WORKING PAPERS

Ana. B. Ania

Learning by Imitation when Playing the Field

DEPARTMENT OF ECONOMICS

UNIVERSITY   OF   VIENNA

Learning by Imitation when Playing the Field

Ana B. Ania[*]

Department of Economics, University of Vienna, Hohenstaufeng. 9, A-1010 Vienna, Austria. Phone: +43 1 4277 37433. Fax: +43 1 4277 9374.
E-mail: ana-begona.ania-martinez@univie.ac.at

We study the properties of learning rules based on imitation in the context of n-player games played among agents within the same population. We find that there are no (non-trivial) rules that increase (average) expected payoffs at each possible state, and for any possible game. The results highlight the complexity of learning by imitation of successful behavior displayed by conspecifics in the presence of strategic considerations within the same population. *Journal of Economic Literature* Classification Numbers: C70, C72.

Key Words: learning, imitation.

1. INTRODUCTION

Informally defined, imitation is a behavioral rule that prescribes to mimic observed behavior—to take an action only if this action has been used before. In order to mimic, imitators need information about the actions taken by others. That implies that imitation can only occur under particular information structures. Imitative rules of behavior are further specified by the type of role model behavior that they follow.

Imitation is the basis of social learning and cultural transmission in general (see e. g. Boyd and Richerson [2]). In economic contexts in particular, experimental work (e. g. Huck, Norman, and Oechssler [5], Pingle and Day [7]) finds that agents making complicated economic decisions also use imitative rules when they lack relevant information and as a mode of economizing behavior in order to save decision-making time. However, decisions taken by imitation may be suboptimal. The question is then why and what type of imitation would prevail in a population, provided that the information conditions that allow for it are fulfilled.

The literature justifies imitation on different grounds. Taking into account that optimizing is costly, imitation may coexist with optimization since it saves the costs of information gathering and processing (Conlisk [3]).

On the other hand, certain imitative rules have been justified on the grounds of their optimality among behavioral rules, according to suitable optimality criteria.

Schlag [8] explores the context of a finite population of agents who face a game against nature. Decisions are made based on own experience and after observing the behavior and experience of *only one* other (randomly sampled) agent in the population. In that context, he finds a particular imitation rule, which he calls *proportional imitation*, that is optimal in the sense of increasing expected payoffs from every state and in *every possible game*. Proportional imitation prescribes to imitate an action observed if it gave higher payoffs than the own action with a probability proportional to the difference in the payoffs observed.

Schlag [9] and Hofbauer and Schlag [4] focus instead on *proportional observation*, which prescribes to imitate another observed individual with a probability proportional to her payoff and independent of own payoff. The analysis is extended here to the case of observing more than one individual and playing a game between two different populations. It is argued that, provided that play in the other population does not change, optimality properties obtained in the case of games against nature extend to this context.

In the present note we focus on situations where each agent in a finite population *plays the field*, meaning that payoffs depend on the actions of *all* agents in the population, and not just on performance in a bilateral encounter against nature, or against a randomly chosen member of a different population.[1] We find that, in general, there are no rules of behavior that are optimal for *all* possible strategic situations. More likely, the properties of imitation cannot be separated from the situation at hand. In our view, the advantage of imitation in general lies in the fact that it minimizes decision costs, and other properties of imitation are strongly dependent on the situation considered (whether it is specifically strategic or not, whether the whole population is observed or not, etc).

The paper is organized as follows. In Section 2 we review the work on optimality of imitation rules. In Section 3 we carry out the analysis for the case of playing the field. In Section 4 we conclude.

2. GAMES AGAINST NATURE

Schlag [8] considers a finite population of agents $I = \{1, \ldots, n\}$ who repeatedly choose actions from a finite set A. Each action $a \in A$ yields an uncertain payoff $x \in [\alpha, \omega] \subset \mathbb{R}$ according to a probability distribution P_a with finite support in the interval $[\alpha, \omega]$. Agents do not know the probability distributions over payoffs; they only know the set of available actions and the range of attainable payoffs.

For any action $a \in A$, let $\pi_a = \sum_x x P_a(x)$ denote its expected payoff. The tuple $\langle A, (P_a)_{a \in A} \rangle$ is then called a *multi-armed bandit*. It constitutes a game against nature, where the agent's realization of payoffs is independent of the actions

---

[1]The term 'playing the field' is due to Maynard Smith [6]

chosen by other agents in the population. Call $\mathcal{G}(A, [\alpha, \omega])$ the set of all possible multi-armed bandits with actions in set A and payoffs in the interval $[\alpha, \omega]$.

At any period t let $\mathbf{a} \in A^n$ be the population state—the vector of actions chosen by each agent. Let $\Delta(A)$ denote the unit simplex over the set A, and let $p = (p_a)_{a \in A} \in \Delta(A)$ be the vector of proportions of agents choosing each action. Then $\overline{\pi}(\mathbf{a}) = \sum_a p_a \pi_a$ is the average expected payoff in the population at state $\mathbf{a}$.

Each period each agent decides what action to choose based on her own immediate previous experience as well as that of *one* other, randomly observed member of the population.

A *behavioral rule F* is a function that maps the actions and payoffs observed to the set of probability distributions over actions. That is, given the action and payoffs observed, it gives a prescription of the probabilities with which actions should be chosen next period.

$$F : A \times A \times [\alpha, \omega] \times [\alpha, \omega] \to \Delta(A).$$

$F_{a''}(a, a', x_a, x_{a'})$ is the probability that an agent chooses action a'' if she chose action a last period, obtained payoff x_a, and observed another agent who chose action a' and obtained payoff $x_{a'}$. A behavioral rule F is called *imitating* if $F_{a''}(a, a', x_a, x_{a'}) = 0$ for all $a'' \in A$, $a'' \neq a$, $a'' \neq a'$.

Two optimality criteria are then proposed for selecting among behavioral rules, one from an individual, although boundedly rational point of view, another one from the point of view of a rational social planner.

Consider first a boundedly rational agent who were to enter this population to randomly replace one of its members, without knowing how long she will stay in. At entering, she would have the information about what her predecessor was doing and the information about what another, randomly selected member of the population was doing. This new member must first decide on a rule to choose actions. One possible decision rule is the 'never switch' rule, which would prescribe not to change with respect to what her predecessor was doing. So at least she can always perform as well as her predecessor in expected terms. Therefore, if she had to choose another decision rule F, she would ask that F lets her at least as well off as the rule 'never switch' for the period in which entry occurs.

A behavioral rule F is called *improving*, if the expected payoff of an agent following F who randomly enters to replace any agent in the population at state $\mathbf{a}$ is higher than or equal to the expected payoff of not switching actions —following 'never switch'— and this for all possible states and all possible bandits—for all possible situations faced.

Let $\pi(F, \mathbf{a})$ be the expected payoff of any agent that enters at state $\mathbf{a}$ and uses rule F.[2] Now define *expected improvement* at state $\mathbf{a}$ when using rule F by

$$EIP_F(\mathbf{a}) = \pi(F, \mathbf{a}) - \overline{\pi}(\mathbf{a}).$$

---

[2] See [8] for the exact definition.

DEFINITION 2.1.   Given A and $[\alpha, \omega]$, a rule F is said to be *improving* if $EIP_F(\mathbf{a}) \geq 0$ for all $\mathbf{a} \in A^n$ and all multi-armed bandits in $\mathcal{G}(A, [\alpha, \omega])$.

Consider now, alternatively, that a social planner had to decide on a behavioral rule to prescribe to all members of this population. The planner would look for a rule that increases average expected payoff in the population. This is the second criterion proposed by Schlag [8] to select among behavioral rules.

Given a multi-armed bandit, a rule is said to be *payoff increasing* in that bandit, if it increases average expected payoff in the population every period. At any state $\mathbf{a}$, let $p_a(F, \mathbf{a})$ be the expected proportion of agents choosing action $a \in A$ next period when all agents decide according to a rule F. The average expected payoff in the population next period is given by $\overline{\pi}(F, \mathbf{a}) = \sum_a p_a(F, \mathbf{a}) \cdot \pi_a$.

DEFINITION 2.2.   A rule F is said to be *payoff increasing* in the bandit $g \in \mathcal{G}(A, [\alpha, \omega])$ if $\overline{\pi}(F, \mathbf{a}) \geq \overline{\pi}(\mathbf{a})$ for all $\mathbf{a} \in A^n$.

It is easy to see that, at any $\mathbf{a}$, $\overline{\pi}(F, \mathbf{a}) = \pi(F, \mathbf{a})$. Hence, a rule F is improving if and only if it is payoff increasing for all $g \in \mathcal{G}(A, [\alpha, \omega])$—both selection criteria turn out to be equivalent. Schlag [8] then concentrates only on improving rules.

Trivially, 'never switch' is improving, since it always leaves expected payoffs unchanged. However, the interest is to find nontrivial improving rules. A first, important result establishes that *all improving rules must be imitating*; in order to assure that a behavioral rule F works well *in all possible games against nature*, F must never prescribe using any new, unobserved action. The reason being that one can always find games where precisely that new action yields very low payoffs. This is crucial when agents lack information about the game they are facing.

However, not all imitative rules are improving. Schlag [8] points out that some imitative rules, e. g. 'imitate if better', cannot distinguish between lucky and certain (or highly probable) payoffs. 'Imitate if better' prescribes to imitate an action $a' \in A$ if it gave higher payoffs than another action $a \in A$ observed.

The following theorem constitutes the main result by Schlag [8] and gives a complete characterization of improving rules. A rule is improving if and only if, first, it is imitating and, second, the probability of switching from one action $a \in A$ to a different one $a' \in A$ is proportional to the difference in the payoffs observed for both actions. These rules are called *proportional imitation rules*.

THEOREM 2.1 (Schlag [8]).   *A behavioral rule F is improving if and only if (1) F is imitating, and (2) for all $a, a' \in A$, $a \neq a'$, there exist $\theta_{aa'} = \theta_{a'a} \in [0, 1/(\omega - \alpha)]$ such that*

$$F_{a'}(a, a', x_a, x_{a'}) - F_a(a', a, x_{a'}, x_a) = \theta_{aa'} \cdot (x_{a'} - x_a) \qquad \forall x_a, x_{a'} \in [\alpha, \omega].$$

Theorem 2.1 still gives a wide range of improving rules depending on $\theta_{aa'}$. It is easy to see that, a behavioral rule is *dominant*, in the sense of achieving maximal expected improvement, if it is improving, and for any $a \neq a'$, $\theta_{aa'} = 1/(\omega - \alpha)$.

This result provides a rationale for imitative rules in the case of games against nature. Being improving, proportional imitation rules are payoff increasing for all possible such games faced. Therefore, if a population faces this type of decision problem, and agents do not know anything about the bandit they are facing, or if the bandit itself is subject to changes—if the environment faced is not stationary— then the conjecture is that evolutionary selection of populations based on average performance would favor those populations using proportional imitation. In this way certain imitative rules turn out to be an optimal way to learn about a decision problem and to make decisions.

The question then is whether those properties of the proportional imitation rule also hold in strategic contexts. Schlag [8] argues that his analysis applies to normal form games in the following way. In a game played between two different populations, one can focus on how one of these two populations learns to play against the other one. Given the payoff matrix, and the strategy profile in the population of opponents, one can reinterpret the situation faced by the players in the learning population as a multi-armed bandit. Provided that play in the other population does not change, the optimal proportional imitation rule defined for the case of bandits will also be the one that maximizes the increase in average expected payoff for the learning population in this case.

Schlag [9] and Hofbauer and Schlag [4] study a different rule, called *proportional observation*, which prescribes to imitate the sampled individual with a probability proportional to her payoff and independent of own payoff.

Schlag [9] studies the framework of an infinite population of agents who decide after sampling two other agents. When agents face a game against nature, he finds that payoff increasing rules must again be imitative, but dominant rules do not exist. Moreover, it is shown that, in this context, an imitation rule consisting of a sequential application of the proportional observation rule is payoff increasing. When two populations play a game and learn according to proportional observation, the population state evolves according to a discrete version of the replicator dynamics of Taylor [10].

Hofbauer and Schlag [4] generalize this approach to the sampling of $k \in \mathbb{N}$ individuals from an explicitly infinite population, and consider imitative rules obtained by sequentially evaluating the proportional observation rule. In the context of two populations playing a two-player game, where agents in each population decide according to sequential proportional observation, they find that in the limit, as the sample size k grows to infinity, the population state evolves according to a discretization of the replicator dynamics of Maynard Smith [6].

The analysis above shows an interesting property of imitation rules when the sample and the population size are large, but does not address the question of what type of imitation would prevail when the population size is finite and each agent observes the whole population.

Moreover, the analysis above for the case of two-player games identifies rules that improve average expected payoff in one population only when play in the other

population does not change. Strictly speaking, the analysis applies neither to the case in which both populations learn simultaneously, nor to the case of playing the field—when there is only one population of n agents playing an n-person game. The reason is that, in a strategic context, it is not only the position of each agent with respect to the environment, but also the position of each agent with respect to her opponents what counts.

In the next section we analyze the case of playing the field, and try to find improving and payoff increasing rules in that case. We find that it is very restrictive to ask for such properties to be fulfilled in all possible situations.

3. PLAYING THE FIELD

In the present section we are interested in studying the properties of imitative rules in a context where the analysis in Schlag [8] does not directly apply. In particular, we are interested in n-person games, played by agents within *the same population*, a main economic example being oligopolies. Furthermore, we want to allow for each agent to make her decision on the basis of the information about actions and payoffs of *all* members of the population. In order to tackle this task we first extend the model presented in the previous section to the case of symmetric n-person games. Then, we extend the notion of a 'behavioral rule' in order to accommodate our information structure in the model. Having done this, we explore the applicability of the properties studied above to our context.

We prove that improving and payoff increasing rules are not equivalent in the context of playing the field. Moreover, only trivial learning rules can have such properties. It is obvious, for example, that a rule that would prescribe to switch actions among players must be trivially payoff increasing for all possible games. Similarly, a rule that prescribes to change actions only if the worst possible payoff is observed must be improving. But we are interested in rules that are nontrivial.

3.1.   Behavioral rules in an n-person game

Consider a finite population of agents $I = \{1, \ldots, n\}$, who are involved in the same symmetric, strategic situation. All of them have the same finite set A of actions available. In this context a *population state* is a vector $\mathbf{a} \in A^n$ specifying one action for each agent. Given a population state, payoffs for all agents are determined by the same payoff function $\pi : A \times A^{n-1} \to [\alpha, \omega]$. Agents do not know this function—they lack information about the exact way payoffs are determined. We further assume that each agent's payoff is determined independently of the names of the opponents choosing specific actions. That is, for all $a \in A$ and for all $(a_1, ..., a_{n-1}) \in A^{n-1}$, $\pi(a, (a_1, ..., a_{n-1})) = \pi(a, (a_{\sigma(1)}, ..., a_{\sigma(n-1)}))$ for all $\sigma \in \Sigma_{n-1}$, where Σ_{n-1} denotes the symmetric group of permutations of $n-1$ elements. Following the notation in Schlag [8], we define $\mathcal{G}(n, A, [\alpha, \omega])$ as the class of all symmetric n-person games with action set A, and payoff function $\pi : A \times A^{n-1} \longrightarrow [\alpha, \omega]$ satisfying the former condition.

Here we extend the notion of a behavioral rule to allow for each agent to base her decision on the information about the actions chosen and payoffs obtained by *all* members of the population in the previous period. This is done by defining a function $F : A \times A^{n-1} \times [\alpha, \omega] \times [\alpha, \omega]^{n-1} \longrightarrow \Delta(A)$. Given the actions chosen and the payoffs realized, the function F specifies how actions will be chosen in the next period. In particular, $F_{a_i'}(a_i, a_{-i}, \pi_i, \pi_{-i})$ is the probability that agent i chooses a_i' next period if she chose a_i today, obtained payoff π_i, and observed $a_{-i} \in A^{n-1}$, and payoffs $\pi_{-i} \in [\alpha, \omega]^{n-1}$ on the part of the other players.[3]

Again, $\overline{\pi}(\mathbf{a})$ denotes the average payoff at state $\mathbf{a}$, and $\pi(F, \mathbf{a})$ the expected payoff of any agent following F, who is in the position to enter at state $\mathbf{a}$ to randomly, uniformly replace any agent in the population. This expected payoff depends here on the payoff function of the game, π, and is given by

$$\pi(F, \mathbf{a}) = \frac{1}{n} \sum_{i=1}^{n} \sum_{a_i' \in A} F_{a_i'}(a_i, a_{-i}, \pi_i, \pi_{-i}) \pi(a_i', a_{-i}) \tag{1}$$

A behavioral rule F is *improving* if $EIP_F(\mathbf{a}, \pi) = \pi(F, \mathbf{a}) - \overline{\pi}(\mathbf{a}) \geq 0$ for all $\mathbf{a} \in A^n$ and for all games in $\mathcal{G}(n, A, [\alpha, \omega])$. The improving rule F is *degenerate* if $EIP_F(\mathbf{a}, \pi) = 0$ for all states and all games. Again, the behavioral rule 'never switch' is obviously a degenerate improving rule. The rule 'never switch' prescribes here $F_{a_i'}(a_i, a_{-i}, \pi_i, \pi_{-i}) = 0$ for all $a_i' \in A \setminus \{a_i\}$. Moreover, any behavioral rule F, with $F_{a_i'}(a_i, a_{-i}, \pi_i, \pi_{-i}) = 0$ for all $a_i' \in A \setminus \{a_i\}$, if $\pi_i \neq \alpha$, is obviously improving. Such rules would prescribe to switch only if the minimum possible payoff was obtained. We call this type of rules *trivially improving*.

Given that the population is in state $\mathbf{a}$, and that all agents follow F, call $\overline{\pi}(F, \mathbf{a})$ the expected average payoff next period. In our framework this is given by

$$\overline{\pi}(F, \mathbf{a}) = \frac{1}{n} \sum_{i=1}^{n} \sum_{a' \in A^n} \left(\prod_{j=1}^{n} F_{a_j'}(a_j, a_{-j}, \pi_j, \pi_{-j}) \right) \pi(a_i', a_{-i}') \tag{2}$$

Given a game in $\mathcal{G}(n, A, [\alpha, \omega])$, a behavioral rule F is *payoff increasing* for that game if $\overline{\pi}(F, \mathbf{a}) \geq \overline{\pi}(\mathbf{a})$ for all $\mathbf{a} \in A^n$.

Let $C(\mathbf{a}) = \{\mathbf{a}' \in A^n \mid a_i' = a_{\sigma(i)}, \ i = 1, ..., n, \ \sigma \in \Sigma_n\}$. Behavioral rules F, such that $\prod_{j=1}^{n} F_{a_j'}(a_j, a_{-j}, \pi_j, \pi_{-j}) = 0$, if $\mathbf{a}' \notin C(\mathbf{a})$ are obviously payoff increasing, since they leave average payoff unchanged. These are rules that prescribe a mere permutation of strategies among the agents. Rules satisfying

---

[3]Since payoffs are determined symmetrically, it seems natural to assume that behavioral rules prescribe the same to all agents who face the same situation independently of their names. Formally, for all actions $a \in A$, and all payoffs $\pi \in [\alpha, \omega]$, and for all $(a_1, ..., a_{n-1}) \in A^{n-1}$, and all $(\pi_1, ..., \pi_{n-1}) \in [\alpha, \omega]^{n-1}$, F is such that $F_{a'}(a, (a_1, ..., a_{n-1}), \pi, (\pi_1, ..., \pi_{n-1})) = F_{a'}(a, (a_{\sigma(1)}, ..., a_{\sigma(n-1)}), \pi, (\pi_{\sigma(1)}, ..., \pi_{\sigma(n-1)}))$ for all $\sigma \in \Sigma_{n-1}$. However, we do not make this assumption in what follows.

$\prod_{j=1}^{n} F_{a'_j}(a_j, a_{-j}, \pi_j, \pi_{-j}) = 0$, if $\overline{\pi}(\mathbf{a}) \neq \alpha$ are also obviously payoff increasing, since they also leave average payoff unchanged, except (maybe) in case *all* agents are obtaining the minimum possible payoff. Moreover, a rule that prescribes $\prod_{j=1}^{n} F_{a'_j}(a_j, a_{-j}, \pi_j, \pi_{-j}) > 0$ if and only if either $\mathbf{a}' \in C(\mathbf{a})$ or $\overline{\pi}(\mathbf{a}) = \alpha$ is also payoff increasing. We call these last rules *trivially payoff increasing*.

It might seem that rules prescribing switching of strategies necessarily imply coordination of individual behavior. The next example shows that this is not the case. Behavioral rules can be specified that do not imply explicit coordination.

EXAMPLE 3.1.   Let $A = \{L, R\}$, and let $n = 4$. Consider the following behavioral rule $F_R(L, (L, R, R), \pi_i, \pi_{-i}) = 1$, $F_L(R, (L, L, R), \pi_i, \pi_{-i}) = 1$ for all $\pi = (\pi_i, \pi_{-i})$ and no change from all other states. That is, the rule prescribes to change only when two agents are playing L and two R, and in that case it prescribes to switch strategy with probability one, which results in a permutation of strategies and does not require explicit coordination.

It does not seem that improving and payoff increasing rules are equivalent in our framework. In fact, as we have seen, trivially payoff increasing rules are different from trivially improving rules. The main difference with the context analyzed in Schlag [8] is that there payoffs are a result of the realization of random variables, which are independent of the proportion of agents choosing each action, while in our case payoffs depend in a deterministic way on the actions chosen by the other players. Let us explore the two properties in our context.

3.2.   Nonexistence of nontrivial rules

We prove now that, in the case of n-player games, it is impossible to find nontrivial rules that are well-behaved in all possible situations. Asking conditions (1) and (2) to be fulfilled from every state and for *all* games is extremely restrictive.

PROPOSITION 3.1.   *The only improving rules are the trivially improving ones.*

Proof.   Assume there exists an improving rule F that is not trivially improving. Then, there exists $\mathbf{a} \in A^n$ and $\pi \in [\alpha, \omega]^n$ with $\pi_j \in (\alpha, \omega]$ for some j, such that $F_{a'_j}(a_j, a_{-j}, \pi_j, \pi_{-j}) > 0$ for some $a'_j \in A \setminus \{a_j\}$. Consider the state $\mathbf{a} \in A^n$ and the game in $\mathcal{G}(n, A, [\alpha, \omega])$ satisfying $\pi(a_i, a_{-i}) = \pi_i$, $\pi(a'_i, a_{-i}) = \alpha$, for all $a'_i \in A \setminus \{a_i\}$, $i = 1, 2, ..., n$. Then,

$$
\begin{aligned}
EIP_F(\mathbf{a}, \pi) &= \frac{1}{n} \sum_{i=1}^{n} \sum_{a'_i \in A} F_{a'_i}(a_i, a_{-i}, \pi_i, \pi_{-i}) \cdot (\pi(a'_i, a_{-i}) - \pi(a_i, a_{-i})) \\
&= \frac{1}{n} \sum_{i=1}^{n} \sum_{a'_i \in A \setminus a_i} F_{a'_i}(a_i, a_{-i}, \pi_i, \pi_{-i}) \cdot (\alpha - \pi_i) \\
&\leq \frac{1}{n} F_{a'_j}(a_j, a_{-j}, \pi_j, \pi_{-j}) \cdot (\alpha - \pi_j) < 0
\end{aligned}
$$

which yields a contradiction. ∎

That is, if F were improving, but not trivially improving, there would be some state where a player j, who was earning payoffs above the minimum α, is prescribed to change action from a_j to a'_j. In that case, we can always define a game such that nothing else changes, but precisely the combination of strategies that results, (a'_j, a_{-j}), yields payoff α, so that at least j's payoffs strictly decrease.

PROPOSITION 3.2. *The only rules that are payoff increasing for all possible games in $\mathcal{G}(n, A, [\alpha, \omega])$ are the trivially payoff increasing ones.*

Proof. Assume the rule F is payoff increasing for all games in $\mathcal{G}(n, A, [\alpha, \omega])$, but not trivially payoff increasing. Then, there exists a state $\mathbf{a} \in A^n$ and a payoff vector $\pi \in [\alpha, \omega]^n$ with $\pi_k \in (\alpha, \omega]$ for some k, such that F prescribes $\prod_{j=1}^{n} F_{a_j^*}(a_j, a_{-j}, \pi_j, \pi_{-j}) > 0$ for some $\mathbf{a}^* \in A^n \setminus C(\mathbf{a})$. Consider now the state $\mathbf{a} \in A^n$ and the game in $\mathcal{G}(n, A, [\alpha, \omega])$ satisfying that $\pi(a_i, a_{-i}) = \pi_i$, and $\pi(a'_i, a'_{-i}) = \alpha$ for all $\mathbf{a}' \in A^n \setminus C(\mathbf{a})$, and all $i = 1, 2, ..., n$.

$$\bar{\pi}(F, \mathbf{a}) - \bar{\pi}(\mathbf{a}) = \frac{1}{n} \sum_{i=1}^{n} \sum_{\mathbf{a}' \in A^n} \prod_{j=1}^{n} F_{a'_j}(a_j, a_{-j}, \pi_j, \pi_{-j})(\pi(a'_i, a'_{-i}) - \pi(a_i, a_{-i}))$$

$$= \frac{1}{n} \sum_{i=1}^{n} \sum_{\mathbf{a}' \in A^n \setminus C(\mathbf{a})} \prod_{j=1}^{n} F_{a'_j}(a_j, a_{-j}, \pi_j, \pi_{-j})(\alpha - \pi_i)$$

$$\leq \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{n} F_{a_j^*}(a_j, a_{-j}, \pi_j, \pi_{-j})(\alpha - \pi_i)$$

$$\leq \frac{1}{n} \prod_{j=1}^{n} F_{a_j^*}(a_j, a_{-j}, \pi_j, \pi_{-j})(\alpha - \pi_k) < 0$$

which yields a contradiction. ∎

That is, if F were payoff increasing for all games, but not trivially payoff increasing, there would be some state where at least some players were earning payoffs above the minimum, and where the rule prescribes to change actions in such a way that the resulting state is not a mere permutation. In that case, we can always define a game such that nothing else changes, but precisely the combination of strategies that results yields minimum payoff, so that some agents' payoffs, and thus the resulting average payoff, strictly decrease.

Therefore, in the framework of a small population of agents involved in a symmetric, strategic situation, only trivial behavioral rules can be improving or payoff increasing for all possible games. These properties turn out to be too strong. In the search for well-behaved behavioral rules that are nontrivial, we must either relax these properties, or concentrate on different ones. The following example tries to illustrate further what underlies the results above, showing how easy it is to find even small classes of games for which neither improving, nor payoff increasing rules exist.

EXAMPLE 3.2.   Consider the class of games $\mathcal{G}(3, \{L, R\}, [-2, 2])$ with three players, actions $A = \{L, R\}$ and payoffs in the interval $[-2, 2]$. Let $\tilde{\mathcal{G}} = \{G_1, G_2, G_3, G_4\}$ be a subclass of $\mathcal{G}$. Assume that in these four games all players get the same payoffs in each state. Payoffs are summarized in the following table.

	LLL	LLR	LRR	RRR
G_1	1	2	-2	1
G_2	1	-2	2	1
G_3	1	2	-2	-2
G_4	-2	-2	2	1

For example, in each of the three states of type (L, L, R), where two players play L, and the third one plays R, all players get payoff 2 in games G_1 and G_3, and payoff -2 in games G_2 and G_4. Note that in $\tilde{\mathcal{G}}$ for each type of state there is always some game where precisely that situation yields the lowest possible payoffs. In the following we show that, in the subclass $\tilde{\mathcal{G}}$ there are neither nontrivial improving, nor nontrivial payoff increasing rules.

Any improving rule must prescribe not to change from states of type (L, L, R) when payoffs $(2, 2, 2)$ are observed – a unilateral change of strategy would take agents to a state of type (R, R, L) or (L, L, L), which in G_1 and G_3 yields a negative expected improvement. Analogously, any improving rule must prescribe not to change from (L, R, R) when payoffs $(2, 2, 2)$ are observed. Moreover, any improving rule must prescribe not to change from state (L, L, L) when payoffs $(1, 1, 1)$ are observed, since a unilateral change of strategy would take agents to a state of type (L, L, R), which in game G_2 would yield payoffs -2 and, thus, a negative expected improvement. Analogously, any improving rule must prescribe not to change from state (R, R, R) when payoffs $(1, 1, 1)$ are observed. This proves that an improving rule may only prescribe to change action from states where payoffs $(-2, -2, -2)$ are observed.

Similarly, if we want a behavioral rule to be payoff increasing for all games in $\tilde{\mathcal{G}}$ the following must hold. The rule must prescribe not to change from states (L, L, R) and (R, R, L) when payoffs $(2, 2, 2)$ are observed.[4] It must also prescribe not to change from state (L, L, L) when payoffs $(1, 1, 1)$ are observed. To see this, let p_0 be the induced joint probability of staying at (L, L, L) when payoffs $(1, 1, 1)$ are observed. Let p_1, p_2, and p_3 be the induced joint probabilities of transition to states (L, L, R), (R, R, L), and (R, R, R) respectively when payoffs $(1, 1, 1)$ are observed. The following conditions must hold simultaneously.

$$p_0 + 2p_1 - 2p_2 + p_3 \geq 1$$
$$p_0 - 2p_1 + 2p_2 + p_3 \geq 1$$
$$p_0 + 2p_1 - 2p_2 - 2p_3 \geq 1$$

---

[4]Note that in this case there exist no individual behavioral rule that induces with probability one a permutation of strategies among agents.

The first two equations imply that $p_0 + p_3 = 1$. Thus, $p_1 = p_2 = 0$. Since p_i, with $i = 0, \ldots, 3$ are joint probabilities, this implies that either $p_0 = 1$, or $p_3 = 1$. However, if $p_3 = 1$, then in G_3 the rule would cause average expected payoff to decrease. Thus, any payoff increasing rule for all games in $\tilde{\mathcal{G}}$ must prescribe not to change from state (L, L, L) when payoffs $(1, 1, 1)$ are observed. Analogously, the same must hold for state (R, R, R) when payoffs $(1, 1, 1)$ are observed. This proves that a payoff increasing rule may only prescribe to change action from states where payoffs $(-2, -2, -2)$ are observed.

In some cases, however, improving rules will exist, if we restrict the class of games sufficiently. In what follows we give an example of a behavioral rule that is improving for a particular class of Bertrand games.

EXAMPLE 3.3.   Consider an industry where identical firms set prices. Firms are likely to know that they are involved in a market with some sort of Bertrand competition. However, it is unlikely that they know the exact demand, and cost functions. In this context they would probably like to use a behavioral rule that works well for an extensive class of Bertrand-type of games, even if it is not appropriate for making other decisions.[5]

For simplicity, consider a market with n identical firms facing constant marginal and average costs c. Assume they set prices and customers buy only from the one with minimum price. Let $D(p)$ be a decreasing demand function. In case of ties, demand splits equally. The profits of a firm i that charges minimum price p_i are $\pi_i(p_i, p_{-i}) = (p_i - c)\frac{D(p_i)}{m}$, where $m \in \{1, 2, \ldots, n\}$ is the number of firms with minimum price p_i. Profits are zero for the firms with higher than minimum price. We want to show that rules of the type 'imitate the best' are improving for this class of Bertrand games. These rules prescribe to mimic with probability one the price charged by the firm with biggest profits. If several firms obtain the same maximum profits, then any of them with positive probability.

Note that only two types of states are relevant. First, states in which all firms set the same price. Here all firms share the market, and obtain the same profits. Imitate the best gives zero expected improvement at these states. Second, states of the form $(p, p, \overset{m}{\ldots}, p, p_1, p_2, \ldots, p_{n-m})$ with $1 \le m < n$ and $p < p_i$, $i = 1, \ldots, n - m$. For these states we must consider the following cases. If $p > c$, i. e. firms with minimum price make profits, then the expected improvement of a firm following imitate the best is

$$\frac{n - m}{n}(p - c)\frac{D(p)}{m + 1} > 0.$$

If $p < c$ ($p = c$) and $m > 1$, i. e. at least two firms set minimum price and these make losses (zero profits), then imitate the best always prescribes to mimic any $p_i > p$ (any $p_i > p$ or $p = c$) with $i = 1, \ldots, n - m$, which yields zero expected

profits and a nonnegative expected improvement. Finally, if $p < c$ ($p = c$) and $m = 1$, i. e. only one firm sets minimum price and this makes losses (zero profits), then again imitate the best always prescribes to mimic any $p_i > p$ (any $p_i > p$ or $p = c$) with $i = 1, \ldots, n-1$. Let $p' = \min\{p_1, \ldots, p_{n-1}\}$ which is the new minimum price after the firm with price p imitates away from p. Call ρ the probability that this firm imitates precisely p', and let $m' \geq 2$ be the number of firms setting p' after imitation. Then the expected improvement of imitate the best is

$$\frac{\rho}{n}(p' - c)\frac{D(p')}{m'} - (p - c)\frac{D(p)}{n},$$

which is obviously positive if $p' \geq c$. If $p' < c$, it follows from the fact that $D(p)$ is decreasing; since $p' > p$

$$\rho\frac{D(p')}{m'} < \frac{D(p')}{m'} < D(p') \leq D(p),$$

and thus

$$\frac{\rho}{n}(p' - c)\frac{D(p')}{m'} > (p' - c)\frac{D(p)}{n} > (p - c)\frac{D(p)}{n}.$$

Imitate the best is also payoff increasing in this case. Note that at any state, average payoff is always $\frac{D(p)}{n}(p - c)$. At states where all firms charge the same price, and at states of type $(p, p, .^m., p, p_1, p_2, \ldots, p_{n-m})$ with $p > c$ imitate the best leaves average payoff unchanged. If $p \leq c$ imitate the best increases average payoff strictly.

4. CONCLUSION

We review the properties of learning by imitation in populations where all agents face the same problem and lack either relevant information, or the capability to decide by optimization. In such contexts, it seems that learning from the experience of other agents who have faced the same problem would allow to decide accurately, and even optimally, without incurring the effort of individually analyzing the problem. Intuitively, the accumulated experience of other agents in the population should contain more and better information than all the experience that each agent can ever gather and process in a reasonable time. Moreover, by imitating, an agent will avoid errors that the population has already leaned to avoid. Actions that are no longer present in the population cannot be imitated.

In games against nature and in games between two different populations, where conspecifics and opponents are separated, it has been shown in the literature that certain imitation rules (proportional imitation and proportional observation) have the property of increasing average expected payoff in the population, when all

agents use these rules, and whenever play in the population of opponents does not change. In those contexts, each agent learns from the members of her own population and then plays against nature or against agents from a different population. This implies that each agent's payoffs are independent of the actions chosen by other agents in the same population. An economic example of this type is that of a population of sellers of a certain good trying to learn the average reservation value of a population of buyers.

Our interest is on a framework where conspecifics and opponents are no longer separated. We have looked at the example of identical firms in a market facing the same demand and cost functions, unknown to them, and trying to decide what price to charge for their output. The question is here to what extent imitation is an efficient way of learning in the presence of strategic considerations within the population.

In this case, we have found that, in general, nontrivial rules that increase average payoffs in all possible situations do not exist. That is, there are no universally well-behaved behavioral rules to be applied in all cases. We have seen that it is easy to find simple examples where only trivial rules will always increase average payoffs. Rules based on switching actions among players, or rules that prescribe to change actions only when the worst possible case is observed, will trivially work. At the same time, however, we have seen in an example that, if we restrict to specific classes of games, we can still find imitative rules that are always payoff increasing. In the example of Bertrand competition, if each firm imitates the price charged by the firm with highest profits, average profits in the industry will never decrease.

REFERENCES

1. C. Alós-Ferrer, A.B. Ania, and K.R. Schenk-Hoppé, An Evolutionary Model of Bertrand Oligopoly, *Games Econ. Behav.* **33** (2000), 1-19.

2. R. Boyd and P.J. Richerson, "Culture and the Evolutionary Process," The University of Chicago Press, Chicago and London, 1985.

3. J. Conlisk, Costly optimizers versus cheap imitators, *J. Econ. Behav. Organ.* **1** (1980), 275-293.

4. J. Hofbauer and K. Schlag, Sophisticated Imitation in Cyclic Games, *J. Evolutionary Econ.* **10** (2000), 523-543.

5. S. Huck, H.-T. Norman, J. Oechssler, Learning in Cournot oligopoly—an experiment, *Econ. J.* **109** (1999), C80-C95.

6. J. Maynard Smith, "Evolution and the Theory of Games," Cambridge University Press, Cambridge, 1982.

7. M. Pingle and R.H. Day, Modes of economizing behavior: Experimental evidence, *J. Econ. Behav. Organ.* **29** (1996), 191-209.

8. K. Schlag, Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits, *J. Econ. Theory* **78** (1998), 130-156.

9. K. Schlag, Which one should I imitate? *J. Math. Econ.* **31** (1999), 493-522.

10. P.D. Taylor, Evolutionary stable strategies with two types of players, *J. Appl. Probability* **16** (1979), 76-83.