

# Lifetime reproductive output: individual stochasticity, variance, and sensitivity analysis

Silke F. van Daalen<sup>1</sup>  · Hal Caswell<sup>1</sup> 

Received: 23 December 2016 / Accepted: 16 March 2017 / Published online: 17 April 2017  
© The Author(s) 2017. This article is published with open access at Springerlink.com

**Abstract** Lifetime reproductive output (LRO) determines per-generation growth rates, establishes criteria for population growth or decline, and is an important component of fitness. Empirical measurements of LRO reveal high variance among individuals. This variance may result from genuine heterogeneity in individual properties, or from individual stochasticity, the outcome of probabilistic demographic events during the life cycle. To evaluate the extent of individual stochasticity requires the calculation of the statistics of LRO from a demographic model. Mean LRO is routinely calculated (as the net reproductive rate), but the calculation of variances has only recently received attention. Here, we present a complete, exact, analytical, closed-form solution for all the moments of LRO, for age- and stage-classified populations. Previous studies have relied on simulation, iterative solutions, or closed-form analytical solutions that capture only part of the sources of variance. We also present the sensitivity and elasticity of all of the statistics of LRO to parameters defining survival, stage transitions, and (st)age-specific fertility. Selection can operate on variance in LRO only if the variance results from genetic heterogeneity. The potential opportunity for selection is quantified by Crow's index  $\mathcal{I}$ , the ratio of the variance to the square of the mean. But variance due to individual stochasticity is only an *apparent* opportunity for selection. In a comparison of a

range of age-classified models for human populations, we find that proportional increases in mortality have very small effects on the mean and variance of LRO, but large positive effects on  $\mathcal{I}$ . Proportional increases in fertility increase both the mean and variance of LRO, but reduce  $\mathcal{I}$ . For a size-classified tree population, the elasticity of both mean and variance of LRO to stage-specific mortality are negative; the elasticities to stage-specific fertility are positive.

**Keywords** Lifetime reproductive output · Matrix population models · Individual stochasticity · Sensitivity analysis · Markov chains with rewards · Opportunity for selection · Inter-individual variance

## Introduction

Like all men in Babylon, I have been proconsul; like all, I have been a slave. I have known omnipotence, ignominy, imprisonment ... I owe this almost atrocious variety to an institution which other republics know nothing about, and which operates among them imperfectly and in secret: the lottery.

Jorge Luis Borges, *The Lottery in Babylon*

Lifetime reproductive output (LRO) is, as the name implies, the total production of offspring over the lifetime of an individual<sup>1</sup> and is one of the most important characteristics of an individual life history. The expectation of LRO, calculated in terms of female offspring per female, is the net reproductive rate  $R_0$ . In ecology, the critical value

✉ Hal Caswell  
h.caswell@uva.nl  
Silke F. van Daalen  
silkevandaalen@hotmail.com

<sup>1</sup> Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, PO Box 94248, 1090 GE, Amsterdam, The Netherlands

<sup>1</sup>This is a slightly more general concept than lifetime reproductive success (Clutton-Brock 1988) because it can accommodate many different operational definitions of reproduction, as are often encountered in ecological and demographic analysis.

$R_0 = 1$  defines the boundary separating population growth and persistence from population decline and extinction. In epidemiology,  $R_0$  for a pathogen determines whether a disease will or will not cause an outbreak. In evolutionary biology,  $R_0$  is a critical component of fitness (sometimes considered to be fitness, although that is sometimes problematic).

Genetic variance in LRO is considered to be the raw material on which natural selection operates. Crow (1958) introduced an index, which now bears his name, of the “opportunity for selection.” If  $X$  denotes some measure of fitness, then the opportunity for selection is

$$\mathcal{I} = \frac{V(X)}{E(X)^2}, \quad (1)$$

also known as the standardized variance. It gives the maximum rate of evolutionary change that could be produced by selection if all the variance in  $X$  were genetic. To rigorously investigate opportunities for selection would require more explicit population genetic models. Even so, the opportunity for selection is widely used in studies of both animal and human populations (e.g., Jones 2009; Brown et al. 2009; Robbins et al. 2011; Moorad et al. 2011; Courtiol et al. 2012).

Empirical studies of individual LRO routinely find large variance and usually a positive skew. Most individuals produce few, or no, offspring, while a few rare individuals produce many offspring (Clutton-Brock 1988; Newton 1989). These differences in LRO have two possible sources. One is heterogeneity—differences in the properties of individuals—including genetic heterogeneity, physiological differences, phenotypic plasticity, and environmental heterogeneity.

However, the differences may also be due to individual stochasticity (Caswell 2009, 2011, 2014; van Daalen and Caswell 2015). Individual stochasticity refers to differences among individuals due to the accumulation of random outcomes of the stochastic processes of mortality, growth, development, breeding, etc. Individual stochasticity would lead to variance among individuals even if they were totally identical and experienced exactly the same demographic rates.<sup>2</sup> Depending on the outcome under consideration, variance due to individual stochasticity can be as great as or even exceed that caused by unobserved heterogeneity (Tuljapurkar et al. 2009; Steiner et al. 2010; Caswell 2011, 2014; Hartemink et al. 2017). Thus, before invoking heterogeneity as the explanation for differences among individuals,

it is important to calculate the variation due to stochasticity as a kind of “neutral model” for variation (Steiner and Tuljapurkar 2012).

Individual stochasticity itself has two components. Consider an individual at some stage in its life cycle (say, at birth). The growth, development, and eventual death of this individual define a path through the stages of the life cycle. The pathways of two or more identical individuals, subject to the same rates at every stage of the life cycle, will differ randomly among themselves, and this variation among pathways is one source of individual stochasticity. At each step along an individual’s path, it may or may not reproduce. If it reproduces, the number of its offspring will be drawn from some probability distribution. This within-pathway stochastic variation in (stage-specific) fertility is the second source of individual stochasticity.

Our goal is to calculate the variance (and other statistics) of LRO, due to individual stochasticity, from basic demographic information, just as  $R_0$  can be calculated from age-classified or stage-classified demographic models (Rhodes 1940; Cushing and Zhou 1994; De Camino-Beck and Lewis 2007; Cushing and Ackleh 2012). To calculate the variance in LRO, we need to account for the generally infinite number of pathways through the life cycle, calculate the probabilities of each path, calculate the distribution of reproductive output at each stage on each path, and then integrate those probabilities to calculate the mean, variance, etc. of LRO.

The calculation of variance in LRO has been approached in several ways. Early studies used simulation to generate random trajectories through stages, including stages defined by reproductive output, to create a sample of lives from which variance in LRO could be calculated (e.g., Tuljapurkar et al. 2009; Steiner et al. 2010). An analytical solution for all the moments of LRO was provided by Caswell (2011) and provides the framework for our results here. That result took the form of an iterative calculation rather than a closed form expression. Steiner and Tuljapurkar (2012) presented a closed form analytical solution for all the moments of one component of LRO, using a moment generating function approach. They also reported simulations of the complete distribution of LRO and explored effects of heterogeneity. However, their solution included only part of the variance in LRO. They assumed the fertility of each age or stage to be a fixed deterministic quantity, neglecting the within-pathway component of variance. For example, the age-specific fertility of a 26-year-old Japanese woman in 1950 was 0.25. The analysis of Steiner and Tuljapurkar (2012) assumes that every 26-year-old woman, without exception, produces one fourth of a baby. In our framework, every 26-year-old woman, without exception, produces one baby with a probability of 0.25 and zero babies with a probability of 0.75 (Caswell 2011). This

<sup>2</sup>Tuljapurkar et al. (2009) independently and at the same time introduced the term “dynamic heterogeneity” to refer to the same random variation. We continue to use individual stochasticity because it more accurately describes the process creating the inter-individual variance and allows us to distinguish heterogeneity that is static from that which changes over the life of an individual.

component of variance is biologically realistic and quantitatively important. For example, analysis of 40 developed countries during the second demographic transition found that, as life expectancy increased, the fraction of variance in LRO due to within-trajectory randomness increased from  $\sim 50\%$  to  $\sim 99\%$  (Van Daalen and Caswell 2015). We show further examples and present the methodology for decomposing variance into these components, below.

All the studies published so far agree in finding that the variance in LRO due to individual stochasticity can comprise a high proportion of the observed variance in LRO, and that to ignore it is to miss a major source of variation (e.g., Tuljapurkar et al. 2009; Steiner et al. 2010; Steiner and Tuljapurkar 2012; Caswell 2011; van Daalen and Caswell 2015).

In this paper, we present exact, closed form, analytical formulae for all the moments of LRO, extending and replacing the iterative formulae of Caswell (2011). We include both the between-trajectory and within-trajectory components of variance and incorporate variance in stage-specific fertility either empirically or by a statistical model. We also present the sensitivity and elasticity analysis of means, variances, and all moments of LRO due to individual stochasticity. Our results can be applied to any age- or stage-classified matrix population model and to constant, periodic, or stochastic environments.

Our results rely on a mathematical model called a Markov chain with rewards, which we describe in “[Markov chains with rewards as a model for LRO](#).” The construction of a Markov chain with rewards requires demographic information on survival, stage transitions, and fertility. Section “[The statistics of LRO](#)” presents the calculation of all the moments and other descriptive statistics of LRO. Section “[Sensitivity analysis of LRO](#)” presents the sensitivity and elasticity analysis of LRO. Section “[Examples](#)” presents examples of age-classified and stage-classified populations, and “[Discussion](#)” concludes with a discussion of results and possible extensions. Proofs and derivations appear in Appendix A.

## Markov chains with rewards as a model for LRO

**Notation** Matrices are denoted by uppercase boldface letters (e.g.,  $\mathbf{P}$ ), and vectors by lowercase boldface letters (e.g.,  $\boldsymbol{\rho}$ ). Vectors are column vectors by default;  $\mathbf{X}^T$  is the transpose of  $\mathbf{X}$ . The vector  $\mathbf{1}_n$  is a  $n \times 1$  vector of ones,  $\mathbf{I}_n$  is the identity matrix of order  $n$ , and  $\mathbf{e}_i$  is the  $i$ th unit vector, with a 1 in the  $i$ th entry and zeros elsewhere. The matrix  $\mathbf{E}$  is a matrix of ones, and  $\mathbf{E}_{ij}$  is a matrix with a 1 in the  $(i, j)$  entry and zeros elsewhere. The diagonal matrix with the vector  $\mathbf{x}$  on the diagonal and zeros elsewhere is denoted  $\mathcal{D}(\mathbf{x})$ . The symbol  $\circ$  denotes the Hadamard, or

element-by-element product, and  $\otimes$  denotes the Kronecker product. The  $\text{vec}$  operator  $\text{vec } \mathbf{X}$  stacks the columns of an  $m \times n$  matrix  $\mathbf{X}$  into an  $mn \times 1$  column vector. The vec-permutation matrix  $\mathbf{K}_{m,n}$  satisfies  $\text{vec } \mathbf{X}^T = \mathbf{K}_{m,n} \text{vec } \mathbf{X}$ . In cases where it will help understanding, we indicate the dimension of displayed matrix expressions.

## The life cycle as a Markov chain

Our approach uses a mathematical model called a Markov chain with rewards. These models have a long history in stochastic process theory (e.g., Howard 1960; Puterman 1994; Sheskin 2010) but have only recently been applied to study lifetime reproduction (Caswell 2011; Van Daalen and Caswell 2015).

The individual life cycle is described by a finite-state, discrete-time, absorbing Markov chain; absorbing states represent death. The population projection matrix is written

$$\mathbf{A} = \mathbf{U} + \mathbf{F} \quad (2)$$

where  $\mathbf{U}$  contains the transition probabilities for extant individuals and  $\mathbf{F}$  contains stage-specific fertilities. Let  $\tau$  be the number of transient (living) states,  $\alpha$  the number of absorbing states, and  $s = \tau + \alpha$  the total number of states. Then the Markov chain transition matrix, including both transient and absorbing states, can be written

$$\mathbf{P} = \begin{pmatrix} \mathbf{U}_{\tau \times \tau} & \mathbf{0}_{\tau \times \alpha} \\ \mathbf{M}_{\alpha \times \tau} & \mathbf{I}_{\alpha \times \alpha} \end{pmatrix} \quad s \times s \quad (3)$$

where the dimensions of the submatrices are noted. The matrix  $\mathbf{P}$  is column stochastic. We assume that the spectral radius of  $\mathbf{U}$  is strictly less than 1; thus, any individual eventually dies with probability 1. The probability of dying while in each of the transient states is contained in the mortality matrix  $\mathbf{M}$ .

**Reproductive rewards** An individual experiences a sequence of transitions according to the probabilities in  $\mathbf{P}$ . Associated with each transition (including the transition of remaining in a stage) is a “reward” representing, in our case, reproductive output. The reproductive reward is a random variable with a specified set of moments. Rewards accumulate over the lifetime of the individual; the total accumulation at the time of death is the LRO. We make the reasonable assumption that individuals in the absorbing state stop accumulating rewards.<sup>3</sup>

<sup>3</sup>If the Markov chain is ergodic, or if it is absorbing but the absorbing states continue to collect rewards, then rewards will continue to accumulate forever. Lifetime accumulated rewards in such cases converge only if a discount factor is imposed, making delayed rewards less valuable than immediate rewards. We will present results for ergodic chains elsewhere.

The reproductive rewards associated with each transition are given by a set of matrices  $\mathbf{R}_k$ ,  $k = 1, 2, \dots$ . The  $(i, j)$  entry of  $\mathbf{R}_k$  is the  $k$ th moment of the reproductive output associated with the transition  $j \rightarrow i$ ; that is,

$$\mathbf{R}_k = \left( E \left[ r_{ij}^k \right] \right). \quad (4)$$

These matrices  $\mathbf{R}_m$  can be obtained in several ways, discussed in detail in Caswell (2011).

**Empirical distribution.** The moments can be calculated empirically from data on individual reproduction. Such data are frequently obtained, but typically only mean reproductive output is reported. In the absence of this information, the reward matrices can be modelled by any probability distribution that is determined by its mean, including the following.

**Bernoulli distribution.** When only a single offspring is produced, mean offspring production equals the probability of reproducing. The matrices of second and third moments satisfy

$$\mathbf{R}_3 = \mathbf{R}_2 = \mathbf{R}_1. \quad (5)$$

**Poisson distribution.** When multiple offspring are produced, the Poisson distribution describes a situation where every individual has the same chances of producing those offspring. The moments satisfy

$$\mathbf{R}_2 = \mathbf{R}_1 + (\mathbf{R}_1 \circ \mathbf{R}_1) \quad (6)$$

$$\mathbf{R}_3 = \mathbf{R}_1 + 3(\mathbf{R}_1 \circ \mathbf{R}_1) + (\mathbf{R}_1 \circ \mathbf{R}_1 \circ \mathbf{R}_1). \quad (7)$$

**Fixed rewards.** It is possible to eliminate variance in fertility by creating reward matrices where every individual in a given stage produces exactly the mean number of offspring, in which case

$$\mathbf{R}_2 = \mathbf{R}_1 \circ \mathbf{R}_1 \quad (8)$$

$$\mathbf{R}_3 = \mathbf{R}_1 \circ \mathbf{R}_1 \circ \mathbf{R}_1. \quad (9)$$

In most demographic models, mean fertility is stage-specific rather than transition-specific. That is, mean fertility is described by a vector  $\mathbf{f}$  that typically appears in the first row of the matrix  $\mathbf{A}$ . An individual in stage  $i$  produces, on average,  $f_i$  offspring per time step, regardless of what transition the individual may make. In this case

$$\mathbf{R}_1 = \left( \begin{array}{ccc|c} f_1 & \cdots & f_\tau & 0_{1 \times \alpha} \\ \vdots & & \vdots & \vdots \\ f_1 & \cdots & f_\tau & 0_{1 \times \alpha} \\ \hline f_1 & \cdots & f_\tau & 0_{1 \times \alpha} \end{array} \right) \quad (10)$$

$$= \mathbf{1}_s (\mathbf{f}^\top | \mathbf{0}_{1 \times \alpha}). \quad (11)$$

If parental survival is required for reproduction, then the last row[s] of  $\mathbf{R}_1$  should be set to zero.

## The statistics of LRO

In this model, every individual is subject to the same rates at any given stage, so there is no heterogeneity. Even so, each individual may experience a different life course. The resulting variation among individuals is due to the individual stochasticity implied by the vital rates in  $\mathbf{P}$  and the fertility process described by the  $\mathbf{R}_i$ . Our task is to derive the statistics of LRO from this information.

Let  $\rho_k$  be a vector (dimension  $s \times 1$ ) whose  $i$ th entry is the  $k$ th moment of the remaining lifetime reproductive output of an individual starting in state  $i$ , i.e.,

$$\rho_k = (E[\rho_i^k]). \quad (12)$$

Because we assume that the dead do not reproduce, we know that the entries of  $\rho_k$  corresponding to absorbing states are zero. Thus, we obtain the complete statistics of LRO if we can find the moment vectors  $\tilde{\rho}_k$ , where  $\tilde{\rho}$  contains the first  $\tau$  entries of  $\rho$ , corresponding to the living stages. This vector is given by

$$\tilde{\rho} = \mathbf{Z}\rho \quad \text{and} \quad \rho = \mathbf{Z}^\top \tilde{\rho} \quad (13)$$

where

$$\mathbf{Z} = (\mathbf{I}_{\tau \times \tau} | \mathbf{0}_{\tau \times \alpha}); \quad (14)$$

in terms of this matrix,  $\rho = \mathbf{Z}^\top \tilde{\rho}$ . We also define  $\tilde{\mathbf{R}}_k$ , the  $\tau \times \tau$  submatrix of  $\mathbf{R}_k$  corresponding to transitions among the transient states:

$$\tilde{\mathbf{R}}_k = \mathbf{Z}\mathbf{R}_k\mathbf{Z}^\top. \quad (15)$$

In terms of these quantities, the moments of LRO are given by the following theorem.

**Theorem 1** *The moment vectors  $\tilde{\rho}$  of lifetime accumulated reproductive output are*

$$\tilde{\rho}_1 = \mathbf{N}^\top \mathbf{Z} (\mathbf{P} \circ \mathbf{R}_1)^\top \mathbf{1}_s \quad (16)$$

$$\tilde{\rho}_2 = \mathbf{N}^\top \left[ \mathbf{Z} (\mathbf{P} \circ \mathbf{R}_2)^\top \mathbf{1}_s + 2 (\mathbf{U} \circ \tilde{\mathbf{R}}_1)^\top \tilde{\rho}_1 \right] \quad (17)$$

$$\tilde{\rho}_3 = \mathbf{N}^\top \left[ \mathbf{Z} (\mathbf{P} \circ \mathbf{R}_3)^\top \mathbf{1}_s + 3 (\mathbf{U} \circ \tilde{\mathbf{R}}_2)^\top \tilde{\rho}_1 + 3 (\mathbf{U} \circ \tilde{\mathbf{R}}_1)^\top \tilde{\rho}_2 \right] \quad (18)$$

and, in general,

$$\tilde{\rho}_m = \mathbf{N}^\top \mathbf{Z} (\mathbf{P} \circ \mathbf{R}_m)^\top \mathbf{1}_s + \sum_{k=1}^{m-1} \binom{m}{k} \mathbf{N}^\top (\mathbf{U} \circ \tilde{\mathbf{R}}_{m-k})^\top \tilde{\rho}_k \quad (19)$$

where  $\mathbf{N} = (\mathbf{I}_\tau - \mathbf{U})^{-1}$  is the fundamental matrix of the Markov chain.

*Proof* See Appendix A.1 □

In terms of the moment vectors provided by Theorem 1, the mean, variance, standard deviation, and coefficient of variation of LRO, and Crow's index  $\mathcal{I}$  of opportunity for selection, are given by

$$E(\tilde{\rho}) = \tilde{\rho}_1 \quad (20)$$

$$V(\tilde{\rho}) = \tilde{\rho}_2 - (\tilde{\rho}_1 \circ \tilde{\rho}_1) \quad (21)$$

$$SD(\tilde{\rho}) = \sqrt{V(\tilde{\rho})} \quad (22)$$

$$CV(\tilde{\rho}) = \mathcal{D}(\rho_1)^{-1} SD(\tilde{\rho}) \quad (23)$$

$$\mathcal{I} = CV(\tilde{\rho}) \circ CV(\tilde{\rho}). \quad (24)$$

We focus in this paper on statistics obtained from the first and second moments; these statistics, such as the variance, describe variability of the distribution. Calculations of the skewness, as a measure of the shape of the distribution, are given in Caswell (2011) and Van Daalen and Caswell (2015). All of the methods we present here can be applied to these statistics.

### Partitioning variance: within and between pathways

Variance in LRO is partly due to individuals following different pathways through the life cycle and partly due to variance in stage-specific fertility along those pathways. The overall variance  $V(\tilde{\rho})$  can be decomposed into these components using the law of conditional variance

$$V(\tilde{\rho}) = V_{\text{within}} + V_{\text{between}} \quad (25)$$

(Rényi 1970, Theorem 1, p. 275). To calculate the variance between trajectories, we eliminate the within-trajectory variance by calculating  $V(\tilde{\rho})$  using the fixed reward model, in which the second moment matrix  $\mathbf{R}_2$  is given by Eq. 8. The within-trajectory variance is obtained by subtraction as  $V_{\text{within}} = V(\tilde{\rho}) - V_{\text{between}}$ .

### Sensitivity analysis of LRO

The statistics of LRO depend on the life history parameters that determine the transition matrices  $\mathbf{U}$  and  $\mathbf{M}$ , and the moment matrices  $\mathbf{R}_1, \mathbf{R}_2, \dots$ . Sensitivity analysis reveals how these parameters affect LRO; our goal is to derive the sensitivity and elasticity of LRO to changes in any parameters. Although sensitivity analysis of the net reproductive rate  $R_0$  has been presented before (Matser et al. 2009; Caswell 2009), there has been no such analysis for the variance or other measures of variation in LRO.

We derive sensitivity formulae using matrix calculus formalism (Caswell 2007, 2008, 2009, 2012; Caswell and Salguero-Gómez 2013). Let  $\theta$  be a vector ( $p \times 1$ ) of parameters of interest; these could be mortalities, transition probabilities, means or variances of fertility, etc. The sensi-

tivity of the moment vector  $\tilde{\rho}_m$  to the parameter vector  $\theta$  is the derivative matrix

$$\frac{d\tilde{\rho}_m}{d\theta^T} = \left( \frac{d\tilde{\rho}_m(i)}{d\theta(j)} \right) \quad \tau \times p. \quad (26)$$

That is, the  $(i, j)$  entry of this sensitivity matrix is the derivative of the  $i$ th entry of  $\tilde{\rho}_m$  with respect to the  $j$ th entry of  $\theta$ .

To write the sensitivity of the  $\tilde{\rho}_m$ , let us define the following matrices.

$$\mathbf{V}_i = (\mathbf{1}_s^T \otimes \mathbf{Z}) \mathbf{K}_{ss} \left[ \mathcal{D}(\text{vec } \mathbf{R}_i) \frac{d\text{vec } \mathbf{P}}{d\theta^T} + \mathcal{D}(\text{vec } \mathbf{P}) \frac{d\text{vec } \mathbf{R}_i}{d\theta} \right] \quad (27)$$

$$\mathbf{W}_{i,j} = (\tilde{\rho}_i^T \otimes \mathbf{I}_\tau) \mathbf{K}_{\tau\tau} \left[ \mathcal{D}(\text{vec } \tilde{\mathbf{R}}_j) \frac{d\text{vec } \mathbf{U}}{d\theta^T} + \mathcal{D}(\text{vec } \mathbf{U}) \frac{d\text{vec } \tilde{\mathbf{R}}_j}{d\theta^T} \right] + (\mathbf{U} \circ \tilde{\mathbf{R}}_j)^T \frac{d\tilde{\rho}_i}{d\theta^T} \quad (28)$$

$$\mathbf{X}_i = (\tilde{\rho}_i^T \otimes \mathbf{I}_\tau) \mathbf{K}_{\tau\tau} \frac{d\text{vec } \mathbf{U}}{d\theta^T}. \quad (29)$$

In Eqs. 27–29,  $\mathbf{K}_{mn}$  is the vec-permutation matrix of order  $(m, n)$ ; see Henderson and Searle (1981) and Magnus and Neudecker (1979).

**Theorem 2** Let  $\mathbf{P}$ ,  $\mathbf{U}$ , and  $\mathbf{M}$  define the absorbing Markov chain in Eq. 3, with  $\tau$  transient states,  $\alpha$  absorbing states, and  $s = \tau + \alpha$  total states. Let  $\mathbf{R}_i$  contain the  $i$ th moments of the reproductive rewards corresponding to each transition. Let  $\theta$  be a vector of parameters. The vector  $\tilde{\rho}_m$  contains the  $m$ th moments of remaining LRO for each of the  $\tau$  transient stages. The sensitivity of  $\tilde{\rho}_m$  to  $\theta$  is

$$\frac{d\tilde{\rho}_m}{d\theta^T} = \mathbf{N}^T \left[ \mathbf{V}_m + \sum_{k=1}^{m-1} \binom{m}{k} \mathbf{W}_{k,m-k} + \mathbf{X}_m \right] \quad m = 1, 2, \dots \quad (30)$$

*Proof* Proof is given in Appendix A.2.  $\square$

The sensitivities of the first three moments of LRO are of particular interest; they are

$$\frac{d\tilde{\rho}_1}{d\theta^T} = \mathbf{N}^T (\mathbf{V}_1 + \mathbf{X}_1) \quad (31)$$

$$\frac{d\tilde{\rho}_2}{d\theta^T} = \mathbf{N}^T (\mathbf{V}_2 + 2\mathbf{W}_{1,1} + \mathbf{X}_2) \quad (32)$$

$$\frac{d\tilde{\rho}_3}{d\theta^T} = \mathbf{N}^T (\mathbf{V}_3 + 3\mathbf{W}_{1,2} + 3\mathbf{W}_{2,1} + \mathbf{X}_3). \quad (33)$$

If, as is often the case, the model includes only a single absorbing state (death), then  $\mathbf{M}$  is a  $(1 \times \tau)$  matrix, given by

$$\mathbf{M} = \mathbf{1}_\tau^T - \mathbf{1}_\tau^T \mathbf{U} \quad 1 \times \tau. \quad (34)$$



In this case, the derivative of  $\mathbf{P}$  that appears in Eq. 27 is

$$\frac{d\text{vec } \mathbf{P}}{d\boldsymbol{\theta}^\top} = [\mathbf{C}_1 - \mathbf{C}_2 (\mathbf{I}_\tau \otimes \mathbf{1}_\tau^\top)] \frac{d\text{vec } \mathbf{U}}{d\boldsymbol{\theta}^\top} \quad (35)$$

where

$$\mathbf{C}_1 = \begin{pmatrix} \mathbf{I}_{\tau \times \tau} \\ \mathbf{0}_{1 \times \tau} \end{pmatrix} \otimes \begin{pmatrix} \mathbf{I}_{\tau \times \tau} \\ \mathbf{0}_{1 \times \tau} \end{pmatrix} \quad (36)$$

$$\mathbf{C}_2 = \begin{pmatrix} \mathbf{I}_{\tau \times \tau} \\ \mathbf{0}_{1 \times \tau} \end{pmatrix} \otimes \begin{pmatrix} \mathbf{0}_{\tau \times 1} \\ \mathbf{I}_{1 \times 1} \end{pmatrix} \quad (37)$$

Caswell and Van Daalen (2016). This permits expressing all the effects of  $\boldsymbol{\theta}$  on the transition matrix  $\mathbf{P}$  in terms of effects on  $\mathbf{U}$  (see Appendix A.2).

### Sensitivity of the statistics of LRO

The moments  $\tilde{\rho}_i$  provide the statistics (21)–(24) describing the inter-individual variability of LRO, including the variance, standard deviation, coefficient of variation, and the scaled variance (Crow's index) (Caswell 2011; Van Daalen and Caswell 2015).

The sensitivities of these quantities are

$$\frac{dV(\tilde{\rho})}{d\boldsymbol{\theta}^\top} = \frac{d\tilde{\rho}_2}{d\boldsymbol{\theta}^\top} - 2\mathcal{D}(\tilde{\rho}_1) \frac{d\tilde{\rho}_1}{d\boldsymbol{\theta}^\top} \quad (38)$$

$$\frac{dSD(\tilde{\rho})}{d\boldsymbol{\theta}^\top} = \frac{1}{2}\mathcal{D}[SD(\tilde{\rho})]^{-1} \frac{dV(\tilde{\rho})}{d\boldsymbol{\theta}^\top} \quad (39)$$

$$\begin{aligned} \frac{dCV(\tilde{\rho})}{d\boldsymbol{\theta}^\top} &= \mathcal{D}(\tilde{\rho}_1)^{-1} \frac{dSD(\tilde{\rho})}{d\boldsymbol{\theta}^\top} \\ &\quad - (SD(\tilde{\rho}) \otimes \mathbf{I}_\tau) \mathcal{D}(\tilde{\rho}_1)^{-2} \frac{d\tilde{\rho}_1}{d\boldsymbol{\theta}^\top} \end{aligned} \quad (40)$$

$$\frac{d\mathcal{I}}{d\boldsymbol{\theta}^\top} = 2\mathcal{D}[CV(\tilde{\rho})] \frac{dCV(\tilde{\rho})}{d\boldsymbol{\theta}^\top}. \quad (41)$$

For derivations, see Appendix A.3.

### Elasticity

The derivatives in Theorem 2 and Eqs. 38–41 measure the effect of an additive perturbation of the parameter vector  $\boldsymbol{\theta}$ . Elasticities, which measure the proportional change resulting from a proportional change in  $\boldsymbol{\theta}$ , are easily calculated. Let  $\xi$  be any quantity calculated from the demographic model. The elasticity of  $\xi$  with respect to  $\boldsymbol{\theta}$  is the matrix

$$\frac{\varepsilon \xi}{\varepsilon \boldsymbol{\theta}^\top} = \mathcal{D}(\xi)^{-1} \frac{d\xi}{d\boldsymbol{\theta}^\top} \mathcal{D}(\boldsymbol{\theta}). \quad (42)$$

As usual, elasticity calculations can be applied only if  $\xi > 0$  and  $\boldsymbol{\theta} \geq 0$ .

### Some special perturbations

Here we consider some perturbations that are of special interest: the survival and transitions in stage-classified models and sensitivity to means and variances of stage-specific fertility.

#### Mortality and transitions in stage-classified models

In an age-classified model, the transient matrix  $\mathbf{U}$  is completely determined by the mortality schedule (see “Age-classified human populations” for age-classified examples). In a stage-classified model, however,  $\mathbf{U}$  depends on both survival and probabilities of transitions among stages. To calculate sensitivities of LRO to survival and transition probabilities, we write

$$\mathbf{U} = \mathbf{G}\boldsymbol{\Sigma} \quad (43)$$

where  $\boldsymbol{\Sigma}$  has the survival probabilities  $\boldsymbol{\sigma}$  on the diagonal; i.e.,  $\boldsymbol{\Sigma} = \mathcal{D}(\boldsymbol{\sigma})$ , and  $\mathbf{G}$  is a matrix of transition probabilities conditional on survival. Differentiating (43) gives

$$d\text{vec } \mathbf{U} = (\mathbf{I}_\tau \otimes \mathbf{G}) d\text{vec } \boldsymbol{\Sigma} + (\boldsymbol{\Sigma} \otimes \mathbf{I}_\tau) d\text{vec } \mathbf{G}. \quad (44)$$

Writing  $\mathcal{D}(d\boldsymbol{\sigma}) = \mathbf{I} \circ (\mathbf{1}_\tau d\boldsymbol{\sigma}^\top)$  and simplifying lead to

$$\frac{d\text{vec } \mathbf{U}}{d\boldsymbol{\sigma}^\top} = (\mathbf{I}_\tau \otimes \mathbf{G}) \mathcal{D}(\text{vec } \mathbf{I}) (\mathbf{I}_\tau \otimes \mathbf{1}_\tau), \quad (45)$$

and since  $\boldsymbol{\sigma} = \exp(-\boldsymbol{\mu})$ ,

$$\frac{d\text{vec } \mathbf{U}}{d\boldsymbol{\mu}^\top} = -\frac{d\text{vec } \mathbf{U}}{d\boldsymbol{\sigma}^\top} \mathcal{D}(\boldsymbol{\sigma}). \quad (46)$$

From (44), the sensitivity of  $\mathbf{U}$  to the growth matrix  $\mathbf{G}$  is

$$\frac{d\text{vec } \mathbf{U}}{d\text{vec }^\top \mathbf{G}} = (\boldsymbol{\Sigma}^\top \otimes \mathbf{I}). \quad (47)$$

However,  $\mathbf{G}$  is column-stochastic; i.e., its columns must sum to 1. An arbitrary perturbation of  $\mathbf{G}$  would result in the loss of the column-stochastic property. The only relevant perturbations are those that maintain the column sums. Thus, perturbations of the entry  $g_{ij}$  must be compensated for by changes in other elements of column  $j$ . Caswell (2001, 2013) presents a method for compensating for perturbation of each entry in a column-stochastic matrix in a way that maintains the proportional structure of the column. In matrix notation, we write

$$\left. \frac{d\text{vec } \mathbf{U}}{d\text{vec }^\top \mathbf{G}} \right|_{\text{comp}} = \frac{d\text{vec } \mathbf{U}}{d\text{vec }^\top \mathbf{G}} \frac{d\text{vec } \mathbf{G}}{d\text{vec }^\top \boldsymbol{\Theta}} \quad (48)$$

where the matrix  $\boldsymbol{\Theta}$  is a matrix of perturbations that include the compensation for the column sums. The matrix of derivatives of  $\mathbf{G}$  with respect to  $\boldsymbol{\Theta}$  is given by Caswell (2013). Let  $\mathbf{g}_i$  denote column  $i$  of  $\mathbf{G}$ ; then

$$\frac{d\text{vec } \mathbf{G}}{d\text{vec }^\top \boldsymbol{\Theta}} = \mathbf{I}_{\tau^2} - \sum_{i=1}^{\tau} \mathbf{E}_{ii} \otimes \left[ \mathcal{D}(\mathbf{g}_i) \mathbf{C} \mathcal{D}(\mathbf{1} - \mathbf{g}_i)^{-1} \right] \quad (49)$$

where  $\mathbf{C} = \mathbf{E} - \mathbf{I}$ .

### Sensitivity to means and variances of fertility

Perturbations of fertility appear in Eqs. 27–29 as derivatives of the reproductive reward matrices  $\mathbf{R}_i$ . When the distributions of stage-specific fertilities are specified by a parametric distribution, the moments may be linked, so that changes in mean fertility also affect the variance (e.g., in the Poisson distribution, the variance is equal to the mean). Sometimes, however, it is of interest to treat the mean and variance of fertility as independent traits and calculate the sensitivity of LRO to the mean, holding the variance fixed, and to the variance, holding the mean fixed. This subtle but important distinction was emphasized by Tuljapurkar et al. (2003) and Haridas and Tuljapurkar (2005) in the context of the elasticity of the stochastic growth rate to the entries of a stochastically varying matrix. One might compute the effect of changing one of the matrix entries (perhaps by a change in energy allocation strategy), recognizing that this would change both the mean and the variance. Or, one might be interested in the effects of variance per se and manipulate the moments to calculate elasticities with respect to the mean and variance independently.

As in Eq. 11, suppose that fertility is defined by a fertility vector  $\mathbf{f}$  with first and second moments  $\mathbf{f}_1$  and  $\mathbf{f}_2$ , and a variance vector

$$\mathbf{v} = \mathbf{f}_2 - (\mathbf{f}_1 \circ \mathbf{f}_1). \quad (50)$$

The first and second moment matrices are given by

$$\mathbf{R}_i = \mathbf{1}_s \mathbf{f}_i^T \mathbf{Z} \quad i = 1, 2. \quad (51)$$

**Sensitivity to mean fertility, variance fixed** From Eq. 50, it follows that

$$d\mathbf{v} = d\mathbf{f}_2 - 2\mathcal{D}(\mathbf{f}_1)d\mathbf{f}_1. \quad (52)$$

To hold the variance fixed, we require  $d\mathbf{v} = 0$ , which implies that

$$\frac{d\mathbf{f}_2}{d\mathbf{f}_1^T} = 2\mathcal{D}(\mathbf{f}_1). \quad (53)$$

To evaluate sensitivity to the mean, we set the parameter vector  $\boldsymbol{\theta} = \mathbf{f}_1$ , subject to Eq. 53, and obtain

$$\left. \frac{d\text{vec } \mathbf{R}_1}{d\boldsymbol{\theta}^T} \right|_{d\mathbf{v}=0} = (\mathbf{Z}^T \otimes \mathbf{1}_s) \quad (54)$$

$$\left. \frac{d\text{vec } \mathbf{R}_2}{d\boldsymbol{\theta}^T} \right|_{d\mathbf{v}=0} = (\mathbf{Z}^T \otimes \mathbf{1}_s) \frac{d\mathbf{f}_2}{d\mathbf{f}_1} \quad (55)$$

$$= 2(\mathbf{Z}^T \otimes \mathbf{1}_s) \mathcal{D}(\mathbf{f}_1). \quad (56)$$

Substituting these expressions into Eqs. 31 and 32, and then into Eq. 38, we obtain the sensitivity of the mean and variance of LRO to changes in mean fertility, with variance in fertility held constant.

**Sensitivity to variance in fertility, mean fixed** To hold the mean fixed, we require that  $d\mathbf{f}_1 = 0$ , in which case (52) implies that

$$\frac{d\mathbf{f}_2}{d\mathbf{v}^T} = \mathbf{I}_\tau. \quad (57)$$

Now we set the parameter vector  $\boldsymbol{\theta} = \mathbf{v}$  to obtain

$$\left. \frac{d\text{vec } \mathbf{R}_1}{d\boldsymbol{\theta}^T} \right|_{d\mathbf{f}_1=0} = \mathbf{0}_{s \times \tau} \quad (58)$$

$$\left. \frac{d\text{vec } \mathbf{R}_2}{d\boldsymbol{\theta}^T} \right|_{d\mathbf{f}_1=0} = (\mathbf{Z}^T \otimes \mathbf{1}_s) \frac{d\mathbf{f}_2}{d\mathbf{v}^T} \quad (59)$$

$$= (\mathbf{Z}^T \otimes \mathbf{1}_s). \quad (60)$$

Substituting these expressions into Eqs. 31 and 32, and then into Eq. 38, gives the sensitivity of the mean and variance of LRO to changes in the variance in fertility, with mean fertility held constant.

### A protocol for the analysis of lifetime reproductive output

The results presented to this point provide a protocol for analysis of lifetime reproductive output, applicable to any matrix population model. A stepwise version of this protocol is given in Table 1. In the next section, we present age-classified and stage-classified examples of the analysis.

### Examples

This section presents examples of the calculation of the statistical properties of lifetime reproductive output and its subsequent sensitivity analysis, for both age-classified and stage-classified population. In “[Age-classified human populations](#),” we analyze a set of age-classified human populations that span a wide range of demographic conditions. In “[A stage-classified tree population](#),” we analyze a size-classified model for Canadian hemlock (*Tsuga canadensis*), a coniferous tree.

#### Age-classified human populations

The transition matrix  $\mathbf{U}$  for an age-classified model contains survival probabilities on the subdiagonal and zeros

**Table 1** A step-by-step protocol for analysis of lifetime reproductive output and its sensitivity, from any stage- or age-classified matrix population model

1. Obtain a transition matrix  $\mathbf{U}$ , perhaps from decomposing a population projection matrix as  $\mathbf{A} = \mathbf{U} + \mathbf{F}$ .
2. Locate reproductive transitions.
  - (a) If fertility is transition specific, identify the transitions (e.g., to reproductive states).
  - (b) If fertility is stage-specific, extract the vector  $\mathbf{f}$  from  $\mathbf{F}$ .
3. Obtain statistical moments of fertility:
  - (a) From empirical measurements of the moments of stage-specific fertility, or
  - (b) From an assumption of Bernoulli [see equation (5)], or Poisson [see Eqs. 6 and 7], or fixed [see Eqs. 8 and 9] reproduction.
4. Construct reward matrices from Eq. 11.
5. Compute desired moments of LRO from Eqs. 16–19.
6. Compute desired statistics of LRO from Eqs. 21–24.
7. Sensitivity analysis
  - (a) Specify parameter vector  $\boldsymbol{\theta}$  of interest
  - (b) Calculate derivatives of  $\mathbf{U}$ , and  $\mathbf{R}_i$  to  $\boldsymbol{\theta}$ . Take advantage of Eqs. 35 or 90–92 to compute derivatives of  $\mathbf{P}$  to  $\boldsymbol{\theta}$ .
  - (c) If the matrix is stage-classified,
    - i. Decompose  $\mathbf{U} = \mathbf{G}\Sigma$ .
    - ii. Use Eq. 46 to compute the derivative of  $\mathbf{U}$  to mortality.
    - iii. Use Eq. 49 to compute the derivative of  $\mathbf{G}$  to  $\boldsymbol{\theta}$ , including compensation to preserve column sums of  $\mathbf{G}$
  - (d) Compute derivatives of the moment vectors  $\tilde{\rho}_i$  for the moments of interest ( $i = 1, 2$  suffice to analyze the variance, standard deviation, CV, and  $\mathcal{I}$ ).
    - i. Compute  $\mathbf{V}$ ,  $\mathbf{W}$ , and  $\mathbf{X}$  using Eqs. 27–29.
    - ii. Compute derivatives of  $\tilde{\rho}_i$  using Theorem 2.
  - (e) Compute sensitivity of desired statistics of LRO using Eqs. 38–41.
  - (f) If desired, compute elasticities of statistics of LRO using Eq. 42.

elsewhere. One absorbing state, death, is included, and the mortality matrix  $\mathbf{M}$  is calculated according to Eq. 34.

We ignore multiple births and treat the entries of the fertility vector  $\mathbf{f}$  as the probability of producing a single offspring. The offspring production is given by the reward  $r_{ij}$ , following a Bernoulli distribution,

$$r_{ij} = \begin{cases} 1 & \text{with probability } f_j \\ 0 & \text{with probability } (1 - f_j) \end{cases} \quad (61)$$

The matrix  $\mathbf{R}_1$  containing the first moment of the reward matrix is built from the fertility vector using Eq. 11. The higher moments of the reward matrix follow from the Bernoulli model of reproduction, in which the higher moments are all equal to the first, as in Eq. 5, so that  $\mathbf{R}_1 = \mathbf{R}_2 = \mathbf{R}_3$ .

We present results for nine populations: the Netherlands (1950 and 2009), Sweden (1891 and 2010), Japan (1947 and 2009), two hunter-gather populations (the Ache of subtropical Paraguay (Gurven and Kaplan 2007; Hill and Hurtado 1996) and the Hadza of the Tanzanian savanna (Blurton Jones 2011)), and the Hutterites of North America. The Netherlands, Sweden, and Japan are typical of developed countries progressing through the demographic transition. The hunter-gatherer populations have higher mortality,

lower life expectancy, and higher fertility than the developed countries. The Hutterites, an Anabaptist religious community in the United States and Canada, are known for having the highest total fertility for any known human population (Eaton and Mayer 1953), but are assumed to have experienced mortality rates similar to that of the USA around 1946–1950.

Data for the Netherlands, Sweden, and Japan were obtained from the Human Mortality Database (2013) and the Human Fertility Database (2013). Data for the Ache were obtained from Gurven and Kaplan (2007) and Hill and Hurtado (1996), and for the Hadza from Blurton Jones (2011). The fertility and mortality schedules for the Hutterites were taken from Eaton and Mayer (1953).

**Variance in LRO** Using Theorem 1 and Eqs. 21–24, and following the protocol in Table 1, we computed the mean, variance (both within and between pathways), standard deviation, coefficient of variation, and Crow's  $\mathcal{I}$  from  $\mathbf{U}$ ,  $\mathbf{R}_1$ , and  $\mathbf{R}_2$ . We also calculated life expectancy as the column sums of the fundamental matrix. The results are given in Table 2.

Not surprisingly, recent populations in developed countries have higher longevity and lower mean lifetime repro-



**Table 2** The statistics of lifetime reproductive output for the Netherlands (NLD), Sweden (SWE), and Japan (JPN), with two points in time for each country, two hunter-gatherer populations, the Hadza and the Ache, and a population of high-fertility Hutterites

Population	Mean	$V$	$V_{\text{between}} (\%)$	$V_{\text{within}} (\%)$	SD	CV	$\mathcal{I}$	Life exp.
NLD 1950	2.96	2.91	12.5	87.5	1.71	0.58	0.33	73.1
NLD 2009	1.78	1.61	1.4	98.6	1.27	0.72	0.51	83.1
SWE 1891	3.00	5.60	55.6	44.4	2.37	0.79	0.62	53.0
SWE 2010	1.97	1.79	1.4	98.6	1.34	0.68	0.46	84.0
JPN 1947	3.50	6.10	54.8	45.2	2.47	0.71	0.50	54.2
JPN 2009	1.35	1.26	0.9	99.1	1.12	0.83	0.69	86.9
Hadza	3.13	11.30	78.9	21.1	3.36	1.07	1.15	34.6
Ache	4.48	17.23	81.1	18.9	4.15	0.93	0.86	38.0
Hutterites	7.53	8.58	41.3	58.7	2.93	0.39	0.15	70.0

ductive output. The hunter-gatherer and Hutterite populations show the highest mean LRO. In each of the developed countries (the Netherlands, Sweden, and Japan), the reductions in mean LRO and increases in life expectancy are accompanied by reductions in the variance in LRO. The Ache and Hadza show the highest variance in LRO, and the variance for the Hutterites is higher than any of the developed countries.

Most of the variance in the hunter-gatherer populations is due to variance among pathways, which in an age-classified model is determined by survival from birth through the reproductive ages. The most recent years in developed countries show extremely low between-pathway variance because almost all women survive through their reproductive years (Van Daalen and Caswell 2015). The Ache and Hadza, with the lowest life expectancy, show very high  $V_{\text{between}}$ . The Hutterites have a long life expectancy, but their high fertility amplifies the effect of differences in longevity, so  $V_{\text{between}}$  is similar to that of Japan in 1947 and Sweden in 1891.

The opportunity for selection  $\mathcal{I}$  varies less (about sevenfold) among these populations than does the variance in LRO (about 13-fold). The Hutterites showed the lowest opportunity for selection of all the populations we included, yet the other high-fertility populations, the Ache and the Hadza, show the highest values for the opportunity for selection. Taking into account the fact that the hunter-gatherer populations have the highest variances in LRO, mostly due to variation in the pathways individuals take through life, we posit that the high opportunity for selection reflects room for improvement in survival rates from birth to reproductive ages.

The variance in LRO documented in Table 2 is calculated on the assumption that every individual experiences the same vital rates at every age and is thus due to individual stochasticity. Crow's  $\mathcal{I}$  is a measure of the potential relative increase in fitness per generation, but the variance here is stochastic, not genetic, so the opportunity for selection is only apparent, not real.

**Sensitivity analysis** The age-classified models for human populations are parameterized by the mortality rate vector  $\boldsymbol{\mu}$  and the mean fertility vector  $\mathbf{f}$ . The sensitivity of the statistics of LRO to these parameters, obtained from Theorem 2 and Eqs. 38–41, requires  $d\text{vec } \mathbf{U}/d\boldsymbol{\mu}^T$  and  $d\text{vec } \mathbf{R}_i/d\mathbf{f}^T$  for  $i = 1, 2$ . The derivative of  $\mathbf{U}$  is

$$\frac{d\text{vec } \mathbf{U}}{d\boldsymbol{\mu}^T} = -\mathcal{D}(\text{vec } \mathbf{Y})(\mathbf{I} \otimes \mathbf{1})\mathcal{D}(\mathbf{P}) \quad (62)$$

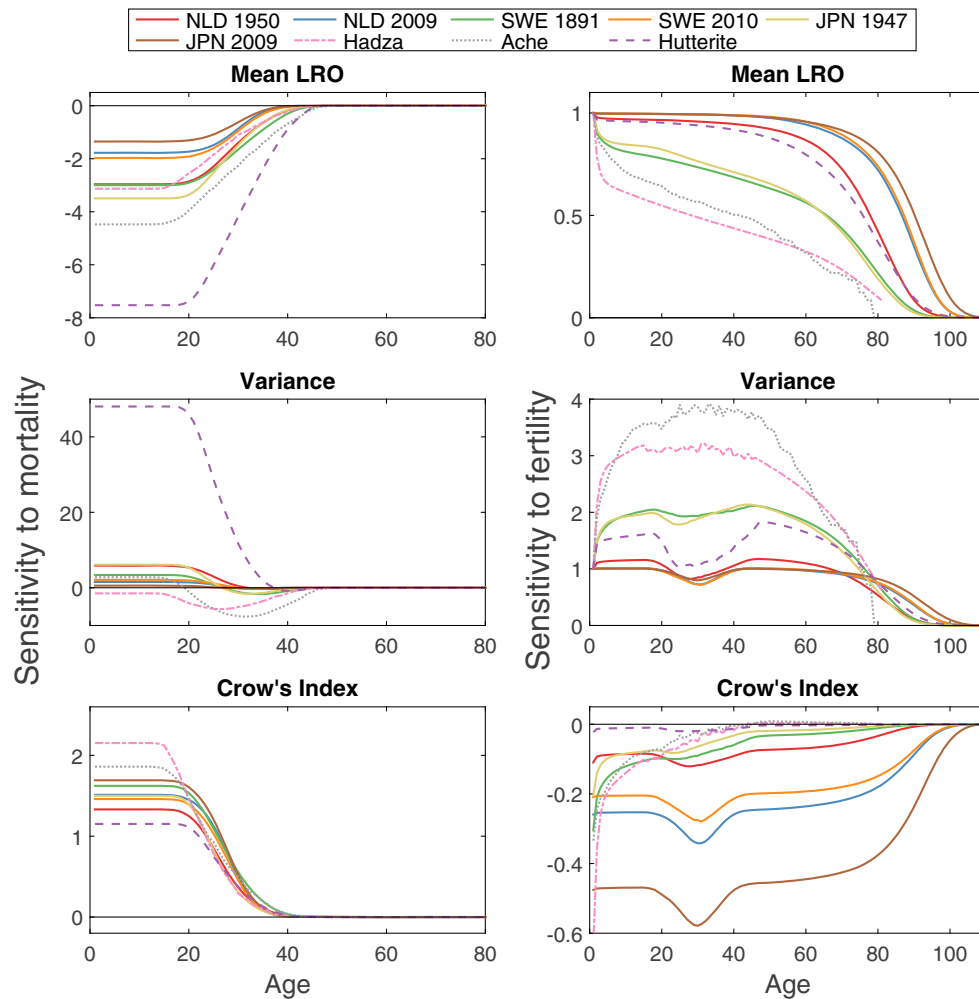
where  $\mathbf{Y}$  is an indicator matrix defining the subdiagonal structure of  $\mathbf{U}$ , and  $\mathbf{P} = \exp(-\boldsymbol{\mu})$  is the vector of survival probabilities. The Bernoulli distribution assumption implies that the derivatives of the first and second (and all other) moments are equal, with

$$\frac{d\mathbf{R}_1}{d\mathbf{f}^T} = (\mathbf{Z}^T \otimes \mathbf{1}_s) \quad (63)$$

(for derivations, see Appendix A.4).

The sensitivity of mean LRO, variance in LRO, and Crow's  $\mathcal{I}$  to mortality and fertility are shown in Fig. 1. Not surprisingly, increased mortality at any age up to the end of the reproductive period reduces mean LRO, while increased fertility increases mean LRO. The mortality effect is greatest for the Hutterites, because their high fertility magnifies the impact of mortality changes, and least for the recent developed countries. The sensitivity of mean LRO to fertility is given by the survivorship function and thus is smallest for the Ache and Hadza. It is highest for recent developed countries in which low mortality means that almost everyone would survive to benefit from an increase in fertility.

The variance in LRO increases with an increase in mortality rate for most populations; the effect is greatest for the Hutterites (Fig. 1). For the developed countries, the sensitivity is positive over most of the reproductive ages. For the



**Fig. 1** Sensitivity of mean LRO, variance in LRO, and Crow's index to changes in age-specific mortality (*left column*) and age-specific fertility (*right column*) for nine human populations

Hadza and the Ache, variance in LRO decreases with higher age-specific mortality rates. Variance in LRO increases with increasing fertility rates for all countries. The Hutterites, Sweden, the Netherlands, and Japan show a reduced sensitivity of variance in LRO to fertility around the reproductive ages.

Crow's opportunity for selection  $\mathcal{I}$  combines both the mean and the variance. Increased mortality during the reproductive period increases  $\mathcal{I}$  in all the populations. It is most sensitive to mortality in the Ache and Hadza populations and least sensitive in the Hutterites. An increase in fertility reduces  $\mathcal{I}$  in all the populations. Thus, the net result of environmental changes that affect both mortality and fertility cannot be predicted a priori.

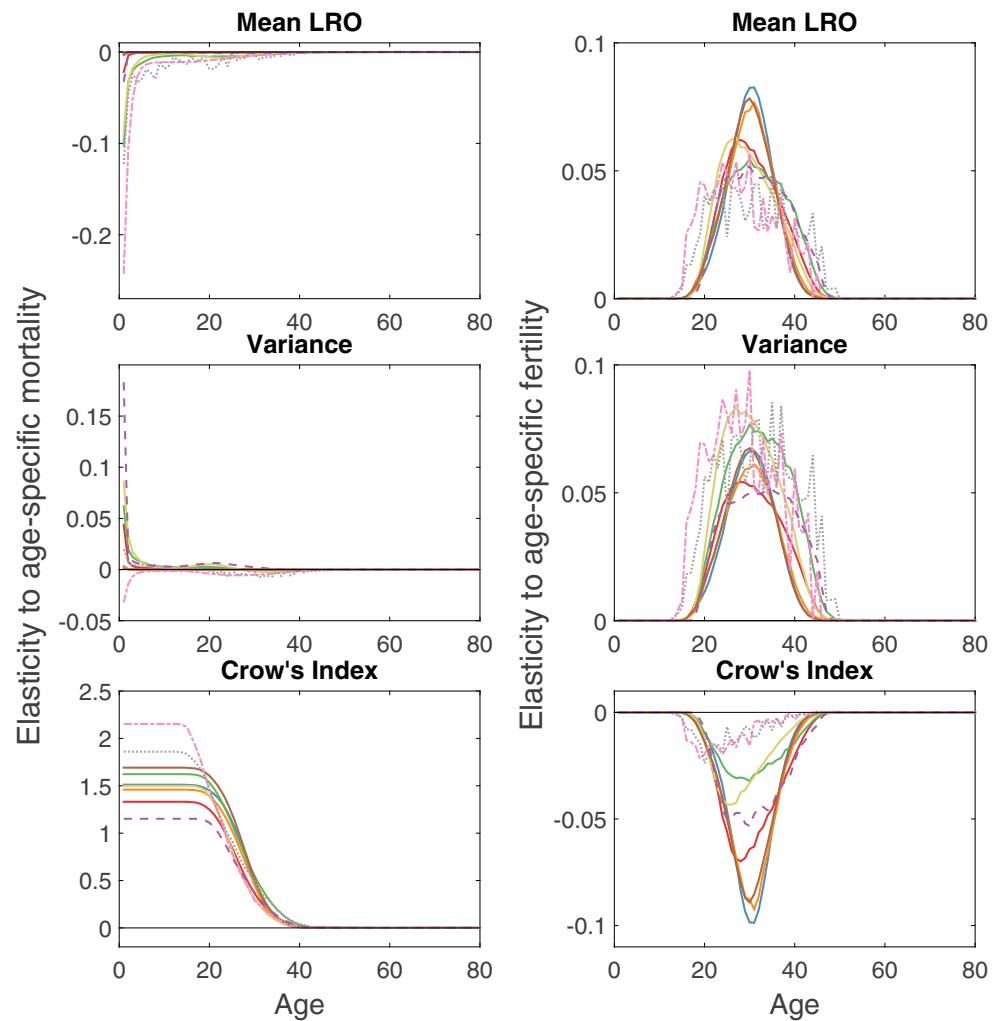
Both mortality and fertility vary widely across ages in these populations, so it may be useful to standardize the responses by calculating elasticities (Fig. 2). The elasticities of the mean and variance of LRO with respect to mortality are generally low, except for effects of infant mortality,

especially for the hunter-gatherer populations, in which infant mortality is high. However, the elasticity of Crow's  $\mathcal{I}$  is large and positive, in fact, the largest of any of the elasticities obtained. The elasticities with respect to fertility are naturally confined to the reproductive ages. Proportional increases in fertility increase the mean and variance of LRO, but reduce the value of Crow's  $\mathcal{I}$ .

### A stage-classified tree population

As an example of a typical stage-classified population, we analyze a model for the Canadian hemlock (*T. canadensis* L.). Lamar and McGraw (2005) developed a model based on six size classes (from  $< 5.0$  cm dbh (diameter at breast height) to  $> 42.5$  cm dbh). They reported population projection matrices for trees in a low-disturbance plot in Shenandoah National Park in the eastern USA, between 1997 and 1999. We obtained  $\mathbf{U}$  and  $\mathbf{F}$  from the mean popu-

**Fig. 2** Elasticity of mean LRO, variance in LRO, and Crow's index to changes in age-specific mortality (left column) and age-specific fertility (right column) for nine human populations



lation matrix **A** obtained from the COMPADRE Plant Matrix Database (COMPADRE Plant Matrix Database 2014).

$$\mathbf{A} = \begin{pmatrix} 0.90 & 0 & 0.30 & 0.77 & 1.96 & 6.03 \\ 0.004 & 0.96 & 0 & 0 & 0 & 0 \\ 0 & 0.012 & 0.97 & 0 & 0 & 0 \\ 0 & 0 & 0.017 & 0.98 & 0 & 0 \\ 0 & 0 & 0 & 0.012 & 0.96 & 0 \\ 0 & 0 & 0 & 0 & 0.018 & 0.99 \end{pmatrix} \quad (64)$$

Reproduction was measured as new recruits (rather than seeds or seedlings) per individual, per year, as a function of size (Lamar and McGraw 2005). In the absence of information on the empirical distribution of size-specific fertility, we use the Poisson distribution to define the reward moment matrices **R**<sub>1</sub> and **R**<sub>2</sub> following Eqs. 11 and 6.

**Results** The mean, variance, and other statistics of lifetime reproductive output for *T. canadensis*, obtained using Theorem 1 and Eqs. 21–24, are shown in Table 3. Because of the high mortality of small trees, the mean LRO is small and life expectancy is only 12 years. However, the variance in

LRO is very high, as is the apparent opportunity for selection  $\mathcal{I}$ . The variance in LRO exceeds that for any of the human populations by two orders of magnitude. The value of  $\mathcal{I}$  is 600 times higher than the highest for human populations in Table 2. Almost all of the variance in LRO is due to variance among pathways. The variance due to stochastic rewards along those pathways is small and approximately equal to the mean, which reflects the assumption of Poisson distributed fertility. It is swamped by the huge differences among pathways of trees that die small and those that become large and reproductive.

**Sensitivity analysis** As in Eq. 43, the transition matrix **U** is the product of the survival matrix **Σ** and a growth matrix **G**. The derivations of the equations below are presented in Appendix A.4. The derivative of the transition matrix **U** with respect to mortality is

$$\frac{d\text{vec } \mathbf{U}}{d\boldsymbol{\mu}^T} = \frac{d\text{vec } \mathbf{U}}{d\text{vec}^T \boldsymbol{\Sigma}} \frac{d\text{vec } \boldsymbol{\Sigma}}{d\boldsymbol{\mu}^T} = -(\mathbf{I} \otimes \mathbf{G}) \mathcal{D}(\text{vec } \mathbf{I})(\mathbf{I} \otimes \mathbf{1}) \mathcal{D}(\mathbf{p}). \quad (65)$$

**Table 3** The statistics of lifetime reproductive output for *T. canadensis*

Mean	V	$V_{\text{between}}$ (%)	$V_{\text{within}}$ (%)	SD	CV	$\mathcal{I}$	Life exp.
1.42	$1.41 \times 10^3$	99.9	0.1	37.54	26.39	696.23	12.16

The sensitivity of  $\mathbf{U}$  to the growth matrix  $\mathbf{G}$ , maintaining the column sums, is

$$\left. \frac{d\text{vec } \mathbf{U}}{d\text{vec}^T \mathbf{G}} \right|_{\text{comp}} = (\boldsymbol{\Sigma} \otimes \mathbf{I}) \frac{d\text{vec } \mathbf{G}}{d\text{vec}^T \boldsymbol{\Theta}}. \quad (66)$$

where  $d\text{vec } \mathbf{G}/d\text{vec}^T \boldsymbol{\Theta}$  is given by Eq. 49.

The sensitivity of LRO to fertility is determined by the derivatives of the  $\mathbf{R}_i$  to fertility. The derivative of the matrix  $\mathbf{R}_1$  of mean fertility is given by Eq. 63. The derivatives of  $\mathbf{R}_2$  and  $\mathbf{R}_3$  with respect to  $\mathbf{f}$  are calculated from the Poisson distribution using Eqs. 6 and 7:

$$\frac{d\text{vec } \mathbf{R}_2}{d\mathbf{f}^T} = [\mathbf{I}_{s^2} + 2\mathcal{D}(\text{vec } \mathbf{R}_1)] (\mathbf{Z}^T \otimes \mathbf{1}_s) \quad (67)$$

$$\frac{d\text{vec } \mathbf{R}_3}{d\mathbf{f}^T} = [\mathbf{I}_{s^2} + 6\mathcal{D}(\text{vec } \mathbf{R}_1) + 3\mathcal{D}(\text{vec } (\mathbf{R}_1 \circ \mathbf{R}_1))] (\mathbf{Z}^T \otimes \mathbf{1}_s). \quad (68)$$

Incorporating (67) and (68) into Eqs. 27–29 and applying Theorem 2 provides the sensitivity of lifetime reproductive output to mortality, fertility, and the growth matrix (see Appendix A.4 for the derivations). To make comparisons across these variables measured on different scales, we calculate the elasticity or proportional sensitivity of LRO using Eq. 42.

An increase in mortality in the first size-class or the last size-class reduces both mean lifetime reproduction of Hemlock individuals and variance in LRO (Fig. 3). Increasing fertility rates increases both mean and variance in LRO for trees in the last size-class. These are the trees with the highest survival and fertility rates.

The elasticity of the mean and variance to the growth matrix  $\mathbf{G}$  is shown in Fig. 4. The results are dominated by the extremely large negative elasticities to the probability of remaining in stage 1 (thereby reducing growth) and the large positive elasticity to the probability of remaining in the largest size class. Increasing the probability of staying in any size-class (again, reducing growth) reduces both mean and variance in LRO.

The mean and variance, the contribution to the variance of different processes, and the sensitivity of these indices to parameters differ between these examples. This reflects the different life history strategies of trees and humans, the difference between age-classified populations with low fertility, and a size-classified population with strongly size-dependent fertility, and the difference between assumptions of Bernoulli or Poisson distributed rewards. Vastly

different life histories can be incorporated into the Markov chain with reward framework, allowing for the investigation of life history in many species, and from different perspectives.

## Discussion

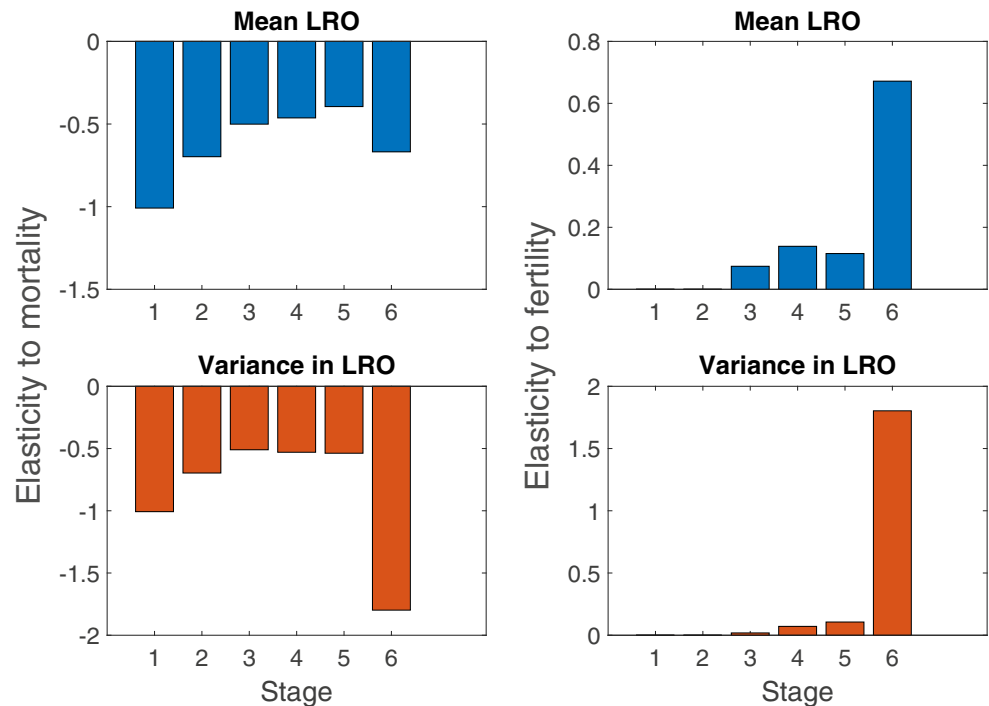
Lifetime reproductive output is an outcome of the life cycle. Any demographic model implies a distribution for LRO, just as it implies more familiar measures such as  $R_0$  and life expectancy. Even in a deterministic environment, the LRO is a random variable; the stochasticity arises from two sources: the random pathway that the individual follows through its life and the random fertility it exhibits at each stage on that pathway.

Our results in “[Markov chains with rewards as a model for LRO](#)” provide the analytical machinery needed to calculate all the statistical properties of LRO that follow from a specified set of stages, a transition matrix, and the moments of stage-specific fertility. These statistical properties include measures of variability among individuals (the variance, CV, skewness, opportunity for selection, etc.). It is important to recognize that this variability does not reflect heterogeneity, genetic or otherwise. Every individual is subject to the same set of vital rates at any stage in the life cycle. Only the outcomes of applying those rates vary; the variation is thus due to individual stochasticity.

The variance, or standardized variance, calculated from demographic models for a variety of species, is large (see also Caswell 2011; van Daalen and Caswell 2015; Steiner et al. 2010). Individual stochasticity creates a large apparent opportunity for selection that is not, in fact, a true opportunity. In this sense, as emphasized by Steiner and Tuljapurkar (2012), the calculations of individual stochasticity provide a neutral model for LRO. A number of comparisons of calculated variance (due to stochasticity) and observed variance (due to some mix of stochasticity and heterogeneity) have shown that stochasticity may explain a significant amount, or all, of the observed variance (Caswell 2011; Van Daalen and Caswell 2015; Steiner et al. 2010).

It is important to remember what neutral model results do, and do not, imply (Caswell 1976). The results show that a certain amount of variance can be accounted for by stochasticity, and hence, that the mere observation of such variation is no evidence for heterogeneity. It does not prove that the observed variance is stochastic, or that there is no heterogeneity, as pointed out by Steiner and Tuljapurkar

**Fig. 3** Elasticity of mean LRO and variance in LRO to changes in stage-specific mortality (*left column*) and stage-specific fertility (*right column*) for *T. canadensis*



(2012). It calls for a comparison with models explicitly incorporating heterogeneity, either observed or unobserved, as suggested by Cam et al. (2016). For examples of this approach, see Caswell (2014); Hartemink et al. (2017); Jenouvrier et al. (2016).

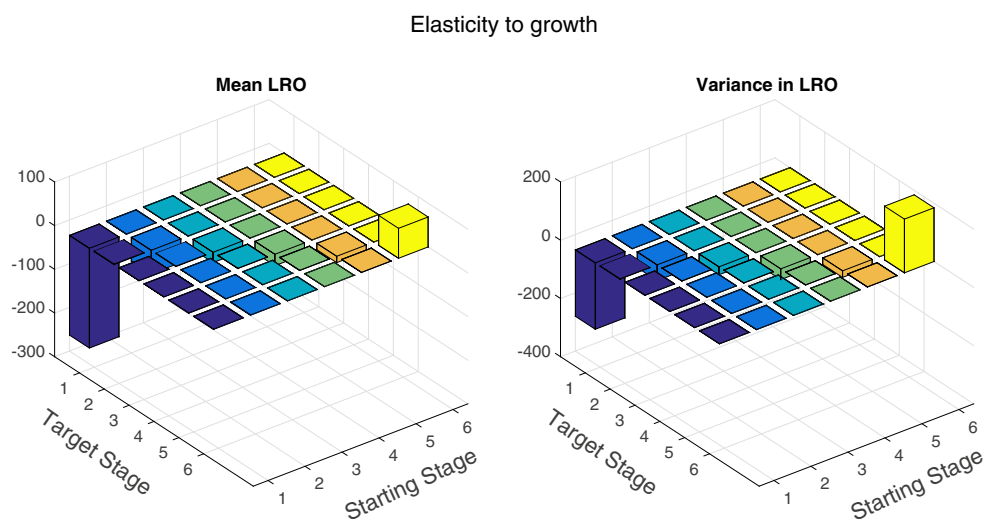
The results of Theorem 1 provide an exact solution to the calculation of the statistics of LRO, including both components (within and between pathways; see “[Partitioning variance: within and between pathways](#)”). We also provide a complete sensitivity analysis for LRO. Theorem 2 makes it possible to calculate the sensitivity and elasticity of all the moments of LRO and all the statistics calculated from those moments, with respect to changes in mortality, transition probabilities, and the moments of stage-specific fertility.

The results include the sensitivity to changes in mean fertility (holding variance constant) and variance in fertility (holding the mean constant).

The formulae for the moments in Theorem 1 and the sensitivities in Theorem 2 are complicated and opaque, because the relationships between LRO and the life cycle structure, the moments of reproduction, the probabilities of survival, and the infinite diversity of pathways through the life cycle, are complicated. Simplifications that permit qualitative generalities are always welcome, and more work on this will be valuable.

We applied sensitivity analysis to several populations of humans and a population of trees. The patterns of sensitivity and elasticity of LRO that we report for these populations

**Fig. 4** Sensitivity of the mean LRO and variance in LRO of *Tsuga canadensis* to the growth matrix (*left and right panels*, respectively)





have not been described before. Some suggestive patterns appear; they warrant further investigation.

In long-lived age-classified populations with low reproductive output, as diverse as the nineteenth century Swedes, mid-twentieth century Hutterites, the twenty-first century Dutch, and Hadza and Ache hunter-gatherers, the sensitivity of mean LRO to mortality is negative. Most populations show a positive sensitivity of variance in LRO to mortality for the first 40 years of life, only showing a small negative sensitivity between ages 20 and 40. The Ache and Hadza have more pronounced negative sensitivity around these ages, with the Hadza showing negative sensitivity across the first 40 years of life. The sensitivity of Crow's  $\mathcal{I}$  to mortality in the first 40 years of life is positive, showing that an increase in mortality would increase the apparent opportunity for selection of lifetime reproduction. Patterns for elasticities are similar, though smaller in magnitude.

All populations show broadly similar patterns for the sensitivity of LRO to fertility. The sensitivities of mean LRO and variance in LRO to fertility are positive. The sensitivity of Crow's  $\mathcal{I}$  to fertility is negative. The elasticity of mean and variance to fertility is positive, but the elasticity of Crow's  $\mathcal{I}$  is negative. According to these results, populations in modern countries would reduce the apparent opportunity for selection of LRO with increasing fertilities, and high-fertility populations such as Hutterites and hunter-gatherers would only slightly reduce the opportunity for selection should fertilities increase.

On average, the variance in LRO is 59% between pathways and 41% within pathways. However, the individual populations differ significantly in these contributions by sources of variance. In the twenty-first century in Japan, Sweden, and the Netherlands, only 1.2% of variance in LRO is between pathways and 98.8% is within pathways. Variance in LRO of the high-fertility populations, the Ache, Hadza, and Hutterites, is 71% between pathways and 29% within pathways.

In contrast, in *T. canadensis*, with size-dependent demography, high fertility, and strongly increasing fertility with size, we find that the elasticities of both mean and variance in LRO to mortality are negative across all size classes. The elasticities of these statistical properties to changes in fertility in the reproductive classes are positive. In case of the elasticity to growth transition rates, mean and variance in LRO once again show similar patterns. Elasticity to stasis, i.e., not growing, is negative for the first five size classes, and positive for the last size class.

Lifetime reproductive output interests population ecologists and epidemiologists (for whom  $R_0$  is a measure of population growth), evolutionary biologists (for whom variance in LRO is a measure of potential selection), and human demographers (for whom declines in LRO following the demographic transition pose serious social policy

challenges). It is thus important that our analysis is not restricted to any one class of population models. It applies equally to age-structured, stage-structured, and multistate models, and to any reproductive strategy. It also applies to periodic and stochastic time-varying models, by applying Theorems 1 and 2 to the models in Caswell (2011). The method can also be applied to rewards other than reproductive output, including health status (Caswell and Zarulli 2015) and economic transfers (Caswell and Kluge 2015).

The transition matrix  $\mathbf{U}$  is part of any population projection matrix; the COMPADRE and COMADRE matrix databases provide many examples (Salguero-Gómez et al. 2016; Salguero-Gómez et al. 2015). The *mean* reproductive reward matrix  $\mathbf{R}_1$  can be obtained from the projection matrix, but the higher moments cannot and require assumptions of a parametric distribution for fertility. We encourage researchers with the appropriate reproduction data to report not only mean fertility but also the higher moments, or even the entire distribution.

Individual stochasticity arises in both reproductive output and in survival or longevity. Our results here complement the analysis of variation in longevity using Markov chain methods, which are widely used in ecology (e.g., Cochran and Ellner 1992, Caswell 2001, 2006, 2009, Horvitz and Tuljapurkar 2008, Tuljapurkar and Horvitz 2006) and have well-developed sensitivity analyses. The Markov chain with reward model now permits a similarly deep analysis of lifetime reproduction and its sensitivity analysis.

**Acknowledgments** This research was supported by ERC Advanced Grant 322989 and NWO Project ALWOP.2015.100. We thank two anonymous reviewers for helpful comments.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## A Derivations

### A.1 Proof of Theorem 1: moments of LRO

We assume a finite state absorbing Markov chain with transition matrix  $\mathbf{P}$  given by Eq. 3, with  $\tau$  transient states,  $\alpha$  absorbing states, and  $s = \tau + \alpha$  total states. The  $(i, j)$  entry of the  $s \times s$  matrix  $\mathbf{R}_k$  is the  $k$ th moment of the reproductive reward associated with the transition from state  $j$  to state  $i$ . We assume that no rewards accrue to individuals in absorbing states, so columns  $\tau + 1$  to  $s$  of  $\mathbf{R}_k$  are zero.

The  $i$ th entry of the  $s \times 1$  vector  $\rho_m$  is the  $m$ th moment or remaining lifetime reproductive output for an individual

in stage  $i$ . Caswell (2011, Proposition 1) derived an iterative approximation for the moment vectors,

$$\rho_m(t+1) = \sum_{k=0}^m \binom{m}{k} (\mathbf{P} \circ \mathbf{R}_{m-k})^\top \rho_k(t) \quad m = 1, 2, \dots \quad (69)$$

which converges to the lifetime accumulation, for each stage, as  $t \rightarrow \infty$ . To obtain an analytical solution for the moment vectors, we solve (69) for its equilibrium.

The equilibrium moment vector  $\rho_m$  satisfies

$$\rho_m = \binom{m}{0} (\mathbf{P} \circ \mathbf{R}_m)^\top \rho_0 + \sum_{k=1}^{m-1} \binom{m}{k} (\mathbf{P} \circ \mathbf{R}_m)^\top \rho_k + \binom{m}{m} (\mathbf{P} \circ \mathbf{R}_0)^\top \rho_m \quad (70)$$

Because  $\rho_0 = \mathbf{1}_s$  and  $\mathbf{R}_0$  is a  $s \times s$  matrix of ones, this simplifies to

$$(\mathbf{I} - \mathbf{P}^\top) \rho_m = (\mathbf{P} \circ \mathbf{R}_m)^\top \mathbf{1}_s + \sum_{k=1}^{m-1} \binom{m}{k} (\mathbf{P} \circ \mathbf{R}_{m-k})^\top \rho_k. \quad (71)$$

Solving (71) directly for  $\rho_m$  is impossible because the matrix  $(\mathbf{I} - \mathbf{P}^\top)$  is singular. However, since no rewards accumulate in the absorbing states, so we need solve only for the vector of remaining LRO from the transient states, which we denote by  $\tilde{\rho}_m$ . Multiplying both sides of Eq. 70 by the matrix  $\mathbf{Z}$  in Eq. 14 gives

$$\tilde{\rho}_m = \mathbf{Z} (\mathbf{P} \circ \mathbf{R}_m)^\top \mathbf{1}_s + \sum_{k=1}^{m-1} \binom{m}{k} \mathbf{Z} (\mathbf{P} \circ \mathbf{R}_{m-k})^\top \mathbf{Z}^\top \tilde{\rho}_k + \mathbf{Z} \mathbf{P}^\top \mathbf{Z}^\top \tilde{\rho}_m \quad (72)$$

Noting that

$$\mathbf{Z} \mathbf{P}^\top \mathbf{Z}^\top = \mathbf{U}^\top \quad (73)$$

$$\mathbf{Z} (\mathbf{P} \circ \mathbf{R}_{m-k})^\top \mathbf{Z}^\top = (\mathbf{U} \circ \tilde{\mathbf{R}}_{m-k})^\top \quad (74)$$

we see that  $\tilde{\rho}_m$  must satisfy

$$\tilde{\rho}_m = \mathbf{Z} (\mathbf{P} \circ \mathbf{R}_m)^\top \mathbf{1}_s + \sum_{k=1}^{m-1} \binom{m}{k} (\mathbf{U} \circ \tilde{\mathbf{R}}_{m-k})^\top \tilde{\rho}_k + \mathbf{U}^\top \tilde{\rho}_m \quad (75)$$

Solving (75) for  $\tilde{\rho}_m$  gives (19) and completes the proof of Theorem 1.

## A.2 Proof of Theorem 2: sensitivity analysis of moments of LRO

To derive the sensitivity result of Theorem 2, we will break the result (19) into three types of terms and differentiate each of these in turn. We rewrite (19) as

$$\underbrace{(\mathbf{N}^\top)^{-1} \tilde{\rho}_m}_A = \underbrace{\mathbf{Z} (\mathbf{P} \circ \mathbf{R}_m)^\top \mathbf{1}_s}_B + \sum_{k=1}^{m-1} \binom{m}{k} \underbrace{(\mathbf{U} \circ \tilde{\mathbf{R}}_{m-k})^\top \tilde{\rho}_k}_C \quad (76)$$

Equation 76 contains three types of terms:

$$A = (\mathbf{N}^\top)^{-1} \tilde{\rho} \quad (77)$$

$$B = \mathbf{Z} (\mathbf{P} \circ \mathbf{R})^\top \mathbf{1}_s \quad (78)$$

$$C = (\mathbf{U} \circ \tilde{\mathbf{R}})^\top \tilde{\rho} \quad (79)$$

We differentiate each of these in turn. Differentiating (77) gives

$$dA = d[(\mathbf{N}^\top)^{-1}] \tilde{\rho} + (\mathbf{N}^\top)^{-1} d\tilde{\rho}. \quad (80)$$

Substituting

$$(\mathbf{N}^\top)^{-1} = \mathbf{I}_s - \mathbf{U}^\top \quad (81)$$

gives

$$dA = -(d\mathbf{U}^\top) \tilde{\rho} + (\mathbf{I} - \mathbf{U}^\top) d\tilde{\rho}, \quad (82)$$

and applying the vec operator yields

$$d\text{vec} A = -(\tilde{\rho}^\top \otimes \mathbf{I}) \mathbf{K}_{\tau\tau} d\text{vec} \mathbf{U} + (\mathbf{I} - \mathbf{U}^\top) d\tilde{\rho} \quad (83)$$

Differentiating term  $B$  gives

$$dB = \mathbf{Z} (d\mathbf{P} \circ \mathbf{R})^\top \mathbf{1}_s + \mathbf{Z} (\mathbf{P} \circ d\mathbf{R})^\top \mathbf{1}_s \quad (84)$$

Applying the vec operator to both sides of Eq. 84 yields

$$d\text{vec} B = (\mathbf{1}^\top \otimes \mathbf{Z}) \mathbf{K}_{ss} [\mathcal{D}(\text{vec} \mathbf{R}) d\text{vec} \mathbf{P} + \mathcal{D}(\text{vec} \mathbf{P}) d\text{vec} \mathbf{R}] \quad (85)$$

Differentiating term  $C$  gives

$$dC = (d\mathbf{U} \circ \tilde{\mathbf{R}})^\top \tilde{\rho} + (\mathbf{U} \circ d\tilde{\mathbf{R}})^\top \tilde{\rho} + (\mathbf{U} \circ \tilde{\mathbf{R}})^\top d\tilde{\rho}. \quad (86)$$

Applying the vec operator yields

$$d\text{vec} C = (\tilde{\rho}^\top \otimes \mathbf{I}) \mathbf{K}_{\tau\tau} [\mathcal{D}(\text{vec} \tilde{\mathbf{R}}) d\text{vec} \mathbf{U} + \mathcal{D}(\text{vec} \mathbf{U}) d\text{vec} \tilde{\mathbf{R}}] + (\mathbf{U} \circ \tilde{\mathbf{R}})^\top d\tilde{\rho} \quad (87)$$

Next, differentiate (76) and substitute (83)–(87) where appropriate; this leads to

$$\begin{aligned}
 (\mathbf{N}^\top)^{-1} d\tilde{\boldsymbol{\rho}}_m &= \underbrace{(\mathbf{1}_s^\top \otimes \mathbf{Z}) \mathbf{K}_{ss} [\mathcal{D}(\text{vec } \mathbf{R}_m) d\text{vec } \mathbf{P} + \mathcal{D}(\text{vec } \mathbf{P}) d\text{vec } \mathbf{R}_m]}_{\mathbf{V}_m} \\
 &+ \sum_{k=1}^{m-1} \binom{m}{k} (\tilde{\boldsymbol{\rho}}_k \otimes \mathbf{I}_\tau) \mathbf{K}_{\tau\tau} \underbrace{\left[ \mathcal{D}(\text{vec } \tilde{\mathbf{R}}_{m-k}) d\text{vec } \mathbf{U} + \mathcal{D}(\text{vec } \mathbf{U}) d\text{vec } \tilde{\mathbf{R}}_{m-k} \right]}_{\mathbf{W}_{k,m-k}} + (\mathbf{U} \circ \tilde{\mathbf{R}}_{m-k})^\top d\tilde{\boldsymbol{\rho}}_k \\
 &+ \underbrace{(\tilde{\boldsymbol{\rho}}_m^\top \otimes \mathbf{I}_\tau) \mathbf{K}_{\tau\tau} d\text{vec } \mathbf{U}}_{\mathbf{X}_m}
 \end{aligned} \quad (88)$$

Using the terms  $\mathbf{V}_m$ ,  $\mathbf{W}_{k,m-k}$ , and  $\mathbf{X}_m$  identified in Eq. 88, multiplying both sides by  $\mathbf{N}^\top$ , and replacing differentials with derivatives with respect to a vector  $\boldsymbol{\theta}$  of parameters give

$$\frac{d\tilde{\boldsymbol{\rho}}_m}{d\boldsymbol{\theta}^\top} = \mathbf{N}^\top \left( \mathbf{V}_m + \sum_{k=1}^{m-1} \binom{m}{k} \mathbf{W}_{k,m-k} + \mathbf{X}_m \right) \quad (89)$$

which completes the derivation of Theorem 2.

The term  $\mathbf{V}_i$  in Eq. 89 contains the derivative  $d\text{vec } \mathbf{P}/d\boldsymbol{\theta}^\top$ . Because  $\mathbf{P}$  is a block structured matrix and  $\mathbf{0}$  and  $\mathbf{I}$  are constants, the derivative of  $\mathbf{P}$  can be written as a linear combination of the derivatives of  $\mathbf{U}$  and  $\mathbf{M}$ ,

$$\frac{d\text{vec } \mathbf{P}}{d\boldsymbol{\theta}^\top} = \mathbf{C}_1 \frac{d\text{vec } \mathbf{U}}{d\boldsymbol{\theta}^\top} + \mathbf{C}_2 \frac{d\text{vec } \mathbf{M}}{d\boldsymbol{\theta}^\top}, \quad (90)$$

where

$$\mathbf{C}_1 = \begin{pmatrix} \mathbf{I}_\tau \\ \mathbf{0}_{\alpha \times \tau} \end{pmatrix} \otimes \begin{pmatrix} \mathbf{I}_\tau \\ \mathbf{0}_{\alpha \times \tau} \end{pmatrix} \quad (91)$$

$$\mathbf{C}_2 = \begin{pmatrix} \mathbf{I}_\tau \\ \mathbf{0}_{\alpha \times \tau} \end{pmatrix} \otimes \begin{pmatrix} \mathbf{0}_{\tau \times \alpha} \\ \mathbf{I}_\alpha \end{pmatrix} \quad (92)$$

(Caswell and van Daalen 2016). In many applications, only a single absorbing state (death) exists, in which case  $\alpha = 1$  and Eqs. 91 and 92 simplify to Eq. 35.

### A.3 Sensitivity analysis of the statistics of LRO

The sensitivity analysis of the statistical properties of LRO is obtained by differentiating and applying the vec operator to Eqs. 21–24. Several matrix calculus results are used frequently in these sensitivity calculations. Roth's theorem, i.e.,

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}^\top \otimes \mathbf{A}) \text{vec } \mathbf{B}, \quad (93)$$

is applied often, as is the rule for taking the vec of a Hadamard product,

$$\text{vec}(\mathbf{A} \circ \mathbf{B}) = \mathcal{D}(\mathbf{A}) \text{vec}(\mathbf{B}) + \mathcal{D}(\mathbf{B}) \text{vec}(\mathbf{A}). \quad (94)$$

For notational clarity, we redefine some of the statistical properties of LRO as follows in this section only:

$$\mathbf{v} = V(\tilde{\boldsymbol{\rho}}) \quad (95)$$

$$\mathbf{s} = SD(\tilde{\boldsymbol{\rho}}) \quad (96)$$

$$\mathbf{C} = CV(\tilde{\boldsymbol{\rho}}) \quad (97)$$

$$\mathcal{I} = \mathcal{I}(\tilde{\boldsymbol{\rho}}) \quad (98)$$

$$\mathbf{D} = \mathcal{D}(\tilde{\boldsymbol{\rho}}_1) \quad (99)$$

**Variance** Differentiating the variance in lifetime reproductive output as given by Eq. 21 gives

$$d\mathbf{v} = d\tilde{\boldsymbol{\rho}}_2 - d(\tilde{\boldsymbol{\rho}}_1 \circ \tilde{\boldsymbol{\rho}}_1). \quad (100)$$

Applying the vec operator results in

$$d\mathbf{v} = d\tilde{\boldsymbol{\rho}}_2 - d\text{vec}(\tilde{\boldsymbol{\rho}}_1 \circ \tilde{\boldsymbol{\rho}}_1), \quad (101)$$

which, using the rule for taking the vec of a Hadamard product, yields

$$d\mathbf{v} = d\tilde{\boldsymbol{\rho}}_2 - \mathbf{D}d\tilde{\boldsymbol{\rho}}_1 - \mathbf{D}d\tilde{\boldsymbol{\rho}}_1. \quad (102)$$

Taking the derivatives with respect to a vector of parameters  $\boldsymbol{\theta}$  and rearranging give

$$\frac{d\mathbf{v}}{d\boldsymbol{\theta}^\top} = \frac{d\tilde{\boldsymbol{\rho}}_2}{d\boldsymbol{\theta}^\top} - 2\mathbf{D} \frac{d\tilde{\boldsymbol{\rho}}_1}{d\boldsymbol{\theta}^\top}, \quad (103)$$

as found in Eq. 38.

**Standard deviation** To obtain the sensitivity of the standard deviation of LRO, we rewrite (22) as

$$\mathbf{s} \circ \mathbf{s} = \mathbf{v} \quad (104)$$

and differentiate this giving us

$$d(\mathbf{s} \circ \mathbf{s}) = d\mathbf{v}. \quad (105)$$

Using the vec operator, we obtain

$$d\text{vec}(\mathbf{s} \circ \mathbf{s}) = d\mathbf{v}, \quad (106)$$

where, after replacing the differentials with derivatives with respect to  $\theta$  and applying the Hadamard-rule on the left-hand side, we find

$$2\mathcal{D}(\mathbf{s}) \frac{d\mathbf{s}}{d\theta^\top} = \frac{d\mathbf{v}}{d\theta^\top} \quad (107)$$

Solving for  $ds/d\theta^\top$  yields Eq. 39.

**Coefficient of variation** The sensitivity of the coefficient of variation is obtained by differentiating

$$\mathbf{C} = \mathbf{D}^{-1}\mathbf{s} \quad (108)$$

from Eq. 23, resulting in

$$d\mathbf{C} = d(\mathbf{D}^{-1})\mathbf{s} + \mathbf{D}^{-1}d(\mathbf{s}). \quad (109)$$

However, for any nonsingular matrix  $\mathbf{Y}$ ,

$$d(\mathbf{Y}^{-1}) = -\mathbf{Y}^{-1}d(\mathbf{Y})\mathbf{Y}^{-1}. \quad (110)$$

Thus, Eq. 109 can be written as

$$d\mathbf{C} = -\mathbf{D}^{-1}d(\mathbf{D})\mathbf{D}^{-1}\mathbf{s} + \mathbf{D}^{-1}d(\mathbf{s}). \quad (111)$$

Applying the vec operator and Roth's theorem yields

$$d\mathbf{C} = -\left(\mathbf{s}^\top [\mathbf{D}^{-1}]^\top \otimes \mathbf{D}^{-1}\right) d\text{vec}(\mathbf{D}) + \mathbf{D}^{-1}d(\mathbf{s}). \quad (112)$$

Noting that

$$\mathbf{s}^\top (\mathbf{D}^{-1})^\top = \mathbf{C}^\top \quad (113)$$

and that  $\mathbf{D}$  can be rewritten as

$$\mathbf{D} = \mathbf{I} \circ (\mathbf{1}_\tau \tilde{\rho}_1^\top) \quad (114)$$

we obtain

$$d\mathbf{C} = -\left(\mathbf{C}^\top \otimes \mathbf{D}^{-1}\right) d\text{vec}\left(\mathbf{I} \circ [\mathbf{1}_\tau \tilde{\rho}_1^\top]\right) + \mathbf{D}^{-1}d(\mathbf{s}). \quad (115)$$

Using the rule for taking the vec of a Hadamard product and applying Roth's theorem once more yields

$$d\mathbf{C} = -\left(\mathbf{C}^\top \otimes \mathbf{D}^{-1}\right) \mathcal{D}(\text{vec } \mathbf{I})(\mathbf{I} \otimes \mathbf{1}_\tau) d(\tilde{\rho}_1) + \mathbf{D}^{-1}d(\mathbf{s}). \quad (116)$$

Replacing the differentials with derivatives with respect to  $\theta$  gives

$$\frac{d\mathbf{C}}{d\theta^\top} = -\left(\mathbf{C}^\top \otimes \mathbf{D}^{-1}\right) \mathcal{D}(\text{vec } \mathbf{I})(\mathbf{I} \otimes \mathbf{1}_\tau) \frac{d\tilde{\rho}_1}{d\theta^\top} + \mathbf{D}^{-1} \frac{d\mathbf{s}}{d\theta^\top}. \quad (117)$$

**Crow's index  $\mathcal{I}$**  The opportunity for selection, also known as Crow's index, can be calculated from the coefficient of variation, as shown in Eq. 24. Differentiating this equation results in

$$d\mathcal{I} = d\text{vec}(\mathbf{C} \circ \mathbf{C}). \quad (118)$$

Applying the vec operator to the Hadamard product yields

$$d\mathcal{I} = \mathcal{D}(\mathbf{C})d(\mathbf{C}) + \mathcal{D}(\mathbf{C})d(\mathbf{C}). \quad (119)$$

Replacing the differentials with derivatives of Crow's  $\mathcal{I}$  with respect to  $\theta$  yields

$$\frac{d\mathcal{I}}{d\theta^\top} = 2\mathcal{D}(\mathbf{C}) \frac{d\mathbf{C}}{d\theta^\top} \quad (120)$$

#### A.4 Deriving sensitivities for the examples

To obtain the sensitivities for mean LRO and other statistics, the following pieces are required:

$$\frac{d\text{vec } \mathbf{U}}{d\theta^\top}, \frac{d\text{vec } \mathbf{P}}{d\theta^\top}, \text{ and } \frac{d\text{vec } \mathbf{R}_i}{d\theta^\top}. \quad (121)$$

In the case of both humans and trees, a single stage of death is incorporated, so that the derivative of  $\mathbf{P}$  with respect to  $\theta$  can be obtained from the derivative of  $\mathbf{U}$  with respect to  $\theta$ , as shown in Eq. 35. As the transition matrix depends on mortality rates and the reward matrix depends on fertility rates, the required pieces for sensitivity analysis become

$$\frac{d\text{vec } \mathbf{U}}{d\mu^\top} \text{ and } \frac{d\text{vec } \mathbf{R}_i}{d\mathbf{f}^\top}. \quad (122)$$

The way these matrices depend on the underlying vectors of parameters differs between species, as shown below for humans and trees.

**Humans** In humans,  $\mathbf{U}$  is a function of the mortality rates  $\mu$  through its survival probabilities  $\sigma$ .  $\mathbf{U}$  depends on  $\sigma$  as

$$\mathbf{U} = \mathbf{Y} \circ (\mathbf{1}\sigma^\top), \quad (123)$$

where  $\mathbf{Y}$  is a  $\tau \times \tau$  matrix with ones on the subdiagonal and zeros elsewhere. Given that

$$\sigma = e^{-\mu}, \quad (124)$$

the derivative of  $\sigma$  becomes

$$d\sigma = -\mathcal{D}(\sigma)d\mu. \quad (125)$$

As such, the complete analytical expression for the sensitivity of the transition matrix to the mortality schedule is

$$\frac{d\text{vec } \mathbf{U}}{d\mu^\top} = -\mathcal{D}(\text{vec } \mathbf{Y})(\mathbf{I} \otimes \mathbf{1})\mathcal{D}(\sigma). \quad (126)$$

The sensitivity of the moments of  $\mathbf{R}$  to a vector of fertility rates depends on the kind of reproductive strategy used. In the case of humans, fertility rate is the *chance* of producing a single offspring. With the subsequent assumption that the moments of the reward matrix follow a Bernoulli distribution,

$$\frac{d\mathbf{R}_1}{d\mathbf{f}^\top} = \frac{d\mathbf{R}_2}{d\mathbf{f}^\top} = \frac{d\mathbf{R}_3}{d\mathbf{f}^\top} \quad (127)$$

Therefore, we only need to derive the sensitivity of  $\mathbf{R}_1$  to  $\mathbf{f}$  to know all the moments of the reward matrix. Equation 11 can be rewritten as

$$\mathbf{R}_1 = \mathbf{I}_s \mathbf{f}^\top \mathbf{Z}. \quad (128)$$

Differentiating this entails taking the derivative on both sides and applying Roth's theorem. Then, the equation for the sensitivity is simply

$$\frac{d\mathbf{R}_1}{d\mathbf{f}^\top} = (\mathbf{Z}^\top \otimes \mathbf{I}_s). \quad (129)$$

**Trees** In trees, the transition matrix consists not only of survival rates but also growth/stasis/regression rates (see Equation 43). By differentiating (43) with respect to  $\Sigma$ , we obtain

$$\frac{d\text{vec } \mathbf{U}}{d\text{vec}^\top \Sigma} = (\mathbf{I}^\top \otimes \mathbf{G}). \quad (130)$$

As  $\mathbf{U}$  now depends on mortality rate through  $\Sigma$ , the sensitivity of the transition matrix to  $\mu$  is

$$\frac{d\text{vec } \mathbf{U}}{d\mu^\top} = \frac{d\text{vec } \mathbf{U}}{d\text{vec}^\top \Sigma} \frac{d\text{vec } \Sigma}{d\mu^\top}. \quad (131)$$

Given that  $\Sigma$  can be written as  $\Sigma = \mathbf{I} \circ (\mathbf{1}\sigma^\top)$ , differentiating, applying the vec operator, using the rule for taking the vec of a Hadamard product, and applying Roth's theorem give

$$\frac{d\text{vec } \Sigma}{d\sigma^\top} = \mathcal{D}(\text{vec } \mathbf{I})(\mathbf{I}^\top \otimes \mathbf{1}). \quad (132)$$

Combining Eqs. 130 and 132 with the derivative of  $\sigma$  to  $\mu$  derived in Eqs. 125 and 131 becomes

$$\frac{d\text{vec } \mathbf{U}}{d\mu^\top} = -(\mathbf{I}^\top \otimes \mathbf{G})\mathcal{D}(\text{vec } \mathbf{I})(\mathbf{I}^\top \otimes \mathbf{1})\mathcal{D}(\sigma). \quad (133)$$

The sensitivity of the transition matrix to the growth matrix is found by differentiating (43) with respect to  $\mathbf{G}$

$$\frac{d\text{vec } \mathbf{U}}{d\text{vec}^\top \mathbf{G}} = (\Sigma^\top \otimes \mathbf{I}). \quad (134)$$

However, a complication arises from the fact that  $\mathbf{G}$  is column-stochastic, that is, the probabilities of staying or moving between stages always sums to 1. Sensitivity analysis works on the premise of a small change in the underlying parameters effecting a change in the model outcome. In the case of  $\mathbf{G}$ , a small additive change would result in the loss of the column-stochastic property. Fortunately, we can instead use a proportional change. Caswell (2001, 2013) presents a method of proportionally compensating for a perturbation of a single entry in a column-stochastic matrix by proportionally subtracting the perturbation from the other entries in that column. In matrix notation, this becomes

$$\frac{d\text{vec } \mathbf{U}}{d\text{vec}^\top \Theta} = \frac{d\text{vec } \mathbf{U}}{d\text{vec}^\top \mathbf{G}} \frac{d\text{vec } \mathbf{G}}{d\text{vec}^\top \Theta} \quad (135)$$

where

$$\frac{d\text{vec } \mathbf{G}}{d\text{vec}^\top \Theta} = \mathbf{I}_{s^2} - \sum_{i=1}^s (\mathbf{E}_{ii} \otimes \mathcal{D}[\mathbf{G}(:, i)] \mathbf{C} \mathcal{D}[\mathbf{I} - \mathbf{G}(:, i)]^{-1}),$$

with  $\mathbf{C} = \mathbf{E} - \mathbf{I}$ . (136)

In trees, reproduction is assumed to adhere to a Poisson distribution as opposed to a Bernoulli distribution. Higher moments of the reward matrix therefore depend on  $\mathbf{R}_1$  as follows:

$$\mathbf{R}_2 = \mathbf{R}_1 + \mathbf{R}_1 \circ \mathbf{R}_1 \text{ and} \quad (137)$$

$$\mathbf{R}_3 = \mathbf{R}_1 + 3\mathbf{R}_1 \circ \mathbf{R}_1 + \mathbf{R}_1 \circ \mathbf{R}_1 \circ \mathbf{R}_1. \quad (138)$$

By differentiating the expression for  $\mathbf{R}_2$  we obtain

$$d\mathbf{R}_2 = d[\mathbf{R}_1] + d[\mathbf{R}_1 \circ \mathbf{R}_1]. \quad (139)$$

Applying the vec operator yields

$$d\text{vec } \mathbf{R}_2 = d\text{vec } \mathbf{R}_1 + 2\mathcal{D}(\text{vec } \mathbf{R}_1)d\text{vec } \mathbf{R}_1. \quad (140)$$

Replacing the differentials with the derivative of  $\mathbf{R}_2$  with respect to  $\mathbf{R}_1$  gives

$$\frac{d\text{vec } \mathbf{R}_2}{d\text{vec } \mathbf{R}_1} = \mathbf{I}_{s^2} + 2\mathcal{D}(\text{vec } \mathbf{R}_1). \quad (141)$$

Then, the derivative of  $\mathbf{R}_2$  with respect to  $\mathbf{f}$  becomes

$$\frac{d\text{vec } \mathbf{R}_2}{d\mathbf{f}^\top} = \frac{d\text{vec } \mathbf{R}_2}{d\text{vec } \mathbf{R}_1} \frac{d\text{vec } \mathbf{R}_1}{d\mathbf{f}^\top}, \quad (142)$$

which, using the expression above and recalling (129), becomes

$$\frac{d\text{vec } \mathbf{R}_2}{d\mathbf{f}^\top} = (\mathbf{I}_{s^2} + 2\mathcal{D}(\text{vec } \mathbf{R}_1))(\mathbf{Z}^\top \otimes \mathbf{I}_s). \quad (143)$$

Similarly, we can obtain the sensitivity of  $\mathbf{R}_3$  to  $\mathbf{f}$  by differentiating (138), resulting in

$$d\mathbf{R}_3 = d\mathbf{R}_1 + 3d(\mathbf{R}_1 \circ \mathbf{R}_1) + d(\mathbf{R}_1 \circ \mathbf{R}_1 \circ \mathbf{R}_1). \quad (144)$$

Applying the vec operator then gives

$$d\text{vec } \mathbf{R}_3 = d\text{vec } \mathbf{R}_1 + 6\mathcal{D}(\text{vec } \mathbf{R}_1)d\text{vec } \mathbf{R}_1 + 3\mathcal{D}(\text{vec}(\mathbf{R}_1 \circ \mathbf{R}_1))d\text{vec } \mathbf{R}_1, \quad (145)$$

where the differentials can be replaced with derivatives of  $\mathbf{R}_3$  with respect to  $\mathbf{R}_1$ , yielding

$$\frac{d\text{vec } \mathbf{R}_3}{d\text{vec } \mathbf{R}_1} = \mathbf{I}_{s^2} + 6\mathcal{D}(\text{vec } \mathbf{R}_1) + 3\mathcal{D}(\text{vec}(\mathbf{R}_1 \circ \mathbf{R}_1)). \quad (146)$$

Given that

$$\frac{d\text{vec } \mathbf{R}_3}{d\mathbf{f}^\top} = \frac{d\text{vec } \mathbf{R}_3}{d\text{vec } \mathbf{R}_1} \frac{d\text{vec } \mathbf{R}_1}{d\mathbf{f}^\top}, \quad (147)$$



and that the derivative of  $\mathbf{R}_1$  with respect to  $\mathbf{f}$  is (129), the sensitivity of  $\mathbf{R}_3$  to  $\mathbf{f}$  is

$$\frac{d\text{vec } \mathbf{R}_3}{d\mathbf{f}^T} = [\mathbf{I}_{3^2} + 6\mathcal{D}(\text{vec } \mathbf{R}_1) + 3\mathcal{D}(\text{vec}(\mathbf{R}_1 \circ \mathbf{R}_1))] (\mathbf{Z}^T \otimes \mathbf{I}_3). \quad (148)$$

## References

- Blurton Jones N (2011) Hadza demography and sociobiology. <http://www.sscnet.ucla.edu/anthro/faculty/blurton-jones/hadza-part-1.pdf>
- Brown GR, Laland KN, Mulder MB (2009) Bateman's principles and human sex roles. *Trends Ecol Evol* 24:297–304
- Cam E, Aubry LM, Authier M (2016) The conundrum of heterogeneities in life history studies. *Trends Ecol Evol* 31:872–886
- Caswell H (1976) Community structure: a neutral model analysis. *Ecol Monogr* 46:327–354
- Caswell H (2001) Matrix population models: construction, analysis and interpretation, 2nd edn. Sunderland, Sinauer Associates
- Caswell H (2006) Applications of Markov chains in demography. In: MAM2006: Markov Anniversary Meeting. Bosen Books, Raleigh, North Carolina, pp 319–334
- Caswell H (2007) Sensitivity analysis of transient population dynamics. *Ecol Lett* 10:1–15
- Caswell H (2008) Perturbation analysis of nonlinear matrix population models. *Demogr Res* 18:59–116
- Caswell H (2009) Stage, age and individual stochasticity in demography. *Oikos* 118:1763–1782
- Caswell H (2011) Beyond  $r_0$ : demographic models for variability of lifetime reproductive output. *PloS ONE* 6:e20809
- Caswell H (2012) Matrix models and sensitivity analysis of populations classified by age and stage: a vec-permutation matrix approach. *Theor Ecol* 5:403–417
- Caswell H (2013) Sensitivity analysis of discrete Markov chains via matrix calculus. *Linear Algebra Appl* 438:1727–1745
- Caswell H (2014) A matrix approach to the statistics of longevity in heterogeneous frailty models. *Demogr Res* 31:553–592
- Caswell H, Kluge FA (2015) Demography and the statistics of lifetime economic transfers under individual stochasticity. *Demogr Res* 32:563–588
- Caswell H, Salguero-Gómez R (2013) Age, stage and senescence in plants. *J Ecol* 101:585–595
- Caswell H, Van Daalen SF (2016) A note on the vec operator applied to unbalanced block-structured matrices. *J Appl Math* 2016:1–3
- Caswell H, Zarulli V (2015) The statistics of health and longevity: a dynamic analysis of prevalence data. Page <http://paa2015.princeton.edu/uploads/153187> in Annual Meeting. Population Association of America
- Clutton-Brock TH (1988) Reproductive success: studies of individual variation in contrasting breeding systems. University of Chicago Press, Chicago
- Cochran ME, Ellner S (1992) Simple methods for calculating age-based life history parameters for stage-structured populations. *Ecol Monogr* 62:345–364
- Compadre Plant Matrix Database (2014) Max Planck Institute for Demographic Research (Germany). <http://www.compadre-db.org/>
- Courtial A, Pettay JE, Jokela M, Rotkirch A, Lummaa V (2012) Natural and sexual selection in a monogamous historical human population. *Proc Natl Acad Sci* 109:8044–8049
- Crow JF (1958) Some possibilities for measuring selection intensities in man. *Hum Biol* 30:1–13
- Cushing JM, Ackleh AS (2012) A net reproductive number for periodic matrix models. *J Biol Dyn* 6:166–188
- Cushing JM, Zhou Y (1994) The net reproductive value and stability in matrix population models. *Nat Resour Model* 8:297–333
- De Camino-Beck T, Lewis MA (2007) A new method for calculating net reproductive rate from graph reduction with applications to the control of invasive species. *Bull Math Biol* 69:1341–1354
- Eaton JW, Mayer AJ (1953) The social biology of very high fertility among the hutterites: the demography of a unique population. *Hum Biol* 25:206–264
- Gurven M, Kaplan H (2007) Longevity among hunter-gatherers: a cross-cultural examination. *Popul Dev Rev* 33:321–365
- Haridas CV, Tuljapurkar S (2005) Elasticities in variable environments: properties and implications. *Am Nat* 166:481–495
- Hartemink N, Missov TI, Caswell H (2017) Stochasticity, heterogeneity, and variance in longevity in human populations. *Theor Popul Biol* 114:107–116
- Henderson HV, Searle SR (1981) The vec-permutation matrix, the vec operator and Kronecker products: a review. *Linear Multilinear Algebra* 9:271–288
- Hill K, Hurtado A (1996) Aché life history: the ecology and demography of a foraging people. *Evolutionary foundations of human behavior series*, Aldine de Gruyter. <http://books.google.nl/books?id=759IQgAACAAJ>
- Horvitz CC, Tuljapurkar S (2008) Stage dynamics, period survival, and mortality plateaus. *Am Nat* 172:203–215
- Howard RA (1960) Dynamic programming and Markov processes. Wiley, New York
- Human Fertility Database (2013) Max Planck Institute for Demographic Research (Germany) and the Vienna Institute of Demography (Austria). [www.humanfertility.org](http://www.humanfertility.org)
- Human Mortality Database (2013) University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). [www.mortality.org](http://www.mortality.org)
- Jenouvrier S, Aubrey L, Barbraud C, Weimerskirch H, Caswell H (2016) Unobserved heterogeneity in vital rates reveals a diversity of life histories in an Antarctic seabird. In prep
- Jones AG (2009) On the opportunity for sexual selection, the Bateman gradient and the maximum intensity of sexual selection. *Evolution* 63:1673–1684
- Lamar WR, McGraw JB (2005) Evaluating the use of remotely sensed data in matrix population modeling for eastern hemlock (*Tsuga canadensis* L.) *For Ecol Manag* 212:50–64
- Magnus JR, Neudecker H (1979) The commutation matrix: some properties and applications. *Ann Stat* 7:381–394
- Matser A, Hartemink N, Heesterbeek H, Galvani A, Davis S (2009) Elasticity analysis in epidemiology: an application to tick-borne infections. *Ecol Lett* 12:1298–1305
- Moorad JA, Promislow DE, Smith KR, Wade MJ (2011) Mating system change reduces the strength of sexual selection in an American frontier population of the 19th century. *Evol Hum Behav* 32:147–155
- Newton I (1989) Lifetime reproductive success in birds. Academic Press, San Diego
- Puterman ML (1994) Markov decision processes: discrete dynamic stochastic programming. Wiley, New York
- Rényi A (1970) Probability theory. North-Holland, Amsterdam
- Rhodes EC (1940) Population mathematics. I. *J R Stat Soc* 103:61–89
- Robbins AM, Stoinski T, Fawcett K, Robbins MM (2011) Lifetime reproductive success of female mountain gorillas. *Am J Phys Anthropol* 146:582–593
- Salguero-Gómez R, Jones OR, Archer CR, Bein C, Buhr H, Farack C, Gottschalk F, Hartmann A, Henning A, Hoppe G, Romer G, Ruoff T, Sommer V, Wille J, Voigt J, Zeh S, Viereggs D, Buckley YM, Che-Castaldo J, Hodgson D, Scheuerlein A, Caswell H, Vaupel JW (2016) COMADRE: A global data base of animal demography. *J Anim Ecol* 85:371–384

- Salguero-Gómez R, Jones OR, Archer CR, Buckley YM, Che-Castaldo J, Caswell H, Hodgson D, Scheuerlein A, Conde DA, Brinks E, et al (2015) The COMPADRE Plant Matrix Database: an open online repository for plant demography. *J Ecol* 103:202–218
- Sheskin TJ (2010) Markov chains and decision processes for engineers and managers. CRC Press
- Steiner UK, Tuljapurkar S (2012) Neutral theory for life histories and individual variability in fitness components. *Proc Natl Acad Sci* 109:4684–4689
- Steiner UK, Tuljapurkar S, Orzack SH (2010) Dynamic heterogeneity and life history variability in the kittiwake. *J Anim Ecol* 79: 436–444
- Tuljapurkar S, Horvitz CC (2006) From stage to age in variable environments: life expectancy and survivorship. *Ecology* 87:1497–1509
- Tuljapurkar S, Horvitz CC, Pascarella JB (2003) The many growth rates and elasticities of populations in random environments. *Am Nat* 162:489–502
- Tuljapurkar S, Steiner UK, Orzack SH (2009) Dynamic heterogeneity in life histories. *Ecol Lett* 12:93–106
- Van Daalen S, Caswell H (2015) Lifetime reproduction and the second demographic transition: stochasticity and individual variation. *Demogr Res* 33:561–588