Monotone imitation*

Carlos Oyarzun
oyarzun@merlin.fae.ua.es
Universidad de Alicante

Johannes Ruf
ruf@stat.columbia.edu
Columbia University

June 30, 2008

Abstract

We analyze the social learning process of a group of individuals who have limited information about the payoff distributions of each action. We say that a behavioral rule is *first-order monotone* (FOM) if the number of individuals who play actions with first-order stochastic dominant payoff distributions is expected to increase in any environment. We provide a characterization of FOM rules. Both Imitate if Better and Schlag's (J Econ Theory 78:130-156, 1998) Proportional Imitation rule are FOM. No FOM rule is *dominant* in the sense of having the best performance in every environment.

Keywords: Imitation, Social learning, Stochastic dominance, First-order monotonicity

JEL Classification: D81, D83

*We received many helpful suggestions from Timothy Gronberg, Brit Grosskopf, Paan Jindapon, Lucas Rentschler, Andrew Schotter, Li Song, Guoqiang Tian, and Megha Watugala. We thank Karl Schlag and anonymous referees for insightful comments. We specially thank Rajiv Sarin for detailed discussions on many previous versions of this paper. Financial support from the Private Enterprise Research Center at Texas A&M University, the Lynde and Harry Bradley Foundation, and the Spanish Ministry of Science and Technology (Project SEJ 2007-62656) is gratefully appreciated.

1 Introduction

People often learn from their own experiences and by observing others. As decision makers, we compare the performance of our decisions with the performance of the actions selected by other people. Actions leading to better results are more likely to be played in the future when the same problem is faced again. In this paper we find conditions, for this manner of learning, under which the number of individuals who play the actions with the best payoff-distribution is expected to increase. We show that this requires the social learning process to exhibit three simple and intuitively appealing features. First, the action each individual plays in the next period is either the action she played in the current period or an action played by another individual she observed. Second, the higher the payoff provided by an action, the more likely this action is to be played in the next period. Third, all the actions are treated in a pairwise symmetric way.

We analyze a finite population of individuals facing the same multi-armed bandit repeatedly. The information context we analyze is the one introduced by Schlag [15]: every period each member of the population (she) has to choose one action out of a fixed set that is common to the entire population. The decision makers do not know the payoff distributions of these actions; the only information available to each of them is the action that she played last period, the action played by some other member of the population (he), and the corresponding payoffs. This information determines the probability with which the individual will play each action in the next period. The function mapping the available information to the probability of choosing each action is what we call a *behavioral rule*, or, simply, the rule.

We study how the behavioral rule used by the individuals determines whether a greater fraction of the population will, in the future, choose actions that are more likely to give higher payoffs. A behavioral rule is said to be *first-order monotone* (FOM) if the fraction of the population who plays the actions whose distributions are first-order stochastically dominant is expected to increase. This is required for every possible set of probability distributions associated to the different actions.

In economics, an important role for learning and evolution theory is providing an analytical framework to asses when experience will lead individuals to behave as rational agents (for example, see the discussion in Börgers [4]). This theory should illuminate our understanding of when it is plausible that simple and natural learning processes may lead decision makers to rational decisions in complex environments. In this paper, we derive the precise characteristics that a social learning process has to satisfy in order to lead the population to choose the actions which fully informed rational agents would choose. Consistency of decisions with the order defined by first-order stochastic dominance may be regarded as one of the most elemental characteristics of rational choice. Our results reveal that very simple and rather natural behavioral rules may display this property. In this sense the results implied by the characterization of FOM rules suggest that, indeed, we could have posed the question in the opposite direction. In other words, we could have asked how simple behavioral rules could be characterized in terms of the dynamics of the fraction of the population who play first-order stochastic dominant actions. The characterization we provide below reveals that the set of FOM rules can be obtained in either way.

In multi-armed bandit problems as the one studied here, payoffs are typically interpreted as units of money. For example, in Rothschild [14], payoffs represent the profits of a store trying to price an item in its inventory; and in the recent experiments of Apesteguia, Huck, and Oechssler [1], payoffs correspond to the profits of oligopolistic firms playing a Cournot game. When payoffs are interpreted as monetary units, our focus on rules that lead to play actions with first-order stochastic dominant payoff distributions is consistent with the developments in the theory of rational choice under risk (expected utility theory, EUT) with regards to stochastic dominance.[1] In EUT, decision makers with increasing Bernoulli utility functions prefer first-order stochastically dominant distributions. However, they do not always prefer payoff distributions with the highest expected value. Therefore, we identify the behavioral rules that lead to decision making consistent with first-order stochastic dominance. The results reveal that such rules are indeed simple and intuitive. Alternatively, Schlag [15] introduces and characterizes *improving* rules; this is the set of rules for which the average payoff of the population is expected to increase in every environment. Using Schlag's [15] characterization for improving rules and the characterization of FOM rules we provide here, it is easy to see that every (non-trivial) improving rule is FOM. In EUT, the class of Bernoulli functions consistent with first-order stochastic dominance contains all the Bernoulli functions that are increasing, not only those that are linear. Analogously, FOM rules do not need to be linear in payoffs; this contrasts with the set of improving rules where linearity in payoffs is a

[1] See, for example, Hanoch and Levy [9].

necessary condition. In other words, our definition of FOM rules allows for a wider range of specifications because it is based on the notion of first-order stochastic dominance and not on the notion of expected values. Besides of its technical importance, this generalization is of interest in economics, since, as we will see below, it allows for behavioral rules that play an important rule in the literature.[2]

Alternatively, payoffs may be (more generally) interpreted as a numerical representation of a preference order over outcomes associated to the played actions. A sufficient condition for this representation is that preferences over outcomes induce a linear order over this set and that the cardinality of the set of outcomes is smaller than the cardinality of the set of payoffs in the analysis.[3] Under this interpretation FOM rules are those leading individuals to choose actions that are more likely to provide preferred outcomes.

After introducing the framework in Section 2, in Section 3 we provide two characterizations for FOM rules. The first characterization reveals that a necessary condition for first-order monotonicity is that the behavioral rule must be *imitative*. A behavioral rule is called imitative if each individual plays either the action that she played in the previous period or the action played by the individual she observed. Furthermore, this characterization shows that when the payoff distribution of one action strictly first-order stochastically dominates the payoff distribution of the other, then the expected net-switch to the first-order dominant action is strictly positive. The analysis also reveals that this condition requires the behavioral rule to be *impartial*. A behavioral rule is called impartial if, when all the distributions of payoffs associated to the different actions are the same, the proportion of individuals who choose each action is expected to remain the same.

In the second characterization we describe FOM rules in terms of their functional form. In particular, besides being imitative, FOM rules can be described in terms of what we call the net-switching functions of a behavioral rule. For each pair of actions a and a', we define the *net-switching function from a to a'*. This is the difference between the probability of playing a' in

[2]The concept of first-order monotonicity may also be motivated by evolutionary considerations, however we have not explored formally that possibility in this paper. We also discuss this in Section 6.

[3]In other words, if it is possible to define an injective function from the set of outcomes to the set of payoffs. If a preference order is indifferent over two or more different outcomes, the set of outcomes can be redefined in such a way that these outcomes are the same in order to obtain a linear order.

the next period by an individual who played a in this period and observed another individual who played a', and the probability of playing action a in the next period by an individual who played a' in this period and observed another individual who played a, when the payoff obtained from each action in both cases is the same. The arguments of this function are the payoffs obtained from playing each action. This characterization reveals that the net-switching functions are symmetric in the obtained payoffs in the sense that if the payoff obtained with action a is substituted for the payoff obtained with action a' and vice versa, then only the sign of the net-switching function changes. The magnitude remains the same. From the proof of this result it is easy to see that this feature is derived from the impartial property. Furthermore we prove that the net-switching function from a to a' is increasing in the payoff obtained with action a'. Likewise, and as a consequence of the symmetry described above, the net-switching function is decreasing in the payoff obtained with action a. Finally, we show that the net-switching functions are strictly positive if the payoff of the action to which the probability is being switched is strictly greater than the payoff of the action from which the probability is being switched. However, the net-switching functions do not need to be strictly increasing in the payoff of the action that receives the probability. A technical lemma that is a corner stone for the proof of this result is relegated to Appendix A. The characterizations of FOM rules that we provide in Section 3 assume that all the individuals in the population use the same behavioral rule. Appendix B shows that it is straightforward to generalize these characterizations to the case of heterogenous behavioral rules.

In Section 4 we discuss a number of examples in the literature that satisfy the properties we described above. It is easy to show that the rule Imitate if Better (IIB) and Schlag's [15] Proportional Imitation rule are FOM. As suggested by its name, the rule IIB prescribes switching (with probability one) to the observed action only if the payoff of that action is higher.[4] The Proportional Imitation rule prescribes switching to the observed action only if the payoff of that action is higher and with a probability proportional to the difference in the payoffs. We define formally these rules in the next section. We also show that no FOM behavioral rule can be said to be dominant in the sense that for all of them we can find a set of distributions for the payoffs such

[4]The rule IIB plays an important role in the literature, see for example Ellison and Fudenberg [8] and Vega-Redondo [20].

that, in that set, another FOM behavioral rule has a greater expected increase in the fraction of the population playing a dominant action. However, a rule may be dominant within a certain class of environments. In particular, we characterize the class of environments for which the rule IIB is dominant.

In Section 5 we study the dynamics of the choices in large populations. We show that if decision makers use a FOM rule, then it is very likely that an arbitrarily large fraction of the population will play a dominant action after a finite number of periods, provided that some initial experimentation takes place.

Section 6 discusses and motivates alternative directions for further research. Some of these extensions are analyzed formally in Appendix C and Appendix D. Appendix C provides the analysis of *individually monotone* behavioral rules. These rules are imitative rules such that when the payoff distribution of one of the actions an individual observes first-order stochastically dominates the payoff distribution of the other, then the expected probability of playing the dominant action is higher than the expected probability of playing the other. Appendix D analyzes *second-order monotone* (SOM) rules. These rules are the analogous of FOM rules with regards to second-order stochastic dominance.

Learning in social contexts has received considerable attention. Ellison and Fudenberg [7], [8] show how simple behavioral rules may lead a population to play optimal actions in the long run. Proceeding in the opposite direction, Schlag [15] and Morales [10] propose and characterize a number of desirable properties for behavioral rules. Our work moves in this direction too, but in contrast to theirs, our paper focuses on properties of performance based on the concept of stochastic dominance rather than on the expected value of the payoffs. Other information contexts and problems are considered in Schlag [16], Bala and Goyal [2], Morales [10], and Offerman and Schotter [11]. Schlag [16] analyzes a model similar to the one analyzed in this paper, but the decision maker is able to observe the actions and obtained payoffs of two other individuals of the population, instead of only one. He identifies behavioral rules that lead an infinite population to the expected payoff maximizing action. Bala and Goyal [2] analyze a model where individuals can learn from their own experience and their neighbors' about wether one of two actions has a better payoff distribution than the other. In their model individuals use a boundedly (rational) Bayesian rule to update their beliefs about the payoff distribution of each action. There are only two possible payoffs, high and low. If one of the actions is more likely to provide the high payoff than the other, then, with probability 1, all the individuals converge to choosing that action after a finite number of periods. In Morales [10] decision makers observe their actions and obtained payoffs and the action played by another member of the population and the corresponding payoff. He identifies the behavioral rules for which the expected payoff of each individual is expected to increase in every environment. Agents' information in Morales [10] is the same as in the model discussed here, but in his model only the probabilities of the played and observed actions can be updated. Offerman and Schotter [11] study an experiment of endogenous sampling where the experimental subjects, after observing everyone's payoff, are allowed to choose the individuals from whom they will observe the played actions. They find that experimental subjects choose to observe the actions played by the individuals who obtained the highest payoffs and then imitate them, even though in their setup, this leads decision makers to choose actions that are not optimal in terms of expected payoffs. Finally, similar information settings have been studied in Vega-Redondo [20] and Apesteguia et al. [1], in the context of Cournot games. In Vega-Redondo's paper, in every period and with positive probability, each firm can modify its action. He assumes that in the next period firms would play any of the actions that obtained the highest profit in the current period and shows that this leads to Walrasian quantities. Apesteguia et al. [1] provide some experimental results that show how the imitation patterns of experimental subjects respond to different information treatments. An important result in their experiments is that the probability of switching to observed actions responds to both sign and magnitude of the difference in payoffs.

2 Framework

We analyze the behavior of a population described as a set of agents W. For most of our analysis the size of the population is finite, so $|W| < \infty$. In every period, each member of the population has to choose an action $a \in A$, where A is the finite set of actions available to the decision maker. The chosen action, a, yields a payoff $x \in [0, 1]$; this payoff is a random variable whose probability measure and distribution are denoted by μ_a and F_a, respectively. We will refer to the vector of distributions as the *environment* and we will denote it by F, i.e. $F := (F_a)_{a \in A}$. We assume that the payoffs obtained from different actions are pairwise independent and independent in time. The fraction of the population that chooses the action a in the current period is denoted by p_a. In what follows, first-order stochastic dominance is abbreviated by F_a fosd $F_{a'}$ and means that $F_a(x) \leq F_{a'}(x)$ for all $x \in [0,1]$; and strict first-order stochastic dominance, denoted by F_a sfosd $F_{a'}$, means that F_a fosd $F_{a'}$ and $F_a(x) < F_{a'}(x)$ for at least one element $x \in [0,1]$. Let $A^* := \{a \in A : F_a \text{ fosd } F_{a'} \; \forall \; a' \in A\}$ and for every set $S \subseteq A$, let $p(S)$ be the fraction of the population that chooses, in the current period, an action contained in this set, i.e. $p(S) := \sum_{a \in S} p_a$.

Each member of the population is able to observe the action that she chose in the current period and the payoff she obtains. She is also able to observe the action chosen in the current period by one of the other members of the population and the payoff he obtains. This is all the information individuals use at the moment of choosing their actions in the next period. Decisions are assumed to be probabilistic, i.e., we assume that the action each decision maker chooses in the next period is random. The available information, however, affects the probability with which each action is chosen. In particular, we assume that the behavior of each individual in the population can be described by the function $L : A \times [0,1] \times A \times [0,1] \to \Delta(A)$. This function maps each quartet (a, x, a', y) to a vector $L(a, x, a', y)$. Here a is the action chosen by the agent in the current period, x is the payoff she obtains, a' is the action chosen by the agent that she observes and y is the payoff he obtains. All of these variables will determine the vector $L(a, x, a', y)$ containing the probabilities of choosing each action in the next period. The element $L(a, x, a', y)_{a''}$ of $L(a, x, a', y)$ denotes the probability with which the decision maker will play action a'' in the next period. $L(a, x, a', y)$ must be contained in $\Delta(A)$, which is the set of all probability distributions over A. Accordingly, the vector-valued function L is called the *behavioral rule* of the individual. For notational convenience, we assume that all the individuals in the population use the same rule. However, as we show in Appendix B, it is easy to generalize the characterization of FOM rules to the case of heterogenous behavioral rules. Note that if we know the behavioral rule, the action chosen by the agent and the action chosen by the agent that she is able to observe, we can compute the expected probability of playing each action in the next period. Let $L^{a''}_{a,a'}$ be the expected probability, before the realization of the payoffs, of choosing action a'' tomorrow by a member of the population with behavioral rule L who played action a and observed another individual who played action a', i.e., $L^{a''}_{a,a'} := \int \int L(a, x, a', y)_{a''} dF_a(x) dF_{a'}(y)$.

For each member of the population, the individual that she is able to observe is determined by a random procedure. Formally, the probability that the agent $c \in W$ observes another agent $d \in W$ will be denoted by $\Pr(c \curvearrowright d)$; thus $\sum_{d \in W \setminus \{c\}} \Pr(c \curvearrowright d) = 1$ for all $c \in W$. It is assumed that the sampling process is symmetric, i.e., $\Pr(c \curvearrowright d) = \Pr(d \curvearrowright c)$ for all $c, d \in W$. Without this assumption the analysis rapidly becomes more complex. Furthermore, we assume that $\Pr(c \curvearrowright d) > 0$ for all c, $d \in W$ such that $c \neq d$. As we will see in the next section, this assumption guarantees that any individual playing a non-dominant action can learn from any other individual playing a dominant action with some positive probability, and therefore, it allows the fraction of the population playing dominant actions to increase in expected terms.

The expected fraction of the population that will play action $a \in A$ in the next period, given the choices of the population in the current period, is denoted by p'_a. Let $s(c)$ be the action $c \in W$ played in the current period. It is easy to see that p'_a can be computed as

$$p'_a = \frac{1}{|W|} \sum_{c \in W} \sum_{d \in W \setminus \{c\}} \Pr(c \curvearrowright d) L^a_{s(c),s(d)}.$$

Likewise, if we denote the expected fraction of the population that will play an action in A^* during the next period by $p'(A^*)$, it is easy to see that

$$p'(A^*) = \sum_{a \in A^*} \frac{1}{|W|} \sum_{c \in W} \sum_{d \in W \setminus \{c\}} \Pr(c \curvearrowright d) L^a_{s(c),s(d)}.$$

In the next section we characterize the behavioral rules that guarantee that $p'(A^*) > p(A^*)$ when $p(A^*) \in (0, 1)$, in every environment.

We end this section with two examples of behavioral rules that are well known in the literature and that will be discussed in detail in Section 4. Our first example is Schlag's [15] Proportional Imitation rule. This rule never switches unless the payoff of the sampled action is strictly greater than the payoff of the action played by the individual. If this is the case, the probability of switching is $y - x$. This rule is defined as follows: for all a, $a' \in A$ let

$$L(a, x, a', y)_{a'} = \begin{cases} y - x & \text{if } y > x \\ 0 & \text{if } y \leq x \end{cases}$$

$$L(a, x, a', y)_a = \begin{cases} 1 - y + x & \text{if } y > x \\ 1 & \text{if } y \leq x. \end{cases}$$

Our second example is a version of the rule IIB. This rule prescribes switching with probability one when $y > x$, with probability $1/2$ if $y = x$, and not to switch otherwise. This rule can be written as

$$L(a, x, a', y)_{a'} = \begin{cases} 1 & \text{if } y > x \\ \frac{1}{2} & \text{if } y = x \\ 0 & \text{if } y < x \end{cases}$$

$$L(a, x, a', y)_{a} = \begin{cases} 0 & \text{if } y > x \\ \frac{1}{2} & \text{if } y = x \\ 1 & \text{if } y < x \end{cases}$$

for all a, $a' \in A$.

Later on, in Section 4, we use the characterizations we provide in the next section for FOM behavioral rules to verify that these rules display this property.

3 First-order monotone behavioral rules

In this section we provide the formal analysis of FOM rules. These behavioral rules lead a population to increase the number of its members that play a dominant action.

Definition 1. A rule L is said to be FOM if $p'(A^*) \geq p(A^*)$, with strict inequality when $p(A^*) \in (0, 1)$, in every environment.

Definition 1 does not impose extra restrictions on the random process that determines how individuals sample each other. Furthermore, it does not specify what the proportions of the population playing each action in the current period are. Therefore $p'(A^*) \geq p(A^*)$ is required to be satisfied for any environment, any symmetric sampling procedure where each individual may observe any of the others with a positive probability, and regardless the actions that are currently being played by the members of the population. Furthermore, in expected terms, the fraction of the population playing a dominant action always increases *strictly* as long as at least one individual plays a dominant action and at least one individual does not. By definition, $p(A^*)$ cannot increase strictly when it is equal to one. The reason why we do not impose $p(A^*)$ to increase strictly when $p(A^*) = 0$ is a little bit more subtle and we discuss about it with more detail below.

We provide two characterizations for FOM rules. The first characterization reveals that these rules need to satisfy a number of restrictions. The first restriction requires that each member of the population will not choose in the next period any action that is not either the action she chose or the one that was chosen by the agent she observed. Schlag [15] calls such rules *imitative.*

Definition 2. A rule L is said to be imitative if for all actions a, a', $a'' \in A$, such that $a'' \notin \{a, a'\}$, we have $L(a, x, a', y)_{a''} = 0$ for all x, $y \in [0, 1]$.

The following preliminary result describes the expected change in the fraction of the population who plays action $a \in A$ for imitative rules.

Lemma 0. *If L is imitative, then*

$$p'_a - p_a = \frac{1}{|W|} \sum_{c \in W,\ s(c) \neq a} \ \sum_{d \in W,\ s(d) = a} \Pr(c \curvearrowright d)(L^a_{s(c),a} - L^{s(c)}_{a,s(c)}).$$

The expression above is just the difference between the expected fraction of the population that will switch from an action $a' \in A \setminus \{a\}$ to action a in the next period and the expected fraction of the population that will switch from action a to a different action in the next period. The proof of Lemma 0 follows from straightforward calculations and the assumption that $\Pr(c \curvearrowright d) = \Pr(d \curvearrowright c)$ for all c, $d \in W$.

We also found that every FOM behavioral rule is not expected to change the proportion of the population which chooses each action in all the environments where the distributions of payoffs associated with each action are the same. We call that property *impartiality.*

Definition 3. We say that a rule is impartial if $p'_a - p_a = 0$ for all $a \in A$, whenever $F_a(x) = F_{a'}(x)$ for all $x \in [0, 1]$ for all $a' \in A$.

We now present a first characterization of FOM rules. This characterization shows that both imitation and impartiality are necessary conditions for a rule to be FOM.

Lemma 1. *A rule L is FOM if and only if it satisfies the following conditions:*

(i) L is imitative.

(ii) $F_{a'}$ *sfosd* $F_a \Rightarrow L^{a'}_{aa'} - L^a_{a'a} > 0$*,* $\forall\, a,\, a' \in A$ *for all environments.*

Proof. Sufficiency:
First note that (ii) implies that $F_{a'}$ *fosd* $F_a \Rightarrow L^{a'}_{aa'} - L^a_{a'a} \geq 0, \forall\, a,\, a' \in A$, for all environments. To see this suppose that for some $a,\, a' \in A$, $F_{a'} = F_a$, but $L^{a'}_{aa'} - L^a_{a'a} < 0$. For $\varepsilon \in (0,1)$ consider the modified environment $\widetilde{F}$ such that for any interval $I \subseteq [0,1)$, $\widetilde{\mu}_{a'}(I) = (1-\varepsilon)\mu_{a'}(I)$, $\widetilde{\mu}_{a'}(1) = \mu_{a'}(1) + \varepsilon\mu_{a'}[0,1)$; and for any interval $I \subseteq (0,1]$, $\widetilde{\mu}_a(I) = (1-\varepsilon)\mu_a(I)$, $\widetilde{\mu}_a(0) = \mu_a(0)+\varepsilon\mu_a(0,1]$. In this modified environment $\widetilde{F}_{a'}$ sfosd $\widetilde{F}_a$. Since $L^{a'}_{a,a'} - L^a_{a',a} < 0$ in the original environment, and $\widetilde{L}^{a'}_{a,a'} - \widetilde{L}^a_{a'a}$ can be written as a continuous function in ε, we obtain that for small enough ε, $\widetilde{L}^{a'}_{aa'} - \widetilde{L}^a_{a'a} < 0$. This is a contradiction because $\widetilde{F}_{a'}$ sfosd $\widetilde{F}_a$.

Now, since L is imitative, Lemma 0 applies. Therefore, for all $a \in A$ we have

$$p'_a - p_a = \frac{1}{|W|} \sum_{c,d \in W; s(c) \neq a, s(d) = a} \Pr(c \frown d) \left(L^a_{s(c),a} - L^{s(c)}_{a,s(c)} \right).$$

Because of (ii), and the argument above, for every $a \in A^*$ and $c \in W$, we have $L^a_{s(c),a} - L^{s(c)}_{a,s(c)} \geq 0$. It follows that $p'_a - p_a \geq 0$ for all $a \in A^*$, therefore $p'(A^*) \geq p(A^*)$. Note that when $p(A^*) \in \{0,1\}$ we have that $p'(A^*) = p(A^*)$, because L is imitative. Now consider $p(A^*) \in (0,1)$. We know that $p'_{a'} - p_{a'} \geq 0$ for all $a' \in A^*$. Furthermore, by Lemma 0, for all $a' \in A^*$, we have $p'_{a'} - p_{a'} > 0$, because $p(A \setminus A^*) > 0$ and $L^{a'}_{a,a'} - L^a_{a',a} > 0$ for all $a \in A \setminus A^*$.

Necessity of (i):
Suppose that there are some $x,\, y \in [0,1]$, $a, a', a'' \in A$, $a'' \notin \{a, a'\}$ such that $L(a, x, a', y)_{a''} > 0$. Furthermore, suppose $p_a + p_{a'} = 1$, $F_a = F_{a'}$, $\mu_a(x) = \mu_a(y) = \mu_a(1) = 1/3$, and $\mu_{a'''}(0) = 1$ for all $a''' \in A \setminus \{a, a'\}$.[5] It follows that $A^* = \{a, a'\}$, therefore $p(A^*) = 1$. Suppose that there are c, $d \in W$ such that $s(c) = a$, and $s(d) = a'$, then $p'(A^*) < 1$. Thus, L has to be imitative.

Necessity of (ii):
Suppose that $F_{a'}$ sfosd F_a, but $L^{a'}_{aa'} - L^a_{a'a} \leq 0$, for some $a,\, a' \in A$. Suppose

[5]If $x \neq y = 1$, then it means that $\mu_a(1) = 2/3$ and $\mu_a(x) = 1/3$; if $x = y = 1$, then $\mu_a(1) = 1$; etc. In what follows we adopt this notation convention.

that $p_a + p_{a'} = 1$, $p_a < 1$, $p_{a'} < 1$ and that $A^* = \{a'\}$. Since L is imitative, Lemma 0 yields

$$p'_{a'} - p_{a'} = \frac{1}{|W|} \sum_{c,d \in W; s(c)=a, s(d)=a'} \Pr(c \curvearrowright d) \left(L^{a'}_{a,a'} - L^{a}_{a',a} \right).$$

Assume that $s(c) = a$ and $s(d) = a'$ for some $c, d \in W$. Then, we have that $p(A^*) \in (0, 1)$, but $L^{a'}_{aa'} - L^{a}_{a'a} \leq 0$ implies $p'(A^*) \leq p(A^*)$. □

Remark 1. Note that Lemma 1 implies that if $F_a(x) = F_{a'}(x)$ for all $x \in [0, 1]$ and L is FOM then $L^{a'}_{aa'} - L^{a}_{a'a} = 0$. From Lemma 0, this implies that every FOM rule is impartial.

As revealed by the proof of Lemma 1, the imitative property of FOM rules follows from the fact that improvement is required in every environment, regardless of the fraction of the population that is playing each action in the current period. In particular, if all the members of the population play an action in A^*, then the expected fraction of the population that plays an action in A^* in the next period will be strictly less than one unless all of them, with probability one, play either the action they played in the current period or the action played by the other member of the population they observed. This requires the behavioral rule to be imitative. This condition implies that when no individual in the population is playing a dominant action the population is unable to learn to play actions in A^*. For this reason we do not impose that $p'(A^*) > p(A^*)$ when $p(A^*) = 0$. Clearly a behavioral rule that satisfies $p'(A^*) > p(A^*)$ when $p(A^*) = 0$ and $p'(A^*) = p(A^*)$ when $p(A^*) = 1$ cannot exist.

The necessity of properties that are similar to our notion of impartiality has been found in many places in the literature. From Schlag's [15] results it follows that, if the average payoff of the population is expected to increase in every environment, then the behavioral rule of the population must verify that, in every environment where all the actions have the same expected payoffs, the expected fraction of the population who plays each action in the next period is the same as today. Similar results for individual learning have been found by Börgers, Morales, and Sarin [5] and Oyarzun and Sarin [12].

The next result provides a characterization of FOM behavioral rules in terms of the shape of $L(a, x, a', y)_{a''}$, for all a, a', $a'' \in A$. In the analysis below, we will recurrently use the concept of *net-switching function* from action a to action a'. We denote this function by $g_{aa'}(x, y)$ and it is defined as $g_{aa'}(x,y) := L(a,x,a',y)_{a'} - L(a',y,a,x)_a$. This function measures how much probability is being shifted from action a to action a' when two agents playing a and a' observe each other.

Proposition 1. *A rule L is FOM if and only if it satisfies the following conditions:*

(i) L is imitative.

(ii) For all a, $a' \in A$, the net-switching functions $g_{aa'}(x,y)$ satisfy

(ii.1) $g_{aa'}(x,y) = -g_{aa'}(y,x), \forall\ x,\ y \in [0,1]$.

(ii.2) $g_{aa'}(x,y)$ is non-decreasing in y, for all $x \in [0,1]$.

(ii.3) $g_{aa'}(x,y) > 0$, $\forall\ x,\ y \in [0,1]$ such that $y > x$.

Proof. Necessity of (ii.1):
We start by proving that $g_{aa'}(x,x) = 0$ for all $x \in [0,1]$ and all a, $a' \in A$. Consider an arbitrary $x \in [0,1]$ and an environment where $\mu_a(x) = \mu_{a'}(x) = 1$. Notice that $F_a = F_{a'}$. Therefore, by Remark 1, $L^{a'}_{aa'} - L^{a}_{a'a} = 0$. But in this environment $L^{a'}_{aa'} - L^{a}_{a'a} = g_{aa'}(x,x)$. It follows that $g_{aa'}(x,x) = 0$.

Now, for any x, $y \in [0,1]$ and some $a, a' \in A$, consider an environment where $\mu_a(x) = \mu_{a'}(x) = \mu_a(y) = \mu_{a'}(y) = 1/2$, and notice that $F_a = F_{a'}$. By Remark 1,

$$\begin{aligned} 0 &= L^{a'}_{aa'} - L^{a}_{a'a} \\ &= \frac{1}{4}(g_{aa'}(x,y) + g_{aa'}(y,x) + g_{aa'}(x,x) + g_{aa'}(y,y)) \\ &= \frac{1}{4}(g_{aa'}(x,y) + g_{aa'}(y,x)). \end{aligned}$$

Therefore, $g_{aa'}(x,y) = -g_{aa'}(y,x)$.

Necessity of (ii.2):
Suppose that for some a, $a' \in A$ and x, y', $y \in [0,1]$ such that $y' > y$, we have that $g_{aa'}(x,y') < g_{aa'}(x,y)$. Consider an environment such that $\mu_a(x) = \mu_{a'}(x) = 1-\varepsilon$, $\mu_a(y) = \mu_{a'}(y') = \varepsilon$, and notice that $F_{a'}$ sfosd F_a. It follows that

$$\begin{aligned} 0 &< L^{a'}_{aa'} - L^{a}_{a'a} \\ &= (1-\varepsilon)\varepsilon g_{aa'}(x,y') + (1-\varepsilon)\varepsilon g_{aa'}(y,x) + (1-\varepsilon)^2 g_{aa'}(x,x) + \varepsilon^2 g_{aa'}(y,y') \\ &= \varepsilon[(1-\varepsilon)\{g_{aa'}(x,y') - g_{aa'}(x,y)\} + \varepsilon g_{aa'}(y,y')]. \end{aligned}$$

The last equality holds because of the necessity of (ii.1). The term inside the brackets $\{.\}$ is negative. Thus, for small enough ε, the term inside the brackets $[.]$ is negative, causing a contradiction.

Necessity of (ii.3):
Suppose that $y > x$ and $g_{aa'}(x, y) \leq 0$. Consider a, $a' \in A$ such that $\mu_a(x) = \mu_{a'}(y) = 1$. Clearly $F_{a'}$ sfosd F_a, but $L^{a'}_{aa'} - L^{a}_{a'a} = g_{aa'}(x, y) \leq 0$, which violates Condition (ii) of Lemma 1.

Sufficiency:
Sufficiency follows directly from Lemma 3. This lemma and its proof are provided in Appendix A. □

Proposition 1 reveals the properties of the functional form of FOM behavioral rules that we described in the introduction. Besides being imitative, these rules treat every pair of actions in a pairwise symmetric way [Condition (ii.1)]. Furthermore, net-switching from action a to action a' must be a non-decreasing function of the payoff of action a' [Condition (ii.2)], and strictly positive if the payoff of a' is greater than the payoff of a [Condition (ii.3)]. This result allows us to check, directly from the functional form of a behavioral rule, whether it is FOM or not. In Section 4 we analyze a number of behavioral rules from the literature and use Proposition 1 to asses whether they are FOM.

The following corollary follows directly from Conditions (ii.1) and (ii.2) in Proposition 1.

Corollary 1. *If L is FOM then $g_{aa'}(x, y)$ is non-increasing in x, for all $y \in [0, 1]$.*

In our definition of A^*, all the distributions in this set are the same. This can be slightly generalized. If A can be partitioned in A^{**} and $A \setminus A^{**}$ in such a way that for all actions $a \in A^{**}$ and $a' \in A \setminus A^{**}$, we have that F_a sfosd $F_{a'}$, then the set of rules that satisfy $p'(A^{**}) \geq p(A^{**})$, with strict inequality when $p(A^*) \in (0, 1)$, in every environment, is equivalent to the set of FOM rules.[6]

[6]Clearly every rule that satisfies $p'(A^{**}) \geq p(A^{**})$, in every environment, with strict inequality when $p(A^{**}) \in (0, 1)$ has to be FOM. The fact that FOM rules satisfy $p'(A^{**}) \geq p(A^{**})$ in every environment, with strict inequality when $p(A^{**}) \in (0, 1)$, follows directly from Lemmas 0, 1, and straightforward algebraic manipulations.

4 Examples

In this section we provide a number of examples of behavioral rules that satisfy the properties studied above. We also show that no FOM rule can be said to be dominant in the sense of having the greatest $p'(A^*)$ in every environment among all the FOM rules. However it is possible to identify classes of environments in which a given rule may be dominant. In particular we characterize the environments for which the rule IIB is dominant. We begin with the analysis of Schlag's [15] Proportional Imitation rule. We described this rule formally in Section 2. Straightforward calculations show that $g_{aa'}(x, y) = y - x$. It follows that this rule is FOM. Schlag [15] shows that this rule has a number of interesting properties. For example, it has the greatest increase in the expected average payoff of the population among all the improving behavioral rules.

Now we analyze the rule IIB we introduced in Section 2. For this rule

$$g_{aa'}^{IIB}(x, y) = \begin{cases} 1 & \text{if } y > x \\ 0 & \text{if } y = x \\ -1 & \text{if } y < x \end{cases}$$

for all a, $a' \in A$.

This rule is FOM, yet, as shown by Schlag [15], IIB is not improving, so it is not expected to increase the expected average payoff of the population in every environment. Can we claim that any FOM rule is better than the others in any sense? Or more important, can we find any rule that is the "best"? A natural notion to assess the performance of different rules is based on the expected increase in the number of individuals that play a dominant action in the next period. Therefore, a rule may be said to be the best if the expected increase in the fraction of the population that plays a dominant action in the next period is bigger than for any other rule, in every environment, regardless the initial state and sampling process. This is captured in the following definition.

Definition 4. A FOM rule L is dominant if $p'_L(A^*) - p(A^*) \geq p'_{L'}(A^*) - p(A^*)$ for every FOM rule L' in every environment.

The next result shows that such a behavioral rule cannot exist.

Proposition 2. *FOM dominant rules do not exist.*

Proof. The proof proceeds in two steps. First, we show that if a dominant rule exists, it has to be the rule Imitate if Better.[7] This rule satisfies

$$g^{IIB}_{aa'}(x,y) = \begin{cases} 1 & \text{if } y > x \\ 0 & \text{if } y = x \\ -1 & \text{if } y < x \end{cases}$$

for all x, $y \in [0,1]$, and a, $a' \in A$. Consider any other FOM rule L' such that $g'_{aa'}(x,y) < 1$ for some $y > x$. Consider an environment such that $\mu_{a'}(y) = \mu_a(x) = 1$. Assume $p_a + p_{a'} = 1$, p_a, $p_{a'} \in (0,1)$, and $A^* = \{a'\}$. Then,

$$\begin{aligned} p'_{L'}(A^*) - p(A^*) &= p'_{a'} - p_{a'} \\ &= \frac{1}{|W|} \sum_{c,d\in W; s(c)=a, s(d)=a'} \Pr(c \curvearrowright d)(L'^{a'}_{aa'} - L'^{a}_{a'a}) \\ &= \frac{1}{|W|} \sum_{c,d\in W; s(c)=a, s(d)=a'} \Pr(c \curvearrowright d) g'_{aa'}(x,y) \\ &< \frac{1}{|W|} \sum_{c,d\in W; s(c)=a, s(d)=a'} \Pr(c \curvearrowright d) \\ &= p'_{IIB}(A^*) - p(A^*). \end{aligned}$$

It follows that if a dominant rule exists, it has to be the rule Imitate if Better. Now, we prove that this rule is not dominant either. Consider an environment such that $\mu_{a'}(.05) = \mu_{a'}(.9) = \mu_a(.05) = \mu_a(.1) = 1/2$, and consider a FOM rule L' such that

$$g'_{aa'}(x,y) = \begin{cases} 1 & \text{if } y \geq 0.8 \text{ and } x \leq 0.2 \\ -1 & \text{if } x \geq 0.8 \text{ and } y \leq 0.2 \\ \frac{1}{2}(1_{\{y>x\}} - 1_{\{x>y\}}) & \text{otherwise.} \end{cases}$$

It is easy to compute that for this rule, $L'^{a'}_{a,a'} - L'^{a}_{a',a} = 3/8$, while $IIB^{a'}_{a,a'} - IIB^{a}_{a',a} = 2/8$. As above, assume that $p_a + p_{a'} = 1$ and $A^* = \{a'\}$. With similar calculations to those in the first step, we obtain

$$p'_{L'}(A^*) - p(A^*) > p'_{IIB}(A^*) - p(A^*).$$

Thus, the rule IIB is not dominant either. □

[7]By IIB rules we mean any behavioral rule with net-switching functions $g_{aa'}(x,y) = g^{IIB}_{aa'}(x,y)$. In particular, $L(a,x,a',x)_{a'}$ does not need to be 1/2 as in our earlier example.

Proposition 2 shows that no FOM rule is dominant. However, the proof of this result suggests that there are classes of environments within which a behavioral rule may be dominant. For example, as the first part of that proof reveals, the rule IIB is dominant within all the environments where the support of the payoff distribution of each action is a singleton. This result can be strengthened further as formalized in the following proposition, which provides a full characterization of the environments for which the rule IIB is dominant.

Proposition 3. *Let Γ be the class of all sets $B \subset [0,1]^2$ which satisfy*

$$(x, y) \in B \Rightarrow (y, x) \in B$$

and for $y > x$

$$(x, y) \in B \Rightarrow (x, z) \in B \text{ and } (z, y) \in B \;\forall\; z \in (x, y).$$

The set of environments for which the rule IIB is dominant corresponds to exactly all those environments for which the condition

$$\int\int_B 1_{\{y>x\}} dF_a(x) dF_{a'}(y) \geq \int\int_B 1_{\{y<x\}} dF_a(x) dF_{a'}(y)$$

is verified for all $B \in \Gamma$, $a \in A \backslash A^*$, and $a' \in A^*$.

Proof. First, consider an environment which does not satisfy the characterization, that is, there are actions $a \in A \backslash A^*$ and $a' \in A^*$ such that, for some set $B \in \Gamma$, we have

$$\int\int_B 1_{\{y>x\}} dF_a(x) dF_{a'}(y) < \int\int_B 1_{\{y<x\}} dF_a(x) dF_{a'}(y).$$

Consider the rule L' with net-switching function $g'_{aa'}(x, y)$ given by

$$g'_{aa'}(x, y) = \begin{cases} \frac{1}{2}(1_{\{y>x\}} - 1_{\{x>y\}}) & \text{if } (x, y) \in B \\ g^{IIB}_{aa'}(x, y) & \text{otherwise.} \end{cases}$$

All the other net-switching functions involving any action $a'' \in A \backslash \{a, a'\}$ (if any) are the same as for the rule IIB. Using Proposition 1 and the conditions on B it can be verified that L' is FOM. Straightforward calculations reveal that

$$\int\int g'_{aa'}(x, y) dF_a(x) dF_{a'}(y) > \int\int g^{IIB}_{aa'}(x, y) dF_a(x) dF_{a'}(y),$$

which implies that L' outperforms IIB.
Now consider any environment which satisfies the hypothesis. Suppose that there exists another FOM rule L', with net switching functions $g'_{aa'}$, which outperforms IIB in this environment. This implies that there exists some $\epsilon > 0$ such that, for some actions $a \in A \backslash A^*$ and $a' \in A^*$, we have that $\int\int g'_{aa'}(x,y)dF_a(x)dF_{a'}(y) - \int\int g^{IIB}_{aa'}(x,y)dF_a(x)dF_{a'}(y) > \epsilon$. Choose n large enough so that $\frac{1}{n} < \epsilon$. For $k = 1, \ldots, n$, let

$$B^k := \left\{(x,y) \in [0,1]^2 : \frac{k-1}{n} < g'_{aa'}(x,y) \leq \frac{k}{n}\right\},$$

$$\tilde{B}^k := \left\{(x,y) \in [0,1]^2 : -\frac{k-1}{n} > g'_{aa'}(x,y) \geq -\frac{k}{n}\right\}$$

and

$$g^0_{aa'} := \sum_{k=1}^{n} \frac{k}{n}(1_{B^k} - 1_{\tilde{B}^k}).$$

Using Proposition 1, $g^0_{aa'}$ can be shown to be a FOM rule. Since $\frac{1}{n} < \epsilon$, we also have that $|g'_{aa'}(x,y) - g^0_{aa'}(x,y)| < \epsilon$ for all $(x,y) \in [0,1]^2$, and thus $\int\int g^0_{aa'}(x,y)dF_a(x)dF_{a'}(y) > \int\int g^{IIB}_{aa'}(x,y)dF_a(x)dF_{a'}(y)$. Now let

$$g^1_{aa'} := g^0_{aa'} + \frac{1}{n}\left(1_{\cup_{k=1}^{n-1} B^k} - 1_{\cup_{k=1}^{n-1} \tilde{B}^k}\right).$$

For $k = 2, \ldots, n-1$, define inductively the functions $g^k_{aa'}$ as

$$g^k_{aa'} := g^{k-1}_{aa'} + \frac{1}{n}\left(1_{\cup_{i=1}^{n-k} B^i} - 1_{\cup_{i=1}^{n-k} \tilde{B}^i}\right).$$

It follows that $g^{n-1}_{aa'} = g^{IIB}_{aa'}$. Using our assumptions about the environment, for all $i = 1, \ldots, n-1$, we have

$$\int\int 1_{\cup_{k=1}^{i} B^k} dF_a dF_{a'} \geq \int\int 1_{\cup_{k=1}^{i} \tilde{B}^k} dF_a dF_{a'}$$

and thus for all $k = 1, \ldots, n-1$,

$$\int\int g^k_{aa'}(x,y)dF_a(x)dF_{a'}(y) \geq \int\int g^{k-1}_{aa'}(x,y)dF_a(x)dF_{a'}(y).$$

Therefore,

$$\begin{aligned}\int\int g^{IIB}_{aa'}(x,y)dF_a(x)dF_{a'}(y) &= \int\int g^{n-1}_{aa'}(x,y)dF_a(x)dF_{a'}(y)\\ &\geq \int\int g^{0}_{aa'}(x,y)dF_a(x)dF_{a'}(y)\\ &> \int\int g^{IIB}_{aa'}(x,y)dF_a(x)dF_{a'}(y),\end{aligned}$$

which is a contradiction. □

Intuitively the condition in Proposition 3 guarantees that given that both payoffs are greater than an arbitrary x and smaller than an arbitrary y, the probability of obtaining the higher payoff with the dominant action is greater than the probability of obtaining the higher payoff with the dominated action. For environments where this condition is not satisfied, rules which do not change with probability one when differences in the payoffs are observed can outperform IIB as the example of the second part of the proof of Proposition 2 shows. It is easy to verify that all the environments where almost surely the payoffs provided by a dominant action are greater than the payoffs provided by a non-dominant action are included in the set of environments characterized above. However when this condition is not verified, the construction of environments which satisfy the characterization is not trivial. An example of an environment which satisfies the characterization, and where the supports of dominant and non-dominant actions exhibit considerable overlapping, is the following: for all actions $a \in A\backslash A^*$ and $a' \in A^*$ we have $\mu_a(0.1) = \mu_a(0.5) = \mu_a(0.6) = 1/3$ and $\mu_{a'}(0.5) = \mu_{a'}(0.6) = 1/2$. The discussion above suggests that if individuals have some information about the environment, then they may use a rule which performs well in such environment. For example, if the payoffs obtained with each action are not random and decision makers are aware of this (either by experience with that particular problem or previous knowledge of this property of the environment), then they may use the rule IIB. An analogous argument could be applied to argue that a different rule may be used in a different environment which satisfies some basic property that allows the individual to identify the rule which performs the best in such a problem. We discuss this point further in Section 6.[8]

[8]We thank a referee for rising this issue.

5 Large population dynamics

In this section we analyze the dynamics of the fraction of the population playing a dominant action in large populations. We provide conditions under which, with a probability arbitrarily close to one, the fraction of the population that plays a dominant action is arbitrarily close to one after a finite number of periods. This can be achieved for a large enough population of individuals who use a FOM rule provided that either a fraction of the population play a dominant action in the first period or some experimentation takes place, but vanishes as time goes to infinity.

In the rest of this section we will assume that $\Pr(c \curvearrowright d) = 1/(|W|-1)$ for all c, $d \in W$ such that $c \neq d$. We will also assume that the sampling process is independent.[9] We use these assumptions below, when we approximate the behavior of a finite population by analyzing the dynamics of a population that is a continuum.

In order to allow for experimentation, in this section we allow behavioral rules to depend on time. In particular, we consider a population in which each individual uses a FOM rule, but with some probability experiments any action with some positive probability. Formally, we define an *experimentation sequence* as a sequence $\{(e_a^t)_{a\in A}\}_{t=0}^{\infty}$ where $e_a^t \in [0,1]$ for all $a \in A$ and $t = 0, 1, ...$ and define $e^t := \sum_a e_a^t \leq 1$. In every period, each individual uses a FOM rule L with probability $(1 - e^t)$, and experiments (plays each action a with probability e_a^t/e^t) with probability e^t. The resulting rule is denoted by $\widetilde{L}^t$, i.e.,

$$\widetilde{L}^t(a', x, a'', y)_a := (1 - e^t)L(a', x, a'', y)_a + e_a^t \quad \forall\, a \in A,\; t = 0, 1, ...$$

We call these rules *experimentally first-order monotone* (EFOM).

The analysis below benefits from studying the dynamical system

$$p_a^{t+1} = \sum_{a'\in A}\sum_{a''\in A} p_{a'}^t p_{a''}^t \int\int \widetilde{L}^t(a', x, a'', y)_a dF_{a'}(x) dF_{a''}(y),\ \forall a \in A,\, t = 0, 1, ...$$

with $p_a^0 \in [0,1]$ given for all $a \in A$ and $\sum_{a\in A} p_a^0 = 1$. This system is called the *dynamics of an infinite population* because this expression can be interpreted as the motion equation of the fraction of the population

[9] Independent sampling means that the joint probability of the events c samples d and d samples c is given by $\Pr(c \curvearrowright d)\Pr(d \curvearrowright c)$ for all c, $d \in W$ such that $c \neq d$.

that plays action a in a population that is a continuum. Here each individual samples another member of the population according to a uniform distribution. For each action $a' \in A$, the expected fraction of the population that plays action a' in this period and a in the next period is given by $p^t_{a'} \sum_{a'' \in A} p^t_{a''} \int \int \widetilde{L}^t(a', x, a'', y)_a dF_{a'}(x) dF_{a''}(y)$. Adding across actions $a' \in A$, we obtain that the total fraction of the population that will play action a in the next period, p^{t+1}_a, corresponds to the expression above. Notice that this system depends on the behavioral rule $\widetilde{L}^t$ used by the individuals of the population. The fraction of the population that plays an action in A^* at time t will be denoted by $p^t(A^*)$. In other words, $p^t(A^*) := \sum_{a \in A^*} p^t_a$. Straightforward calculations reveal the following result.

Remark 2. The dynamics of an infinite population associated to an EFOM behavioral rule $\widetilde{L}^t$ can be simplified to

$$p^{t+1}_a = [p^t_a + p^t_a \sum_{a' \in A} p^t_{a'} (L^a_{a'a} - L^{a'}_{aa'})](1 - e^t) + e^t_a, \qquad \forall\, a \in A,\ t = 0, 1, ...$$

It follows then that, for all $t = 0, 1, ...$, we have

$$p^{t+1}(A^*) - p^t(A^*) = \sum_{a \in A^*} p^t_a \sum_{a' \in A \setminus A^*} p^t_{a'} (L^a_{a'a} - L^{a'}_{aa'})(1 - e^t) + e^t \left(\frac{\sum_{a \in A^*} e^t_a}{e^t} - p^t(A^*) \right).$$

In the rest of this section we assume that $e^t > 0$ for some $t = 0, 1, 2...$, $\lim_{t \to \infty} e^t = 0$, and $e^t > 0$ implies $e^t_a > 0$ for all $a \in A$ and $t = 0, 1....$ This means that individuals experiment and, furthermore, each action is played with positive probability when experimentation takes place. Experimentation rates, however, vanish as time goes to infinity. An example of such an experimentation can be obtained by setting $e^t_a = \frac{1}{|A| t}$ for all $a \in A$ and $t = 0, 1, ...$

The next result shows that, if the behavioral rule used by the population is EFOM, then, $p^t(A^*)$ converges to one for the dynamics of an infinite population. If a FOM rule is used, the same result may be obtained assuming that in the first period a fraction of the population plays a dominant action.

Lemma 2. *The dynamics of an infinite population of individuals who use an EFOM rule, or use a FOM rule with* $p^0(A^*) > 0$, *satisfy* $\lim_{t \to \infty} p^t(A^*) = 1$.

Proof. Suppose $\alpha := \limsup_{t\to\infty} p^t(A^*) < 1$ and let $p^* := \alpha + \frac{1-\alpha}{2}$. Also, define

$$\gamma := \min_{a\in A^*, a'\in A\setminus A^*} \{L^a_{a'a} - L^{a'}_{aa'}\} > 0.$$

If $e^t > 0$ infinitely often then, $p^t(A^*) > 0$ infinitely often. If $e^t > 0$ finitely many times (or never) then $p^t(A^*) > 0$ infinitely often because after the last period in which $e^t > 0$, $\widetilde{L}^t$ becomes a FOM rule and thus $p^t(A^*)$ becomes a monotonely increasing sequence.

Since $p^t(A^*) > 0$ infinitely often, $\lim_{t\to\infty} e^t = 0$, and $p^* > \alpha$, there exists a t^* such that $p^{t^*}(A^*) > 0$; and $(1-p^*)\gamma(1-e^{t^*}) - e^{t^*} > \frac{\gamma(1-p^*)}{2}$ and $p^t(A^*) < p^*$ for all $t \geq t^*$. Then, we obtain

$$\begin{aligned} p^{t^*+1}(A^*) &> p^{t^*}(A^*) + p^{t^*}(A^*)((1-p^*)\gamma(1-e^t) - e^t) \\ &> p^{t^*}(A^*) + \frac{p^{t^*}(A^*)\gamma(1-p^*)}{2}. \end{aligned}$$

Thus, inductively, we obtain $p^{t^*+n}(A^*) > p^{t^*}(A^*) + np^{t^*}(A^*)\gamma(1-p^*)/2$, $n = 1, 2...$, which leads to a contradiction.

Now we prove that $\lim_{t\to\infty} p^t(A^*) = 1$. For any $\epsilon \in (0,1)$ let $p^* := 1 - \frac{\epsilon}{2}$ and define γ as above. There exists a t^* such that for all $t \geq t^*$ we have $e^t < \min\{\frac{\epsilon}{4}, \frac{(1-p^*)\gamma}{2(1+(1-p^*)\gamma)}\}$. Since $\limsup_{t\to\infty} p^t = 1$, there exists $\tilde{t} > t^*$ such that $p^{\tilde{t}} > p^*$. Now observe that for all $t > t^*$ we have that p^{t+1} is bounded below by $p^t - \frac{\epsilon}{4}$ and the inequality $p^{t+1}(A^*) > p^t(A^*)$ holds (by the same argument as above) whenever $p^t(A^*) \leq p^*$. It follows that $p^{\tilde{t}+1} \geq p^*$ or $p^{\tilde{t}+1} \in (1-\epsilon, p^*)$, which implies that $p^{\tilde{t}+2} > p^{\tilde{t}+1}$. In either case we have that $p^t(A^*) > 1-\epsilon$ for all $t \geq \tilde{t}$. □

In our next result, we use Theorem 3 in Schlag [15] and Lemma 2 to prove that, with a probability arbitrarily close to one, the fraction of the population that plays a dominant action is arbitrarily close to one after a finite number of periods, provided that the population is large enough, and either (i) decision makers use an EFOM rule, or (ii) at least one individual in the population plays a dominant action in the initial period and decision makers use a FOM rule. Let $p^{t,N}_a$ denote the fraction of the population that, at time t, plays a in a population of size N and let $p^{t,N}(A^*) := \sum_{a\in A^*} p^{t,N}_a$.

First, we provide a generalized version of Theorem 3 in Schlag [15].[10]

Theorem. *(Schlag [15]) Assume that all individuals of a population use the rule* $\widetilde{L}^t$ *and* $p_a^{0,N} = p_a^0$ *for all* $a \in A$*. Then, for every* ε, $\delta > 0$, $T \in \{1, 2, ...\}$, *and* $(p_a^{0,N})_{a\in A}$, *there exists* $N_0 \in \{1, 2, ...\}$ *such that for any population of size* $N > N_0$, *the event* $\{\left[\sum_{a\in A}(p_a^{T,N} - p_a^T)^2\right]^{1/2} > \delta\}$ *occurs with probability less than* ε.

Proposition 4. *If the population uses an EFOM rule, or* $p^{0,N}(A^*) > 0$ *and the population uses a FOM rule, then for all* $\varepsilon, \delta > 0$, *there exist* N_0, $T < \infty$ *such that for all* $N > N_0$, $P(1 - p^{T,N}(A^*) > \delta) < \varepsilon$.

Proof. By Lemma 2, we know that there exists a natural number T such that $1 - p^T(A^*) < \delta/2$. Now, by Theorem 3 in Schlag [15], for this T there exists N_0 such that for all $N > N_0$ we have

$$P(|p^T(A^*) - p^{T,N}(A^*)| > \delta/2) < \varepsilon.$$

Then, by the triangle inequality, we obtain that $P(1-p^{T,N}(A^*) > \delta) < \varepsilon$. □

It is worth noting that the force driving the population to choose a first-order dominant action is first-order monotonicity. Nevertheless, and since FOM rules are imitative, initial experimentation plays a key role by allowing a part of the population to play a dominant action, so that the rest of the population can imitate it later.[11]

Finally, we shall point out that the FOM rules L which we use to construct EFOM rules do not need to be time-homogeneous. The following result formalizes this.

Remark 3. Consider a population of individuals who use a rule $\widetilde{L}^t$ given by

$$\widetilde{L}^t(a', x, a'', y)_a := (1 - e^t)L^t(a', x, a'', y)_a + e_a^t \quad \forall\, a \in A,\ t = 0, 1, ...$$

where L^t is a FOM rule for $t = 0, 1, ...$and $\{(e_a^t)_{a\in A}\}_{t=0}^{\infty}$ is an experimentation sequence which satisfies

$$\sum_{t=1}^{\infty}\left(c(1 - e^t)\gamma^t - e^t\right) = \infty \qquad \forall\, c > 0,$$

[10]It is straightforward to verify that this theorem generalizes to the case of time dependent rules as the ones we consider in this section.

[11]We thank an anonymous referee who suggested that we introduce experimentation as an alternative to just assuming $p^{N,0}(A^*) > 0$.

with $\gamma^t := \min_{a\in A^*,a'\in A\setminus A^*}\{L^{ta}_{a'a} - L^{ta'}_{aa'}\}$, for all environments. Then for all $\varepsilon, \delta > 0$, there exist N_0, $T < \infty$ such that for all $N > N_0$, $P(1 - p^{T,N}(A^*) > \delta) < \varepsilon$.[12]

The condition of this remark is satisfied, for example, if finitely many rules are used. Intuitively, it means that if the monotone dynamic effect of the sequence of FOM rules vanishes, it does it slowly enough.

6 Discussion

In this section we consider possible directions for future research. The motivation for the construction of FOM rules is based on the performance of these rules in terms of the fraction of the population that is expected to play a dominant action in the next period. The analysis provided so far has not considered how the way in which each individual uses the information she gets helps her improve her own payoffs. Can the construction of FOM rules be derived from the analysis of individual performance? A first observation is that FOM rules are imitative. As revealed by the analysis, this feature follows from fundamental characteristics of the social learning process and the properties, in terms of social performance, to be obtained: imitation is required to avoid that a population, already playing dominant actions, experiment non-dominant choices. In the analysis of individual choices, such an argument does not apply. Each individual must be exposed to go from dominant to non-dominant choices when learning takes place. As a consequence, when studying improved performance properties at the individual level, imitation cannot be derived from the analysis as we did in the case of improved performance at the social level. Therefore, in order to derive FOM rules from properties related to individual level performance, imitation must be *assumed* –not *derived*. In this sense, an individual level derivation of FOM rules takes imitative behavior as given, but provides the specific ways in which imitation has to be implemented in order to satisfy improved

[12]Note that in order to obtain this result and Proposition 4 we rely on the fact that $\limsup_{t\to\infty} p^t(A^*) = 1$. In particular, it is not necessary that $\lim_{t\to\infty} p^t(A^*) = 1$. We can also obtain that $\lim_{t\to\infty} p^t(A^*) = 1$ under the assumptions of Remark 3 if we additionally assume that for all $c > 0$, there exists t^*_c such that $(c(1 - e^t)\gamma^t - e^t) \geq 0$ for all $t > t^*_c$ and all environments.

performance at the individual level.[13] Formally, we can characterize a family of imitative rules such that, if the payoff distribution of one of the observed actions first-order stochastically dominates the distribution associated with the other observed action, then the expected probability of playing the dominant action is higher than the expected probability of playing the dominated action. We call such rules *individually monotone* (IM). The characterization of these rules reveals that, indeed, they are a refinement of the family of FOM rules. The formal analysis of IM rules is provided in Appendix C.

A natural extension of the analysis in this paper relates to the concept of second-order stochastic dominance. When payoffs are monetary, very often decision makers are concerned about risk. Usually this concern leads them to refrain from playing actions providing higher expected payoffs when they are associated with a higher risk. Are there specific features that one can impose on behavioral rules in order to lead the population to choices consistent with risk aversion? In order to introduce such concerns, a similar analysis to the one provided for first-order stochastic dominance and FOM rules can be developed for second-order stochastic dominance. This analysis reveals the conditions that need to be imposed on rules so that they lead the population to make safer choices. Formally, we can characterize a family of rules such that, in environments where all the actions have the same expected payoffs, the expected fraction of the population playing in the next period an action that second-order stochastically dominates all the other actions in the set is higher than that fraction in the current period. We call such rules *second-order monotone* (SOM) and their characterization is provided in Appendix D. The characterization of SOM rules is analogous to the characterization of FOM rules, except that the net-switching functions are concave instead of increasing in the payoff of the action that receives the probability. It follows that an interesting subset of FOM rules is the one with rules that are both FOM and SOM at the same time, i.e., those with increasing, concave net-switching functions. These rules lead to rational choice in the sense of first-order monotonicity and, at the same time, are consistent with risk-averse decision making.

As discussed in the introduction, our motivation to study FOM rules was identifying a set of behavioral rules that lead individuals to choose actions in a way that is consistent with the decisions of fully informed rational agents.

[13]Indeed, imitative behavior is sometimes considered an intrinsic characteristic of the human decision process. For example, see Cubitt and Sugden [6].

However, one can think about a motivation for FOM rules based on arguments related to evolutionary theory. For example, Schlag's [15] improving rules are motivated by their performance under selection pressure. If survival of behavioral rules depends on their average payoff, then a successful rule must be able to lead decision towards expected payoff maximizing actions. A formal analysis of these ideas is provided in Björnerstedt and Schlag [3]. An interesting direction for further research is finding the conditions under which FOM rules may prevail under selection pressure. More specifically, it would be interesting to find the specific ways (if any) in which survival must be determined in order to allow FOM rules to be dominant. The analysis in Robson [13] and To [19] is also related.

A restrictive feature in our analysis is the assumption that the payoff-distributions associated to the different actions are the same for all individuals in the population. In many settings one would like to allow for different payoff distributions for agents with different characteristics. In Ellison and Fudenberg [7] heterogeneity of the population makes it harder for a social learning process to lead the population to play the action that provides the highest expected payoff. Future research could extend the analysis to allow for this possibility and try to identify rules that lead to (expected) better decisions in the future in heterogenous environments. Future research could also identify and characterize properties based on notions analogous to the ones studied in our paper, but suitable to the information contexts of different models, for example when two or more other individuals are sampled.

We have shown that no FOM rule performs better than any other FOM rule in every environment. Yet, it is possible to identify classes of environments within which a given FOM rule is dominant. In particular we characterized the class of environments for which the rule IIB is dominant within the FOM rules. Future research could identify other classes of environments where a given rule is dominant. Another interesting question concerns the circumstances under which individuals may be able to figure out that the environment they face belongs to a class within which a given rule is dominant. For example, individuals may be aware that they face an environment in which payoffs are not random within each action and therefore they may use the rule IIB. Alternatively, such knowledge may arise from experience.[14] Fu-

[14] A related work in this direction is Stahl [17], [18]. In these papers, individuals consider alternative rules to update their behavior and these rules are reinforced according to performance. However, in the information context of those papers, individuals receive much more feedback than in the model studied here, specially concerning to forgone payoffs.

ture research could study how, in the limited feedback context of this model, agents may keep track of the performance of a behavioral rule and eventually switch to the rule that performs the best in the environment they are facing. Alternatively, a rule may be selected by evolution as in the work of Björnerstedt and Schlag [3] discussed above. Future research could also attempt to identify alternative ways in which the probabilities of survival can be determined so that in a given class of environments the corresponding dominant rule is selected.

Our concept of FOM rules serves to provide a foundation for a number of behavioral rules in the literature, including very simple rules such as IIB and relatively more sophisticated rules such as the Proportional Imitation rule. We show that among all the FOM rules, none of them outperforms all the others in every environment. Therefore we do not single out a best rule. This non-uniqueness parallels non-uniqueness in rational choice. In our view, a next step in order to bound the set of behavioral rules should be based on experimental evidence. Future research could test if experimental subjects use behavioral rules that are FOM and try to find out what are the specific shapes of the rules they use. The hypothesis of risk-averse imitation described above may be of interest as well.

A Technical lemma

Lemma 3. *Consider a function* $g : [0,1]^2 \to [-1,1]$ *that satisfies*

(i) $g(x,y) = -g(y,x)$, $\forall\ x,\ y \in [0,1]$.

(ii) $g(x,y)$ *is non-decreasing in* y, *for all* $x \in [0,1]$.

(iii) $g(x,y) > 0$, $\forall\ x,\ y \in [0,1]$ *such that* $y > x$.

Then, for any two independent random variables X *and* Y, *taking values in* $[0,1]$ *such that* Y *sfosd* X, *we have that* $E(g(X,Y)) > 0$.

Proof. The proof proceeds in two steps:

Step 1: We first prove that Y sfosd X implies $P(Y > X) > P(Y < X)$. Let $M := \{z \in [0,1] : F_X(z) > F_Y(z)\}$ denote the set of all points in $[0,1]$ where the distribution function of X is strictly larger than the distribution function of Y. M is non-empty since Y sfosd X. We show $P_X(M) > 0$, where P_X is the measure induced by X. If $0 \in M$ then $P_X(M) \geq F_X(0) > 0$. So, without loss of generality we will assume $0 \notin M$. Let $\tilde{z} \in M$. Define $z_0 := \sup\{z : z \leq \tilde{z}, F_X(z) = F_Y(z)\} \geq 0$. If $z_0 \in M$, then $\lim_{z \to z_0-} F_X(z) = \lim_{z \to z_0-} F_Y(z)$ because of the definition of z_0 and

$$\begin{aligned} P_X(M) &\geq P(X = z_0) \\ &= F_X(z_0) - \lim_{z \to z_0-} F_X(z) \\ &= F_X(z_0) - \lim_{z \to z_0-} F_Y(z) \\ &\geq F_X(z_0) - F_Y(z_0) \\ &> 0. \end{aligned}$$

If $z_0 \notin M$, then $F_X(z) > F_Y(z)$ for all $z \in (z_0, \tilde{z}] =: N$ and

$$\begin{aligned} P_X(M) &\geq P_X(N) \\ &= F_X(\tilde{z}) - F_X(z_0) \\ &= F_X(\tilde{z}) - F_Y(z_0) \\ &\geq F_X(\tilde{z}) - F_Y(\tilde{z}) \\ &> 0. \end{aligned}$$

We define now $h(a,b) := 1_{\{b>a\}}$ and $\triangle h(a,b) := h(a,b) - h(b,a)$. For all $a \in M$, we have the inequality

$$\begin{aligned} E(\triangle h(a,Y)) &= 1 - F_Y(a) - F_Y(a-) \\ &> 1 - F_X(a) - F_X(a-) \\ &= E(\triangle h(a,X)). \end{aligned}$$

The inequality holds weakly for $a \notin M$. Let $\tilde{X} \stackrel{\mathrm{d}}{=} X$ be a random variable which is independent from X and Y. Then we get

$$\begin{aligned} E(\triangle h(X,Y)) &= E(E(\triangle h(X,Y)|X)) \\ &> E(E(\triangle h(X,\tilde{X})|X)) \\ &= E(\triangle h(X,\tilde{X})) \\ &= 0 \end{aligned}$$

and therefore, using the definition of $\triangle h$, we obtain $P(Y > X) > P(Y < X)$.

Step 2: Because of the monotonicity of the probability measure P, there exists $\epsilon > 0$ such that $P(Y - X > \epsilon) > P(X > Y)$. Let $D := \{(x,y) : x, y \in [0,1], y - x > \epsilon\}$. Now, we prove that

$$c := \inf\{g(x,y) : x, y \in D\} > 0.$$

Let $k := \inf\{n \in \mathbb{N} : n > 2/\epsilon\}$ and $I_i := [0, \frac{i-1}{k}] \times [\frac{i}{k}, 1]$ for all $i \in \{2, ..., k-1\}$. It is easy to see that $D \subset \bigcup_{i=2}^{k-1} I_i$. Then $c_i := \inf\{g(x,y) : x, y \in I_i\} = g(\frac{i-1}{k}, \frac{i}{k}) > 0$ for all $i \in \{2, ..., k-1\}$ because of the monotonicity of g and assumption (iii). Then $c \geq \min\{c_i : i \in \{2, ..., k-1\}\} > 0$.

Let

$$g_1(x,y) := (g(x,y) - c) \cdot 1_{\{g(x,y) \geq c\}} + (g(x,y) + c) \cdot 1_{\{g(x,y) \leq -c\}},$$

and

$$g_2(x,y) := c \cdot (1_{\{g(x,y) \geq c\}} - 1_{\{g(x,y) \leq -c\}}) + g(x,y) \cdot 1_{\{-c < g(x,y) < c\}}.$$

We have $g(x,y) = g_1(x,y) + g_2(x,y)$ where g_1, g_2 satisfy Conditions (i) and (ii) of the hypothesis. It follows that $E(g_1(X,Y)) \geq 0$ because

$$\begin{aligned} E(g_1(X,Y)) &= \int\int g_1(x,y) dF_Y(y) dF_X(x) \\ &\geq \int\int g_1(x,y) dF_X(y) dF_X(x) \\ &\geq \int\int g_1(x,y) dF_Y(x) dF_X(y) \\ &= -\int\int g_1(y,x) dF_Y(x) dF_X(y) \\ &= -E(g_1(X,Y)), \end{aligned}$$

where we have used the independence of X and Y and the fact that $g_1(x,y)$ is non-decreasing in the second argument and non-increasing in the first argument. On D, we have $g_2(x,y) = c$ and on $C := \{(x,y) : x > y\}$, we have $g_2(x,y) \geq -c$. Thus, we get

$$E(g_2(X,Y)) \geq c \cdot P(D) - c \cdot P(C) > 0.$$

Summarized, we have $E(g(X,Y)) > 0$. □

B Heterogeneous behavioral rules

This appendix provides a characterization of first-order monotonicity when behavioral rules may be different across individuals and each individual may imitate different observed individuals in different ways.

For each $c \in W$, a behavioral rule is a function $L(c) : A \times [0,1] \times W \times A \times [0,1] \to \Delta(A)$. These functions map each tuple (a, x, d, a', y) to a vector $L(c)(a, x, d, a', y)$. Here a is the action chosen by c in the current period, x is the payoff she obtains, a' is the action chosen by d (who is observed by c) and y is the payoff he obtains. The element $L(c)(a,x,d,a',y)_{a''}$ of $L(c)(a,x,d,a',y)$ denotes the probability with which c will play action a'' in the next period. A vector of behavioral rules $L := (L(c))_{c\in W}$ is called a *behavioral process.* Let $L^{a''}_{a,a'}(c,d)$ be the expected probability, before the realization of the payoffs, of choosing action a'' tomorrow by c when she played action a and observed d, who played action a', i.e., $L^{a''}_{a,a'}(c,d) := \int\int L(c)(a,x,d,a',y)_{a''} dF_a(x) dF_{a'}(y)$.

The expected fraction of the population that will play action $a \in A$ in the next period, p'_a, can be computed as

$$p'_a = \frac{1}{|W|}\sum_{c\in W}\sum_{d\in W\setminus\{c\}} \Pr(c \curvearrowright d) L^a_{s(c),s(d)}(c,d).$$

Likewise, $p'(A^*)$ is given by

$$p'(A^*) = \sum_{a\in A^*}\frac{1}{|W|}\sum_{c\in W}\sum_{d\in W\setminus\{c\}} \Pr(c \curvearrowright d) L^a_{s(c),s(d)}(c,d).$$

The concept of imitative behavioral rules generalizes in the obvious way to the case of heterogenous behavioral rules. In the same way, the concept of impartial and FOM behavioral rules generalizes to the concept of impartial and FOM processes. Now we proceed to characterize FOM processes.

Lemma 0*. *If $L(c)$ is imitative for all $c \in W$, then*

$$p'_a - p_a = \frac{1}{|W|}\sum_{c\in W,\ s(c)\neq a}\ \sum_{d\in W,\ s(d)=a} \Pr(c \curvearrowright d)(L^a_{s(c),a}(c,d) - L^{s(c)}_{a,s(c)}(d,c)).$$

Lemma 1*. *An imitation process L is FOM if and only if it satisfies the following conditions:*

(i) $L(c)$ is imitative for all $c \in W$.

(ii) $F_{a'}$ fosd $F_a \Rightarrow L^{a'}_{aa'}(c,d) - L^a_{a'a}(d,c) \geq 0$, $\forall$ a, $a' \in A$, $c, d \in W$, for all environments.

(iii) For all a, $a' \in A$, $F_{a'}$ sfosd $F_a \Rightarrow L^{a'}_{aa'}(c,d) - L^a_{a'a}(d,c) > 0$, for some $c \in V$ and $d \in W \setminus V$, for all $V \subset W$ such that V is not empty and $V \neq W$, in every environment.

Proof. Sufficiency:

Since $L(c)$ is imitative for all $c \in W$, Lemma 0* applies. Therefore, for all $a \in A$ we have

$$p'_a - p_a = \frac{1}{|W|} \sum_{c,d \in W; s(c) \neq a, s(d)=a} \Pr(c \curvearrowright d)(L^a_{s(c),a}(c,d) - L^{s(c)}_{a,s(c)}(d,c)).$$

Now, because of (ii), for every $a \in A^*$ and $c \in W$, we have $L^a_{s(c),a}(c,d) - L^{s(c)}_{a,s(c)}(d,c) \geq 0$. It follows that $p'_a - p_a \geq 0$ for all $a \in A^*$, therefore $p'(A^*) \geq p(A^*)$. Now consider $p(A^*) \in (0,1)$. We know that $p'_{a'} - p_{a'} \geq 0$ for all $a' \in A^*$. Now, let V be the set of individuals who play an action in $A \setminus A^*$, therefore $L^{a'}_{a,a'}(c,d) - L^a_{a',a}(d,c) > 0$ for some c, $d \in W$ such that $s(c) = a \in A \setminus A^*$ and $s(d) = a' \in A^*$. Then, by Lemma 0*, for some $a' \in A^*$, we have $p'_{a'} - p_{a'} > 0$.

Necessity of (i):

This proof is identical to the proof of necessity of (i) in Lemma 1.

Necessity of (ii):

Suppose that $F_{a'}$ sfosd F_a, but $L^{a'}_{aa'}(c,d) - L^a_{a'a}(d,c) < 0$, for some a, $a' \in A$. Suppose that $p_a + p_{a'} = 1$, $A^* = \{a'\}$, and $s(d) = a'$ and $s(e) = a$ for all $e \in W \setminus \{d\}$. Since $L(c)$ is imitative for all $c \in W$, Lemma 0* yields

$$p'_{a'} - p_{a'} = \frac{1}{|W|} \sum_{e \in W \setminus \{d\}} \Pr(e \curvearrowright d)(L^{a'}_{a,a'}(e,d) - L^a_{a',a}(d,e)).$$

Then, if the sampling process is such that $\Pr(e \curvearrowright d)$ is small enough for all $e \in W \setminus \{c,d\}$ then $L^{a'}_{aa'}(c,d) - L^a_{a'a}(d,c) < 0$ implies that $p'_{a'} - p_{a'} < 0$ and therefore $p'(A^*) < p(A^*)$. If $F_{a'} = F_a$, as in the proof of Lemma 1, the environment F can be modified and turned into an environment $\widetilde{F}$ such that $\widetilde{F}_{a'}$ sfosd $\widetilde{F}_a$, but still $\widetilde{L^{a'}_{aa'}}(c,d) - \widetilde{L^a_{a'a}}(d,c) < 0$, so the analysis above applies in the same way.

Necessity of (iii):

Suppose that there exists a, $a' \in A$ with $F_{a'}$ sfosd F_a and $A^* = \{a'\}$. Suppose further that there exists a non-empty set $V \subset W$, such that $V \neq W$, $L^{a'}_{aa'}(c,d) - L^a_{a'a}(d,c) = 0$ for all $c \in V$ and $d \in W \setminus V$. Suppose also that $s(c) = a$ and $s(d) = a'$ for all $c \in V$ and $d \in W \setminus V$. Lemma 0* leads to the result. □

The next result provides a characterization of FOM behavioral processes in terms of the shape of $L(c)(a,x,d,a',y)_{a''}$, for all a, a', $a'' \in A$. In the analysis below, we will use the concept of *net-switching function* from action a to action a' for each c, $d \in W$. We denote this function by $g_{aa'}(c,x,d,y)$ and it is defined as $g_{aa'}(c,x,d,y) := L(c)(a,x,d,a',y)_{a'} - L(d)(a',y,c,a,x)_a$. This function measures how much probability is being shifted from action a to action a' when c and d play a and a' respectively and observe each other.

Proposition 1*. *A behavioral process L is FOM if and only if it satisfies the following conditions:*

(i) *$L(c)$ is imitative for all $c \in W$.*

(ii) *For all a, $a' \in A$, the net-switching functions $g_{aa'}(c,x,d,y)$ satisfy*

- *(ii.1)* $g_{aa'}(c,x,d,y) = -g_{aa'}(c,y,d,x)$, $\forall\ x,\ y \in [0,1]$, *for all $c \in W$ and $d \in W \setminus \{c\}$.*
- *(ii.2)* $g_{aa'}(c,x,d,y)$ *is non-decreasing in y, for all $x \in [0,1]$ for all $c \in W$ and $d \in W \setminus \{c\}$.*
- *(ii.3)* $g_{aa'}(c,x,d,y) > 0$, $\forall\ x,\ y \in [0,1]$ *such that $y > x$ for some $c \in V$ and $d \in W \setminus V$ for all $V \subset W$ such that V is not empty and $V \neq W$.*

Proof. The proof of necessity is identical to the proof of Proposition 1, except for the fact that for Condition (ii.2) the strict inequality $0 < L^{a'}_{aa'} - L^{a}_{a'a}$ turns into the weak inequality $0 \leq L^{a'}_{aa'}(c,d) - L^{a}_{a'a}(d,c)$. Sufficiency for Condition (iii) of Lemma 1* follows directly from Lemma 3 and sufficiency for Condition (ii) of Lemma 1* follows from the same argument used to prove that $E(g_1(X,Y)) \geq 0$ in Step 2 of the proof of Lemma 3. □

Remark 4. The behavioral process associated to a heterogenous population of individuals who use FOM behavioral rules is not necessarily FOM.

For example, it is easy to check that the behavioral process of a population where some individuals use Schlag's [15] Proportional Imitation rule and some others use the rule IIB is not FOM, because Condition (ii.1) in Proposition 1* is not satisfied.

C Individual monotonicity

This appendix provides the formal analysis of individually monotone (IM) rules that we discussed in Section 6.

Definition 5. An imitative rule is said to be individually monotone if $F_{a'}$ sfosd $F_a \Rightarrow$

(i) $L^{a'}_{aa'} > L^{a}_{aa'}$

(ii) $L^{a'}_{a'a} > L^{a}_{a'a}$.[15]

We can characterize individually monotone rules using the mathematical structures we discovered in the analysis of net-switching functions for FOM rules. Specifically, we find that the probability of switching, $L(a, x, a', y)_{a'}$, must be increasing in the payoff of the sampled action and symmetric with respect to payoffs, in the sense that, if payoffs of the actions are interchanged, the probabilities of playing them are also interchanged.

Proposition 5. *An imitative rule L is individually monotone if and only if, for all a, $a' \in A$, $L(a, x, a', y)_{a'}$ satisfies*

(i) $L(a, x, a', y)_{a'} = L(a, y, a', x)_a$, $\forall\ x,\ y \in [0, 1]$.

(ii) $L(a, x, a', y)_{a'}$ *is non-decreasing in* y, *for all* $x \in [0, 1]$.

(iii) $L(a, x, a', y)_{a'} > 1/2$ *for all* $x,\ y \in [0, 1]$ *such that* $y > x$.

Proof. Suppose $F_{a'}$ fosd F_a. We need to verify

$$\int\int L(a, x, a', y)_{a'} dF_a(x) dF_{a'}(y) > \int\int L(a, x, a', y)_a dF_a(x) dF_{a'}(y).$$

But, since L is imitative, this is equivalent to

$$\int\int L(a, x, a', y)_{a'} dF_a(x) dF_{a'}(y) > 1/2.$$

Let $h_{aa'}(x, y) := 2L(a, x, a', y)_{a'} - 1$. Then, the last expression is equivalent to

$$\int\int h_{aa'}(x, y) dF_a(x) dF_{a'}(y) > 0.$$

[15]It might seem more attractive to find rules L such that, when $F_{a'}$ sfosd F_a, the lower bound for $L^{a'}_{aa'}$ is greater than 1/2 for any environment. However, an argument similar to the one that leads to the first statement in Remark 1 shows that such rules cannot exist.

The arguments in the proof of Proposition 1 can be used to prove that a necessary and sufficient condition for the last statement to be true is that $h_{aa'}(x, y)$ satisfies

(i) $h_{aa'}(x, y) = -h_{aa'}(y, x), \forall xy \in [0, 1]$.

(ii) $h_{aa'}(x, y)$ is non-decreasing in y, for all $x \in [0, 1]$.

(iii) $h_{aa'}(x, y) > 0$ for all x, $y \in [0, 1]$ such that $y > x$.

These conditions are equivalent to Conditions (i)–(iii) in the hypothesis. □

For example, the rule IIB discussed in Section 4 is IM if and only if $L(a, x, a', x) = 1/2$ for all a, $a' \in A$ and for all $x \in [0, 1]$. The Proportional Imitation rule is not IM, but the rule given by

$$\begin{aligned} L(a, x, a', y)_a &= \frac{1}{2}(1 + x - y) \\ L(a, x, a', y)_{a'} &= \frac{1}{2}(1 + y - x) \end{aligned}$$

is both improving and IM.

If every member of the population uses an IM rule, which may be different across individuals, the resulting behavioral process is FOM. This result follows from Propositions 5 and 1* of Appendix B.

Proposition 6. *If every individual in the population uses an IM rule, then the corresponding behavioral process is FOM.*

D Second-order monotonicity

This appendix provides the analysis of SOM rules as discussed in Section 6. These rules may be interpreted as capturing risk averse attitudes of the individuals. In order to isolate risk attitudes, in this appendix we consider properties of rules in environments where the expected payoff of the distributions of different actions is the same. In what follows, second-order stochastic dominance, abbreviated by F_a sosd $F_{a'}$, means that $\int_0^x (F_{a'}(t) - F_a(t))dt \geq 0$ for all $x \in [0, 1]$. We start by defining the set of actions that would be preferred for risk-averse decision makers, i.e., the set of actions whose payoff distribution second-order stochastically dominates the distributions of the other actions in the set. Let $A_* := \{a \in A : F_a \text{ sosd } F_{a'} \ \forall \ a' \in A\}$.

Definition 6. A rule L is said to be SOM if $p'(A_*) \geq p(A_*)$ in every environment where all the actions have the same expected payoff.

Since the structure of the definition of SOM rules is analogous to that of FOM rules, they can be characterized in an analogous way. This characterization is provided in Lemma 4 and Proposition 7.[16]

Lemma 4. *A rule L is SOM if and only if it satisfies the following conditions:*

(i) L *is imitative.*

(ii) $F_{a'}$ *sosd* $F_a \Rightarrow L^{a'}_{aa'} - L^{a}_{a'a} \geq 0$, $\forall$ a, $a' \in A$, *in every environment.*

Proof. Sufficiency is proved analogously to the proof of sufficiency in Lemma 1.

Necessity of (i):

Consider the SOM rule L. Let x, $y \in [0,1]$; a, a', $a'' \in A$, $a'' \notin \{a, a'\}$. Suppose $L(a, x, a', y)_{a''} > 0$. Consider an environment where $F_a = F_{a'}$, $\mu_a(x) = \mu_a(y) = \mu_a(1/2) = 1/3$, and for all $a''' \in A \setminus \{a, a'\}$, $\mu_{a'''}(x) = \mu_{a'''}(y) = 1/3$ and $\mu_{a'''}(1/4) = \mu_{a'''}(3/4) = 1/6$. It follows that $A_* = \{a, a'\}$. Assume that $p(A_*) = 1$. Suppose that there are c, $d \in W$ such that $s(c) = a$, and $s(d) = a'$, then $p'(A^*) < 1$. Thus, L has to be imitative.

Necessity of (ii):

Suppose that for some a, $a' \in A$, $F_{a'}$ sosd F_a, but $L^{a'}_{aa'} - L^{a}_{a'a} < 0$. Suppose $a' \in A^*$. Consider $\varepsilon \in (0,1)$. If $\mu_a(1) = 1$ for all $a \in A$, then consider the modified environment $\widehat{F}$ such that $\widehat{\mu}_a(1) = 1 - \varepsilon$ and $\widehat{\mu}_a(y) = \varepsilon$ for some $y \in [0,1)$, for all $a \in A$. If $\mu_a(0) = 1$ for all $a \in A$, then consider the modified environment $\widehat{F}$ such that $\widehat{\mu}_a(0) = 1 - \varepsilon$ and $\widehat{\mu}_a(y) = \varepsilon$ for some $y \in (0,1]$, for all $a \in A$. Otherwise just let $\widehat{F} = F$. Denote by π the expected value of the distributions in the environment $\widehat{F}$. Now consider the new modified environment $\widetilde{F}$ such that for any interval $I \subseteq [0,1] \setminus \{\pi\}$, $\widetilde{\mu}_{a'}(I) = (1-\varepsilon)\widehat{\mu}_{a'}(I)$, $\widetilde{\mu}_{a'}(\pi) = \widehat{\mu}_{a'}(\pi) + \varepsilon\widehat{\mu}_{a'}([0,1] \setminus \{\pi\})$; for any interval $I \subseteq (0,1)$, $\widetilde{\mu}_{a'''}(I) = (1-\varepsilon)\widehat{\mu}_{a'''}(I)$, $\widetilde{\mu}_{a'''}(1) = (1-\varepsilon)\widehat{\mu}_{a'''}(1) + \varepsilon\pi$, and $\widetilde{\mu}_{a'''}(0) = (1-\varepsilon)\widehat{\mu}_{a'''}(0) + \varepsilon(1-\pi)$ for all $a''' \in A \setminus \{a'\}$. In this new environment $\widetilde{A}_* = \{a'\}$. Now, conclude as in Lemma 1. □

[16] As in EUT, the analysis of SOM rules is easier to carry on when properties are stated in terms of weak inequalities.

¿From Lemma 4, it is clear that SOM rules are impartial.

Proposition 7. *L is SOM if and only if it satisfies the following conditions:*

(i) L is imitative.

(ii) For all $a, a' \in A$, the net-switching function $g_{aa'}(x, y)$ satisfies

(ii.1) $g_{aa'}(x, y) = -g_{aa'}(y, x)$, $\forall\ x,\ y \in [0, 1]$.

(ii.2) $g_{aa'}(x, y)$ is concave in y, for all $x \in [0, 1]$.

Proof. Sufficiency and the necessity of the Condition (ii.1) can be proven in the same way as in Proposition 1. Now we prove that $g_{aa'}(x, y)$ is concave with respect to y for SOM rules. Suppose that for some x, y', y, $\lambda \in [0, 1]$, and $y'' := \lambda y + (1 - \lambda)y'$ we have that

$$g_{aa'}(x, y'') < \lambda g_{aa'}(x, y) + (1 - \lambda)g_{aa'}(x, y').$$

Consider an environment where $\mu_a(x) = \mu_{a'}(x) = 1 - \varepsilon$, $\mu_a(y) = \lambda\varepsilon$, $\mu_a(y') = \varepsilon(1 - \lambda)$, $\mu_{a'}(y'') = \varepsilon$, and notice that $F_{a'}$ sosd F_a. Lemma 4 and Condition (ii.1) imply

$$\begin{array}{rcl} 0 & \leq & L^{a'}_{aa'} - L^{a}_{a'a} \\ & = & (1-\varepsilon)^2 g_{aa'}(x, x) + (1-\varepsilon)\varepsilon g_{aa'}(x, y'') + (1-\varepsilon)\lambda\varepsilon g_{aa'}(y, x) \\ & & +(1-\varepsilon)\varepsilon(1-\lambda)g_{aa'}(y', x) + \lambda\varepsilon^2 g_{aa'}(y, y'') + (1-\lambda)\varepsilon^2 g_{aa'}(y', y'') \\ & = & \varepsilon[\lambda\varepsilon g_{aa'}(y, y'') + (1-\lambda)\varepsilon g_{aa'}(y', y'') \\ & & +(1-\varepsilon)\{g_{aa'}(x, y'') - \lambda g_{aa'}(x, y) - (1-\lambda)g_{aa'}(x, y')\}]. \end{array}$$

By our hypothesis, the term inside the brackets {.} is negative. Thus, for small enough ε, the term inside the brackets [.] is negative, causing a contradiction. □

The following corollary completes the analogy between the characterizations of FOM and SOM rules and it follows directly from Condition (ii) of Proposition 7.

Corollary 2. *If L is SOM then $g_{aa'}(x, y)$ is convex in x, for all $y \in [0, 1]$.*

References

[1] Apesteguia, J., Huck, S., Oechssler, J.: Imitation –theory and experimental evidence. J. Econ. Theory 136, 217-235 (2007)

[2] Bala, V., Goyal, S.: Conformism and diversity under social learning. Econ. Theory 17, 101-120 (2001)

[3] Björnerstedt, J., Schlag, K.: On the evolution of imitative behavior. Discussion Paper B-378, University of Bonn (1996)

[4] Börgers, T.: On the relevance of learning and evolution to economic theory. Econ. J. 106, 1375-1385 (1996)

[5] Börgers, T., Morales A., Sarin, R.: Expedient and monotone learning rules. Econometrica 72, 383-405 (2004)

[6] Cubitt, R., Sugden, R.: The selection of preferences through imitation. Rev. Econ. Stud. 65, 761-771 (1998)

[7] Ellison, G., Fudenberg, D.: Rules of thumb for social learning. J. Polit. Economy 101, 612-643 (1993)

[8] Ellison, G., Fudenberg, D.: Word of mouth communication and social learning. Quart. J. Econ. 110, 93-125 (1995)

[9] Hanoch, G., Levy, H.: The efficiency analysis of choices involving risk. Rev. Econ. Stud. 36, 335-346 (1969)

[10] Morales, A.: Absolutely expedient imitative behavior. Int. J. Game Theory 31, 475-492 (2002)

[11] Offerman, T., Schotter, A.: Sampling for information or sampling for imitation: an experiment on social learning with information on ranks. New York University (2005) (manuscript)

[12] Oyarzun, C., Sarin, R.: Learning and risk aversion. Texas A&M University (2007) (manuscript)

[13] Robson, A.: A biological basis for expected and non-expected utility preferences. J. Econ. Theory 68, 397-424 (1996)

[14] Rothschild, M.: A two-armed bandit theory of market pricing. J. Econ. Theory 9, 185-202 (1974)

[15] Schlag, K.: Why imitate, and if so, how? A bounded rational approach to multi-armed bandits. J. Econ. Theory 78, 130-156 (1998)

[16] Schlag, K.: Which one should I imitate. J. Math. Econ. 31, 493-522 (1999)

[17] Stahl, D.: Boundedly rational rule learning in a guessing game. Game. Econ. Behav. 16, 303-330 (1996)

[18] Stahl, D.: Population rule learning in symmetric normal-form games: theory and evidence. J. Econ. Behav. Organ. 45, 19-35 (2001)

[19] To, T.: Risk and evolution. Econ. Theory 13, 329-343 (1999)

[20] Vega-Redondo, F.: The evolution of Walrasian behavior. Econometrica 65, 375-384 (1997)