,\alpha)$ in (11) we obtain $h(\beta,\gamma)=u(\beta)(\gamma-\beta)$. It follows that

$h(\beta,\rho)=u(\beta)(\rho-\beta)$ for all $\kappa^A\leq\rho\leq\zeta^A$ and $\kappa^A<\beta<\zeta^A$ such that $\beta\neq\rho$. (12)

Setting $\alpha=\gamma$ in (10) it follows that $(\delta-\alpha)h(\beta,\alpha)=(\beta-\alpha)h(\delta,\alpha)$. Consequently there exists $w(\alpha)$ such that
$h(\rho,\alpha)=w(\alpha)(\rho-\alpha)$ for all $\kappa^A\leq\rho\leq\zeta^A$ and $\kappa^A<\alpha<\zeta^A$ such that $\alpha\neq\rho$. (13)

From (12) and (13) we obtain that there exists $\sigma\in\Re$ such that $h(\chi,\psi)=\sigma(\psi-\chi)$ for all $\kappa^A<\chi<\zeta^A$, $\kappa^A<\psi<\zeta^A$ such that $\chi\neq\psi$. (14)

Using (10) the statement in (14) can be extended to the case $\chi=\psi$ and to the border values of I^A. From (7) it now follows that
$r(A^2,A^1)_1-r(A^1,A^2)_2=\sigma[\pi^A(A^1,x^B)-\pi^A(A^2,x^B)]$. (15)

Finally with lemma 1 and the fact that $F^A(A^i,\chi,A^j,\psi)_j\in[-1,1]$ we obtain that $0\leq\sigma\leq 1/(\zeta^A-\kappa^A)$. □

As seen above, the characterization of updating rules that are improving is independent of the exact features of the sequence of symmetric sampling rules $\{z_N\}_N$. From condition ii) in theorem 2 it follows that the boundedness of the set of feasible payoffs I^A is a necessary condition for improving rules for type A to exist with $\sigma_{ij}>0$ for some $A^i,A^j\in S^A$.

The following corollary supplements the characterization of improving rules given in theorem 2.

COROLLARY 3:

Condition ii) in theorem 2 holds if and only if the following condition holds:
ii') if $i\neq j$ then either $\sigma_{ij}=0$ and $F^A(A^i,\chi,A^j,\psi)=A^j)=F^A(A^j,\psi,A^i,\chi)=A^i)$ for all $\chi,\psi\in I^A$ or there exists $\sigma_{ij}>0$ and a function $g_{ij}:I^A\times I^A\rightarrow\Re$ such that $-\min\{\chi,\psi\}\leq g_{ij}(\chi,\psi)\leq-\max\{\chi,\psi\}+1/\sigma_{ij}$, $F^A(A^i,\chi,A^j,\psi)_j=\sigma_{ij}(\psi+g_{ij}(\chi,\psi))$ and $F^A(A^j,\chi,A^i,\psi)_i=\sigma_{ij}(\psi+g_{ij}(\psi,\chi))$ holds for all $\chi,\psi\in I^A$.

PROOF:

It is easy to see that ii') implies ii).

Conversely let $A^i, A^j \in S^A$, $i \neq j$ and let F^A satisfy ii). If $\sigma_{ij}=0$ then $F^A(A^i,\chi,A^j,\psi)_j = F^A(A^j,\psi,A^i,\chi)_i$ for all $\chi,\psi \in I^A$ which together with (15) implies ii').

Assume now that $\sigma_{ij}>0$. Let $g_{ij}(\chi,\psi)=[F^A(A^i,\chi,A^j,\psi)_j/\sigma_{ij}]-\psi$. It follows that $-\psi \leq g_{ij}(\chi,\psi) \leq -\psi+1/\sigma_{ij}$ and $F^A(A^i,\chi,A^j,\psi)_j=\sigma_{ij}(\psi+g_{ij}(\chi,\psi))$. Together with ii) we obtain $F^A(A^j,\psi,A^i,\chi)_i=F^A(A^i,\chi,A^j,\psi)_j-\sigma_{ij}(\psi-\chi)=\sigma_{ij}(\chi+g_{ij}(\chi,\psi))$. This implies $-\chi \leq g_{ij}(\chi,\psi) \leq -\chi+1/\sigma_{ij}$ which completes the proof of condition ii'). ☐

5. Maximal Improvement

In this section we will select among the improving rules ones with special properties.

From (5) it follows that the expected improvement of an improving rule only depends on the values σ_{ij}, $i,j\in\{1,..,n_A\}$, $i\neq j$. If $\sigma_{ij}=0$ for some $i\neq j$ (i.e, $F^A(A^i,\chi,A^j,\psi)_j=F^A(A^j,\psi,A^i,\chi)_i$ for all $\chi,\psi\in I^A$) then the adjustment behavior when using A^i and sampling A^j or vice versa does not contribute to improving expected payoffs. Especially, if $k^A_i+k^A_j=N$ then $EIP^A=0$. On the other hand, $\sigma_{ij}>0$ induces a strictly positive expected improvement given that agents using A^i sample ones using A^j with positive probability and A^i and A^j achieve different expected payoffs.

Notice that maximizing expected improvement is equivalent to maximizing payoffs in the next round. It follows immediately that since (5) is increasing in σ_{ij}, the improving rules that perform "best" (i.e., obtain the maximal expected payoffs in the next round among the improving rules) are precisely the ones with maximal σ_{ij}. Among these rules we single out a unique rule. The proportional imitation rule with appropriate adjustment rate is the unique "best" rule that never adapts strategies that achieved lower payoffs. Therefore it is most preferred among the "best" rules anticipating that payoffs might be deterministic (and evaluating this situation according to the worst circumstances). An alternative justification is that it is the unique rule that minimizes the probability of switching strategies among the "best" rules.

THEOREM 4:

Under the assumptions of lemma 1, let *M* be the set of improving rules for type A with the following property. If $F\in M$ then there is no improving rule F' such that in some environment and some state F' performs better than F, i.e., EIP(F')>EIP(F). Then

i) $F^A\in M$ if and only if F^A is an improving rule and $\sigma_{ij}=1/(\zeta^A-\kappa^A)$ for $A^i,A^j\in S^A$, $i\neq j$ where σ_{ij} is given by theorem 2.

ii) the proportional imitation rule with rate $1/(\zeta^A-\kappa^A)$ is the unique rule in *M* that never imitates a strategy that achieved a lower payoff.

iii) the proportional imitation rule with rate $1/(\zeta^A-\kappa^A)$ is the unique rule that minimizes the probability of switching among the set of all rules in *M*.

PROOF:

With $I^A=[\kappa^A,\zeta^A]$ it follows from theorem 2 that $\sigma_{ij}\leq 1/(\zeta^A-\kappa^A)$. Moreover the proportional imitation rule with rate $1/(\zeta^A-\kappa^A)$ is well defined. Therefore any improving rule with $\sigma_{ij}=1/(\zeta^A-\kappa^A)$ maximizes the expected improvement among all improving rules.

Statements ii) and iii) follow easily from corollary 2 since the proportional imitation rule is the unique rule in *M* with $g_{ij}(\chi,\psi)=-\min\{\chi,\psi\}$.□

In the following we will present the "best" improving rule that requires minimal information. For $0<\sigma\leq 1/(\zeta^C-\kappa^C)$ consider the updating rule F^A defined by $F^A(A^i,\chi,A^j,\psi)_j=\sigma(\zeta^A-\chi)$ for all $A^i,A^j\in S^A$ and $\chi,\psi\in I^A$. We will call F^A the <u>proportional reviewing rule with rate σ</u>. The proportional reviewing rule with rate $1/(\zeta^A-\kappa^A)$ is used in Binmore, Gale and Samuelson [1993] who interpret it on the basis of random aspiration levels. In each round an aspiration level is chosen from a uniform distribution on the set of feasible payoffs $[\kappa^A,\zeta^A]$. If the individual obtains a payoff below her current aspiration level then she samples an agents and switches to the sampled agent's strategy.

The following theorem characterizes a unique rule for the individual to use when the payoff of the sampled agent is not observable.

THEOREM 5:

Under the assumptions of theorem 4, the proportional reviewing rule with rate $1/(\zeta^A-\kappa^A)$ is the unique rule in *M* that is independent of the sampled agent's payoff.

PROOF:

Assume that $F^A\in M$ does not depend on the sampled agent's payoff. For $A^i,A^j\in S^A$, $i\neq j$ let $f_{ij}(x)=F^A(A^i,x,A^j,y)$. From theorems 2 and 4 it follows that $f_{ij}(x)+\sigma x=f_{ji}(y)+\sigma y$ where $\sigma=1/(\zeta^A-\kappa^A)$. Therefore there exists $c\in\Re$ such that $f_{ij}(x)=c-\sigma x$. Since $f_{ij}(x)\in[0,1]$, $c=\zeta^A/(\zeta^A-\kappa^A)$. Hence F^A is the proportional reviewing rule with rate $1/(\zeta^A-\kappa^A)$.□

6. Short Run Behavior of the Population

In the previous sections we analyzed which updating rules an individual might choose in a specific matching and sampling scenario. Especially we assumed that this choice was made ignoring any possible decisions by other agents. As a result we obtained a unique "best" rule (see theorem 4). Consequently, given that individuals of the same type have identical preferences, agents of the same type would choose independently of each other the same rule. In this section we will analyze how this individual choice of rules effects the evolution of the distribution of strategies played in the two sub populations. Notice that the common a priori assumption that all agents use the same rule is here justified on an individual level.

In the following we will assume that each agent uses the same "best" rule given by part i) of theorem 4. In this context let $\sigma^C=1/(\zeta^C-\kappa^C)$, C=A,B. Especially the agents might be using the proportional imitation rule singled out by theorem 4 or the proportional reviewing rule (see theorem 5). We will analyze the short run dynamic evolution of the frequencies of the strategies played when the population is large. There will be a strong connection to the replicator dynamics in two type populations (see Taylor [1979]) when agents employ random sampling.

Assume that the population $W^A \cup W^B$ is in state s at time t ($t \in \mathbb{N}$). Let z be a symmetric sampling rule. What is the expected number of agents of type C playing C^i in round t+1, denoted by $Ek^{C'}_i$? Using symmetry of sampling,

$$Ek^{C'}_i=k^C_i+\sum_{\{u:u\neq i\}}\ \sum_{\substack{\{c:s(c)=C^i\}\\ \{d:s(d)=C^u\}}} P(csd)[r(C^u,C^i)_i-r(C^i,C^u)_u],\ 1\le i\le n_A.$$

Using (4) we obtain

$$Ek^{A'}_i=k^A_i+\sigma^A\sum_{\{u:u\neq i\}}\ \sum_{\substack{\{a:s(a)=A^i\}\\ \{c:s(c)=A^u\}}} P(asc)[\pi^A(A^i,x^B)-\pi^A(A^u,x^B)],\ i=1,..,n_A,$$

and similarly

$$Ek^{B'}_j=k^B_j+\sigma^B\sum_{\{r:r\neq j\}}\ \sum_{\substack{\{b:s(b)=B^j\}\\ \{d:s(d)=B^r\}}} P(bsd)[\pi^B(x^A,B^j)-\pi^B(x^A,B^r)],\ j=1,..,n_B.$$

For the rest of this section we will consider the special case where z is a random sampling procedure. Letting $Ex^{C\prime}_{i}$ denote the expected frequency of C^i in round t+1 ($1 \leq i \leq n_C$), i.e., $Ex^{C\prime}_{i}=Ek^{C\prime}_{i}/N$, we obtain

$$Ex^{A\prime}_{i}=x^{A}_{i}+\sigma^{A}\frac{N}{N\;1}\sum_{r:r\neq i}[\pi^{A}(A^{i},x^{B})-\pi^{A}(A^{r},x^{B})]x^{A}_{i}x^{A}_{r} \text{ and hence}$$

$$Ex^{A\prime}_{i}=x^{A}_{i}+\sigma^{A}\frac{N}{N\;1}[\pi^{A}(A^{i},x^{B})-\pi^{A}(x^{A},x^{B})]x^{A}_{i},\ i=1,..,n_{A} \text{ and}$$

$$Ex^{B\prime}_{j}=x^{B}_{j}+\sigma^{B}\frac{N}{N\;1}[\pi^{B}(x^{A},B^{j})-\pi^{B}(x^{A},x^{B})]x^{B}_{j},\ j=1,..,n_{B}. \tag{16}$$

Notice for a given game as the range of the payoffs for type C (measured by $\zeta^{C}-\kappa^{C}=1/\sigma^{C}$) increases, strategies among agents of type C are expected to change at a slower rate.

Now consider the following deterministic adjustment process in discrete time defined on $\Delta S^{A}\times\Delta S^{B}$:

$$y^{A,0}=\overline{y^{A}},\ y^{B,0}=\overline{y^{B}},$$

$$y^{A,t+1}_{i}=y^{A,t}_{i}+\sigma^{A}[\pi^{A}(A^{i},y^{B,t})-\pi^{A}(y^{A,t},y^{B,t})]y^{A,t}_{i},\ i=1,..,n_{A},$$

$$y^{B,t+1}_{j}=y^{B,t}_{j}+\sigma^{B}[\pi^{B}(y^{A,t},B^{j})-\pi^{B}(y^{A,t},y^{B,t})]y^{B,t}_{j},\ j=1,..,n_{B},\ t=0,1,2,..,$$

where $(\overline{y^{A}},\overline{y^{B}})\in\Delta S^{A}\times\Delta S^{B}$ is the initial state. (17)

The following continuous time version of (17) was first established by Taylor [1979]. The continuous replicator dynamics for two type populations matched to play the game $\Gamma(S^{A},S^{B},E^{A},E^{B})$ (see Taylor [1979]) are defined as follows:

$$z^{A}(0)=\overline{z^{A}},\ z^{B}(0)=\overline{z^{B}},$$

$$\frac{d}{dt}z^{A}_{i}=[E^{A}(A^{i},z^{B})-E^{A}(z^{A},z^{B})]z^{A}_{i},\ i=1,..,n_{A} \text{ and}$$

$$\frac{d}{dt}z^{B}_{j}=[E^{B}(z^{A},B^{j})-E^{B}(z^{A},z^{B})]z^{B}_{j},\ j=1,..,n_{B},\ t\geq 0 \tag{18}$$

where $(\overline{z^{A}},\overline{z^{B}})\in\Delta S^{A}\times\Delta S^{B}$ is the initial state, $z^{C}=z^{C}(t)\in\Delta S^{C}$ is the mean strategy and z^{C}_{i} is the proportion of agents using $C^{i}\in S^{C}$ among the agents of type C in the population at time t (C=A,B).

The continuous replicator dynamics are derived from applying a "law of large numbers" argument to a large population random matching and reproduction scenario. It is easily shown that the trajectories of (17) approximate the trajectories of (18) for finite time horizons when σ^A and σ^B are small and $E^C()=\sigma^C\pi^C()$, C=A,B.

We are interested in the relationship between our (stochastic) process and the deterministic process given in (17), more specifically if a "law of large numbers" type of result applies when the population is large but finite.

Given random sampling it is sufficient to specify a population state by the frequencies of the strategies that are present. Therefore we will identify states s with the associated population shares (x^A,x^B). Given $N\in\mathbb{N}$ let $\Delta S^C_N=\{x\in\Delta S^C \text{ s.t. } Nx_i\in\mathbb{N} \text{ for } 1\le i\le n_C\}$ be the set of feasible population shares in the population W^C of size N.

Let $\|..\|$ be the supremums norm on $\Delta S^A\times\Delta S^B$, i.e.,
$\|(x^A,x^B)\|=\max\{|x^A{}_i|,|x^B{}_j|,\ 1\le i\le n_A,\ 1\le j\le n_B\}$.

The following theorem states that in a sufficiently large population under random sampling where agents of the same type use the same "best" improving rule, the frequencies of the strategies evolve in the short run with high probability approximately according to the discrete time version of the replicator dynamics for two type populations given in (17). Especially the trajectories move with high probability approximately in the direction of the continuous replicator dynamics given by (18).

THEOREM 6:

Let the assumptions of lemma 1 hold. Assume for each N that z_N is a random sampling rule. Assume that in each population $W^A_N\cup W^B_N$ agents of the same type use the same "best" improving rule F^C (C=A,B). Then for every $\delta>0$, $\epsilon>0$ and $T\in\mathbb{N}$ there exists an N_0 such that for any $N>N_0$ and any $(\overline{x}^A,\overline{x}^B)\in\Delta S^A_N\times\Delta S^B_N$, $P(\|(x^{A,T}(N),x^{B,T}(N))-(y^{A,T},y^{B,T})\|>\delta)<\epsilon$,

where $(x^{A,t}(N),x^{B,t}(N))$ is the random state of the population (W^A_N,W^B_N) at time t given the initial state $(x^{A,0}(N),x^{B,0}(N))=(\overline{x}^A,\overline{x}^B)$ and $(y^{A,t},y^{B,t})$ evolves according to (17) with $(y^{A,0},y^{B,0})=(\overline{x}^A,\overline{x}^B)$.

The statement of theorem 6 is graphically illustrated in figure 1 where $\tilde{x}(t,N)=(\tilde{x}^A(t,N),\tilde{x}^B(t,N))$ is a "typical" path of the population $W^A \cup W^B$ starting in $x(0)=(x^{A,0}(N),x^{B,0}(N))$.

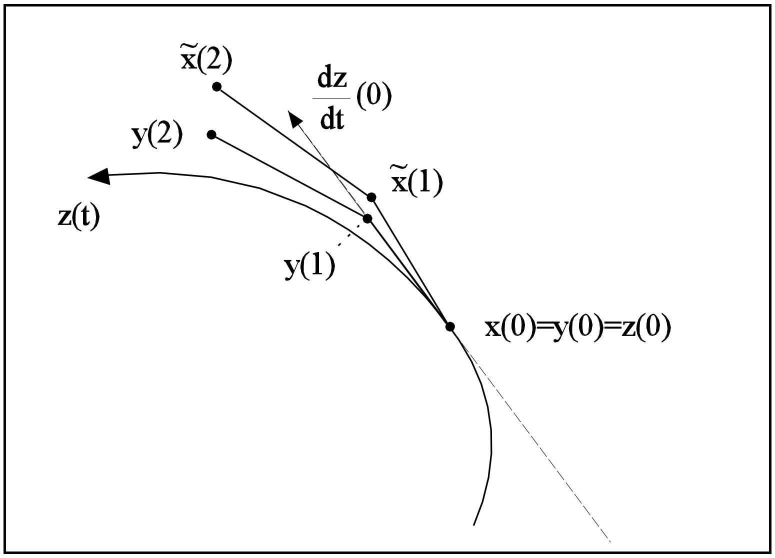

Figure 1: Graphic illustration to theorem 6.

The intuition behind the proof is as follows. We show that the variance of our process from one round to the next is arbitrarily small for sufficiently large populations (the "law of large numbers" effect). Therefore the state at time one will be close to the expected state at time one (which is approximately $(y^{A,1},y^{B,1})$, compare (16) and (17)). Using the fact that trajectories starting close stay close when evolving according to (17), the state of our process at time two will therefore also be close to the trajectory of (17) at time 2, and so on.

PROOF:

Let the assumptions of theorem 6 hold. We will first prove the statement for T=1. For notational simplicity we will often omit N from the notation, e.g., $x^{A,t}=x^{A,t}(N)$.

Fix N. For $1 \le i \le n_A$ we will show that $Var(x^{A,1}_i)$ is of the order 1/N. For $a,c \in W^A_N$ and $b,d \in W^B_N$ let [a,b,c,d] be the event that 'a' is matched against 'b', 'c' is matched against 'd' and 'a' samples 'c'. Let Z(a,b,c,d)=1 if [a,b,c,d] occurs and 'a' plays A^i in the next round, otherwise

Z(a,b,c,d)=0. Then $x^{A,1}_i = \frac{1}{N} \sum_{\substack{a,c \in W^A \\ b,d \in W^B}} Z(a,b,c,d)$.

If P([a,b,c,d])=0 then VarZ(a,b,c,d)=0. If P([a,b,c,d])>0 then $EZ(a,b,c,d) = \frac{1}{N(N-1)^2}\mu$ and VarZ(a,b,c,d)=EZ(a,b,c,d)[1−EZ(a,b,c,d)] where

μ=P(a plays A^i in next round|[a,b,c,d]).

Furthermore

$$Var(x^{A,1}_i) = \frac{1}{N} \sum_{\substack{a,c \in W^A \\ b,d \in W^B}} VarZ(a,b,c,d) + \frac{1}{N^2} \sum_{*} Cov(Z(a,b,c,d),Z(a',b',c',d')) \qquad (19)$$

where the second summation (*) is over all $a,c,a',c' \in W^A_N$ and $b,d,b',d' \in W^B_N$ such that $(a,b,c,d) \neq (a',b',c',d')$.

There are of order N^4 events [a,b,c,d] and for each event the variance VarZ(a,b,c,d) is of order $1/N^3$. Therefore the first summation in (19) is of order 1/N and hence converges to 0 as $N \to \infty$.

We will now investigate the second summation. Since Cov(Z,Z')=EZZ'−EZEZ' and $EZZ' \geq 0$ we will only consider terms in (19) such that EZZ'>0. For $a,c,a',c' \in W^A_N$ and $b,d,b',d' \in W^B_N$ let λ be the probability that [a,b,c,d] and [a',b',c',d'] occur simultaneously. Then $EZZ' \leq \lambda$. Consider the terms in (19) such that (a',b',c',d')=(c,d,a,b). There are of order N^4 such pairs and λ is of order $1/N^3$. Therefore the sum of these terms (including the factor $1/N^2$) converges to 0 as $N \to \infty$. All other terms in the second summation are of the type $\{a,b,c,d\} \cap \{a',b',c',d'\} = \emptyset$. There are of order N^8 such pairs. For these terms we must give a better bound than λ for the covariance. $EZZ' - EZEZ' \leq \frac{1}{N(N-1)^2}\left[\frac{1}{(N-2)^2(N-3)} - \frac{1}{N(N-1)^2}\right]$ which is of order $1/N^7$. It follows that the sum over these pairs is also of order 1/N. We therefore have shown that the second summation in (19) also converges to 0 as N goes to ∞. Hence $Varx^{A,1}_i$ is of order 1/N for each $i \in \{1,..,n_A\}$. Similarly it can be shown for $1 \leq j \leq n_B$ that $Varx^{B,1}_j$ is also of order 1/N.

Fix $\epsilon>0$ and $\delta>0$. Since the above calculations were independent of the initial state $(\overline{x}^A,\overline{x}^B)$ there exists N_0 such that for any $(\overline{x}^A,\overline{x}^B)\in\Delta S^A\times\Delta S^B$ and $N>N_0$, $\sum_i \mathrm{Var}x^{A,1}_i+\sum_j \mathrm{Var}x^{B,1}_j<\epsilon\delta^2/4$ when $(x^{A,0},x^{B,0})=(\overline{x}^A,\overline{x}^B)$.

From (16) and (17) it follows that there exists $N_1\geq N_0$ such that for any $N>N_1$, any initial state $(\overline{x}^A,\overline{x}^B)$, any $1\leq i\leq n_A$ and $1\leq j\leq n_B$, $|Ex^{A,1}_i-y^{A,1}_i|<\delta/2$ and $|Ex^{B,1}_j-y^{B,1}_j|<\delta/2$ when $(x^{A,0},x^{B,0})=(y^{A,0},y^{B,0})=(\overline{x}^A,\overline{x}^B)$. So $P(|x^{A,1}_i-y^{A,1}_i|>\delta)\leq P(|x^{A,1}_i-Ex^{A,1}_i|>\delta/2)$. We now obtain using Tschebyscheff's inequality that for any $N>N_1$,

$$P(\|(x^{A,1},x^{B,1})-(y^{A,1},y^{B,1})\|>\delta)\leq\sum_i P(|x^{A,1}_i-y^{A,1}_i|>\delta)+\sum_j P(|x^{B,1}_j-y^{B,1}_j|>\delta)$$

$$\leq\sum_i P(|x^{A,1}_i-Ex^{A,1}_i|>\delta/2)+\sum_j P(|x^{B,1}_j-Ex^{B,1}_j|>\delta/2)\leq\frac{4}{\delta^2}\sum_i \mathrm{Var}x^{A,1}_i+\frac{4}{\delta^2}\sum_j \mathrm{Var}x^{B,1}_j<\epsilon \text{ when}$$

$(x^{A,0},x^{B,0})=(y^{A,0},y^{B,0})=(\overline{x}^A,\overline{x}^B)$.

This concludes the proof of the theorem when T=1.

We will now show the statement for T=2. Fix $\delta>0$ and $\epsilon>0$. Let $f:\Delta S^A\times\Delta S^B\rightarrow\Delta S^A\times\Delta S^B$ be such that $f=(f^A,f^B)$, $f^A_i(w)=w^A_i+\sigma^A[\pi^A(A^i,w^B)-\pi^A(w^A,w^B)]w^A_i$, $1\leq i\leq n_A$ and $f^B_j(w)=w^B_j+\sigma^B[\pi^B(w^A,B^j)-\pi^B(w^A,w^B)]w^B_j$, $1\leq j\leq n_B$ for $w=(w^A,w^B)\in\Delta S^A\times\Delta S^B$. It follows that $(y^{A,t+1},y^{B,t+1})=f(y^{A,t},y^{B,t})$, $t\geq 0$. Since f is continuous and $\Delta S^A\times\Delta S^B$ is compact, for each $\delta>0$ there exists $\alpha\in(0,\delta/2)$ such that $\|f(w)-f(w^\circ)\|<\delta/2$ if $\|w-w^\circ\|<\alpha$. Let η be such that $(1-\eta)^2=1-\epsilon$. Using the fact that we have proven the theorem for T=1, let N_0 be such that for any $N>N_0$, $P(\|(x^{A,1},x^{B,1})-(y^{A,1},y^{B,1})\|\leq\alpha)\geq 1-\eta$ when $(x^{A,0},x^{B,0})=(y^{A,0},y^{B,0})$. Applying the theorem for T=1 again we obtain

$$P(\|(x^{A,2},x^{B,2})-(y^{A,2},y^{B,2})\|\leq\delta)$$
$$\geq P(\|(x^{A,2},x^{B,2})-f(w)\|\leq\alpha \text{ for some } w\in\Delta S^A\times\Delta S^B \text{ s.t. } \|w-(y^{A,1},y^{B,1})\|\leq\alpha)$$
$$\geq(1-\eta)^2=1-\epsilon.$$

This concludes the proof of statement for T=2. Using induction it is easy to extend the above argument to the case of T>2 and hence this part of the proof is omitted. ☐

7. An Entry and Exit Scenario

In this section we present an entry and exit scenario that gives additional motivation to the assumption made in the introduction that in any round and in any state the individual expects to be equally likely in the position of any agent of her own type, referred to as the uniform prior assumption.

Periodically a new agent (individual) of either type emerges that replaces a randomly selected agent (chosen with equal probability) among the agents of her own type in the population. When entering the population the new agent learns the strategy and the payoff of the agent she replaces. In this setup an entering agent must have besides her updating rule an initial strategy selection rule. This rule determines the strategy the agent plays in the first matching round she takes part in as a function of the strategy and payoff of the agent she replaced.

Given the above entry and exit scenario the decision problem of an agent can be formulated explicitly. Before entering the population each agent must choose an initial strategy selection rule she will use on entry and an updating rule she will use throughout the matching. At the time of this decision she knows her own type but has no other information besides the basic assumptions of the model. Furthermore, the agent ignores the fact that other agents might adjust and chooses a rule that is expected to improve payoffs in each state of the population and each feasible game. On entering the population the increase in expected payoffs is relative to the expected payoff of the agent that is replaced.

Following the above scenario an agent enters equally likely in the position of any agent of her own type. This immediately implies that the unique improving initial strategy selection rule is to adapt the strategy of the replaced agent. This follows just as in lemma 1. Therefore, the uniform prior assumption is correct in the entering agent's first matching round. Moreover, since agents ignore the effect their own rule (and those of others) has on future population distributions, an agent will believe a priori to entering the population that she will be equally likely in the position of any agent of her own type in any state and in any round. This is precisely the uniform prior assumption made in the introduction.

8. Discussion

In this section we will discuss some assumptions made in the model and will refer to the related literature.

At first we will go through the basic informational assumptions to show that they are minimal to ensure that the individual strictly prefers some rule to not switching both in the general model and in the case where all contests are degenerate.

Central to the matching and sampling setup is that an agent learns from other agents that are in the same matching situation. In this respect, we can allow for a general class of sampling rules. In fact the assumptions on the sampling rules can be slightly weakened. One might want to include the case in which agents do not necessarily sample in each round. Given a sampling rule let $\theta^C>0$ be the probability that an agent of type C makes an observation. Let $\{z_N\}_N$ be a sequence of sampling rules such that θ^C is independent of N. Then theorem 2 can easily be adjusted to these more general symmetric sampling rules by adding to condition i) that an agent without a sample does not adjust. Similarly, σ^C must be replaced by $\theta^C\sigma^C$ in (17) in order for theorem 6 to hold for these more general sampling rules.

Observing the sampled agent's payoff is necessary to find a rule that is better than "never switch" when all contests are degenerate. Agents have minimal memory since the updating rules only depend on the action and the payoff of the last period. Updating rules with minimal memory are simple and therefore plausible under complexity restrictions. Additionally, changing payoffs can make rules based solely on information from the last period plausible.

Concerning the game structure, agents have no idea about what kind of game they are playing besides the fact that they know the set of strategies and an interval containing the payoffs. It is necessary to assume that agents know a bounded interval that contains the payoffs (see the note made after the proof of theorem 2). It is not even necessary to assume that the agents in our model are aware of the fact that they are playing a game. Additionally, the "optimal" behavior does not require that agents know the set of strategies. We only need to assume that agents can recognize and adapt strategies of sampled agents. This can for example be applied to a model in which new strategies (e.g., technologies) emerge exogenously in the sense that a few agents enter using these strategies. Whether the new technology takes over or dies out in the short run can be analyzed with the help of the replicator dynamics (see theorem 6). A follow up paper to this one under way will provide the long run analysis.

We make two assumptions that an agent bases her calculations on, namely that she will be the only one to adjust (the ignorance assumption) and that she will be equally likely in the position of any agent in any round and any state (the uniform prior assumption).

The ignorance assumption is correct with high probability if the probability of receiving information between rounds θ^C is small (given the extended model mentioned above in which an agent might not sample in each round). Only agents that receive information will adjust. Therefore, with high probability an agent that receives information will be the only one who received information.

The uniform prior assumption is motivated in section 7 by an entry and exit scenario. A conceivable alternative to this assumption would be to prefer rules that maximize expected payoffs under the worst circumstances given the current choice of strategy. However this condition is too strong and leads to a trivial result. The individual will never change her strategy under such a condition since the possibility that she might be playing a current best response will deter her from ever switching.

Notice that from theorem 2 it follows that we must allow for the individual to use a randomizing device when constructing an updating rule that is better than not switching.

A central assumption driving the results are the preferences of the individual. Maximizing expected payoffs makes sense when the contest (game) yields a reproductive fitness. In the context of decision theory, it may be argued that payoffs should be identified with von Neumann Morgenstern utilities. Either framework however does not contradict the assumption that payoffs are observable. A fitness or utility can be calculated if we assume that the individual observes the outcome of the sampled agent's match. Of course the observability of payoffs (utilities) is not at all necessary for our analysis. In theorem 4 we show that the proportional reviewing rule a "best" rule although it does not depend on the payoff of the sampled agent.

Apart from the uniform prior assumption we chose a "distribution free" approach when selecting the optimal updating rule. Any alternative (plausible) preference relation of the individual invariably leads to making assumptions on the individuals beliefs about the unobservable variables in the model.

We will now refer to some of the related literature.

Kandori, Mailath and Rob [1993] justify individual behavior in some examples in order to motivate the class of dynamics they consider in their random matching model. However the informational assumptions in these examples are quite drastic.

In one of these scenarios an agent's knowledge of the game is limited but each agent is matched against each agent of the opposite type as in a tournament. Kandori, Mailath and Rob [1993] informally propose for the individual to imitate the strategy sampled if and only if it achieved a higher payoff in the previous round. If we were to consider tournament play in our model then this rule "imitate if better" would clearly become the most preferred rule. However, as shown in section 3, "imitate if better" is no longer improving when agents are only matched once.

In another scenario of Kandori, Mailath and Rob [1993] matching is random but agents know which game they are playing. Moreover, between matching rounds each agent that may adjust observes the entire population distribution and therefore learns (given her knowledge of the game) which strategy would have been "best" in the last round. The proposal of Kandori, Mailath and Rob [1993] to play a best response when adjusting is in terms of our analysis the unique most preferred rule. In section 3 we show that this rule however can fail to be improving when the best response is taken to a sample.

Motivated by assumptions on the preferences of an individual we restrict our attention to updating rules that increase expected payoffs under any circumstances. A slightly stronger condition would be to a priori demand that expected payoffs improve strictly whenever possible, i.e., if not all strategies used by agents of the same type achieve the same expected payoff. This condition is closely related to the concept of "absolute expediency" used by Sarin [1993]. Updating rules satisfying this strict improving condition would require additionally to the characterization in theorem 2 for $\sigma_{ij}>0$ for all $A^i,A^j \in S^A$. Sarin [1993] uses absolute expediency to axiomatize Cross' learning rule (see Cross [1973]). This rule applies to a setup in which two players face each other and adapt their mixed action according to the action they take and the payoff they receive. Börgers and Sarin [1993] show that the resulting dynamics when both players use Cross' learning rule approximate the replicator dynamics.

Although related, the general approach of Sarin [1993] is different in spirit from ours because their axioms concern the updating behavior of a player and are not formulated in terms of preferences.

Originally the replicator dynamics were only interesting for the biological literature. Lately various models of individual behavior leading to the replicator dynamics have emerged. Cabrales [1993] previously derived the replicator dynamics using the rule we call the proportional imitation rule. They motivate the rule by idiosyncratic uniformly distributed costs of changing strategies. Binmore, Gale and Samuelson [1993] derive the replicator dynamics (in an infinite population) using the "best" proportional reviewing rule, interpreting it as a rule based on random aspiration levels. Björnerstedt and Weibull [1993] also derive the replicator dynamics with (infinitely many) individuals on average using a proportional reviewing rule.

In the above models, the individual rules seem more or less arbitrary and thus the replicator dynamics too. However, both in our model and in the model used in Sarin [1993] and Börgers and Sarin [1993], the replicator dynamics result from individual behavior determined by preferences or axioms. These papers show that the replicator dynamics have a distinguished role in learning models.

Random matching models with very limited observations to base behavior on can be found in Friedman [1991] in an example and in the experiment of Malawski [1989].

Malawski [1989] sets up a laboratory experiment for the same matching and sampling model as presented in section one (they assume random sampling). There were 5-6 subjects of each type that were randomly matched against each other to play in each of 200 rounds one of three games. The games had disjoint strategy sets. Participants were told in each round their current strategy set and observed the previous strategy and payoff of a randomly sampled player of the same type. They were not aware that they were playing a game. Malawski [1989] relates the data to two specific learning rules. One is the updating rule "imitate if better" (which he refers to as "learning by observing"). Although a definite reaction of the participants to the sample is observed, imitation alone is not able to explain the data. Malawski [1989] prefers to explain the data with a rule that determines behavior according to a constant aspiration level and independent of the data observed in the sample.

Our model is not an attempt to explain behavior in such an experiment. Rules derived from maximizing the outcomes in the worst cases should be understood as a tool to mix with other rules like experimentation or just as a benchmark when the individual is able to gather more information. None-the-less, a thorough analysis of Malawski's [1989] experiments with respect to our findings are of interest and are currently under way.

References:

Banerjee, A. V. (1992). "A Simple Model of Herd Behavior," *Quart. J. Econ.* **107**, 797-818.

Binmore, K. G., Gale, J., Samuelson, L. (1993). "Learning to be Imperfect: the Ultimatum Game," forthcoming in *Games Econ. Beh.*.

Börgers, T., Sarin, R. (1993). "Learning Through Reinforcement and Replicator Dynamics," Disc. Paper No. 93-19, University College of London.

Björnerstedt, J., Weibull, J. (1994). "Nash Equilibrium and Evolution by Imitation," The Industrial Institute for Economic and Social Research, WP 407.

Cabrales, A. (1993). "Stochastic Replicator Dynamics," mimeo, University of California, San Diego.

Cross, J. G. (1973). "A Stochastic Learning Model of Economic Behavior," *Quart. J. Econ.* **87**, 239-266.

Friedman, D. (1991). "Evolutionary Games in Economics," *Econometrica* **59**, 637-666.

Kandori, M., Mailath, G., Rob, R. (1993). "Learning, Mutation and Long Run Equilibria in Games," *Econometrica* **61**, 29-56.

Malawski, M. (1989). *Some Learning Processes in Population Games*, Inaugural-Dissertation, University of Bonn.

Sarin, R. (1993). "An Axiomatization of the Cross Learning Dynamic," Mimeo, University of California, San Diego.

Taylor, P. (1979). "Evolutionarily Stable Strategies With Two Types of Players," *J. Applied Prob.* **16**, 76-83.