

Evolutionary Foundations of Rational Choice

Olivier Gossner* and Christoph Kuzmics†

January 25, 2009

Abstract

We study the potential evolutionary appeal of rationality in a model in which different populations differ with respect to their experimentation over new rules of behavior. We find that experimentation over the set of (strictly) rational rules dominates any other form of experimentation. This evolutionary advantage of strict rationality, furthermore, is substantial when learning takes place over a limited amount of time, or when the environment is stochastically changing.

Keywords: rationality, evolution, weak axiom of revealed preferences, strict preference, adaptation

Journal of Economic Literature classification numbers: C73, D01, D11

1 Introduction

The aim of this paper is to study the potential evolutionary appeal of rationality. We are particularly interested in actual rational behavior as opposed to “as if” rational behavior.

*Paris School of Economics, Paris and London School of Economics, London, gossner@ens.fr

†Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, c-kuzmics@kellogg.northwestern.edu

In order to do so we distinguish between behavior that is rational and behavior that is what we call **adapted** to the environment it interacts with. Rationality is in fact a property of individual behavior which is independent of the environment, and the degree to which a particular behavior is adapted to a given environment is a priori a distinct question as to whether or not this behavior is rational. The framework we use to make this distinction is the context of choice from sets of alternatives.

As in consumer choice theory, we describe an agent's behavior by a choice rule (in fact a correspondence). This rule tells us what the agent chooses when presented with any subset of the set of all possible alternatives. An agent is **rational** if her behavior is consistent with a preference relation over alternatives that is a weak order, i.e. satisfies completeness and transitivity. When choice sets contain all possible finite sets, which is the framework of this paper, Arrow (1959)'s well known result shows that an agent is rational in the above sense if and only if her choice rule exhibits the **weak axiom of revealed preferences** of Samuelson (1938). A preference is **strict** if it exhibits no ties, and a corresponding choice rule is then **strictly rational**.

In our model there are multiple populations with a fixed and constant number of individuals. Each individual lives only for one period, and has offsprings who live in the period after that. Individuals face various decision problems multiple times over their lifetime according to some pre-specified distribution. A choice rule characterizes choices made by the agent. The environment describes the distribution over choice sets available to agents, and assigns a fitness to each alternative. Each combination of a choice rule and an environment induces an average fitness over all decision problems.

We model the evolution of individuals' rules through experimentation and selection. At the beginning of each period of time, all individuals of one population use the rule that gave the best average fitness in the last period. One of the individuals in this population experiments by randomly drawing a rule according to this population's fixed distribution. At the end of this

period all individuals use the rule of these two that provides the best fitness given the environment in this period. Thus, all populations have the same fast selection process. Populations differ (only) by the distributions according to which new rules are experimented, called their **sampling distributions**.

We have in mind an environment which may evolve over time, and does not necessarily leave enough time for the learning process to converge. In particular, and in contrast to most evolutionary models, our primary interest is not in the asymptotic properties of the process over rules used by individuals in a fixed environment when the number of generations becomes large. We rather focus our attention to the properties of the processes after a finite (thought as being small) number of generations, both when the environment is fixed across time, or when the environment is itself stochastic.

Looking first at the case of a fixed environment, our first set of results establishes a dominance of strictly rational rules over other choice rules. For every dynasty characterized by some sampling distribution, our Rational Dominance Theorem (Theorem 1) shows the existence of another dynasty which samples only from the set of all choice rules which are **strictly rational** with the property that the expected fitness of the latter dynasty is no less than the expected fitness of the former one. Furthermore, the Universal Rational Dominance Theorem (Theorem 2) shows that a dynasty that experiments uniformly over the strictly rational rules obtains a fitness no less than the one obtained by any other dynasty using any symmetric sampling distribution over the set of rules. We stress the fact that these comparisons hold after any number of generations, be it small or large.

The superiority of strictly rational rules is strict under mild conditions on the environment. In particular, and, perhaps surprisingly, strictly rational rules perform better than rational rules even if several alternatives provide the same level of fitness.

Our second set of results quantifies the adaptation level of different dynasties after a fixed number of generations and the fitness-wedge between

them. The most striking result is achieved when comparing a **non-rational** dynasty which experiments uniformly from the set of all choice rules and a **strictly rational** dynasty which experiments uniformly from the set of strictly rational rules. We prove that there is a significant payoff-wedge in favor of the strictly rational dynasty. This advantage appears even after just one generation and persists to at least a number of generations which is a double exponential in the total number of alternatives. Even with a rather moderate number of alternatives in mind, this number of generations is a very large number.

Keeping the environment fixed, all dynasties (as long as the the sampling distribution has sufficient support) do equally well in the ultra-long run as they all eventually use a fitness-maximizing choice rule. This means that in the ultra-long run every individual in these dynasties will behave “as if” she is rational and “as if” she knows the exact fitness-function, i.e. will be perfectly adapted. Now, introducing the possibility that environments might change, no matter whether this happens very frequently or very rarely, will make this payoff-wedge between the strictly rational dynasty and the non-rational dynasty persist forever. This is the message of our Universal Rational Dominance Theorem for changing environments, Theorem 4.

Our study thus shows that a population endowed with a ‘gene of rationality’ that forces its experimentation to the subset of strictly rational rules has an evolutionary advantage over any other population. In a set-up that allows to quantify this advantage, we see that it is substantial for a wide range of environments. In the case of changing environments, even if there is a cost associated to rationality, this cost might well be counterbalanced by the persistent payoff-wedge if environments change with some frequency. In this respect our paper offers an evolutionary foundation of rational choice, or rather, more precisely, of strictly rational choice.

The paper is organized as follows. Section 2 introduces the model. Section 3 discusses the distribution of fitness of randomly selected rules. For fixed

environments, section 4 establishes the superiority of rational rules, Section 5 studies the expected time to reach perfect adaptation, and Section 6 the time to reach partial adaptation. Section 7 addresses changing environments. We discuss related literature in Section 8 and finally, conclude in Section 9.

2 Model

2.1 Choice

Let $K = \{1, \dots, K\}$, $K > 1$ be the set of all possible **alternatives**, subsets of which a decision maker may be presented with at one time or another. Let $\mathcal{L} = \mathcal{P}(K) \setminus \{\emptyset\}$ denote the set of all non-empty subsets of K . We call an element in \mathcal{L} a **choice set**, as we think of these sets as the possible sets of choices an individual might at one point or another be faced with and be asked to make a choice from.

Definition 1 *A choice rule is a function $R : \mathcal{L} \rightarrow \mathcal{L}$ such that $R(L) \subseteq L$ for all $L \in \mathcal{L}$. Let \mathcal{R} denote the set of all such choice rules.*

Following Uzawa (1956) and Arrow (1959) (see also Chapter 1.B in Mas-Collel, Whinston, and Green (1995)), let \succeq denote a binary (preference) relation over elements in K with the interpretation that when $i \succeq j$ an agent holding this preference relation weakly prefers i over j . The relation \succeq is complete if for any two $i, j \in K$, $i \succeq j$ or $j \succeq i$ (or both), it is transitive if $i \succeq j$ and $j \succeq l$ imply $i \succeq l$. A complete and transitive relation is called **rational** (see e.g. Definition 1.B.1 in Mas-Collel, Whinston, and Green (1995)). In this paper a special case of rational preferences plays a prominent role, namely, strict preferences. A relation \succ is irreflexive if, for all $i \in K$, we do not have $i \succ i$. We call a preference relation **strictly rational** if it satisfies completeness, transitivity, and irreflexivity.

These definitions extend from preference relations to the corresponding agent's behavior.

Definition 2 A choice rule $R \in \mathcal{R}$ is **rational** if there exists a complete and transitive preference relation \succeq such that, for every L , $R(L)$ is the set of maximal elements in L for \succeq . Let \mathcal{R}^r denote the set of rational rules. It is **strictly rational** if it is rational and $R(L)$ is a singleton for all $L \in \mathcal{L}$. Let \mathcal{R}^s denote the set of all strictly rational rules.

As an alternative definition, a strictly rational rule is one based on a strictly rational preference relation.

It is interesting to know how the rationality of a choice rule expresses itself in terms of the choices made by the agent, or, in other words, what property a choice rule has to have such as to be derivable from a rational preference. The answer to this question was given by Arrow (1959) (see also Mas-Collel, Whinston, and Green (1995)[Proposition 1.D.2]). Namely, the choice rule R , has to satisfy the weak axiom of revealed preferences, first stated by Samuelson (1938), see e.g. [Definition 1.C.1 in Mas-Collel, Whinston, and Green (1995)].¹

Our paper provides an evolutionary or learning model in which agents experiment rules from the set \mathcal{R} of choice rules. Nature first fixes the environment, which consists of two ingredients, the frequency distribution with which individuals face each $L \in \mathcal{L}$ and the function that determines how much **fitness** or **material payoff** any particular choice provides (this is discussed in detail in section 2.3). For a given environment (or a stochastic process over environments), we study the evolution of a population characterized by a distribution, over the set of all rules, it samples from. We allow

¹This is true since we assume the domain of R is \mathcal{L} , and remains true, as shown by Arrow (1959), as long as this domain includes all finite subsets of K . Arrow (1959) also shows that, in this case, R satisfies the weak axiom of revealed preferences if and only if it satisfies Ville (1946)'s and Houthakker (1950)'s strong axiom of revealed preferences. Houthakker (1950) demonstrates that this strong axiom of revealed preferences is sufficient for a choice (or demand) function to be rationalizable by a rational preference relation for the case where the domain of R is the set of all budgets (see e.g. Mas-Collel, Whinston, and Green (1995)[Chapter 3.J]), a case which does not meet the requirement of Arrow's (1959) result.

this sampling probability to be any distribution q over the set of all rules \mathcal{R} . Given the assumptions that individuals do not a priori know the environment nature chooses and that nature works on individuals independently of its choice of environment, we find those distributions q over \mathcal{R} of particular interest, in which no element of K plays a special role. These distributions are characterized by a symmetry property.

Let Π be the set of all permutations over K . For $\pi \in \Pi$ and $L \in \mathcal{L}$ we let $\pi(L) = \{j \in K \mid \exists i \in L : j = \pi(i)\}$ be the set-wise extension of π . For $\pi \in \Pi$ and $R \in \mathcal{R}$ let $R^\pi \in \mathcal{R}$ be such that $R^\pi(L) = \pi^{-1}(R(\pi(L)))$ for all $L \in \mathcal{L}$.

Definition 3 *A distribution q over \mathcal{R} is **symmetric** if for every R such that $q(R) > 0$ and every $\pi \in \Pi$, $q(R^\pi) = q(R)$.*

For symmetric distributions of choice rules, the names of alternatives plays no role. I.e. if one permutes alternatives, the distribution over choice rules remains the same. One example of a non-symmetric distribution over decision rules is one that puts probability one on the decision rule induced by some strict ranking of K . One example of a symmetric distribution over decision rules is the uniform distribution over the set of all rules. Other symmetric distributions of special interest are the uniform distribution over the set of all rational rules, and the uniform distribution over the set of all strictly rational rules.

There are, however, many other symmetric distributions over decision rules. Consider for instance the set of rules which are all such that they designate a best option, which can be any one element of K , but otherwise does not impose any more restrictions. The best option is chosen whenever it is available. Let

$$\mathcal{R}^o = \{R \in \mathcal{R} \mid \exists j \in K : R(L) = \{j\} \forall L \text{ with } j \in L\}.$$

denote this set, which we refer to as the set of **optimistic** rules. Another set of rules of interest is what we call the set of **pessimistic** rules, given by

$$\mathcal{R}^p = \{R \in \mathcal{R} \mid \exists j \in K : j \notin R(L) \forall L \text{ except } L = \{j\}\}.$$

These rules are such that they dedicate a particular choice, which is to be avoided at all cost, without imposing any restrictions other than that.

The uniform distributions both on the set of optimistic and pessimistic rules are symmetric. If there are at least three alternatives ($K \geq 3$), the sets of optimistic and pessimistic rules differ, and are neither the set of all rules nor the set of rational or strictly rational rules. In fact, for $K \geq 3$ these sets are proper subsets of the set of all rules, while the set of strictly rational rules is properly contained in them.

The only symmetric distribution over choice rules which has a singleton as support is the one which puts probability one on the rule R^0 with the property that $R^0(L) = L$ for all $L \in \mathcal{L}$. This is the rational rule for an agent who is completely indifferent between all choices, and we refer to R^0 as the **zero rule**. We discuss more symmetric distributions over rules in section 5.

2.2 The environment

Recall that the set of alternatives $K = \{1, \dots, K\}$ is fixed, as is the set of all choice sets $\mathcal{L} = \mathcal{P}(K) \setminus \{\emptyset\}$, the set of all non-empty subsets of K . Nature chooses the environment, which consists of two components.

First $p: \mathcal{L} \rightarrow \mathbb{R}$ denotes a probability mass function over all such choice sets. It describes the frequency with which choice sets are accessible to agents. In some cases, it is useful to consider **neutral** distributions, for which all alternatives play the same role.

Definition 4 *A distribution p over choice sets is **neutral** if, for every permutation π of K , and every choice set $L \subseteq K$, $p(L) = p(\pi(L))$.*

Obviously, the uniform distribution is neutral. Other examples of neutral distributions over choice sets are the uniform distributions over choice sets of fixed size l , for $1 \leq l \leq K$. Other results call for another (mild) assumption on p .

Definition 5 *A distribution p over choice sets has **full support** if it puts positive probability on every choice set that contains at least two elements, i.e. $p(L) > 0$ for all $L \in \mathcal{L}$ with $|L| \geq 2$.*

Second $u: K \rightarrow \mathbb{R}_+$ is a function from the set of all possible choices to non-negative real numbers with the interpretation that $u(i)$ is the **fitness**, or **material payoff**, an individual receives when choosing $i \in K$. We extend any fitness function to the set \mathcal{L} of choice sets by setting

$$u(L) = \frac{1}{|L|} \sum_{k \in L} u(k)$$

for $L \in \mathcal{L}$, with the natural interpretation that $u(L)$ is the expected fitness for the agent when L is the set of accepted alternatives, using the fact that each element in L is then eventually chosen by the agent with equal probability. The pair $e = (u, p)$ is called an **environment**. Some results will call for a (mild) assumption on the fitness function u .

Definition 6 *A fitness function u is **discriminatory** if it is injective, i.e. $u(i) \neq u(j)$ for all $i \neq j$, i.e. every choice in K provides a distinct level of fitness.*

In order to define the average fitness of a rule R in the environment $e = (u, p)$, we need to first specify the choices realized by the agent when facing the choice set L . If $R(L)$ is a singleton, the agent ends up getting the alternative $R(L)$, and obtains a fitness (for this period) of $u(R(L))$. If $R(L)$ is not a singleton we assume that the agent accepts all alternatives in $R(L)$

equally, and ends up with each of them with equal probabilities.² Given our set extension of u , the average fitness received by the agent is also, in this case, $u(R(L))$.

In the environment $e = (u, p)$, the (average) fitness or material payoff of any rule $R \in \mathcal{R}$ is then given by

$$U_e(R) = \mathbb{E}_p [u(R(L))] = \sum_{L \in \mathcal{L}} p(L) u(R(L)).$$

We study evolutionary processes under two assumptions on the environments. One in which there is a unique fixed environment which never changes, which is discussed in sections 4, 5, and 6, and one in which the environment follows a stochastic process, studied in section 7. When dealing with a fixed environment e we suppress the dependence on the environment e in $U_e(R)$ and simply write $U(R)$.

2.3 The evolutionary or learning process

In a given environment (or stochastic process over environments), a population evolves according to a process of experimentation and selection.

At each generation, the population experiments a random rule selected according to a distribution q on the set of rules. This distribution, which is fixed across time and is a characteristic of the population, is called the population's sampling distribution. The selection process is fast: at the end of each generation, the rule providing the best fitness in the current environment is selected, and is adopted by all agents in the population.

The evolutionary model thus has the two central features any model of evolution should have, mutation (or experimentation) and selection (or learning). Here mutations are fixed at a rate of one per generation and population

²We believe this to be an innocuous assumption, which, however, provides us with the property that the set of all decision rules is finite. We do not believe that any additional insight can be gained by relaxing this assumption.

and selection is fast. The model allows us to simply analyze each population by following its representative individual, one for each period t . We thus speak of the different populations as different dynasties of these representative individuals.

Our model thus makes a distinction between the rule prevailing in the population at some generation t from the type of sampling distribution used by this population. In particular, a population may, at some moment in time, adopt a rational rule although its sampling distribution q does not put weight on rational distributions only. In such situations, the rationality of the population at time t is the result of a learning process among a larger set of rules, and is only temporary in case an offspring samples a non-rational but superior rule. On the other hand, if a population's sampling distribution puts weight on rational rules only, agents in this population use rational rules at every point of the evolutionary process.

Each sampling distribution q defines a different dynasty, with the understanding that at time 0 the representative agent uses a rule randomly chosen according to probability distribution q , while all descendants do the same when experimenting. We focus on the comparison of evolutionary performances of different dynasties. In particular, we refer to the dynasty sampling according to the uniform distribution q^r over the set \mathcal{R}^r as the **rational dynasty**, and to the dynasty sampling according to the uniform distribution q^s over \mathcal{R}^s as the **strictly rational dynasty**, and to the dynasty sampling according to the uniform distribution q^u over the whole set of rules \mathcal{R} as the **non-rational dynasty**.

3 On the fitness distribution of sampled rules

In this section we study and compare the distributions of fitness of rules selected according to different sampling distributions. The results obtained here are the building blocks of the Rational Dominance Theorems of Sections

4 and 7. The reader mostly interested in the comparison of dynasties in the dynamic set-up may jump to section 4. The main results in sections 4 and 7 can be read without reading this section. Their proofs, however, are based on the results presented here.

The ingredients of this section are a fixed environment $e = (u, p)$ and various sampling distributions q as set down in section 2.2.3.

We study the distribution of fitness $\tilde{U}^q = U(R) = U_e(R)$ induced by choosing rule R randomly according to q . This is the random fitness any member of a dynasty sampling according to q encounters from her random experimentation. We call such \tilde{U}^q , with distribution depending on q , the **sampling fitness** of the corresponding dynasty.

The results in this section are as follows. First, the expected sampling fitness is the same for all symmetric sampling distributions. Therefore, in terms of expected fitness, and in the absence of learning, all dynasties that experiment according to symmetric distributions perform equally well.

Second, for a dynasty that experiments according to any q , symmetric or not, there is another dynasty q' with support contained in the set of strictly rational rules such that the distribution of the sampling fitness induced by q' is a mean-preserving spread of the sampling fitness induced by q . Moreover, and importantly, the sampling distribution q' , whose fitness is a mean-preserving spread of that induced by q , is the same for all environments (with same p), i.e. it does not depend on the fitness function u . This is saying that we can replace any rule with a convex combination of strictly rational rules such that the expected fitness of the latter coincides with the fitness of the former for every fitness-function u .

Third, if p is neutral, there exists a unique sampling distribution q' , whose fitness is a mean-preserving spread of that induced by any symmetric distribution q , and this q' is the sampling distribution of the strictly rational dynasty, i.e. it is q^s , the uniform distribution over all strictly rational rules.

Finally, we provide conditions on p and u under which the mean preserv-

ing spreads thus constructed are “strict” in the sense that the two distributions of fitness differ.

All results in this section are proven in Appendix C.

Lemma 1 *Let the environment $e = (u, p)$ be arbitrary. The expected value of the sampling fitness \tilde{U}^q according to any symmetric sampling distribution q is independent of the sampling distribution considered and is given by*

$$\mathbb{E}_q \tilde{U}^q = \mathbb{E}_p u(L).$$

Although the expected sampling fitness associated to any two symmetric distributions over choice rules coincide, this is typically not at all true of their variances. A key to a better understanding of any agent’s sampling fitness is to look at the probability distribution over the agent’s choices induced by the distribution over alternatives, p , and some choice rule R . Given p over choice sets \mathcal{L} , and R , let $\lambda_p(R)(k)$ denote the overall probability with which an element $k \in K$ is selected under the rule R , it is given by

$$\lambda_p(R)(k) = \sum_{L:k \in R(L)} \frac{p(L)}{|R(L)|}.$$

We call $\lambda_p(R)$ the **choice distribution** associated to R . For any fitness function u (and given p), the average fitness of rule R can be expressed as $U(R) = \sum_k \lambda_p(R)(k)u(k)$, so that a rule’s average fitness is entirely determined by its choice distribution.

For instance, consider the case in which p is uniform on \mathcal{L} , and let R be the strictly rational rule that is induced by the preference relation which selects the least element available (1 is strictly preferred to 2, which is strictly preferred to 3, etc.). There are in total $2^K - 1$ choice sets (all subsets of K except the empty set are choice sets), and 2^{K-1} of them contain the preferred choice (1). This preferred choice is chosen by the agent in all choice sets that contain it, so that $\lambda_p(R)(1) = \frac{2^{K-1}}{2^K - 1}$. The second preferred choice is

selected whenever the first preferred choice is unavailable and the second preferred is available. This is the case for 2^{K-2} choice sets, and consequently $\lambda_p(R)(2) = \frac{2^{K-2}}{2^K-1}$. More generally, $\lambda_p(R)(l) = \frac{2^{K-l}}{2^K-1}$ for every l . Strictly rational rules span all permutations of this probability vector.

For a given distribution p over choice sets, let Λ_p denote the set of all choice distributions, i.e. $\Lambda_p = \{\lambda_p(R), R \in \mathcal{R}\}$. Similarly, denote by Λ_p^s the set of choice distributions induced by strictly rational rules, i.e. $\Lambda_p^s = \{\lambda_p(R), R \in \mathcal{R}^s\}$. Obviously $\Lambda_p^s \subseteq \Lambda_p$. A graphical depiction of these sets for $K = 3$ and p the uniform distribution is given in Figure 1. The following result locates the choice distributions induced by rational rules as extreme points in the set of choice distributions.

Lemma 2 *Given any distribution p over choice sets, every choice distribution in Λ_p is a convex combination of choice distributions in Λ_p^s .*

Lemma 2 shows that for any realization of \tilde{U}^q , the underlying state, a rule $R \in \mathcal{R}$, can be replaced by a lottery over strictly rational rules in \mathcal{R}^s with the same expected choice distribution. An important consequence is that, for every fitness function u , the lottery over rules in \mathcal{R}^s achieves the same expected fitness as the rule R . If we fix a distribution q on \mathcal{R} and replace each rule of \mathcal{R} by its corresponding lottery over rules in \mathcal{R}^s , we obtain a distribution q' over \mathcal{R}^s such that, for every fitness function u , the distribution of $\tilde{U}^{q'}$ when rules R are drawn according to q' is a mean preserving spread of the distribution of \tilde{U}^q when rules R are drawn according to q . We thus have the following lemma.

Lemma 3 *Let p be any distribution over choice sets. For every sampling distribution q , there exists a sampling distribution q' with $q(\mathcal{R}^s) = 1$ such that, given any fitness function u , the sampling fitness $\tilde{U}^{q'}$ induced by q' in the environment (u, p) is a mean preserving spread of the sampling fitness \tilde{U}^q induced by q .*

Note the order of quantifiers in Lemma 3. For a given distribution over choice sets p , the sampling distribution over exclusively strictly rational rules such that $\tilde{U}^{q'}$ is a mean preserving spread of \tilde{U}^q is the same for all fitness functions u . Yet, the distribution q' still depends on q . Interestingly, under some symmetry assumptions, q' can be taken as the uniform distribution on \mathcal{R}^s , and the dependence on q disappears.

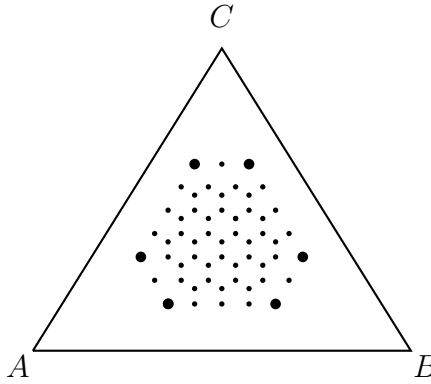


Figure 1: Choice distributions of strictly rational rules (big dots) and of all rules (small dots), for $K = 3$ when p is the uniform distribution over choice sets. The corners of the simplex represent the infeasible case of choosing one and the same of the three alternatives in $K = \{A, B, C\}$ in all decision problems. The convex hull of choice distributions induced by strictly rational rules clearly includes all feasible choice distributions.

Lemma 4 *Assume p is neutral and q is a symmetric sampling distribution. Then the distribution q' of Lemma 3 can be taken as q^s .*

Note that both the assumption that p is neutral as well as the assumption that q is symmetric are necessary in Lemma 4, in the sense that we can find counterexamples to the Lemma for a non-symmetric q and also for a non-neutral p , separately. First, consider the distribution q which puts probability 1 on the single rule which so happens to be a best rule. For almost every p

there is no symmetric distribution which is a mean-preserving spread of this distribution. To see that neutrality is needed consider the following example.

Example: Consider 3 alternatives, $K = \{A, B, C\}$, and the following 8 choice rules (we omit their definitions on singletons).

	R_1	R_2	R_{AB}^s	R_{AC}^s	R_{BA}^s	R_{BC}^s	R_{CA}^s	R_{CB}^s
ABC	ABC	ABC	A	A	B	B	C	C
AB	A	B	A	A	B	B	A	B
BC	B	C	B	C	B	B	C	C
AC	C	A	A	A	A	C	C	C

The rules R_1 and R_2 are not rational, and the distribution q that puts equal weight on each of them is symmetric. To see this, note that for all permutations π which switch exactly 2 elements in K , i.e. these three: $\pi(A) = A, \pi(B) = C, \pi(C) = B$; $\pi(A) = B, \pi(B) = A, \pi(C) = C$; and $\pi(A) = C, \pi(B) = B, \pi(C) = A$ we have that $R_1^\pi = R_2$ and $R_2^\pi = R_1$, while for the other three permutations we have that $R_1^\pi = R_1$ and $R_2^\pi = R_2$.

The other 6 rules are the strictly rational rules. Let p put probability $1/2$ each on $\{A, B\}$ and $\{B, C\}$. This p is not neutral (as neutrality would here require that $\{A, C\}$ receives the same probability as the other two sets). Under p , the distribution $\lambda_p(R_1)$ induced over choices in K is given by $\lambda_p(R_1) = \frac{1}{2}A + \frac{1}{2}B$. With R_2 , the corresponding distribution $\lambda_p(R_2)$ is given by $\lambda_p(R_2) = \frac{1}{2}B + \frac{1}{2}C$. The choice distributions induced by the 6 rational rules are $\lambda_p(R_{AB}^s) = \frac{1}{2}A + \frac{1}{2}B$, $\lambda_p(R_{AC}^s) = \frac{1}{2}A + \frac{1}{2}C$, $\lambda_p(R_{BA}^s) = B$, $\lambda_p(R_{BC}^s) = B$, $\lambda_p(R_{CA}^s) = \frac{1}{2}A + \frac{1}{2}C$, and $\lambda_p(R_{CB}^s) = \frac{1}{2}B + \frac{1}{2}C$.

Attach the following material payoff to the three alternatives. Let $u(A) = 0$, $u(B) = 2$, $u(C) = 4$. The expected fitness for both distributions in the environment (u, p) equals 2. The variance of \tilde{U}^q induced by q is 1, while the variance of \tilde{U}^{q^s} induced by the uniform distribution q^s over strictly rational rules is $\frac{1}{3}$. Hence, the uniform distribution over strictly rational choice rules is not a mean preserving spread of the distribution of $\lambda_p(R)$ induced by the symmetric distribution q .

The following Lemmata provide conditions under which the mean preserving spread in Lemma 3 is strict.

Lemma 5 *Given p and q , there exists a distribution q' as in Lemma 3 such that the obtained mean preserving spread is strict whenever either i) p has full support, u is discriminatory, and $q(\mathcal{R}^s) < 1$ or ii) p, q have full support and u is not constant.*

Note that the provided proof in the appendix shows that part ii) of Lemma 5 holds under the weaker condition that $q(R^0) > 0$, where R^0 is the zero-rule. This suggests that the conditions of the Lemma can be weakened further.

Under the symmetry assumptions of Lemma 4, a strict mean preserving spread obtains under minimal conditions on u , and q .

Lemma 6 *Let the distribution over choice rules p be neutral and have full support and u be non-constant. Then \tilde{U}^{q^s} , the sampling fitness of the strictly rational dynasty, sampling according to the uniform distribution q^s over \mathcal{R}^s , is a strict mean preserving spread of \tilde{U}^q , for any symmetric sampling distribution q with $q(\mathcal{R}^s) < 1$.*

Note that both conditions, that u is not constant and that $q(\mathcal{R}^s) < 1$ are also necessary.

4 Ranking of dynasties in a fixed environment

We use the results of the previous section to show that for any dynasty with sampling distribution q there is another dynasty with sampling distribution q' that puts weight 1 on \mathcal{R}^s (i.e. it uses strictly rational rules only), such that for any time t the expected fitness of the latter dynasty exceeds that of the former. When q is symmetric, q' can be taken as the uniform distribution q^s over strictly rational rules.

Furthermore, under mild assumptions on the environment, at any time $t > 1$ the strictly rational dynasty has strictly higher expected fitness than any symmetric dynasty.

It is important to understand how the expected fitness of a dynasty at generation t is related to the sampling distribution of that dynasty. Let q be any sampling distribution, and R_t denote the choice rule experimented by generation t of the dynasty. The sampling fitness of generation t is $\tilde{U}_t = U(R_t)$, and $\tilde{U}_1, \dots, \tilde{U}_t$ are i.i.d. according to the distribution of the sampling fitness. Because each generation adopts the best rule between the one prevalent at the previous generation and the experimented one, the obtained fitness for generation t is given by $Z_t = \max(\tilde{U}_1, \dots, \tilde{U}_t)$.

Theorem 1 (Rational Dominance) *Let p be arbitrary, and let q be any sampling distribution. There exists a sampling distribution q' such that $q'(\mathcal{R}^s) = 1$ with the property that, for any fitness function u , letting Z_t and Z_t^s be the fitness of the t -th generations of dynasties with sampling distribution q and q' respectively,*

$$\mathbb{E}Z_t^s \geq \mathbb{E}Z_t \text{ for all } t \geq 0$$

with strict inequality for all $t > 0$ if either i) p has full support, u is discriminatory, and $q(\mathcal{R}^s) < 1$ or ii) p, q have full support and u is not constant.

Proof: By Lemma 3 there exists a sampling distribution q' whose associated random fitness \tilde{U}_t^s is a mean-preserving spread of the random fitness \tilde{U}_t associated to q . Under i) or ii) this mean-preserving spread can be taken to be strict by Lemma 5. Note that all \tilde{U}_t^s and \tilde{U}_t are independent and that all \tilde{U}_t , being in fact i.i.d., have the same support. The result (both the weak and the strict inequality) thus follows from Proposition 7 in the Appendix. QED

Theorem 2 (Universal Rational Dominance) *Let p be neutral and let q be any symmetric sampling distribution. Letting Z_t be the fitness of the t -th generations of dynasties with sampling distribution q and Z_t^s be the fitness of*

the t -th generation of the strictly rational dynasty (sampling according to the uniform distribution q^s over \mathcal{R}^s),

$$\mathbb{E}Z_t^s \geq \mathbb{E}Z_t \text{ for all } t \geq 0$$

with strict inequality for all $t > 0$ if u is non-constant, p has full support, and $q(\mathcal{R}^s) < 1$.

Proof: Let \tilde{U}_t^s and \tilde{U}_t be the random fitness associated to the uniform distribution on strictly rational rules and to q , respectively. By Lemma 4, \tilde{U}_t^s is a mean-preserving spread of \tilde{U}_t . If u is non-constant, p has full support, and $q(\mathcal{R}^s) < 1$, moreover, this mean-preserving spread can be taken to be strict by Lemma 6. Note that all \tilde{U}_t^s and \tilde{U}_t are independent and that all \tilde{U}_t , being in fact i.i.d., have the same support. The result (both the weak and the strict inequality) thus follows from Proposition 7 in the Appendix. QED

A corollary of this Theorem is that, provided p is neutral and has full support, the strictly rational dynasty (with uniform sampling q^s over \mathcal{R}^s) performs strictly better than the rational dynasty (with uniform sampling q^r over \mathcal{R}^r) as long as the fitness function u is non-constant. Note that the fitness function can well have ties, i.e. two elements $i, j \in K$ have $u(i) = u(j)$. One might think that the rational dynasty (which allows for indifferences) might have an advantage in this situation. Yet, this is not the case.

Let us summarize the findings of this section. For any distribution p and any dynasty, there exists a dynasty, which samples over strictly rational rules only, that performs at least as well as the former, independently of the fitness function u . The latter does strictly better than the former under mild assumptions. In this sense, strictly rational dynasties are always superior to non-rational (or even just simply rational) ones. Furthermore, a much stronger result obtains if p is neutral. In this case the dynasty that samples uniformly over strictly rational rules does at least as well as, and in many cases strictly better than, any other symmetric dynasty. Hence, under sym-

metry and neutrality, the strictly rational dynasty is superior to any other dynasty, independently of the environment. In fact, under very mild conditions on the environment, the strictly rational dynasty is strictly superior to even the rational dynasty.

5 The long-run dynamics

In this section we address the long-run convergence of the dynamics. We study what rules are to be observed after a large number of generations, and whether these rules are necessarily rational or not. We also question whether full adaptation can be reached by a given dynasty, i.e. whether a dynasty will eventually use a rule that maximizes fitness, and compare the expected time to full adaptation for different dynasties.

5.1 Full adaptation

Consider any environment $e = (u, p)$, and a dynasty with sampling distribution q . Among the rules that are sampled with positive probability, we let \mathcal{R}_q^m denote those with maximal fitness under (u, p) , i.e.

$$\mathcal{R}_q^m = \left\{ R \in \mathcal{R} \mid R \in \arg \max_{R' \in \mathcal{R}: q(R') > 0} U(R') \right\}.$$

The learning process only allows rules with higher fitness to replace rules with lower fitness. Therefore, in a fixed environment, prevalent rules yield higher and higher fitness, until a rule in \mathcal{R}_q^m is reached and the learning process reaches a steady state.

We say that full adaptation is reached when the dynasty uses a rule that maximizes fitness over the whole set of rules. Let \mathcal{R}^* be this set of fitness-maximizing rules, i.e. $\mathcal{R}^* = \{R \in \mathcal{R} \mid R \in \arg \max_{R' \in \mathcal{R}} U(R')\}$.

Therefore, if $\mathcal{R}_q^m \cap \mathcal{R}^*$ is non-empty, the dynasty reaches full adaptation in

the long-run. In this case, can we say if rational rules prevail in the long-run? Given that the prevalent rule in the long run can be any element of $\mathcal{R}_q^m \cap \mathcal{R}^*$, the answer depends on whether this set contains only rational rules or not.

Proposition 1 *\mathcal{R}^* always contains a strictly rational rule. It is a singleton containing this strictly rational rule if and only if p has full support and u is discriminatory.*

Proof: Assume, without loss of generality, that $u(1) \geq u(2) \geq \dots \geq u(K)$. The strictly rational rule R^* that corresponds to the strict preferences $1 \succ 2 \succ \dots \succ K$ maximizes fitness choice set per choice set, hence maximizes $U(R)$ over \mathcal{R} .

If p has full support and u is discriminatory, then any fitness-maximizing rule must maximize fitness in any choice set L with at least two elements (since all such choice sets occur with positive probability) and the only fitness maximizing choice in L is the one chosen by R^* (since u has no ties). Hence R^* is the only fitness-maximizing choice rule.

If u is not discriminatory then there exist two alternatives $l, l + 1$ such that $u(l) = u(l + 1)$, and replacing l by $l + 1$ in $R^*(\{l, l + 1\})$ gives another fitness-maximizing rule. If p does not have full support then there exists a choice set L containing at least two alternatives with $p(L) = 0$, and altering $R^*(L)$ also leaves the fitness unchanged. QED

In the very-long run, every dynasty is perfectly adapted as long as it samples a payoff-maximizing rule with positive probability. If one is interested in the long-run dynamics, sampling over the set of strictly rational rules \mathcal{R}^s and sampling over any superset of \mathcal{R}^s perform equally well.

It is meaningful to compare the evolutionary performance of a strictly rational dynasty endowed with a gene that restricts it to using strictly rational rules only (its sampling distribution is q^s , uniform over \mathcal{R}^s), and of a non-rational dynasty that samples all rules equally (its sampling distribution is q^u , uniform over \mathcal{R}), in an environment $e = (u, p)$, where p has full support

and u is discriminatory. According to Theorem 2, the non-rational dynasty performs better at any generation t than the rational dynasty. However, in the long-run, the two dynasties converge to the same rule, and both are fully adapted. It is interesting, though, to notice that, on the path to full adaptation, only the strictly rational dynasty uses rational rules at every generation. The non-rational dynasty may use non-rational rules during some transitory period.

Thus, restricting sampling to strictly rational rules makes no difference if learning may take place over a large number of generations. It may therefore seem that a “gene of rationality” would provide only a transitory, and possibly not very strong, advantage to a population embedded with it. But how long is the long run in this case and is it realistic to assume that learning can take place over a large enough number of generations in a stable environment? This question is the topic of the next subsection.

5.2 How long is the long run

Let us focus on a fixed environment $e = (u, p)$, where p has full support and u is discriminatory, and let $R^* \in \mathcal{R}$ denote the unique rule which is fully adapted to the environment (u, p) .

For a dynasty that samples according to q , let $T(q)$ denote the number of generations needed to reach full adaptation, i.e. the number of generations after which the rule R^* prevails in the dynasty. $T(q)$ is possibly infinite, and if we let R_t denote the rule sampled by generation t we have that $T(q) = \inf\{t, R_t = R^*\}$.

Proposition 2 *$T(q)$ is infinite if $q(R^*) = 0$, and otherwise, $T(q)$ follows a geometric distribution with parameter $q(R^*)$. In particular, in the latter case, the expected time before full adaptation is given by $\mathbb{E}T(q) = \frac{1}{q(R^*)}$.*

If the sampling distribution q is uniform over its support $\tilde{\mathcal{R}} \subset \mathcal{R}$ then $q(R^*) = \frac{1}{|\tilde{\mathcal{R}}|}$, where $|\tilde{\mathcal{R}}|$ is the number of rules in $\tilde{\mathcal{R}}$. This means that for

uniform sampling distributions, the expected time a dynasty will achieve full adaptation and, hence, maximal fitness, is just given by the cardinality of the set of rules this dynasty samples from, provided the optimal rule R^* is among the eventually sampled ones.

The set of rules with the shortest expected time to full adaptation is, of course, the singleton set including only the one rule which is induced by the correct ranking given by the environment u . The (expected) time to full adaptation is then simply 1. Among all symmetric distributions, the quickest to achieve full adaptation is the uniform distribution over the set of strictly rational rules.

Theorem 3 *Assume that p has full support and u is discriminatory. Let q^s be the uniform distribution over \mathcal{R}^s , and let $q \neq q^s$ be any other symmetric sampling distribution. Then $\mathbb{E}T(q) > \mathbb{E}T(q^s)$.*

Proof: Let $q \neq q^s$ be symmetric. If $q(R^*) = 0$ then $T(q) = \infty > T(q^s)$ almost surely. If $q(R^*) > 0$, by the symmetry of q , $q(R^*) = q(R')$ for all $R' \in \mathcal{R}^s$. But then $q(R^*) = \frac{1}{|\mathbb{I}|}q(\mathcal{R}^s) < \frac{1}{|\mathbb{I}|} = q^s(R^*)$, which proves the result by Proposition 2. QED

Hence, strict rationality provides some evolutionary advantage in the sense that it minimizes the expected time to full adaptation among all symmetric distributions.

Note that Theorem 3 actually extends to all environments, in which case not all fully adapted rules are strictly rational. In this case, we conjecture that the uniform distribution among all strictly rational rules still minimizes (maybe not strictly) the expected time to full adaptation among all symmetric distributions.

The size of some remarkable symmetric sets of rules is given in section A. The following table compares the expected time to full adaptation for the non-rational and rational dynasties for different values of the complexity of the environment.

Complexity (K)	Strictly rational	Non-rational
3	6	24
5	120	10^{11}
10	10^6	10^{658}

Expected time to perfect adaptation

Consider for instance a very fast reproducing organism. Assuming a time between generations of 1h and for a complexity of just 5, the expected time for full adaptation already of the order of 10 million years. Even for the strictly rational population, which is the fastest symmetric one to reach full adaptation, if the complexity of the environment is sufficiently large, it may not be realistic to assume that the environment is stable for a long enough number of generations so that full adaptation can be reached.

The next section measures the evolutionary advantage of the strictly rational population over the non-rational population when only partial adaptation can be reached.

6 Time to partial adaptation

As shown in the previous section, full adaptation cannot be achieved in a reasonable convergence time, even for the strictly rational population, when the complexity of the environment is not small, but not necessarily very large. In this section, giving up on full adaptation, we estimate the time necessary for different dynasties in order to achieve some degree of partial adaptation.

In particular, we address the question of the superiority of the rational dynasty over other symmetric dynasties after a small number of generations. We thus estimate the expected fitness of different dynasties in the short run.

Although we know from Theorem 2 that the strictly rational dynasty achieves no less than any other symmetric dynasty, the difference in expected fitness between the two dynasties after a fixed number of generations depends

on several parameters.

First, this difference must depend on the environment itself. If under u , all alternatives yield close fitness, the expected fitness of all rules, and hence all dynasties, are very close to one another. For this reason, we study the more interesting case in which u 's values are evenly spread, taking for u a bijection from K to itself. Also, the probabilities, defined by p , over sets of alternatives, play an important role. If p is concentrated over one choice set, or on a set of few choice sets, the behavior of choice rules outside of this set have a small impact on the overall fitness. For this reason, our benchmark is the uniform distribution over all choice sets.

Second, the difference in expected fitness depends on the dynasties. Our main focus is on the comparison between the uniform distribution over strictly rational rules and the uniform distribution over all rules. This is particularly meaningful, as we know that the strictly rational rules possess an advantage over any other set of rules, and, in the absence of rationality, we can assume the rules sampled to be distributed uniformly over the whole set of rules. We also derive results for dynasties sampling uniformly over optimistic rules, and over deterministic rules.

Let \tilde{U}^q denote the random fitness of a rule $R \in \mathcal{R}$ randomly drawn according to some symmetric distribution q . From Lemma 1, we know that the expected fitness of a rule drawn according to a symmetric distribution does not depend on the distribution at hand. Given u is a bijection from K to K we can, as an application of Lemma 1, in fact, calculate this expectation. It is given by $\mathbb{E}_q \tilde{U}^q = \frac{K+1}{2}$. In particular, this is the fitness of the zero rule, which always accepts the whole choice set offered ($R^0(L) = L$ for every $L \in \mathcal{L}$).

For this environment we also obtain simple bounds for the fitness of any rule. The maximal payoff (fitness) any rule can obtain is denoted by V and

is equal to

$$\begin{aligned} V &= \frac{2^{K-1}}{2^K - 1}K + \frac{2^{K-2}}{2^K - 1}(K - 1) + \dots + \frac{1}{2^K - 1}1 \\ &= \frac{1}{2^K - 1} \sum_{l=0}^{K-1} (K - l)2^{K-(l+1)} = \frac{2^K}{2^K - 1} \left[(K - 1) + \frac{1}{2^K} \right]. \end{aligned}$$

Let us renormalize payoffs, and define the performance of a rule R as $\rho(R) = \frac{U(R) - \frac{K+1}{2}}{V - \frac{K+1}{2}}$. Thus, a performance of 0 is achieved by the zero rule, whereas a performance of 1 corresponds to full adaptation. The expected time to reach full adaptation was studied in section 5.2. We now investigate the expected time to reach partial adaptation for different populations, defined as the expected time to reach a performance of e.g. 1%, 10%, or 40%.

For a given dynasty, let $\rho(\tilde{R})$ be the performance of a randomly sampled rule \tilde{R} . For $0 \leq \rho \leq 1$, let T_ρ be the random variable corresponding to the first generation reaching a performance of ρ or more. The same logic as in the case of full adaptation shows that T_ρ follows a geometric distribution with parameter $q(\rho(\tilde{R}) \geq \rho)$ if $q(\rho(\tilde{R}) \geq \rho) > 0$, and is infinite otherwise. In the former case, the expected time to reach a performance of ρ is thus $\mathbb{E}T_\rho = \frac{1}{q(\rho(\tilde{R}) \geq \rho)}$. Hence, the expected time to partial adaptation is closely related to how concentrated the distribution of $\rho(\tilde{R})$ is around 0. We investigate this question in the following subsections, for different dynasties. We first investigate the dynasties defined by the uniform distributions over all rules as well as the set of all singleton rules, and then consider the dynasties defined by the uniform distributions over all optimistic rules as well as the set of all strictly rational rules. The final subsection contrasts the two sets of results obtained.

6.1 On all rules

Let q be the uniform distribution over either \mathcal{R} (the set of all rules) or the set of all singleton rules, denoted \mathcal{R}^f . The key point of this subsection is to

recognize that, when R drawn uniformly in either the whole set of rules or in the whole set of single-valued rules, the choices made in different choice sets are independent. We recall the following result (Theorem A.1.18 in Alon and Spencer (2000)) from the theory of large deviations.

Lemma 7 *Let X_1, \dots, X_n be a family of mutually independent random variables with each $\mathbb{E}X_i = 0$ and no two values of any X_i are ever more than 1 apart. Then, for $a > 0$,*

$$P(X_1 + \dots + X_n > a) \leq \exp\left(-\frac{2a^2}{n}\right).$$

An application of Lemma 7 allows us to derive the following lower bound on the expected time to reach a performance of δ .

Proposition 3 *For the dynasties defined by the uniform distribution either over the whole set of rules, or over the set of single-valued rules, the expected time to reach a performance of δ is at least $\exp(\delta^2 2^{K-4})$ for $K \geq 4$.*

Proof: Recall that for a randomly drawn rule R its random fitness is given by $U = \sum_{L \in \mathcal{L}} p(L)u(R(L))$, where $p(L) = \frac{1}{2^{K-1}}$. Let $Z_L = \frac{u(R(L)) - \mathbb{E}_q u(R(L))}{K-1}$. Then (in both cases) the family $(Z_L)_L$ is mutually independent, $\mathbb{E}_q Z_L = 0$ for each L and no two values of Z_L are more than 1 apart. Since $\sum_{L \in \mathcal{L}} Z_L = \frac{2^K-1}{K-1}(U - \mathbb{E}_p U)$ an application of Lemma 7 shows that for $a > 0$ we have $q\left(\frac{2^K-1}{K-1}(U - \frac{K+1}{2}) > a\right) \leq \exp\left(-\frac{2a^2}{2^K-1}\right)$. Substituting for $a = \frac{2^K-1}{K-1}(V - \frac{K+1}{2})\delta$, we obtain

$$q\left(\frac{U - \frac{K+1}{2}}{V - \frac{K+1}{2}} > \delta\right) \leq \exp\left(-\frac{2\left(\frac{2^K-1}{K-1}\left(V - \frac{K+1}{2}\right)\delta\right)^2}{2^K-1}\right).$$

Given $\frac{V - \frac{K+1}{2}}{K-1} > \frac{1}{4}$ for $K \geq 4$, the previous inequality allows us to derive the bound $q(\rho(R) > \delta) < \exp\left(-\frac{2^K-1}{8}\delta^2\right) < \exp(-\delta^2 2^{K-4})$, where the last inequality follows from the fact that, for $K \geq 4$, $\frac{2^K-1}{8} > 2^{K-4}$. QED

6.2 On optimistic and strictly rational rules

Proposition 4 *For the rational and for the optimistic dynasties, the expected time to reach a performance of $\delta = \frac{2j-K-2}{2K-2}$, $j = 1, \dots, K$ is no more than $\frac{4}{1-2\delta}$.*

Note that Proposition 4 only yields (interesting) bounds on performance levels, δ , between 0 and $\frac{1}{2}$. One could, in principle, modify the proof such as to cover performance levels above $\frac{1}{2}$. As we here are interested in the short run performance of these dynasties we do not pursue this extension here.

Let q now be the uniform distribution over either \mathcal{R}^o (the set of optimistic rules) or \mathcal{R}^s (the set of strictly rational rules). Let U^o and U^s , respectively, denote the random fitness of a rule drawn uniformly from \mathcal{R}^o and \mathcal{R}^s . We first prove the following lemma, expressed in terms of payoffs instead of performances.

Lemma 8 *For both $U = U^o$ and $U = U^s$ we have $q(U \geq \frac{j}{2} + \frac{K}{4}) \geq \frac{K-j+1}{2K}$.*

Proof: It is sufficient to consider the case in which the environment u is the identity mapping from K to K . Let R be a randomly chosen optimistic or strictly rational rule. Let $R(K)$ denote its preferred element in K (which is unique in both cases). Since R is chosen uniformly, the distribution of $R(K)$ is uniform in K . i.e. $q(R(K) = j) = \frac{1}{K}$ for all $j \in K$ and $q(R(K) \geq j) = \frac{K-j+1}{K}$ for all $j \in K$. For $j \in K$ let A^j be the event $R(K) = j$, i.e. the event (set of rules) with the property that the agent chooses j if presented with choice set K . Conditional on the agent's rule being in A^j , 2^{K-1} of all choice sets, those $L \in \mathcal{L}$ with $j \in L$, also must provide fitness j . The agent's choice in the other $2^{K-1} - 1$ choice sets, those $L \in \mathcal{L}$ with $j \notin L$, is completely independent of A^j . For these $2^{K-1} - 1$ choice sets the agent's random payoff, denoted U' , is that derived from the payoff of a choice rule from a symmetric distribution when there are $K - 1$ choices and payoffs are in the set $\{1, 2, \dots, j - 1, j + 1, \dots, K\}$. This random variable U' is first-order stochastically dominated by another random variable \hat{U} which is the payoff obtained by a rule over $K - 1$ choices

drawn according to the same symmetric distribution when payoffs are in the set $\{1, 2, \dots, K - 1\}$. This last distribution is symmetric around its mean $\frac{K}{2}$. Hence,

$$q\left(U \geq \frac{2^{K-1}}{2^K - 1}j + \frac{2^{K-1} - 1}{2^K - 1} \frac{K}{2} \mid A^j\right) \geq \frac{1}{2}.$$

Now, for any x , $q(U \geq x) = \sum_{j=1}^K q(U \geq x \mid A^j) q[A^j]$. Also for any $j' \geq j$

$$q\left(U \geq \frac{j}{2} + \frac{K}{4} \mid A^{j'}\right) \geq q\left(U \geq \frac{2^{K-1}}{2^K - 1}j + \frac{2^{K-1} - 1}{2^K - 1} \frac{K}{2} \mid A^{j'}\right) \geq \frac{1}{2}.$$

Hence,

$$\begin{aligned} q\left(U \geq \frac{j}{2} + \frac{K}{4}\right) &= \sum_{j'=1}^K q\left(U \geq \frac{j}{2} + \frac{K}{4} \mid A^{j'}\right) q\left(A^{j'}\right) \\ &\geq \sum_{j' \geq j} q\left(U \geq \frac{j}{2} + \frac{K}{4} \mid A^{j'}\right) q\left(A^{j'}\right) \\ &\geq \sum_{j' \geq j} \frac{1}{2} \frac{1}{K} = \frac{K - j + 1}{2K}. \end{aligned}$$

QED

Proof of Proposition 4: Let ρ denote the performance of either a uniformly chosen optimistic rule or a uniformly chosen rational rule. Lemma 8 shows that $q\left(\rho \geq \frac{\frac{j}{2} + \frac{K}{4} - \frac{K+1}{2}}{V - \frac{K+1}{2}}\right) \geq \frac{K-j+1}{2K}$. Using that $V \leq K$ we obtain $q\left(\rho \geq \frac{2j-K-2}{2K-2}\right) \geq \frac{K-j+1}{2K}$. Setting $\delta = \frac{2j-K-2}{2K-2}$, this becomes $q(\rho \geq \delta) \geq \frac{1}{4} - \frac{K-1}{2K-2}\delta \geq \frac{1-2\delta}{4}$. QED

6.3 The advantage of rationality for partial adaptation

Using the results of the previous subsections, we compare the performances of the strictly rational dynasty with that of the non-rational dynasty. When the complexity of the environment is large enough, full adaptation cannot be reached in a reasonable number of generations, and the adequate comparison should be on the time needed to reach partial adaptation.

Proposition 3 shows us that the expected time to reach a performance of δ for the non-rational dynasty (single valued or not) is a double exponential in the complexity of the environment, K . As shown by the following table, the number of generations needed to reach a performance level grows very fast with the complexity of the environment. Hence, unless the environment is very stable, and adaptation can take place over a large number of periods, the non-rational population possesses no significant advantage over a dynasty that would simply implement the zero-rule at every generation.

$K \setminus \delta$	1%	5%	10%
5	1.08	8	4000
6	2.5	10^9	10^{39}
7	7300	10^{96}	10^{386}

Lower bound on time to a performance of δ for various levels of complexity K ; non-rational population

The results obtained for the non-rational dynasty contrast sharply with the results obtained for the rational (or optimistic) dynasty, for which Proposition 4 provides an upper bound on the expected time to reach a performance which is independent of the complexity of the environment. As shown by the following table, even after a small number of generations, the performance of the rational dynasty is already substantial.

$K \setminus \delta$	1%	25%	40%
Any	5	8	20

Upper bound on time to a performance of δ (independent of complexity K); strictly rational population

7 Changing environments

In this section we investigate the evolutionary performance of populations in a changing environment. The environment now follows a stochastic process,

and we let $e_t = (u_t, p_t)$ be the environment faced by generation t .

The environment is renewed at stochastic times, and $\tau = (\tau_1, \tau_2, \dots)$ where $\tau_k < \tau_{k+1}$ describes the process of renewal times. An interesting example is the case in which the intervals $\tau_{k+1} - \tau_k$ between renewal times follow exponential processes with same parameter, but our results cover the more general case where the probability that the environment is renewed at a given time is history dependent.³ In particular, the rate of renewal of the environment needs not be constant over time. Periods of high instability of the environments can be followed by highly stable periods.

Environments are drawn according to a distribution Q over the set \mathcal{E} of environments. The first environment, e_1 , is drawn according to Q . If the environment is renewed at stage $t > 1$ ($t = \tau_k$ for some k), a new environment e_t is drawn according to Q . Otherwise, the environment stays the same and $e_t = e_{t-1}$.

Recall that, for an environment $e = (u, p)$ and a choice rule $R \in \mathcal{R}$ the rule's average fitness in that environment is given by $U_e(R) = \mathbb{E}_p u(R(L)) = \sum_{L \in \mathcal{L}} p(L) u(R(L))$.

We turn to the evolution of rules adopted by agents. Each generation t samples a rule \hat{R}_t where the family of sampled rules is i.i.d. according to the dynasty's sampling distribution q . The prevalent (or adopted) rule at generation t is denoted R_t . The first generation adopts the sampled rule $R_1 = \hat{R}_1$. Generation $t+1$ adopts the best rule between the prevalent rule at time t and the sampled rule at time $t+1$, where both rules are evaluated in the environment in place at stage $t+1$. Thus, $R_{t+1} = \arg \max \left\{ U_{e_{t+1}}(R_t), U_{e_{t+1}}(\hat{R}_{t+1}) \right\}$. The induced process over agents' fitness is given by $Z_t = U_{e_t}(R_t)$.

A distribution Q over environments in \mathcal{E} is **symmetric** if for every $e = (u, p) \in \mathcal{E}$, p is neutral and for every permutation $\pi : K \rightarrow K$ we have that $Q(e) = Q(e^\pi)$, where $e^\pi = (u^\pi, p^\pi)$ and $u^\pi(L) = u(\pi(L))$ and $p^\pi(L) =$

³It is, however, not allowed to depend on rules or the environment prevalent at the time. We are thus ruling out cases in which the environment is influenced by the rules individuals follow, i.e. we are ruling out phenomena like human induced global warming.

$p(\pi(L))$ for all $L \in \mathcal{L}$.

The following result shows that, if the distribution of environments is symmetric, the strictly rational dynasty performs better than any dynasty with a symmetric sampling rule, after any number of generations.

Theorem 4 (Universal Dominance Theorem for Changing Environments)

Let Q be a symmetric distribution over \mathcal{E} and let q be a symmetric sampling distribution over \mathcal{R} . Letting Z_t be the fitness of the t -th generation of the dynasty with sampling distribution q and Z_t^s be the fitness of the t -th generation of the strictly rational dynasty,

$$\mathbb{E}Z_t^s \geq \mathbb{E}Z_t \text{ for all } t \geq 0$$

with strict inequality for all $t > 1$ if there exists an environment $e = (u, p)$ in the support of Q such that u is non-constant and p has full support, and $q(\mathcal{R}^s) < 1$.

Proof: We prove a stronger result, which is that the inequality holds conditional on any realization of the process τ of renewals of the environment. We first prove that the weak inequality also holds conditional on the current environment e_t . Let $\tau_k \leq t < \tau_{k+1}$. We then have $e_t = e_{t-1} = \dots = e_{\tau_k}$, and denote this environment by e . Conditional on τ and on e the expected fitness of generation t for the non-rational dynasty is given by

$$\mathbb{E}[\max\{U_e(R_{\tau_k-1}), U_e(\hat{R}_{\tau_k}), \dots, U_e(\hat{R}_t)\} | \tau, e].$$

Note that the random variables $U_e(R_{\tau_k-1}), U_e(\hat{R}_{\tau_k}), \dots, U_e(\hat{R}_t)$ are independent given τ, e , and that they all share the same support. Given the symmetries of Q and q , the distribution of R_{τ_k-1} given e, τ is symmetric, and so are the distributions of $\hat{R}_{\tau_k}, \dots, \hat{R}_t$. Also $U_e(R_{\tau_k-1}^s), U_e(\hat{R}_{\tau_k}^s), \dots, U_e(\hat{R}_t^s)$ are

independent given τ, e . From Lemma 4 and Proposition 7 we obtain

$$\begin{aligned} \mathbb{E}[\max\{U_e(R_{\tau_k-1}), U_e(\hat{R}_{\tau_k}), \dots, U_e(\hat{R}_t)\} | \tau, e] &\leq \\ \mathbb{E}[\max\{U_e(R_{\tau_k-1}^s), U_e(\hat{R}_{\tau_k}^s), \dots, U_e(\hat{R}_t^s)\} | \tau, e] & \end{aligned}$$

If there exists an environment $e = (u, p)$ in the support of Q such that u is non-constant and p has full support, and $q(\mathcal{R}^s) < 1$ then this inequality is strict by the fact that $U_e(\hat{R}_t^s)$ is a strict mean preserving spread of $U_e(\hat{R}_t)$ due to Lemma 6 and the appropriate appeal to Proposition 7. The Theorem then follows from averaging over all τ, e . QED

Theorem 4 shows that the main message of the discussion of evolution in a fixed environment remains unchanged in a changing environment. If the distribution generating environments is symmetric then a dynasty using a symmetric sampling distribution is fitness-dominated in every generation by the strictly rational dynasty. We conjecture that we can relax either (but not both) of the two symmetry assumptions and yet obtain the same result.

One may wonder how the results of Section 6 that quantify the advantage of strictly rational rules over the whole set of deterministic rules can be generalized from the case of a fixed environment to the case of a changing environment. Instead of adapting or mimicking the logic of the proofs of section 6, we show that the expected fitness of a symmetric dynasty t stages after a change in environment has natural bounds in term of the expected fitness of the same dynasty after t generations.

Fix a realization of τ , and let R_{τ_k-1} denote the inherited rule at stage τ_k . Assume that $\tau_k + t \leq \tau_{k+1}$. The fitness of generation $\tau + t$ (thus t stages after the last environment change) is $Z_{\tau_k+t} = \max\{U_e(R_{\tau_k-1}), U_e(\hat{R}_{\tau_k}), \dots, U_e(\hat{R}_{\tau_k+t})\}$, where $\hat{R}_{\tau_k}, \dots, \hat{R}_t$ are i.i.d. according to Q . Let $Z_t^f = \max(U_e(\hat{R}_{\tau_k}), \dots, U_e(\hat{R}_{\tau_k+t}))$ denote the fitness of the dynasty after t stages in the fixed environment e . Bounds on Z_t^f are obtained in section 6.

Proposition 5 *Assume that Q, q are symmetric, fix realizations of the re-*

newal times τ and consider a stage t such that $\tau_k + t \leq \tau_{k+1}$. We have

$$\mathbb{E}[Z_t^f | \tau, e] \leq \mathbb{E}[Z_{\tau_k+t}^f | \tau, e] \leq \mathbb{E}[\max\{U_e(\hat{R}^s), Z_t^f\} | \tau, e]$$

where \hat{R}^s is a uniformly drawn strictly rational rule.

Proof: The left-hand inequality is immediate. For the right-hand inequality, note that given τ and e , $U_e(\hat{R}_{\tau_k}), \dots, U_e(\hat{R}_{\tau_k+t})$ are independent and the distribution of \hat{R}_{τ_k} is symmetric, and apply Lemma 4 and Proposition 7. QED

This result demonstrates two things. A dynasty does better t stages after the last change of environment than after t stages in the fixed environment case (simply, due to the inheritance of a rule from the previous environment). Nevertheless, this advantage is bounded by the advantage that would be provided by adding one strictly rational uniformly chosen rule to the rules sampled by the dynasty.

This shows that, analogous to results that quantify the adaptation level of a dynasty in the fixed environment, very similar bounds can be provided for the expected fitness of the same dynasty in a changing environment.

8 Related literature

The literature offers many evolutionary models in several different contexts. A great survey is Robson (2001b). The closest to our paper is perhaps Robson (2001a) who shows that individuals evolve to evaluate gambles (two-armed bandits) according to the “correct” expected utility criterion. The individuals in Robson (2001a)’s model choose at each point in time among two arms of a multi-armed bandit. Each arm provides a lottery over a given and fixed finite set of consumption levels, each of which in turn provides a fixed Poisson distributed number of offspring which only depends on the consumption level. While the set of consumption levels and its link to offspring

is fixed, the distribution of the two arms changes at discrete points in time. Individuals, however, always have enough time to correctly figure out these two distributions. One might say these individuals are perfectly adapted to the distribution over consumption levels of the two arms but not necessarily adapted to the correct link between consumption levels and expected number of offspring.

Robson (2001a) shows that evolution then favors those individuals who are adapted in both senses, i.e. adapted to the correct distribution of consumption levels for the two arms of the bandit, as they must be given the model, as well as adapted to the correct link between consumption levels and expected number of offspring. Thus evolution in Robson (2001a)'s model favors individuals who behave "as if" they are expected utility maximizers with the correct utility function.

The additional insight our paper provides is to demonstrate the evolutionary value of being "rational" even when not perfectly adapted. We could have studied the question of the evolutionary value of "rationality" in Robson (2001a)'s model. In order to find analogue results to ours in Robson (2001a)'s model, we would have to investigate the (possible) evolutionary advantage of being an expected utility maximizer without having correct beliefs, i.e. we would compare individuals who just follow some rule of thumb with individuals who have a consistent theory in their mind about the link between consumption levels and expected number of offspring. In order to see the exact advantage of "rationality" even without perfect "adaptedness" more clearly we chose to study the more basic problem of consumer choice, i.e. without uncertainty, which we also find of interest in its own right. We conjecture, though, that similar results to ours can be shown in Robson (2001a)'s context of decision making under uncertainty, i.e. that being rational even without perhaps ever being adapted is strongly favored by evolution over those who are not rational, where rational is in the sense of revealed preferences. In fact we consider our results, while of interest in its own right

as they pertain to consumer choice, also to be a metaphor to illustrate the general evolutionary advantage of consistent behavior according to some true regularity properties about the world the individual lives in. Whatever these regularities are, populations using rules of behavior that are suited to this property have an evolutionary edge over those who do not.

Many other well-known models of evolution do not directly allow the discussion whether or not evolution favors actual rationality over “as if” rationality. For instance the literature on the evolution of preferences in games, based on the indirect evolutionary approach of Güth and Yaari (1992) and Güth (1995), perhaps culminating in Dekel, Ely, and Yilankaya (2007), assumes that individuals always hold consistent preferences, even if not necessarily the correct ones. I.e. these individuals are all rational, perhaps not adapted. Most of evolutionary game theory also does not allow the distinction between rational and “as if” rational. Typically (see e.g. Weibull (1995) for a textbook treatment of evolutionary game theory) individuals are programmed to play certain strategies and just may or may not eventually disappear. The evolutionary models which share these characteristics and, hence, for which the discussion below applies, includes static concepts of evolutionary stability such as the concept of an evolutionary stable strategy (ESS) of Maynard Smith and Price (1973), deterministic dynamic models of evolution such as the replicator dynamics of Taylor and Jonker (1978), as well as stochastic models such as that of Kandori, Mailath, and Rob (1993). If such an evolutionary process leads to a convergence point of its dynamics or an in some sense stable outcome, which then typically constitutes a Nash equilibrium, it looks “as if” individuals are rational. In fact this comes with the added implication that also their beliefs about nature as well as about their opponents are perfectly correct, i.e. they are adapted. Again they appear “as if” they are rational as well as “as if” they know the environment they live and act in, while they are not actually rational, being simply programmed to their strategy choice.

Interesting exceptions to this literature are the models of Dekel and Scotchmer (1992) and Banerjee and Weibull (1995). In addition to the above-mentioned programmed individuals playing the game, Dekel and Scotchmer (1992) allow for individuals who are programmed to rules, such as the rule “play a best response to the previous period’s population”, while Banerjee and Weibull (1995) allow individuals of a type, termed *homo oeconomicus*, who, when called upon to play, always knows either the exact behavior of its opponent or at least the correct distribution of behaviors and plays a best response to its information. In either case one could call these additional types more rational than the programmed ones. Yet, the notion of rationality in both Dekel and Scotchmer (1992) and Banerjee and Weibull (1995) is obviously somewhat different than that in our paper as theirs pertains mostly to acting on beliefs about (or knowledge of) opponent behavior. Thus also the results are different. Both Dekel and Scotchmer (1992) as well as Banerjee and Weibull (1995) show that, while such rational individuals typically survive evolution, also programmed individuals can survive evolution for a wide class of games.

A lack of adaptedness is also central in Samuelson and Swinkels (2006) and Ely (2007). In Samuelson and Swinkels (2006), in a model close to Robson (2001a), nature is restricted in that she is not able to endow agents with the correct information-processing. Thus individuals, by assumption cannot be completely adapted to their environment. Samuelson and Swinkels (2006) show how nature then makes up for this inability by attaching utility to, fitness-irrelevant, intermediate actions. The resulting utility then has some interesting features such as choice-set dependence and its induced self-control problems. Ely (2007) provides a natural model, in which agents with positive probability never achieve perfect adaptation, even though the environment is, in essence, fixed. The present paper demonstrates the value of consistency, a vital part of rationality, especially in cases where individuals can never hope to be perfectly adapted to the world they live in.

Finally, while Campbell (1978) argues that rational choice functions are easy to compute, Kalai (2003) that they are easy to learn, Rubinstein (1996) that they are easy to describe, and Salant (2007) that they are procedurally simple, this paper argues that rational choice functions are likely to survive, i.e. have a strong evolutionary appeal.

9 Conclusion

In this paper we study a possible foundation of a crucial aspect of rationality (consistency in choice behavior) based on evolutionary selection arguments. Put simply, the questions we address are the following. 1) Would a “gene of rationality” provide an evolutionary advantage to a population of individuals who carry it, 2) which environmental conditions are the most favorable to the emergence of such a gene, and 3) how strong is this evolutionary advantage if there is one?

In a model in which different populations of individuals differ in the way they experiment when trying new rules of behavior, the answer to the first question is positive. The Rational Dominance Theorem demonstrates that any sampling distribution individuals use can be replaced by a sampling distribution over strictly rational rules with no loss in fitness to these individuals. The Universal Rational Dominance Theorem, furthermore, shows that uniform sampling over strictly rational rules does at least as well as any sampling distribution for which no particular choice plays a special role. Moreover, in both cases, the strictly rational population performs strictly better than the other under mild conditions on the environment. This strict superiority in fitness of the strictly rational population is present at every generation, both if the environment is constant over time (Theorem 1) and when the environment changes (Theorem 4).

To answer the second question, we find that this form of rationality has a particular advantage if there is sufficient variability on the side of the environ-

ment. Indeed, if the environment is very stable, and there is sufficient time for the evolutionary processes to reach convergence, then in the ultra-long run, both “rational” and “non-rational” populations achieve perfect adaptation and, thus, behavior that seems rational. It is worth noting, however, that in this case the time to perfect adaptation is much smaller for the rational population than for the non-rational one (see Theorem 3, Proposition 2 and Section A). For a very large range of parameters between (not including) complete instability and (including) complete instability, the advantage of rationality is quite substantial, which thus leads us to our answer to the third question.

The rational population reaches a significant degree of adaptation (although not necessarily perfect) in any finite number of generations larger than one, independently of the complexity of the environment (see Proposition 4). This contrasts sharply with the non-rational population, whose degree of adaptation stays desperately close to none for a large number of generations up to a point in time, which is a double exponential in the number of alternatives. Thus the rational population has a very strong evolutionary advantage over the non-rational one for a very large range of parameters.

Our findings are consistent with evidence from anthropology and cognitive sciences. Richerson, Bettinger, and Boyd (2005) show that periods of higher variability of the environment are closely followed by higher degrees of adaptation of cognitive skills of early humanoids. They argue that advanced cognitive skills permit social learning, and that this form of learning is suited for rapid adaptation (of the order of a few dozen generations), which is much faster than biological adaptation (of the order of 100,000 years). Another important advantage of highly developed cognitive skills is that they may be suited for the implementation of rational rules. In view of our results, social learning may permit fast selection of better suited rules in the population, but this fast selection alone is insufficient if sampling is done over the set of non-rational rules. But if both rationality and social learning are signif-

icantly enhanced by higher cognitive skills, they, in fact, complement each other in order to achieve fast adaptation in a changing environment.

Egan, Santos, and Bloom (2007) study sequential choices among alternatives by four year-old children and capuchin monkeys. The experiments are designed in such a way that the subjects are a priori indifferent between all alternatives. Yet, observed behavior is compatible with strictly rational choice rules, but allow to reject the hypothesis that all rational choice rules may be used (see also Chen (2008)). It is quite striking that subjects seem to “break indifferences” between alternatives they are initially indifferent from. Such a propensity to break indifferences between any pair of alternatives is, in fact, one of the implications of this paper. We show that strictly rational rules perform better than merely rational rules, even when the environment exhibits indifferences.

A final caveat is in order. Even though our results argue in favor of rational behavior, we do not predict individuals to always make the correct choices. The sense in which we understand rationality here is that it is a form of regularity, or consistency, of the agent’s behavior. In the particular case of the choice between alternatives we are interested in, this regularity takes the form of the weak axiom of revealed preferences and of single choices, equivalently represented by a strict preference relation. We do not make the case, however, that an agent’s preference should always be the best suited given the environment the agent lives in. Quite on the contrary, we show that a population may never reach perfect adaptation if the environment is not extremely stable over time. In other words, our results are not inconsistent with empirical findings that agents may make choices that are not in their best self-interest.

What we offer is a reinterpretation of the source of maladaptation of choices to the environment as coming from a difficulty in reaching perfect adaptation when the agent is confronted with new situations, rather than from a lack of rationality per se on the agent’s side. Looking at deviations of

Homo Economicus from the classical paradigm of optimal decision making in view of this new light (i.e. while maintaining some consistency assumptions on the agents' part) might provide fruitful directions of future research.

A On the number of rules

To facilitate a more detailed comparison of the time to full adaptation for some sets of rules we here count the number of rules in each of these sets. The number of rules are given in the following table for the set of all rules \mathcal{R} , the set of all singleton-value rules \mathcal{R}^f , the set of all rational rules \mathcal{R}^r , and the set of all strictly rational rules \mathcal{R}^s .

$$\begin{array}{cccc}
\mathcal{R} & \mathcal{R}^f & \mathcal{R}^r & \mathcal{R}^s \\
\prod_{L \in \mathcal{L}} (2^{|L|} - 1) & \prod_{L \in \mathcal{L}} |L| & & \\
= \prod_{l=0}^{k-1} (2^{k-l} - 1) \binom{k}{l} & = \prod_{l=0}^{k-1} (k-l) \binom{k}{l} & \in [k!, (k!)2^{k-1}] & k!
\end{array}$$

In addition, we have $|\mathcal{R}^o| = k \prod_{L \in \mathcal{L}; j \notin L} (2^{|L|} - 1)$ for any fixed $j \in K$, and $|\mathcal{R}^p| = k \prod_{L \in \mathcal{L}; j \notin L} (2^{|L|} - 1) \prod_{L \in \mathcal{L}; j \in L} (2^{|L|-1} - 1)$, or alternatively, $|\mathcal{R}^o| = k \prod_{l=0}^{k-2} (2^{k-1-l} - 1) \binom{k-1}{l}$ and $|\mathcal{R}^p| = k \left[\prod_{l=0}^{k-2} (2^{k-1-l} - 1) \binom{k-1}{l} \right]^2$. For $k = 3$ and $k = 4$, for instance, we have

	\mathcal{R}	\mathcal{R}^f	\mathcal{R}^r	\mathcal{R}^s	\mathcal{R}^o	\mathcal{R}^p
$k = 3$	189	24	$19 \in [6, 24]$	6	9	27
$k = 4$	26245935	20736	$\in [24, 192]$	24	756	142884.

B Mean Preserving Spreads

Let X, Y be random variables with supports included in the finite sets \mathcal{X} and \mathcal{Y} respectively, subsets of \mathbb{R} or \mathbb{R}^n . Recall that X is a **mean preserving spread** of Y if there exists a collection of vectors of weights $\alpha = \{\alpha_y\}_{y \in \mathcal{Y}}$ with $\alpha_y \in \Delta(\mathcal{X})$ such that $\sum_x \alpha_y(x)x = y$ for every $y \in \mathcal{Y}$ and such that

$P(X = x) = \sum_{y \in Y} P(Y = y) \alpha_y(x)$. If furthermore there exists y such that $P(Y = y) > 0$ and α_y does not put probability 1 on y , we say that X is a **strict mean preserving spread** of Y . Equivalently, X is a strict mean preserving spread of Y if and only if it is a mean preserving spread of Y and the distributions of X, Y differ.

Note that in creating X as a mean preserving spread of Y , we are “replacing” each y in the support of Y with a distribution over x ’s, such that its mean is exactly y . Alternatively one could say that we are “splitting” y into a distribution over x ’s, leaving the mean the same. We return to “splittings” in Appendix C. Since mean preserving spreads depend on random variables only through their distributions, we also say that the distribution of X is a (strict) mean preserving spread of the distribution of Y .

Proposition 6 *Let X, Y, Z be real valued random variables with finite support such that X is a mean preserving spread of Y , and let Z is independent of X and of Y . Then*

$$\mathbb{E} \max(X, Z) \geq \mathbb{E} \max(Y, Z)$$

Furthermore, the inequality is strict if X is a strict mean preserving spread of Y and the support of Z contains the support of Y .

Proof: For each pair of values y, z , Jensen’s inequality implies that $\sum_x \alpha_y(x) \max(x, z) \geq \max(y, z)$ with strict inequality if α_y does not put probability 1 on y and $z = y$ (as the mapping $y \mapsto \max(y, z)$ is strictly convex at the point $y = z$). Summing over all values of Y , multiplying each inequality by $P(Y = y)$ yields $\mathbb{E} \max(X, z) \geq \mathbb{E} \max(Y, z)$ with strict inequality if $P(Y = z) > 0$ and $\alpha_z(z) < 1$. Now multiplying each inequality by $P(Z = z)$ and summing over values of z gives the result. QED

Proposition 7 *Let $(X_i)_{i=1, \dots, n}$ and $(Y_i)_{i=1, \dots, n}$ be two families of independent random variables such that for each $i = 1, \dots, n$ let X_i be a mean preserving-*

spread of Y_i . Then

$$\mathbb{E} \max(X_1, X_2, \dots, X_n) \geq \mathbb{E} \max(Y_1, Y_2, \dots, Y_n).$$

The inequality is strict if at least one of the mean preserving spreads is strict and Y_1, \dots, Y_n have same support.

Proof: Assume (X_i) and (Y_i) are independent. For $k \in \{1, \dots, n\}$, Lemma 6 gives the inequality

$$\begin{aligned} \mathbb{E} \max(X_1, \dots, X_k, Y_{k+1}, \dots, Y_n) &= \mathbb{E} \max(X_k, \max(X_1, \dots, X_{k-1}, Y_{k+1}, Y_n)) \\ &\geq \mathbb{E} \max(Y_k, \max(X_1, \dots, X_{k-1}, Y_{k+1}, Y_n)) \\ &= \mathbb{E} \max(X_1, \dots, X_{k-1}, Y_k, \dots, Y_n) \end{aligned}$$

If Y_1, \dots, Y_n all have the same support then $\max(X_1, \dots, X_{k-1}, Y_{k+1}, Y_n)$ also has this support (the minimum of the support of X_i is no larger than the minimum of the support of Y_i), so that, if X_i is a strict mean preserving spread of Y_i , the above inequality is strict by the ‘‘strict’’ part of Lemma 6. The result then follows from a finite chain of such inequalities. QED

C Proofs of Results in Section 3

C.1 Proof of Lemma 1

By definition

$$\mathbb{E}_q \tilde{U}^q = \sum_{R \in \mathcal{R}} q(R) U(R) = \sum_{R \in \mathcal{R}} q(R) \mathbb{E}_p u(R(L)).$$

Changing the order of summation we have

$$\mathbb{E}_q \tilde{U}^q = \mathbb{E}_p \sum_{R \in \mathcal{R}} q(R) u(R(L)).$$

For any choice set L and any permutation π of K , the symmetry of q implies

$$\sum_{R \in \mathcal{R}} q(R)u(R(L)) = \sum_{R \in \mathcal{R}} q(R)u(\pi(R(L))).$$

By averaging the above equality on all permutations π , we deduce that

$$\begin{aligned} \sum_{R \in \mathcal{R}} q(R)u(R(L)) &= \sum_{R \in \mathcal{R}} q(R) \sum_{\pi \in \Pi} \frac{1}{|\Pi|} u(R(\pi(L))) \\ &= \sum_{R \in \mathcal{R}} q(R)u(L) \\ &= u(L), \end{aligned}$$

and, hence, the result. QED

C.2 Proof of Lemmata 2 and 3

Proof of Lemma 2: It is enough to prove that, for any vector $v = (v(k))_k \in \mathbb{R}^K$, $\max_{\lambda_p \in \Lambda_p} \sum_k \lambda_p(k)v(k)$ is attained for some $\lambda_p \in \Lambda_p^s$. Interpret $v(k)$ as a “utility” for the choice k . For $L \subseteq K$, let $v(L) = \frac{1}{|L|} \sum_{l \in L} v(l)$. Let π be a permutation of K that orders the coordinates of v such that $v(\pi(1)) \geq v(\pi(2)) \geq \dots \geq v(\pi(k))$. Maximizing $\sum_k \lambda_p(k)v(k)$ over $\lambda_p \in \Lambda_p$ is equivalent to maximizing the expected “utility” $\sum_{L \in \mathcal{L}} p(L)v(R(L))$ over all rules.

The rule R^π that selects the least element according to π in every choice set, $R(L) = \min\{l, \pi(l) \in L\}$, maximizes each term of the sum $\sum_{L \in \mathcal{L}} p(L)v(R(L))$, so it maximizes the sum. Also, R^π is strictly rational, since it is the rule that corresponds to the preference relation $\pi(1) \succ \pi(2) \succ \dots \succ \pi(k)$. Hence, $\lambda_p(R^\pi)$ belongs to Λ_p^s , and achieves $\max_{\lambda_p \in \Lambda_p} \sum_k \lambda_p(k)v_k$. QED

Proof of Lemma 3: Lemma 2 implies that for any $R \in \mathcal{R}$ there exist non-negative coefficients $\{\alpha_{R,R^s}\}_{R^s \in \mathcal{R}^s}$ summing to 1 such that $\lambda_p(R) = \sum_{R^s} \alpha_{R,R^s} \lambda_p(R^s)$. Let q' be the distribution over \mathcal{R}^s defined by $q'(R^s) =$

$\sum_R \alpha_{R,R^s} q(R)$. It follows by construction that the distribution of $\lambda_p = \lambda_p(R)$ under q' is a mean preserving spread of the distribution of λ_p under q . The result follows, since for any u , the application that associates to a choice distribution λ_p the corresponding sampling fitness $\sum_{k \in K} \lambda_p(k)u(k)$ is linear. QED

C.3 Proof of Lemma 4

As in the proof of Lemma 3, it suffices to prove that the distribution of λ_p under q^s is a mean preserving spread of the distribution of λ_p under q . Recall that, with $q'(R^s) = \sum_R \alpha_{R,R^s} q(R)$, the distribution of λ_p under q' is a mean preserving spread of the distribution of λ_p under q . For any permutation π of K , with $q'^\pi(R^s) = \sum_R \alpha_{R^\pi,R^{s\pi}} q(R^\pi) = \sum_R \alpha_{R^\pi,R^{s\pi}} q(R)$, the distribution of λ_p under q'^π is also a mean preserving spread of the distribution of λ_p under q . Hence, with $q'' = \frac{1}{|\Pi|} \sum_\pi q'^\pi$, the distribution of λ_p under q'' is, again, a mean preserving spread of the distribution of λ_p under q . We now complete the proof by showing that $q'' = q^s$, and for this, it is enough to show that q'' is symmetric. Indeed, for every permutation π' ,

$$q''(R^{s,\pi'}) = \frac{1}{|\Pi|} \sum_R \left(\sum_\pi \alpha_{R^{\pi\pi'},R^{s\pi'}} \right) q(R) = \frac{1}{|\Pi|} \sum_R \left(\sum_\pi \alpha_{R^\pi,R^s} \right) q(R) = q''(R^s).$$

QED

C.4 Proofs of Lemmata 5 and 6

We here identify conditions under which for some sampling distribution q , there is another distribution q' with $q'(\mathcal{R}^s) = 1$ such that, not only is the distribution of $\tilde{U}^{q'}$ a mean preserving spread of \tilde{U}^q , but also these two distributions are not identical. This means that the mean preserving spread is strict.

Note that the distribution q' as in Lemma 3 is not necessarily unique. The mean preserving spread may be strict for some q' , but not for some other q' . In order to tackle this issue, we look for mean preserving spreads with “maximal” support, in the sense that each rule R with $q(R) > 0$ is itself “split” (as in the construction of the proof of Lemma 3) into a maximal subset of \mathcal{R}^s .

There are two main reasons why, for a given sampling distribution q , we might not be able to construct a strict mean-preserving spread of it.

First, there might be a rule $R \in \mathcal{R}$ such that for all α_{R,R^s} , summing to one, with $\lambda_p(R) = \sum_{R^s \in \mathcal{R}^s} \alpha_{R,R^s} \lambda_p(R^s)$ we have that $\lambda_p(R^s) = \lambda_p(R)$ for all R^s with $\alpha_{R,R^s} > 0$. Thus all replacements of R with rules in \mathcal{R}^s , which induce the same expected choice distribution, are such that all rules used in the replacement induce the exact same choice distribution. But then they all induce the same fitness. This is for example the case when the distribution p over choice sets puts probability only on singleton choice sets.

Second, even if, for every rule $R \in \mathcal{R} \setminus \mathcal{R}^s$ we can find a replacement that is composed of strictly rational rules inducing different choice distributions, we might still find a rule $R \in \mathcal{R}$ such that for all such replacements, i.e. for all α_{R,R^s} collections, summing to one, with $\lambda_p(R) = \sum_{R^s \in \mathcal{R}^s} \alpha_{R,R^s} \lambda_p(R^s)$ we have that $\lambda_p(R^s) \neq \lambda_p(R)$ for some R^s with $\alpha_{R,R^s} > 0$ and, yet, $U(R^s) = U(R)$ for all R^s with $\alpha_{R,R^s} > 0$. This is for example the case when the fitness function u is a constant function.

To guarantee that for every sampling distribution q with $q(\mathcal{R}^s) < 1$ there is a strict mean preserving spread q' with $q'(\mathcal{R}^s) = 1$ we thus need simultaneous conditions on the distribution over choice sets, p , as well as on the fitness function, u .

Fixing p , for a given rule R , a collection of weights $\alpha = \{\alpha_{R^s}\}_{R^s \in \mathcal{R}^s}$ is a **splitting of R** (or of $\lambda_p(R)$) if $\alpha_{R^s} \geq 0$ for all $R^s \in \mathcal{R}^s$, $\sum_{R^s} \alpha_{R^s} = 1$, and $\sum_{R^s} \alpha_{R^s} \lambda_p(R^s) = \lambda_p(R)$. Splittings form a convex set. If α, α' are splittings, so is any convex combination of α and α' . Hence, there exists a splitting

R with maximal support. Its support includes the support of any other splitting. Let this maximal support be denoted by $S(R)$ and, thus, given by $S(R) = \{R^s \in \mathcal{R}^s, \exists \text{ splitting } \alpha \text{ such that } \alpha_{R^s} > 0\}$. The maximal support, $S(R)$, of splittings of a rule R has a useful geometric characterization.

Lemma 9 *For every p and R , there exists a vector $v: K \rightarrow \mathbb{R}$ such that $S(R)$ is the set of maximizers of $\sum_k v(k)\lambda_p(R^s)(k)$ over $R^s \in \mathcal{R}^s$.*

Proof: Consider the minimal face F containing $\lambda_p(R)$ in the convex polyhedron whose vertices are the elements of Λ_p^s , i.e. in the convex hull of Λ_p^s . The maximal support for a splitting of $\lambda_p(R)$ is $F \cap \Lambda_p^s$. This support hence consists of the points of Λ_p^s that maximize some linear functional. QED

Note that for a rule R with a choice distribution $\lambda_p(R)$ that is in the strict interior of the convex hull of Λ_p^s the minimal face, in the proof of Lemma 9, is the whole convex polyhedron. Thus, the vector v must be constant in this case.

For any q, q' , let, abusing notation slightly, a collection of weights $\alpha = \{\alpha_{R,R^s}\}_{R,R^s}$ be a **splitting of q into q'** if each sub-collection $(\alpha_{R,R^s})_{R^s}$ is a splitting of each R and for every R^s , $q'(R^s) = \sum_R \alpha_{R,R^s}q(R)$. A splitting of q into q' has maximal support if each $(\alpha_{R,R^s})_{R^s}$ is a splitting of R with maximal support.

Given p and a fitness function u , the following characterizes splittings that induce strict mean preserving spreads in fitness. Recall that given a distribution q , \tilde{U}^q denotes the random variable $U(R)$, where $R \sim q$.

Lemma 10 *Let α be a splitting of q into q' . The distribution of $\tilde{U}^{q'}$ is a strict mean preserving spread of the distribution of \tilde{U}^q if and only if there exists R, R^s with $q(R)\alpha_{R,R^s} > 0$ and $U(R) \neq U(R^s)$.*

Proof: To prove the “if” part, suppose that there exists R, R^s with $q(R)\alpha_{R,R^s} > 0$ and $U(R) \neq U(R^s)$. Thus this R can be replaced by a strict mean preserving spread. Also all other rules R' such that $q(R') > 0$ can be replaced

by “weak” mean preserving spreads. Hence, $\tilde{U}^{q'}$ is a strict mean preserving spread of the distribution of \tilde{U}^q by integration over the support of q . To prove the “only if” part, note that $U(R) = U(R^s)$ whenever $q(R)\alpha_{R,R^s} > 0$ immediately implies that the distributions \tilde{U}^q and $\tilde{U}^{q'}$ are the same. QED

Lemma 11 *Assume that p has full support and u is discriminatory. Then for every $R \notin \mathcal{R}^s$ there exists $R^s \in S(R)$ such that $U(R) \neq U(R^s)$.*

Proof: Let v be as in Lemma 9. If v is injective, since p has full support, there is only one rule R' achieving the maximum of $\sum_k v(k)\lambda_p(R')(k)$, so that $S(R) = \{R\}$ and thus $R \in \mathcal{R}^s$, a contradiction. Assume then wlog that $v(1) \geq \dots \geq v(k) = v(k+1) \geq \dots \geq v(K)$ for some $k \geq 1$. Then $S(R)$ contains both rules R' and R'' corresponding to the preference orders $1 \succ \dots \succ k-1 \succ k \succ k+1 \succ k+2 \succ \dots \succ K$ and $1 \succ \dots \succ k-1 \succ k+1 \succ k \succ k+2 \succ \dots \succ K$ respectively. But then, $\lambda_p(R)(i) = \lambda_p(R')(i)$ for $i \neq k, k+1$. From this fact and the fact that $\sum_l \lambda_p(R)(l) = 1$ implying $\sum_l \lambda_p(R)(l) = \sum_l \lambda_p(R')(l)$ we obtain $\lambda_p(R)(k) + \lambda_p(R)(k+1) = \lambda_p(R')(k) + \lambda_p(R')(k+1)$, and, hence,

$$\lambda_p(R)(k) - \lambda_p(R')(k) = \lambda_p(R')(k+1) - \lambda_p(R)(k+1).$$

Note that $\lambda_p(R)(k) - \lambda_p(R')(k)$ is given by the sum of $p(L)$ over all $L \in \mathcal{L}$ with $k, k+1 \in L$ and $k = \min\{L\}$. Thus $\lambda_p(R)(k) - \lambda_p(R')(k) > 0$ since p has full support. This implies $U(R') \neq U(R'')$ since u , being discriminatory, satisfies $u(k) \neq u(k+1)$. QED

Lemma 12 *Assume that p has full support and u is non constant. For the zero rule R^0 , there exists $R^s \in S(R^0)$ such that $U(R^0) \neq U(R^s)$.*

Proof: Let v define $S(R^0)$ as in Lemma 9. Since R^0 maximizes $\sum_k \lambda_p(R)(k)v(k)$, and p has full support, v must be constant. Hence $S(R^0) = \Lambda_p^s$. Now, if $U(R^0) = U(R^s)$ for all R^s , u must be constant, a contradiction. QED

Lemma 13 *Assume p is neutral and has full support, q is symmetric with $q(\mathcal{R}^s) < 1$, and u is non-constant. Then there exists $R \in \mathcal{R}$ with $q(R) > 0$ and $R^s \in S(R)$ such that $U(R) \neq U(R^s)$.*

Proof: Let $R \notin \mathcal{R}^s$ such that $q(R) > 0$, and let v define $S(R)$ as in Lemma 9. Since $R \notin \mathcal{R}^s$ and p has full support, v must admit a tie. I.e. there exist i, j such that $v(i) = v(j)$. Supposing $U(R^s) = U(R)$ for all $R^s \in S(R)$ we deduce that $u(i) = u(j)$. Now, consider any permutation $\pi : K \rightarrow K$. By the symmetry of q we must have $q(R^\pi) = q(R) > 0$. By the neutrality of p , though, the set of convex combinations of strictly rational rules giving rise to $U(R^\pi)$ must be symmetric to the set of convex combinations, above, giving rise to $U(R)$. But then by the symmetric argument we must have $u(\pi(i)) = u(\pi(j))$. As the choice of π is arbitrary we thus obtain that u must be constant, a contradiction. QED

Proof of Lemma 5: Fixing p and q , let α be a splitting of q into some q' with maximal support. Assume i) u is discriminatory and p has full support and $q(\mathcal{R}^s) < 1$. For $R \notin \mathcal{R}^s$, by Lemma 11 there exists $R^s \in \mathcal{R}^s$ such that the pair R, R^s satisfies the conditions of Lemma 10, hence $\tilde{U}^{q'}$ is a strict mean preserving spread of \tilde{U}^q . Now if ii) p, q have full support and u is non-constant, by Lemma 12 there exists $R^s \in \mathcal{R}^s$ such that the pair R^0, R^s satisfies the conditions of Lemma 10, hence the result. QED

Proof of Lemma 6: Let p be neutral, q be symmetric, and α be a splitting of q with maximal support into some q' . The splitting constructed in the proof of Lemma 4 using all permutations of α has a support that includes that of α , hence it is also maximal. Thus, we have a splitting α' with maximal support of q into q^s , the uniform distribution over \mathcal{R}^s . Choosing $R \notin \mathcal{R}^s$ such that $q(R) > 0$, Lemma 13 shows there exists $R^s \in \mathcal{R}^s$ such that R, R^s satisfies the conditions of Lemma 10 for α' , hence the result. QED

References

- ALON, N., AND J. H. SPENCER (2000): *The Probabilistic Method*. Wiley-Interscience, John Wiley and Sons.
- ARROW, K. J. (1959): “Rational choice functions and orderings,” *Economica*, 26, 121–27.
- BANERJEE, A., AND J. W. WEIBULL (1995): “Evolutionary selection and rational behavior,” in *Learning and Rationality in Economics*, ed. by A. Kirman, and M. Salmon, pp. 343–63. Blackwell, Oxford, UK and Cambridge, USA.
- CAMPBELL, D. E. (1978): “Realization of choice functions,” *Econometrica*, 46, 171–80.
- CHEN, K. M. (2008): “Rationalization and Cognitive Dissonance: Do Choices Affect or Reflect Preferences?,” mimeo.
- DEKEL, E., J. C. ELY, AND O. YILANKAYA (2007): “Evolution of Preferences,” *Review of Economic Studies*, forthcoming.
- DEKEL, E., AND S. SCOTCHMER (1992): “On the Evolution of Optimizing Behavior,” *Journal of Economic Theory*, 57, 392–406.
- EGAN, L. C., L. R. SANTOS, AND P. BLOOM (2007): “The origins of cognitive dissonance: Evidence from children and monkeys,” *Psychological Science*, 18, 978–83.
- ELY, J. C. (2007): “Kludged,” mimeo.
- GÜTH, W. (1995): “An evolutionary approach to explaining cooperative behavior by reciprocal incentives,” *International Journal of Game Theory*, 24, 323–44.

- GÜTH, W., AND M. YAARI (1992): “Explaining reciprocal behavior in a simple strategic game,” in *Explaining Process and Change-Approaches to Evolutionary Economics*, pp. 23–24. University of Michigan Press.
- HOUTHAKKER, H. S. (1950): “Revealed preference and the utility function,” *Economica*, 17, 159–74.
- KALAI, G. (2003): “Learnability and rationality of choice,” *Journal of Economic Theory*, 113, 104–17.
- KANDORI, M., G. MAILATH, AND R. ROB (1993): “Learning, mutation, and long-run equilibria in games,” *Econometrica*, 61, 29–56.
- MAS-COLLEL, A., M. D. WHINSTON, AND J. R. GREEN (1995): *Microeconomic Theory*. Oxford University Press, Oxford, UK.
- MAYNARD SMITH, J., AND G. R. PRICE (1973): “The logic of animal conflict,” *Nature*, 246, 15–18.
- RICHERSON, P. J., R. BETTINGER, AND R. BOYD (2005): “Evolution on a restless planet: Were environmental variability and environmental change major drivers of human evolution?,” in *Handbook of Evolution*, ed. by F. M. Wuketits, and F. J. Ayala, vol. 2, pp. 223–42.
- ROBSON, A. J. (2001a): “Why would nature give individuals utility functions?,” *Journal of Political Economy*, 109, 900–14.
- (2001b): “The biological basis of economic behavior,” *Journal of Economic Literature*, 38, 11–33.
- RUBINSTEIN, A. (1996): “Why are certain properties of binary relations relatively more common in natural language,” *Econometrica*, 64, 343–56.
- SALANT, Y. (2007): “Procedural analysis of choice rules with applications to bounded rationality,” mimeo, GSB Stanford University.

- SAMUELSON, L., AND J. SWINKELS (2006): “Information, evolution and utility,” *Theoretical Economics*, 1, 119–42.
- SAMUELSON, P. (1938): “A note on the pure theory of consumer’s behaviour,” *Economica*, 5, 61–71, 353–54.
- TAYLOR, P., AND L. JONKER (1978): “Evolutionary stable strategies and game dynamics,” *Mathematical Biosciences*, 40, 145–56.
- UZAWA, H. (1956): “Note on preference and axioms of choice,” *Annals of the Institute of Statistical Mathematics*, 8, 35–40.
- VILLE, J. (1946): “Sur les conditions d’existence d’une ophélimité totale et d’un indice du niveau des prix,” *Annales de l’Université de Lyon*, 9, A(3), 32–39.
- WEIBULL, J. W. (1995): *Evolutionary Game Theory*. MIT Press, Cambridge, Mass.