


Evolved Attitudes to Idiosyncratic and Aggregate Risk in Age-Structured Populations*

Arthur J. Robson  Larry Samuelson

Department of Economics, Simon Fraser University, Burnaby, BC Canada
V5A 1S6

and Department of Economics, Yale University, New Haven CT 06520 USA
robson@sfu.ca and Larry.Samuelson@yale.edu

February 11, 2019

Abstract We examine the evolutionary foundations of risk attitudes in age-structured populations. The effect of idiosyncratic risk concerning fertility or mortality rates is captured by the corresponding mean fertility or mortality rate. The effect of aggregate risk, relative to the mean, varies with the type of risk and age. We establish conditions under which aggregate risk in fertility rates at young ages increases the population growth rate (compared to the corresponding mean fertility), but such risk in fertility rates at old ages reduces the growth rate. On the other hand, aggregate risk in mortality at young as well as old ages leads to increased growth rates. *JEL* codes B52 and D81.

Contents

1	Introduction	1
1.1	Idiosyncratic and Aggregate Risk	1
1.2	Preview	4
2	Idiosyncratic Risk	5
2.1	The Model	6
2.2	Risk Preferences	7
3	Aggregate Risk	8
3.1	An Approximation	8
3.2	Preferences Over Aggregate Risk	10
3.3	An Example	11
3.4	Age and Risk Attitudes	14
3.4.1	Rare Aggregate Shocks to Fertility	15
3.4.2	Rare Aggregate Shocks to Mortality	18
4	The Allocation of Resources	21
4.1	Optimal Resource Use	21
4.2	Shocks to Income	23
4.2.1	Idiosyncratic Shocks to Income	23
4.2.2	Aggregate Shocks to Income	24
5	Discussion	25
5.1	Life Beyond Fertility	25
5.2	Attitudes Toward Risk	25
6	Appendices	26
6.1	A Discrete Foundation	26
6.1.1	The Discrete Model	26
6.1.2	Taking the Limit	30
6.1.3	The Criterion for Evolutionary Success	32
6.2	Proof of Proposition 1	33
6.3	Calculations for Section 3.3	35
6.4	Relation to Robatto and Szentes	37

*We thank Phil Reny, Bill Sandholm, Balazs Szentes and Ken Wachter for helpful discussions and an associate editor and four referees for their comments. The paper also benefited from audience feedback during presentations at CERGE-EI, Penn State, SFU, the University of Lund, and Duke-UNC. Robson gratefully acknowledges support from a Canada Research Chair and from the Social Sciences and Humanities Research Council of Canada; Samuelson from the National Science Foundation (SES-1459158).

1 Introduction

A fruitful approach to sharpening our understanding of preferences is to consider their evolutionary foundations (see Robson and Samuelson [19] for a survey). We consider in this paper how aggregate and idiosyncratic risks may have distinct implications for risk aversion.

1.1 Idiosyncratic and Aggregate Risk

Robson [17] drew attention to evolutionary foundations by demonstrating that evolution may select for greater aversion to aggregate (correlated) risk than to idiosyncratic (independent) risk. To illustrate his argument, suppose that in each period each individual takes a draw from a random variable generating $x \in \{0, 1, \dots, X\}$ offspring with probability $m(x)$ and then dies. If these draws are independent across agents—the case of idiosyncratic risk—and the population is sufficiently large, then the law of large numbers will ensure that the average number of offspring each period is close to $\bar{x} = \sum_{x=0}^X xm(x)$. The size $N(t)$ of the population at time t will then be approximately

$$N(t) = N(0)\bar{x}^t$$

and the population growth rate will be given by

$$\lim_{t \rightarrow \infty} \frac{\ln N(t)}{t} = \ln \bar{x}.$$

Evolution will then select for preferences that maximize the expected number of offspring \bar{x} . We summarize this result by saying that idiosyncratic risks to the number of offspring are evolutionarily neutral.

Now suppose that in each period, the draws governing offspring are perfectly correlated across agents—the case of aggregate risk. Then the size $N(t)$ of the population at time t will be

$$N(t) = N(0) \prod_{x=0}^X x^{t(x)},$$

where $t(x)$ is the number of periods (from among $1, 2, \dots, t$) in which x is the number of offspring drawn. Then the growth rate is

$$\lim_{t \rightarrow \infty} \frac{\ln N(t)}{t} = \lim_{t \rightarrow \infty} \sum_{x=0}^X \frac{t(x)}{t} \ln x = \sum_{x=0}^X m(x) \ln x,$$

with the final equality again following from the law of large numbers. In this case, evolution will select for lotteries that maximize the expected log of off-

spring, building risk aversion to aggregate risks into preferences.¹ Aggregate risks to numbers of offspring are thus disadvantageous, in the sense that a population subject to such risk will grow more slowly than an analogous population in which the number of offspring is fixed at the mean of the aggregate risk.

To interpret the argument concerning growth rates in terms of individual choices, suppose, for example, that an individual faces a choice between two lotteries over consumption. Suppose that expected offspring is a function of consumption given by $u(c)$ for each consumption level c . In either of these lotteries, there is probability p of receiving bundle c_1 , and probability $1 - p$ of c_2 . In the first lottery, the risk is idiosyncratic, independent across individuals in each period; in the second lottery, the risk is aggregate, so all individuals receive the same consumption outcome. The growth rate of a population whose members all choose the idiosyncratic lottery is then $\ln[p u(c_1) + (1 - p) u(c_2)]$, which (since \ln is concave) is higher than the growth rate of a population whose members all choose the aggregate lottery, namely $p \ln u(c_1) + (1 - p) \ln u(c_2)$. Hence, evolution will select for preferences that are more averse to lotteries that entail aggregate risk than to those that entail idiosyncratic risk. More generally, it follows that the coefficient of absolute risk aversion for aggregate risks is given by $R_A(c) + \frac{u'(c)}{u(c)} > R_A(c)$ where $R_A(c) = -\frac{u''(c)}{u'(c)}$ is the coefficient of absolute risk aversion for idiosyncratic risks.²

Robatto and Szentes [16] casts a surprising and revealing light on these results. Suppose the population is described by a continuous rather than discrete time process. To keep things simple, at each integer point in time, a growth rate is drawn from $\{\lambda_0, \dots, \lambda_X\}$ according to probabilities $m(0), \dots, m(X)$, and the entire population—in the case of aggregate risk—grows continuously at this rate until the next draw. Then the size $N(t)$ of the population at time t will be

$$N(t) = N(0) \prod_{x=0}^X e^{\lambda_x t(x)},$$

where $t(x)$ is the number of periods (from among $1, 2, \dots, t$) in which growth rate λ_x is drawn. Then the growth rate is

$$\lim_{t \rightarrow \infty} \frac{\ln N(t)}{t} = \lim_{t \rightarrow \infty} \sum_{x=0}^{\infty} \frac{t(x)}{t} \lambda_x = \sum_{x=0}^X m(x) \lambda_x.$$

Hence, evolution will select for preferences that maximize the expected growth rate, even in the face of aggregate risk.

¹Robson [17] presents a more general argument, including the possibility of extinction. (See also Sinn and Weichenrieder [20].) Robson and Samuelson [18] extend this simple model to allow a more general age structure, showing that aggregate shocks to mortality can generate a higher pure rate of time preference than can be derived from idiosyncratic mortality alone.

²Evolution will select among aggregate risks to as to maximize the expected value of $\ln(u(c)) =: U(c)$, and

$$-\frac{U''(c)}{U'(c)} = -\frac{u''(c)}{u'(c)} + \frac{u'(c)}{u(c)} = R_A(c) + \frac{u'(c)}{u(c)}.$$

Suppose, for example, an individual faces a choice between an idiosyncratic lottery or an aggregate lottery. Each of these entails either a *rate* of consumption given by c_1 , with probability p , or a *rate* of consumption c_2 , with probability $1 - p$. For the idiosyncratic lottery, the outcomes are independent across individuals; for the aggregate lottery, the outcomes are the same across individuals in a given period. Suppose that the rate of production of expected offspring is $u(c)$ for each rate of consumption, c . In either case, the growth rate induced is $pu(c_1) + (1 - p)u(c_2)$, so the individual should have identical attitudes to idiosyncratic and aggregate risk.

Robatto and Szentes [16] establish a quite general version of this result, showing that in continuous-time models, evolution selects for preferences that are neutral with respect to aggregate risk. Although aggregate risk generates a growth path that is observably different from that of idiosyncratic risk, the corresponding long-run growth rates are identical, and evolution should then select individuals to have identical attitudes to aggregate and idiosyncratic risk.

A key observation is that the mathematical reason for the difference in these results is not whether the growth process is cast in continuous or discrete time, nor is there any conflict between these results. The relevant difference lies in how the risk is described, which determines the riskless default to which the risk is compared. To see this, suppose the population grows continuously. At each integer time, a random draw occurs that determines the growth of the population until the next draw. In one description, each random draw chooses equiprobably between the growth rates λ_1 and λ_2 , with the population growing at this rate until the next draw, at which point the population is either e^{λ_1} or e^{λ_2} times its current size. Then Robatto and Szentes' [16] analysis ensures that evolution will select for preferences that maximize the expected growth rate, and hence aggregate uncertainty concerning growth rates is neutral.³

In an equivalent description, the random draw chooses between the growth factors Λ_1 and Λ_2 (where $\Lambda_k = e^{\lambda_k}$), with the population being either Λ_1 or Λ_2 times its current size after one period. Then Robson's [17] analysis ensures that evolution will select for preferences that maximize the expected log of the growth factor, and hence aggregate uncertainty concerning growth factors is disadvantageous.⁴ But these are the same lotteries and hence the same preferences. It is simply a matter of whether we describe the risks in terms of factors or rates.

Despite these observations, Robatto and Szentes' [16] continuous time formulation naturally and forcefully directs attention to rates. Their striking result is then that as long as we consider continuous-time models with a simple age structure and focus on risks to rates, aggregate risks (like idiosyncratic risks) are evolutionarily neutral. The present paper shows, nevertheless, that a distinction

³In this case, the salient riskless mean is a growth rate of $\frac{1}{2}(\lambda_1 + \lambda_2)$. Since $e^{\lambda_1}e^{\lambda_2} = e^{\frac{1}{2}(\lambda_1 + \lambda_2)}e^{\frac{1}{2}(\lambda_1 + \lambda_2)}$, it follows that it is irrelevant whether the population switches between growth at rates λ_1 and λ_2 , where these are equally likely, or grows risklessly at the mean rate.

⁴Now, the salient riskless mean is a growth factor $\frac{1}{2}(\Lambda_1 + \Lambda_2)$. Since $\Lambda_1\Lambda_2 < [\frac{1}{2}(\Lambda_1 + \Lambda_2)]^2$, the population facing aggregate uncertainty grows more slowly than population growing risklessly at the mean rate.

between aggregate and idiosyncratic risk reappears in a continuous-time model with a richer age structure. For a simple example, suppose that mortality is constant across all ages, but fertility is zero before menarche at age $M > 0$ and constant thereafter. Idiosyncratic risk to fertility rates again yields the same growth rate as does the mean fertility rate. However, a population whose after-menarche fertility rates are perfectly correlated but random generates a lower long-run growth rate than that generated by the mean fertility rate. Intuitively, the lag built into the period length in discrete time is mirrored as the lag from birth to menarche in continuous time.⁵

1.2 Preview

Section 2 examines the effect of idiosyncratic risk in a continuous-time setting, culminating in the Euler-Lotka equation, which provides the central tool for our analysis. Appendix 6.1 provides a rigorous foundation for this continuous-time approach. We consider a discrete-time setting with finite numbers of agents facing independent risks to fertility and mortality. We take an appropriate limit to obtain the continuous-time Euler-Lotka equation in which fertility and mortality can be interpreted as means of idiosyncratic distributions. In the process, we confirm the property (built into the model of Section 2) that in general, idiosyncratic shocks to fertility and mortality rates are captured by the mean fertility and mortality rates.

Section 3 turns to aggregate risk. Analyzing the effect of age-varying aggregate shocks to demographic parameters is complex. We accordingly develop an approximation, based on the assumption that shocks to population parameters are rare. Formally, population parameters are determined by an underlying state that follows an ergodic Markov process with rare transitions. We use the resulting approximation to show that the effect of aggregate risk, relative to the mean, varies with the type of risk and age. Rare aggregate transitions in fertility rates at young ages are advantageous (compared to the corresponding mean fertility), but such shocks to fertility rates at old ages are disadvantageous. On the other hand, rare aggregate transitions in mortality at young as well as old ages are advantageous.

We are ultimately interested in attitudes toward risk in economic variables. Section 4 therefore develops the model further by supposing that both fertility and mortality are driven by a stream of economic resources. In equilibrium, resources are optimally allocated so that variations in resource “income” show

⁵A referee raised the question of how our analysis related to the capital asset pricing model, where random returns of individual securities (like idiosyncratic risks) have no effect, while market risk (like aggregate risk) does, all in the absence of an age structured or discrete time framework. The CAPM begins by assuming that individuals maximize the expectation of the discounted present value of a strictly concave utility function, thus building into the model risk attitudes that we seek to derive from evolutionary considerations. Given these risk preferences, risks to individual securities do not matter because the investor can diversify them, in much the same way that idiosyncratic risks to fertility or mortality are fully diversified from an evolutionary point of view. Non-diversifiable market risk plays a role because the utility function according to which they are evaluated is assumed to exhibit risk aversion.

up only as variations in fertility.⁶ Since resource transitions are then reflected as fertility transitions, the model implies that evolution will select for preferences that seek aggregate risk to income at young ages, but become averse to it as age increases. Hence, we might expect people to ease out of the stock market in their later years, a result that is empirically plausible but is not an obvious consequence of standard economic theory.

The body of Robatto and Szentes' [16] paper works with a population without an age structure, yielding results consistent with those of the current paper. In an online appendix, Robatto and Szentes' [16] examine a more general model, allowing arbitrary age structures, and suggest there is still a sense in which aggregate risks are neutral. Appendix 6.4 explains how their analysis relates to that in the current paper.

2 Idiosyncratic Risk

This section examines a population subject to *idiosyncratic risk*. The number of offspring that a surviving individual has at each age, and whether the individual continues to survive, are determined by random variables that are independent across individuals.

We consider a population growing in continuous time. We assume that a law-of-large-numbers argument, applied to the random fertility and mortality draws, ensures that the population growth is governed by the mean fertility and mortality rates. We examine the steady-state growth path for the population, in which the population grows at a constant rate and the age distribution remains constant over time. We summarize this steady-state growth path with the Euler-Lotka equation given in (5) below.

Appendix 6.1 provides a foundation for the model, replacing our law-of-large-numbers intuition and our implicit assumption that a steady-state growth path exists with rigorous arguments. Section 6.1.1 constructs a discrete-age and discrete-time branching process in which the idiosyncratic randomness in mortality and fertility appears explicitly. This section culminates in Proposition 5, showing that the population may suffer extinction, but that conditional on survival, the population converges (as time progresses) to a steady-state growth path in which the growth rate is constant and the age distribution remains constant over time. This steady state is a function of the mean fertility and mortality rates, and is characterized by a growth factor that satisfies the discrete version of the Euler-Lotka equation, given in (22). Section 6.1.2 then examines the limit of the discrete model as the length of a period approaches zero. Proposition 6 shows that the discrete version of the Euler-Lotka equation given by (22) converges to the Euler-Lotka equation given by (5).

⁶We assume there is a linear technology for converting resources to increased fertility, but a convex technology for converting resources to reduced mortality.

2.1 The Model

We consider a population of agents, each of whom is characterized by an age $a \in [0, A]$ where A is finite. We could extend the analysis to infinite A at the cost of additional technical complication, and we simplify the calculations in some examples by letting A be infinite. At each time t , the age structured population is $N_a(t)$, where the total population is then $N(t) = \int_0^A N_a(t) da$. At time 0, there is an initial age-structured population $N_a(0)$.

The mean fertility rate is given by the continuous function μ_a which is assumed positive for $a \in (M, A)$, for some $M \in [0, A]$.⁷ The rate of offspring production at time t is then $N_a(t)\mu_a$. The mean mortality rate is given by the continuous function ρ_a on $[0, A]$, so that ρ_a is the rate at which agents of age a die. The probability of survival to age a is then given by the decreasing function p_a , where

$$p_a = e^{-\int_0^a \rho_{\bar{a}} d\bar{a}}. \quad (1)$$

Given these fertility and mortality rates, the basic equations governing the population are twofold. First, for any age $a > 0$ and time $t \geq a$, the number of agents of age a at time t equals the number who were born at time $t - a$ and have survived until time t , or

$$N_a(t) = N_0(t - a)p_a, \quad (2)$$

assuming for simplicity that $t \geq a$.⁸ Next, for any time t , the number of agents born (i.e., of age 0) at time t is given by

$$N_0(t) = \int_0^A N_a(t)\mu_a da. \quad (3)$$

Each pair of fertility and mortality functions (μ_a, ρ_a) is associated with a steady-state growth path, in which the relative age structure of the population remains constant over time, with the population as a whole characterized by a growth rate λ . Our analysis focusses on these steady-state growth rates.

In the steady state, we must have, for all a and $t > \tau$

$$N_a(t) = e^{\lambda(t-\tau)} N_a(\tau).$$

This captures the steady-state notion that the number of agents of each age a at time t is the growth factor $e^{\lambda(t-\tau)}$ times the number of such agents at time $t - \tau$. As a special case of this steady-state condition, we have $N_0(t) = e^{\lambda a} N_0(t - a)$. We can insert this into (2) to obtain the steady-state age distribution

$$\frac{N_a(t)}{N_0(t)} = e^{-\lambda a} p_a. \quad (4)$$

⁷It is without much loss of generality to assume that agents do not live beyond the end of fertility. Such extended life, that is, has no effect on the growth rate.

⁸It is straightforward to allow for the initial range where $t < a$.

This implies that if $\lambda > 0$ and hence the population is growing, then as we move through older agents, the proportion of such agents in the population declines for two reasons. The first is that age- a agents were born at time $t - a$, and the population has since grown by factor $e^{\lambda a}$, and the second is that only p_a of the original cohort survives.

We can now insert the expression (4) for the steady-state age distribution into (3), describing the number of agents born at time t , to obtain

$$N_0(t) = \int_0^A N_0(t) e^{-\lambda a} p_a \mu_a da.$$

Canceling the term $N_0(t)$ from both sides, we have

$$1 = \int_0^A e^{-\lambda a} p_a \mu_a da. \tag{5}$$

This relationship, known as the Euler-Lotka equation, gives the population growth rate as a function of the age distributions μ_a of fertility and p_a of survival. This will be our basic tool for examining risk preferences.

2.2 Risk Preferences

A first implication of (5) is intuitive. Increases in mean fertility increase the population growth rate, as do decreases in mean mortality.

To see the implications for risk preferences, suppose that at each age, each individual in the population must choose between lotteries over fertility and mortality rates. In keeping with our focus on idiosyncratic risk, these lotteries are independent across individuals. We can think of two equivalent ways in which these choices might be implemented. Evolution might endow agents with utility functions, with the agents then choosing utility-maximizing lotteries. In this case, we are interested in characterizing the utility functions, and in particular the attendant risk attitudes, selected by evolution. Alternatively, evolution might simply hard-wire the appropriate choices. In this case, we are interested in the utility function that a revealed-preference analyst would use to describe the hard-wired choices.

Of course, evolution does not *purposefully* design agents, with either utility functions or choices. Instead, a process of random mutations introduces agents with different utility functions or different sets of hard-wired choices into the environment. The different utility functions or choice functions induce different rates of growth, leading to a process of selection ensuring that the environment will eventually be dominated that the utility function or choice function inducing the largest steady-state growth rate. Hence, evolution will select for the utility or choice function that leads to the fertility and mortality schedules that maximize the growth rate determined by (5).⁹

⁹This result is an implication of the Fundamental Theorem of Natural Selection (Hofbauer and Sigmund [12, p. 15]), and is implicit in Robson [17] and Robatto and Szentes [16].

The important observation now is that the growth rate determined by (5) is maximized by choosing between lotteries over fertility so as to maximize mean fertility, and choosing between lotteries over mortality so as to minimize mean mortality. Hence, evolution will select for risk neutrality with respect to idiosyncratic risk in fertility and mortality. Of course, we rarely think of people choosing directly between lotteries over fertility rates, and our ultimate interest is in risk attitudes over economic variables such as income. These are derived from the underlying risk attitudes over fertility and mortality, a link to which we turn in Section 4.

3 Aggregate Risk

We now incorporate aggregate risk concerning the demographic variables into the model. We invoke the results of Section 2 and assume that all idiosyncratic demographic variables are replaced by deterministic rates set at their means.

If a population subject to aggregate risk in a random variable grows faster than would an equivalent deterministic population where the aggregate random variable has been replaced by its mean, then we say that aggregate risk in that variable is advantageous. If a population subject to aggregate risk in a random variable grows more slowly than would an equivalent deterministic population where the aggregate random variable has been replaced by its mean, then we say that aggregate risk in that variable is disadvantageous. If neither is the case, then aggregate risk is neutral.

3.1 An Approximation

To capture aggregate risk, we assume that at time t , the environment is characterized by a state ω_t drawn from the set of states Ω . The state determines the fertility and mortality schedules μ_a and ρ_a applicable to the population. In each state, the mean fertility rate is given by the continuous function μ_a which is positive for $a \in (M, A)$, and the mean mortality rate is given by the continuous function ρ_a on $[0, A]$. The state is determined by a continuous-time Markov process which admits an ergodic distribution. For technical simplicity, we assume Ω is finite.

We simplify the analysis by focussing on the case in which changes in the state are rare. In this case, the system will spend a small fraction of time in short run adjustments of the population age structure in response to changes in the state. The population will spend most of its time characterized by (or, more precisely, very closely approximated by) an age structure that would prevail if the current state had *always* been the state. The overall growth rate will then be approximately the expectation λ^* of a random growth rate λ , where each realization of λ is derived from a system in which there is no aggregate risk and whose parameters correspond to one of the states in Ω . This is the analog of Robson and Samuelson's [18] focus on the case in which aggregate shocks are

frequent, but arbitrarily small. Ben-Porath, Dekel and Rustichini [4] examine a similar approximation.

For example, suppose that the population is characterized by one of two sets of demographic parameters, that prevail in states 1 and 2, respectively. Suppose that a population *always* in state 1 would have long-run growth rate λ_1 , and that a population *always* in state 2 would have long-run growth rate λ_2 . Let us further suppose that state 1 arises with ergodic probability p_1 and state 2 with probability p_2 . If changes between states happen frequently, then the population may often be growing at rates that are not close to either λ_1 or λ_2 . However, if changes between states are rare, then in most periods the population will have been in its current state for a very long time, and will have a growth rate very close to either λ_1 or λ_2 , depending on which is the current state. In that case, the total size of the population at time t will be given approximately by

$$N(t) = N(0)(e^{\lambda_1})^{t_1}(e^{\lambda_2})^{t-t_1},$$

where $t_1 \in [0, t]$ is the total length of time that the first state is in effect and N_0 is the initial population size. The rate of growth of this expression, denoted by λ^* , satisfies

$$\frac{\ln N(t)}{t} + \frac{t_1}{t}\lambda_1 + \frac{t-t_1}{t}\lambda_2 \longrightarrow p_1\lambda_1 + p_2\lambda_2 = \lambda^*$$

with convergence as $t \rightarrow \infty$. The approximation in this limiting growth rate becomes more precise as we look at state distributions with rarer transitions.

More formally, let the states in Ω be denoted by $1, \dots, N$ and indexed by n or m . The state evolves according to an irreducible continuous-time Markov chain with the instantaneous rate of transition from state m to n given by p_{mn} for all $m, n \in \Omega = \{1, \dots, N\}$, $m \neq n$. These transition rates generate an invariant distribution given by $\{\pi_n\}_{n=1}^N$ which satisfies, for each state n (Durrett [6, Theorem 3.2, p. 170])

$$\sum_{m, m \neq n} \pi_m p_{mn} = \left(\sum_{m, m \neq n} p_{nm} \right) \pi_n.$$

We now capture our interest in rare transitions by letting the instantaneous transition probabilities be given by p_{mn}/k , where $k \geq 1$, so that transitions become rarer as k increases. This rescaling of the transition probabilities has no effect on the invariant distribution $\{\pi_n\}_{n=1}^N$. However, each time the process enters state n , the length of its “sojourn” in state n is given by a random variable $\tilde{\tau}_n$ whose expected value is $1/(\sum_{m, m \neq n} \frac{p_{nm}}{k})$. Hence, as k increases, transitions between states become increasingly rare. The random variables $\tilde{\tau}_n$ are scaled upwards, and sojourn lengths become longer.¹⁰

¹⁰The rate at which a sojourn in state n ends is given by $\sum_{m, m \neq n} (p_{nm}/k)$. Conditional on being in state n at time 0, the probability of not having left state n by t is $e^{-(\sum_{m, m \neq n} (p_{nm}/k))t}$ and the expected sojourn is $1/(\sum_{m, m \neq n} (p_{nm}/k))$.

Let λ_n be the dominant eigenvalue (cf. Section 6.1) for a hypothetical population *permanently* in state n , and hence subject to no aggregate uncertainty, and let π be the stationary distribution of the states inducing the aggregate uncertainty. Appendix 6.2 proves:

Proposition 1 *For any given $\varepsilon > 0$, there exists a $K > 0$ such that if $k \geq K$, then the growth factor $P(T) = N(T)/N(0)$ of the population at time T almost surely satisfies*

$$\lim_{T \rightarrow \infty} \frac{1}{T} \ln P(T) \in \left[\sum_n \pi_n \lambda_n - \varepsilon, \sum_n \pi_n \lambda_n + \varepsilon \right], \quad (6)$$

The key implication of this result is that if we let

$$\lambda^* = \sum_n \lambda_n \pi_n, \quad (7)$$

then in the limit, as transitions between states become arbitrarily rare, the long-run growth rate of the population is given by λ^* .

Robatto and Szentes [16] derive a general expression for the growth rate of a population subject to aggregate risk. The criterion given by (7) is the special case of their criterion applicable to the limit of arbitrarily rare transitions.

3.2 Preferences Over Aggregate Risk

To see the implications for risk preferences, we once again suppose that at each age, each individual in the population must choose between lotteries over fertility and mortality rates. In this case, however, we assume the lotteries are perfectly correlated.

We again assume that a process of random mutations introduces agents with different utility functions or different sets of hard-wired choices into the environment. The different utility functions or choice functions induce different rates of growth, leading to a process of selection ensuring that the environment will eventually be dominated by the utility function or choice function inducing the largest steady-state growth rate. In this case, a lottery specifies fertility and mortality schedules for each state $n \in \{1, \dots, N\}$, with probabilities π_1, \dots, π_N reflecting the ergodic distribution over states. These lotteries in turn give rise to growth rates $\lambda_1, \dots, \lambda_N$, and are evaluated according to (7), with evolution selecting for choices that maximize (7).

We then gain insight into the effects of aggregate risk by examining how the growth rate λ , associated with a population subject to no risk, varies as the underlying demographic parameters vary. Applying Jensen's inequality to (7) allows us to conclude that if λ is concave in a parameter, then (rare) aggregate variation in that parameter will be disadvantageous. In that case, evolution will select for utility or choice functions that exhibit risk aversion with respect to aggregate risk in the parameter. If λ is convex in a parameter, then (rare) aggregate variation in that parameter will be advantageous, and evolution will select for utility or choice functions that are risk seeking in the parameter.

3.3 An Example

To illustrate the effect of aggregate risk in age-structured populations, consider a population in which there is always constant mortality at rate $\rho > 0$. The rate of fertility is 0 up to an age of menarche given by M , after which the rate of fertility is a constant $\mu > 0$, until a terminal age A at which fertility ceases.

The Euler-Lotka equation for a population subject to no aggregate risk, which uniquely determines the growth rate λ , is

$$1 = \int_M^A \mu e^{-(\lambda+\rho)a} da.$$

Performing the integration, we obtain λ as the implicit solution of

$$\mu = \frac{\lambda + \rho}{e^{-(\lambda+\rho)M} - e^{-(\lambda+\rho)A}}. \quad (8)$$

Clearly, the population growth rate must be increasing in the fertility rate μ and decreasing in the mortality rate ρ .

We now introduce aggregate risk by supposing that the fertility rate μ , or the mortality rate, ρ , is subject to rare random shocks. Considering mortality first, note that the sum $\lambda + \rho$ is determined once we fix the value of μ . It follows immediately that $\frac{d\lambda}{d\rho} = -1$ and hence, $\frac{d^2\lambda}{d\rho^2} = 0$, and so aggregate shocks to (constant across ages) mortality are neutral.

When examining fertility we simplify the notation by considering the case in which $\rho = 0$. This sacrifices no generality, given the previous paragraph's observation that the sum $\lambda + \rho$ is determined by μ —the implications of aggregate variations in μ for λ when ρ is fixed at zero are identical to those arising when ρ is fixed at a positive level. We must only remember to interpret the variables appropriately when constructing numerical estimates of the coefficients of relative risk aversion.

We identify an aggregate state ω with a fertility rate $\mu(\omega)$. We are then interested in the expected growth rate $\lambda^* = \sum_{\omega \in \Omega} \lambda(\mu(\omega))\pi(\omega)$, where $\lambda(\mu)$ is given by (8) (with $\rho = 0$) and is the growth rate of a population whose fertility rate is (permanently) fixed at μ . As we have noted, evaluating the effect of aggregate fertility risk on this growth rate involves assessing the curvature of the function $\lambda(\mu)$.

First, suppose that the population has a particularly simple age structure, in which $M = 0$ and $A = \infty$, so that maturation is immediate and there is no terminal age for fertility. Then equation (8) (with $\rho = 0$) gives the linear function $\lambda(\mu) = \mu$. In this case, aggregate risk in the fertility rate μ is also neutral—a population subject to rare shocks in the fertility rate μ grows at the same rate as would a population subject to no aggregate risk and with the fertility rate fixed at the mean fertility rate. Aggregate shocks to mortality and fertility rates are evolutionarily neutral as in Robatto and Szentes [16].

Now suppose that $M > 0$ but $A = \infty$. The implicit definition (8) of the population growth rate now becomes

$$\mu = \lambda(\mu)e^{\lambda(\mu)M}. \quad (9)$$

We can differentiate this expression and simplify to obtain

$$\frac{d\lambda}{d\mu} = \frac{\lambda}{\mu(1 + \lambda M)}.$$

Of course, $d\lambda/d\mu > 0$ —a population with a higher fertility rate grows faster. Differentiating again and rearranging yields

$$\mu(1 + \lambda M) \frac{d^2\lambda}{d\mu^2} = -\lambda M \frac{d\lambda}{d\mu} - \mu M \left(\frac{d\lambda}{d\mu} \right)^2$$

and hence (using our solution for $d\lambda/d\mu$)

$$\frac{d^2\lambda}{d\mu^2} = \frac{-\lambda M \frac{\lambda}{\mu(1 + \lambda M)} - \mu M \left(\frac{\lambda}{\mu(1 + \lambda M)} \right)^2}{\mu(1 + \lambda M)}.$$

The right side is clearly negative, and hence λ is concave in μ . In this case, aggregate risk is disadvantageous—rare aggregate variations in μ give a population growth rate lower than the growth rate of a population whose fertility is always given by the corresponding mean of μ . Intuitively, the lag to menarche has an effect akin to the lag dictated by a discrete time formulation.¹¹

How risk averse are the induced preferences over aggregate variations on fertility? Equation (9) does not give rise to an explicit solution for $\lambda(\mu)$, but we can use our expressions for the first and second derivatives to calculate the Arrow-Pratt measure of relative risk aversion $R_R(\mu)$, giving (after some simplification)

$$R_R(\mu) = -\mu \frac{\frac{d^2\lambda}{d\mu^2}}{\frac{d\lambda}{d\mu}} = \frac{\lambda M}{(1 + \lambda M)} \left(\frac{2 + \lambda M}{1 + \lambda M} \right). \quad (10)$$

This result has intuitively appealing features— $R_R(\mu)$ is 0 when $M = 0$, as discussed before, giving results analogous to those of Robatto and Szentes [16]. It also increases in M , converging to 1 as M grows to infinity.¹² That is, in this limiting case the criterion corresponds to the logarithmic criterion derived by Robson [17].¹³

What numerical implications does (10) have? Suppose, for example, that $M = 15$ and that $\mu = 0.1$, as is roughly consistent with poorer modern societies.¹⁴ It follows from (9) that $\lambda = 0.048$.¹⁵ It follows that $R_R(\mu) = 0.66$, a

¹¹Appendix 6.3 presents a discrete-time example, confirming that a delay in reaching reproductive age increases aversion to aggregate risk in that setting as well.

¹²In the limit as $M \rightarrow \infty$, (9) implies that $\lambda \rightarrow 0$ and $\lambda M \rightarrow \infty$.

¹³However, the criterion here is defined over growth rates and that in Robson [17] is over offspring levels.

¹⁴This is based on a total fertility rate, of 6, that is, 6 lifetime offspring. Of these offspring, 3 are female. We also suppose the rate of reproduction is constant over 30 years, from age 15 to age 45. Attention is limited to females, since they are the scarce factor in reproduction. Assuming $A = \infty$ does not have a huge effect, since this end of life is heavily discounted.

¹⁵This (calculated) value of λ may appear excessive when applied to the bulk of human evolutionary history, but the value of 0.048 is the sum $\lambda + \rho$ of the population growth rate and the mortality rate. This estimate is then consistent with evolutionarily plausible population growth rates and mortality rates.

plausible value.

Alternatively, suppose that $M = 0$ but A is finite. The implicit definition (8) of the population growth rate is now

$$\mu = \frac{\lambda(\mu)}{1 - e^{-\lambda(\mu)A}}. \quad (11)$$

A straightforward calculation shows that $d^2\mu/d\lambda^2 > 0$, so λ is a concave function of μ .¹⁶ Once again, aggregate risk is disadvantageous—rare aggregate variations in μ give a population growth rate lower than the growth rate of a population whose fertility is always given by the corresponding mean of μ . We can again derive the Arrow-Pratt measure of relative risk aversion, but the lack of realism in assuming $M = 0$ renders its numerical value of less interest.¹⁷

Of these two simple cases (the first with $M > 0$ and $A = \infty$, and the second with $M = 0$ and $A < \infty$), the first is more relevant. That is, with population growth, it is numerically more crucial when reproduction starts than when it finishes, or even whether it finishes at all. To obtain an explicit economic interpretation, suppose then the individual faces risk over consumption streams c that are constant over $[M, \infty)$. Suppose the rate of production of expected

¹⁶In particular, we have

$$\frac{d\mu}{d\lambda} = \frac{1 - e^{-\lambda A} - \lambda A e^{-\lambda A}}{(1 - e^{-\lambda A})^2} = \frac{1 - (1 + \lambda A)e^{-\lambda A}}{(1 - e^{-\lambda A})^2} > 0.$$

The second derivative is

$$\begin{aligned} \frac{d^2\mu}{d\lambda^2} &= \frac{(1 - e^{-\lambda A})^2[(1 + \lambda A)Ae^{-\lambda A} - Ae^{-\lambda A}] - [1 - (1 + \lambda A)e^{-\lambda A}]2(1 - e^{-\lambda A})Ae^{-\lambda A}}{(1 - e^{-\lambda A})^4} \\ &= \frac{Ae^{-\lambda A}[(1 - e^{-\lambda A})\lambda A - 2(1 - (1 + \lambda A)e^{-\lambda A})]}{(1 - e^{-\lambda A})^3}. \end{aligned}$$

The sign of $d\mu^2/d\lambda^2$ is given by the sign of

$$(1 - e^{-\lambda A})\lambda A - 2(1 - (1 + \lambda A)e^{-\lambda A}).$$

$$z - ze^{-z} - 2 + 2e^{-z} + 2ze^{-z} = z + ze^{-z} + 2e^{-z} - 2,$$

which is positive, and hence λ is a concave function of μ .

¹⁷If we write $\mu = f(\lambda)$, then we have $1 = f'(d\lambda/d\mu)$, and hence $d\lambda/d\mu = 1/f'$. Differentiating $1 = f'(d\lambda/d\mu)$, we have $0 = f''(d\lambda/d\mu)^2 + f'(d^2\lambda/d\mu^2)$, which we solve for $d^2\lambda/d\mu^2 = -f''(d\lambda/d\mu)^2/f' = -f''/(f')^3$. We thus have

$$\begin{aligned} -\frac{\frac{d^2\lambda}{d\mu^2}}{\frac{d\lambda}{d\mu}} &= \frac{f''}{(f')^2} = \frac{\frac{d^2\mu}{d\lambda^2}}{\left(\frac{d\mu}{d\lambda}\right)^2} = \frac{d^2\mu}{d\lambda^2} \left(\frac{d\mu}{d\lambda}\right)^{-2} \\ &= \frac{Ae^{-\lambda A}[(1 - e^{-\lambda A})\lambda A - 2(1 - (1 + \lambda A)e^{-\lambda A})]}{(1 - e^{-\lambda A})^3} \left(\frac{(1 - e^{-\lambda A})^2}{1 - (1 + \lambda A)e^{-\lambda A}}\right)^2 \end{aligned}$$

Hence, since $\mu(1 - e^{-\lambda A}) = \lambda$, the coefficient of relative risk aversion is

$$R_R(\mu) = -\mu \frac{\frac{d^2\lambda}{d\mu^2}}{\frac{d\lambda}{d\mu}} = \frac{\lambda Ae^{-\lambda A}[(1 - e^{-\lambda A})\lambda A - 2(1 - (1 + \lambda A)e^{-\lambda A})]}{(1 - (1 + \lambda A)e^{-\lambda A})^2}.$$

offspring is given by $u(c)$ for each c . If the consumption is specified by an idiosyncratic lottery \tilde{c} , the rate of growth $\lambda(\mathbb{E}u(\tilde{c}))$ is determined by the expected rate of reproduction $\mathbb{E}u(\tilde{c})$. If the lotteries are aggregate, on the other hand, the appropriate criterion is the mean growth rate $\mathbb{E}\lambda(u(\tilde{c}))$. Since λ is concave, it follows that the individual is more averse to aggregate risk than to precisely comparable idiosyncratic risk. It follows, indeed, that the coefficient of absolute risk-aversion for aggregate risks is given by

$$R_A(c) - \frac{\lambda''(u(c))u'(c)}{\lambda'(u(c))} > R_A(c),$$

where $R_A(c) = -\frac{u''(c)}{u'(c)}$ is the coefficient of absolute risk aversion for idiosyncratic risks.¹⁸

More generally, Appendix 6.3 shows that when $M > 0$ and $A < \infty$, aggregate risk in fertility is disadvantageous. Some intuition as to why a limited range of fertility leads to aversion to aggregate risk, in contrast to Robatto and Szentes [16], is as follows. Suppose fertility is $\mu_a = \frac{\tilde{\varepsilon}}{A-M} \mathbf{1}_{a \in [M, A]}$, where $\mathbf{1}$ denotes the set indicator function. Fertility is subject to an aggregate shock, $\tilde{\varepsilon}$, as in Robatto and Szentes, but over a limited age range. In the limit as $A \downarrow M$, fertility becomes $\tilde{\varepsilon} \mathbf{d}(a - M)$, where $\mathbf{d}(a - M)$ is the Dirac delta function. It then follows readily that the growth rate of the line of descendants of a single individual is $\frac{1}{M} \mathbb{E} \ln(\tilde{\varepsilon}) < \frac{1}{M} \ln \mathbb{E}(\tilde{\varepsilon})$, by a calculation analogous to that required in discrete time. In this limit, individuals are averse to aggregate risk, with the logarithmic criterion of Robson [17].

This example shows that overall risk aversion to aggregate risk does not neatly decompose across age ranges. Consider the case that $M = C > 0$ and $A = \infty$ and the case that $M = 0$ and $A = C < \infty$, for a given $C > 0$. These two subranges partition $[0, \infty)$. Nevertheless, despite strict aversion to aggregate risk over constant fertility defined over either subrange, with zero fertility outside that subrange, there is risk neutrality with respect to aggregate shocks to fertility over $[0, \infty)$. That is, the effect of aggregate risk at a particular set of ages depends crucially on what occurs outside that set. Similarly, Proposition 2 below shows that shocks to fertility that are sufficiently concentrated are advantageous at young ages, while Proposition 3 shows that sufficiently concentrated shocks to mortality are advantageous at all ages, even in this example, where there is risk neutrality with respect to aggregate shocks to fertility and mortality over $[0, \infty)$.

3.4 Age and Risk Attitudes

We are especially interested in how risk attitudes vary by age. Hence, instead of examining the effects of aggregate risk that affects the fertility of every age, as in Section 3.3, we would like to examine a shock that affects the fertility or mortality only of agents of some age b .

¹⁸Note that, by (10), $-\frac{\lambda''}{\lambda'} = \frac{\lambda M}{\mu(1+\lambda M)} \left(\frac{2+\lambda M}{1+\lambda M} \right)$, increases in growth λ or menarche M increase risk aversion.

3.4.1 Rare Aggregate Shocks to Fertility

We consider a population fertility schedule $\mu_{(\cdot)} : [0, A] \rightarrow \mathbb{R}_+$, where A is finite, given by

$$\mu_a + \varepsilon \delta(a, b, c),$$

where the continuous function $\delta \geq 0$ has the following properties

$$\int_b^{b+c} \delta(a, b, c) da = 1 \text{ and } \delta(a, b, c) = 0 \text{ whenever } a \leq b \text{ or } a \geq b + c.$$

Thus ε is an aggregate shock to the fertility rate of agents in the population of ages $[b, b + c]$. We are interested in how the effects of shocks to fertility at different ages compare, and so consider a narrow range of ages $[b, b + c]$. The integrated magnitude of the fertility shock is ε , regardless of b or c .

As explained in Sections 3.1–3.3, we now identify a state with the value of the shock ε , and view $\varepsilon \in \Omega$ as being governed by a Markov process with rare transitions and ergodic distribution π . Let $\lambda(\varepsilon)$ denote the growth rate of population facing a fixed value of ε . We then fix the ergodic distribution π and examine cases in which transitions between states become arbitrarily rare, allowing us to rewrite the long-run growth rate λ^* given by (7) as $\lambda^* = \int_{\Omega} \lambda(\varepsilon) d\pi(\varepsilon)$. We then examine the second derivative of $\lambda(\varepsilon)$ with respect to ε . We sharpen the results by focussing on the case of small shocks, implemented by evaluating the second derivative at $\varepsilon = 0$. We further simplify the resulting expression by taking the limit as $c \rightarrow 0$. The technique is equivalent to subjecting fertility to an aggregate shock at age b so that the fertility schedule becomes

$$\mu_a + \varepsilon \mathbf{d}(a - b),$$

where \mathbf{d} is the Dirac delta function. Although $\mu_a + \varepsilon \delta(a, b, c)$ is a continuous and bounded function, for all $c > 0$, the Dirac delta function is not. Nevertheless, this limit is formally useful to sharply characterize how risk attitudes vary with age.

Our first task is then to examine $\lambda(\varepsilon)$. The appropriate Euler-Lotka equation is

$$1 = \int_0^A [\mu_a + \varepsilon \delta(a, b, c)] p_a e^{-\lambda a} da = \int_0^A \mu_a p_a e^{-\lambda a} da + \varepsilon \int_b^{b+c} \delta(a, b, c) p_a e^{-\lambda a} da.$$

Taking the derivative of the Euler-Lotka equation with respect to ε gives

$$-\frac{d\lambda}{d\varepsilon} \int_0^A a \mu_a p_a e^{-\lambda a} da - \frac{d\lambda}{d\varepsilon} \varepsilon \int_b^{b+c} \delta(a, b, c) a p_a e^{-\lambda a} da + \int_b^{b+c} \delta(a, b, c) p_a e^{-\lambda a} da = 0. \quad (12)$$

To make this derivative easier to interpret, we consider its value at $\varepsilon = 0$, giving

$$-\frac{d\lambda}{d\varepsilon} \int_0^A a \mu_a p_a e^{-\lambda a} da + \int_b^{b+c} \delta(a, b, c) p_a e^{-\lambda a} da = 0.$$

Finally, we take the limit as $c \rightarrow 0$ to obtain

$$-\frac{d\lambda}{d\varepsilon} \int_0^A a\mu_a p_a e^{-\lambda a} da + p_b e^{-\lambda b} = 0.$$

Notice that $\frac{d\lambda}{d\varepsilon} > 0$, giving the expected result that increases in fertility increase the growth rate.

We apply the same technique to derive a useful expression for the second derivative of λ . That is, we return to the expression in (12), take the derivative with respect to ε , then evaluate this derivative at $\varepsilon = 0$, and take the limit as $c \rightarrow 0$. This yields

$$-\frac{d^2\lambda}{d\varepsilon^2} \int_0^A a\mu_a p_a e^{-\lambda a} da + \left(\frac{d\lambda}{d\varepsilon}\right)^2 \int_0^A a^2 \mu_a p_a e^{-\lambda a} da - 2\frac{d\lambda}{d\varepsilon} b p_b e^{-\lambda b} = 0.$$

Evaluating the derivative at $\varepsilon = 0$ gives us a sharper conclusion which holds as long as the shocks to fertility are not too large, leading to the qualification “small” in Proposition 2. Taking the limit as $c \rightarrow 0$ corresponds to focussing on an increasingly narrow range of ages that are affected by the aggregate fertility shock. This simplifies the expression, and makes clear the meaning of referring to fertility shocks “at age b .”

After some algebra, it can be shown that

$$\frac{\lambda''}{\lambda'} = \frac{p_b e^{-\lambda b}}{\left(\int_0^A a\mu_a p_a e^{-\lambda a} da\right)^2} \left(\int_0^A a^2 \mu_a p_a e^{-\lambda a} da - 2b \int_0^A a\mu_a p_a e^{-\lambda a} da \right). \quad (13)$$

It follows immediately from this expression that $\frac{\lambda''}{\lambda'} > 0$ for $b < \bar{b} = \frac{\int_0^A a^2 \mu_a p_a e^{-\lambda a} da}{2 \int_0^A a\mu_a p_a e^{-\lambda a} da}$, but $\frac{\lambda''}{\lambda'} < 0$ for $b > \bar{b}$. Hence, we have established:

Proposition 2 *There exists an age $\bar{b} > 0$ such that small, rare aggregate shocks to fertility at age $b < \bar{b}$ are advantageous, while small, rare aggregate shocks to fertility at age $b > \bar{b}$ are disadvantageous.*

If evolution has shaped our risk preferences in light of these considerations, then we can expect people to be averse to aggregate risk to fertility at older ages, but to seek such risk at younger ages.

Suppose, for example, that $\mu_a = \mu$, so that the fertility rate is constant, that $p_a = e^{-\rho a}$, where the mortality rate ρ is also constant, and $A = \infty$. It can then readily be shown that $\bar{b} = \frac{1}{\rho + \lambda}$, which is the average age of the steady state population. Hence rare aggregate shocks to fertility are advantageous at any age younger than the average age, and are disadvantageous at any age older than the average age.

To build some intuition for Proposition 2, let us continue with this special case. If fertility is subject to shocks at ages in the range $[b, b + c]$, then the Euler-Lotka equation becomes

$$1 = \int_0^\infty \mu e^{-(\lambda + \rho)a} da + \varepsilon \int_b^{b+c} \delta(a, b, c) e^{-(\lambda + \rho)a} da,$$

which becomes, in the limit as $c \rightarrow 0$,

$$\rho + \lambda(\varepsilon) = \mu + \varepsilon(\rho + \lambda(\varepsilon))e^{-(\rho + \lambda(\varepsilon))b}. \quad (14)$$

The function $n_b(\varepsilon) = (\rho + \lambda(\varepsilon))e^{-(\rho + \lambda(\varepsilon))b}$ describes the effect on population structure of the shock ε at age b .

It is obvious that the growth rate λ is an increasing function of ε —higher fertility leads to faster growth. Suppose first (and counterfactually) that $n_b(\varepsilon)$ were constant in ε . Then it is apparent from (14) that $\lambda(\varepsilon)$ would increase linearly in ε , and hence aggregate shocks to fertility at age b would be evolutionarily neutral. Instead, however, $n_b(\varepsilon)$ is an increasing function of ε when $b < \bar{b}$ and a decreasing function if $b > \bar{b}$. This reflects the fact that faster growth increases the relative number of young individuals and decreases the relative number of old individuals. It follows that $\lambda(\varepsilon)$ is convex in ε when $b < \bar{b}$, and so aggregate variations in fertility at age $b < \bar{b}$ are advantageous. On the other hand, $\lambda(\varepsilon)$ is concave in ε , when $b > \bar{b}$ and so aggregate variations in fertility at age $b > \bar{b}$ are disadvantageous.¹⁹

We can examine the magnitude of the risk aversion induced by the variations in fertility examined here. In Section 3.3, when examining variations in a constant fertility rate μ or mortality rate ρ , the coefficient of absolute risk aversion has the same dimension as $1/\mu$ or $1/\rho$, and we examined coefficients of relative risk aversion in order to obtain a dimensionless measure. In the current case, the factor ε has no population or time (or other) dimension. The coefficient of absolute risk aversion is similarly already dimensionless, and hence is a suitable measure. Continuing with our simple case of constant fertility, μ , and constant mortality, ρ , it follows readily from (13) that²⁰

$$R_A(\mu, b) = -\frac{\lambda''}{\lambda'} = 2e^{-(\lambda + \rho)b}(b(\lambda + \rho) - 1). \quad (15)$$

This gives $R_A(\mu, 0) = -2$ so that there is noticeable degree of risk *seeking* at $b = 0$. The maximum of $R_A(\mu, b)$ obtains at $b = 2/(\lambda + \rho)$ where it has the value $2e^{-2} < 1$, giving risk aversion, at less dramatic rates, when $b > \bar{b}$.

¹⁹These results follow from writing (14) as $\rho + \lambda(\varepsilon) = \mu + \varepsilon n_b(\varepsilon)$, and then noting that

$$\begin{aligned} \frac{d\lambda}{d\varepsilon} &= n_b(\varepsilon) + \varepsilon \frac{dn_b(\varepsilon)}{d\varepsilon} \quad \text{and} \\ \frac{d^2\lambda}{d\varepsilon^2} &= 2 \frac{dn_b(\varepsilon)}{d\varepsilon} + \varepsilon \frac{d^2n_b(\varepsilon)}{d\varepsilon^2} = 2 \frac{dn_b(\varepsilon)}{d\varepsilon} \quad \text{if } \varepsilon = 0. \end{aligned}$$

²⁰Apply the results $\int_0^\infty e^{-(\lambda + \rho)a} da = 1/(\lambda + \rho)$, $\int_0^\infty a e^{-(\lambda + \rho)a} da = 1/(\lambda + \rho)^2$, and $\int_0^\infty a^2 e^{-(\lambda + \rho)a} da = 2/(\lambda + \rho)^3$ to obtain

$$-\frac{\lambda''}{\lambda} = -\frac{e^{-(\lambda + \rho)b}}{\frac{\mu^2}{(\lambda + \rho)^4}} \left(\frac{2\mu}{(\lambda + \rho)^3} - \frac{2b\mu}{(\lambda + \rho)^2} \right)$$

and simplify, using the fact that when $\varepsilon = 0$, we have $\mu = \lambda + \rho$.

3.4.2 Rare Aggregate Shocks to Mortality

We now turn our attention to rare aggregate shocks in mortality. Again, we can smooth out variation in mortality. That is, we can consider aggregate risk about mortality rates in a small range of ages $[b, b + c]$, analogous to the argument of Section 3.4.1. For simplicity, however, we suppose directly that the mortality schedule is given by $\rho_a + \varepsilon \mathbf{d}(a - b)$, where \mathbf{d} is the Dirac delta function, so that we have an aggregate shock to mortality of magnitude ε at age b .

A straightforward derivation gives the Euler-Lotka equation

$$1 = \int_0^A \mu_a \exp \left[- \int_0^a \rho_{\bar{a}} d\bar{a} - \varepsilon H(a - b) \right] e^{-\lambda a} da,$$

where H is the Heaviside step function. We again analyze the effect of rare shifts in aggregate mortality by examining the curvature of the function $\lambda(\varepsilon)$.

We can differentiate the Euler-Lotka equation to obtain:

$$\begin{aligned} 0 = & -\frac{d\lambda}{d\varepsilon} \int_0^A a \mu_a \exp \left[- \int_0^a \rho_{\bar{a}} d\bar{a} - \varepsilon H(a - b) \right] e^{-\lambda a} da \\ & - \int_0^A \mu_a \exp \left[- \int_0^a \rho_{\bar{a}} d\bar{a} - \varepsilon H(a - b) \right] H(a - b) e^{-\lambda a} da. \end{aligned} \quad (16)$$

Notice that $\frac{d\lambda}{d\varepsilon} < 0$, giving the expected result that increases in mortality increase the growth rate. We take another derivative with respect to ε , and evaluate this derivative at $\varepsilon = 0$, to obtain

$$\begin{aligned} 0 = & -\frac{d^2\lambda}{d\varepsilon^2} \int_0^A a \mu_a p_a e^{-\lambda a} da + \left(\frac{d\lambda}{d\varepsilon} \right)^2 \int_0^A a^2 \mu_a p_a e^{-\lambda a} da \\ & + 2 \frac{d\lambda}{d\varepsilon} \int_0^A a \mu_a p_a H(a - b) e^{-\lambda a} da + \int_0^A \mu_a p_a (H(a - b))^2 e^{-\lambda a} da. \end{aligned} \quad (17)$$

After some algebra, it follows that²¹

$$-\frac{\lambda''}{\lambda'} = \frac{D}{\left(\int_0^A a \mu_a p_a e^{-\lambda a} da \right)^2}, \quad (18)$$

²¹Let

$$D_1 = \int_0^A a \mu_a p_a e^{-\lambda a} da, \quad D_2 = \int_0^A a^2 \mu_a p_a e^{-\lambda a} da, \quad D_3 = \int_0^A a \mu_a p_a H(a - b) e^{-\lambda a} da, \quad D_4 = \int_0^A \mu_a p_a H(a - b) e^{-\lambda a} da.$$

Then, noting that $(H(a - b))^2 = H(a - b)$, we can rewrite equation (17) as

$$0 = -\lambda'' D_1 + (\lambda')^2 D_2 + 2\lambda' D_3 + D_4.$$

From (16) evaluated at $\varepsilon = 0$, we have $0 = -\lambda' D_1 - D_4$, so that

$$0 = -\lambda'' D_1 - \lambda' \frac{D_2 D_4}{D_1} + 2\lambda' D_3 - \lambda' D_1.$$

Rearranging gives

$$-\frac{\lambda''}{\lambda'} = \frac{D_2 D_4 - 2D_1 D_3 + (D_1)^2}{(D_1)^2} = \frac{D}{(D_1)^2},$$

which is (18).

where

$$D = \left(\int_0^A \mu_a p_a H(a-b) e^{-\lambda a} da \right) \left(\int_0^A a^2 \mu_a p_a e^{-\lambda a} da \right) - 2 \left(\int_0^A a \mu_a p_a e^{-\lambda a} da \right) \left(\int_0^A a \mu_a p_a H(a-b) e^{-\lambda a} da \right) + \left(\int_0^A a \mu_a p_a e^{-\lambda a} da \right)^2.$$

Recall that $\lambda' < 0$, so that the sign of λ'' is the sign of D . If $b = 0$, D is the variance of age in the steady state distribution and so is positive. It is clear too that D must be positive in the limit as $b \rightarrow \infty$. It is easily shown that D reaches a minimum at $\hat{b} = \frac{\int_0^A a^2 \mu_a p_a e^{-\lambda a} da}{2 \int_0^A a \mu_a p_a e^{-\lambda a} da}$. More generally, we have:

Proposition 3 *Small, rare aggregate shocks to mortality are advantageous at all ages $b \geq 0$.*

Proof Because $\lambda' < 0$, it suffices for the result to show that D is positive, for all $b \geq 0$. Let f denote the distribution $\mu_a p_a e^{-\lambda a}$ of ages on $[0, A]$, and assume (normalizing if needed) that this is a probability density with cumulative distribution F . Then we can rewrite D as

$$(1 - F(b))\mathbb{E}\{a^2\} - 2\mathbb{E}\{a\}\mathbb{E}\{a|a > b\}(1 - F(b)) + (\mathbb{E}\{a\})^2,$$

where the expectations are taken with respect to the distribution f . Expanding the expressions for $\mathbb{E}\{a\}$ and $\mathbb{E}\{a^2\}$ it follows that the desired inequality holds if and only if

$$\begin{aligned} & (1 - F(b))[F(b)\mathbb{E}\{a^2|a \leq b\} + (1 - F(b))\mathbb{E}\{a^2|a > b\}] \\ & + [F(b)\mathbb{E}\{a|a \leq b\} + (1 - F(b))\mathbb{E}\{a|a > b\}]^2 \\ & > 2[F(b)\mathbb{E}\{a|a \leq b\} + (1 - F(b))\mathbb{E}\{a|a > b\}]\mathbb{E}\{a|a > b\}(1 - F(b)). \end{aligned}$$

Eliminating some common terms, this inequality holds, in turn, if and only if

$$\begin{aligned} & (1 - F(b))F(b)\mathbb{E}\{a^2|a \leq b\} + (1 - F(b))^2\mathbb{E}\{a^2|a > b\} + F(b)^2[\mathbb{E}\{a|a \leq b\}]^2 \\ & > (1 - F(b))^2[\mathbb{E}\{a|a > b\}]^2, \end{aligned}$$

which holds (noting that $\text{Var}[a|a > b] = \mathbb{E}[a^2|a > b] - [\mathbb{E}[a|a > b]]^2 \geq 0$) if and only if

$$(1 - F(b))F(b)\mathbb{E}\{a^2|a \leq b\} + F(b)^2[\mathbb{E}\{a|a \leq b\}]^2 + (1 - F(b))^2\text{Var}[a|a > b] > 0,$$

which is obvious. ■

Again, if evolution has shaped our risk preferences in light of these considerations, then we can expect people to seek aggregate risk to mortality. This

contrasts with the results for aggregate shocks to fertility, which (in this setting) are advantageous at young ages but disadvantageous for those over the average age.

Consider the case that $\mu_a = \mu$, a constant, $p_a = e^{-\rho a}$, where ρ is also constant and $A = \infty$. After some algebra, it can be shown that the sign of D is the sign of $-2e^{-(\lambda+\rho)b}(\lambda+\rho)b+1$. This is minimized at $\hat{b} = \frac{1}{(\lambda+\rho)}$ in which case it has value $1 - 2/e > 0$. As expected, rare random aggregate shocks to mortality are advantageous at all ages.

Continuing with this simple case, we seek an idea of the strength of the preference for risk. Since ε is again dimensionless, we again examine the coefficient of absolute risk aversion. Using (18), we have²²

$$R_A(\rho, b) = \frac{\lambda''}{\lambda'} = [2\mu b e^{-\mu b} - 1].$$

(We define the coefficient of absolute risk aversion, $R_A(\rho, b)$, like this to preserve the convention about the meaning of its sign, given $\lambda' < 0$.) Hence, $R_A(\rho, b)$ is -1 at $b = 0$, reaches a maximum of $-1 + 2e^{-1} \in (-1, 0)$ at $b = 1/\mu$ and tends to -1 as $b \rightarrow \infty$.

To build some intuition for Proposition 3, we consider a simple discrete-time two-age model. Survival from age 0 to age 1 is certain, but individuals survive from age 1 to age 2 with probability $1 - \varepsilon$, and then die at the end of age 2. Each individual of age 1 and each surviving individual of age 2 has probability u of having one offspring. We then consider aggregate variation in ε , i.e., aggregate variation in survival.

The Euler-Lotka equation then becomes

$$1 = \frac{u}{\lambda} + \frac{(1 - \varepsilon)u}{\lambda^2},$$

which we can solve for

$$\varepsilon = \frac{u\lambda - \lambda^2 + u}{u}.$$

It is apparent that ε is a strictly decreasing (since $\lambda > u$) and strictly concave function of λ . It then follows that λ is also a strictly decreasing and strictly concave function of ε . Hence, aggregate shocks to survival from age 1 to age 2 are disadvantageous. The forces behind this result are reminiscent of those that cause aggregate shocks to fertility at later ages to be disadvantageous.

What then accounts for the risk *loving* in Proposition 3? We have described the aggregate variations in our two-period example as affecting the survival probability $1 - \varepsilon$, whereas Proposition 3 is expressed in terms of mortality rates. To recast our example in terms of mortality rates, we write the probability of survival from period 1 to period 2 as $e^{-\varepsilon}$, and then consider aggregate variations in the mortality rate ε . The Euler-Lotka equation now becomes

$$1 = \frac{u}{\lambda} + \frac{ue^{-\varepsilon}}{\lambda^2},$$

²²Apply the results $\int_b^\infty e^{-(\lambda+\rho)a} da = e^{-(\lambda+\rho)b}/(\lambda+\rho)$, $\int_b^\infty ae^{-(\lambda+\rho)a} da = e^{-(\lambda+\rho)b} [1 + (\lambda+\rho)b]/(\lambda+\rho)^2$, which hold for all $b \geq 0$, and $\int_0^\infty a^2 e^{-(\lambda+\rho)a} da = 2/(\lambda+\rho)^3$.

which we can solve for

$$\varepsilon = -\ln[\lambda^2 - u\lambda] + \ln u.$$

It follows readily that

$$\begin{aligned} \frac{d\varepsilon}{d\lambda} &= \frac{-2\lambda + u}{\lambda^2 - u\lambda} < 0 \\ \frac{d^2\varepsilon}{d\lambda^2} &= \frac{\lambda^2 + (\lambda - u)^2}{(\lambda^2 - u\lambda)^2} > 0. \end{aligned}$$

Since ε is a strictly decreasing and strictly convex function of λ , it follows that λ is also a strictly decreasing and strictly convex function of ε . Aggregate shocks to mortality rates are advantageous.

We obtain different results for aggregate variations in survival and mortality because the probability of surviving from period 1 to period 2 is linear in the survival probability (i.e., $1 - \varepsilon$ is linear in ε), but is convex in the mortality rate (i.e., $e^{-\varepsilon}$ is convex in ε). When examining fertility, we found that aggregate variations at later ages were disadvantageous. We find an analogous result here for survival probabilities. However, the convexity of the survival probability in the mortality rate builds in a force pushing causing aggregate variations in the mortality *rate* to be advantageous. We see again that attitudes to risk are affected by whether the demographic variables are expressed as factors or as rates.

4 The Allocation of Resources

We usually think of risk as affecting economic variables such as income, rather than affecting mortality or fertility directly. We would accordingly like to examine the evolutionary implications of aggregate shocks to economic variables such as income. We construct a simple model in which such effects appear.

4.1 Optimal Resource Use

Suppose that each agent receives an age-dependent flow of income, with y_a specifying income at age a . Income can be divided between income used to increase fertility (c_a) and income used to decrease mortality (s_a), so that $y_a = c_a + s_a$. There is no saving. This aids tractability, but it might also be a reasonable approximation of hunter-gatherer economies. Fertility at age a is endogenous, given by $\mu_a = u_a(c_a)$, where u_a is an age-dependent increasing concave production function. Similarly, mortality is endogenous, given by $\rho_a = r_a(s_a)$, where r_a is an age-dependent decreasing convex production function. Survival to age a is given by p_a , which connects to mortality as $\frac{dp_a}{da} = -p_a r_a(s_a)$. The Euler-Lotka equation is

$$\int_0^A e^{-\lambda a} p_a u_a(c_a) da = 1.$$

The basic problem is then to find the growth rate $\hat{\lambda}$ that solves

$$\begin{aligned} \max_{c_a, s_a} \lambda & \quad \text{subject to} \\ 1 &= \int_0^A e^{-\lambda a} p_a u_a(c_a) da \\ y_a &= c_a + s_a \\ \frac{dp_a}{da} &= -p_a r_a(s_a), \quad p_0 = 1. \end{aligned}$$

It is not hard to show that we can equivalently find $\hat{\lambda}$ by first solving the following problem, for every value of λ ,

$$\begin{aligned} \max_{c_a, s_a} \int_0^A e^{-\lambda a} p_a u_a(c_a) da & \quad \text{subject to} \\ y_a &= c_a + s_a \\ \frac{dp_a}{da} &= -p_a r_a(s_a), \quad p_0 = 1 \end{aligned}$$

and then noting that $\hat{\lambda}$ is the unique value of λ that then solves $1 = \int_0^A e^{-\lambda a} p_a u_a(c_a) da$.

The Hamiltonian for this second maximization problem is

$$\mathcal{H} = e^{-\lambda a} p_a u_a(c_a) - \eta_a p_a r_a(s_a),$$

which is to be maximized subject to $y_a = c_a + s_a$. Here, η_a is the costate variable for the state variable p_a , where $\eta_a e^{\lambda a}$ can be interpreted as the contemporaneous reproductive value of an individual of age a . To see this, note that the adjoint equation is $\frac{d\eta_a}{da} = -\frac{\partial \mathcal{H}}{\partial p} = -e^{-\lambda a} u_a(c) + \eta_a r_a(s)$. Using $\frac{dp_a}{da} = -p_a r_a(s_a)$, it follows then that $\frac{d}{da} (p_a \eta_a) = -p_a e^{-\lambda a} u_a(c_a)$ so that

$$\eta_a e^{\lambda a} = \frac{\int_a^A p_{\bar{a}} e^{-\lambda \bar{a}} u_{\bar{a}}(c_{\bar{a}}) d\bar{a}}{p_a e^{-\lambda a}}, \quad (19)$$

using the transversality condition that $\eta_A = 0$. This equation gives the reproductive value of an individual of age a . This is the appropriately discounted integral of future fertility, conditional on being alive at age a .

The solution to the problem of maximizing the Hamiltonian, \mathcal{H} , is characterized by the Kuhn-Tucker conditions. In particular, the first-order condition for an interior solution is

$$p_a e^{-\lambda a} \frac{du_a(c)}{dc_a} = -\eta_a p_a \frac{dr_a(s)}{ds_a},$$

so that

$$-\eta_a e^{\lambda a} \frac{dr_a(s)}{ds_a} = \frac{du_a(c)}{dc_a}. \quad (20)$$

This equates the marginal gain from using resources to promote fertility, $\frac{du_a(c)}{dc_a}$, to the marginal gain from using resources to reduce mortality, $-\eta_a e^{\lambda a} \frac{dr_a(s)}{ds_a}$.

The latter expression is the marginal gain from mortality reduction, $-\frac{dr_a(s)}{ds_a}$, times the contemporaneous value of an individual alive at date a , namely $\eta_a e^{\lambda a}$.

Existence of an optimal solution follows readily. Further, the maximized Hamiltonian is concave in the state p_a , so that the above conditions characterize the optimum (Mangarasian [14]).

4.2 Shocks to Income

How would individuals react to rare aggregate shocks to income y_b at age b ? We have examined the effects of aggregate shocks to fertility and mortality, and the effect of aggregate shocks to income will depend on how these translate into shocks to fertility and mortality. In general, the result will involve a complex weighted average of the responses to aggregate shocks to fertility and to mortality. We accordingly restrict attention to a simple special case. Tractability plays an important role in our choices, though we argue at the end of this section that this is a plausible first approximation.

Suppose that fertility is a linear function, so that

$$u_a(c) = \alpha_a c$$

for some $\alpha_a > 0$. For an interior solution, the first order condition given by (20) is then

$$\alpha_a = -\eta_a e^{\lambda a} \frac{dr_a(s)}{ds_a}.$$

If y_a is sufficiently small for some age a , then there might well be a corner solution to the problem of maximizing the Hamiltonian, \mathcal{H} , with $c_a = 0$, so that all resources are devoted to survival. In this case, there is no fertility at age a , and all variation in $y_a = s_a$ induces variation in mortality. This might be the case over some initial period of childhood. For simplicity, however, we limit attention to the case that y_a is sufficiently large that there is an interior solution to the first-order condition. Our interpretation is either that this is the case throughout an agent's lifetime, or that we are focussing on shocks to income during the period of "adulthood" in which this is the case. We then have $c_a > 0$. The equivalence of the marginal products in the first-order condition fixes the contribution s_a to reducing mortality, and hence all variation in y_a will translate directly into variation of c_a , with no variation in s_a . In particular, there will exist an equilibrium level s_a^* , with fertility μ_a then given by

$$\mu_a = \alpha_a (y_a - s_a^*). \tag{21}$$

4.2.1 Idiosyncratic Shocks to Income

Equation (21) indicates that in this simple model, variations in income at age b translate linearly into variations in fertility at age b , while having no effect on mortality. However, we have already established that idiosyncratic variations in fertility lead to population growth rates identical to those that would prevail if fertility is fixed at its mean level. Evolution will then select for preferences that are neutral with respect to idiosyncratic variations in income.

4.2.2 Aggregate Shocks to Income

Now consider aggregate variations in income. In particular, suppose we replace y_a by $y_a + \varepsilon \mathbf{d}(a - b)$. How does the growth rate λ now depend on ε ? As before, we interpret this as an approximation to a case in which there is large but finite shock to income over a narrow range of ages, anchored at age b and conveniently captured by a impulse shock to income at age b . In the model, aggregate impulse variations in income again leave mortality unaffected and translate linearly into aggregate variations in fertility. We can thus apply the results from Section 3.4.1 to conclude that aggregate variations in income are advantageous at early adult ages and disadvantageous at later ages:

Proposition 4 *There exists an age \bar{b} such that small, rare aggregate shocks to income at age $b < \bar{b}$ are advantageous, while small, rare aggregate shocks to income at age $b > \bar{b}$ are disadvantageous.*

We can thus expect adults to be risk-preferring at relatively young ages, for aggregate shocks to income, but to eventually become risk-averse and remain so. For example, investments in the stock market entail risks that are correlated across investors, and so we might then not be surprised to see younger individuals participate in the market, while older individuals often divest themselves of stocks, while continuing to save using other (but more idiosyncratic) risky instruments.²³

To make this economic prediction more explicit, reconsider the simple case with constant fertility, μ , and mortality ρ , over $[0, \infty)$. Suppose also that the conversion of resources to fertility is age-independent so that $\alpha_a = \alpha$ in (21). From (15), it follows that $R_A(y, b) = \alpha R_A(\mu, b) = 2\alpha e^{-(\lambda+\rho)b}(b(\lambda + \rho) - 1)$, where $R_A(y, b)$ is the coefficient of absolute risk-aversion over income at age b . This gives us an explicit prediction, for this special case, of how $R_A(y, b)$ depends on age b .

There is evidence suggesting that our model is not an unreasonable first approximation. The popular press has given considerable attention to research, involving animal experiments (often rats) showing that the restriction of food intake may prolong life. The effect of calorie restriction on human longevity has not been empirically established, despite the enormous inherent interest, but there is a view that such an effect might be small (e.g., Page and Rose [15]).

²³There is some evidence that bears out this prediction. For example, Dohmen, Falk, Golsteyn, Huffman and Sunde [5, Abstract] state that "...willingness to take risks decreases over the life course, linearly until age 65 after which the slope becomes flatter". However, they do not break out stock holdings, perhaps the leading example of modern aggregate risk. Ameriks and Zeldes [1, Abstract] claim more specifically that "We find no evidence of a gradual reduction in portfolio shares with age. There is some tendency for older individuals to shift completely out of the stock market around the time of annuitizations and withdrawals." Ameriks and Zeldes also point out that simple standard economic theory does not predict a reduction in the portfolio share of risky assets just because there are fewer periods left. If utility has constant relative risk aversion, for example, and there is no additional future income, the share of savings held in risky assets should be constant, even though the amount saved declines. Hence the current model may help understand a phenomenon that, although having almost the status of conventional wisdom, is not robustly predicted by standard theory.

A model in which variations in resources have only a small effect on mortality might then be a reasonable approximation of our evolutionary past. Lost in the publicity surrounding the effect of nutrition on mortality is the observation that variations in food intake have a larger effect on fertility than mortality. Indeed, the experiments show that a reduction in food intake often causes a *dramatic* reduction in fertility. (For an early paper in this literature, see Ball, Barnes and Visscher [3].) Once again, and for obvious reasons, these experiments have concentrated on nonhuman subjects (often rats), and the effects of calorie restriction on human fertility are relatively unexplored. However, if humans once experienced effects of food intake on fertility similar to those found now for other mammals, we are close to the current formulation.²⁴ That is, it seems likely that variations in food intake translate at least mostly into variations in fertility.

Of course, there is much work to be done in exploring variations on this model. Notice in particular that other technologies for converting income into fertility and mortality may give rise to nontrivial risk attitudes toward idiosyncratic risks.

5 Discussion

We have examined the simplest setting in which we can make the point that age-structures can give rise to differences in attitudes toward idiosyncratic and aggregate risk. Several extensions immediately suggest themselves.

5.1 Life Beyond Fertility

Life beyond menopause could not be evolutionarily optimal in the current model. A large literature has struggled to reconcile the obvious existence of post-menopausal survival with evolutionary arguments, with the grandmother hypothesis being the leading explanation (cf. Hawkes, O’Connell, Blurton Jones, Alvarez and Charnov [11]). Capturing the grandmother hypothesis would require the incorporation of intergenerational resource transfers into the model. We view this as a promising topic for further work, but postpone it because such transfers are most effectively incorporated in connection with simplifications that mask the risk considerations examined here.

5.2 Attitudes Toward Risk

Our basic finding is that evolution need not select for neutral attitudes toward aggregate risk. Psychological studies of risk attitudes suggest that a feeling of control is important in inducing people to be comfortable with risk.²⁵ Risks aris-

²⁴A mechanism that dramatically reduces fertility while holding constant or even reducing mortality might have been selected for its ability to defer reproduction from bad times until times are good again (see Halliday [9]).

²⁵See Slovic, Fischhoff and Lichtenstein [22] for an early contribution and Slovic [21] for a more recent discussion.

ing out of situations in which people feel themselves unable to affect the outcome cause considerably more apprehension than risks arising out of circumstances people perceive themselves to control. “Control” may be an evolutionarily convenient stand-in for an idiosyncratic risk. If so, then our seemingly irrational fear of uncontrolled risk may be a mechanism inducing an evolutionarily rational fear of aggregate risk.

We find here that, as expected, idiosyncratic risk to fertility or mortality is evolutionarily neutral. Aggregate shocks to fertility are advantageous when young and disadvantageous when old, while aggregate shocks to mortality are always advantageous. We obtained these results by examining rare and small shocks. This serves to illustrate a central set of possibilities, but much remains to be done in extending the results to a wider range of settings.

We are interested in evolutionary foundations for attitudes toward risk in an economic variable, income in particular. The simplification we adopt, that fertility is linear in income, implies risk neutral attitudes toward *idiosyncratic* risk in income. Another worthwhile extension would be to consider the more general case where fertility is a strictly concave function of income so that individuals are risk-averse with respect to idiosyncratic gambles in income.

6 Appendices

6.1 A Discrete Foundation

The derivation in Section 2 effectively assumes that a population subject to idiosyncratic risk grows at the same rate as would an equivalent deterministic population where each random variable is replaced by its mean, and assumes that the population will converge to steady-state growth. This section provides foundations.

6.1.1 The Discrete Model

Time is discrete and is measured in periods of length Δ . Section 6.1.2 will consider the case in which Δ approaches zero, but in this section we let $\Delta = 1$ and suppress it in the notation. As noted, we find it convenient to consider the case in which the upper bound A on the set of ages is finite. In this case, the results in this section are straightforward applications of results from Harris [10, Chapter II] and Athreya and Ney [2, Chapter V]. With somewhat more cumbersome notation, the results extend to the case in which age is unlimited (Harris [10, Chapter III]).

At the beginning of each period $t \in \{0, 1, 2, \dots\}$, we take a census of the population, which is characterized by a vector $\mathbf{N}(t) = (N_1(t), \dots, N_A(t))$ identifying the number of individuals of each age $1, \dots, A$.

After the census in each period, reproduction occurs, giving rise to newborns. An individual of each age $a = 1, \dots, A$ has probability $m_a(x)$ of having $x \in \{0, 1, \dots, X\}$ offspring, where this is independent across individuals. Each individual then either survives until the next period, growing one period older

in the process, or dies. An individual of age $a = 0, \dots, A-1$ survives to age $a+1$ with probability $\pi_a \in (0, 1]$, while all individuals of age A die.

We then arrive at period $t+1$ and the next census. We assume for simplicity that each individual's survival is independent of that individual's fertility. Further, and more importantly, all individuals take independent draws from the relevant birth and death distributions.

For each age $a \in \{1, \dots, A\}$, we can think of an A -dimension random variable taking realizations of the form (N_1, \dots, N_A) , specifying the number of members of each age emanating in the next period from a single individual of age a in the current period. (The notation N_a here has distinct meaning from that in effect elsewhere as the age structured population.) These random variables are independent across ages, individuals, and time periods, and for each age are identically distributed across individuals and time periods.

It is convenient to describe the random variables in terms of a collection of A moment generating functions $f_a : [0, 1]^A \rightarrow \mathbb{R}_+$, one for each age $1, \dots, A$, where

$$f_a(s_1, \dots, s_A) = \sum_{(N_1, \dots, N_A)} r_a(N_1, \dots, N_A) \prod_{k=1}^A s_k^{N_k},$$

where $r_a(N_1, \dots, N_A)$ is the probability that an age- a individual gives rise to N_k individuals of age k for $k \in \{1, \dots, A\}$. (The (s_1, \dots, s_A) are formal variables.) In this particular case, an individual of age a can give rise only to individuals of age $a+1$ (by surviving until the next period, if $a < A$) and of individuals of age 1 (by having offspring that survive). Hence, we can write our generating functions as

$$\begin{aligned} f_1(s_1, s_2) &= \sum_{N_1, N_2} r_1(N_1, N_2) s_1^{N_1} s_2^{N_2} \\ f_2(s_1, s_3) &= \sum_{N_1, N_3} r_2(N_1, N_3) s_1^{N_1} s_3^{N_3} \\ &\vdots \\ f_{A-1}(s_1, s_A) &= \sum_{N_1, N_A} r_{A-1}(N_1, N_A) s_1^{N_1} s_A^{N_A} \\ f_A(s_1) &= \sum_{N_1} r_A(N_1) s_1^{N_1}. \end{aligned}$$

Inserting the relevant expressions for the probabilities and rearranging, the moment generating functions can be shown to reduce to the convenient forms:

$$\begin{aligned} f_a(s_1, s_{a+1}) &= (\pi_a s_{a+1} + (1 - \pi_a)) \sum_{x=0}^X m_a(x) \{\pi_0 s_1 + (1 - \pi_0)\}^x, \quad a = 1, \dots, A-1 \\ f_A(s_1) &= \sum_{x=0}^X m_A(x) \{\pi_0 s_1 + (1 - \pi_0)\}^x. \end{aligned}$$

These expressions reflect the requirement that the coefficient of $(s_1)^{N_1}(s_{a+1})^{N_{a+1}}$ in the moment generating function $f_a(s_1, s_{a+1})$ is the probability that there are N_1 age 1 individuals and N_{a+1} age $a + 1$ individuals one period after starting with one age a individual. For example, consider the coefficient of $s_1 s_{a+1}$ in the expression for $f_a(s_1, s_{a+1})$, which, for $a = 1, \dots, A - 1$, is

$$\pi_a \{m_a(1)\pi_0 + m_a(2)2\pi_0(1-\pi_0) + m_a(3)3\pi_0(1-\pi_0)^2 + \dots + m_a(X)X\pi_0(1-\pi_0)^{X-1}\}.$$

This is the product of probabilities that the age a individual survives (given by π_a) and that exactly one of her offspring survives as well (the term in braces).

The moment generating functions provide a complete description of the stochastic population process, in a form that allows us to extract various types of useful information. Each moment generating function describes the implications of a single individual of a certain age. However, using the independence across individuals and ages, we can obtain a description of the entire population. If the current population is given by $(N_1(t), \dots, N_A(t))$ at time t , the moment generating function for the state of the population in the next period is

$$\left[\prod_{a=1}^{A-1} (f_a(s_1, s_{a+1}))^{N_a(t)} \right] (f_A(s_1))^{N_A(t)},$$

Iterating this map allows one to describe how the population evolves over time.

Our first application of the moment generating function allows us to describe the probability of extinction. Consider a population that initially consists of a single individual of age a , so that $N_a = 1$ and $N_{\tilde{a}} = 0$ for $\tilde{a} \neq a$. Define the probability of extinction for this population to be q_a . These probabilities are characterized by the equations

$$q_a = f_a(q_1, q_{a+1}) = [\pi_a q_{a+1} + (1 - \pi_a)] \sum_{x=0}^X m_a(x) (\pi_0 q_1 + (1 - \pi_0))^x, \quad a = 1, \dots, A-1,$$

and

$$q_A = f_A(q_1) = \sum_{x=0}^X m_A(x) (\pi_0 q_1 + (1 - \pi_0))^x.$$

To interpret these expressions, note that two things must occur if the line of descendants of the individual of age $a = 1, \dots, A - 1$ is to suffer extinction. First, all offspring must give rise to lines of descendants that suffer extinction. The probability of this event is given by the sum on the right side of the equation. Second, the individual in question must either die immediately, or survive until the next period, and then give rise to a line of descendants that suffers extinction. The probability of this combination of events is given by the term in square brackets on the right side of the equation. These events are independent, and so the overall probability of extinction is given by the product of these probabilities.

In general, again using the independence across individuals, if there is a population given by (N_1, \dots, N_A) , the probability of extinction is

$$\prod_{a=1}^A (q_a)^{N_a}.$$

Next, derivatives of moment generating functions identify expected numbers of descendants. Define an $A \times A$ matrix L as

$$L_{a,1} = \frac{\partial f_a(1,1)}{\partial s_1} = \sum_{x=1}^X m_a(x)x\pi_0 =: \mu_a\pi_0, \quad a = 1, \dots, A$$

and

$$L_{a,a+1} = \frac{\partial f_a(1,1)}{\partial s_{a+1}} = \pi_a, \quad a = 1, \dots, A-1,$$

where all other terms $L_{a,\bar{a}}$ are 0. Then we can interpret $L_{a,1}$ as the expected number of surviving offspring in the next period, from an individual of age a in the current period, and can interpret $L_{a,a+1}$ as the expected number of age $a+1$ individuals next period as a result of the survival of an individual who is age a in the current period. Putting these terms together, we have the ‘‘Leslie’’ matrix

$$L = \begin{bmatrix} \mu_1\pi_0 & \pi_1 & \dots & 0 \\ \mu_2\pi_0 & 0 & \dots & 0 \\ & & \vdots & \\ \mu_{A-1}\pi_0 & 0 & \dots & \pi_{A-1} \\ \mu_A\pi_0 & 0 & \dots & 0 \end{bmatrix}.$$

We say that the process is positively regular if there exists K such that $L^K \gg 0$. One sufficient condition for this is that $\mu_1 > 0$ and $\mu_A > 0$. We say that the process is nonsingular if it is not the case that all of the functions f_a are linear.

The central result in the theory of branching processes is that the matrix L of the derivatives of the generating functions determines the fate of the population. Let $\Lambda > 0$ be the Frobenius root (or dominant eigenvalue) of L .

Proposition 5 *Let the population process be positively regular and nonsingular. Then:*

[5.1] *If $\Lambda \leq 1$ then $q_a = 1$ for all a , so extinction occurs with probability 1 given any initial population.*

[5.2] *If $\Lambda > 1$, then $q_a < 1$ for all a , so that extinction has probability strictly less than 1 for any initial population. Furthermore, every finite nonzero state is transient. Finally, almost surely,*

$$\frac{\mathbf{N}(t)}{\Lambda^t} \rightarrow vW, \text{ as } t \rightarrow \infty,$$

where $v \gg 0$ is the left eigenvector of L and $W \geq 0$ is a random variable with the property that $\Pr(W > 0) > 0$.

Proof See Athreya and Ney [2, Theorem 3.2, p. 186 and Theorem 6.1, p. 192] ■

When $\Lambda > 1$, the process *may* go extinct, with the probability of extinction given by the probability that the random variable W equals zero. However,

with strictly positive probability (i.e., with $\Pr(W > 0)$) it does not go extinct. It then explodes, growing to infinity with growth factor Λ and settling down to a steady population distribution given by v . The population is eventually described by a deterministic growth path exhibiting the constant growth rate $\lambda = \ln \Lambda$ and a constant age structure. The level of this growth path depends upon chance events that affect the size of the population when small, which are captured by the random variable W , while the rate of growth does not depend on such events.

The growth factor Λ satisfies the characteristic equation of the matrix M ,

$$1 = \sum_{a=1}^A \frac{(\prod_{\tilde{a}=0}^{a-1} \pi_{\tilde{a}}) \mu_a}{\Lambda^a}. \quad (22)$$

This is the counterpart of the Euler-Lotka equation given by (5).

We have thus established the result that in the discrete model, the long-run growth rate of a population subject to idiosyncratic risk concerning fertility or survival matches that of a population growing deterministically at mean rates.

6.1.2 Taking the Limit

We now show that the Euler-Lotka equation (5) for a continuous-age model is the limit of the Euler-Lotka equations (22) for a converging sequence of discrete-age models.

Our candidate limit for such a sequence is a continuous-age model with continuous fertility and mortality schedules μ_a and ρ_a defined on the interval $[0, A]$. We index the elements in the sequence of discrete-age models by Δ , interpreted as the length of a period, and consider the limit as $\Delta \rightarrow 0$. We then refer to an element of our sequence as “model Δ .”

The finite set of ages for model Δ is given by $G(\Delta) = \{\Delta, 2\Delta, \dots, 1, 1 + \Delta, 1 + 2\Delta, \dots, A\}$. That is, we limit attention to Δ that divide evenly into 1 and hence A . We now think of the fertility rate μ_a from the continuous specification as applying for each age $a \in G(\Delta)$, as a constant rate, over the interval $[a, a + \Delta]$. This generates a fertility level

$$\mu_a \Delta$$

for each age $a \in G(\Delta)$ in model Δ . For mortality, we suppose similarly that ρ_a applies, as a constant rate, over the interval $[a, a + \Delta]$ for each $a \in G(\Delta)$. This implies that the survival probability from age a to age $a + \Delta$ is now given by

$$\pi_a(\Delta) = e^{-\rho_a \Delta}$$

for each age $a \in G(\Delta)$ in model Δ .

Each discrete model Δ exhibits idiosyncratic risk—the fertility and mortality outcomes of the agents in model Δ are outcomes of independent draws from distributions, as in Section 6.1, summarized by $\mu_a \Delta$ and $\pi_a(\Delta)$. Taking the limit as $\Delta \rightarrow 0$ entails a larger number of such independent draws.

We are interested in the resulting steady-state configurations. Proposition 5 indicates that the steady-state configuration for model Δ is described by the Euler-Lotka equation (22) for model Δ , which we can write as

$$1 = \sum_{a \in G(\Delta)} \frac{(\prod_{\bar{a}=0}^{\bar{a}=a-1} \pi_{\bar{a}}(\Delta)) \mu_a \Delta}{\Lambda^a}. \quad (23)$$

In this expression, Λ denotes the growth factor that applies over an interval of fixed length 1. Our task is then to show that the sequence of Euler-Lotka equations given by (23) converges (as $\Delta \rightarrow 0$) to the continuous-time Euler-Lotka equation given by (5).

A straightforward technical result for evaluating the limit as $\Delta \rightarrow 0$ is:

Lemma 1 *Suppose that $g(a, \Delta)$ is a twice continuously differentiable function $g : [0, A] \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$, with $g(a, 0) = 0$. Consider a grid $G(\Delta) = \{\Delta, 2\Delta, \dots, A\}$, for any $\Delta > 0$ and $A > 0$. Then*

$$\sum_{a \in G(\Delta)} g(a, \Delta) \rightarrow \int_0^A \frac{\partial g(a, 0)}{\partial \Delta} da$$

as $\Delta \rightarrow 0$.

Proof. Taylor's Theorem implies that

$$g(a, \Delta) = \Delta \frac{\partial g(a, 0)}{\partial \Delta} + \frac{\Delta^2}{2} \frac{\partial^2 g(a, \chi)}{\partial \Delta^2},$$

where $\chi \in (0, \Delta)$. Because $g(a, \Delta)$ is a twice continuously differentiable function, we have $\sum_{a \in G(\Delta)} \Delta \frac{\partial g(a, 0)}{\partial \Delta} \rightarrow \int_0^A \frac{\partial g(a, 0)}{\partial \Delta} da$ and $\sum_{x \in G(\Delta)} \Delta \frac{\partial^2 g(a, \chi)}{\partial \Delta^2} \rightarrow \int_0^A \frac{\partial^2 g(a, 0)}{\partial \Delta^2}$. Hence $\Delta \sum_{a \in G(\Delta)} \Delta \frac{\partial^2 g(a, \chi)}{\partial \Delta^2} \rightarrow 0$ and so we have $\sum_{a \in G(\Delta)} g(a, \Delta) \rightarrow \int_0^A \frac{\partial g(a, 0)}{\partial \Delta} da$. \blacksquare

Consider now the limit of our sequence of discrete-time models. By construction, the survival factors in the discrete-time model are given by $\pi_a(\Delta) = e^{-\rho_a \Delta}$. It follows that the probability of surviving to age a is given by p_a where $\ln p_a = \sum_{\bar{a} \in G(\Delta), \bar{a} \leq a} \ln(\pi_{\bar{a}}(\Delta)) = \sum_{a \in G(\Delta), \bar{a} \leq a} \ln(e^{-\rho_{\bar{a}} \Delta})$. If we let $g(a, \Delta) = \ln(e^{-\rho_a \Delta})$, then $g(a, \Delta)$ satisfies the conditions of Lemma 1. Since $\frac{\partial g(a, 0)}{\partial \Delta} = -\rho_a$, it follows from Lemma 1 that $\ln p_a \rightarrow -\int_0^a \rho_{\bar{a}} d\bar{a}$ and hence $\prod_{\bar{a}=0}^{\bar{a}=a-1} \pi_{\bar{a}}(\Delta) = p_a \rightarrow e^{-\int_0^a \rho_{\bar{a}} d\bar{a}}$.

From this, it follows that:

Proposition 6 *The discrete Euler-Lotka equation converges to its continuous counterpart, i.e.,*

$$\sum_{a \in G(\Delta)} \frac{(\prod_{\bar{a}=0}^{\bar{a}=a-1} \pi_{\bar{a}}(\Delta)) \mu_a \Delta}{\Lambda^a} \longrightarrow \int_0^\infty \exp \left\{ -\int_0^a \rho_{\bar{a}} d\bar{a} \right\} \mu_a e^{-\lambda a} da,$$

where $\lambda = \ln \Lambda$.

Using (23) and (1), this gives us the Euler-Lotka equation (5).

Much as in the discrete case, the population governed by the Euler-Lotka equation (5) grows if $\lambda > 0$. That is, for an arbitrary initial age distribution, the population has a deterministic growth path that converges to a steady-state population configuration with the constant growth rate λ and a constant age structure. See Frauenthal [8, pp. 131–134] and Kot [13, 356–357] for details.

6.1.3 The Criterion for Evolutionary Success

Our criterion for evolutionary success is to maximize the “ultimate” growth rate, by which we mean the long-run growth rate conditional on escaping extinction. If a type escapes extinction, then it eventually grows without bound at this rate. This criterion is standard in the literature, but raises three questions.

First, what if there are two types, one with a higher ultimate growth rate, but also a higher probability of extinction? This issue is addressed in Robson (1996), which supposes that there is some small probability of each type mutating into the other.²⁶ Once we bring these mutations into the model, the various types are bound together—either they all go extinct or none of them do. Further, if mutations are rare, the growth rate of the integrated population is close to the growth rate of the that type with the highest ultimate growth rate, and this type dominates the population. In this sense, then, evolution selects the type with the highest ultimate growth rate, as captured by our model.

Second, populations clearly cannot increase without bound, contrary to the model. How does this affect the current analysis? In principle, a more realistic but also more complex model would incorporate a feedback from population size to the probability distribution of offspring, so that the expected number of offspring would sink to 1 as the carrying capacity of the environment is approached.²⁷ It remains true, of course, that the type with the maximum ultimate growth at carrying capacity will dominate the population. Hence, as long as our model captures the ranking of ultimate growth rates that prevails near the carrying capacity of the environment, our results apply, and it is convenient to simplify the analysis by removing considerations of carrying capacity.

Third, there is a more subtle issue that carrying capacity introduces. Suppose that the type that dominates the population at carrying capacity does so with an offspring distribution that features a positive probability of zero offspring, as would generally be the case. Then there is only one possible long run outcome—eventually all individuals in a given generation will obtain the zero outcome and the type will become extinct. Surely this would introduce some degree of aversion to idiosyncratic risk? If so, carrying capacity would have to enter the model.

How important this consideration is depends on how large the carrying capacity is. There are currently estimated to be only three northern white

²⁶This assumption is biologically plausible, as long as the time considered is long. Indeed, the very existence of a type confirms the existence of such mutations.

²⁷In general, the differential success of different types might vary as the population size approaches carrying capacity.

rhinoceros in the world. Idiosyncratic fertility and mortality draws pose a very real extinction threat to such animals, but it is unlikely that the population will remain positive but so small long enough for this to play a role in the evolutionary selection of risk attitudes. On the other hand, it is possible, in principle, that for the next century, there are no female births among humans, so driving the population extinct. Given the billions of reproductive age women in the world, this is prodigiously unlikely, and little harm is done by neglecting it. In general, a population that survives long enough for evolution to shape risk attitudes will be large enough that idiosyncratic realizations pose little risk of extinction. The phenomena of interest here are thus well approximated in a model that allows populations to grow without bound, while conditioning on escaping extinction.

6.2 Proof of Proposition 1

The population growth factor. The growth factor $P(T) = N(T)/N(0)$ is the product of a number of growth factors, where each factor corresponds to a sojourn in some particular state n . Let T_n be the total amount of time, in the interval $[0, T]$, that the population spends in state n (so $\sum_{n=1}^N T_n = T$). This time will be comprised of W_n sojourns in state n of lengths $\tau_n^1, \dots, \tau_n^{W_n}$ (where $\tau_n^1 + \dots + \tau_n^{W_n} = T_n$). If we let P_n^w denote the growth factor associated with sojourn w in state n , then we can write (suppressing the dependence of variables in the final term on T , to keep the notation uncluttered)

$$P(T) = \prod_{n=1}^N P_n(T) = \prod_{n=1}^N \prod_{w=1}^{W_n} P_n^w,$$

where $P_n(T)$ denotes the product $\prod_{w=1}^{W_n} P_n^w$ of the growth factors associated with the sojourns in state n . That is, the growth factor P_n^w is $N(\bar{t}_n^w)/N(\underline{t}_n^w)$ where $[\underline{t}_n^w, \bar{t}_n^w]$ is the w th sojourn in state n in $[0, T]$.

In order to verify (6), fix a value ε .

Fixed state n . We now fix a state n and examine $P_n(T)$. Throughout this portion of the argument, whenever possible, we omit the state subscript on variables, again to keep the notation uncluttered.

Within any sojourn w in state n , growth proceeds deterministically. It follows that there exists an upper bound on growth rates, so that

$$\frac{\ln P^w}{\tau^w} \leq m \tag{24}$$

for $m > 0$. This bound holds uniformly in the initial population distribution, and we can choose a bound that holds across all states.²⁸

²⁸Suppose that $m = \max_{n,a} \mu_a^n$ where μ_a^n is fertility at age a in state n . Consider a population where all individuals have fertility m and there is no mortality. Clearly the growth rate of this population, which is always m , regardless of age structure, exceeds that of the population in any state.

Next, choose a value ε' such that

$$\lambda_n - \varepsilon \leq (\lambda_n - \varepsilon')(1 - \varepsilon') \leq (\lambda_n + \varepsilon')(1 + \varepsilon') + m\varepsilon' \leq \lambda_n + \varepsilon \quad (25)$$

holds for all states.

Invoking the arguments in Feller [7, Ch XI], we can choose $\hat{\tau}$ such that

$$\frac{\ln P^w}{\tau^w} \in [\lambda_n - \varepsilon', \lambda_n + \varepsilon'] \quad (26)$$

for all sojourn lengths $\tau^w \geq \hat{\tau}$, uniformly in the initial population distribution and across states.²⁹

Let $\bar{\tau}$ be the expected length of a sojourn in state n . Notice that $\bar{\tau}$ increases without bound as k increases. Hence, we can choose K such that for all $k \geq K$,

$$\Pr\{\tilde{\tau} \leq \hat{\tau}\} \leq \varepsilon' \quad (27)$$

$$\hat{\tau} \leq \bar{\tau} \quad (28)$$

$$1 \leq \frac{\mathbb{E}\{\tilde{\tau} | \tilde{\tau} \geq \hat{\tau}\}}{\bar{\tau}} \leq 1 + \varepsilon'. \quad (29)$$

We can again ensure that this holds for all states.

We can rewrite the equality $P_n(T) = \prod_{w=1}^{W_n} P_n^w$ as

$$\frac{\ln P_n(T)}{T_n} = \sum_{\tau^w < \hat{\tau}} \frac{\ln P^w}{\tau^w} \frac{\tau^w}{T_n} + \sum_{\tau^w \geq \hat{\tau}} \frac{\ln P^w}{\tau^w} \frac{\tau^w}{T_n}. \quad (30)$$

Using (24) allows us to bound the second term on the right side by

$$0 \leq \sum_{\tau^w < \hat{\tau}} \frac{\ln P^w}{\tau^w} \frac{\tau^w}{T_n} \leq \frac{mW}{T_n} \sum_{\tau^w < \hat{\tau}} \frac{\tau^w}{W}.$$

Taking a limit and using (27) and (28) for the second inequality gives

$$0 \leq \lim_{T \rightarrow \infty} \sum_{\tau^w < \hat{\tau}} \frac{\ln P^w}{\tau^w} \frac{\tau^w}{T_n} \leq \frac{m}{\bar{\tau}} \mathbb{E}\{\tilde{\tau} | \tilde{\tau} < \hat{\tau}\} \Pr\{\tilde{\tau} < \hat{\tau}\} \leq m\varepsilon',$$

almost surely. Turning to the first term on the right side of (30), we have (using (26) for the first inequality and (29) for the second)

$$\lim_{T \rightarrow \infty} \sum_{\tau^w \geq \hat{\tau}} \frac{\ln P^w}{\tau^w} \frac{\tau^w}{T_n} \leq \frac{1}{\bar{\tau}} (\lambda_n + \varepsilon') \mathbb{E}\{\tilde{\tau} | \tilde{\tau} \geq \hat{\tau}\} \Pr\{\tilde{\tau} \geq \hat{\tau}\} \leq (\lambda_n + \varepsilon')(1 + \varepsilon'),$$

almost surely, and that (again using (26) for the first inequality and (29) for the second)

$$\lim_{T \rightarrow \infty} \sum_{\tau^w \geq \hat{\tau}} \frac{\ln P^w}{\tau^w} \frac{\tau^w}{T_n} \geq \frac{1}{\bar{\tau}} (\lambda_n - \varepsilon') \mathbb{E}\{\tilde{\tau} | \tilde{\tau} \geq \hat{\tau}\} \Pr\{\tilde{\tau} \geq \hat{\tau}\} \geq (\lambda_n - \varepsilon')(1 - \varepsilon'),$$

²⁹A detailed proof of this assertion is available at <http://www.sfu.ca/~robson/RSA> and at <https://sites.google.com/site/larryatyale/home/papers>.

almost surely. Our choice of ε' in (25) then ensures that

$$\lim_{T_n} \frac{1}{T_n} \ln P_n(T) \in [\lambda_n - \varepsilon, \lambda_n + \varepsilon],$$

almost surely, for all states $n = 1, \dots, N$.

Combining states. The growth factor at time T is then

$$P(T) = P_1(T) \dots P_N(T)$$

Hence, recalling that $T_1 + \dots + T_N = T$,

$$\frac{1}{T} \ln P(T) = \sum_n \frac{T_n}{T} \left(\frac{\ln P_n(T)}{T_n} \right),$$

so that (using Durrett [6, Theorem 3.1, p. 169] for the first equality)

$$\lim_{T} \frac{1}{T} \ln P(T) = \sum_n \pi_n \lim_{T_n} \frac{\ln P_n(T)}{T_n} \in \left[\sum_n \pi_n \lambda_n - \varepsilon, \sum_n \pi_n \lambda_n + \varepsilon \right],$$

almost surely, which gives (6).

6.3 Calculations for Section 3.3

We first consider how a lag to first reproduction also induces greater aversion to aggregate risk concerning fertility in a discrete model. First, suppose there is no age structure. In each period $t = 1, 2, \dots$, each individual in the population has μ offspring and survives herself. Letting $N(t)$ be the population size at time t , with $N(0) = 1$, we have

$$N(t) = (1 + \mu)N(t-1) \text{ so that } N(t) = (1 + \mu)^t \text{ for } t = 0, 1, 2, \dots$$

The population growth rate is then given by

$$\lambda = \frac{\ln N(t)}{t} = \ln(1 + \mu).$$

The growth rate λ is thus a concave function of the fertility rate μ , indicating that aggregate risk in fertility is disadvantageous. Aggregate risk has an effect here, despite the lack of an age structure, because of the discrete formulation. The Arrow-Pratt measure of absolute risk aversion in this case is given by

$$R_A^N = -\frac{\frac{d^2 \lambda}{d\mu^2}}{\frac{d\lambda}{d\mu}} = \frac{1}{(1 + \mu)}.$$

Now suppose that reproduction does not begin until the second period of one's life. The population is then governed by the difference equation

$$N(t) = N(t-1) + \mu N(t-2),$$

with initial conditions $N(0) = N(1) = 1$. This is a straightforward second-order difference equation, whose solution is given by

$$N(t) = \frac{1}{\sqrt{1+4\mu}} \left[\left(\frac{1+\sqrt{1+4\mu}}{2} \right)^{t+1} - \left(\frac{1-\sqrt{1+4\mu}}{2} \right)^{t+1} \right],$$

with growth rate

$$\begin{aligned} \lim_{t \rightarrow \infty} \lambda &= \lim_{t \rightarrow \infty} \frac{\ln N(t)}{t} = \lim_{t \rightarrow \infty} \frac{\ln \left[\left(\frac{1+\sqrt{1+4\mu}}{2} \right)^{t+1} - \left(\frac{1-\sqrt{1+4\mu}}{2} \right)^{t+1} \right]}{t} \\ &= \ln \frac{1+\sqrt{1+4\mu}}{2}. \end{aligned}$$

This gives us the growth rate λ as a concave function of μ , and hence aggregate risk in fertility is disadvantageous. We can now calculate:

$$\begin{aligned} \frac{d\lambda}{d\mu} &= \frac{2}{1+4\mu+\sqrt{1+4\mu}} \\ \frac{d^2\lambda}{d\mu^2} &= -\frac{8+4(1+4\mu)^{-1/2}}{(1+4\mu+\sqrt{1+4\mu})^2} \\ R_A^L &= -\frac{\frac{d^2\lambda}{d\mu^2}}{\frac{d\lambda}{d\mu}} = \frac{4\sqrt{1+4\mu}+2}{1+4\mu+(1+4\mu)\sqrt{1+4\mu}}. \end{aligned}$$

Comparing to the case of no age structure, it follows that, as long as μ is not large, the measure of risk aversion is larger for the case of the structure with a reproductive lag.³⁰ As $\mu \rightarrow \infty$, both measures of risk aversion approach 0.

We now turn to the example in Section 3.3 and show that aversion to aggregate risk arises in the case $M > 0$ and $A < \infty$. We have the relationship

$$\mu = \frac{\lambda}{e^{-\lambda M} - e^{-\lambda A}} =: \frac{\lambda}{f(\lambda)}.$$

To establish that λ is a concave function of μ , and hence that aggregate risk is disadvantageous, we show that $d^2\mu/d\lambda^2 > 0$. We first calculate, with abbreviated notation,

$$\begin{aligned} \frac{d\mu}{d\lambda} &= \frac{f - f'\lambda}{f^2} \\ \frac{d^2\mu}{d\lambda^2} &= \frac{f(-f''\lambda) - (f - f'\lambda)2f'}{f^3}. \end{aligned}$$

We thus need to show

$$2(f')^2\lambda > ff''\lambda + 2ff'.$$

³⁰If $x = \sqrt{1+4\mu}$, it follows readily that $R_A^L > R_A^N$ if and only if $x < 3 + 2\sqrt{3}$ or if and only $\mu < 10.2$, approximately, a bound that includes all plausible human levels of fertility. For larger values of μ , the comparison is reversed.

We then note that

$$\begin{aligned} f &= e^{-\lambda M} - e^{-\lambda A} \\ f' &= Ae^{-\lambda A} - Me^{-\lambda M} \\ f'' &= M^2e^{-\lambda M} - A^2e^{-\lambda A}. \end{aligned}$$

Inserting these terms and simplifying, it follows that we need

$$\lambda(A-M)^2e^{-\lambda(A+M)} + \lambda(Ae^{-\lambda A} - Me^{-\lambda M})^2 + 2(Me^{-2\lambda M} + Ae^{-2\lambda A}) > 2(A+M)e^{-\lambda(A+M)}.$$

This inequality holds for all $\lambda > 0$ and $0 < M < A < \infty$.³¹

6.4 Relation to Robatto and Szentes

Robatto and Szentes [16] argue that aggregate shocks to fertility and mortality rates, like idiosyncratic shocks, are evolutionarily neutral, in the sense that a population subject to such shocks exhibits the same long-run growth rate as a population characterized by the corresponding mean fertility and mortality rates but subject to no shocks. For a population in which mortality and fertility are independent of age, their results are consistent with the current paper. That is, our analysis echoes their striking result that, although aggregate risk generally has observable (stochastic) consequences, it induces the same relevant long-run growth rate as comparable idiosyncratic risk.

³¹After some algebra, this inequality can be shown to be equivalent to

$$\lambda A^2 + \lambda M^2 + \lambda A^2 e^{\lambda(M-A)} + \lambda M^2 e^{\lambda(A-M)} + 2M e^{\lambda(A-M)} + 2A e^{\lambda(M-A)} > 4\lambda AM + 2A + 2M (*).$$

This holds as an equality in the limit when $\lambda = 0$. It then suffices that, for $A > M$ and $\lambda > 0$, the first derivative of the left hand side of (*) in λ exceeds the first derivative of the right hand side in λ . This inequality is

$$\begin{aligned} &A^2 + M^2 + A^2 e^{\lambda(M-A)} + M^2 e^{\lambda(A-M)} \\ &+ \lambda A^2 (M-A) e^{\lambda(M-A)} + \lambda M^2 (A-M) e^{\lambda(A-M)} \\ &+ 2M(A-M) e^{\lambda(A-M)} + 2A(M-A) e^{\lambda(M-A)} > 4AM. \end{aligned}$$

Again, this holds as an equality in the limit when $\lambda = 0$. It then suffices to show that the second derivative of the left hand side of (*) is strictly positive when $\lambda > 0$ (the second derivative of the right hand side is clearly zero). The second derivative of the left hand side of (*) is

$$\begin{aligned} &2A^2(M-A) e^{\lambda(M-A)} + 2M^2(A-M) e^{\lambda(A-M)} \\ &+ \lambda A^2 (M-A)^2 e^{\lambda(M-A)} + \lambda M^2 (A-M)^2 e^{\lambda(A-M)} \\ &+ 2M(A-M)^2 e^{\lambda(A-M)} + 2A(M-A)^2 e^{\lambda(M-A)}. \end{aligned}$$

This expression has the form $T_1 e^{\lambda(M-A)} + T_2 e^{\lambda(A-M)}$, where $T_2 > 0$. Furthermore, it follows readily that $T_1 + T_2 > 0$. If $T_1 \geq 0$, the second derivative of the left hand side of (*) is positive and we have the desired result. If $T_1 < 0$ then

$$T_2 e^{\lambda(A-M)} > T_2 > -T_1 > -T_1 e^{\lambda(M-A)}$$

and we also have the result.

In an online Appendix Robatto and Szentes [16] argue that there is a more general sense in which aggregate risk is equivalent to idiosyncratic risk, no matter what the population's age structure. Since we do not obtain such a general equivalence for an age-structured population in continuous time, this appendix relates the two approaches.

We consider the model of Section 4.2, in which we found that aggregate shocks to income are advantageous when young and disadvantageous when old. It is convenient to define a state here as a value of the parameter ε scaling the shock to income at age b . Let the set of such states be Ω . With each state $\varepsilon \in \Omega$, we can associate a growth rate $\lambda(\varepsilon)$, and our (approximate) growth rate is given by

$$\lambda^* = \int_{\Omega} \lambda(\varepsilon) dG(\varepsilon), \quad (31)$$

where G is the ergodic distribution over states ε .

Let n be a steady-state age distribution, with value $n_a = \frac{N_a}{\int_0^A N_a da}$ at age a , and let \mathcal{N} be the set of all such age distributions n . Then the measure G will induce a distribution \tilde{G} over the set $\Omega \times \mathcal{N}$. Robatto and Szentes ([16], equation (A.9)) show that, under these assumptions, the long-run growth rate is

$$\lambda^* = \int_{\Omega \times \mathcal{N}} \int_0^A (\mu_a(\varepsilon) - \rho_a(\varepsilon)) n_a da d\tilde{G}(\varepsilon, n), \quad (32)$$

where $\mu_a(\varepsilon)$ is the fertility rate of an agent of age a in state ε and $\rho_a(\varepsilon)$ the mortality rate of an agent of age a in state ε . Notice that the population growth rate is given by the mean net fertility rate, where the mean is calculated with respect to the distribution over ages and states, as in the right side of (32).

We first show that the criterion (32), derived from Robatto and Szentes [16], is equivalent to our expression (31), in a simple case. Suppose that income is a constant y for all ages $a \in [0, \infty)$, except for the shock at age b , so that $y_a = y + \varepsilon \mathbf{d}(a - b)$. Section 4.2 shows that the mortality rate will remain unchanged, at rate we denote by ρ , so that

$$p_a = e^{-\rho a}$$

Fertility will be given by $\mu + \varepsilon \mathbf{d}(a - b)$, assuming that $\alpha_a = 1$, for simplicity, and the Euler-Lotka equation for a population permanently in state ε is then given by

$$1 = \int_0^{\infty} \mu e^{-(\lambda(\varepsilon) + \rho)a} da + \varepsilon e^{-(\lambda(\varepsilon) + \rho)b}.$$

We can solve this for

$$\mu = (\lambda(\varepsilon) + \rho)(1 - \varepsilon e^{-(\lambda(\varepsilon) + \rho)b}). \quad (33)$$

Because we are working with the limiting case of rare aggregate shocks, we can associate a unique age distribution $n_a(\varepsilon)$ with each state ε , where

$$n_a(\varepsilon) = (\lambda(\varepsilon) + \rho) e^{-(\lambda(\varepsilon) + \rho)a}.$$

Using the approximation that state ε appears only when coupled with age distribution $n_a(\varepsilon)$, condition (32) then becomes

$$\lambda^* = \int_{\Omega} \left(\int_0^{\infty} (\mu - \rho)(\lambda(\varepsilon) + \rho)e^{-(\lambda(\varepsilon) + \rho)a} da + \varepsilon(\lambda(\varepsilon) + \rho)e^{-(\lambda(\varepsilon) + \rho)b} \right) dG(\varepsilon). \quad (34)$$

We can solve for

$$\lambda^* = \int_{\Omega} \left((\mu - \rho) + \varepsilon(\lambda(\varepsilon) + \rho)e^{-(\lambda(\varepsilon) + \rho)b} \right) dG(\varepsilon) \quad (35)$$

$$= \int_{\Omega} \lambda(\varepsilon) dG(\varepsilon), \quad (36)$$

where we obtain (35) by performing the inner integration in (34),³² and then inserting the expression for μ given by (33) into (35) and simplifying to obtain (36). But this is precisely our expression for the population growth rate given by (31). Since the reverse implication is immediate, this establishes the equivalence of our expression for the growth rate and that derived by Robatto and Szentes.

We can illuminate the differences between the two approaches by examining the induced risk attitudes towards aggregate shocks to income. Suppose agents receive a draw from an aggregate random variable determining income as $y_a(\varepsilon) = y + \varepsilon \mathbf{d}(a - b)$, where ε has distribution given by G , and $\int_{\Omega} \varepsilon dG(\varepsilon) = 0$. We have shown that evolution will select for agents who, when young (i.e., when b is small), prefer to receive this random income rather than receive the mean income $\int_{\Omega} y(\varepsilon) dG(\varepsilon) = y$; but when they are old, they prefer the mean income y . Hence, evolution will select for agents who seek aggregate risk when young, but are averse to such risk when old.

On the other hand, Robatto and Szentes [16] argue that their condition given here as (32) implies agents are risk neutral with respect to this random income. We interpret this as the observation that individuals would be indifferent between this random income and receiving the mean income given by $\int_{\Omega} \left(\int_0^{\infty} y(\varepsilon) n_a(\varepsilon) da \right) dG(\varepsilon)$. This expression calculates the mean with respect to the joint distribution over population age distributions and the state. To compare to our results, we note that this expression is equal to

$$\int_{\Omega} \int_0^{\infty} (y + \varepsilon \mathbf{d}(a - b)) (\lambda(\varepsilon) + \rho) e^{-(\lambda(\varepsilon) + \rho)a} da dG(\varepsilon) = \int_{\Omega} T(\varepsilon) dG(\varepsilon),$$

where

$$T(\varepsilon) = y + \varepsilon(\lambda(\varepsilon) + \rho)e^{-(\lambda(\varepsilon) + \rho)b}.$$

If λ did not depend on ε , we would have $\int_{\Omega} T(\varepsilon) dG(\varepsilon) = y$, so that the random income is compared with the mean income y . But λ does depend on ε , and hence the requirement that the mean of ε with respect to G is 0 generally implies that the mean of ε with respect to the joint distribution over population distributions and the state is not 0.

³² $\int_0^{\infty} (\mu - \rho)(\lambda(\varepsilon) + \rho)e^{-(\lambda(\varepsilon) + \rho)a} da = \mu - \rho$.

Indeed, for any $b \geq 0$, it follows readily that $T''(0) = 2\lambda' e^{-(\lambda+\rho)b}(1 - (\lambda + \rho)b)$, so T is convex for $b < 1/(\lambda + \rho)$ but concave for $b > 1/(\lambda + \rho)$. Hence $\int_{\Omega} T(\varepsilon)dG(\varepsilon) > y$, for all $b < 1/(\lambda + \rho)$ but $\int_{\Omega} T(\varepsilon)dG(\varepsilon) < y$, for all $b > 1/(\lambda + \rho)$, for ε with sufficiently small support. This precisely reflects our results in Propositions 2 and 4.

Our analysis implies that aggregate shocks to income are advantageous relative to the mean of $\int_{\Omega} y(\varepsilon)dG(\varepsilon) = y$ at young ages, but then disadvantageous at older ages. Robatto and Szentes' [16] condition (32) equivalently implies that aggregate shocks to income are neutral when compared to a mean $\int_{\Omega} T(\varepsilon)dG(\varepsilon)$, which is higher than the mean income y at young ages, and lower at older ages. We consider a mean taken with respect only to the distribution of income, whereas Robatto and Szentes [16] could be interpreted as considering a mean taken with respect to the distribution over income and the endogenously generated equilibrium age distribution.

Notice that, in any case, there must be a discrepancy between the effect of idiosyncratic and aggregate risk. For example, adding aggregate risk with mean 0 to income y at $b = 0$ is strictly preferred to y ; however, adding idiosyncratic risk with mean 0 to y is equivalent to y .

References

- [1] John Ameriks and Stephen P. Zeldes. How do household portfolio shares vary with age. Working paper, Vanguard Group and Graduate School of Business, Columbia University, 2004.
- [2] K. B. Athreya and P. E. Ney. *Branching Processes*. Dover Publications, Mineola, New York, 2004.
- [3] Zelda B. Ball, Richard H. Barnes, and Maurice B. Visscher. The effects of dietary caloric restriction on maturity and senescence, with particular reference to fertility and longevity. *American Journal Physiology*, 150(3):511–519, 1947.
- [4] Elchanan Ben-Porath, Eddie Dekel, and Aldo Rustichini. On the relationship between mutation rates and growth rates in a changing environment. *Games and Economic Behavior*, 5(4):576–603, 1993.
- [5] Thomas Dohmen, Armin Falk, Bart H. H. Golsteyn, David Huffman, and Uwe Sunde. Risk attitudes across the life course. *Economic Journal*, 127(605):F95–F116, 2017.
- [6] Rick Durrett. *Essentials of Stochastic Processes*. Springer, New York, 1999.
- [7] William Feller. *An Introduction to Probability Theory and Its Applications, Vol. II*. John Wiley, New York, 2nd edition, 1971.

- [8] James C. Frauenthal. Analysis of age-structure models. In Thomas G. Hallam and Simon A. Levin, editors, *Mathematical Ecology*, pages 117–147. Springer-Verlag, New York, 1986.
- [9] Robin Halliday. Food, reproduction and longevity: Is the extended lifespan of calorie-restricted animals an evolutionary adaptation? *BioEssays*, 10(4):125–127, 1989.
- [10] Theodore E. Harris. *The Theory of Branching Processes*. Dover Publications, New York, 1989.
- [11] K. Hawkes, J. F. O’Connell, N. G. Blurton Jones, H. Alvarez, and E. L. Charnov. Grandmothering, menopause, and the evolution of human life histories. *Proceedings of the National Academy of Sciences of the United States of America*, 95(3):1336–1339, 1998.
- [12] J. Hofbauer and K. Sigmund. *The Theory of Evolution and Dynamical Systems*. Cambridge University Press, Cambridge, 1988.
- [13] Mark Kot. *Elements of Mathematical Ecology*. Cambridge University Press, Cambridge, 2001.
- [14] O. L. Mangasarian. Sufficient conditions for the optimal control of nonlinear systems. *SIAM Journal on Control*, 4(1):139–152, 1966.
- [15] John P. Phelan and Michael R. Rose. Why dietary restriction substantially increases longevity in animal models but won’t in humans. *Ageing Research Reviews*, 4:339–350, 2005.
- [16] Roberto Robatto and Balazs Szentes. On the biological foundation of risk preferences. *Journal of Economic Theory*, 172:410–422, 2017.
- [17] Arthur J. Robson. A biological basis for expected and non-expected utility. *Journal of Economic Theory*, 68(2):397–424, 1996.
- [18] Arthur J. Robson and Larry Samuelson. The evolution of time preference with aggregate uncertainty. *American Economic Review*, 99(5):1925–1953, 2009.
- [19] Arthur J. Robson and Larry Samuelson. The evolutionary foundations of preferences. In Alberto Bisin Jess Benhabib and Matthew Jackson, editors, *The Social Economics Handbook*, pages 21–310. North Holland, Amsterdam, 2011.
- [20] Hans-Werner Sinn and Alfrons J. Weichenrieder. The biological selection of risk preferences. In Bayerische Rückversicherung, editor, *Risk is a Construct - Perceptions of Risk Perception*. Knesebeck, Munich, 1993.
- [21] Paul Slovic. *The Perception of Risk*. Earthscan Publications, London, 2000.
- [22] Paul Slovic, Baruch Fishhoff, and Sarah Lichtenstein. Why study risk perception? *Risk Analysis*, 2(2):83–93, 1982.