

A Biological Basis for Expected and Non-expected Utility*

ARTHUR J. ROBSON

Department of Economics, University of Western Ontario, London, Ontario, Canada N6A 5C2

Received April 16, 1993; revised March 15, 1995

A biological model is developed here to determine the fittest attitude to risk. With a fixed environment, the type maximizing expected offspring is selected. This yields the expected utility theorem when translated into a criterion for evaluating gambles over commodities. With a random environment, however, the type selected is strictly less averse to idiosyncratic risk than to risk which is correlated across all individuals. The implied criterion for choice over gambles does not satisfy the expected utility theorem and may induce choice of a gamble which is first-order stochastically dominated. *Journal of Economic Literature* Classification Number: D81. © 1996 Academic Press, Inc.

1. INTRODUCTION

The central thesis of this paper is that modern attitudes to risk are, at least in part, the product of biological evolution. This argument could be applied to preferences in general and contrasts with the usual axiomatic approach.¹ Consider, for example, preferences for sweet and rich foods. The tastes for these foods were presumably formed when food was not plentiful so that such concentrated sources of calories were valuable for survival and, even more basically, for the production of offspring. These tastes are less appropriate in wealthy modern societies where overindulgence is readily possible and consumption of these foods seems unlikely to increase the number of offspring. That some preferences seem inappropriate in our

* I appreciate the support of the Social Sciences and Humanities Research Council of Canada for this research. I am grateful to Eddie Dekel and to Peter Streufert for helpful remarks on earlier versions of the paper. This was also improved by a number of presentations, especially those at the Hebrew University of Jerusalem, Tel Aviv University, the conference "Evolution and Economics" at the University of Western Ontario, the Summer Meetings of the Econometric Society at Boston University, and the workshop "Evolutionary Game Dynamics in Biology and Economics" at IIASA, Laxenburg, Austria. Finally, I thank a referee and an associate editor for their comments.

¹ The desirability of considering the biological formation of preferences was advocated, for example, by Hirshleifer [11].

current environment may be evidence that such preferences were hard-wired by evolution. It is hypothesized here that our attitudes to gambles over commodities were hard-wired and so continue to be exhibited now despite the apparent severing of the link between commodities and offspring.

Take then the following abstract model of human evolutionary history. The population is composed of individuals of a number of different types, where each type has a characteristic probability distribution of offspring. Types are basically hereditary and there are an unlimited number of generations.² The key question addressed here is: Which type will be selected? That is, which type (if any) will ultimately dominate the population? The existence of a favored type has more familiar implications as follows. Imagine that every individual of each type takes a particular action and that it is this action which determines the characteristic probability distribution of offspring for that type. Selecting a type then yields an associated favored action and illuminates the preferences to be encountered in the population.

The model here is general in that there are a small number of individuals initially and extinction is possible. It is shown that a favored type exists under each of two scenarios concerning the statistical dependence of the gambles taken by individuals. In the first of these scenarios, which is considered in Sections 2 and 3, each type of individual faces an offspring distribution that is fixed from generation to generation. Furthermore, these gambles are statistically independent across all generations and individuals. Now the type maximizing expected offspring ultimately dominates the population, whenever this population is not extinct. In other words, the evolutionarily most successful attitude to risk is risk neutrality in offspring. It is shown in Section 4 that this result induces preferences for gambles over commodities representable by expected utility. The von Neumann–Morgenstern utility function here is the biological production function relating offspring to commodities. The selected attitudes to risk involving gambles over commodity bundles then derive from the shape of this production function. If hard-wired, these attitudes to risk would still be exhibited in modern individuals.

The second scenario, which is considered in Section 5, is more general in assuming that there is an environmental dependence among the gambles taken by individuals in a given generation. In each period, there is an i.i.d. draw of an environmental parameter to determine an offspring distribution for each type, and independent draws from these distributions are then made for each individual. Section 6.1 shows that the result concerning the type selected implies a non-expected utility criterion for evaluating gambles

² Types are “basically” hereditary in that small probabilities of mutation are present.

over commodities. This criterion incorporates an essential distinction between idiosyncratic risk given the environment and aggregate uncertainty concerning the environment itself, implying a greater aversion toward the latter. This is because the present model generates a form of automatic *biological* insurance via the law of large numbers against idiosyncratic risk, whereas this insurance is inoperative in the same sense against aggregate uncertainty.³ Section 6.2 discusses how the present criterion might lead to the choice of one gamble involving idiosyncratic risk over another involving aggregate uncertainty, despite the stochastic domination of the first gamble by the second from an individual viewpoint. It also discusses the evolutionary rationale for this and presents some relevant empirical observations.

Section 7 concludes by relating the present results to previous work in biology and economics.

2. TWO INTRODUCTORY EXAMPLES

The first example here introduces the concept of a branching process, with a single type and independent offspring distributions. It assumes there is a small number of individuals initially, but discusses how the type with the highest mean offspring level ultimately dominates the population, regardless of the other characteristics of the various distributions.

EXAMPLE 1. Suppose that there are two types of individuals, 1 and 2. Each type 1 individual produces 1 offspring with probability $1-p > 0$ and 3 offspring with probability $p > 0$, as an independent event. The mean number of offspring produced by each type 1 individual is then $m_1 = 1 + 2p$. It is assumed that $p > 1/2$, so that $m_1 > 2$. On the other hand, each type 2 individual produces 2 offspring with certainty, so that the mean is $m_2 = 2$.⁴ Each type breeds true in the sense that offspring are of the same type as their parent. Suppose there is initially 1 individual of each type and there are an infinite number of generations. Define $z_T(i)$ as the population of type $i = 1, 2$, in generation $T = 0, 1, \dots$. Of course, $z_T(2) = 2^T$, $T = 0, 1, \dots$. It is clear that $z_T(1) \rightarrow \infty$, with probability 1, as $T \rightarrow \infty$. The theory of branching processes applied in the present paper proves, more precisely, that $z_T(1)/(m_1)^T \rightarrow W$, as $T \rightarrow \infty$, with probability 1, for some non-negative random variable W . In this case, since extinction is impossible,

³ The present result should be contrasted with a somewhat analogous distinction between idiosyncratic risk and aggregate uncertainty stemming from their *market* insurability. This more familiar distinction arises in the presence of standard preferences that, in themselves, do not differentiate between idiosyncratic risk and aggregate uncertainty.

⁴ Such degenerate distributions are ruled out in the general model, but are especially useful in simple examples.

W is strictly positive with probability 1.⁵ Hence $z_T(1)/z_T(2) \rightarrow \infty$, with probability 1, as $T \rightarrow \infty$, and type 1 ultimately dominates type 2 in a straightforward sense. Thus the greater mean of the offspring distribution of a type 1 individual ultimately renders irrelevant its greater variance.

The above example conveys the main point of the version of the present model without aggregate uncertainty, but it has the special feature that both types have strictly positive levels of offspring. Allowing the possibility of zero offspring is realistic but requires an extension of the model. Example 2 is obtained from Example 1 by reducing all offspring levels by one, so that type 1 now has a positive probability of zero offspring. A difficulty brought out by Example 2 is that type 2 will dominate type 1 when type 1 becomes extinct. As is also illustrated by Example 2, a key property of a (non-degenerate) branching process is that the probability of ultimate extinction is less than one if and only if the mean is greater than one.

EXAMPLE 2. Each type 1 individual now produces 0 offspring with probability $1-p$ and 2 offspring with probability p , so the mean number of offspring is $m_1 = 2p$. Each type 2 individual produces 1 offspring with certainty, so that $m_2 = 1$.⁶ Again, there is initially 1 individual of each type. Regardless of $p \in (0, 1)$, even if $m_1 > m_2$, there is a possibility that type 2's dominate the population simply because type 1's are subject to extinction but type 2's are not. For example, there is probability $q = 1-p > 0$ that extinction of the type 1's occurs at $T=1$.

Whether it is *inevitable* that type 1 ultimately becomes extinct depends on the size of m_1 relative to 1. Suppose the probability that type 1 is extinct at date T , given there is 1 individual of type 1 at $T=0$, is $r_T > 0$. By definition, $r_1 = q$. The probability of ultimate extinction is then $\lim_{T \rightarrow \infty} r_T$, where this limit exists since $r_{T+1} \geq r_T$, $T=1, 2, \dots$. The independence of the type 1 process implies that the probability of extinction in T periods given there are two individuals initially is $(r_T)^2$. Since the probability that there are two individuals at $T=1$ is p whereas the probability of immediate extinction at $T=1$ is q , it then follows that $r_{T+1} = q + p(r_T)^2$, $T=0, 1, \dots$ ⁷ Consideration of this difference equation yields that $r_T \uparrow r$, as $T \rightarrow \infty$, where r is the smallest root of the equation $r = q + pr^2$. If $p \leq 1/2$, so that $m_1 \leq 1$, it

⁵ See Athreya and Ney [1, p. 9, Theorems 1 and 2]. Intuitively, W must be a random variable since a string of good or bad luck at the beginning would have a permanent multiplicative effect on the level of the population, although it would not affect the ultimate growth rate.

⁶ As before, such a degenerate distribution is ruled out in the general model.

⁷ This difference equation for the probability that $z_T(1) = 0$ is implied by the general result that the moment generating function of $z_T(1)$ is the T th functional iterate of the moment generating function of $z_1(1)$. See Athreya and Ney [1, p. 2].

follows that $r=1$. If $p>1/2$, so that $m_1>1$, on the other hand, $r=q/p\in(0,1)$. That is, type 1 avoids extinction forever with probability $1-q/p>0$.

When $m_1>1$, it is again true that there exists a non-negative random variable W such that $z_T(1)/(m_1)^T\rightarrow W$, as $T\rightarrow\infty$, with probability 1. Further, $W=0$ essentially if and only if extinction occurs. If extinction does not occur it follows that the population grows to infinity at the same rate as $(m_1)^T$, with probability 1.⁸

The possibility that the type 2's dominate the population due to the extinction of the type 1's is artificial. One resolution of the difficulty would be to consider infinitely large populations directly. The growth rate of each type would be its mean offspring level, by the law of large numbers. The alternative resolution adopted here is to introduce small but strictly positive mutation rates. This maintains detailed descriptions of small populations and so is capable of explaining how such small populations can become large.⁹ Although strictly positive mutation rates do not prevent the extinction of the entire population, they do essentially prevent the extinction of any given type while another type survives. If the highest mean is greater than one, the next section shows that the type with this highest mean ultimately dominates, conditional only on non-extinction of the entire population.

3. INDEPENDENT GAMBLES

3.1. *The Model*

Consider $G\geq 2$ types of individuals, where each type has a number of members that varies over the generations $T=0,1,\dots$. Type $i=1,\dots,G$ is characterized by a distribution of offspring represented by $p_0^i,\dots,p_B^i\geq 0$, where $\sum_{j=0}^B p_j^i=1$. That is, p_j^i is the probability that each type i individual in any generation has j offspring, for $j=1,\dots,B$.¹⁰ Each individual reproduces independently.

There is a probability $\lambda_{ij}>0$ that each such offspring of a type i individual is of type j , for all $j\neq i$, where with remaining probability, $A_i=(1-\sum_{j\neq i}\lambda_{ij})>0$, each such offspring remains of type i . The vector of

⁸ See Athreya and Ney [1, p. 9, Theorems 1 and 2].

⁹ Such detail also permits a rigorous examination in Sections 5 and 6 of the distinction between idiosyncratic risk and environmental uncertainty.

¹⁰ To avoid trivialities it is assumed, as in [1, p. 3], that $p_0^i+p_1^i<1$ and $p_j^i\neq 1$, for all i and j . It is biologically reasonable and technically convenient to place the upper bound, B , on the number of offspring produced by a single individual in a single generation.

all such λ_{ij} , $i, j = 1, \dots, G$, $i \neq j$, is denoted by $\lambda \in \mathbf{R}_{++}^{G(G-1)}$. These mutations are also independent events.

The presence of mutation implies that the evolution of the various types must be considered together as a multitype branching process. Define

$$M_{ij}(\lambda) = \begin{cases} \lambda_{ij}m^i, & \text{if } i \neq j, i, j = 1, \dots, G \\ A_i m^i, & \text{if } i = j = 1, \dots, G, \end{cases}$$

where m^i is the mean total number of offspring for a single individual of type i , $m^i = \sum_{j=1}^G j \cdot p_j^i \in (0, \infty)$, $i = 1, \dots, G$. Thus $M_{ij}(\lambda)$ is the expected number of type j offspring produced in a single generation by a single individual of type i . It is remarkable that key properties of multitype branching processes rely only on this matrix.¹¹

For example, suppose that the random (row) vector representing the population of G types at date $T = 1, 2, \dots$ is $z_T = (z_T(1), \dots, z_T(G))$ and that the initial population vector is $z_0 = (z_0(1), \dots, z_0(G))$. It follows that the mean of z_T is related to $M(\lambda)$ and z_0 by $E(z_T | z_0) = z_0 M(\lambda)^T$, for $T = 1, 2, \dots$

A preliminary result is quoted for reference.

THEOREM 0 (Perron–Frobenius). *Suppose a matrix $M(\lambda)$ has all its entries strictly positive. Then $M(\lambda)$ has a simple real eigenvalue, $m(\lambda)$, say, which strictly exceeds the modulus of any other (perhaps complex) eigenvalue of $M(\lambda)$. Furthermore, there is a unique normalized left eigenvector associated with $m(\lambda)$, $v(\lambda)$, say, such that $v(\lambda) > 0$.¹²*

Proof. See Gantmacher [8, Vol. II, pp. 53–66].

3.2. The Maximization of Expected Offspring

Theorem 1 below shows that, if the probability of extinction is less than one, then the population grows to infinity with the remaining positive probability. Furthermore, when growth to infinity occurs, the proportions of the various types in the population converge to a (non-stochastic) mixture. As the mutation rates converge to zero, this long run mixture converges to a degenerate mixture having only the type with the highest mean offspring level.

This predicts the outcome of the following experiment. Suppose that a species with the potential for unlimited growth is examined T periods after its creation. One possibility is that it is extinct. Given that it is not extinct, it can be made arbitrarily likely that the proportion of the type with highest mean offspring is arbitrarily close to one. This requires choosing

¹¹ See Athreya and Ney [1, Chapter V].

¹² The elements of $v(\lambda)$ are strictly positive and sum to one.

the mutation rates small enough and the time T large enough. In this appealing sense, evolution ensures the selection of a particular type.¹³

THEOREM 1. *Consider the model of Section 3.1. Suppose that $m^1 > m^i$, $i = 2, \dots, G$, so type 1 is the unique type with maximal expected offspring.¹⁴*

(i) *Positive mutation rates lower the asymptotic growth rate in that $m(\lambda) < m^1$, for all mutation vectors $\lambda > 0$.¹⁵*

(ii) *If $m^1 \leq 1$ then $m(\lambda) < 1$, for all mutation vectors $\lambda > 0$. Now the probability of ultimate extinction of the entire population is 1, for all $\lambda > 0$ and all initial vectors, z_0 .*

(iii) *Suppose instead that $m^1 > 1$. Now, if the mutation vector $\lambda > 0$ is small enough, then $m(\lambda) > 1$ and the population becomes extinct with probability strictly less than 1, for any initial vector $z_0 \neq 0$. With the remaining positive probability, the population grows asymptotically at the same rate as $m(\lambda)^T$, with the vector of proportions of the types converging to the associated (non-stochastic) normalized left eigenvector of $M(\lambda)$, $v(\lambda)$. That is, if $m(\lambda) > 1$, then $\lim_{T \rightarrow \infty} z_T / (m(\lambda))^T = v(\lambda) W$, with probability 1, where W is a non-degenerate non-negative random variable. In addition, the event that $W = 0$ essentially coincides with the event of ultimate extinction.¹⁶ Finally, it follows that $v(\lambda) \rightarrow e_1 = (1, 0, \dots, 0)$, as $\lambda \rightarrow 0$.*

Proof. See the Appendix.

3.3. An Example with Two Favored Types

Theorem 1 shows that a type with *strictly* higher expected offspring will ultimately be selected. A minor issue is: What if two types attain the maximum level of expected offspring? The following example shows that the long run proportions of the two types in the population are then sensitive to the ratio of the two mutation rates.

EXAMPLE 3. Suppose the mean matrix is given by

$$M = \begin{bmatrix} 2(1-\lambda) & 2\lambda \\ 2\mu & 2(1-\mu) \end{bmatrix},$$

¹³ This implication of the result of Theorem 1 is shown in Theorem 2 to hold in more general circumstances.

¹⁴ Example 3 below considers the secondary issue of non-uniqueness.

¹⁵ " $\lambda > 0$ " means $\lambda_{ij} > 0$, for all $i \neq j$.

¹⁶ The event of ultimate extinction is clearly essentially a subset of the event that $W = 0$. In addition, the event that $W = 0$ but non-extinction also occurs has probability zero.

where $\lambda = \lambda_{12}$ and $\mu = \lambda_{21}$, in the notation introduced above. The maximal eigenvalue here is 2 and the associated normalized left eigenvector v , say, is $v = (\mu, \lambda)/(\mu + \lambda)$. Thus the long-run proportion of type 1's and type 2's in v is arbitrary, determined by the ratio λ/μ , and the limit of v as $(\lambda, \mu) \rightarrow (0, 0)$ does not exist.

3.4. *Extensions*

The present model abstracts from various aspects of biological reality. For example, it is assumed that each individual has a single parent and type is inherited except for the small probabilities of mutation. The actual process of inheritance in a diploid sexual species such as humans is more complex. In addition, it is assumed here that each individual lives for a single period. A more realistic model would involve individual consumption and reproduction occurring over a number of periods. Although the present model should serve to illuminate some of the central properties of more general models, these may also have properties not found here.¹⁷

The present branching processes have a property of extinction or explosion. This instability means they must be interpreted with care as models of actual biological populations, given that explosion should then be ruled out by carrying capacity limits. Even if the model were modified so that the feedback affect of the aggregate population were strong enough to bound its size, however, this would merely ensure that extinction would occur with probability equal to one.¹⁸

This ultimate certainty of extinction is not of great practical significance. After all, at any point in time at which a large carrying capacity is an effective constraint, the probability of extinction in each period would be extremely small and the expected remaining life of the population extremely long. For the modified model, it seems reasonable to conjecture that any type having the highest mean at *all* levels of the aggregate population will still ultimately dominate the population, given that this population is not extinct. As before, this will require mutation rates to be small and the time of observation to be far into the future. As an additional condition, however, the carrying capacity will now need to be large.

¹⁷ Note that evolutionary game theory as applied in biology typically makes both abstractions here.

¹⁸ When the offspring distributions depend on population but remain independent, Athreya and Ney [1, p. 8] claim that instability will still occur. There is an uninteresting way to avoid this instability, which is to assume that each individual has one offspring in the limit. However, in a slightly different model it can be shown that there is no other way [1, pp. 266–267].

4. EXPECTED UTILITY AND WELFARE

4.1. *Expected Utility*

The present section shows how the maximization of expected offspring translates into the maximization of expected utility. Thus the linearity of expecting offspring in probabilities is inherited by the criterion for evaluating gambles over commodities. Properties of risk aversion or risk preference for gambles over commodities then derive from a biological production function relating expected offspring to commodities.

Take the consumption space over commodities to be \mathbf{R}_+^n . Suppose that a bundle $x \in \mathbf{R}_+^n$ induces for each individual of any type a distribution over offspring represented by $p_0(x), p_1(x), \dots, p_B(x) \geq 0$, where $\sum_{j=0}^B p_j(x) = 1$. Thus $p_j(x)$ is the probability that this bundle x results in j (surviving) offspring.¹⁹ These gambles are independent over all individuals. Define $\psi: \mathbf{R}_+^n \rightarrow \mathbf{R}_+$ by $\psi(x) = \sum_{j=1}^B j \cdot p_j(x) \in (0, \infty)$, for each $x \in \mathbf{R}_+^n$, as the production function for expected offspring from commodities.²⁰ Suppose now that each particular type of individual takes a particular simple gamble over commodities. Suppose that $q_k^i \in [0, 1]$ is the probability that bundle $x_k^i \in \mathbf{R}_+^n$ is obtained in the gamble taken by type i . This is for $k = 1, \dots, K$, where $\sum_{k=1}^K q_k^i = 1$, and for $i = 1, \dots, G$.²¹ If these gambles over commodity bundles are independent over all individuals, Theorem 1 implies that the type which is selected maximizes the following expression for overall expected offspring:²²

$$\sum_{k=1}^K \sum_{j=1}^B j q_k^i p_j(x_k^i) = \sum_{k=1}^K q_k^i \psi(x_k^i), \quad \text{over } i = 1, \dots, G.$$

Hence the evolutionarily induced preferences for gambles over commodities satisfy the expected utility theorem, where von Neumann–Morgenstern utility is the production function for expected offspring, ψ . The attitudes to risk stemming from the concavity and convexity properties of this von Neumann–Morgenstern utility function are then biological in origin. For example, if the production of expected offspring from commodities were described by a concave function, this implies risk aversion toward gambles over commodities.

¹⁹ It is assumed that $p_0(x) + p_1(x) < 1$ and $p_j(x) \neq 1$, for all j and all $x \in \mathbf{R}_+^n$.

²⁰ Robson [17] examines an evolutionary model in which commodities are not the sole determinants of offspring, but status also plays a role.

²¹ K is independent of i , without loss of generality.

²² It must also be assumed that this maximum is greater than 1. Affine transformations of ψ have, as usual, no effect on attitudes to risk but remove this significance of the value 1.

4.2. *The Interpretation of Expected Utility*

It is not claimed that typical modern individuals maximize expected offspring. For that matter, the desire for offspring need never have been explicit. What is hypothesized is that preferences over commodities were formed during human evolution as a means for maximizing expected offspring. It is also hypothesized that these preferences over commodities were hard-wired and remain largely unchanged now, despite not having their former implications for offspring in many societies. The legacy of biological evolution in the present circumstances would then be the maximization of expected utility, where the von Neumann–Morgenstern utility function ψ (or an affine transformation of ψ) would appear as a primitive.

Note that Karni and Schmeidler [12] previously proposed maximization of the probability of survival as the basis of an evolutionary explanation of attitudes to risk. Although expected offspring is a more fundamental biological goal than is the probability of survival, the expected offspring criterion also reinforces the expected utility theorem.²³

4.3. *Bentham–Harsanyi Welfare*

It is shown here that large societies which maximize the sum of von Neumann–Morgenstern utilities also maximize the growth rate of total population and thus achieve evolutionary success relative to other societies.²⁴

Consider then the implications of various (non-stochastic) distributions of commodities. Suppose there are N individuals with individual i given bundle $x^i \in \mathbf{R}_+^n$, $i = 1, \dots, N$. These N individuals are otherwise identical. Each such bundle x^i induces expected offspring $\psi(x^i)$, as in Section 4.1. To make this society large, the R -fold replication is considered. Suppose, that is, that individuals $N + i, \dots, (R - 1)N + i$ are replicates of individual i , and are also assigned bundle x^i , $i = 1, \dots, N$. Replicates have identical but independent distributions of offspring. The offspring produced by individual i is a random variable denoted by z^i . It follows that the growth rate of the society is given by

$$(1/N) \sum_{i=1}^N \frac{z^i + z^{N+1} + \dots + z^{(R-1)N+i}}{R} \rightarrow (1/N) \sum_{i=1}^N \psi(x^i), \quad \text{as } R \rightarrow \infty,$$

²³ Note, however, that Sections 5 and 6 show that maximization of expected offspring is not evolutionarily appropriate in more general circumstances than those considered so far.

²⁴ The idea that ethical systems are not products of an independent moral sense but rather have evolved can be found in Hayek [10], for example.

with probability 1, by the strong law of large numbers. That is, the growth rate of a large society obtained in this fashion by replication is, in the limit, the Bentham–Harsanyi welfare criterion composed of the sum of von Neumann–Morgenstern utilities.²⁵

5. A RANDOM ENVIRONMENT

The model of the previous two sections raises a natural question: Why would reproduction actually be precisely independent across individuals? This is not a generic assumption. The weather, for example, seems likely to imply correlation in the reproduction distribution functions. If so, the weather is a random environment in the present statistical sense. The notion of a branching process with a random environment is introduced by means of the following examples.

5.1. *Two Examples with a Random Environment*

The key implication of the present version of the model is that uncertainty due to a random environment has distinct evolutionary consequences from risk given the environment. Both the following examples illustrate this, where the first is simpler than the second. In this first example, a type taking correlated risk will become extinct with probability 1 whereas a type taking similar but independent risk has a positive probability of growing to infinity.

EXAMPLE 4. Suppose that the environment has two states, 1 and 2, say. These states occur independently in each generation $T=0, 1, \dots$, with probabilities $2/3$ and $1/3$, respectively. Both types breed true. Type 1 individuals are sensitive to the environment, producing 2 offspring in state 1 and 0 offspring in state 2. Type 2 individuals are not affected by the environment, but independently produce 2 offspring with probability $2/3$ and 0 offspring with probability $1/3$. Both types have therefore a mean offspring level of $4/3 > 1$ and there is, indeed, no difference between types 1 and 2 from an individual viewpoint. Suppose there is a single individual of each type at $T=0$. As in the model of a fixed environment, there is a positive probability that type 2 never becomes extinct.²⁶ On the other hand, the type 1 population will become extinct at the first occurrence of

²⁵ Harsanyi [9] presents an axiomatic derivation of this welfare criterion.

²⁶ Type 2 avoids extinction with probability $1/2$, as shown for Example 2. In this event, it grows to infinity at the same rate as $(4/3)^T$, with probability 1.

state 2, so its ultimate extinction occurs with probability 1.²⁷ If the combined population is examined at a date sufficiently far into the future, it can then be made arbitrarily likely that only type 2 will exist, given there are any individuals at all.

In the above example, despite the ultimate dominance achieved by type 2, the two types have identical *expected* generation sizes at all dates.²⁸ This common expected generation size is clearly $(4/3)^T$, at each date T . The crucial difference between the two types concerns how this mean is made up. For type 1 individuals, only two extreme possibilities exist at date T . Either there are 2^T individuals, which occurs with probability $(2/3)^T$, or there are none, which occurs with probability $1 - (2/3)^T$. Thus, although the mean tends to infinity, the probability of extinction of type 1 tends to 1, as $T \rightarrow \infty$.²⁹ Less extreme population levels are important for type 2. For example, consider $T=2$. With probability $(2/3)^3 < (2/3)^2$, there are 4 individuals; with probability $2(2/3)^2 (1/3)$, there are 2; and with probability $1/3 + (2/3)(1/3)^2 < 1 - (2/3)^2$, there are 0. That is, the independence of the reproduction of type 2 individuals removes probability from the extremes of the type 1 distribution and shifts it toward the mean. This shifting of probability toward the mean allows the probability of ultimate extinction of type 2 to be less than 1 so that it can ultimately dominate type 1.

The next example shows that the ultimate domination achieved by type 2 is not an artifact of the certainty of ultimate extinction of type 1. It also introduces the measure of the long run growth rate applied here in the presence of both environmental uncertainty and idiosyncratic risk.

EXAMPLE 5. The environment has two states, 1 and 2, each occurring independently in each generation $T=0, 1, \dots$, with probability 1/2. Both types still breed true. Type 1 individuals produce 2 offspring in state 1 and 1 offspring in state 2. Type 2 individuals each independently produce 2 or 1 offspring with probability 1/2. Hence the offspring distributions are again identical from an individual viewpoint. Suppose that there is a single individual of each type at $T=0$. Define $z_T(i)$ as the random population of

²⁷ In the model of the next section, type 1 is ruled out. The degeneracy of this type helps to make the present expositional point, however.

²⁸ This is also true for Example 5 below and demonstrates that the expected size of a population is not always appropriate as a criterion of evolutionary success.

²⁹ This is perhaps reminiscent of a familiar betting paradox. Suppose an infinitely rich individual makes fair bets repeatedly on the flip of a fair coin, quitting at the first win. If the stakes are doubled after each loss, the bettor gains the initial stake with probability 1. This is despite the fairness of the overall bet at every stage. Suppose the process is examined at date T . The probability that the bettor is losing tends to zero, but the amount of this loss tends to infinity and there is a non-vanishing contribution to the mean, as $T \rightarrow \infty$.

type $i = 1, 2$ at time T . Clearly, $z_T(i) \rightarrow \infty$, with probability 1, as $T \rightarrow \infty$, for $i = 1, 2$.

For type 2, the model with a fixed environment implies that $z_T(2)/(3/2)^T \rightarrow W$, with probability 1, as $T \rightarrow \infty$. The random variable W is strictly positive with probability 1.³⁰ This implies that $(1/T) \ln(z_T(2)) \rightarrow \ln(3/2)$, with probability 1, as $T \rightarrow \infty$, so that $\ln(3/2)$ is the long run average growth rate.

What about the long run average growth rate of type 1? Note that $z_T(1) = 2^{n_1(T)}$, where $n_1(T) \in \{0, \dots, T\}$ is the random number of state 1's observed in T periods. The strong law of large numbers implies that $n_1(T)/T \rightarrow 1/2$, with probability 1, as $T \rightarrow \infty$. Hence $(1/T) \ln(z_T(1)) \rightarrow \ln(\sqrt{2})$, with probability 1, as $T \rightarrow \infty$.

This measure of the long run growth rate for the two types has the appropriate implication. That is, it follows that $(1/T) \ln(z_T(2)/z_T(1)) \rightarrow \ln(3/(2\sqrt{2})) > 0$, with probability 1, as $T \rightarrow \infty$, so that $z_T(2)/z_T(1) \rightarrow \infty$, with probability 1. Thus type 2 ultimately dominates the population in this compelling sense.

Despite this result, it is again true in Example 5 that the expected sizes of the two populations are equal, both being $(3/2)^T$ in period T . Note that, although the number of occurrences of state 1 is symmetrically distributed, the type 1 population has a skewed distribution, with its right-hand tail contributing disproportionately to the expectation. This occurs because the size of the type 1 population is exponential in the number of occurrences of state 1. Thus the median of the type 1 distribution, for example, falls increasingly below the mean as T increases. In the limit, remarkably, *all* of the probability mass of the type 1 distribution falls below its own mean and, indeed, also below all of the probability mass of the type 2 distribution.

The possibility of extinction in Example 4 served to dramatize the difference between the two types. More generally, however, the possibility of extinction can still cause the complication which arose in Example 2. That is, the type which performs best in the large number case might accidentally become extinct, despite the survival of other types. Small probabilities of mutation are again introduced into the model to rule this out.

5.2. The Model with a Random Environment

Consider a general model incorporating environmental uncertainty. Suppose there is an infinite sequence of environmental random variables $\xi \equiv (\xi_0, \xi_1, \dots)$. For simplicity, ξ_t , $t = 0, 1, \dots$, are i.i.d. random variables

³⁰ See Athreya and Ney [1, p. 9, Theorems 1 and 2].

with a fixed finite set of realizations $\xi_t \in \{\xi^1, \dots, \xi^S\}$. Here ξ^s has probability $\pi^s > 0$, $s = 1, \dots, S$, and $\sum_{s=1}^S \pi^s = 1$.³¹

There are $G \geq 2$ types of individuals, where the distribution of offspring at time $t+1$ produced by an individual of each type at time $t=0, 1, \dots$ is determined by the realization of ξ_t . Conditional on any realization ξ^s , for $s \in \{1, \dots, S\}$, the type $i \in \{1, \dots, G\}$ has an offspring distribution described by $p_0^i(\xi^s), \dots, p_B^i(\xi^s) \geq 0$ where $\sum_{j=0}^B p_j^i(\xi^s) = 1$. Thus $p_j^i(\xi^s)$ is the probability that type i has j (surviving) offspring in the environment ξ^s .³² Define $m^i(\xi^s) = \sum_{j=1}^B j \cdot p_j^i(\xi^s) \in (0, \infty)$ as the expected number of offspring produced by type i in environment ξ^s , for $i \in \{1, \dots, G\}$ and $s \in \{1, \dots, S\}$. The reproduction of each individual is independent of that of other individuals *conditional* on the environment.

The main result of the present version of the model is that the type selected is that with the maximum long run potential growth rate according to $\rho^i = \sum_{s=1}^S \pi^s \ln(m^i(\xi^s))$, $i = 1, \dots, G$. It is then important to outline the derivation of this criterion as it applies to each type i considered in isolation.³³ Note first that

$$\frac{1}{T} \sum_{t=0}^{T-1} \ln(m^i(\xi_t)) = \sum_{s=1}^S (n^s(T)/T) \ln(m^i(\xi^s)),$$

where $n^s(T)$ is the number of occurrences of the state ξ^s in the first T periods, $s = 1, \dots, S$, so that $\sum_{s=1}^S n^s(T) = T$. The strong law of large numbers implies that $(n^s(T)/T) \rightarrow \pi^s$, with probability 1, as $T \rightarrow \infty$, for $s = 1, \dots, S$. Hence

$$\frac{1}{T} \sum_{t=0}^{T-1} \ln(m^i(\xi_t)) \rightarrow \rho^i \quad \text{as } T \rightarrow \infty,$$

for almost all environmental sequences $\xi = (\xi_0, \xi_1, \dots)$.

The value of ρ^i has direct implications for the probability of extinction.³⁴ In the subcritical or critical cases where $\rho^i < 0$ or $\rho^i = 0$, respectively, the probability of extinction is one. However, in the supercritical case where $\rho^i > 0$, the probability of extinction is strictly less than one. In this case,

³¹ The theory of multitype branching processes with a random environment typically assumes merely that the sequence $\xi = (\xi_0, \xi_1, \dots)$ is stationary and ergodic. It also allows an infinite number of environmental states.

³² It is assumed that $p_0^i(\xi^s) + p_1^i(\xi^s) < 1$ and $p_j^i(\xi^s) \neq 1$, for all i, j and ξ^s .

³³ "In isolation" means in the absence of mutation. Note that, in the special case where $S=1$, $\rho^i = \ln(m^i)$, and the present criterion is equivalent to that already obtained in the fixed environment model.

³⁴ These implications are analogous to those deriving from the mean for a single type branching process with a fixed environment.

which is central for the present purpose, non-extinction implies that growth to infinity occurs.³⁵

Consider then the case that $\rho^i > 0$ in greater detail. Consider the population of type i in generation T , $z_T(i)$, say. In analogy to the results for a fixed environment, it can be shown that $z_T(i)/\prod_{t=0}^{T-1} m^i(\xi_t) \rightarrow W$, with probability 1, as $T \rightarrow \infty$. Here W is again a non-negative random variable and $W=0$ essentially if and only if extinction occurs.³⁶ Conditional on non-extinction, it follows that

$$\frac{1}{T} \ln(z_T(i)) \rightarrow \rho^i = \lim_{T \rightarrow \infty} \left[\frac{1}{T} \sum_{t=0}^{T-1} \ln(m^i(\xi_t)) \right],$$

with probability one, as $T \rightarrow \infty$. That is, ρ^i is a measure of the potential long run growth rate of type $i = 1, \dots, G$.

Type i might dominate type j , despite $\rho^j > \rho^i$ and $\rho^j > 0$, because type j happens to become extinct but type i does not. To prevent this, mutation is now introduced. There is again a probability $\lambda_{ij} > 0$ that each offspring of a type i individual mutates to each type $j \neq i$. With remaining probability $A_i = (1 - \sum_{j \neq i} \lambda_{ij}) > 0$, each offspring remains of type i . The vector of all such λ_{ij} , $i, j = 1, \dots, G$, $i \neq j$, is again denoted by $\lambda \in \mathbf{R}_{++}^{G(G-1)}$. These mutations are independent of all earlier events.

All types must now be considered together as a multitype branching process with a random environment. It is again remarkable that the matrices of the means for this process suffice for the present purposes. Given the realization ξ^s , the associated matrix is

$$M_{ij}(\xi^s) = \begin{cases} \lambda_{ij} m^i(\xi^s), & \text{if } i \neq j, i, j = 1, \dots, G, \\ A_i m^i(\xi^s), & \text{if } i = j, i = 1, \dots, G, \end{cases}$$

where $m^i(\xi^s) = \sum_{j=1}^G j \cdot p_j^i(\xi^s) \in (0, \infty)$, for all i and s . The random vector representing the population of the G types in generation T is $z_T = (z_T(1), \dots, z_T(G))$, for $T = 1, 2, \dots$. It follows that the mean of z_T conditional on the sequence $\xi = (\xi_0, \xi_1, \dots)$ and the initial population vector z_0 , is $E(z_T | \xi, z_0) = z_0 M(\xi_0) \cdots M(\xi_{T-1})$.

For this multitype branching process with a random environment, the following elegant generalizations of the results for the single type case are available. Where the norm $\|\cdot\|$ is the sum of the elements of a (non-negative) matrix, it follows that there exists a (finite) number $E(\lambda)$,

³⁵ See Athreya and Ney [1, p. 255, 1a) and 1b)].

³⁶ See Athreya and Ney [1, p. 256, Theorem 4].

say, such that $\lim_{T \rightarrow \infty} (1/T) \ln \|M(\xi_0) \cdots M(\xi_{T-1})\| = E(\lambda)$, with probability 1.³⁷ Furthermore, $E(\lambda)$ is the potential long-run growth rate of the total population. If $E(\lambda) < 0$ then ultimate extinction is with probability 1. If $E(\lambda) > 0$, then there exists a positive probability that the population never becomes extinct. In the event of non-extinction when $E(\lambda) > 0$, $\lim_{T \rightarrow \infty} (1/T) \ln \|z_T\| = E(\lambda)$ with probability 1.³⁸

5.3. An Example—Mutation Can Increase Growth

Before presenting the main result of this section, Theorem 2, it is of interest to note that one result of Theorem 1 no longer holds here. That is, mutation may serve the interest of the species as a whole by raising the potential long run growth rate.³⁹

EXAMPLE 6. Suppose there are two types and there are two environments, each occurring with probability $1/2$. Suppose that the mean matrices are given by

$$M(\xi^1) = \begin{bmatrix} 32(1-\lambda) & 32\lambda \\ 2\mu & 2(1-\mu) \end{bmatrix} \quad \text{and} \quad M(\xi^2) = \begin{bmatrix} 2(1-\lambda) & 2\lambda \\ 32\mu & 32(1-\mu) \end{bmatrix},$$

where $\lambda = \lambda_{12}$ and $\mu = \mu_{21}$. When the mutation rates $\lambda = \mu = 0$, it follows that the long run growth rate is $E(0, 0) = \rho^1 = \rho^2 = \ln(8)$. On the other hand, if $\lambda = \mu = 1/2$, the product of any sequence of T matrices of the above two types has some element growing at least like 16^T . Hence now the constant $E(1/2, 1/2) \geq \ln(16) > \ln(8)$.

5.4. The Selection of the Favored Type

Natural selection again concerns the composition of the population, when unlimited growth occurs and given that mutation rates are small. There is now no analogous result to that of Theorem 1 concerning convergence of the vector of proportions to a non-stochastic eigenvector. Environmental uncertainty implies that the vector of proportions is typically stochastic even in the long run. Nevertheless, Theorem 2 shows that the type with the highest value of $\rho^i > 0$ ultimately dominates the population, conditional on non-extinction and when the mutation rates are small.

³⁷ That is, this limit exists and is constant over ξ , for almost all ξ . This was shown by Furstenberg and Kesten. See Kingman [13] and Tanny [21, Section 7].

³⁸ See Tanny [21, Theorem 9.10]. If $E(\lambda) = 0$, an additional hypothesis is needed to ensure that the probability of extinction is 1.

³⁹ Ben Porath *et al.* [2] make an analogous point in a related model.

THEOREM 2. *Consider the model of Section 5.2. Assume that $\rho^1 > \rho^j$, for all $j \neq 1$.⁴⁰*

(i) *If $\rho^1 < 0$ and the mutation vector $\lambda > 0$ is small enough, it follows that $E(\lambda) < 0$. The population then ultimately becomes extinct with probability 1, for any initial population vector, z_0 .⁴¹*

(ii) *If $\rho^1 > 0$ and the mutation vector $\lambda > 0$ is small enough, then $E(\lambda) > 0$. Now there exists a positive probability that the population avoids extinction forever, for any initial population vector, z_0 . Further, $\lim_{T \rightarrow \infty} (1/T) \ln \|z_T(\xi)\| = E(\lambda)$, with probability 1, given that $z_T(\xi) \neq 0$ for $T = 1, \dots$*

(iii) *Given $\rho^1 > 0$, it also follows that type 1 is selected when mutation rates are small, a long time has elapsed, and the population is found to have avoided extinction. To be precise, for all $\bar{\varepsilon} > 0$, $\bar{\delta} > 0$, there exists $\bar{\lambda} \in (0, 1)$ such that, for all mutation rates $\lambda_{ij} \in (0, \bar{\lambda})$, $i, j = 1, \dots, G$, $i \neq j$, there exists an integer $\bar{T}(\lambda)$ such that $T \geq \bar{T}(\lambda)$ implies*

$$\Pr \left\{ \frac{z_T(2) + \dots + z_T(G)}{z_T(1) + \dots + z_T(G)} < \bar{\delta} \mid z_T \neq 0 \right\} > 1 - \bar{\varepsilon},$$

where $z_T = (z_T(1), \dots, z_T(G))$ is the population vector at time T .

Proof. See the Appendix. Although (iii) is a plausible result, the possibility of type 1 mutating into the other slower growing types makes it more difficult to prove.

6. NON-EXPECTED UTILITY

6.1. Gambles over Commodities

This section shows how the result of the previous section translates into preferences for gambles over economic commodities. The interesting case involves the environmental parameter being part of the description of the gambles over commodities.

First suppose that possession of any particular commodity bundle $x \in \mathbf{R}_+^n$ induces for any individual a distribution over offspring represented by $p_0(x), p_1(x), \dots, p_B(x) \geq 0$, where $\sum_{j=0}^B p_j(x) = 1$. That is, $p_j(x)$ is the probability that bundle x entails j (surviving) offspring.⁴² All types are

⁴⁰ Example 3 still applies to demonstrate the possible consequences of non-uniqueness of the type maximizing ρ^i .

⁴¹ In the non-generic case that $\rho^1 = 0$, it can be shown that the probability of ultimate extinction tends to 1 as $\lambda \rightarrow 0$.

⁴² It is again assumed that $p_0(x) + p_1(x) < 1$ and $p_j(x) \neq 1$, for all j and all $x \in \mathbf{R}_+^n$.

identical in this respect and these distributions are independent over individuals. Define the expected offspring production function, again, as $\psi: \mathbf{R}_+^n \rightarrow \mathbf{R}_+$, $\psi(x) = \sum_{j=1}^B j \cdot p_j(x)$, for each $x \in \mathbf{R}_+^n$.

Suppose now that various gambles over commodity bundles are available and each type of individual is characterized by some such gamble. These gambles exhibit statistical dependence as follows. First, the environmental parameter, ξ_t , $t=0, 1, \dots$, is generated. Recall that ξ_t is i.i.d. having realization ξ^s with probability $\pi^s > 0$, for $s=1, \dots, S$, where $\sum_{s=1}^S \pi^s = 1$. Given each realization of this parameter, each type of individual faces a particular finite gamble over commodity bundles. These gambles are independent across all individuals *given* the environment. Suppose that $q_k^{i,s} \in [0, 1]$ is the probability that bundle $x_k^{i,s} \in \mathbf{R}_+^n$ is attained in environment ξ^s by type i . This holds for $k=1, \dots, K$, where $\sum_{k=1}^K q_k^{i,s} = 1$, for $s=1, \dots, S$ and $i=1, \dots, G$.⁴³ Consider the criterion:

$$\sum_{s=1}^S \pi^s \ln \left[\sum_{k=1}^K \sum_{j=1}^B j q_k^{i,s} p_j(x_k^{i,s}) \right] = \sum_{s=1}^S \pi^s \ln \left[\sum_{k=1}^K q_k^{i,s} \psi(x_k^{i,s}) \right], \quad i=1, \dots, G.$$

Theorem 2(iii) shows that the type taking the gamble over commodities which maximizes this criterion will ultimately dominate the population.⁴⁴

The left-hand side above explicitly concerns the distribution of offspring, as dependent on environmental uncertainty as well as on risk in each environment. Recall that it is hypothesized here that preferences were hard-wired over gambles involving commodities, rather than formed directly over offspring. Then the preferences of modern individuals would be represented by the right-hand side, in which the function ψ (or a multiple of ψ , perhaps) would appear as a primitive.

6.2. The Interpretation of Non-expected Utility

The criterion of the previous section is not a function of the marginal probabilities facing an individual and so is an unusual form of non-expected utility. The essential distinction here between idiosyncratic risk and environmental uncertainty derives from their differing implications for individuals *other* than the given individual. The present biological approach then leads to results which may be inconsistent with individual rationality. The following example involving gambles over commodities indeed shows that the present criterion implies choice of a first-order stochastically dominated gamble.

⁴³ K can be taken to be independent of i and s .

⁴⁴ This highest value must be greater than zero. However, this property would be lost if the function ψ were multiplied by a constant, which has no effect on observed attitudes to risk.

EXAMPLE 7. All individuals select one of two gambles. The first gamble yields a wealth level of either w_1 or w_2 , where $w_1 > w_2$, say. In this gamble, w_1 occurs with probability p and w_2 occurs with probability $(1 - p)$, independently across all individuals. The second of these gambles also entails either w_1 or w_2 , but this consequence is the same for all individuals at a given date. This common consequence is chosen independently at each date with a probability of $1/2$ for each possibility. Suppose that the function ψ , as in Section 6.1, is a function of wealth alone. If $p = 1/2$, Jensen's inequality yields that

$$\ln[(1/2) \psi(w_1) + (1/2) \psi(w_2)] > (1/2) \ln(\psi(w_1)) + (1/2) \ln(\psi(w_2)),$$

given only that ψ is strictly monotonic. Individuals are thus selected to opt for the first gamble involving idiosyncratic risk instead of the second gamble involving aggregate uncertainty, despite their equivalence from an individual viewpoint.

This example can be sharpened as follows. Since the inequality here is strict, it still holds if p is less than $1/2$, but remains sufficiently close to $1/2$. Choice of the first gamble is then still dictated, despite the strict first-order stochastic domination of the first gamble by the second gamble from an individual viewpoint. Hence evolutionarily induced attitudes to risk need not conform to a weak criterion of individual rationality.

A crucial issue here is whether actions are determined genetically so that an individual's choice is tied to those of her descendants. After all, if an individual chooses *independently* of her offspring, the stochastic dominance of the second gamble remains a compelling reason to select it. All individuals would then choose this way. When these choices are genetically tied together, on the other hand, it is as if the gamble involving idiosyncratic risk is selected by an individual because this maximizes the ultimate growth rate of the descendants of that individual. It is enough for there to be the *possibility* of genetic control, whether or not it is also possible to choose independently. That is, a population choosing the second gamble for any reason could be invaded by mutants whose choice of the first gamble was genetically determined.

Recall that interactions within the family produce a roughly analogous conflict between individual rationality and genetic control.⁴⁵ Consider a simple model in which each individual produces a variable number of offspring who play a generalized one-shot prisoner's dilemma among themselves. If the choice of a given individual is independent of those of her descendants, non-cooperation would generate more of these descendants, for any fixed pattern of choice by them. Non-cooperation is

⁴⁵ The strategic element in family interactions is lacking in the present model, however.

then individually rational in this sense. Nevertheless, maximal long-run growth of the descendants of a particular individual is ensured by cooperation among them. Cooperation is then the better choice if behavior is genetically determined so the choice of a given individual must match that of her descendants. Since some degree of cooperative behavior among close relatives is a stylized fact, genetic determination of such family interactions should be taken seriously.

There is no obvious reason why genetic determination of attitudes to risk could not also occur. There is an empirical issue here: Is any behavior toward risk observed which seems inconsistent with individual rationality but which might be consistent with long-run biological advantage?

Attitudes to health risks may provide such examples. Such risks have direct biological effects and require no translation into gambles over commodities. Viscusi [22], for example, suggests that the Tylenol tampering incidents of the early 1980's provoked an overreaction in the form of reduced sales, whereas individuals underreact to the possibility of automobile accidents by failing to wear seat belts often enough. Since the raw probability of harm in the Tylenol case seems to be smaller than that in the seat belt case, this can be described by Viscusi in terms of a systematic flattening of the relationship between perceived and actual probabilities. However, the Tylenol case involved significant aggregate uncertainty, since some possible mechanisms for altering the bottles would have had simultaneous implications for a large number of people. This would be much less so in the seat belt case. If the risk in the Tylenol case was largely aggregate but that for the seatbelt case was largely idiosyncratic, the present approach could explain the divergence in behavior.

There are other comparisons involving health risks which might have a similar rationale. Consider, for example, the heightened concern with AIDS, as compared to heart disease, say, when AIDS first came to public attention. Although the risk from heart disease seems largely idiosyncratic, there was initially an important component of aggregate uncertainty about AIDS, since it was unclear how infectious it would be. Hence the present approach is again capable of resolving the apparent anomaly.

The present approach might also shed light on some examples which seem to suggest a Knightian distinction between risk and uncertainty. For example, in the cases discussed above of Tylenol and of AIDS, the aggregate uncertainty might be reinterpreted as a lack of precision concerning the risk facing an individual. The greater aversion to aggregate uncertainty than to idiosyncratic risk derived under the present approach is then consistent with Knight's suggestion that individuals would be averse to ignorance of relevant probabilities. (See Knight [14].)

The empirical importance of the present distinction between aggregate uncertainty and idiosyncratic risk remains to be fully explored. Perhaps

the most significant point here is theoretical. This is that the notion that biological evolution will *automatically* provide a foundation for individually rational attitudes to risk seems to be mistaken.

7. FURTHER RELATED WORK

First, note that the present work can be interpreted in terms of evolutionary theory as an investigation of the validity of the expected offspring criterion. Maynard Smith [16], for example, adopts this criterion as the basis of evolutionary game theory. It would be useful to extend the current approach to a strategic setting. This would again involve small probabilities of mutation and a population which could be small at the outset. It might also involve a fully random pairing mechanism. This approach would then contrast with Boylan [4], who examines directly the properties of an infinite population model.

The first result of the present paper, that the expected utility theorem holds in a fixed environment, is related to work by Cooper [6]. He considers a number of environments, where the choice of an idiosyncratic gamble is made *after* the environmental outcome is realized. Fitness in each environment is identified with expected surviving offspring, on the basis of an implicit appeal to the law of large numbers. It follows readily that the favored type will be that which maximizes fitness in each and every environment.⁴⁶ Cooper shows how this criterion might be interpreted as satisfying most of the Savage axioms.

Cooper and Kaplan [7] consider a simple model in which individuals also face environmental uncertainty in the present sense. The criterion they apply is identical to that derived here—the expectation over the environment of the logarithm of expected offspring conditional on the environment. Cooper and Kaplan are concerned to show that such a criterion might favor a type of individual which generated its own idiosyncratic risk. For example, suppose there are originally two types of animal—one which turns white in the fall and is optimal for a winter with snow and another type which does not change color and is then optimal for a winter without snow. Both of these types are then subject to aggregate uncertainty concerning the weather. Suppose that the probability of winter with or without snow is $1/2$, for example. It follows that a new type which randomly and independently chooses whether to change color in the fall, making each choice with probability $1/2$, will do the best of all three types.

⁴⁶ It is assumed that such a type exists, that is, which it need not in the model of the present paper.

Such a third type benefits from the replacement of the original aggregate uncertainty with idiosyncratic risk.

Within the literature on attitudes to risk, the present work is also in the spirit of Marshak [15] who conjectured that expected utility might be motivated as satisfying a "rule of long run success." Marshak's approach involves consumption of infinite sequences of prizes. Camacho [5] proves a result along these lines, where the law of large numbers is used to obtain the linearity of expected utility in probabilities. This is then mathematically reminiscent of the approach here. On the other hand, Marshak and Camacho postulate that preferences are given as a primitive over the infinite sequences, whereas the present model considers the evolutionary *formation* of preferences.

Related results have also been obtained in models of investment over time. Suppose the stock market generates aggregate uncertainty but no idiosyncratic risk and the goal is simply to maximize wealth at a date far into the future. The appropriate myopic criterion involves a logarithmic von Neumann–Morgenstern utility function. Equivalently, it is the expected growth rate of the portfolio. Samuelson [18] derives this result, but also points out the artificial nature of the objective in the light of general intertemporal preferences. More recently, Blume and Easley [3] reexamine a model which is an elaboration and refinement of Samuelson's. They show, for example, that the long-run behavior of prices is driven by investors who maximize the expected growth rate. This follows since these investors will asymptotically hold all the wealth.

It might, finally, be noted that the present criterion is formally similar to work by Segal [19]. He relaxes the reduction to simple gambles assumption of expected utility theory and considers an anticipated utility representation of preferences. There are two sequential stages, but usually an assumption of time neutrality removes effects due to the timing of the resolution of uncertainty. This assumption is not formally satisfied in the present case if it is assumed that the first stage corresponds to the selection of the environmental parameter and the second stage to the occurrence of the risk given the environment. The distinction here is conceptually unrelated to any distinction based on sequential stages, arising instead from the public or private incidence of the uncertainty or risk.

APPENDIX

Proof of Theorem 1. (i) This follows since $m(\lambda)$ is strictly less than the maximal row sum, given only that not all row sums are identical. (See Takayama [20, p. 388, Theorem 4.C.10].)

(ii) Since $m(\lambda) < m^1 \leq 1$, that the probability of ultimate extinction is 1 follows from Athreya and Ney [1, Chapter V, Theorem 3.2.].

(iii) Suppose now that $m^1 > 1$. Take the characteristic equation for the matrix $M(\lambda)$, as, say, $P(m, \lambda) \equiv m^G + a_{G-1}(\lambda) m^{G-1} + \dots + a_0(\lambda) = 0$. Clearly, $P(m, 0) \equiv \prod_{i=1}^G (m - m^i) = 0$ has a simple real root, m^1 , strictly exceeding all the other real roots m^i , $i = 2, \dots, G$. It follows readily that $m(\lambda) \rightarrow m^1$, as $\lambda \rightarrow 0$, so that $m(\lambda) > 1$, if λ is small enough. With the exception of the last assertion, the conclusions of (iii) follow from Athreya and Ney [1, Chapter V, Theorems 3.2, 6.1, and 6.2].

Consider any sequence $\lambda_n \rightarrow 0$ such that $v(\lambda_n) \rightarrow \bar{v}$, where $v(\lambda_n)$ and \bar{v} belong to the unit simplex. Since $m(\lambda_n) \rightarrow m^1$, as $\lambda_n \rightarrow 0$, it follows readily that $\bar{v} = e_1$, as required.

Proof of Theorem 2. (i) Recall that $E(\lambda) = \lim_{T \rightarrow \infty} (1/T) \ln \|M(\xi_0, \lambda) \dots M(\xi_{T-1}, \lambda)\|$, which limit exists and is constant with probability 1, for all $\lambda \geq 0$.⁴⁷ The norm $\|\cdot\|$ is the sum of the absolute values of the elements of matrices (or vectors) and is submultiplicative in that, for all conformable matrices A and B , $\|AB\| \leq \|A\| \cdot \|B\|$. It follows that, for all integers R and T ,

$$\begin{aligned} & 1/(RT) \ln \|M(\xi_0, \lambda) \dots M(\xi_{RT-1}, \lambda)\| \\ & \leq (1/R) \{ (1/T) \ln \|M(\xi_0, \lambda) \dots M(\xi_{T-1}, \lambda)\| \\ & \quad + \dots + (1/T) \ln \|M(\xi_{(R-1)T}, \lambda) \dots M(\xi_{RT-1}, \lambda)\| \}. \end{aligned}$$

Hence, taking the limit as $R \rightarrow \infty$, the strong law of large numbers implies that $E(\lambda) \leq E\{(1/T) \ln \|M(\xi_0, \lambda) \dots M(\xi_{T-1}, \lambda)\|\}$, since ξ_t is i.i.d. Moreover, $(1/T) \ln \|M(\xi_0, 0) \dots M(\xi_{T-1}, 0)\| \rightarrow \rho^1 < 0$, as $T \rightarrow \infty$, with probability 1. Hence there exists a T such that $E\{(1/T) \ln \|M(\xi_0, 0) \dots M(\xi_{T-1}, 0)\|\} < 0$. Now $\|M(\xi_0, \lambda) \dots M(\xi_{T-1}, \lambda)\|$ is continuous in λ , uniformly in ξ_0, \dots, ξ_{T-1} , so that, if the mutation vector $\lambda > 0$ is sufficiently small, $E(\lambda) \leq E\{(1/T) \ln \|M(\xi_0, \lambda) \dots M(\xi_{T-1}, \lambda)\|\} < 0$. The desired conclusion then follows from Tanny [21, Theorem 9.10(i)].

(ii) Suppose that $\rho^1 > 0$. If λ is small enough,

$$E(\lambda) \geq \lim_{T \rightarrow \infty} (1/T) \ln [A_1 m^1(\xi_0) \dots A_1 m^1(\xi_{T-1})] = \ln(A_1) + \rho^1 > 0.$$

The result then also follows from Tanny [21, Theorem 9.10(iii)].

⁴⁷ This limit is less than ∞ since the matrices $M(\xi_t, \lambda)$ are bounded. It is greater than $-\infty$ since $A_i > 0$, for $i = 1, \dots, G$. The dependence of $M(\xi_t, \lambda)$ on λ is noted here, and it is now convenient to allow $\lambda = 0$.

(iii) Consider $\bar{\xi}$ as the formal extension of ξ over all integers, Z . Define the $(G-1) \times (G-1)$ matrix $R_T(\bar{\xi}, \lambda) = I + \sum_{s=1}^{\infty} Q(\xi_{T-s}, \lambda) \cdots Q(\xi_{T-1}, \lambda)$, where I is the $(G-1) \times (G-1)$ identity and where $Q(\xi_t, \lambda)$ is also $(G-1) \times (G-1)$ and is given by

$$Q_{ij}(\xi_t, \lambda) = \begin{cases} \lambda_{ij} m^i(\xi_t) / A_1 m^1(\xi_t), & \text{for all } i \neq j, i, j = 2, \dots, G \\ A_i m^i(\xi_t) / A_1 m^1(\xi_t), & \text{for all } i = j, i = 2, \dots, G. \end{cases}$$

The distribution of $R_T(\bar{\xi}, \lambda)$ is clearly independent of T .

CLAIM 1. *If λ is small enough, then $\Pr\{R_T(\bar{\xi}, \lambda)_{ij} < \infty\} = 1$, for all $i, j = 2, \dots, G$.⁴⁸*

Proof. Kingman [13] implies that, since the matrices $Q(\xi_t, \lambda)$ are i.i.d. for all $\lambda \geq 0$, then $(1/s) \ln \|Q(\xi_{T-s}, \lambda) \cdots Q(\xi_{T-1}, \lambda)\| \rightarrow \tilde{E}(\lambda)$, as $s \rightarrow \infty$, with probability 1. That is, this limit exists and is independent of $\bar{\xi}$, for almost all $\bar{\xi}$, and all $\lambda \geq 0$.⁴⁹ If $\lambda > 0$, the matrices $Q(\xi_t, \lambda)$ have strictly positive entries and it also follows that

$$\frac{1}{s} \ln [(Q(\xi_{T-s}, \lambda) \cdots Q(\xi_{T-1}, \lambda))_{ij}] \rightarrow \tilde{E}(\lambda),$$

as $s \rightarrow \infty$, for all $i, j = 2, \dots, G$, with probability 1. If $\lambda = 0$, it can be shown that $\tilde{E}(0) = \max_{j \neq 1} \{\rho^j - \rho^1\} < 0$. Now an analogous argument to that in the proof of (i) yields that $\tilde{E}(\lambda) < 0$, if $\lambda > 0$ is small enough, and the desired result follow by the root test. ■

From (ii), it follows that $\Pr\{\|z_t\| \rightarrow \infty\}$ is bounded away from zero as $\lambda \rightarrow 0$. Claim 1 implies that there exists $\bar{\lambda} \in (0, 1)$ such that, if $\lambda_{ij} \in (0, \bar{\lambda})$, for $i, j = 1, \dots, G$, $i \neq j$, then, for all T ,

$$1 \geq \Pr\{\ell R_T(\bar{\xi}, \lambda) < \bar{\delta} u / (2(G-1))\} > 1 - \bar{\varepsilon} \cdot \Pr\{\|z_t\| \rightarrow \infty\} / 2,$$

where $u = (1, \dots, 1)$, the $(G-1)$ row-vector of 1's, ℓ is the $(G-1)$ row-vector with $\ell_i = \lambda_{1i} / A_1$, $i = 2, \dots, G$, and $\bar{\varepsilon}$ and $\bar{\delta}$ are as in the statement of Theorem 2(iii). Since $\{\|z_t\| \rightarrow \infty\} \subseteq \{z_T \neq 0\}$, for all T , it follows that $\Pr\{\ell R_T(\bar{\xi}, \lambda) < \bar{\delta} u / (2(G-1)) \mid z_T \neq 0\} > 1 - \bar{\varepsilon} / 2$. Fix from now on any particular such mutation vector $\lambda > 0$.

Lemma 1 then serves to complete the proof of Theorem 2(iii).⁵⁰

⁴⁸ Explicit reference to the underlying sample space, Ω , say, is omitted in this Appendix. For example, $\{R_T(\bar{\xi}, \lambda)_{ij} < \infty\}$ is an abbreviation for $\{\omega \in \Omega: R_T(\bar{\xi}, \lambda)_{ij}(\omega) < \infty\}$.

⁴⁹ This limit is less than ∞ since the $Q(\xi_t, \lambda)$ are bounded above. It is greater than $-\infty$ since $A_i > 0$, for $i = 1, \dots, G$. It is again convenient to allow $\lambda = 0$.

⁵⁰ Set $\varepsilon = \bar{\varepsilon} / 2$ and $\delta = \bar{\delta} / [2(G-1)]$ in Lemma 1. Note that $f_T < \delta u / (G-1)$ implies $\sum_{i=2}^G f_T(i) < \delta$.

LEMMA 1. For all $\varepsilon > 0$, $\delta > 0$, there exists an integer $\bar{T}(\lambda)$ such that $T \geq \bar{T}(\lambda)$ implies $\Pr \{f_T - \ell R_T(\bar{\xi}, \lambda) < \delta u \mid z_T \neq 0\} > 1 - \varepsilon$, where u and ℓ are as above and $f_T(i) = z_T(i)/z_T(1) \geq z_T(i)/(z_T(1) + \dots + z_T(G))$, $i = 2, \dots, G$.

Proof. This is broken into the following claims.

CLAIM 2. For all $z = (z(1), \dots, z(G)) \geq 0$ there exists an integer $T_1(z)$ such that, if $T \geq T_1(z)$, then $\Pr\{z_T \geq z \mid z_T \neq 0\} > 1 - \varepsilon/5$.

Proof. This follows readily from Tanny [21, Theorem 9.10(iii)], since $\lambda > 0$ and $M_{ij}(\xi_T, \lambda)$, for $i, j = 1, \dots, G$ are then strictly positive.⁵¹

CLAIM 3. For any integer Δ , there exists $T_2(\Delta) \geq \Delta$ such that $T \geq T_2(\Delta)$ implies⁵² $\Pr\{z_T \neq 0, f_T - \bar{f}_{T, T-\Delta} < \delta u/2 \mid z_{T-\Delta} \neq 0\} > 1 - 2\varepsilon/5$, where $u = (1, \dots, 1)$ and $\bar{f}_{T, S}(i) = E(z_T(i) \mid \xi, z_S)/E(z_T(1) \mid \xi, z_S)$, for $T > S$ and $i = 2, \dots, G$.

Proof. Note that $z_T(i) = z_T^1(i) + \dots + z_T^G(i)$, where, for $i, j = 1, \dots, G$, $z_T^j(i)$ is the number of type i descendants at date $T \geq \Delta$ of type j 's at $T - \Delta \geq 0$. Hence,

$$\frac{z_T^j(i)}{z_{T-\Delta}(j)} \rightarrow_P e_j M(\xi_{T-\Delta}) \cdots M(\xi_{T-1}) e'_i,$$

as $z_{T-\Delta}(j) \rightarrow \infty$, $i, j = 1, \dots, G$, by the weak law of large numbers.⁵³ It follows that, for $i = 1, \dots, G$,

$$\frac{z_T(i)}{\|z_{T-\Delta}\|} \rightarrow_P \frac{z_{T-\Delta} M(\xi_{T-\Delta}) \cdots M(\xi_{T-1}) e'_i}{z_{T-\Delta}(1) + \dots + z_{T-\Delta}(G)},$$

as $(z_{T-\Delta}(1), \dots, z_{T-\Delta}(G)) \rightarrow (\infty, \dots, \infty)$.⁵⁴ Hence, for $i = 2, \dots, G$,

$$f_T(i) = \frac{z_T(i)/\|z_{T-\Delta}\|}{z_T(1)/\|z_{T-\Delta}\|} \rightarrow_P \frac{z_{T-\Delta} M(\xi_{T-\Delta}) \cdots M(\xi_{T-1}) e'_i}{z_{T-\Delta} M(\xi_{T-\Delta}) \cdots M(\xi_{T-1}) e'_1} = \bar{f}_{T, T-\Delta}(i),$$

⁵¹ Since $\lambda > 0$ is now fixed, it is convenient in the following to suppress it as an independent variable.

⁵² The notation " $z_T \neq 0, f_T - \bar{f}_{T, T-\Delta} < \delta u/2$ " refers, for example, to the event $\{z_T \neq 0\} \cap \{f_T - \bar{f}_{T, T-\Delta} < \delta u/2\}$.

⁵³ " e_i " denotes the unit row vector with 1 in the i th position and zeros elsewhere. Primes denote associated column vectors and " \rightarrow_P " denotes convergence in probability.

⁵⁴ The limit here is bounded away from 0 and ∞ since $M_{ij}(\xi^s) > 0$, for all $i, j = 1, \dots, G$ and $s = 1, \dots, S$. Convergence is uniform in the realizations of the random variables ξ and $z_{T-\Delta}$.

as $(z_{T-\Delta}(1), \dots, z_{T-\Delta}(G)) \rightarrow (\infty, \dots, \infty)$. Hence there exists $\bar{z} = (\bar{z}(1), \dots, \bar{z}(G))$ such that $\Pr\{z_T \neq 0, f_T - \bar{f}_{T, T-\Delta} < \delta u/2 \mid z_{T-\Delta} \geq \bar{z}\} > 1 - \varepsilon/5$. If $T \geq T_2(\Delta) = T_1(\bar{z}) + \Delta$, Claim 2 implies $\Pr\{z_{T-\Delta} \geq \bar{z} \mid z_{T-\Delta} \neq 0\} > 1 - \varepsilon/5$, so that $\Pr\{z_T \neq 0, f_T - \bar{f}_{T, T-\Delta} < \delta u/2 \mid z_{T-\Delta} \neq 0\} > 1 - 2\varepsilon/5$.⁵⁵ ■

CLAIM 4. *There exists an integer T_3 and $k < \infty$ such that $T \geq T_3$ implies $\Pr\{z_T \neq 0, f_T < ku \mid z_{T-1} \neq 0\} > 1 - 2\varepsilon/5$, where $u = (1, \dots, 1)$.*

Proof. Argue as in the proof of Claim 3 that, for $i = 2, \dots, G$, $f_T(i) \rightarrow_P (z_{T-1} M(\xi_{T-1}) e'_i) / (z_{T-1} M(\xi_{T-1}) e'_1)$, as $(z_{T-1}(1), \dots, z_{T-1}(G)) \rightarrow (\infty, \dots, \infty)$. It follows that there exists $k < \infty$ such that $((z_{T-1} M(\xi_{T-1}) e'_i) / (z_{T-1} M(\xi_{T-1}) e'_1)) \leq k/2$, say, for all z_{T-1} and all ξ_{T-1} , $i = 2, \dots, G$. Hence there exists $\tilde{z} = (\tilde{z}(1), \dots, \tilde{z}(G))$ such that $\Pr\{z_T \neq 0, f_T < ku \mid z_{T-1} \geq \tilde{z}\} > 1 - \varepsilon/5$. By Claim 2, if $T \geq T_1(\tilde{z}) + 1 = T_3$, say, then $\Pr\{z_{T-1} \geq \tilde{z} \mid z_{T-1} \neq 0\} > 1 - \varepsilon/5$ so that $\Pr\{z_T \neq 0, f_T < ku \mid z_{T-1} \neq 0\} > 1 - 2\varepsilon/5$.⁵⁶

CLAIM 5. *There exists an integer \bar{A} such that for all $T \geq T_3 + \bar{A}$,*

$$\Pr\{z_{T-\bar{A}} \neq 0, \bar{f}_{T, T-\bar{A}} - \ell R_T(\bar{\xi}) < \delta u/2 \mid z_{T-\bar{A}-1} \neq 0\} > 1 - 3\varepsilon/5,$$

where $\bar{f}_{T, S}$ is as in the statement of Claim 3, for all $T > S$, where T_3 is as in Claim 4, $u = (1, \dots, 1)$ and $\ell_i = \lambda_{1i}/A_1$, $i = 2, \dots, G$.

Proof. Take any $S \geq T_3$ and $T > S$. Note that, for $i = 2, \dots, G$, $\bar{f}_{T, S}(i) = (z_S M(\xi_S) \cdots M(\xi_{T-1}) e'_i) / (z_S M(\xi_S) \cdots M(\xi_{T-1}) e'_1)$, so that $\bar{f}_{T+1, S}(i) = ([1, \bar{f}_{T, S}] M(\xi_T) e'_i) / ([1, \bar{f}_{T, S}] M(\xi_T) e'_1)$ and $\bar{f}_{T+1, S} \leq \ell + \bar{f}_{T, S} Q(\xi_T)$, where the matrix $Q(\xi_t) \equiv Q(\xi_t, \lambda)$ is defined prior to Claim 1. Also $\bar{f}_{S+1, S} \leq \ell + f_S Q(\xi_S)$. It follows that $\bar{f}_{T, S} \leq \ell R_T(\bar{\xi}) + f_S Q(\xi_S) \cdots Q(\xi_{T-1})$, for all $T > S$, where $R_T(\bar{\xi}) \equiv R_T(\bar{\xi}, \lambda)$ is also defined prior to Claim 1. Fix $\bar{A} = T - S \geq 1$ such that, in the light of Claim 1,

$$\Pr\{ku Q(\xi_{T-\bar{A}}) \cdots Q(\xi_{T-1}) \leq \delta u/2\} > 1 - \varepsilon/5.$$

From Claim 4, since $T - \bar{A} \geq T_3$, then

$$\Pr\{z_{T-\bar{A}} \neq 0, f_{T-\bar{A}} < ku \mid z_{T-\bar{A}-1} \neq 0\} > 1 - 2\varepsilon/5.$$

⁵⁵ Note that $\{z_{T-\Delta} \geq \bar{z}\} \subseteq \{z_{T-\Delta} \neq 0\}$.

⁵⁶ Note that $\{z_{T-1} \geq \tilde{z}\} \subseteq \{z_{T-1} \neq 0\}$.

From the independence of ξ ,

$$\begin{aligned} & \Pr\{kuQ(\xi_{T-\bar{A}}) \cdots Q(\xi_{T-1}) < \delta u/2 \mid z_{T-\bar{A}} \neq 0, f_{T-\bar{A}} < ku\} \\ &= \Pr\{kuQ(\xi_{T-\bar{A}}) \cdots Q(\xi_{T-1}) < \delta u/2\}. \end{aligned}$$

It follows that, if $T \geq T_3 + \bar{A}$, then⁵⁷

$$\begin{aligned} & \Pr\{z_{T-\bar{A}} \neq 0, \bar{f}_{T, T-\bar{A}} - \ell R_T(\bar{\xi}) < \delta u/2 \mid z_{T-\bar{A}-1} \neq 0\} \\ & \geq \frac{\Pr\{z_{T-\bar{A}} \neq 0, f_{T-\bar{A}} < ku, kuQ(\xi_{T-\bar{A}}) \cdots Q(\xi_{T-1}) < \delta u/2\}}{\Pr\{z_{T-\bar{A}-1} \neq 0\}} \\ &= \frac{\Pr\{kuQ(\xi_{T-\bar{A}}) \cdots Q(\xi_{T-1}) < \delta u/2\} \cdot \Pr\{z_{T-\bar{A}} \neq 0, f_{T-\bar{A}} < ku\}}{\Pr\{z_{T-\bar{A}-1} \neq 0\}} \\ & > 1 - 3\varepsilon/5. \quad \blacksquare \end{aligned}$$

Finally, then, define $\bar{T}(\lambda) = \max\{T_2(\bar{A}), T_3 + \bar{A}\}$, as in Claims 3 and 5. These two claims then imply that, for all $T \geq \bar{T}(\lambda)$,

$$\begin{aligned} & \Pr\{f_T - \ell R_T(\bar{\xi}) < \delta u \mid z_T \neq 0\} \\ & \geq \Pr\{f_T - \bar{f}_{T, T-\bar{A}} < \delta u/2, \bar{f}_{T, T-\bar{A}} - \ell R_T(\bar{\xi}) < \delta u/2 \mid z_T \neq 0\} \\ & \geq \frac{\Pr\{z_T \neq 0, f_T - \bar{f}_{T, T-\bar{A}} < \delta u/2, \bar{f}_{T, T-\bar{A}} - \ell R_T(\bar{\xi}) < \delta u/2\}}{\Pr\{z_{T-\bar{A}} \neq 0\}} \\ & \geq -1 + \Pr\{z_T \neq 0, f_T - \bar{f}_{T, T-\bar{A}} < \delta u/2\} / \Pr\{z_{T-\bar{A}} \neq 0\} \\ & \quad + \Pr\{z_{T-\bar{A}} \neq 0, \bar{f}_{T, T-\bar{A}} - \ell R_T(\bar{\xi}) < \delta u/2\} / \Pr\{z_{T-\bar{A}-1} \neq 0\} \\ & > 1 - \varepsilon, \end{aligned}$$

as required to prove Lemma 1 and hence to complete the proof of Theorem 2.⁵⁸

REFERENCES

1. K. B. ATHREYA AND P. E. NEY, "Branching Processes," Springer-Verlag, Berlin, 1972.
2. E. P. BEN PORATH, E. DEKEL, AND A. RUSTICHINI, On the relationship between mutation rates and growth rates in a changing environment, *Games Econ. Behav.* **5** (1992), 576–603.
3. L. BLUME AND D. EASLEY, Evolution and market behavior, *J. Econ. Theory* **58** (1992), 9–40.

⁵⁷ Note that $\{z_{T-\bar{A}} \neq 0\} \subseteq \{z_{T-\bar{A}-1} \neq 0\}$.

⁵⁸ Note that $\{z_T \neq 0\} \subseteq \{z_{T-\bar{A}} \neq 0\} \subseteq \{z_{T-\bar{A}-1} \neq 0\}$.

4. R. T. BOYLAN, Laws of large numbers for dynamical systems with randomly matched individuals, *J. Econ. Theory* **57** (1992), 473–504.
5. A. CAMACHO, Maximizing expected utility and the rule of long run success, in “Expected Utility Hypotheses and the Allais Paradox” (M. Allais and O. Hagen, Eds.), Reidel, Dordrecht, Holland, 1979.
6. W. S. COOPER, Decision theory as a branch of evolutionary theory: A biological derivation of the Savage axioms, *Psych. Rev.* **4** (1987), 395–411.
7. W. S. COOPER AND R. H. KAPLAN, Adaptive “coin-flipping”: A decision-theoretic examination of natural selection for random individual variation, *J. Theor. Biol.* **94** (1982), 135–151.
8. F. R. GANTMACHER, “Matrix Theory, Vol. II,” Chelsea, New York, 1964.
9. J. C. HARSANYI, Cardinal welfare, individualistic ethics and interpersonal comparisons of utility, *J. Polit. Econ.* **63** (1955), 309–321.
10. F. A. HAYEK, “Law, Legislation, and Liberty, Vol. 3,” Univ. of Chicago Press, Chicago, 1979.
11. J. HIRSHLEIFER, Natural economy versus political economy, *J. Soc. Biol. Structures* **1** (1978), 319–337.
12. E. KARNI AND D. SCHMEIDLER, Self-preservation as a foundation of rational behavior under risk, *J. Econ. Behav. Organ.* **7** (1986), 71–82.
13. J. F. C. KINGMAN, Subadditive ergodic theory, *Ann. Prob.* **1** (1973), 883–909.
14. F. H. KNIGHT, “Risk, Uncertainty and Profit,” Univ. of Chicago Press, Chicago, 1921.
15. J. MARSHAK, Why “should” statisticians and businessmen maximize “moral expectation,” in “Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability” (J. Neyman, Ed.), Univ. of California, Berkeley, 1951.
16. J. MAYNARD SMITH, “Evolution and the Theory of Games,” Cambridge Univ. Press, Cambridge, 1982.
17. A. J. ROBSON, The evolution of attitudes to risk: Lottery tickets and relative wealth, *Games Econ. Behav.*, forthcoming, 1996.
18. P. A. SAMUELSON, The “fallacy” of maximizing the geometric mean in long sequences of investing or gambling, *Proc. Nat. Acad. Sci. USA* **68** (1971), 2493–2496.
19. U. SEGAL, Two-stage lotteries without the reduction axiom, *Econometrica* **58** (1990), 349–377.
20. A. TAKAYAMA, “Mathematical Economics,” Dryden Press, Hinsdale, Illinois, 1974.
21. D. TANNY, On multitype branching processes in a random environment, *Adv. Appl. Prob.* **13** (1981), 464–497.
22. W. K. VISCUSI, Sources of inconsistency in societal response to health risks, *Amer. Econ. Rev.* **80** (1990), 257–261.