
Variety Is the Spice of Life: Irrational Behavior as Adaptation to Stochastic Environments

Thomas J. Brennan

Harvard Law School, Cambridge, MA, USA
tbrennan@law.harvard.edu

Andrew W. Lo*

MIT Sloan School of Management
MIT Laboratory for Financial Engineering
MIT Computer Science and Artificial Intelligence Laboratory
Cambridge, MA 02142-1347, USA
alo-admin@mit.edu

Ruixun Zhang

MIT Laboratory for Financial Engineering
Cambridge, MA, USA
zhangruixun@gmail.com

Published 5 April 2018

The debate between rational models of behavior and their systematic deviations, often referred to as “irrational behavior”, has attracted an enormous amount of research. Here, we reconcile the debate by proposing an evolutionary explanation for irrational behavior. In the context of a simple binary choice model, we show that irrational behaviors are necessary for evolution in stochastic environments. Furthermore, there is an optimal degree of irrationality in the population depending on the degree of environmental randomness. In this process, mutation provides the important link between rational and irrational behaviors, and hence the variety in evolution. Our results yield widespread implications for financial markets, corporate behavior, and disciplines beyond finance.

: Behavioral bias; behavioral finance; adaptive markets hypothesis; evolutionary game theory.

JEL Classifications: D01, D91, E71, G12, G41

*Corresponding author.

1. Introduction

Rationality is the key assumption behind standard economic models of human behavior. The idea that individuals maximize their own self-interest subject to resource constraints has led to numerous breakthroughs including expected utility theory (von Neumann and Morgenstern, 1944), game theory (von Neumann and Morgenstern, 1944; Nash, 1950), rational expectations (Lucas Jr, 1972), the efficient markets hypothesis (Samuelson, 1965; Fama, 1970), and option pricing theory (Black and Scholes, 1973; Merton, 1973). The influence of this paradigm goes far beyond academia—it underlies current macroeconomic and monetary policies, and has become an integral part of the rules and regulations that govern financial markets today (Hu, 2012).

On the other hand, psychologists and economists have documented many violations of rational models in human behavior, often referred to as “cognitive biases”. These systematic deviations from rational behaviors are hard to reconcile with the standard economic models, and are therefore considered irrational behaviors. Representatives of these cognitive biases include probability matching, the tendency to choose randomly between heads and tails when asked to guess the outcomes of a series of independent biased-coin tosses, where the randomization matches the probability of the biased coin (Grant et al., 1951; Herrnstein, 1961); loss aversion, the tendency to take greater risk when choosing between two potential losses and less risk when choosing between two potential gains (Tversky and Kahneman, 1974; Tom et al., 2007); uncertainty effect, where a risky prospect is valued less than its worst possible outcome (Gneezy et al., 2006); and confirmation bias, the tendency to search for or interpret information in a way that confirms one’s preconceptions (Mahoney, 1977). Such anomalous behaviors have also been observed in many non-human subjects ranging from bacteria to primates (Harder and Real, 1987; Kirman, 1993; Smallwood, 1996; Chen et al., 2006; Ben-Jacob, 2008; Santos and Chen, 2009), which suggests that they may have a common and ancient origin, and an evolutionary role that belies their apparent shortcomings.

The debate between rational models of behavior and their systematic deviations has attracted an enormous amount of research in economics, psychology, and evolutionary biology (Becker, 1962; Stanovich and West, 2000; Rabin and Thaler, 2001; McKenzie, 2003; Burnham, 2013; Gneezy and List, 2013). For instance, bounded rationality (Simon, 1955) and prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992)

provide alternative perspectives for understanding human behavior beyond the maximization of expected utility. At the same time, numerous empirical studies are devoted to understanding the relationship between individual rationality and decision-making in the real world (Hsu

principle vividly with the example of a single ant traversing a mixed terrain of sand, rocks, and grass. The ant's path seems highly complex, but the complexity is due more to the environment than the ant's navigational algorithm.

Much of the rationality debate among economists and psychologists focuses on whether the rational models can help people make better inferences and decisions in the real world (McKenzie, 2003). Instead, our framework provides an evolutionary explanation for seemingly irrational behaviors and different degrees of irrationality in the population. The results have widespread implications for asset pricing and financial markets, corporate behavior, and disciplines beyond finance such as science, management, and public policy.

For example, it is well-known that irrational traders persist in financial markets (De Long et al., 1990, 1991; Biais and Shadur, 2000; Hirshleifer et al., 2006; Kogan et al., 2006) and that behaviors such as herding prevail, especially during crisis (Bowe and Domuta, 2004; Drehmann et al., 2005; Hirshleifer and Teoh, 2009). These behaviors can affect asset prices and create bubbles and crashes. From the corporate finance perspective, managers do

the entire population, and therefore, the important link between rational and irrational behaviors. Over time, only a certain degree of mutation and irrationality in the population will persist.

The roles of stochastic environments (Ishii [et al.](#), 1989; Kussell and Leibler, 2005; Acar [et al.](#), 2008; Gaal [et al.](#), 2010; Frank, 2011) and mutation (King, 1972; Taddei [et al.](#), 1997; Drake [et al.](#), 1998) have been extensively studied by evolutionary biologists. Several quantitative models have been developed to understand the magnitude of mutation rates (Kimura, 1960; Levins, 1967; Leigh Jr., 1970; Gillespie, 1981; Travis and Travis, 2002; Desai and Fisher, 2011; Liberman [et al.](#), 2011). While some of our results will be familiar to evolutionary biologists, they do not appear to be widely known in an economic context. For completeness, we derive them from first principles and provide the link between mutation and rationality.

By studying the impact of selection on behavior rather than on genes, we are able to derive evolutionary implications that cut across species, physiology, and genetic origins. In the same way that different magnifications of a microscope reveal different details of a specimen, applying evolutionary principles to behavioral variations leads to different insights that may be more relevant for economics, psychology, and behavioral ecology. Our focus on behavior as the object of selection is a different lens through which the effects of evolution may be studied.

In the remainder of this paper, we first describe the binary choice model with mutation. Then, we show that mutation and irrational behaviors are essential in evolution. Furthermore, the degree of irrationality is determined by evolution to match the degree of environmental stochasticity. We conclude with a brief discussion and provide additional technical details and proofs in [Appendix A](#).

2. Binary Choice Model with Mutation

We begin with the binary choice model (Brennan and Lo, 2011). Consider a population of individuals that live for one period, produce a random number of offspring asexually, and then die. During their lives, individuals make only one decision: they choose from two actions, 0 and 1 , and this results in one of two corresponding random numbers of offspring, o_0 and o_1 , described by some well-behaved probability distribution function, $\Phi(x, y)$. We assume that o_0 and o_1 are not perfectly correlated, and

Assumption 1. o_0 and o_1 are bounded non-negative random variables, and $\mathbb{P}(o_0 = o_1 = 0) = 0$.

Assumption 2. (θ_i, ω_i) is independent and identically distributed (IID) over time and identical for all individuals in a given generation.

Note that Assumption 1 simply rules out the degenerate case that no offspring is produced at all. Now suppose that each individual chooses θ with some probability $\theta \in [0, 1]$ and ω with probability $1 - \theta$, denoted by the Bernoulli variable θ_i , hence the offspring of an individual is given by

$$x_i = \theta_i x + (1 - \theta_i) y, \quad \theta_i = \begin{cases} 1 & \text{with prob } \theta \\ 0 & \text{with prob } 1 - \theta \end{cases}$$

We shall henceforth refer to θ as the individual’s behavior since it completely determines how the individual chooses between x and y . Assume for the moment that there is no mutation, so that offspring from a type θ individual are still of type θ . The following proposition (Brennan and Lo, 2011) summarizes the population dynamics of the binary choice model without mutation.

Proposition 1. θ^* is the unique solution to (1)–(3),

$$-\theta^* \log \theta^* - (1 - \theta^*) \log(1 - \theta^*) = \mu(\theta^*) = \mathbb{E}[\log(\theta + (1 - \theta) \omega)]. \tag{1}$$

$$\theta^* = \begin{cases} 1 & \text{if } \mathbb{E}[x/y] > 1 \text{ and } \mathbb{E}[x/\omega] < 1 \\ \text{solution to (3)} & \text{if } \mathbb{E}[x/y] \geq 1 \text{ and } \mathbb{E}[x/\omega] \geq 1 \\ 0 & \text{if } \mathbb{E}[x/y] < 1 \text{ and } \mathbb{E}[x/\omega] > 1, \end{cases} \tag{2}$$

$$\theta^* = \mathbb{E} \left[\frac{\theta^*}{\theta^* + (1 - \theta^*) \omega} \right] = \mathbb{E} \left[\frac{\theta^*}{\theta^* + (1 - \theta^*) \omega} \right], \tag{3}$$

$$(1)–(3) \quad \Phi(\theta, \omega).$$

The growth-optimal behavior, θ^* , is a function of the particular environment $\Phi(\theta, \omega)$. The role of Φ is critical in our framework, as it represents the entirety of the implications of an individual’s actions for reproductive success. Embedded in Φ is the biological machinery that is fundamental to evolution, i.e., genetics. However, this machinery is of less interest to economists than the link between behavior and reproductive success, which is summarized compactly by Φ . The specification of Φ also captures the fundamental distinction between traditional models of population genetics (Levins, 1968;

Wilson and Bossert, 1971; Dawkins, 1976) and more recent applications of evolution to behavior (Hamilton, 1964; Trivers, 1971; Wilson, 1975; Maynard Smith, 1982); the former focuses on the natural selection of traits (determined by genetics), whereas the latter focuses on the natural selection of behavior. Although behavior is obviously linked to genetics, the specific genes involved, their loci, and the mechanisms by which they are transmitted from one generation to the next are of less relevance to economic analysis than the ultimate implications of behavior for reproduction, i.e., Φ . In the jargon of econometrics, Φ may be viewed as a “reduced form” representation of an individual’s biology.

This simple and general model generates a remarkably rich set of behaviors (see Brennan and Lo (2011)). For example, the three possible behaviors in (2) are a generalization of the “adaptive coin-flipping” strategies of Cooper and Kaplan (1982). The behavior $*$ that emerges through the forces of natural selection is quite distinct from the neoclassical economic framework of expected utility in one important respect: expected utility theory implies deterministic behavior. Furthermore, intelligence has a natural definition in this framework—any type of behavior that is positively correlated with reproductive success—and bounds on the level of intelligence arise organically from physiological and environmental constraints on this correlation (Brennan and Lo, 2012). By considering different sources of randomness in reproductive success, risk aversion can be derived in this framework as a consequence of systematic reproductive risks (Zhang et al., 2014b). An extension of Proposition 1 may also be interpreted as a primitive form of group selection, in which natural selection appears to operate at the group level instead of, or in addition to, the level of individuals, traits, or genes (Zhang et al., 2014a).

In this stylized model of evolution, we are able to derive behaviors purely from evolution. We do not need any assumption on individual utility functions. In fact, the growth-optimal behavior in Proposition 1 does not always align with individually-optimal behavior if individuals maximize their expected number of offspring (Brennan and Lo, 2011). What is optimal from the evolutionary perspective, or what we call “rational behaviors”, depends on the environment. This is an important distinction from the neoclassical economic framework which assumes certain exogenous utility functions and derives behaviors given utility functions.

2.1. Mutation: A link between optimal and suboptimal behavior

Now, we add mutation to the binary choice model. In general, mutation implies that the offspring of type- individuals are not necessarily of type- ,

but assume a probability distribution over all possible types. We consider a simple form of mutation: an offspring of type- i mutates equally likely to all types.

To be more specific, consider a discrete type space. Let i take values in a finite set $\{1, 2, \dots, +1\}$ (for example, $\{0, \frac{1}{2}, \frac{2}{2}, \dots, \frac{-1}{2}, 1\}$) where $+$ is a positive integer. The world has $+$ + 1 types in total. In addition to Assumptions 1 and 2, we further assume that:

Assumption 3. Each type- i individual mutates with a small probability $\epsilon > 0$ to type $j \neq i$. Once it mutates, it mutates with equal probability $\frac{\epsilon}{+}$ to any type $j \in \{1, 2, \dots, +1\} \setminus \{i\}$.

Note that Assumption 3 is a simple and special form of mutation. From the behavioral point of view, it is general enough to capture the most important characteristics of mutation, which is to provide the link between different behaviors. With this particular structure, we are able to parametrize the degree of mutation with a single parameter ϵ .

We would like to emphasize that each individual lives for only one period in our model, and therefore its mutant offspring may be viewed as “new entrants” in the next generation’s population because they represent different behaviors than their predecessors. Also, there is no intelligence or volition ascribed to behavior i ; we are simply providing a formal representation for it, and then investigating its evolutionary implications. To that end, individuals choosing between i and j according to the same π may be viewed as consisting of the same “type”, where types are indexed by i and range continuously from 0 to 1, including the endpoints. In this manner, we are able to study the evolutionary dynamics of each type of individual over many generations.

Once mutation is introduced into the population, it is no longer possible to analyze the population dynamics of each type i separately. The entire system is a multi-type branching process in random environments (Smith and Wilkinson, 1969; Tanny, 1981). Let $\mathbf{n} = (n_1, \dots, n_{+1})'$ be the column vector of the number of individuals of all $+$ + 1 types in generation t . The following proposition describes the population dynamics between two generations.

Proposition 2. $\mathbf{n}_t = \mathbf{A} \cdot \mathbf{n}_{t-1}$ a.s. (4)

$$\mathbf{A} := \mathbf{M} \cdot \mathbf{F} \quad , \quad \mathbf{M} \quad :$$

$$\mathbf{M} = \begin{pmatrix} 1 - \epsilon & \epsilon & \cdots & \epsilon \\ \epsilon & 1 - \epsilon & \cdots & \epsilon \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon & \epsilon & \cdots & 1 - \epsilon \end{pmatrix},$$

$$\mathbf{F} \quad :$$

$$\mathbf{F} = \begin{pmatrix} 1 & , & + (1 - 1) & , & \cdots & & 0 \\ & \vdots & & \ddots & & & \vdots \\ & 0 & & \cdots & +1 & , & + (1 - +1) & , \end{pmatrix},$$

$$0 = 1 < 2 < \cdots < +1 = 1.$$

Equation (4) gives the fundamental relationship between individuals in two consecutive generations. With probability 1, \mathbf{n} can be written as the product of two matrices and \mathbf{n}_{-1} . \mathbf{F} represents the reproducibility of different types of individuals, and \mathbf{M} represents a re-distribution of types as a result of mutation. The natural question is: How does \mathbf{n} behave in the limit? We summarize the asymptotic behavior of a population with mutation in the following proposition.

Proposition 3 (Growth rate). 1-3,

$$\mu_\epsilon$$

$$\mu_\epsilon = \lim_{\rightarrow \infty} \frac{1}{\log \mathbf{c}' \mathbf{n}} = \lim_{\rightarrow \infty} \frac{1}{\log \|\mathbf{A} \mathbf{A}_{-1}\|},$$

proportion of the offspring from the behavior that grows the fastest. Therefore, the ratio of the individuals of any two behaviors can be lower bounded by some positive constant, and no single behavior can grow exponentially faster than any other behaviors. Note that μ_ϵ is called the ϵ -dominant eigenvalue of matrix \mathbf{A} in the probability literature, and Corollary 1 in the next sub-section gives an estimate of μ_ϵ .

Another difference between the mutation and non-mutation populations is the asymptotic ratio between different types of populations. Without mutation, $\mu(\cdot)$ is different for different \mathbf{A} , and therefore the ratio $\mu(\mathbf{A}_1) / \mu(\mathbf{A}_2)$ converges to zero for any $\mathbf{A}_1 \neq \mathbf{A}_2$ (see Proposition 1). However, μ_ϵ is the same with mutation for all types of \mathbf{A} , and the ratio $\mu_\epsilon(\mathbf{A}_1) / \mu_\epsilon(\mathbf{A}_2)$ is typically stochastic even in the long run as t increases without bound. We have ergodic theorems to characterize the asymptotic behavior of this ratio in the next sub-section.

2.2. Asymptotic population dynamics

Under Assumptions 1–3, let $\mathbf{n} = \mathbf{1}'\mathbf{n}$ be the total population size at time t and

$$\mathbf{y} := \frac{\mathbf{n}}{n} = \left(\frac{n_1}{n}, \dots, \frac{n_{+1}}{n} \right)' \tag{5}$$

be the normalized population vector in generation t . Because of the dynamics (4) between two consecutive generations, $\{\mathbf{y}\}_{t=0}^\infty$ is a vector-valued Markov process, with a compact state space

$$\mathcal{Y} := \left\{ \mathbf{y} = (n_1, \dots, n_{+1})' \mid \mathbf{y} \geq 0, \sum_{i=1}^{+1} y_i = 1 \right\}.$$

The one-step transition probability for $\mathbf{y} \in \mathcal{Y}$ and $\mathbf{z} \in \mathcal{Y}$ is:

$$P_1(\mathbf{y}, \mathbf{z}) := \mathbb{P}_\Phi \left(\frac{\mathbf{A}\mathbf{y}}{\|\mathbf{A}\mathbf{y}\|} \in \mathbf{z} \right). \tag{6}$$

Without mutation, because different behaviors grow at different exponential rates, \mathbf{y} converges almost surely to a basis vector $\mathbf{e} = (0, \dots, 1, \dots, 0)$ as $t \rightarrow \infty$. In the case of positive mutation rates, similar results exist only for non-random matrices \mathbf{F} in (4), in which case the long run proportion vector converges to the eigenvector of \mathbf{F} (see models in Robson (1996a) and Gaal et al. (2010) for examples). In the case of positive mutation rates when \mathbf{F} are random matrices, environmental uncertainty implies that \mathbf{y} is typically stochastic even in the long run (see simulation results in the Appendix).

However, the following ergodic theorem (Tuljapurkar, 1990) allows us to characterize the asymptotic behavior of \mathbf{y} :

Proposition 4 (Stochastic ergodic theorem). 1–3,

$$\mathcal{L}(\cdot) \quad \mathbf{y}, \quad \mathcal{L}(\cdot) \quad :$$

$$\mathcal{L}(\cdot) \quad \lim_{\rightarrow \infty} \mathcal{L} = \mathcal{L}.$$

Proposition 4 asserts that the proportion vector, \mathbf{y} , converges weakly as $\rightarrow \infty$. In addition, by basic properties of Markov chains, the stationary distribution, \mathcal{L} , satisfies the following equation:

$$\mathcal{L}(\cdot) = \int_{\mathcal{Y}} \mathbf{1}(\mathbf{y}, \cdot) \mathcal{L}(\mathbf{y})$$

for any $\subseteq \mathcal{Y}$. An important application of Proposition 4 is that it provides a formula to estimate the exponential growth rate μ_ϵ . Note that the total population size

$$= \mathbf{1}' \mathbf{n} = \mathbf{1}' \mathbf{M} \mathbf{F} \mathbf{n}_{-1} = \mathbf{1}' \mathbf{F} \mathbf{n}_{-1} = \mathbf{1}' \mathbf{F} \mathbf{y}_{-1},$$

so the log-geometric-average growth rate μ_ϵ can be expressed as

$$\mu_\epsilon = \mathbb{E}[\log(\mathbf{1}' \mathbf{F} \mathbf{y}_{-1})], \tag{7}$$

where the expectation is taken over the joint stationary distribution of $(\mathbf{F}, \mathbf{y}_{-1})$.

Corollary 1 (Bounds of growth rate). *

$$\left(\mathbf{1} \right) \tag{1} \tag{1–3},$$

$$* \in \{1, 2, \dots, +1\},$$

$$\mu(*) - |\log(1 - \epsilon)| \leq \mu_\epsilon \leq \mu(*) \tag{8}$$

Corollary 1 asserts that the growth rate μ_ϵ is slightly less than the optimal population growth rate without mutation. We will identify the cases where mutation does speed up growth in non-stationary environments in Section 3.

Appendix A gives additional results for population dynamics with mutation. In particular, we give the asymptotic distribution of total population size, $\mathbf{1}' \mathbf{n}$; the rate of convergence for the limit distribution, $\mathcal{L}(\cdot)$; and the optimal behavior with mutation in the probabilistic sense.

2.3. Extinction probability

When the population is extinct in evolution, stochastic processes \mathbf{n} and \mathbf{y} become degenerate. Therefore, all results so far are implicitly conditional on

non-extinction sample paths. However, extinction is important in evolution, and particularly of interest with mutation. In this sub-section, we investigate the extinction probability of different behaviors in different environments $\Phi(\mu, \sigma)$.

Consider a specific behavior $\mu \in \{\mu_1, \mu_2, \dots, \mu_{+1}\}$ starting with an initial population $x_0 > 0$, where the type μ is defined as $\mu = 0$ for some $x_0 > 0$, and $\mu > 0$ otherwise. In terms of extinction, there are two scenarios in which the number of generation increases without bound

- (i) $\lim_{x_0 \rightarrow \infty} \mathbb{P}(x_0 > 0) = 0$: the population is extinct with probability 1;
- (ii) $\lim_{x_0 \rightarrow \infty} \mathbb{P}(x_0 > 0) > 0$: the population survives with positive probability.

Note that in case (ii), if $\lim_{x_0 \rightarrow \infty} \mathbb{P}(x_0 > 0) < 1$, then the extinction probability depends on the initial population, x_0 . However, when x_0 is relatively large, the survival probability is close to 1. To be more specific, we define that the type μ is

So far we have considered stationary environments generating IID fecundities across time. In this case, mutation does not help increase the speed of population growth (Corollary 1). This brings us to the next topic, where non-stationary environments are considered and mutation can indeed speed up growth.

3. Optimal Degree of Irrationality

Corollary 2.

$$\epsilon^* = \frac{1}{\mathbb{E}[\tau_1 + \tau_2]}, \tag{9}$$

By Proposition 6 and Corollary 2, the optimal mutation rate is simply the reciprocal of the expected length of a regime. In the long run, the more stable the environment, the less irrational behaviors are present in the population; the more frequently environmental changes happen, the more irrational behaviors prevail in the population. The mutation rate and the amount of irrational behaviors are not exogenous variables given by Nature. They are not only necessary, but also important quantities that are selected by Nature in evolution to match the degree of environmental instability. In this sense, natural selection shapes the degree of irrationality in the population.

This also implies that the optimal amount of new entrants into the population is determined by the degree of environmental stability. For example, one would expect a relatively small number of new entrants in areas with relatively stable market conditions, such as the automobile industry; and relatively high turnover rates in areas with relatively volatile market conditions, such as the hedge fund industry.

3.2. Generalization and simulation experiments

The implications from the two-behavior example with a special fecundity structure above can be generalized to any number of types and any fecundity structures. We use simulation experiments to demonstrate the generality of the optimal degree of mutation and irrationality. In this section, we consider eight different environments, and derive the optimal degree of mutation for each.

In the following experiments, the lengths of regimes τ_1 and τ_2 are independent random variables with expectation $\mathbb{E}[\tau_1]$ and $\mathbb{E}[\tau_2]$, respectively, ranging from 10 to 37. For a given expectation $\mathbb{E}[\tau_1]$, τ_1 is uniformly distributed in the interval $[0.8 \times \mathbb{E}[\tau_1], 1.2 \times \mathbb{E}[\tau_1]]$, rounding to the nearest integer. τ_2 is distributed in the same way.

For a given pair $(\mathbb{E}[\tau_1], \mathbb{E}[\tau_2])$, 11 types of behavior from $\{0, \frac{1}{10}, \frac{2}{10}, \dots, 1\}$, starting with one individual, each evolve for 700 to 1,000 generations. The optimal degree of mutation in each pair of regimes is calculated by taking the average over 200 to 500 simulation paths.

Table 1. Probability table for the simulation of optimal mutation rates: Environments 1–8.

Environment 1						Environment 2							
Regime 1			Regime 2			Regime 1			Regime 2				
Prob.	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	Prob.	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
	3	2	1	0	0	0		3	2	1	1	1	1
	0	0	0	3	2	1		1	1	1	3	2	1
Environment 3						Environment 4							
Regime 1			Regime 2			Regime 1		Regime 2					
Prob.	0.8	0.2	0.8	0.2	Prob.	0.8	0.2	0.8	0.2				
	3	0	0	3		3	1	1	3				
	0	3	3	0		1	3	3	1				
Environment 5					Environment 6								
Regime 1			Regime 2		Regime 1			Regime 2					
Prob.	0.8	0.2	0.8	0.2	Prob.	0.8	0.2	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$			
	3	0	1	3		3	1	1	1	0			
	0	3	3	1		1	3	3	2	1			
Environment 7					Environment 8								
Regime 1			Regime 2		Regime 1			Regime 2					
Prob.	0.8	0.2	0.8	0.2	Prob.	0.8	0.2	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$			
	3	0	3	1		3	1	3	2	1			
	0	3	1	3		1	3	1	1	0			

Table 1 gives eight different environmental conditions, for which we plot the optimal degree of mutation and the optimal log-geometric-average growth rate as a function of $\mathbb{E}[x^1]$ and $\mathbb{E}[x^2]$ in Fig. 1. In these figures, the colored plane shows the optimal mutation rates; the transparent surface, for which the height is indicated by the z -axis, shows the optimal log-geometric-average growth rate associated with that optimal mutation rate.

Symmetric regimes. Environment 1 assumes that one of the actions in each regime leads to no offspring. Results are consistent with the example of two behaviors: the optimal degree of mutation is inversely proportional to $\mathbb{E}[x^1] + \mathbb{E}[x^2]$. However, the growth rate is proportional to $\mathbb{E}[x^1]$ and $\mathbb{E}[x^2]$: the longer the length of a regime, the faster the population grows.

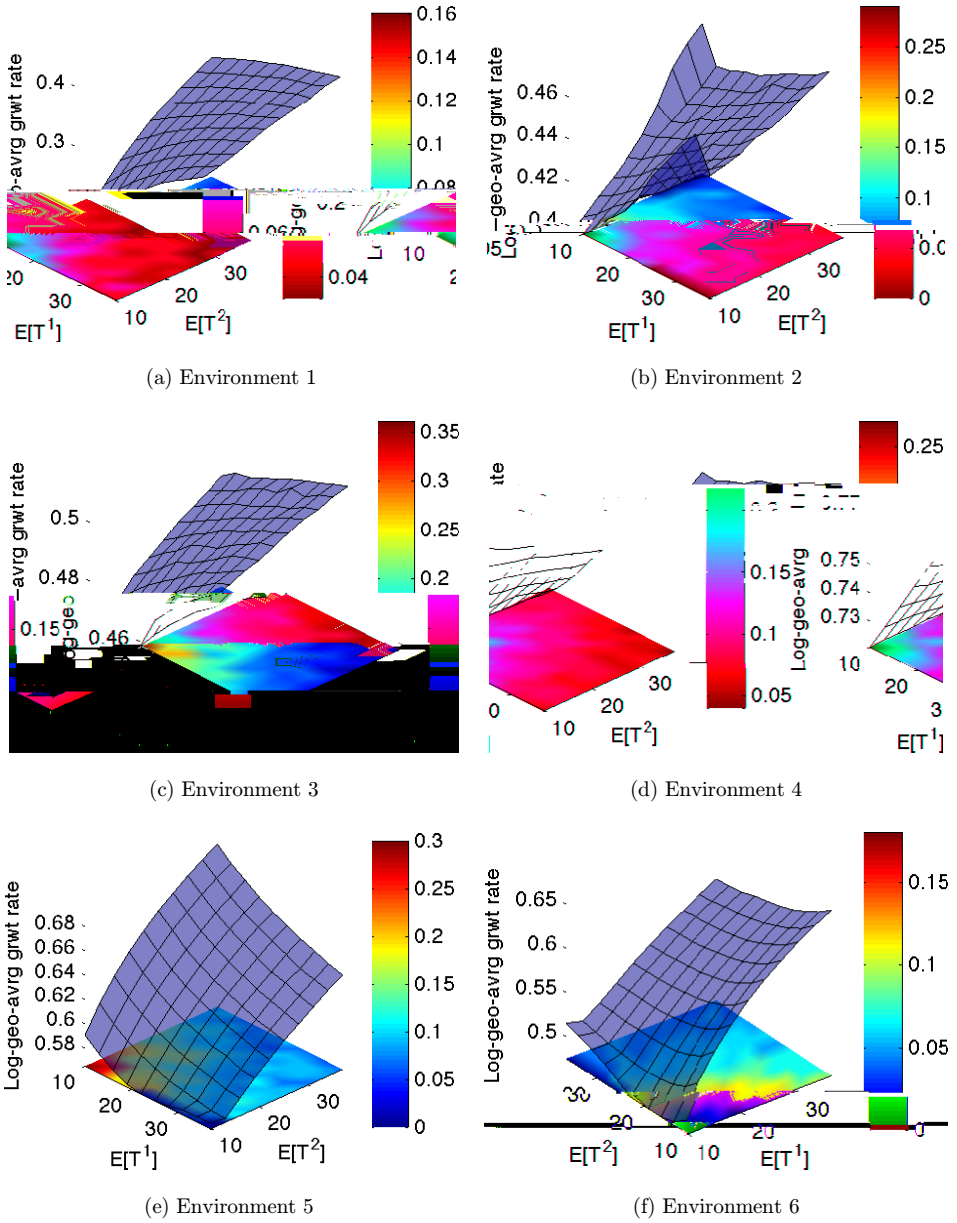


Fig. 1. Optimal degree of mutation and optimal log-geometric-average growth rate as a function of regime lengths $\mathbb{E}[T^1]$ and $\mathbb{E}[T^2]$. The subfigures summarize the simulation results of the environments in Table 1. The colored plane with the colorbar shows the optimal mutation rates; the transparent surface, for which the height is indicated by the z -axis, shows the optimal log-geometric-average growth rate associated with that optimal mutation rate.

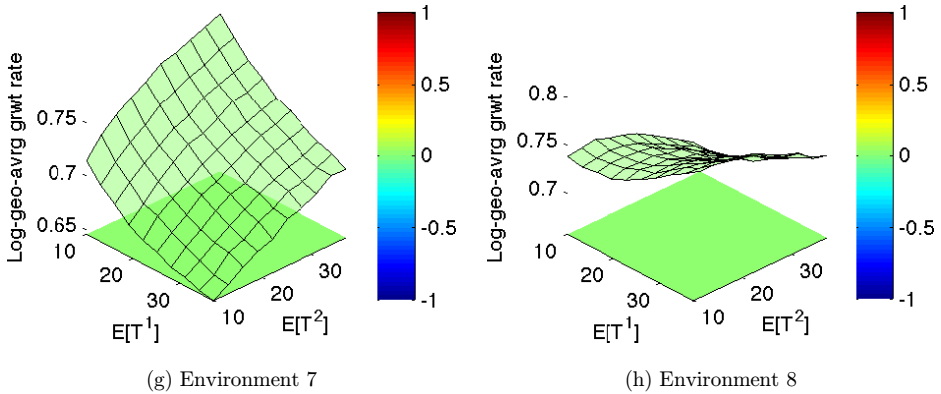


Fig. 1. ()

Environment 2 considers the case where actions μ and ν both produce a positive number of offspring. As expected, the growth rates are much higher than those in environment 1. The optimal degree of mutation is inversely proportional to the length of a regime, except for two regions where the length of one regime is much larger than that of the other (the region $\mathbb{E}[T^1] > 25, \mathbb{E}[T^2] < 12$, and the region $\mathbb{E}[T^1] < 12, \mathbb{E}[T^2] > 25$). In these two regions, the optimal degree of mutation drops to nearly zero because one regime is significantly shorter than the other and therefore it is not worth sacrificing growth in one regime for the other by mutation.

Environments 3 and 4 add dependency between μ and ν in each regime. In those two cases, simulation results are similar to environment 1.

Asymmetric regimes. The four experiments considered so far are symmetric in terms of the two regimes. In other words, the second regime is simply a copy of the first regime with μ and ν reversed. As a consequence, all results are expected to be symmetric with respect to the line $\mathbb{E}[T^1] = \mathbb{E}[T^2]$. In this part we consider asymmetric regimes and investigate how this changes the optimal mutation rates and growth rates.

Environment 5 is a mixture of environments 3 and 4: regime 1 is from environment 3 and regime 2 is from environment 4. In this case the optimal behavior is $\mu_1^* = 0.8$ in regime 1 and $\mu_2^* = 0$ in regime 2. There are several interesting observations. First of all, both the optimal degree of mutation and the growth rate are no longer symmetric with respect to $\mathbb{E}[T^1]$ and $\mathbb{E}[T^2]$. Secondly, the growth rate increases as $\mathbb{E}[T^2]$ increases and decreases as $\mathbb{E}[T^1]$ increases. This is because regime 2 has a larger geometric-mean fitness than regime 1, and the growth rate increases as the proportion of generations in

regime 2 increases. Thirdly, a phenomenon similar to that in environment 2 with zero mutation appears when $\mathbb{E}[r_1]$ is large and $\mathbb{E}[r_2]$ is small.

Environment 6 makes the two regimes more asymmetric. The optimal behavior is $\mu_1^* = 1$ in regime 1 and $\mu_2^* = 0$ in regime 2. These results are similar to those of environment 5.

When mutation is not desirable. Mutation is desirable because the environment is non-stationary and the two regimes favor different actions. When these conditions change, mutation is no longer desirable.

Environment 7 reverses actions μ_1 and μ_2 in the second regime of environment 5. The shape of the transparent surface indicating growth rates is similar to that of environment 5. However, the optimal degree of mutation is zero for any combination of $\mathbb{E}[r_1]$ and $\mathbb{E}[r_2]$ because the optimal behavior is $\mu_1^* = 0.8$ in regime 1 and $\mu_2^* = 1$ in regime 2. They are close to each other, and both of them grow relatively fast in both regimes.

Environment 8 reverses actions μ_1 and μ_2 in the second regime of environment 6. The optimal behavior is $\mu^* = 1$ in both regimes and therefore the optimal mutation rate is 0.

3.3. Optimal degree of irrationality

It is clear that there exists a balance between growth without mutation and robustness with mutation. The simulation results confirm the inverse relation between the optimal degree of mutation and expected lengths of regimes derived analytically in the simple two-behavior model with special fecundity structure (Proposition 6 and Corollary 2). The relation is robust across a variety of environmental conditions.

For symmetric regimes, the optimal degree of mutation is inversely proportional to $\mathbb{E}[r_1] + \mathbb{E}[r_2]$; the growth rate is proportional to both $\mathbb{E}[r_1]$ and $\mathbb{E}[r_2]$. For asymmetric regimes, the growth rate increases as the proportion of the regime that has a larger geometric-mean fitness increases. The relative magnitude of the two regimes matters.

The optimal degree of mutation could be zero if one regime is significantly shorter than the other because it is not worth sacrificing growth in one regime for the other as long as the inferior behavior does not die out in the shorter regime. The optimal degree of mutation could also be zero if the optimal behaviors in two regimes are similar to each other, and both of them grow relatively fast in both regimes.

The length of regime, or equivalently the frequency of change, is one aspect of the nature of environmental change. The intensity of each environmental

also produce more sophisticated behaviors such as overconfidence (Johnson and Fowler, 2011), altruism and self-deception (Trivers, 1971; Becker, 1976), and state-dependent strategies like the Hawk–Dove game (Maynard Smith, 1984), which emerge as a result of more complex environmental conditions. In our framework, if we assume that one individual's action is correlated with the reproductive success of another individual, individuals engaging in strategic behavior will reproduce more quickly than those with simpler behaviors such as probability matching. If the actions of individuals in the current generation can affect the reproductive success of individuals in future generations, even more complex dynamics are likely to emerge as in the well-known overlapping generations model (Samuelson, 1958). In a resource-constrained environment in which one individual's choice can affect another individual's reproductive success, strategic interactions such as reciprocity and cooperation will likely emerge within and across generations (Trivers, 1971; Nowak and Highfield, 2011).

We have modeled mutation in a simple way in this article. There may be other more complicated forms of mutation one can introduce to the evolutionary framework, including mutation rates that are correlated with the environment. This would correspond to individual intelligence and arise when individuals have memory and therefore are able to adapt to the environment given what has happened in the past.

Much of the rationality debate among economists and psychologists focuses on whether the rational models can help people make better inferences and decisions in the real world (McKenzie, 2003). Instead, our framework provides an evolutionary explanation of irrational behaviors and different degrees of irrationality in the population. The results suggest that irrational behaviors are necessary even if they are seemingly inefficient in the current environment, and the nature of the stochastic environment determines the degree of irrationality and the amount of new entrants into the population.

From an application perspective, our results underscore the importance of addressing different human behaviors in different environments. For example, the financial market is considered to be efficient most of the time (Samuelson, 1965; Fama, 1970), and participants with irrational beliefs constitute a minimal part in the market. However, in periods of economic turbulence and financial crisis, irrational behaviors are much more prevalent than usual. Irrational traders persist and behaviors such as herding prevail. These behaviors can affect asset prices as well as create bubbles and crashes.

From the corporate finance perspective, managers do not always form beliefs logically, nor do these beliefs convert to decisions in a consistent and

rational manner. Both the economic and regulatory environments can affect managers' behaviors. Our model suggests that these behaviors are not necessarily "irrational"; they are simply the result of adaptation. In fact, behaviors normally regarded as "irrational" such as overconfidence might even be beneficial in certain market environments. From this perspective, a stable environment would help reduce the amount of irrational behavior in the population, and yield higher economic growth.

Our results also highlight the importance of the entry of new actors into the market even if they appear suboptimal in the current context, and suggest that the optimal amount of new entrants depends on the degree of environmental stability. On the other hand, if not properly managed, volatile environments can lead to increases in the degree of irrationality, implying higher social costs and lower economic growth.

Finally, our results also highlight the potential dangers of sustained government intervention, which can become a source of systematic risk and cause volatile environments in its own right (Acharya et al., 2011; Lucas, 2011).

Acknowledgments

We are grateful for helpful comments and suggestions from the editor, Jean Helwege. We also benefited from discussions with Arthur J. Robson and other participants at the Allied Social Sciences Association annual meeting in Philadelphia, 2018. We thank Jayna Cummings for editorial assistance. Research support from the MIT Laboratory for Financial Engineering is gratefully acknowledged.

Appendix A

In Appendix A, we provide additional technical details and proofs for the main results of the paper.

A.1. Birkhoff's contraction coefficient

The definition and properties of the Birkhoff's contraction coefficient can be found in Caswell (2001, pp. 370–372) or Ipsen and Selee (2011, p. 159). Let \mathbf{x} and \mathbf{y} be positive vectors. The Hilbert pseudo-metric distance between \mathbf{x} and \mathbf{y} is defined as:

$$d(\mathbf{x}, \mathbf{y}) := \log \left(\frac{\max_i \frac{x_i}{y_i}}{\min_i \frac{x_i}{y_i}} \right) = \max_i \log \left(\frac{x_i}{y_i} \right).$$

It measures the distance between two vectors in a way that depends only on their proportional composition, independent of their absolute size. It satisfies the following conditions:

$$\begin{aligned} (\mathbf{x}, \mathbf{y}) &\geq 0, \\ (\mathbf{x}, \mathbf{y}) &= (\mathbf{y}, \mathbf{x}), \\ (\mathbf{x}, \mathbf{y}) &\leq (\mathbf{x}, \mathbf{z}) + (\mathbf{z}, \mathbf{y}), \\ (\mathbf{x}, \mathbf{y}) &= 0 \text{ if and only if } \mathbf{x} = \mathbf{y}, \\ (\mathbf{x}, \mathbf{y}) &= (\mathbf{x}, \mathbf{y}) \text{ for } \mathbf{x}, \mathbf{y} > 0. \end{aligned}$$

The Birkhoff's contraction coefficient of a non-negative matrix \mathbf{A} is defined as

$$\tau(\mathbf{A}) = \sup \frac{(\mathbf{A}\mathbf{x}, \mathbf{A}\mathbf{y})}{(\mathbf{x}, \mathbf{y})}, \tag{A.1}$$

where the supremum is taken over all vectors $\mathbf{x} > 0$ and $\mathbf{y} > 0$ that are not multiples of each other. Note that because $\tau(\mathbf{A})$ is invariant with respect to the absolute magnitude of vectors, the supremum can be taken over a compact subset equivalently, say $\|\mathbf{x}\|_1 = \|\mathbf{y}\|_1 = 1$.

If \mathbf{A} is a strictly positive matrix, then $\tau(\mathbf{A}) < 1$ (see Caswell (2001, p. 372) for example). Under Assumptions 1-3, the matrix \mathbf{A} might not be strictly positive. However, there is at least one positive entry in each row of \mathbf{A} , so $\tau(\mathbf{A}) \leq 1$ (see Hajnal (1976) for a discussion on "row allowable" matrices). We will prove in Lemma 1 that $\tau(\mathbf{A}_t)$ is indeed strictly less than 1.

Lemma 1 (Contraction properties of \mathbf{A}_t). 1-3,

$$\begin{aligned} \tau(\mathbf{A}_t) &< 1 \\ \mathbb{P}(\tau(\mathbf{A}_t) < 1) &= 1. \end{aligned}$$

Because of Assumption 1, there are only finitely many possible random matrices \mathbf{A}_t if τ and ϵ are integers. Therefore, the Birkhoff's contraction coefficient $\tau(\mathbf{A}_t)$ is uniformly less than some positive constant $\delta < 1$. But Lemma 1 is enough for the analysis henceforth.

A.2. Additional results for population dynamics

Lemma 2 (Decomposition of population vector).

$$\mathbf{n}_0 = \mathbf{1} + \epsilon \mathbf{1} + \dots + \epsilon^{t-1} \mathbf{1} + \epsilon^t \mathbf{n}_t$$

$$\begin{aligned} \mathbf{n} &= \tilde{\mathbf{F}}_t \tilde{\mathbf{F}}_{t-1} \dots \tilde{\mathbf{F}}_1 \mathbf{n}_0 \\ &+ \frac{\epsilon}{1-\delta} (\tilde{\mathbf{F}}_t \dots \tilde{\mathbf{F}}_2 \mathbf{1} + \tilde{\mathbf{F}}_t \dots \tilde{\mathbf{F}}_3 \mathbf{1} + \dots + \tilde{\mathbf{F}}_t \mathbf{1} + \mathbf{1}) \end{aligned}$$

$$\tilde{\mathbf{F}} = (1 - (1 + \frac{1}{\sigma})\epsilon)\mathbf{F}$$

Lemma 2 provides a decomposition of the population vector \mathbf{n} into a linear combination of a non-mutation vector and σ vectors for shorter generational spans, each of which is weighted by ϵ and stands for evolution starting from a certain time in the past.

Proposition 7 (Asymptotic population distribution).

$$1-3, \quad \sigma \quad = \mathbf{1}'\mathbf{n}$$

$$\frac{\log \frac{\mathbf{n} - \mu_\epsilon}{\sigma\sqrt{\cdot}}}{\sigma\sqrt{\cdot}} \Rightarrow \text{Normal}(0, 1)$$

$$\rightarrow \infty.$$

By Proposition 7, the asymptotic distribution of total population is log-normal, and the mean and variance of $\log \mathbf{n}$ both increase linearly with time.

Proposition 8 (Rate of convergence).

$$1-3, \quad \mathcal{L}(\cdot)$$

$$\{ \mathbf{y} \}_{1}^{\infty} \quad \mathcal{L}$$

$$\rho < 1 \quad < \infty, \quad \| \cdot \|$$

$$\| \mathcal{L}(\cdot) - \mathcal{L}(\cdot) \| \leq \rho, \quad = 1, 2, 3, \dots$$

Proposition 8 asserts that the rate of convergence in Proposition 4 is exponential. Therefore, one would expect that the convergence of $\mathcal{L}(\cdot)$ to the stationary distribution $\mathcal{L}(\cdot)$ is very fast on an evolutionary time-scale.

Proposition 9 (Selection of the optimal behavior (Robson, 1996a)).

$$1). \quad 1-3, \quad \mu(\cdot) > 0$$

$$\bar{\delta} > 0, \quad \mathbf{y}^* \in (0, 1), \quad \bar{\epsilon} > 0, \quad \epsilon \in (0, \bar{\epsilon}),$$

$$\mathbb{P}_{\mathcal{L}}(\cdot \geq 1 - \bar{\delta}) > 1 - \bar{\epsilon}$$

$$\mathcal{L}.$$

¹The support of $\mathcal{L}(\cdot)$ is defined to be the set of all points $\mathbf{y} \in \mathcal{Y}$ for which every open neighborhood of \mathbf{y} has positive measure.

Proposition 9 asserts an important property of \mathcal{L} : \ast without mutation again dominates the population with mutation in evolution with arbitrarily high probability, provided that the mutation rate is small enough. However, explicit calculation of the stationary distribution is difficult. Section A.3 discusses a simulation experiment to understand the limit stationary distribution \mathcal{L} .

A.3. Simulation for the limit distribution of population proportions

Let’s consider an example to show how the limit distribution of population proportions behaves. With $\Phi(\cdot, \cdot)$ given in Table A.1, we study a system with six behaviors $\in \{0, \frac{1}{5}, \frac{2}{5}, \dots, 1\}$.

Figures A.1(a)–A.1(c) show the proportion of each behavior in the entire population as the number of generation increases in one simulation. Without mutation, the proportion of different behaviors converges almost surely. With positive mutation rates, the population proportion vector is stochastic even in the long run.

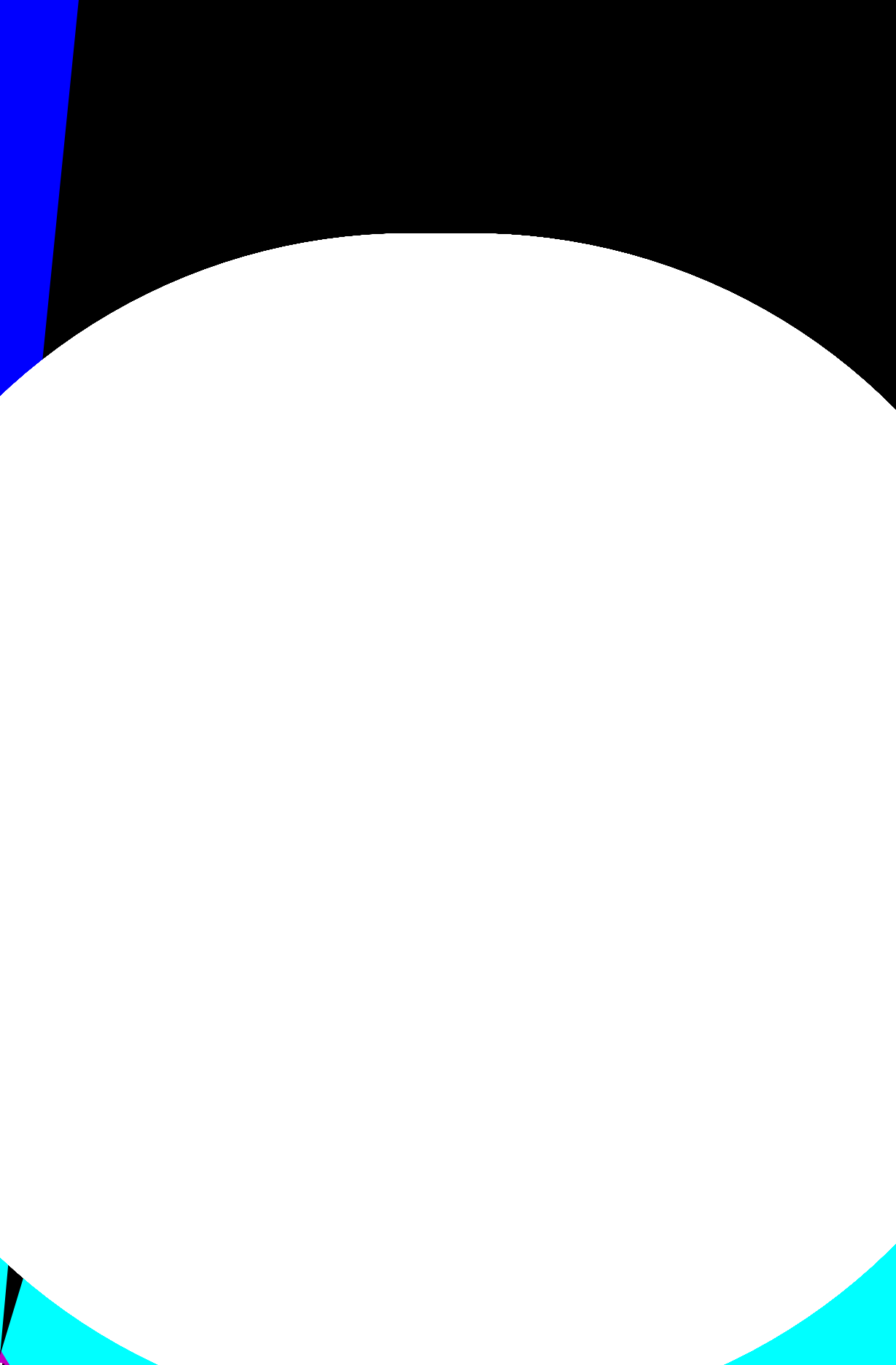
Figures A.1(d) and A.1(e) show the limit distribution of population proportions for mutation rates $\epsilon = 0.01$ and 0.05 . Each subplot shows the histogram of three behaviors in the last generation $= 500$ with 1,000 simulation paths: the optimal behavior $\ast = 0.8$, and two suboptimal behaviors $= 0.6$, $= 1$. We only plot three representative behaviors for simplicity. From the histogram, it is clear that $\ast = 0.8$ corresponds to the optimal behavior. As the mutation rate gets smaller, the probability that $\ast = 0.8$ dominates the entire population gets closer to 1.

Furthermore, the final stationary distribution does not behave like normal because of the heavy tailness observed in the simulation. In particular, a Kolmogorov–Smirnov test of normality on the distribution of $\ast = 0.8$ proportion gives $-value = 1.08 \times 10^{-50}$, rejecting the normality hypothesis.

Finally, although it is hard to show the evolution of the complete distribution of normalized population vector (5) in one figure, Figures A.1(f) and

Table A.1. Probability table for the simulation of asymptotic population dynamics.

	State 1	State 2
Action	prob. $= 0.8$	prob. $1 - = 0.2$
	$= 3$	$= 0$
	$= 0$	$= 3$



A.1(g) show the proportion of each behavior in the entire population averaged over 1,000 simulation paths. It is clear that the expectation of the distribution converges quickly, and the optimal behavior without mutation again dominates the population in expectation. This supports our results in Propositions 8 and 9.

A.4. Proofs

Lemma 1. (2001)

$$\begin{aligned}
 & \text{371–372. } \mathbf{A} = (\quad)_{(\ +1) \times (\ +1)} \\
 & \text{1–3. } \mathbf{A} \quad , \quad \tau(\mathbf{A}) < 1. \quad \mathbf{A} \\
 & \quad , \quad \mathbf{P}(\ > 0 \text{ or } \ > 0) = 1, \quad \mathbf{A} \\
 & \quad \quad \quad (\ +1) \quad (\quad) \\
 & \quad \quad \quad \mathbf{A} \quad 0 \\
 & \quad , \quad \tau(\mathbf{A}) < 1 \quad .
 \end{aligned}$$

Now let $\mathbf{x}(\) = (\ (\ +1))_{=1}^{+1}$ and $\mathbf{y}(\) = (\ (\ +1))_{=1}^{+1}$ be positive vectors that are not proportional to each other, and $\mathbf{x}(\ +1) = \mathbf{A} \cdot \mathbf{x}(\)$ and $\mathbf{y}(\ +1) = \mathbf{A} \cdot \mathbf{y}(\)$. Then

$$\frac{(\ +1)}{(\ +1)} = \frac{\sum (\)}{\sum (\)} = \sum \left(\frac{(\)}{\sum (\)} \right) \frac{(\)}{(\)} = \sum \frac{(\)}{(\)},$$

where $\sum = 1$. A careful examination of yields that for any ,

$$\quad = 0, \quad \text{and} \quad > 0 \quad \text{for} \quad = 2, 3, \dots, \ +1.$$

Therefore, $\frac{(\ +1)}{(\ +1)}$ is a positive weighted average of $\left\{ \frac{(\)}{(\)} \right\}_{=2}^{+1}$, and this is true for all . Because $\mathbf{x}(\)$ and $\mathbf{y}(\)$ are not proportional to each other, there are two possibilities:

- (1) The ratios in $\left\{ \frac{(\)}{(\)} \right\}_{=2}^{+1}$ are all the same, but different from $\frac{1}{1(\)}$. In this case exactly one of the following must be true:

$$\min \frac{(\)}{(\)} < \frac{(\ +1)}{(\ +1)} \leq \max \frac{(\)}{(\)}, \quad \text{for all } ,$$

or

$$\min \frac{(\)}{(\)} \leq \frac{(\ +1)}{(\ +1)} < \max \frac{(\)}{(\)}, \quad \text{for all } .$$

(2) The ratios in $\left\{\frac{(\cdot)}{(\cdot)}\right\}_{=2}^{+1}$ are not all the same. In this case, we have

$$\min \frac{(\cdot)}{(\cdot)} < \frac{(\cdot + 1)}{(\cdot + 1)} < \max \frac{(\cdot)}{(\cdot)}, \quad \text{for all } \cdot.$$

In both (1) and (2), we have

$$(\mathbf{x}(\cdot + 1), \mathbf{y}(\cdot + 1)) < (\mathbf{x}(\cdot), \mathbf{y}(\cdot)).$$

That is, each multiplication by \mathbf{A} contracts the distance between the two vectors. Because the supremum in Birkhoff's contraction coefficient (A.1) can be taken over a compact set, we have $\tau(\mathbf{A}) < 1$ with probability 1.

Lemma 2.

\mathbf{M}

$$\mathbf{M} = \left(1 - \left(1 + \frac{1}{\cdot}\right)\epsilon\right)\mathbf{I}_{\cdot+1} + \frac{\epsilon}{\cdot}\mathbf{1}_{\cdot+1}\mathbf{1}'_{\cdot+1},$$

$$\mathbf{I}_{\cdot+1} \quad (\cdot + 1) \times (\cdot + 1).$$

$$\mathbf{n} \stackrel{\text{a.s.}}{=} \mathbf{M}\mathbf{F}\mathbf{n}_{-1} = \left(1 - \left(1 + \frac{1}{\cdot}\right)\epsilon\right)\mathbf{F}\mathbf{n}_{-1} + \frac{\epsilon}{\cdot}\mathbf{1}_{\cdot+1}. \quad (\text{A.2})$$

$$\mathbf{1}'\mathbf{F}\mathbf{n}_{-1} = \mathbf{1}'\mathbf{M}\mathbf{F}\mathbf{n}_{-1} = \quad (\text{A.2}). \quad (\text{A.2}),$$

Proposition 1.

(2011) 1.

$$(2011) \quad 1).$$

Proposition 2.

(2011).

(2011), 1

“ ” → :

$$= \begin{cases} 1 & \text{with prob } 1 - \epsilon \\ 0 & \text{with prob } \epsilon, \end{cases} \rightarrow = \begin{cases} 1 & \text{with prob } \frac{\epsilon}{\cdot} \\ 0 & \text{with prob } 1 - \frac{\epsilon}{\cdot}. \end{cases}$$

In generation \cdot , type \cdot individuals come from type \cdot individuals without mutation and type $(\neq \cdot)$ individuals with mutation in generation $\cdot - 1$.

Consider them separately. From type :

$$\sum_{=1}^{-1} \rightarrow = \left(\sum_{=1}^{-1} \rightarrow, \rightarrow \right) + \left(\sum_{=1}^{-1} \rightarrow, (1 - \rightarrow) \right),$$

$$\stackrel{\text{a.s.}}{=} (1 - \epsilon) \rightarrow_{-1}(\rightarrow, + (1 - \rightarrow) \rightarrow),$$

as \rightarrow_{-1} increases without bound. From type (\neq)

$$\sum_{\neq} \sum_{=1}^{-1} \rightarrow = \sum_{\neq} \left[\left(\sum_{=1}^{-1} \rightarrow, \rightarrow \right) + \left(\sum_{=1}^{-1} \rightarrow, \rightarrow (1 - \rightarrow) \right) \right],$$

$$\stackrel{\text{a.s.}}{=} \frac{\epsilon}{\neq} \sum_{\neq} \rightarrow_{-1}(\rightarrow, + (1 - \rightarrow) \rightarrow),$$

as \rightarrow_{-1} increases without bound. Note that

$$= \sum_{=1}^{-1} \rightarrow + \sum_{\neq} \sum_{=1}^{-1} \rightarrow,$$

$$\stackrel{\text{a.s.}}{=} (1 - \epsilon) \rightarrow_{-1}(\rightarrow, + (1 - \rightarrow) \rightarrow) + \frac{\epsilon}{\neq} \sum_{\neq} \rightarrow_{-1}(\rightarrow, + (1 - \rightarrow) \rightarrow).$$

(4) simply rewrites the above equation in matrix form.

Proposition 3.

$$\log_+ \|\mathbf{A}_1\| = \max\{0, \log \|\mathbf{A}_1\|\}. \tag{4.2.1}$$

(1990), . 26 ().

Proposition 4.

$$\log_+ \|\mathbf{A}_1\| = \max\{0, \log \|\mathbf{A}_1\|\}. \tag{4.2.1, 4.2.3, 4.2.6}$$

(1990), . 29 ().

Proposition 5. ()

Proposition 6.

Φ^1

Φ^2

, Φ^1 ,

Φ^2

.

A.2

Table A.2. Population dynamics for the two-behavior regime-switching model.

Cycle	Regime	Generation	# of individuals = 0	# of individuals = 1
		0	1	1
1	Φ^1	\vdots	\vdots	\vdots
		$\frac{1}{1}$	$\frac{1}{1} \prod_{=1}^1 1$	$\epsilon \cdot \frac{1}{1}$
		$\frac{1}{1} + 1$	$\epsilon^2 \frac{1}{1-1} \prod_{=1}^1 1 \cdot \frac{2}{1}$	$\epsilon \frac{1}{1-1} \prod_{=1}^1 1$
1	Φ^2	\vdots	\vdots	\vdots
		$\frac{1}{1} + \frac{2}{1}$	$\epsilon^2 \frac{1}{1+2-2} \prod_{=1}^1 1 \cdot \prod_{=1}^2 2$	$\epsilon \frac{1}{1+2-1} \prod_{=1}^1 1 \cdot \prod_{=1}^2 2$
		$\frac{1}{1} + \frac{2}{1} + 1$	$\epsilon^2 \frac{1}{1+2-1} \prod_{=1}^1 1 \cdot \prod_{=1}^2 2 \cdot \frac{1}{1+1}$	$\epsilon^3 \frac{1}{1+2-2} \prod_{=1}^1 1 \cdot \prod_{=1}^2 2 \cdot \frac{1}{1+1}$
2	Φ^1	\vdots	\vdots	\vdots
		$\frac{1}{1} + \frac{2}{1} + \frac{1}{2}$	$\epsilon^2 \frac{1}{1+2+1-2} \prod_{=1}^1 1 \cdot \prod_{=1}^2 2$	$\epsilon^3 \frac{1}{1+2+1-3} \prod_{=1}^1 1 \cdot \prod_{=1}^2 2$
		$\frac{1}{1} + \frac{2}{1} + \frac{1}{2} + 1$	$\epsilon^4 \frac{1}{1+2+1-3} \prod_{=1}^1 1 \cdot \prod_{=1}^2 2 \cdot \frac{2}{1+1}$	$\epsilon^3 \frac{1}{1+2+1-2} \prod_{=1}^1 1 \cdot \prod_{=1}^2 2 \cdot \frac{2}{1+1}$
2	Φ^2	\vdots	\vdots	\vdots
		$\frac{1}{1} + \frac{2}{1} + \frac{1}{2} + \frac{2}{2}$	$\epsilon^4 \frac{1}{1+2+1+2-4} \prod_{=1}^1 1 \cdot \prod_{=1}^2 2$	$\epsilon^3 \frac{1}{1+2+1+2-3} \prod_{=1}^1 1 \cdot \prod_{=1}^2 2$
\vdots	\vdots	\vdots	\vdots	\vdots
\vdots	\vdots	\vdots	\vdots	\vdots
k	Φ^2	$\sum_{=1}^k (1 + \frac{2}{2})$	$\epsilon^2 \sum_{=1}^k (1 + \frac{2}{2})^{-2} \prod_{=1}^k 1 \cdot \prod_{=1}^k 2$	$\epsilon^{2-k} \sum_{=1}^k (1 + \frac{2}{2})^{-(2-k)} \prod_{=1}^k 1 \cdot \prod_{=1}^k 2$
\vdots	\vdots	\vdots	\vdots	\vdots

$\pi'(0^+) > 0, \pi'(1^-) < 0$, which implies that $\pi(\epsilon)$ has a unique maximum in

$$\alpha_{-1} = \sum_{=1}^{+1} \text{ }_{-1}(\text{ }), \quad \beta_{-1} = \sum_{=1}^{+1} \text{ }_{-1}(\text{ })(1 - \text{ }),$$

$$\alpha_{-1} + \beta_{-1} = 1. \quad \mathbf{F} \quad \mathbf{y}_{-1} \quad (\text{ }), \quad \alpha_{-1}$$

$$\beta_{-1} \quad \mathbf{y}_{-1}, \quad :$$

$$\mu_\epsilon = \mathbb{E}_{\mathcal{Z}}\{\mathbb{E}_{\Phi}[\log(\mathbf{1}'\mathbf{F} \mathbf{y}_{-1})|\mathbf{y}_{-1}]\} = \mathbb{E}_{\mathcal{Z}}\{\mathbb{E}_{\Phi}[\log(\alpha_{-1} \text{ } , + \beta_{-1} \text{ } ,)|\mathbf{y}_{-1}]\}$$

$$\leq \mathbb{E}_{\mathcal{Z}}\{\mathbb{E}_{\Phi}[\log(\text{ }^* \text{ } , + (1 - \text{ }^*) \text{ } ,)|\mathbf{y}_{-1}]\} = \mathbb{E}_{\mathcal{Z}}\{\mu(\text{ }^*)|\mathbf{y}_{-1}\} = \mu(\text{ }^*),$$

- Ben-Jacob, E., 2008, Social Behavior of Bacteria: From Physics to Complex Organizations, 65(3), 315–322.
- Biais, B., and R. Shadur, 2000, Darwinian Selection Does Not Eliminate Irrational Traders, 44(3), 469–490.
- Black, F., and M. Scholes, 1973, The Pricing of Options and Corporate Liabilities, 81(3), 637–654.
- Blume, L., and D. Easley, 1992, Evolution and Market Behavior, 58(1), 9–40.
- Bowe, M., and D. Domuta, 2004, Investor Herding during Financial Crisis: A Clinical Study of the Jakarta Stock Exchange, - 12(4), 387–418.
- Brennan, T. J., and A. W. Lo, 2011, The Origin of Behavior, 1, 55–108.
- Brennan, T. J., and A. W. Lo, 2012, An Evolutionary Model of Bounded Rationality and Intelligence, 7(11), e50310.
- Burnham, T. C., 2013, Toward a Neo-Darwinian Synthesis of Neoclassical and Behavioral Economics, 90, S113–S127.
- Camerer, C. F., and E. Fehr, 2006, When Does “Economic Man” Dominate Social Behavior? 311(5757), 47–52.
- Campbell, D. T., 1986, Rationality and Utility from the Standpoint of Evolutionary Biology, 59(4), S355–S364.
- Campbell, J. R., 1998, Entry, Exit, Embodied Technology, and Business Cycles, 1(2), 371–408.
- Caswell, H., 2001, , 2nd Edition, Sinauer Associates, Sunderland, MA.
- Chen, M. K., V. Lakshminarayanan, and L. R. Santos, 2006, How Basic Are Behavioral Biases? Evidence from Capuchin Monkey Trading Behavior, 114(3), 517–537.
- Chen, R., and Y. Chen, 2011, The Potential of Social Identity for Equilibrium Selection, 101(6), 2562–2589.
- Cooper, W. S., and R. H. Kaplan, 1982, Adaptive Coin-Flipping: A Decision-Theoretic Examination of Natural Selection for Random Individual Variation, 94(1), 135–151.
- Daniel, K., and S. Titman, 1999, Market Efficiency in an Irrational World, 55(6), 28–40.
- Dawkins, R., 1976, , Oxford University Press, Oxford, UK.
- De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann, 1990, Noise Trader Risk in Financial Markets, 98(4), 703–738.
- De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann, 1991, The Survival of Noise Traders in Financial Markets, 64(1), 1–19.
- Desai, M. M., and D. S. Fisher, 2011, The Balance between Mutators and Non-Mutators in Asexual Populations, 188(4), 997–1014.
- Drake, J. W., B. Charlesworth, D. Charlesworth, and J. F. Crow, 1998, Rates of Spontaneous Mutation, 148(4), 1667–1686.
- Dreber, A., C. L. Apicella, D. T. Eisenberg, J. R. Garcia, R. S. Zamore, J. K. Lum, and B. Campbell, 2009, The 7R Polymorphism in the Dopamine Receptor D₄

- Gene (DRD4) Is Associated with Financial Risk Taking in Men,
30(2), 85–92.
- Drehmann, M., J. Oechssler, and A. Roeder, 2005, Herding and Contrarian Behavior in Financial Markets: An Internet Experiment,
95(5), 1403–1426.
- Fama, E. F., 1970, Efficient Capital Markets: A Review of Theory and Empirical Work,
25(2), 383–417.
- Fershtman, C., U. Gneezy, and J. A. List, 2012, Equity Aversion: Social Norms and the Desire to Be Ahead,
4(4), 131–144.
- Frank, S. A., 2011, Natural Selection. II. Developmental Variability and Evolutionary Rate,
24, 2310–2320.
- Gaal, B., J. W. Pitchford, and A. J. Wood, 2010, Exact Results for the Evolution of Stochastic Switching in Variable Asymmetric Environments,
184(4), 1113–1119.
- Geroski, P. A., 1995, What Do We Know about Entry?
13(4), 421–440.
- Gillespie, J. H., 1981, Mutation Modification in a Random Environment,
35, 468–476.
- Gilovich, T., D. Griffin, and D. Kahneman, 2002,
, Cambridge University Press, Cambridge, UK.
- Gneezy, U., and A. Imas, 2014, Mazerazzi Effect and the Strategic Use of Anger in Competitive Interactions,
111(4), 1334–1337.
- Gneezy, U., and J. A. List, 2006, Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments,
74(5), 1365–1384.
- Gneezy, U., and J. A. List, 2013,
, PublicAffairs, New York City.
- Gneezy, U., J. A. List, and G. Wu, 2006, The Uncertainty Effect: When a Risky Prospect Is Valued Less Than Its Worst Possible Outcome,
121(4), 1283–1309.
- Grant, D. A., H. W. Hake, and J. P. Hornseth, 1951, Acquisition and Extinction of Verbal Conditioned Responses with Differing Percentages of Reinforcement,
42(1), 1–5.
- Hajnal, J., 1976, On Products of Non-Negative Matrices,
79(3), 521–530.
- Hamilton, W. D., 1964, The Genetical Evolution of Social Behavior. I and II,
7(1), 1–52.
- Harder, L. D., and L. A. Real, 1987, Why Are Bumble Bees Risk Averse?
68(4), 1104–1108.
- Herrnstein, R. J., 1961, Relative and Absolute Strength of Responses as a Function of Frequency of Reinforcement,
4(3), 267–272.
- Hirshleifer, D., and G. Y. Luo, 2001, On the Survival of Overconfident Traders in a Competitive Securities Market,
4(1), 73–84.

- Hirshleifer, D., A. Subrahmanyam, and S. Titman, 2006, Feedback and the Success of Irrational Investors, *Journal of Economic Behavior & Organization*, 81(2), 311–338.
- Hirshleifer, D., and S. H. Teoh, 2009, Thought and Behavior Contagion in Capital Markets, in T. Hens, and K. Schenk-Hoppe (editors), *Contagion in Financial Markets*, Elsevier, pp. 1–46.
- Hirshleifer, J., 1977, Economics from a Biological Viewpoint, *Journal of Economic Literature*, 20, 1–52.
- Houston, A. I., J. M. McNamara, and M. D. Steer, 2007, Do We Expect Natural Selection to Produce Rational Behaviour? *Evolutionary Economics*, 362(1485), 1531–1543.
- Hsu, M., M. Bhatt, R. Adolphs, D. Tranel, and C. F. Camerer, 2005, Neural Systems Responding to Degrees of Uncertainty in Human Decision-Making, *Journal of Neuroscience*, 310(5754), 1680–1683.
- Hu, H. T., 2012, Efficient Markets and the Law: A Predictable Past and an Uncertain Future, *Journal of Economic Surveys*, 4(1), 179–214.
- Ipsen, I. C., and T. M. Selee, 2011, Ergodicity Coefficients Defined by Vector Norms, *Journal of Economic Theory*, 32(1), 153–200.
- Ishii, K., H. Matsuda, Y. Iwasa, and A. Sasaki, 1989, Evolutionarily Stable Mutation Rate in a Periodically Changing Environment, *Journal of Theoretical Biology*, 121(1), 163–174.
- Johnson, D. D., and J. H. Fowler, 2011, The Evolution of Overconfidence, *Journal of Economic Psychology*, 477(7364), 317–320.
- Kahneman, D., and A. Tversky, 1979, Prospect Theory: An Analysis of Decision under Risk, *Econometrica*, 47(2), 263–291.
- Kahneman, D., and A. Tversky, 2000, *Thinking Fast and Slow*, Cambridge University Press, Cambridge, UK.
- Kenrick, D. T., V. Griskevicius, J. M. Sundie, N. P. Li, Y. J. Li, and S. L. Neuberg, 2009, Deep Rationality: The Evolutionary Economics of Decision Making, *Journal of Economic Psychology*, 27(5), 764–785.
- Kimura, M., 1960, Optimum Mutation Rate and Degree of Dominance as Determined by the Principle of Minimum Genetic Load, *Journal of Theoretical Biology*, 57(1), 21–34.
- King, J. L., 1972, The Role of Mutation in Evolution, in L. M. Le Cam, J. Neyman, and E. L. Scott (editors), *Proceedings of the National Academy of Sciences*, Vol. V, University of California Press, Berkeley, CA, pp. 69–100.
- Kirman, A., 1993, Ants, Rationality, and Recruitment, *Journal of Economic Behavior & Organization*, 108(1), 137–156.
- Klepper, S., and E. Graddy, 1990, The Evolution of New Industries and the Determinants of Market Structure, *Journal of Economic Behavior & Organization*, 21(1), 27–44.
- Kogan, L., S. A. Ross, J. Wang, and M. M. Westerfield, 2006, The Price Impact and Survival of Irrational Traders, *Journal of Financial Economics*, 61(1), 195–229.
- Kussell, E., and S. Leibler, 2005, Phenotypic Diversity, Population Growth, and Information in Fluctuating Environments, *Journal of Theoretical Biology*, 309(5743), 2075–2078.
- Kyle, A., and F. A. Wang, 1997, Speculation Duopoly with Agreement to Disagree: Can Overconfidence Survive the Market Test? *Journal of Economic Theory*, 52(5), 2073–2090.
- Leigh, E. G. Jr., 1970, Natural Selection and Mutability, *Journal of Theoretical Biology*, 104, 301–305.

- Levins, R., 1967, Theory of Fitness in a Heterogeneous Environment. VI. The Adaptive Significance of Mutation, *Evolution* 56(1), 163–178.
- Levins, R., 1968, *Evolution in the Two Species Area*, (Vol. 2), Princeton University Press, Princeton, NJ.
- Liberman, U., J. Van Cleve, and M. W. Feldman, 2011, On the Evolution of Mutation in Changing Environments: Recombination and Phenotypic Switching, *Evolution* 187(3), 837–851.
- Lo, A. W., 2004, The Adaptive Markets Hypothesis: Market Efficiency from an Evolutionary Perspective, *Journal of Applied Corporate Finance* 30(5), 15–29.
- Lo, A. W., 2017, *Adaptive Markets: Financial Markets as a Natural Phenomenon*, Princeton University Press, Princeton, NJ.
- Lo, A. W., H. A. Orr, and R. Zhang, 2018, The Growth of Relative Wealth and the Kelly Criterion, *Journal of Applied Corporate Finance* 20(1), 49–68.
- Lucas, D., 2011, Government as a Source of Systemic Risk, Working Paper, MIT Center for Finance and Policy.
- Lucas, R. E. Jr, 1972, Expectations and the Neutrality of Money, *Journal of Monetary Economics* 4(2), 103–124.
- Mahoney, M. J., 1977, Publication Prejudices: An Experimental Study of Confirmatory Bias in the Peer Review System, *Journal of the American Academy of Business* 1(2), 161–175.
- Mata, J., P. Portugal, and P. Guimaraes, 1995, The Survival of New Plants: Start-Up Conditions and Post-Entry Evolution, *Journal of Business Venturing* 13(4), 459–481.
- Maynard Smith, J., 1982, *Evolution and the Theory of Games*, Cambridge University Press, Cambridge, UK.
- Maynard Smith, J., 1984, Game Theory and the Evolution of Behaviour, *Journal of Theoretical Biology* 7(1), 95–101.
- McDermott, R., J. H. Fowler, and O. Smirnov, 2008, On the Evolutionary Origin of Prospect Theory Preferences, *Journal of Experimental Psychology* 70(2), 335–350.
- McKenzie, C. R., 2003, Rational Models as Theories — not Standards — of Behavior, *Journal of Applied Corporate Finance* 7(9), 403–406.
- Merton, R. C., 1973, Theory of Rational Option Pricing, *Journal of Financial Economics* 4(1), 141–183.
- Meyn, S. S. P., and R. L. Tweedie, 2009, *Stochastic Processes: Limit Theory and Applications*, Cambridge University Press, Cambridge, UK.
- Nash, J. F., 1950, Equilibrium Points in N-Person Games, *Econometrica* 36(1), 48–49.
- Nowak, M., and R. Highfield, 2011, *Evolutionary Dynamics: Explaining the Success of Strategies*, Simon and Schuster, New York City.
- Okasha, S., and K. Binmore, 2012, *Evolutionary Game Theory*, Cambridge University Press, Cambridge, UK.
- Rabin, M., and R. H. Thaler, 2001, Anomalies: Risk Aversion, *Journal of Applied Corporate Finance* 15(1), 219–232.
- Robson, A. J., 1996a, A Biological Basis for Expected and Non-Expected Utility, *Journal of Economic Theory* 68, 397–424.

- Robson, A. J., 1996b, The Evolution of Attitudes to Risk: Lottery Tickets and Relative Wealth, 14, 190–207.
- Rogers, A. R., 1994, Evolution of Time Preference by Natural Selection, 84, 460–481.
- Ross, C., and A. Wilke, 2011, Past and Present Environments, 9(3), 275–278.
- Samuelson, L., 2001, Introduction to Evolution of Preferences, 97, 225–230.
- Samuelson, P. A., 1958, An Exact Consumption–Loan Model of Interest with or without the Social Contrivance of Money, 66(6), 467–482.
- Samuelson, P. A., 1965, Proof That Properly Anticipated Prices Fluctuate Randomly, 6(2), 41–49.

- Tversky, A., and D. Kahneman, 1974, Judgment under Uncertainty: Heuristics and Biases, *Journal of Experimental Psychology: Applied*, 185(4157), 1124–1131.
- Tversky, A., and D. Kahneman, 1992, Advances in Prospect Theory: Cumulative Representation of Uncertainty, *Journal of Economic Literature*, 5(4), 297–323.
- von Neumann, J., and O. Morgenstern, 1944, *The Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ.
- Waksberg, A. J., A. B. Smith, and M. Burd, 2009, Can Irrational Behaviour Maximise Fitness? *Journal of Economic Surveys*, 63(3), 461–471.
- Waldman, M., 1994, Systematic Errors and the Theory of Natural Selection, *Journal of Economic Surveys*, 84, 482–497.
- Wilson, E. O., 1975, *Consilience: The Unity of Knowledge*, Harvard University Press, Cambridge, MA.
- Wilson, E. O., and W. H. Bossert, 1971, *Group Selection: A Critique*, Sinauer Associates Sunderland, Sunderland, MA.
- Zhang, R., T. J. Brennan, and A. W. Lo, 2014a, Group Selection as Behavioral Adaptation to Systematic Risk, *Journal of Economic Surveys*, 9(10), e110848.
- Zhang, R., T. J. Brennan, and A. W. Lo, 2014b, The Origin of Risk Aversion, *Journal of Economic Surveys*, 111(50), 17777–17782.