RESEARCH ARTICLE

Ecology and Evolution
WILEY

Age-specific sensitivity analysis of stable, stochastic and transient growth for stage-classified populations

Stefano Giaimo [ID] | Arne Traulsen [ID]

Department of Evolutionary Theory, Max Planck Institute for Evolutionary Biology, Plön, Germany

Correspondence
Stefano Giaimo, Department of Evolutionary Theory, Max Planck Institute for Evolutionary Biology, August-Thienemann-Straße 2, 24306 Plön, Germany.
Email: giaimo@evolbio.mpg.de

Abstract

Sensitivity analysis in ecology and evolution is a valuable guide to rank demographic parameters depending on their relevance to population growth. Here, we propose a method to make the sensitivity analysis of population growth for matrix models solely classified by stage more fine-grained by considering the effect of age-specific parameters. The method applies to stable population growth, the stochastic growth rate, and transient growth. The method yields expressions for the sensitivity of stable population growth to age-specific survival and fecundity from which general properties are derived about the pattern of age-specific selective forces molding senescence in stage-classified populations.

KEYWORDS
aging, elasticity analysis, population growth, senescence, sensitivity analysis, stage, stochastic growth rate, transient growth

TAXONOMY CLASSIFICATION
Demography, Evolutionary ecology, Life history ecology, Theorectical ecology

1 | INTRODUCTION

Matrix models are a widely used tool for demographic analysis. They require individuals in the population to be grouped into separate classes and the individual dynamics through the classes to be quantified (Caswell, 2001). A basic application of matrix models is demographic projection, which consists in updating population abundances starting from a given initial state. A salient feature of these models is their being amenable to the sensitivity analysis of virtually any function, for example, the population growth rate, to the demographic parameters they contain (Caswell, 2019).

Age is a classic descriptor for individual classification. It is at the basis of the Leslie matrix model (Keyfitz & Caswell, 2005), where individuals are classified into separate age classes and their age-specific survival and fecundity are quantified. Stage is another

classic descriptor. A stage may be a size class (Calvo & Horvitz, 1990; Hoffmann, 1999; Pfeifer et al., 2006), a developmental phase (Kaneko & Takada, 2014), an occupied territory (Stephens et al., 2002), the position of the individual within a network (Giaimo et al., 2018), or any other individual state variable. More sophisticated demographic classifications are possible. For example, lately there has been a rise in the need of classifying individuals by both age and stage (Caswell, 2012; Caswell et al., 2018; Steiner et al., 2014) or through multiple descriptors of individual state (Roth & Caswell, 2016).

However, in matrix models purely based on stage, which are the focus of the present work, age is not absent. It merely is implicit (Cochran & Ellner, 1992). With each update of stage abundances, individuals either die or survive while getting one time unit older. Leveraging this ineluctable demographic fact, methods have been proposed to extract from stage-classified matrix models a variety of

age-related quantities. These include, for example, stage-specific life expectancy, average age at parenthood, or average age-specific fecundity (Caswell, 2001, ch. 5, 2009; Cochran & Ellner, 1992; Horvitz & Tuljapurkar, 2008; Tuljapurkar & Horvitz, 2006). The reason that age-from-stage methods exist is that they are a natural response to the need of analyzing along the "age" dimension the numerous matrix models that have been already constructed exclusively on the basis of stage and not of age. Moreover, matrix models only based on stage are arguably more prevalent than matrix models based on both age and stage, as age×stage models require more data for their parameterization and these additional data may not always be available or easy to collect. The presence of environmental variability may further exacerbate this unbalance as multiple matrix models need to be parameterized for the same population when the environment fluctuates.

A form of age-from-stage method is the sensitivity analysis of population growth for stage-classified populations with respect to age-specific perturbations. In a more general work on age×stage matrix models, Caswell (2012) proposed a way to perform this analysis for the stable growth rate of matrix models of populations classified by stage and living under constant environmental conditions. Caswell's method is based on the construction, from the stage-based matrix model, of another, larger matrix model that includes both age and stage classification to then perform age-specific sensitivity analysis of population growth using matrix calculus on the derived model. An application of this method is the quantification of selection gradients on age-specific mortality and fecundity from matrix models of populations classified solely by stage (Caswell, 2012; Caswell & Salguero-Gómez, 2013).

The general framework for the analysis of age×stage matrix models upon which Caswell's method for the analysis of stage-based models relies is very general and so is the matrix calculus upon which it is based (Caswell, 2019; Caswell et al., 2018). This framework covers way more than the sensitivity analysis of population growth. It is also possible to envisage that such framework could be extended to capture the sensitivity analysis of transient population growth and long-run population growth to age-specific parameters for stage-classified models under environmental stochasticity. Notably, an invitation to extend the framework for the analysis of age×stage matrix models to the case of changing environments is contained in Caswell et al. (2018, p. 581). This appears as a potentially important extension, as environmental fluctuations are a very common condition for most populations (Bernhardt et al., 2020). But, to the best of our knowledge, there has been so far no visible progress along the front of explicitly applying age-specific sensitivity analysis of the growth rate for stage-classified populations under environmental stochasticity.

In this work, we do note embark into the broad endeavor of extending the general analysis of age×stage matrix models. We do, however, go in the direction of expanding along the "age" dimension the scope of sensitivity analysis of population growth for matrix models solely classified by stage. In doing so, we hope to contribute to the mentioned need of analyzing the already available matrix models that happen to be based on stage only, and not on age. Here, we propose a method that is alternative to, and larger in scope than, that proposed by Caswell (2012) for the sensitivity analysis of population growth in models only classified by stage with respect to age-specific perturbations. Our method is alternative to that of Caswell (2012) because we do not require the separate construction and analysis of an age×stage matrix model (although a formal equivalence between the two methods is shown in the Appendix A.2). Our method is larger in scope because it directly applies to stable population growth in a constant environment as well as to transient and long-run population growth under environmental stochasticity. We propose two applications of our method to evidence what we believe are some of its advantages. The first application is the analysis of age-specific selective forces molding senescence in stage-classified populations. We show how some properties of these forces, which were made observable in Caswell and Salguero-Gómez (2013), can now be predicted analytically via our method. The second application is the decomposition of the elasticities of transient and stochastic growth into age-specific components. This application shows how our method can yield a fine-grained version of the sensitivity analysis of population growth for stage-classified models.

Finally, we would like to stress that our method for age-specific sensitivity analysis only applies to matrix models structured solely by stage. The method has no say on matrix models that are structured by both age and stage.

2 | SENSITIVITY ANALYSIS OF POPULATION GROWTH

Before starting with the analysis, we stipulate a convention. In the main text, we prefer to limit ourselves to account for the influence of changes in a single parameter on population growth. For this reason, we find convenient to adopt the formalism according to which the derivative $\partial \mathbf{A} / \partial \theta$ of a matrix $\mathbf{A} = (a_{i,j})$ with respect to a parameter θ is a matrix of the same dimensions as $\mathbf{A}$ and the (i, j) entry of $\partial \mathbf{A} / \partial \theta$ is $\partial a_{i,j} / \partial \theta$. Similarly, the derivative $\partial \mathbf{x} / \partial \theta$ of a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)^{\top}$, where the superscript $\top$ indicates transposition, is the vector $\partial \mathbf{x} / \partial \theta = (\partial x_1 / \partial \theta, \partial x_2 / \partial \theta, \ldots, \partial x_n / \partial \theta)^{\top}$. This is the formalism in Caswell (2001). However, it is sometimes useful to consider the effect on population growth of multiple parameters. While this can be done by iterating one-parameter formulas for each of the parameters, another possibility is to do all this at once. In Section A.7 of the Appendix, we generalize results from the main text to account for the effect on growth of changes in multiple age-specific parameters. Therein, we describe and use a different, more apt formalism for the derivative of a vector or a matrix with respect to a vector of parameters that follows the conventions of the matrix calculus introduced to ecology by Caswell (2007, 2009). We assume throughout that the matrix model denoted with $\mathbf{A}$, or $\mathbf{A}(t)$ in the time-varying case, is classified solely by stage.

2.1 | No age specificity

Here, we recall known results from age-independent sensitivity analysis of population growth. A column vector $\mathbf{n}(t)$ describes stage-specific abundances in the population at time t. The component i of this vector is the population abundance in stage i at time t. In one time unit, each individual in stage j contributes $a_{i,j}$ individuals to stage i. When the environment is constant, this contribution is constant too. The dynamics of $\mathbf{n}(t)$ are then determined by the projection matrix $\mathbf{A} = (a_{i,j})$, which is assumed to be classified by stage, as

$$\mathbf{n}(t+1) = \mathbf{A}\mathbf{n}(t). \tag{1}$$

At demographic stability, the total population size grows every time step by a factor corresponding to the dominant eigenvalue λ of $\mathbf{A}$ and the relative fraction in stage j remains constant and equal to the component w_j of the right dominant eigenvector $\mathbf{w}$ of $\mathbf{A}$. This eigenvector is assumed to be normalized so that its components add up to 1. The left dominant eigenvector $\mathbf{v}$ contains reproductive values, which capture the relative importance of individuals to future population growth depending on their stage.

Suppose $\mathbf{A}$ depends on some parameter θ with current value θ^*. A classic result by Caswell (1978) is that the sensitivity of stable population growth on the log scale, $\ln\lambda$, to this parameter is

$$\frac{\partial \ln\lambda}{\partial \theta} = \frac{\mathbf{v}^\top \frac{\partial \mathbf{A}}{\partial \theta} \mathbf{w}}{\lambda \mathbf{v}^\top \mathbf{w}}, \tag{2}$$

where the derivative should be understood as evaluated at $\theta = \theta^*$.

An analogous result holds when the environment can be in different states that affect the demography. Assume transitions among environmental states obey an ergodic time-homogeneous Markov chain and to each environmental state there corresponds a projection matrix. The update of stage-specific abundances then takes the form

$$\mathbf{n}(t+1) = \mathbf{A}(t)\mathbf{n}(t), \tag{3}$$

where $\mathbf{A}(t)$ is the projection matrix for the environmental state at t. As most authors, we assume that in the long run and independently of the initial stage distribution, population growth is characterized by the stochastic growth rate $\ln\lambda_s$ (which is valid under the assumptions discussed by Cohen (1977a, 1977b); Tuljapurkar (1990)),

$$\ln\lambda_s = \lim_{t \to \infty} \frac{1}{t} \ln\left[\frac{N(t)}{N(0)}\right], \tag{4}$$

where $N(t)$ is population size at t.

Tuljapurkar (1990) pioneered sensitivity analysis of the stochastic growth rate. He defined three main quantities: the time-specific stage distribution $\mathbf{w}(t)$, the time-specific reproductive value vector $\mathbf{v}(t)$, which are vectors scaled so that the components of each add up to 1 (for the computation of these quantities see Section A.3 of the Appendix), and the time-specific growth

$$\lambda_t = \frac{\mathbf{e}^\top \mathbf{n}(t+1)}{\mathbf{e}^\top \mathbf{n}(t)}, \tag{5}$$

where $\mathbf{e}$ is a vector of 1s. The quantity λ_t is the factor by which population size changes between t and $t+1$. Suppose that the matrices for demographic projection depend on some parameter θ. Expanding upon the original analysis of Tuljapurkar (1990), Caswell (2005) showed that the sensitivity of long-run population growth to a parameter θ (with current value θ^*) upon which some demographic rates depend is

$$\frac{\partial \ln\lambda_s}{\partial \theta} = \lim_{L \to \infty} \frac{1}{L} \sum_{t=0}^{L-1} \frac{\mathbf{v}^\top(t+1) \frac{\partial \mathbf{A}(t)}{\partial \theta} \mathbf{w}(t)}{\lambda_t \mathbf{v}^\top(t+1)\mathbf{w}(t+1)}, \tag{6}$$

which is a stochastic analog of Equation 2 and where the derivative should be understood as evaluated at $\theta = \theta^*$. The chief approach to estimate $\partial \ln\lambda_s / \partial \theta$ is via extensive stochastic simulations of the demographic process (Caswell, 2001; Morris & Doak, 2002).

Finally, time-specific growth (sometimes called the transient population growth rate) $\ln\lambda_t$ is susceptible of sensitivity analysis too. Combining and differentiating Equations 3 and 5, Caswell (2007, eq. 43) found that

$$\frac{\partial \ln\lambda_t}{\partial \theta} = \frac{\mathbf{e}^\top}{N(t+1)} \frac{\partial \mathbf{n}(t+1)}{\partial \theta} - \frac{\mathbf{e}^\top}{N(t)} \frac{\partial \mathbf{n}(t)}{\partial \theta}, \tag{7}$$

where the vectors $\partial \mathbf{n}(t) / \partial \theta$ are computed via the recursion

$$\frac{\partial \mathbf{n}(t+1)}{\partial \theta} = \frac{\partial \mathbf{A}(t)}{\partial \theta} \mathbf{n}(t) + \mathbf{A}(t) \frac{\partial \mathbf{n}(t)}{\partial \theta}, \tag{8}$$

from given initial vectors $\mathbf{n}(0)$ and $\partial \mathbf{n}(0) / \partial \theta$. Caswell (2007) gives information on how to set the latter vector. A particularly relevant case is when θ has no effect on the initial population and $\partial \mathbf{n}(0) / \partial \theta$ is the zero vector. The derivatives in Equations 7 and 8 should be understood as evaluated at θ's current value.

2.2 | Adding age specificity

Suppose we are given a stage-classified matrix model $\mathbf{A}$. In performing sensitivity analysis of the stable population growth for this matrix, Equation 2 presupposes that, when θ changes, all individuals with demographic rates that depend on θ will be equally affected regardless of the age of these individuals. This is because the entries of $\mathbf{A}$ express stage-specific quantities. Although the given model is classified by stage, changes in the population are theoretically conceivable so that they only affects individuals of age j (Caswell, 2012, p. 408). And we may want to find out the

resulting effect on population growth. Supposing that no additional age-specific data are available or can be collected for the target population, to analyze the age-specific sensitivity of population growth one can only try to squeeze information out of the given stage-classified model. To get the relevant information for this purpose, Caswell (2012) proposed the construction of an age×stage matrix model from the entries of the stage-classified model. Age-specific sensitivity analysis can then be performed on the obtained age×stage model and not anymore on the original stage-classified model. Here, we propose an alternative way of addressing the problem of getting age-specific sensitivity of population growth from models that were constructed using a stage-based classification. We deem our proposal alternative because it does not pass through the construction of an age×stage matrix model. (However, we can prove an essential equivalence with Caswell's approach, see Section A.2 of the Appendix.) Our proposal in essence consists in only slightly modifying the results of the age-independent sensitivity analysis reviewed in the previous section to make them age-specific.

We start from the case of a constant environment. To emphasize that we now analyze a scenario characterized by age-specific effects, we write the parameter as θ_j. Since $\mathbf{A}$ is assumed classified by stage, it is very hard to envisage how we could, if ever, introduce in it a perturbation parameter for age-specific effects. The matrix $\mathbf{A}$ governs the demography of individuals depending on their stage and independently of their age. Modifying entries of $\mathbf{A}$ implies modifying demographics in an age-independent fashion. To solve this problem, we propose the following construction. First, rewrite Equation 1 as

$$\mathbf{n}(t+1) = \sum_{k=1}^{\infty} \mathbf{A}_k \mathbf{n}_k(t),\qquad(9)$$

where $\mathbf{A}_k = \mathbf{A}$ for $k = 1, 2, \ldots$, while $\mathbf{n}_k(t)$ contains stage-specific abundances of individuals aged k, that is, the i component of this vector is the number of individuals in stage i and of age k at t and

$$\mathbf{n}(t) = \sum_{k=1}^{\infty} \mathbf{n}_k(t).$$

Equation 9 separately projects each age class using the same demographic rates. Using the construction in Equation 9, we can specifically target the demographic rates of individuals aged j by supposing that the matrix $\mathbf{A}_j$ depends on some parameter θ_j with current value θ_j^*. We then rewrite Equation 9 as

$$\mathbf{n}(t+1) = \sum_{\substack{k=1 \\ k \neq j}}^{\infty} \mathbf{A}_k \mathbf{n}_k(t) + \mathbf{A}_j\left(\theta_j^*\right)\mathbf{n}_j(t),\qquad(10)$$

where all matrices are still identical with one another but we have singled out $\mathbf{A}_j$ to highlight its dependence on a parameter upon which the other matrices do not depend.

Next, we note that Equation 2 makes the mechanism through which a change in θ affects stable population growth transparent. The quantity $a_{i,j}$ is the demographic contribution of an individual in stage j to stage i per time step. The (i, j)-entry of the matrix $\partial \mathbf{A} / \partial \theta$ is the sensitivity of $a_{i,j}$ to θ. In Equation 2, this sensitivity is weighted to contribute toward $\partial \ln \lambda / \partial \theta$. The weight is the product $v_i w_j$, which accounts both for the relative number w_j of individuals whose demographic rates are directly dependent on θ, that is, the individuals in stage j, and for the relevance, given by the reproductive value v_i, to population growth of a change that alters the demographic influx to stage i. Equation 10 and the logic behind Caswell's result suggest a straightforward way of getting an expression for the age-specific sensitivity of $\ln \lambda$. We should modify Equation 2 in two ways: we should substitute the matrix $\partial \mathbf{A} / \partial \theta$ of sensitivities of age-independent demographic rates with the matrix $\partial \mathbf{A}_j / \partial \theta_j$ of age-specific sensitivities and we should substitute the stable stage distribution $\mathbf{w}$ with the fraction $\mathbf{w}_j$ of the stable stage distribution aged j. The i component of $\mathbf{w}_j$ is the stable population fraction in stage i and of age j. Thus

$$\frac{\partial \ln \lambda}{\partial \theta_j} = \frac{\mathbf{v}^{\top} \frac{\partial \mathbf{A}_j}{\partial \theta_j} \mathbf{w}_j}{\lambda \mathbf{v}^{\top}\mathbf{w}},\qquad(11)$$

where, as in the corresponding age-independent formula, the derivative is evaluated at θ_j's current value. In this way, the sensitivity of $\ln \lambda$ specifically accounts for the relative size of that part of the population that has demographic rates that depend on θ_j. A formal derivation of Equation 11 is in Section A.1 of the Appendix. An essential equivalence with the approach of Caswell (2012) is shown in Section A.2 of the Appendix.

The reasoning leading to Equation 11 smoothly extends to the case of environmental variability. Equation 3 for projection under a changing environment can be decomposed into the projection of separate age classes as Equation 1. The only difference is that projection matrices and stage distribution now depend on time t. Accordingly, an age-specific version of the sensitivity formula in Equation 6 is

$$\frac{\partial \ln \lambda_s}{\partial \theta_j} = \lim_{L \to \infty} \frac{1}{L} \sum_{t=0}^{L-1} \frac{\mathbf{v}^{\top}(t+1)\frac{\partial \mathbf{A}_j(t)}{\partial \theta_j}\mathbf{w}_j(t)}{\lambda_t \mathbf{v}^{\top}(t+1)\mathbf{w}(t+1)},\qquad(12)$$

A more formal derivation of this equation is in Section A.3 of the Appendix.

Finally, the same strategy yields the age-specific sensitivity of time-specific (or transient) growth,

$$\frac{\partial \ln \lambda_t}{\partial \theta_j} = \frac{\mathbf{e}^{\top}}{N(t+1)}\frac{\partial \mathbf{n}(t+1)}{\partial \theta_j} - \frac{\mathbf{e}^{\top}}{N(t)}\frac{\partial \mathbf{n}(t)}{\partial \theta_j},\qquad(13)$$

where the vectors $\partial \mathbf{n}(t) / \partial \theta_j$ in Equation 13 are computed via the recursion

$$\frac{\partial \mathbf{n}(t+1)}{\partial \theta_j} = \frac{\partial \mathbf{A}_j(t)}{\partial \theta_j}\mathbf{n}_j(t) + \mathbf{A}(t)\frac{\partial \mathbf{n}(t)}{\partial \theta_j},\qquad(14)$$

from given initial vectors $\mathbf{n}(0)$ and $\partial \mathbf{n}(0) / \partial \theta_j$. A derivation of Equation 14 is in Section A.4 of the Appendix.

2.3 | Age-specific stage distribution

To obtain $\mathbf{w}_j(t)$ in Equation 12, we represent the projection matrix as a sum $\mathbf{A}(t) = \mathbf{U}(t) + \mathbf{F}(t)$, where $\mathbf{U}(t) = (u_{i,j}(t))$ is the transition matrix and $\mathbf{F}(t) = (f_{i,j}(t))$ is the fecundity matrix. The quantity $u_{i,j}(t)$ is the probability that an individual in stage j at t is in stage i at $t + 1$, while $f_{i,j}(t)$ is the number of new recruits in stage i at $t + 1$ per individual in stage j at t. Age 1 is assigned to new recruits at their first census. Age is updated at demographic projection so that an individual of age j at t is of age $j + 1$ at $t + 1$. Thus, the stage distribution aged j at time $t \geq j$ is

$$\mathbf{w}_j(t) = \frac{\overset{\mathbf{w}_1(t) = \lambda_{t-1}^{-1}\mathbf{F}(t-1)\mathbf{w}(t-1)}{\mathbf{U}(t-1)\mathbf{U}(t-2)\ldots\mathbf{U}(t-j+1)\mathbf{F}(t-j)\mathbf{w}(t-j)}}{\prod_{m=t-j}^{t-1}\lambda_m}, \quad j = 2, 3, \ldots \quad (15)$$

because $\lambda_{t-1}^{-1}\mathbf{F}(t-1)\mathbf{w}(t-1)$ is the fraction of the population at t of those aged 1 (new recruits), the matrix $\mathbf{U}(t-1)\mathbf{U}(t-2)\ldots\mathbf{U}(t-j+1)$ governs the survival and stage transitioning up to t and through age $j \geq 2$ of those who were new recruits at $t - j + 1$. When they reach age j at t, the population has grown by a factor $\lambda_{t-1}\lambda_t\ldots\lambda_{t-j+1}$ since their first census.

When $t < j$, we should consider individuals that were already present in the initial population and are still alive at t. However, we do not really require an initial age distribution within the stages to compute $\partial \ln \lambda_s / \partial \theta_j$. Recall the weak ergodic theorem in demography: two populations experiencing the same sequence of projection matrices converge to the same time-varying stage distribution even if they have different initial stage distributions (Cohen, 1979). A corollary of weak ergodicity is that the two populations also converge to the same age distribution even if their initial age distributions are different. In the long run, all initial individuals are dead and do not count directly anymore toward the age distribution. Moreover, the shared sequence of projection matrices and the convergence to a common stage distribution $\mathbf{w}(t)$ imply that both the production of new recruits $\mathbf{F}(t)\mathbf{w}(t)$ and their subsequent survival, via $\mathbf{U}(t+1)\mathbf{F}(t)\mathbf{w}(t)$, $\mathbf{U}(t+2)\mathbf{U}(t+1)\mathbf{F}(t)\mathbf{w}(t)$, $\ldots$, converge to the same levels in the two populations. The age distribution of a population is determined by the inflow of new individuals and their subsequent survival (Arthur, 1982). Hence, we expect the two populations to converge to the same time-varying age distribution. Appendix B reports a numerical example of this convergence.

We can then estimate $\partial \ln \lambda_s / \partial \theta_j$ from long stochastic simulations starting from arbitrary age and stage distributions. Even more conveniently, we can outright discard a sufficient number of initial iterations of the simulated process so that any influence of the initial age distribution is negligible. Truncation of simulation results is already customary in the sensitivity analysis of $\ln \lambda_s$, for example, in Tuljapurkar et al. (2003), to minimize transient effects due to the initial stage distribution.

When the environment is constant and the population demographically stable, the recursion for $\mathbf{w}_j$ in Equation 11 is

$$\mathbf{w}_{j+1} = \lambda^{-1}\mathbf{U}\mathbf{w}_j, \quad \text{with} \quad \mathbf{w}_1 = \lambda^{-1}\mathbf{F}\mathbf{w} \quad (16)$$

as Cochran and Ellner (1992) showed, where $\mathbf{w}$ is the stable stage distribution.

3 | APPLICATION 1—AGE-SPECIFIC SELECTION, STAGES AND SENESCENCE

Hamilton (1966) kick-started age-specific sensitivity analysis of population growth. Equating $\ln \lambda$ with fitness, he proved that, in age-classified stable populations, selection against mortality always declines with adult age, and so does selection on fecundity under mild assumptions (e.g., the population is not shrinking, $\lambda \geq 1$, and mortality is never zero). His results are key to understand the evolution of senescence, an age-related decline in biological functioning (Baudisch, 2005; Caswell & Shyu, 2017; Charlesworth, 1994; Giaimo & Traulsen, 2022b; Partridge & Barton, 1993). However, Hamilton's results are based on an age-classified model. To understand the stage-classified case, Caswell and Salguero-Gómez (2013) studied selection against mortality and on fecundity within stages for a $s \times s$ stage-classified matrix model $\mathbf{A}$ with dominant eigenvalue λ by constructing from this matrix an age×stage matrix model comprising ω age classes. In the stage-classified case, the selection force against mortality $\mu_{k,j}$ at age j within stage k is the sensitivity $\varphi(\mu_{k,j})$ of $\ln \lambda$ to a proportional change of the same magnitude in survival (possibly accompanied by stage transitioning) for all individuals aged j in stage k in the population. This amounts to a proportional change of all entries in column k of $\mathbf{U}_j = \mathbf{U}$, where $\mathbf{U}_j$ is the transition matrix for individuals aged j. In the stage-classified case, such matrix is equal to the age-independent transition matrix $\mathbf{U}$. The selection force on fecundity $m_{k,j}$ at age j within a reproductive stage k is the sensitivity $\varphi(m_{k,j})$ of $\ln \lambda$ to an additive change of the same magnitude in the fecundities of all individuals aged j in stage k in the population. This amounts to an additive change of the same magnitude in all positive entries of column k of $\mathbf{F}_j = \mathbf{F}$, where $\mathbf{F}_j$ is the fecundity matrix for individuals aged j. In the stage-classified case, such matrix is equal to the age-independent transition matrix $\mathbf{F}$. The stage-independent quantities $\varphi(\mu_j)$ and $\varphi(m_j)$ are the sum of $\varphi(\mu_{k,j})$ and the sum of $\varphi(m_{k,j})$, respectively, over all stages $k = 1, \ldots, s$. The quantity $\varphi(\mu_j)$ is the selection force against mortality μ_j for all individuals aged j in the population regardless of their stage. The quantity $\varphi(m_j)$ is the selection force on fecundity m_j for all individuals aged j independently of their stage. Let $\boldsymbol{\varphi}_\mu(j) = (\varphi(\mu_{1j}), \varphi(\mu_{2j}), \ldots, \varphi(\mu_{sj}))$ and $\boldsymbol{\varphi}_m(j) = (\varphi(m_{1j}), \varphi(m_{2j}), \ldots, \varphi(m_{sj}))$. Caswell and Salguero-Gómez (2013, Eqs. 30-1) proposed the following formulas for these quantities:

$$\boldsymbol{\varphi}_\mu(j) = \frac{1}{\lambda \mathbf{v}^\mathsf{T}\mathbf{w}}(\mathbf{w}^\mathsf{T} \otimes \mathbf{v}^\mathsf{T})(\mathbf{I}_{s\omega} \otimes \mathbf{K}^\mathsf{T}\mathbb{D}_U\mathbf{K})(\mathbf{E}_{jj} \otimes \mathbf{K} \otimes \mathbf{I}_s)(\text{vec}(\mathbf{I}_\omega) \otimes \mathbf{I}_{s^2})(\mathbf{I}_s \otimes \mathbf{G})\mathcal{D}(\text{vec}(\mathbf{I}_s))(\mathbf{I}_s \otimes \mathbf{e}_s)\mathcal{D}(\boldsymbol{\sigma})$$

$$\boldsymbol{\varphi}_m(j) = \frac{1}{\lambda \mathbf{v}^\mathsf{T}\mathbf{w}}(\mathbf{w}^\mathsf{T} \otimes \mathbf{v}^\mathsf{T})(\mathbf{I}_{s\omega} \otimes \mathbf{K}^\mathsf{T}\mathbb{D}_F\mathbf{K})(\mathbf{E}_{jj} \otimes \mathbf{K} \otimes \mathbf{I}_s)(\text{vec}(\mathbf{I}_\omega) \otimes \mathbf{I}_{s^2})(\mathbf{I}_s \otimes \mathbf{z}).$$

$$(17)$$

In Equation 17, $\otimes$ is the Kronecker product, $\mathbf{I}_n$ is the $n \times n$ identity matrix, $\mathbf{K}$ is the $s\omega \times s\omega$ vec-permutation matrix—see box 2 of Caswell et al. (2018) for more details on this matrix—, $\mathbb{D}_U$ is a block diagonal matrix where each block is a $\omega \times \omega$ matrix with 1s in the whole subdiagonal and in the (ω, ω)-entry and zeros everywhere else, $\mathbb{D}_F$ is a block diagonal matrix where each block is a $\omega \times \omega$ matrix with 1s in the first row and zeros everywhere else, $\mathbf{E}_{jj}$ is a matrix with a 1 in the (j, j)-entry and zeros everywhere else, the $\text{vec}(\bullet)$ operator applies to a matrix and returns the columns of this stacked one on top of the other, $\mathbf{G}$ is the matrix $\mathbf{U}$ after normalizing each column so that its components add up to 1, $\mathbf{e}_s$ is a $s \times 1$ vector of 1s, D is the operator that from a vector argument returns a diagonal matrix, σ is the vector of column sums of $\mathbf{U}$ and $\mathbf{z}$ is a $s \times 1$ vector where component i is 0 if row i of $\mathbf{F}$ is zero and 1 otherwise. Caswell and Salguero-Gómez (2013) give more details on these expressions. They also noted that "[w]hile these expressions are impressive at first, they are easily evaluated in matrix-oriented languages such as MATLAB" (p. 594). Using Equation 17, they numerically explored the trajectories of selective forces over age within stages in a sample of several plant species and observed that within-stage selection can be contra-senescent, that is, increasing over some ages, and not only pro-senescent, that is, decreasing with age.

However, the remarkable observation by Caswell and Salguero-Gómez (2013) of contra-senescent selection has remained unexplained, as it is unclear what the determinants of contra-senescent selection may be and when we should expect contra- or pro-senescent selection. One way to get more insight would be to attack Equation 17 analytically. But we preferred to avoid facing the impressive expressions therein and we looked for an alternative approach based on our method. First of all, recall that the selection force against mortality μ_j at age j is the sensitivity of $\ln\lambda$ to a proportional change of the same magnitude in survival for all individuals aged j in the population regardless of their stage. Thus, we set a perturbed matrix as $\mathbf{A}_j = \mathbf{F}_j + (1 + \theta_j)\mathbf{U}_j$, where $\mathbf{U}_j$ is the transition matrix for age class j. This matrix is identical to the overall transition matrix $\mathbf{U}$ in the absence of perturbations ($\theta_j = 0$). From Equation 11, we then get

$$\varphi(\mu_j) = \frac{\partial \ln\lambda}{\partial \theta_j} = \frac{\mathbf{v}^\top \mathbf{U} \mathbf{w}_j}{\lambda \mathbf{v}^\top \mathbf{w}}. \tag{18}$$

Similarly, recall that the selection force on fecundity m_j at age j is the sensitivity $\varphi(m_j)$ of $\ln\lambda$ to an additive increase of the same magnitude in the fecundities of all individuals aged j independently of their stage. Thus, we set a perturbed matrix as $\mathbf{A}_j = \mathbf{U}_j + \mathbf{F}_j + \theta_j \text{sgn}(\mathbf{F}_j)$, where $\mathbf{F}_j$ is the fecundity matrix for age class j, which is identical to the overall fecundity matrix $\mathbf{F}$ in the absence of perturbations ($\theta_j = 0$), while $\text{sgn}(\mathbf{F}_j)$ is the matrix obtained by applying the signum function to $\mathbf{F}_j$ entry-wise, which returns a matrix with 1 in the (i, j)-entry whenever the corresponding entry of $\mathbf{F}_j$ is positive and zeros everywhere else (Section A.8 notes that alternative perturbation patterns are possible). From Equation 11, we then get

$$\varphi(m_j) = \frac{\partial \ln\lambda}{\partial \theta_j} = \frac{\mathbf{v}^\top \text{sgn}(\mathbf{F})\mathbf{w}_j}{\lambda \mathbf{v}^\top \mathbf{w}}. \tag{19}$$

Using Equations 18 and 19, we were able to derive analytic results that the present authors were not able to derive directly from Equation 17. To start putting our expressions at work, we employed Equation 16 and noted that death eventually completely erodes any newborn cohort. Therefore, $\mathbf{w}_j$ tends to the zero vector as $j \to \infty$, see Section A.5 in the Appendix for a proof. The selective forces in Equations 18 and 19 have the same asymptotic behavior,

$$\begin{aligned}\lim_{j\to\infty} \varphi(\mu_j) &= 0, \\ \lim_{j\to\infty} \varphi(m_j) &= 0,\end{aligned} \tag{20}$$

that is, selection vanishes at very late age in stage-classified populations. This is hardly surprising. But, to our knowledge, a mathematical derivation of this fact within the formalism of stage-classified populations was apparently missing in the literature and it might be seen as evidence of the analytic power of our method.

Section A.6 in the Appendix shows another result that can be obtained directly from Equation 18,

$$\varphi(\mu_j) > \varphi(\mu_{j+1}), \qquad \text{with } j \geq \alpha - 1 \tag{21}$$

where α is the earliest reproductive age. This expression indicates that the force of selection against mortality steadily declines with reproductive age. This exactly mirrors a key result by Hamilton (1966) for the age-classified case. No equivalent of Equation 21 exists for fecundity. Hence, the asymptotic behavior of $\varphi(m_j)$ does not rule out that, differently from Hamilton's original finding, selection on fecundity can increase over some ages before eventually waning (more on this below).

After having revisited some results of Hamilton (1966) for the stage-classified case, we use Equations 18 and 19 to better understand age-specific selection within stages. Since $\varphi(\mu_j)$ and $\varphi(m_j)$ have limit zero as $j \to \infty$, $\varphi(\mu_{j,k})$ and $\varphi(m_{j,k})$ must have the same asymptotic behavior. Evolutionarily, this means that selection within each stage has an overall pro-senescent pattern at late ages. As for contra-senescent selection, Section A.9 of the Appendix shows that

$$\varphi(\mu_{k,j}) = b_k w_{k,j} \quad \text{and} \quad \varphi(m_{k,j}) = c_k w_{k,j}, \tag{22}$$

where b_k and c_k are positive constants that depend on stage and not on age, whereas $w_{k,j}$ is the stable population fraction of age j and in stage k. Hence, within a stage, the age-trajectory of selection against mortality and the age-trajectory of selection on fecundity are both proportional to the stable age distribution within that stage.

At demographic stability, the overall age distribution monotonically decreases with age when the population is not going extinct because death progressively erodes any newborn cohort. The stable age distribution within a stage, instead, is not always monotonic, depends on details of the species life cycle and the shape of this distribution is not always obviously guessed (Boucher, 1997). However, some qualitative insights can be gained by reasoning upon the specific matrix model of one's interest. Take, for example, *Dipsacus sylvestris*, the wild teasel, as modeled by Caswell (2001, p. 60). Its life
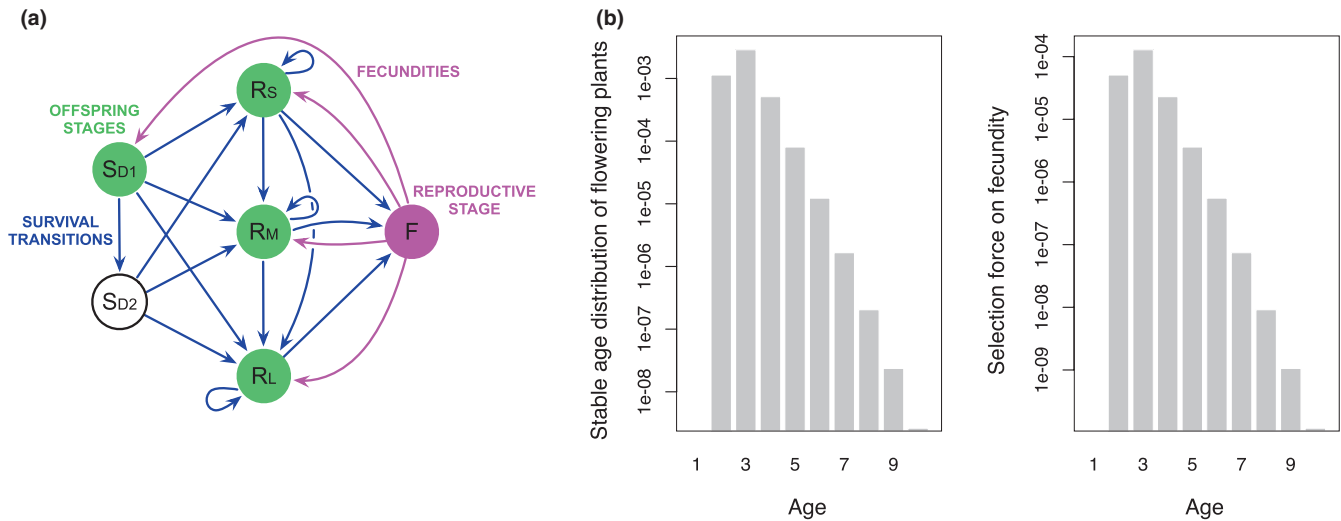
FIGURE 1 (a) Life cycle graph of *Dipsacus sylvestris*. Nodes represent stages. Blue arrows indicate survival transitions and correspond to entries of the **U** matrix. Magenta arrows indicate fecundities and correspond to entries of the **F** matrix. There is a single reproductive stage (magenta), Flowering plant, that can be accessed from any Rosette stage (small, medium and large). There are two Seed stages (dormant year 1 and dormant year 2). Offspring stages (green) are those where new recruits are first censused. (b) Stable age distribution of flowering plants and selection on their age-specific fecundity in *Dipsacus*. Theoretical considerations (see main text) suggest that: age distribution and selection are proportional to one another, flowering plants can only be aged ≥ 2 and the age distribution initially increases and eventually decreases with age. Data for this analysis are from Caswell (2001, p. 60). For this model, stable population growth is $\lambda = 2.33$ and individuals of every stage experience some nonzero mortality. The Supporting Information (Giaimo & Traulsen, 2022a) contains code to generate this panel.

cycle is in Figure 1. It takes a minimum of two demographic projections to go from offspring stages, that is, stages where new recruits are found at their first census, to the unique reproductive stage (flowering plants). The reproductive stage is accessible from multiple stages where stasis, that is, permanence in the same stage upon projection, is possible. In turn, these stages can be accessed via multiple pathways. Hence, the age distribution within the reproductive stage should start increasing only at age 2 being progressively fueled by the arrival to that stage of individuals becoming reproductive at different ages. By proportionality between age distribution within a stage and selection on age-specific fecundity within that stage, we expect selection on fecundity in *Dipsacus* to be initially contra-senescent (Figure 1). Since there is a single reproductive stage, this also shows that stage-independent selection on fecundity in this species can increase with adult age, in contrast to Hamilton's original finding.

Another key insight of Equation 22 is that, within a reproductive stage, selection against mortality and selection on fecundity are proportional to one another. In fact, Equation 22 immediately leads to

$$\frac{\varphi(\mu_{k,j})}{\varphi(m_{k,j})} = \frac{b_k}{c_k}, \tag{23}$$

where b_k / c_k is a quantity that depends on stage (k), yet not on age (j). This means that, within a stage, the magnitude of the ratio of selection against age-specific mortality to selection on age-specific fecundity remains constant independently of age. An illustration of this fact is in Figure 2, where we quantified selection within female stages of

Arisaema serratum (Thunb.) Schott (Araceae), a perennial herb, as modeled by Kinoshita (1987). We also quantified selection with the method of Caswell (2012), which Caswell and Salguero-Gómez (2013, Eqs. 30-1) applied to the same dataset. The exact match with our results serves as a validation of our method. It should be noted that, through Equation 17 by Caswell and Salguero-Gómez (2013), one can also plot the ratio of these two selective forces (Caswell, personal communication). However, as with the other results above, we were not able to prove as directly from Equation 17 the age independence of this ratio.

It should also be remarked that our findings about age-specific selective forces apply to the case of models only classified by stage. To derive these results, we have leveraged the fact that the same transition matrix and the same fecundity matrix apply to each age class in these models. In age×stage matrix models, which we do not study here, the transition matrix and the fecundity matrix typically differ between age classes so that our results do not obviously generalize to these models.

4 | APPLICATION 2—AGE-SPECIFIC ELASTICITIES

Elasticities are proportional sensitivities (Caswell, 2001). They are much used quantities in ecological analysis, for example (Coutts et al., 2016; Csergő et al., 2017; Kayal et al., 2018; Koons et al., 2016; Struckman et al., 2019; Tredennick et al., 2018). Here, we show how to decompose the elasticities of population growth for a stage-classified matrix model into their age-specific components. We shall
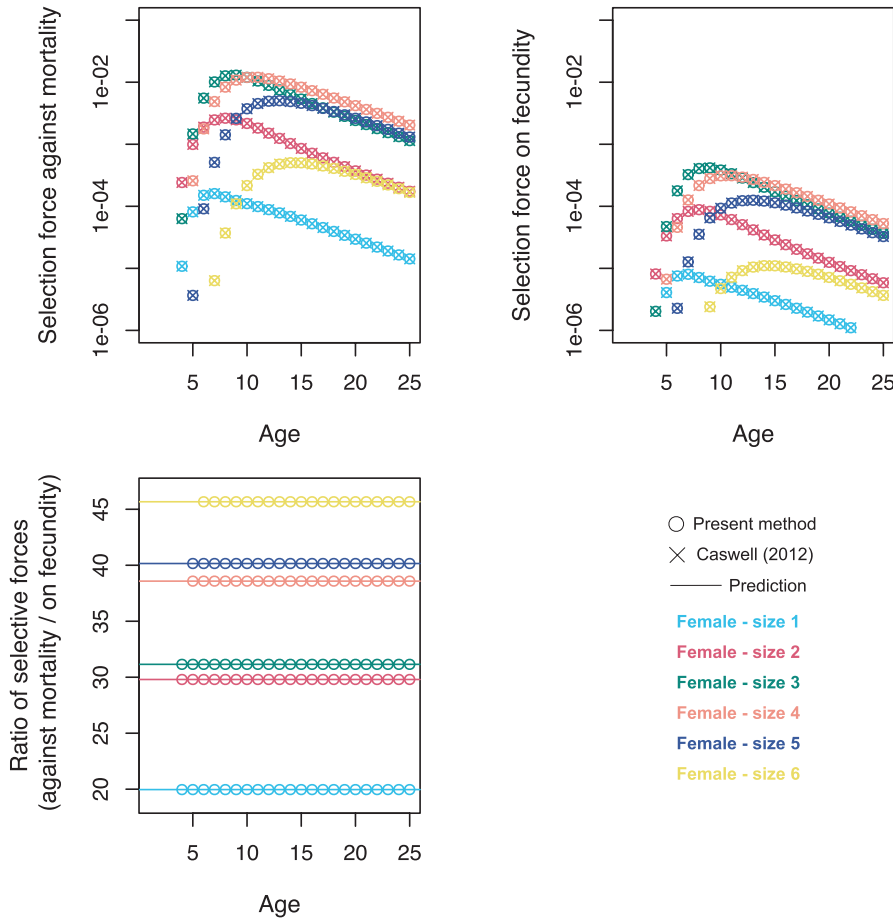
FIGURE 2 Age-specific selective forces within females stages of *Arisaema serratum*. These forces are computed using the method proposed here (circles) and a previous method (diagonal crosses) by Caswell (2012) that Caswell and Salguero-Gómez (2013, their fig. 3) employed to compute the same quantities. The methods are in perfect agreement. The distinctive analytical power of the proposed method, however, allows us to make a step further and predict (horizontal lines) from Equation 22 the existence of (and not only the theoretical possibility of visualizing) an age-independent ratio between the force of selection against mortality and the force of selection on fecundity at each age within a single stage. The Supporting Information (Giaimo & Traulsen, 2022a) contains code to generate this figure.

concentrate on elasticities of population growth to entries of the matrix model, and not to lower level parameters.

We are interested in the elasticity of λ to a subset of the entries of the stage-classified matrix model **A**. Let $\mathbf{A}_0$ be the matrix obtained from **A** by zeroing the entries that are not of our interest. Creating a perturbed matrix $\mathbf{A} + \theta\mathbf{A}_0$ with the unperturbed case given by $\theta = 0$ and using Equation 2, the relevant elasticity is

$$\varepsilon(\lambda, \mathbf{A}_0) = \frac{\mathbf{v}^\top \mathbf{A}_0 \mathbf{w}}{\lambda \mathbf{v}^\top \mathbf{w}}. \tag{24}$$

Using Equation 11, making this elasticity age-specific is very simple

$$\varepsilon_j(\lambda, \mathbf{A}_0) = \frac{\mathbf{v}^\top \mathbf{A}_0 \mathbf{w}_j}{\lambda \mathbf{v}^\top \mathbf{w}}, \tag{25}$$

This is the sensitivity of $\ln\lambda$ to the entries of **A** that are of our interest when these are proportionally perturbed only for individuals of age j.

By a similar reasoning as before applied to Equation 6, the age-independent elasticity of the stochastic growth rate to a subset of the entries of the matrix model is

$$\varepsilon(\lambda_s, \mathbf{A}_0) = \lim_{L \to \infty} \frac{1}{L} \sum_{t=0}^{L-1} \frac{\mathbf{v}^\top(t+1)\mathbf{A}_0(t)\mathbf{w}(t)}{\lambda_t \mathbf{v}^\top(t+1)\mathbf{w}(t+1)}. \tag{26}$$

It has proven important for some analyses to decompose this stochastic elasticity in different ways. Tuljapurkar et al. (2003) and Haridas and Tuljapurkar (2005) distinguished the elasticity of λ_s to the averages of the relevant entries of the matrix model

$$\varepsilon(\lambda_s, \overline{\mathbf{A}}_0) = \lim_{L \to \infty} \frac{1}{L} \sum_{t=0}^{L-1} \frac{\mathbf{v}^\top(t+1)\overline{\mathbf{A}}_0\mathbf{w}(t)}{\lambda_t \mathbf{v}^\top(t+1)\mathbf{w}(t+1)}, \tag{27}$$

from the elasticity of λ_s to the standard deviations of these entries

$$\varepsilon(\lambda_s, \mathbf{A}_0 - \overline{\mathbf{A}}_0) = \lim_{L \to \infty} \frac{1}{L} \sum_{t=0}^{L-1} \frac{\mathbf{v}^\top(t+1)(\mathbf{A}_0(t) - \overline{\mathbf{A}}_0)\mathbf{w}(t)}{\lambda_t \mathbf{v}^\top(t+1)\mathbf{w}(t+1)}, \tag{28}$$

where $\overline{\mathbf{A}}_0$ is the long-run average of $\mathbf{A}_0(t)$. Note that $\varepsilon(\lambda_s, \overline{\mathbf{A}}_0) + \varepsilon(\lambda_s, \mathbf{A}_0 - \overline{\mathbf{A}}_0) = \varepsilon(\lambda_s, \mathbf{A}_0)$, see Tuljapurkar et al. (2003). Hence, the elasticity in Equation 26 can be seen as an elasticity to both the average and the standard deviation of the entries of interest.

The other decomposition of Equation 26 that we will consider here is into environment-specific elasticities. Let ξ be a subset of the environmental states and let $\mathbf{1}_\xi(t)$ be an indicator function that takes value 1 at t when the environment is in a state in ξ and value 0 otherwise. The elasticity

$$\varepsilon(\lambda_s, \mathbf{1}_\xi \mathbf{A}_0) = \lim_{L \to \infty} \frac{1}{L} \sum_{t=0}^{L-1} \frac{\mathbf{v}^\top(t+1)\mathbf{1}_\xi(t)\mathbf{A}_0(t)\mathbf{w}(t)}{\lambda_t \mathbf{v}^\top(t+1)\mathbf{w}(t+1)}, \tag{29}$$

is the sensitivity of $\ln\lambda_s$ to a proportional perturbation of the entries of our interest in the matrix model only when the environment is in a state in ξ (Åberg et al., 2009; Caswell, 2005; Horvitz et al., 2005). When all environmental states are in ξ, the elasticity in Equation 29 is equal to the environment-independent elasticity in Equation 26. When ξ contains a single environmental state, we refer to the resulting elasticity as the elasticity of λ_s to that environment. However, note that this name should not suggest in any way a change in the transition probabilities of the (assumed constant) environmental Markov chain. For perturbations of this chain, see Steinsaltz et al. (2011).

Making elasticities of the stochastic growth rate age-dependent is just as easy as it was for the stable growth rate. Using Equation 12, all it takes is to replace $\mathbf{w}(t)$ in the numerator of Equations 26–29 with $\mathbf{w}_j(t)$,

$$\varepsilon_j(\lambda_s, \mathbf{A}_0) = \lim_{L \to \infty} \frac{1}{L} \sum_{t=0}^{L-1} \frac{\mathbf{v}^{\top}(t+1)\mathbf{A}_0(t)\mathbf{w}_j(t)}{\lambda_t \mathbf{v}^{\top}(t+1)\mathbf{w}(t+1)} \quad (30a)$$

$$\varepsilon_j(\lambda_s, \overline{\mathbf{A}}_0) = \lim_{L \to \infty} \frac{1}{L} \sum_{t=0}^{L-1} \frac{\mathbf{v}^{\top}(t+1)\overline{\mathbf{A}}_0\mathbf{w}_j(t)}{\lambda_t \mathbf{v}^{\top}(t+1)\mathbf{w}(t+1)} \quad (30b)$$

$$\varepsilon_j(\lambda_s, \mathbf{A}_0 - \overline{\mathbf{A}}_0) = \lim_{L \to \infty} \frac{1}{L} \sum_{t=0}^{L-1} \frac{\mathbf{v}^{\top}(t+1)(\mathbf{A}_0(t) - \overline{\mathbf{A}}_0)\mathbf{w}_j(t)}{\lambda_t \mathbf{v}^{\top}(t+1)\mathbf{w}(t+1)} \quad (30c)$$

$$\varepsilon_j(\lambda_s, \mathbf{1}_\xi \mathbf{A}_0) = \lim_{L \to \infty} \frac{1}{L} \sum_{t=0}^{L-1} \frac{\mathbf{v}^{\top}(t+1)\mathbf{1}_\xi(t)\mathbf{A}_0(t)\mathbf{w}_j(t)}{\lambda_t \mathbf{v}^{\top}(t+1)\mathbf{w}(t+1)}. \quad (30d)$$

These elasticities have the same meaning as their age-independent counterparts except that they refer to perturbations that only affect individuals of age j.

Finally, we consider elasticities of transient growth. From Equations 7 and 8, the elasticity of λ_t to a subset of the entries of $\mathbf{A}$ is

$$\varepsilon(\lambda_t, \mathbf{A}_0) = \frac{\mathbf{e}^{\top}}{N(t+1)}\frac{\partial \mathbf{n}(t+1)}{\partial \theta} - \frac{\mathbf{e}^{\top}}{N(t)}\frac{\partial \mathbf{n}(t)}{\partial \theta}, \quad \text{with} \quad \frac{\partial \mathbf{n}(t+1)}{\partial \theta} = \mathbf{A}_0\mathbf{n}(t) + \mathbf{A}(t)\frac{\partial \mathbf{n}(t)}{\partial \theta}. \quad (31)$$

From Equations 13 and 14, the age-specific version of this quantity simply is

$$\varepsilon_j(\lambda_t, \mathbf{A}_0) = \frac{\mathbf{e}^{\top}}{N(t+1)}\frac{\partial \mathbf{n}(t+1)}{\partial \theta_j} - \frac{\mathbf{e}^{\top}}{N(t)}\frac{\partial \mathbf{n}(t)}{\partial \theta_j}, \quad \text{with} \quad \frac{\partial \mathbf{n}(t+1)}{\partial \theta_j} = \mathbf{A}_0\mathbf{n}_j(t) + \mathbf{A}(t)\frac{\partial \mathbf{n}(t)}{\partial \theta_j}. \quad (32)$$

Since the sum of $\mathbf{w}_j(t)$ over all ages j is $\mathbf{w}(t)$, all age-specific elasticities we have considered above decompose the corresponding age-independent elasticities
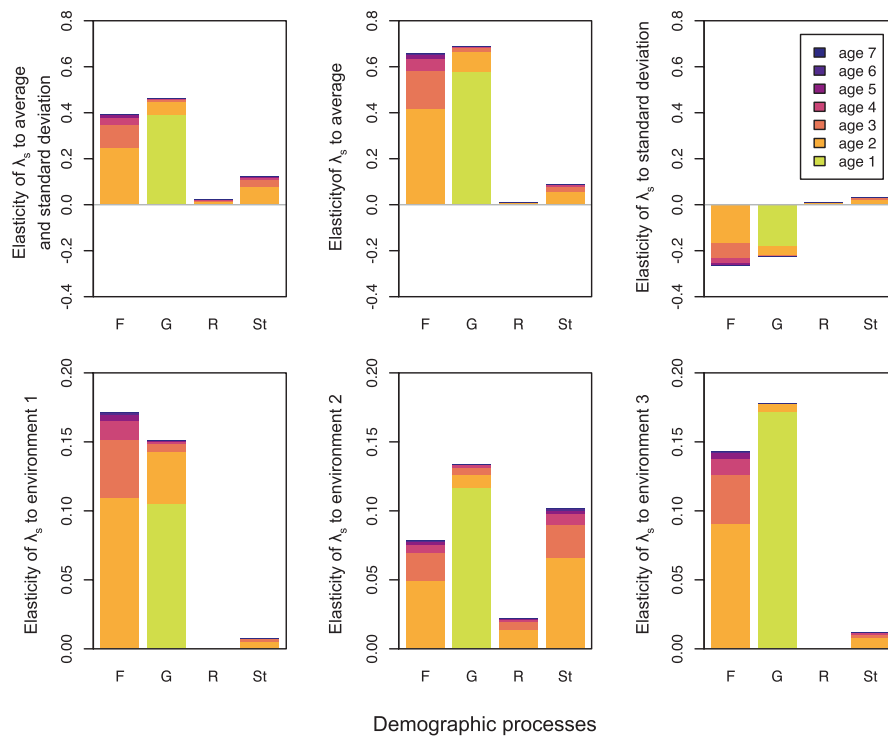
$$\sum_{j=1}^{\infty} \varepsilon_j = \varepsilon. \quad (33)$$

FIGURE 3 Age-specific elasticities of stochastic growth λ_s. We initiated a population with random stage distribution. We demographically updated the populations for 70,000 time-steps using a randomly generated sequence of projection matrices. We computed elasticities of λ_s to different demographic processes: fecundity (F), individual growth (G), retrogression (R), and stasis (St). Age-specific elasticities were computed for each case using Equation 30. Only age-specific elasticities up to age 7 are reported here, as elasticities for later ages contributed only minimally to the overall age-independent elasticity for this dataset. In computing elasticities, we removed the first and last 10,000 steps to minimize transient effects. The analysis is based on demographic data for population E of kidney vetch in Davison et al. (2010). The Supporting Information (Giaimo & Traulsen, 2022a) contains code to generate this figure.

Assuming, as it seems reasonable, that even under fluctuating environment, death stills erodes to the end any initial newborn cohort so that $\mathbf{w}_j(t) \to 0$ as $j \to \infty$, then Equations 30 and 32 imply that $\varepsilon_j \to 0$ as $j \to \infty$. Hence, we can approximate an age-independent elasticity ε by adding a sufficient number of its age-specific components ε_j.

To put age-specific stochastic elasticities at work, we took demographic data about one of the populations of *Anthyllis vulneraria*, the common kidney vetch, analyzed by Davison et al. (2010). In their model, there are three projection matrices corresponding to three environmental states assumed independent and identically distributed. Four stages were distinguished: seedling, vegetative adult, small flowering adult, and large flowering adult. Projection matrices were of the form

$$\mathbf{A}(t) = \begin{pmatrix} 0 & 0 & F_{1,3}(t) & F_{1,4}(t) \\ G_{2,1}(t) & St_{2,2}(t) & R_{2,3}(t) & R_{2,4}(t) \\ G_{3,1}(t) & G_{3,2}(t) & St_{3,3}(t) & R_{3,4}(t) \\ G_{4,1}(t) & G_{4,2}(t) & G_{4,3}(t) & St_{4,4}(t) \end{pmatrix}. \tag{34}$$

Entries in this matrix are grouped on the basis of the general demographic process they contribute to: fecundity (F), individual growth (G), retrogression (R), and stasis (St). To obtain the elasticity of λ_s to one of these processes, we set $\mathbf{A}_0(t)$ by zeroing the entries in Equation 34 that do not pertain to this process. Figure 3 reports stochastic elasticities to each demographic process. Elasticities are represented as sums of age-specific components. Figure 3 makes visible how, overall, individual growth at age 1 is the age-specific demographic process to which λ_s is the most elastic.

Using the same dataset, we computed age-specific elasticities of transient growth to each general demographic process (Figure 4). Age-independent elasticities of transient growth can take different signs at different time points (see elasticity to stasis, Figure 4). Our analysis shows that the elasticity of transient growth at a single time point can have age-specific components of different signs. For example, looking at time $t = 14$ in Figure 4, the elasticity to fecundity at age 2 is positive, while the elasticity to fecundity at age 3 is negative.
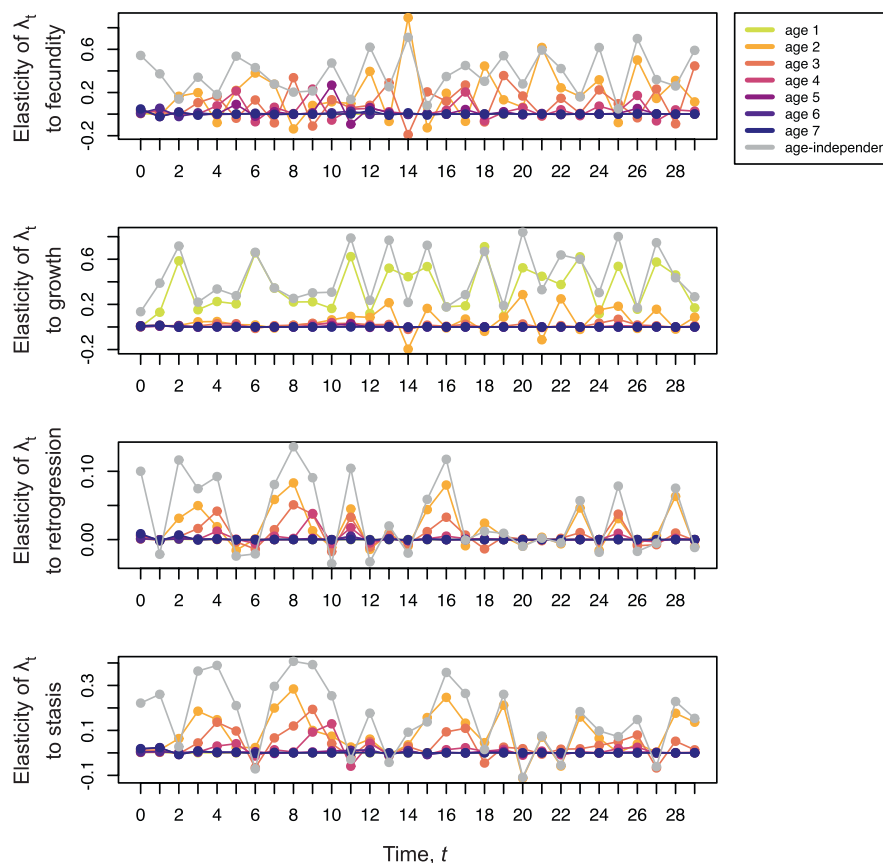

FIGURE 4 Age-specific elasticities of transient growth λ_t. We initiated a population with uniform stage distribution and random age distribution with maximum initial age of 30. We demographically updated the populations for 30 time-steps using a randomly generated sequence of projection matrices. We computed elasticities of λ_t to different demographic processes: fecundity (F), individual growth (G), retrogression (R), and stasis (St). Age-specific elasticities were computed for each case using Equation 32 with initial vectors $\partial \mathbf{n}(0) / \partial \theta$ set equal to the zero vector. Only age-specific elasticities up to age 7 are reported here, as elasticities for later ages contributed only minimally to the overall age-independent elasticity for $t \geq 6$ in these simulations. Note that elasticities are on different scales. The analysis is based on demographic data for population E of kidney vetch in Davison et al. (2010). The Supporting Information (Giaimo & Traulsen, 2022a) contains code to generate this figure.

5 | SUMMARY

The growth rate of a population is a quantity of primary interest in ecology and evolution. A well-established machinery rooted in matrix models exists to compute the sensitivity of population growth to any underlying demographic variable (Caswell, 2001, 2005, 2007, 2019; Caswell et al., 2018; Horvitz et al., 2005; Tuljapurkar et al., 2003). Here, we have proposed a method that, by only minimally modifying some results of such machinery, can make sensitivity analysis age-specific. The modification essentially consists in replacing the overall stage distribution that appears in sensitivity formulas for stage-classified populations with the age-specific stage distribution. The latter is easily computed by distinguishing transition/survival events from reproductive events in the matrix model. The method may be useful to those who are interested in analyzing along the "age" dimension the already existing matrix models that are classified solely by stage.

The proposed method applies to stable population growth, the stochastic growth rate and transient growth. In the age-specific sensitivity analysis of stable growth for populations solely classified by stage, our method offers a possible alternative to the established method that requires for this analysis the separate construction of an age×stage matrix model (Caswell, 2012; Caswell et al., 2018). This latter method led to observe contra- and pro-senescent selection within life stages (Caswell & Salguero-Gómez, 2013). We regard as an advantage of our method its ability to explain these observations by proving the existence of a proportionality between age-specific selective forces within a stage and the age distribution within that stage. In the case of stochastic population growth, to our knowledge there is no other method that has been applied explicitly yet for the age-specific sensitivity analysis of matrix models only classified by stage.

More generally, our method contributes to studies in ecology and evolution by decomposing usual quantities in the sensitivity analysis of population growth, like elasticities, into their age-specific components. Thus, the proposed method makes sensitivity analysis more fine-grained.

OPEN RESEARCH BADGES

This article has earned Open Data and Open Materials badges. Data and materials are available at https://doi.org/10.5061/dryad.6m905qg3m.

ORCID

Stefano Giaimo https://orcid.org/0000-0003-0421-3065
Arne Traulsen https://orcid.org/0000-0002-0669-5267

REFERENCES

Åberg, P., Svensson, C. J., Caswell, H., & Pavia, H. (2009). Environment-specific elasticity and sensitivity analysis of the stochastic growth rate. *Ecological Modelling*, *220*(5), 605–610.

Arthur, W. B. (1982). The ergodic theorems of demography: A simple proof. *Demography*, *19*(4), 439–445.

Baudisch, A. (2005). Hamilton's indicators of the force of selection. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 8263–8268.

Bernhardt, J. R., O'Connor, M. I., Sunday, J. M., & Gonzalez, A. (2020). Life in fluctuating environments. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *375*(1814), 20190454.

Boucher, D. H. (1997). General patterns of age-by-stage distributions. *Journal of Ecology*, *85*(2), 235–240.

Calvo, R. N., & Horvitz, C. C. (1990). Pollinator limitation, cost of reproduction, and fitness in plants: A transition-matrix demographic approach. *The American Naturalist*, *136*(4), 499–516.

Caswell, H. (1978). A general formula for the sensitivity of population growth rate to changes in life history parameters. *Theoretical Population Biology*, *14*(2), 215–230.

Caswell, H. (2001). *Matrix population models* (2nd ed.). Sinauer Associates.

Caswell, H. (2005). Sensitivity analysis of the stochastic growth rate: Three extensions. *Australian & New Zealand Journal of Statistics*, *47*(1), 75–85.

Caswell, H. (2007). Sensitivity analysis of transient population dynamics. *Ecology Letters*, *10*(1), 1–15.

Caswell, H. (2009). Stage, age and individual stochasticity in demography. *Oikos*, *118*(12), 1763–1782.

Caswell, H. (2012). Matrix models and sensitivity analysis of populations classified by age and stage: A vec-permutation matrix approach. *Theoretical Ecology*, *5*, 403–417.

Caswell, H. (2019). *Sensitivity analysis: Matrix methods in demography and ecology*. Springer.

Caswell, H., de Vries, C., Hartemink, N., Roth, G., & van Daalen, S. (2018). Age×stage-classified demographic analysis: A comprehensive approach. *Ecological Monographs*, *88*, 560–584.

Caswell, H., & Salguero-Gómez, R. (2013). Age, stage and senescence in plants. *Journal of Ecology*, *101*(3), 585–595.

Caswell, H., & Shyu, E. (2017). Senescence, selection gradients, and mortality. In *The evolution of senescence in the tree of life* (pp. 56–82). Cambridge University Press.

Charlesworth, B. (1994). *Evolution in age-structured populations* (2nd ed.). Cambridge University Press.

Cochran, M. E., & Ellner, S. (1992). Simple methods for calculating age-based life history parameters for stage-structured populations. *Ecological Monographs*, 62(3), 345–364.

Cohen, J. E. (1977a). Ergodicity of age structure in populations with markovian vital rates. II. General states. *Advances in Applied Probability*, 9(1), 18–37.

Cohen, J. E. (1977b). Ergodicity of age structure in populations with markovian vital rates, III: Finite-state moments and growth rate; an illustration. *Advances in Applied Probability*, 9(3), 462–475.

Cohen, J. E. (1979). Ergodic theorems in demography. *American Mathematical Society*, 1(2), 275–295.

Coutts, S., Salguero-Gómez, R., Csergő, A., & Buckley, Y. (2016). Extrapolating demography with climate, proximity and phylogeny: Approach with caution. *Ecology Letters*, 19, 1429–1438.

Csergő, A., Salguero-Gómez, R., Broennimann, O., Coutts, S., Guisan, A., Angert, A., Welk, E., Stott, I., Enquist, B., McGill, B., Svenning, J.-C., Violle, C., & Buckley, Y. (2017). Less favourable climates constrain demographic strategies in plants. *Ecology Letters*, 20, 969–980.

Davison, R., Jacquemyn, H., Adriaens, D., Honnay, O., De Kroon, H., & Tuljapurkar, S. (2010). Demographic effects of extreme weather events on a short-lived calcareous grassland species: Stochastic life table response experiments. *Journal of Ecology*, 98, 255–267.

Giaimo, S., Arranz, J., & Traulsen, A. (2018). Invasion and effective size of graph-structured populations. *PLoS Computational Biology*, 14, e1006559.

Giaimo, S., & Traulsen, A. (2022a). Supporting Information – Age-specific sensitivity analysis of stable, stochastic and transient growth for stage-classified populations. *Dryad*, Dataset: https://doi.org/10.5061/dryad.6m905qg3m

Giaimo, S., & Traulsen, A. (2022b). The selection force weakens with age because ageing evolves and not vice versa. *Nature Communications*, 13, 686.

Hamilton, W. D. (1966). The moulding of senescence by natural selection. *Journal of Theoretical Biology*, 12, 12–45.

Haridas, C. V., & Tuljapurkar, S. (2005). Elasticities in variable environments: Properties and implications. *The American Naturalist*, 166(4), 481–495 PMID: 16224704.

Hoffmann, W. (1999). Fire and population dynamics of woody plants in a neotropical savanna: Matrix model projections. *Ecology*, 80, 1354–1369.

Horn, R. A., & Johnson, C. R. (2013). *Matrix analysis*. Cambridge University Press.

Horvitz, C. C., & Tuljapurkar, S. (2008). Stage dynamics, period survival, and mortality plateaus. *The American Naturalist*, 172(2), 203–215.

Horvitz, C. C., Tuljapurkar, S., & Pascarella, J. B. (2005). Plant-animal interactions in random environments: Habitat-stage elasticity, seed predators, and hurricanes. *Ecology*, 86(12), 3312–3322.

Kaneko, Y., & Takada, T. (2014). Pair-wise analyses of the effects of demographic processes, vital rates, and life stages on the spatio-temporal variation in the population dynamics of the riparian tree *Aesculus turbinata* blume. *Population Ecology*, 56, 161–173.

Kayal, M., Lenihan, H., Brooks, A., Holbrook, S., Schmitt, R., & Kendall, B. (2018). Predicting coral community recovery using multi-species population dynamics models. *Ecology Letters*, 21, 1790–1799.

Keyfitz, N., & Caswell, H. (2005). *Applied mathematical demography*. Springer.

Kinoshita, E. (1987). Sex change and population dynamics in *Arisaema* (araceae) I. *Arisaema serratum* (thunb.) schott. *Plant Species Biology*, 2, 15–28.

Koons, D., Iles, D., Schaub, M., & Caswell, H. (2016). A life-history perspective on the demographic drivers of structured population dynamics in changing environments. *Ecology Letters*, 19, 1023–1031.

Morris, W., & Doak, D. F. (2002). *Quantitative conservation biology*. Sinauer Associates.

Partridge, L., & Barton, N. H. (1993). Optimality, mutation and the evolution of ageing. *Nature*, 362(6418), 305–311. https://doi.org/10.1038/362305a0

Pfeifer, M., Wiegand, K., Heinrich, W., & Jetschke, G. (2006). Long-term demographic fluctuations in an orchid species driven by weather: Implications for conservation planning. *Journal of Applied Ecology*, 43, 313–324.

Roth, G., & Caswell, H. (2016). Hyperstate matrix models: Extending demographic state spaces to higher dimensions. *Methods in Ecology and Evolution*, 7, 1438–1450.

Steiner, U., Tuljapurkar, S., & Coulson, T. (2014). Generation time, net reproductive rate, and growth in stage-age-structured populations. *The American Naturalist*, 183, 771–783.

Steinsaltz, D., Tuljapurkar, S., & Horvitz, C. (2011). Derivatives of the stochastic growth rate. *Theoretical Population Biology*, 80(1), 1–15.

Stephens, P., Frey-roos, F., Arnold, W., & Sutherland, W. (2002). Model complexity and population predictions. The alpine marmot as a case study. *Journal of Animal Ecology*, 71, 343–361.

Struckman, S., Couture, J., LaMar, M., & Dalgleish, H. (2019). The demographic effects of functional traits: An integral projection model approach reveals population-level consequences of reproduction-defence trade-offs. *Ecology Letters*, 22, 1396–1406.

Tredennick, A., Teller, B., Adler, P., Hooker, G., & Ellner, S. (2018). Size-by-environment interactions: A neglected dimension of species' responses to environmental variation. *Ecology Letters*, 21, 1757–1770.

Tuljapurkar, S. (1990). *Population dynamics in variable environments*. Springer.

Tuljapurkar, S., & Horvitz, C. (2006). From stage to age in variable environments: Life expectancy and survivorship. *Ecology*, 87, 1497–1509.

Tuljapurkar, S., Horvitz, C. C., & Pascarella, J. B. (2003). The many growth rates and elasticities of populations in random environments. *The American Naturalist*, 162(4), 489–502.

APPENDIX A

MATHEMATICAL DERIVATIONS

A.1 | DERIVATION OF EQUATION 11

We start from the eigenvector equation that relates the stable stage distribution $\mathbf{w}$ (with components that add up to 1), the stable growth rate λ and the projection matrix $\mathbf{A}$ of the population:

$$\lambda \mathbf{w} = \mathbf{A}\mathbf{w}. \tag{A1}$$

We rewrite the right hand side of Equation A1 as a series

$$\mathbf{A}\mathbf{w} = \sum_{k=1}^{\infty} \mathbf{A}_k \mathbf{w}_k, \tag{A2}$$

where the component i of $\mathbf{w}_k$ is the stable population fraction that is both in stage i and of age k. Age 1 is assigned to new recruits at their first census. Age is updated at demographic projection so that an individual of age j at t is of age $j+1$ at $t+1$. Thus, $\mathbf{w}$ can be thought of as the row sums of an array with as many rows as stages and an infinite number of columns each representing one age. The sum over all entries of the array is 1. The age index in Equation A2 runs to infinity. To set a maximum age ω, it is sufficient to let all vectors $\mathbf{w}_j$ with $j > \omega$ be equal to the zero vector. In Equation A2, $\mathbf{A}_k = \mathbf{A}$ for all $k = 1, 2, \ldots$ so that $\mathbf{A}_k \mathbf{w}_k$ demographically projects individuals aged k over one time step.

In this scheme, individuals of each age are all subject to the same demographic rates. But suppose there is a parameter x_j on which the demographic rates of individuals of age j in the population depend. The parameter currently takes value x_j^*. Since the demography of the population depends on x_j, both the stable stage distribution $\mathbf{w}$ and the stable growth rate λ are functions of this parameter. As for demographic projection, only $\mathbf{A}_j$ is a function of x_j, while all other matrices $\mathbf{A}_k$ with $k \neq j$ in Equation A2 are not, as these matrices contain the demographic rates to project individuals of age different from j. However, note that there presently is no difference between these projection matrices so that $\mathbf{A}_j(x_j^*) = \mathbf{A}_k$ with $k \neq j$. To express the dependencies of λ, $\mathbf{w}$, and $\mathbf{A}_j$ from x_j we can rewrite Equation A1 using Equation A2

$$\lambda(x_j)\mathbf{w}(x_j) = \sum_{\substack{k=1 \\ k \neq j}}^{\infty} \mathbf{A}_k \mathbf{w}_k(x_j) + \mathbf{A}_j(x_j)\mathbf{w}_j(x_j). \tag{A3}$$

Note that, when $x_j = x_j^*$, the right hand side of Equation A3 converges to $\mathbf{A}\mathbf{w}$. We assume that $\sum_{k=1}^{\infty} \partial \mathbf{w}_j / \partial x_j$ uniformly converges on some interval containing x_j^*. Let us differentiate Equation A3 with respect to x_j

$$\frac{\partial \lambda(x_j)}{\partial x_j}\mathbf{w}(x_j) + \lambda(x_j)\frac{\partial \mathbf{w}(x_j)}{\partial x_j} = \sum_{\substack{k=1 \\ k \neq j}}^{\infty} \mathbf{A}_k \frac{\partial \mathbf{w}_k(x_j)}{\partial x_j} + \frac{\partial \mathbf{A}_j(x_j)}{\partial x_j}\mathbf{w}_j(x_j) + \mathbf{A}_j(x_j)\frac{\partial \mathbf{w}_j(x_j)}{\partial x_j}. \tag{A4}$$

Multiplying $\mathbf{v}$, the reproductive value vector, on the left of Equation A4,

$$\frac{\partial \lambda(x_j)}{\partial x_j}\mathbf{v}^\top(x_j)\mathbf{w}(x_j) + \lambda(x_j)\mathbf{v}^\top(x_j)\frac{\partial \mathbf{w}(x_j)}{\partial x_j} = \mathbf{v}^\top(x_j)\sum_{\substack{k=1 \\ k \neq j}}^{\infty} \mathbf{A}_k \frac{\partial \mathbf{w}_k(x_j)}{\partial x_j}$$
$$+ \mathbf{v}^\top(x_j)\frac{\partial \mathbf{A}_j(x_j)}{\partial x_j}\mathbf{w}_j(x_j) + \mathbf{v}^\top(x_j)\mathbf{A}_j(x_j)\frac{\partial \mathbf{w}_j(x_j)}{\partial x_j}. \tag{A5}$$

Evaluating Equation A5 at $x_j = x_j^*$,

$$\frac{\partial \lambda(x_j)}{\partial x_j}\bigg|_{x_j=x_j^*}\mathbf{v}^\top(x_j^*)\mathbf{w}(x_j^*) + \lambda(x_j^*)\mathbf{v}^\top(x_j^*)\frac{\partial \mathbf{w}(x_j)}{\partial x_j}\bigg|_{x_j=x_j^*} = \mathbf{v}^\top(x_j^*)\sum_{\substack{k=1 \\ k \neq j}}^{\infty} \mathbf{A}_k \frac{\partial \mathbf{w}_k(x_j)}{\partial x_j}\bigg|_{x_j=x_j^*}$$
$$+ \mathbf{v}^\top(x_j^*)\frac{\partial \mathbf{A}_j(x_j)}{\partial x_j}\bigg|_{x_j=x_j^*}\mathbf{w}_j(x_j^*)$$
$$+ \mathbf{v}^\top(x_j^*)\mathbf{A}_j(x_j^*)\frac{\partial \mathbf{w}_j(x_j)}{\partial x_j}\bigg|_{x_j=x_j^*}. \tag{A6}$$

Using the fact that the reproductive value vector is the left eigenvector of the population projection matrix when $x_j = x_j^*$, we have that

$$\lambda(x_j^*)\mathbf{v}^\top(x_j^*) = \mathbf{v}^\top(x_j^*)\mathbf{A}_j(x_j^*) = \mathbf{v}^\top(x_j^*)\mathbf{A}_k, \quad k \neq j \tag{A7}$$

Using Equation A7 into Equation A6, simplifying and rearranging leads to

$$\frac{\partial \lambda}{\partial x_j}\bigg|_{x_j=x_j^*} = \frac{\mathbf{v}^\top(x_j^*)\frac{\partial \mathbf{A}_j(x_j)}{\partial x_j}\big|_{x_j=x_j^*}\mathbf{w}_j(x_j^*)}{\mathbf{v}^\top(x_j^*)\mathbf{w}_j(x_j^*)}, \tag{A8}$$

dividing this equation by $\lambda(x_j^*)$ yields

$$\frac{\partial \ln\lambda}{\partial x_j}\bigg|_{x_j=x_j^*} = \frac{\mathbf{v}^\top(x_j^*)\frac{\partial \mathbf{A}_j(x_j)}{\partial x_j}\big|_{x_j=x_j^*}\mathbf{w}_j(x_j^*)}{\lambda(x_j^*)\mathbf{v}^\top(x_j^*)\mathbf{w}_j(x_j^*)}, \tag{A9}$$

which corresponds to Equation 11 in the main text.

A.2 | EQUIVALENCE WITH CASWELL (2012)

Here, we show that the approach of Caswell (2012) to the age-specific sensitivity analysis of stable population growth for stage-classified models and our approach are essentially equivalent. Given a $q \times q$, time-independent stage-classified model $\mathbf{A}$, Caswell (2012) suggests to decompose it into a transition matrix $\mathbf{U} = (u_{i,j})$ and a fecundity matrix $\mathbf{F} = (f_{i,j})$, where $u_{i,j}$ is the probability that an individual in stage j at t is alive and observed in stage i at $t+1$ and $f_{i,j}$ is the number of new recruits in the population observed in stage i at $t+1$ per individual in stage j at t, so that $\mathbf{A} = \mathbf{U} + \mathbf{F}$. Then, fixing a maximum number ω of age classes, one defines the age-specific matrices $\mathbf{U}_i = \mathbf{U}$ and $\mathbf{F}_i = \mathbf{F}$ for $i = 1, 2, \ldots, \omega$, where $\mathbf{U}_i + \mathbf{F}_i = \mathbf{A}_i$ with $\mathbf{A}_i$ defined as in Section 2.2 of the main text. Using these matrices and a construction approach of extremely high generality, Caswell (2012) builds the age $\times$ stage matrix $\tilde{\mathbf{A}}$ in blocks. It would seem that Caswell's construction leads to:

$$
\tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{F}_1 & \mathbf{F}_2 & \mathbf{F}_3 & \dots & \mathbf{F}_\omega \\ \mathbf{U}_1 & & & & \\ & \mathbf{U}_2 & & & \\ & & \ddots & & \\ & & & \mathbf{U}_{\omega-1} & \mathbf{U}_\omega \end{pmatrix}. \tag{A10}
$$

This matrix updates a population state vector $\tilde{\mathbf{n}}(t)$ of the form

$$
\tilde{\mathbf{n}}(t) = \begin{pmatrix} \tilde{\mathbf{n}}_1(t) \\ \tilde{\mathbf{n}}_2(t) \\ \vdots \\ \tilde{\mathbf{n}}_{\geq\omega}(t) \end{pmatrix}, \tag{A11}
$$

where $\tilde{\mathbf{n}}_k(t) = \left(\tilde{n}_{1,k}(t), \tilde{n}_{2,k}(t), \dots, \tilde{n}_{q,k}(t)\right)^\top$ with $k = 1, 2, \dots, \omega - 1$ is the vector containing the population abundance $\tilde{n}_{i,k}(t)$ in stage i and of age k at t, while the vector $\tilde{\mathbf{n}}_{\geq\omega}(t) = \left(\tilde{n}_{1,\geq\omega}(t), \tilde{n}_{2,\geq\omega}(t), \dots, \tilde{n}_{q,\geq\omega}(t)\right)^\top$ contains the population abundance $\tilde{n}_{i,\geq\omega}(t)$ in stage i and of age at least ω at t. Note, in particular, that the shape of $\tilde{\mathbf{A}}$ implies that new recruits at t are all in $\tilde{\mathbf{n}}_1(t)$ and that individuals in $\tilde{\mathbf{n}}_k(t)$ who survive (and possibly transition of stage) over one time step are all in $\tilde{\mathbf{n}}_{k+1}(t+1)$ for $k = 1, 2, \dots, \omega - 2$, those who are in $\tilde{\mathbf{n}}_{\omega-1}(t)$ and survive over one time step are all in $\tilde{\mathbf{n}}_{\geq\omega}(t+1)$ with the remaining part of those in $\tilde{\mathbf{n}}_{\geq\omega}(t+1)$ being composed of those who are in $\tilde{\mathbf{n}}_{\geq\omega}(t)$ and survive over one time step. This explains the sense in which the matrices $\mathbf{U}_i$ and $\mathbf{F}_i$ are age-specific. Caswell (2012) then proposes to get age-specific sensitivity information about the initial, stage-classified matrix $\mathbf{A}$ by performing the analysis on the age×stage matrix $\tilde{\mathbf{A}}$ instead. The assumption is that $\mathbf{A}$ and $\tilde{\mathbf{A}}$ have the same dominant eigenvalue. As Caswell (2012, p. 409) says, since "the vital rates do not depend on age, the dominant eigenvalues of $\mathbf{A}$ and $\tilde{\mathbf{A}}$ should be identical." Caswell (2012, p. 413) also notes that, in general, $\tilde{\mathbf{A}}$ may be reducible. "This means that one must ascertain that the eigenvalues and eigenvectors [of the age×stage model] under analysis correspond to initial conditions of interest" (Caswell, 2012, p. 413), because asymptotic dynamics may depend on initial conditions. However, Caswell (2012, pp. 413, 415–6) shows that it is both necessary and sufficient that the left eigenvector corresponding to the dominant eigenvalue of $\tilde{\mathbf{A}}$ is positive component-wise for population growth to be described by this eigenvalue regardless of initial conditions. Caswell (2012) appears to suggest that one should numerically check for the specific matrix model of one's interest whether the condition of positivity for the left dominant eigenvector of $\tilde{\mathbf{A}}$ is satisfied.

Before showing equivalence with our approach, we first expand on Caswell's approach. We assume that $\mathbf{A}$ is irreducible. By irreducibility and nonnegativity, the matrix $\mathbf{A}$ has a dominant eigenvalue λ, which is real, positive, with geometric multiplicity of 1 and greater in magnitude than all other eigenvalues of $\mathbf{A}$ (Horn & Johnson, 2013, Theorem. 8.4.4). Let $\mathbf{w}$ and $\mathbf{v}$ be the unique (up to a constant scalar) right and left eigenvectors of $\mathbf{A}$, respectively, corresponding to λ. By Perron–Frobenius theorem, these eigenvectors are both positive

component-wise (ibid.). Recall that the matrix $\tilde{\mathbf{A}}$ is nonnegative and possibly reducible. Let us form the equations $x\mathbf{y} = \tilde{\mathbf{A}}\mathbf{y}$ and $x\mathbf{z}^\top = \mathbf{z}^\top\tilde{\mathbf{A}}$ with unknowns x (scalar), $\mathbf{y}$ and $\mathbf{z}$ (vectors). Recall Equation A16 from the main text,

$$
\begin{cases} \mathbf{w}_1 = \lambda^{-1}\mathbf{F}\mathbf{w} \\ \mathbf{w}_{j+1} = \lambda^{-1}\mathbf{U}\mathbf{w}_j, \end{cases} \tag{A12}
$$

where $\mathbf{w}$ is the dominant right eigenvector of $\mathbf{A}$, that is, the stable stage distribution, $\mathbf{A} = \mathbf{U} + \mathbf{F}$ and $\mathbf{w}_j$ is the fraction of the stable stage distribution aged j so that $\sum_{i=1}^\infty \mathbf{w}_i = \mathbf{w}$. Define $\mathbf{w}_{\geq\omega} = \sum_{i=\omega}^\infty \mathbf{w}_j$ and observe that

$$
\mathbf{U}(\mathbf{w}_{\omega-1} + \mathbf{w}_{\geq\omega}) = \lambda\mathbf{w}_\omega + \mathbf{U}\sum_{i=\omega}^\infty \mathbf{w}_i = \lambda\mathbf{w}_\omega + \sum_{i=\omega}^\infty \mathbf{U}\mathbf{w}_i = \lambda\mathbf{w}_\omega + \sum_{i=\omega}^\infty \lambda\mathbf{w}_{i+1} = \lambda\left(\mathbf{w}_\omega + \sum_{i=\omega+1}^\infty \mathbf{w}_i\right) = \lambda\mathbf{w}_{\geq\omega}. \tag{A13}
$$

We can see that $x = \lambda$ and $\mathbf{y} = \left(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{\geq\omega}\right)^\top$ are solutions of the equation $x\mathbf{y} = \tilde{\mathbf{A}}\mathbf{y}$, because from Equations A10–A13, $\mathbf{U}_i = \mathbf{U}$ and $\mathbf{F}_i = \mathbf{F}$,

$$
\tilde{\mathbf{A}} \begin{pmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \mathbf{w}_3 \\ \vdots \\ \mathbf{w}_{\geq\omega} \end{pmatrix} = \begin{pmatrix} \mathbf{F}_1\mathbf{w}_1 + \mathbf{F}_2\mathbf{w}_2 + \dots + \mathbf{F}_\omega\mathbf{w}_{\geq\omega} \\ \mathbf{U}_1\mathbf{w}_1 \\ \mathbf{U}_2\mathbf{w}_2 \\ \vdots \\ \mathbf{U}_{\omega-1}\mathbf{w}_{\omega-1} + \mathbf{U}_\omega\mathbf{w}_{\geq\omega} \end{pmatrix} = \begin{pmatrix} \mathbf{F}(\mathbf{w}_1 + \mathbf{w}_2 + \dots + \mathbf{w}_{\geq\omega}) \\ \mathbf{U}\mathbf{w}_1 \\ \mathbf{U}\mathbf{w}_2 \\ \vdots \\ \mathbf{U}(\mathbf{w}_{\omega-1} + \mathbf{w}_{\geq\omega}) \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \mathbf{w}_3 \\ \vdots \\ \mathbf{w}_{\geq\omega} \end{pmatrix}. \tag{A14}
$$

Similarly, $x = \lambda$ and $\mathbf{z} = \left(\mathbf{v}^\top, \mathbf{v}^\top, \dots, \mathbf{v}^\top\right)^\top$ are solutions of the equation $\lambda\mathbf{z}^\top = \mathbf{z}^\top\tilde{\mathbf{A}}$, because $\mathbf{v}$ is the left eigenvector corresponding to the eigenvalue λ of $\mathbf{A} = \mathbf{U} + \mathbf{F}$ (with $\mathbf{U} = \mathbf{U}_i$ and $\mathbf{F} = \mathbf{F}_i$) and, therefore, using Equation A10

$$
\begin{aligned} (\mathbf{v}^\top, \mathbf{v}^\top, \dots, \mathbf{v}^\top)\tilde{\mathbf{A}} &= (\mathbf{v}^\top(\mathbf{F}_1 + \mathbf{U}_1), \mathbf{v}^\top(\mathbf{F}_2 + \mathbf{U}_2), \dots, \mathbf{v}^\top(\mathbf{F}_\omega + \mathbf{U}_\omega)) \\ &= (\mathbf{v}^\top\mathbf{A}, \mathbf{v}^\top\mathbf{A}, \dots, \mathbf{v}^\top\mathbf{A}) \\ &= \lambda(\mathbf{v}^\top, \mathbf{v}^\top, \dots, \mathbf{v}^\top). \end{aligned} \tag{A15}
$$

Using these results and the fact that $\mathbf{v}$ is positive component-wise, Theorem 8.3.4 in Horn and Johnson (2013) implies that λ is a dominant eigenvalue of $\tilde{\mathbf{A}}$ and a corresponding left eigenvector is component-wise positive. Letting then $\mathbf{I}_{q\omega}$ be the $q\omega \times q\omega$ identity matrix, we can see from Equation A10 that $\lambda\mathbf{I}_{q\omega} - \tilde{\mathbf{A}}$ has at least some positive principal minor, for example, its $(q+1, q+1)$-entry. Therefore, by Problem 8.3.P14 in Horn and Johnson (2013), the eigenvalue λ of $\tilde{\mathbf{A}}$ has geometric multiplicity at most 1, that is, its left and right eigenvectors are unique. We can then strengthen the informal considerations and results of Caswell (2012) by concluding that the irreducibility of $\mathbf{A}$ and the construction of $\tilde{\mathbf{A}}$ from it are sufficient both to identify the dominant eigenvalue λ of $\mathbf{A}$ with the dominant eigenvalue of $\tilde{\mathbf{A}}$ and to guarantee that this eigenvalue describes population growth when the population is updated via $\tilde{\mathbf{A}}$ independently of initial conditions, that is, if $\mathbf{A}$ is irreducible, there is no need to numerically check the positivity of the left dominant eigenvector of $\tilde{\mathbf{A}}$, as this positivity is already guaranteed by the irreducibility of $\mathbf{A}$.

We can now directly address the equivalence problem. Let us denote with $\tilde{\mathbf{w}}$ and $\tilde{\mathbf{v}}$, respectively, the right and left eigenvectors of $\tilde{\mathbf{A}}$ and note that, using the above results about these eigenvectors,

$$\tilde{\mathbf{v}}^\top \tilde{\mathbf{w}} = \sum_{i=1}^{\omega-1} \mathbf{v}^\top \mathbf{w}_i + \mathbf{v}^\top \mathbf{w}_{\geq\omega} = \sum_{i=1}^{\infty} \mathbf{v}^\top \mathbf{w}_i = \mathbf{v}^\top \sum_{i=1}^{\infty} \mathbf{w}_i = \mathbf{v}^\top \mathbf{w}. \quad (A16)$$

Suppose now that we are interested in the age-specific sensitivity of the natural logarithm of the eigenvalue λ of $\tilde{\mathbf{A}}$. Since the blocks $\mathbf{U}_j$ and $\mathbf{F}_j$ for $j < \omega$ contain the demographic rates specific to individuals of age j, we rewrite the matrix $\tilde{\mathbf{A}}$ as

$$\tilde{\mathbf{A}}(\theta_j) = \begin{pmatrix} \mathbf{F}_1 & \mathbf{F}_2 & \dots & \mathbf{F}_j(\theta_j) & \dots & \dots & \mathbf{F}_\omega \\ \mathbf{U}_1 & & & & & & \\ & \mathbf{U}_2 & & & & & \\ & & \ddots & & & & \\ & & & \mathbf{U}_j(\theta_j) & & & \\ & & & & \ddots & & \\ & & & & & \mathbf{U}_{\omega-1} & \mathbf{U}_\omega \end{pmatrix}, \quad (A17)$$

to highlight that only demographic parameters for individuals of age j depend on θ_j. Using (while somewhat abusing notation) the general result of Caswell (1978) in Equation 2 of the main text, Equation A16 and the results above about $\tilde{\mathbf{w}}$ and $\tilde{\mathbf{v}}$, we have

$$\frac{\partial \ln\lambda}{\partial \theta_j} = \frac{\tilde{\mathbf{v}}^\top \frac{\partial \tilde{\mathbf{A}}}{\partial \theta_j} \tilde{\mathbf{w}}}{\lambda \tilde{\mathbf{v}}^\top \tilde{\mathbf{w}}} = \frac{\mathbf{v}^\top \frac{\partial \mathbf{F}_j}{\partial \theta_j} \mathbf{w}_j}{\lambda \mathbf{v}^\top \mathbf{w}} + \frac{\mathbf{v}^\top \frac{\partial \mathbf{U}_j}{\partial \theta_j} \mathbf{w}_j}{\lambda \mathbf{v}^\top \mathbf{w}} = \frac{\mathbf{v}^\top \frac{\partial \mathbf{A}_j}{\partial \theta_j} \mathbf{w}_j}{\lambda \mathbf{v}^\top \mathbf{w}}, \quad (A18)$$

which coincides with Equation A11 in the main text. This proves the equivalence of the two approaches.

A.3 | DERIVATION OF EQUATION 12

Here, we derive Equation 12 in the main text by adapting the original derivation by Tuljapurkar (1990), with later extension by Caswell (2005), of Equation 6 in the main text. A derivation of Equation 6 that is more rigorous, yet more complex, than that in Tuljapurkar (1990) concerning limit interchanges is in Steinsaltz et al. (2011). For simplicity, we here adapt the original derivation of Tuljapurkar (1990) as presented in Caswell (2001, sec. 14.4.1) to obtain Equation 12.

We introduce four quantities. First, the time-specific stage distribution

$$\mathbf{w}(t+1) = \frac{\mathbf{A}(t)\mathbf{w}(t)}{\mathbf{e}^\top \mathbf{A}(t)\mathbf{w}(t)}, \quad (A19)$$

where $\mathbf{e}$ is a vector of 1s. Second, the time-specific growth

$$\lambda_t = \mathbf{e}^\top \mathbf{A}(t)\mathbf{w}(t). \quad (A20)$$

Third, the time-specific reproductive value vector

$$\mathbf{v}^\top(t) = \frac{\mathbf{v}^\top(t+1)\mathbf{A}(t)}{\mathbf{v}^\top(t+1)\mathbf{A}(t)\mathbf{e}}, \quad (A21)$$

Fourth, the time-specific backward growth of reproductive value

$$\eta_t = \mathbf{v}^\top(t)\mathbf{A}(t-1)\mathbf{e}. \quad (A22)$$

Then, we recall from Tuljapurkar (1990, Eq. 11.2.3) that the stochastic growth rate in Equation 4 in the main text can be expressed as

$$\ln\lambda_s = \lim_{L\to\infty} \frac{1}{L} \ln\left[\mathbf{v}^\top(L)\mathbf{A}(L-1)\mathbf{A}(L-2)\dots\mathbf{A}(0)\mathbf{w}(0)\right], \quad (A23)$$

where $\mathbf{w}(0)$ and $\mathbf{v}(L)$, that is, the initial stage distribution and the final reproductive value vector, are two independent vectors each having arbitrary nonnegative components that add up to 1. We let

$$\beta_L = \mathbf{v}^\top(L)\mathbf{A}(L-1)\mathbf{A}(L-2)\dots\mathbf{A}(0)\mathbf{w}(0), \quad (A24)$$

be the quantity within brackets in Equation A23.

Initially, we consider an age-independent perturbation. Repeatedly using Equations A19–A22, we write β_L by making explicit the contribution of projection of the population stage distribution at time t toward the computation of the stochastic growth rate

$$\beta_L = \left(\prod_{i=0}^{t-1}\lambda_i\right)\left(\prod_{i=t+2}^{L}\eta_i\right)\mathbf{v}^\top(t+1)\mathbf{A}(t)\mathbf{w}(t). \quad (A25)$$

We first consider the effect of a time-specific, age-independent change in β_L. In particular, we assume that the projection matrix $\mathbf{A}(t)$ containing demographic rates at time t depends on some parameter x_t with current value x_t^* so that we write this matrix as $\mathbf{A}(t) = \mathbf{A}(x_t^*, t)$. Note that $\beta(x_t)$ also is a function of x_t. Using Equation A25, the sensitivity of β_L to x_t at x_t^* is

$$\left.\frac{\partial \beta_L}{\partial x_t}\right|_{x_t=x_t^*} = \left(\prod_{i=0}^{t-1}\lambda_i\right)\left(\prod_{i=t+2}^{L}\eta_i\right)\mathbf{v}^\top(t+1)\left.\frac{\partial \mathbf{A}(x_t,t)}{\partial x_t}\right|_{x_t=x_t^*}\mathbf{w}(t) \quad (A26)$$

Suppose now that the change is no longer time specific and all matrices in β_L may depend on the parameter x. The sensitivity of β_L to x at x^* is the sum of time-specific sensitivities each having the same form as Equation A26. Thus,

$$\left.\frac{\partial \beta_L}{\partial x}\right|_{x=x^*} = \sum_{t=0}^{L-1}\left(\prod_{i=0}^{t-1}\lambda_i\right)\left(\prod_{i=t+2}^{L}\eta_i\right)\mathbf{v}^\top(t+1)\left.\frac{\partial \mathbf{A}(x,t)}{\partial x}\right|_{x=x^*}\mathbf{w}(t) \quad (A27)$$

Using Equations A19–A25 and A27, the sensitivity of $\ln\lambda_s$ to x at x^* then is

$$\begin{aligned}\left.\frac{\partial \ln\lambda_s}{\partial x}\right|_{x=x^*} &= \lim_{L\to\infty}\frac{1}{L\beta_L(x^*)}\sum_{t=0}^{L-1}\left(\prod_{i=0}^{t-1}\lambda_i\right)\left(\prod_{i=t+2}^{L}\eta_i\right)\mathbf{v}^\top(t+1)\left.\frac{\partial \mathbf{A}(x,t)}{\partial x}\right|_{x=x^*}\mathbf{w}(t) \\ &= \lim_{L\to\infty}\frac{1}{L}\sum_{t=0}^{L-1}\frac{\left(\prod_{i=0}^{t-1}\lambda_i\right)\left(\prod_{i=t+2}^{L}\eta_i\right)\mathbf{v}^\top(t+1)\frac{\partial \mathbf{A}(x,t)}{\partial x}\Big|_{x=x^*}\mathbf{w}(t)}{\left(\prod_{i=0}^{t-1}\lambda_i\right)\left(\prod_{i=t+2}^{L}\eta_i\right)\mathbf{v}^\top(t+1)\mathbf{A}(x^*,t)\mathbf{w}(t)} \\ &= \lim_{L\to\infty}\frac{1}{L}\sum_{t=0}^{L-1}\frac{\mathbf{v}^\top(t+1)\frac{\partial \mathbf{A}(x,t)}{\partial x}\Big|_{x=x^*}\mathbf{w}(t)}{\lambda_t\mathbf{v}^\top(t+1)\mathbf{w}(t+1)},\end{aligned} \quad (A28)$$

where we have assumed the validity of interchange of limits, that is, the time limit and the derivative with respect to x. Equation A28,

which is due to a variation of Caswell (2005) on the original result of Tuljapurkar (1990), corresponds to Equation 6 in the main text.

We now consider age- and time-specific changes in demographic parameters. We use the fact that Equation A25 singles out the role of demographic projection at time t in the computation of $\ln \lambda_s$ to proceed similarly to Section A.1 and write

$$\beta_L = \left(\prod_{i=0}^{t-1} \lambda_i \right) \left(\prod_{i=t+2}^{L} \eta_i \right) \mathbf{v}^\top(t+1) \sum_{k=1}^{M+t} \mathbf{A}_k(t)\mathbf{w}_k(t), \qquad \text{(A29)}$$

where M is the maximum age in the initial population, that is, at $t = 0$, and $\mathbf{A}_k(t) = \mathbf{A}(t)$, while the vector $\mathbf{w}_k(t)$ is the fraction of the stage distribution of age k at t so that

$$\sum_{k=1}^{M+t} \mathbf{w}_k(t) = \mathbf{w}(t). \qquad \text{(A30)}$$

Equation A29 makes conspicuous the demographic projection of each age class in the population at t. Similarly to the previous section and above, we model time- and age-specific dependency of demographic rates by assuming that the matrix $\mathbf{A}_j(t)$ depends on some parameter $x_{j,t}$ with current value $x_{j,t}^*$ so that $\mathbf{A}_j\left(x_{j,t}^*, t\right) = \mathbf{A}_k(t)$ for $k \neq j$. The sensitivity of β_L to $x_{j,t}$ at $x_{j,t}^*$ is

$$\left.\frac{\partial \beta_L}{\partial x_{j,t}}\right|_{x_{j,t}=x_{j,t}^*} = \left(\prod_{i=0}^{t-1} \lambda_i \right) \left(\prod_{i=t+2}^{L} \eta_i \right) \mathbf{v}^\top(t+1) \left.\frac{\partial \mathbf{A}_j(x_{j,t}, t)}{\partial x_{j,t}}\right|_{x_{j,t}=x_{j,t}^*} \mathbf{w}_j(t) \qquad \text{(A31)}$$

Proceeding as in Equations A27 and A28 to remove time specificity and look at the sensitivity of $\ln \lambda_s$ to x_j at x_j^*, we find that

$$\left.\frac{\partial \ln \lambda_s}{\partial x_j}\right|_{x_j=x_j^*} = \lim_{L \to \infty} \frac{1}{L} \sum_{t=0}^{L-1} \frac{\mathbf{v}^\top(t+1) \left.\frac{\partial \mathbf{A}_j(x_j,t)}{\partial x_j}\right|_{x_j=x_j^*} \mathbf{w}_j(t)}{\lambda_t \mathbf{v}^\top(t+1)\mathbf{w}(t+1)}, \qquad \text{(A32)}$$

which corresponds to Equation 12 in the main text.

A.4 | DERIVATION OF EQUATIONS 13 AND 14

As in the previous sections, let x_j be a parameter with current value x_j^* upon which the demographic rates of individuals aged j depend. We start by decomposing Equation 3 in the main text, where projection is via the time-dependent matrix $\mathbf{A}(t)$, as we did for Equation 1 in the main text to get to Equation 10 in the main text. Thus, we obtain

$$\mathbf{n}(t+1) = \sum_{\substack{k=1 \\ k \neq j}}^{M} \mathbf{A}_k(t)\mathbf{n}_k(t) + \mathbf{A}_j\left(t, x_j^*\right)\mathbf{n}_j(t). \qquad \text{(A33)}$$

where M is some maximum age. This equation expresses population projection for each age class at t. Recall that $\mathbf{A}_j\left(t, x_j^*\right) = \mathbf{A}_k(t) = \mathbf{A}(t)$ with $k \neq j$. However, Equation A33 shows that if the population fraction aged j at some point depends on x_j (right hand side), then

the whole population at the next time point (left hand side) depends on x_j as well. Hence, we can write Equation A33 more generally as

$$\mathbf{n}(t+1, x_j) = \sum_{\substack{k=1 \\ k \neq j}}^{M} \mathbf{A}_k(t)\mathbf{n}_k(t, x_j) + \mathbf{A}_j(t, x_j)\mathbf{n}_j(t, x_j). \qquad \text{(A34)}$$

Taking then the derivative with respect to x_j,

$$\frac{\partial \mathbf{n}(t+1, x_j)}{\partial x_j} = \sum_{\substack{k=1 \\ k \neq j}}^{M} \mathbf{A}_k(t)\frac{\partial \mathbf{n}_k(t, x_j)}{\partial x_j} + \mathbf{A}_j(t, x_j)\frac{\partial \mathbf{n}_j(t, x_j)}{\partial x_j} + \frac{\partial \mathbf{A}_j(t, x_j)}{\partial x_j}\mathbf{n}_j(t, x_j), \qquad \text{(A35)}$$

and evaluating this derivative at $x_j = x_j^*$, when $\mathbf{A}_j\left(t, x_j^*\right) = \mathbf{A}_k(t) = \mathbf{A}(t)$ with $k \neq j$

$$\begin{aligned}
\left.\frac{\partial \mathbf{n}(t+1, x_j)}{\partial x_j}\right|_{x_j=x_j^*} &= \sum_{k=1}^{M} \mathbf{A}_k(t) \left.\frac{\partial \mathbf{n}_k(t, x_j)}{\partial x_j}\right|_{x_j=x_j^*} + \left.\frac{\partial \mathbf{A}_j(t, x_j)}{\partial x_j}\right|_{x_j=x_j^*} \mathbf{n}_j\left(t, x_j^*\right) \\
&= \mathbf{A}(t) \sum_{k=1}^{M} \left.\frac{\partial \mathbf{n}_k(t, x_j)}{\partial x_j}\right|_{x_j=x_j^*} + \left.\frac{\partial \mathbf{A}_j(t, x_j)}{\partial x_j}\right|_{x_j=x_j^*} \mathbf{n}_j\left(t, x_j^*\right) \\
&= \mathbf{A}(t) \left.\frac{\partial \mathbf{n}(t, x_j)}{\partial x_j}\right|_{x_j=x_j^*} + \left.\frac{\partial \mathbf{A}_j(t, x_j)}{\partial x_j}\right|_{x_j=x_j^*} \mathbf{n}_j\left(t, x_j^*\right),
\end{aligned} \qquad \text{(A36)}$$

which corresponds to Equation 14 in the main text.

A.5 | COHORT EROSION BY DEATH

Under standard assumptions in demography about $\mathbf{A}$, we have that λ is the Perron root of this matrix and the corresponding eigenvectors $\mathbf{v}$ and $\mathbf{w}$ are entrywise positive. Hence, from Equation 2, $\partial \ln \lambda / \partial a_{i,j} = v_i w_j / \left(\mathbf{v}^\top \mathbf{w}\right) > 0$ so that λ is an increasing function of the entries of $\mathbf{A}$. Hence, we have that $\rho\left(\lambda^{-1}\mathbf{U}\right) < \rho\left(\lambda^{-1}\mathbf{A}\right) = 1$, where ρ is the spectral radius because $\mathbf{U} \leq \mathbf{A}$ entrywise with at least some strict inequality. Therefore, $\lambda^{-1}\mathbf{U}$ converges to the zero matrix. From Equation 16, $\mathbf{w}_{j+1} = \left(\lambda^{-1}\mathbf{U}\right)^j \mathbf{w}_1$ with $j \geq 1$ tends to the zero vector as j goes to infinity. We refer to this as the fact that death eventually completely erodes any newborn cohort.

A.6 | DERIVATION OF EQUATION 21

Using Equation 16 in the main text, we can rewrite Equation 18 in the main text as

$$\varphi(\mu_j) = \frac{\mathbf{v}^\top \mathbf{w}_{j+1}}{\mathbf{v}^\top \mathbf{w}}. \qquad \text{(A37)}$$

We then recall the left dominant eigenvector equation for reproductive values

$$\mathbf{v}^\top = \lambda^{-1}\mathbf{v}^\top \mathbf{A}. \qquad \text{(A38)}$$

Multiplying $\mathbf{w}_j$ on the right of both sides of Equation A38, using the decomposition $\mathbf{A} = \mathbf{U} + \mathbf{F}$ and Equation 16 from the main text

$$\mathbf{v}^\top \mathbf{w}_j = \lambda^{-1}\mathbf{v}^\top \mathbf{A}\mathbf{w}_j = \lambda^{-1}\mathbf{v}^\top \mathbf{U}\mathbf{w}_j + \lambda^{-1}\mathbf{v}^\top \mathbf{F}\mathbf{w}_j = \mathbf{v}^\top \mathbf{w}_{j+1} + \lambda^{-1}\mathbf{v}^\top \mathbf{F}\mathbf{w}_j. \qquad \text{(A39)}$$

Let α be the first age at which reproduction is possible. This is the shortest path length from any offspring stage to any reproductive stage in the graph of the transition matrix $\mathbf{U}$ (Cochran & Ellner, 1992). Offspring stages are those corresponding to the nonzero rows of the fecundity matrix $\mathbf{F}$, while reproductive stages are those corresponding to the nonzero columns of $\mathbf{F}$. All components of $\mathbf{w}_j$ corresponding to reproductive stages are equal to zero for $1 \leq j < \alpha$. Some of the components of $\mathbf{w}_j$ corresponding to reproductive stages are greater than zero for $j \geq \alpha$. Then, we have that $\lambda^{-1}\mathbf{v}^\top \mathbf{F}\mathbf{w}_j = 0$ for $j < \alpha$ and $\lambda^{-1}\mathbf{v}^\top \mathbf{F}\mathbf{w}_j > 0$ for $j \geq \alpha$. As a consequence, we can establish the following inequalities between the left-most side of Equation A39 and the first term in the right-most side of this equation

$$
\begin{aligned}
\mathbf{v}^\top \mathbf{w}_j &= \mathbf{v}^\top \mathbf{w}_{j+1}, & j < \alpha, \\
\mathbf{v}^\top \mathbf{w}_j &> \mathbf{v}^\top \mathbf{w}_{j+1}, & j \geq \alpha.
\end{aligned}
\tag{A40}
$$

Dividing through the second line of Equation A40 by $\mathbf{v}^\top \mathbf{w}$ and using Equation A37, we obtain Equation 21 in the main text.

A.7 | MORE THAN ONE AGE-SPECIFIC PARAMETER

In the main text, we have explored the sensitivity of population growth to a single age-specific parameter. Here, we give formulas for the sensitivity of population growth to multiple age-specific parameters. Suppose that demographic rates for individuals aged j are a function of a vector

$$
\theta(j) = \begin{pmatrix} \theta_1(j) \\ \theta_2(j) \\ \vdots \\ \theta_p(j) \end{pmatrix}
\tag{A41}
$$

of p age-specific parameters. We want to obtain the sensitivity of (the natural logarithm of) stable population growth λ, which is the dominant eigenvalue of the $m \times m$ stage-classified projection matrix $\mathbf{A}$, to this vector. We recall that $\mathbf{A}_j$ is the projection matrix for individuals of age j. We adopt the formalism of matrix calculus introduced to ecology by Caswell (2007, 2009), see there for more details. In this formalism, the $\mathrm{vec}(\bullet)$ operator takes a matrix as argument and returns a column vector containing the matrix columns stacked one above the other, the $\mathrm{vec}^\top(\bullet)$ operator takes a matrix as argument and returns the transpose of a column vector containing the matrix columns stacked one above the other and $\otimes$ is the Kronecker product. Adapting this formalism to our case, the generalization of Equation 11 in the main text to multiple age-specific parameters is

$$
\frac{\mathrm{d}\ln\lambda}{\mathrm{d}\theta^\top(j)} = \frac{\mathrm{d}\ln\lambda}{\mathrm{dvec}^\top(\mathbf{A}_j)}\frac{\mathrm{dvec}(\mathbf{A}_j)}{\mathrm{d}\theta^\top(j)} = \frac{\mathbf{w}_j^\top \otimes \mathbf{v}^\top}{\lambda \mathbf{v}^\top \mathbf{w}}\frac{\mathrm{dvec}(\mathbf{A}_j)}{\mathrm{d}\theta^\top(j)}
\tag{A42}
$$

where $\mathrm{d}\lambda / \mathrm{d}\theta^\top(j)$ is a $1 \times p$ vector with component i equal to the partial derivative of λ with respect to component i of $\theta(j)$, $\mathrm{d}\ln\lambda / \mathrm{dvec}^\top(\mathbf{A}_j)$ is a $1 \times m^2$ vector with $\partial\lambda / \partial(\mathbf{A})_{k,l}$ as component $k + (l-1)m$, while $\mathbf{w}_j$, $\mathbf{w}$ and $\mathbf{v}$ are, as in the main text, the stable fraction of the stage

distribution aged j (see Equation 16 in the main text), the stable stage distribution (right dominant eigenvector), and the reproductive value vector (left dominant eigenvector) of $\mathbf{A}$, respectively. The partial derivative of the (k,l)-entry of $\mathbf{A}_j$ with respect to component s of $\theta(j)$ is found as the $(k + ml - m, s)$ entry of the $m^2 \times p$ matrix $\mathrm{dvec}(\mathbf{A}_j) / \mathrm{d}\theta^\top(j)$.

Similarly, Equation 12 in the main text generalizes to

$$
\frac{\mathrm{d}\ln\lambda_s}{\mathrm{d}\theta^\top(j)} = \lim_{L \to \infty} \frac{1}{L} \sum_{t=0}^{L-1} \frac{\mathbf{w}_j^\top(t) \otimes \mathbf{v}^\top(t+1)}{\lambda_t \mathbf{v}^\top(t+1)\mathbf{w}(t+1)}\frac{\mathrm{dvec}(\mathbf{A}_j(t))}{\mathrm{d}\theta^\top(j)},
\tag{A43}
$$

while Equations 13 and 14 in the main text generalize to

$$
\frac{\mathrm{d}\ln\lambda_t}{\mathrm{d}\theta^\top(j)} = \frac{\mathbf{e}^\top}{N(t+1)}\frac{\partial\mathbf{n}(t+1)}{\partial\theta^\top(j)} - \frac{\mathbf{e}^\top}{N(t)}\frac{\partial\mathbf{n}(t)}{\partial\theta^\top(j)},
\tag{A44}
$$

where $\partial\mathbf{n}(t) / \partial\theta^\top(j)$ in Equation A44 is a $m \times p$ matrix with (k,l)-entry equal to $\partial n_k(t) / \partial\theta_l(j)$, which is computed via the recursion

$$
\frac{\mathrm{d}\mathbf{n}(t+1)}{\mathrm{d}\theta^\top(j)} = \left(\mathbf{n}_j^\top(t) \otimes \mathbf{I}_m\right)\frac{\mathrm{dvec}(\mathbf{A}_j(t))}{\mathrm{d}\theta^\top(j)} + \mathbf{A}(t)\frac{\mathrm{d}\mathbf{n}(t)}{\mathrm{d}\theta^\top(j)}.
\tag{A45}
$$

with $\mathbf{I}_m$ the $m \times m$ identity matrix.

A.8 | PERTURBATION PATTERN OF FECUNDITIES

We could substitute $\mathrm{sgn}(\mathbf{F})$ with another matrix with a different pattern of 0s and 1s, that is, representing a different way of additively perturbing fecundity, and still obtain a null sequence. These alternative perturbation patterns correspond to what is deemed biologically realistic in a specific context. For example, one can also consider the possibility of additively incrementing all entries in those rows of $\mathbf{F}_j$ that contain at least one positive entry. Different nonzero rows of $\mathbf{F}_j$ identify different offspring stages, where offspring can differ, for example, in their size at birth, that are born to parents aged j. Altering all entries in one of these rows for sensitivity analysis means including in the analysis the effect of altering the fecundity of all stages. But there may be stages that, for biological reasons, cannot be reproductive, for example, a stage comprising individuals of too small size. Moreover, not all reproductive stages may be able to produce all types of offspring, for example, a reproductive stage with small-sized individuals may not produce large-sized offspring. For these reasons, we prefer to only look at the effect of additive increments in already existing fecundities.

A.9 | DERIVATION OF EQUATION 22

First, we derive the sensitivity $\varphi(\mu_{k,j})$ of $\ln\lambda$ to a proportional change of the same magnitude exclusively in the entries of $\mathbf{U}_j = \mathbf{U}$ that are in column k. By a reasoning analogous to that brought us to Equation 18 in the main text

$$
\varphi(\mu_{k,j}) = \frac{\mathbf{v}^\top \mathbf{U}\mathbf{E}_{k,k}\mathbf{w}_j}{\lambda \mathbf{v}^\top \mathbf{w}} = \frac{\sum_i v_i u_{i,k} w_{k,j}}{\lambda \mathbf{v}^\top \mathbf{w}} = w_{k,j}\underbrace{\frac{\sum_i v_i u_{i,k}}{\lambda \mathbf{v}^\top \mathbf{w}}}_{=b_k},
\tag{A46}
$$

where $\mathbf{E}_{k,k}$ is a matrix of the same dimensions as $\mathbf{U}$ and with (k,k) entry equal to 1 and 0 everywhere else. Then, we derive the sensitivity $\varphi(m_{j,k})$ of $\ln \lambda$ to an additive change of the same magnitude exclusively in the positive entries of $\mathbf{F}_j$ that are in column k assuming that this column is not zero. Adapting the reasoning that led us to Equation 19 in the main text,

$$\varphi(m_{k,j}) = \frac{\mathbf{v}^\top \mathrm{sgn}(\mathbf{F})\mathbf{E}_{k,k}\mathbf{w}_j}{\lambda \mathbf{v}^\top \mathbf{w}} = \frac{\sum_i v_i \mathrm{sgn}(f_{i,k})w_{k,j}}{\lambda \mathbf{v}^\top \mathbf{w}} = w_{k,j}\underbrace{\frac{\sum_i v_i \mathrm{sgn}(f_{i,k})}{\lambda \mathbf{v}^\top \mathbf{w}}}_{=c_k}. \quad (A47)$$

APPENDIX B

NUMERICAL SIMULATIONS

We made use of the data of Davison et al. (2010) to numerically illustrate a theoretical idea discussed in Section 2.3 of the main text. There, we noted that two initial populations with different age and stage distributions will eventually converge to a common time-varying age distribution when the same sequence of projection matrices is applied to both populations. Figure B1 gives an example of this phenomenon. As it can be seen, convergence is quite fast and can occur within a few steps of demographic projection.
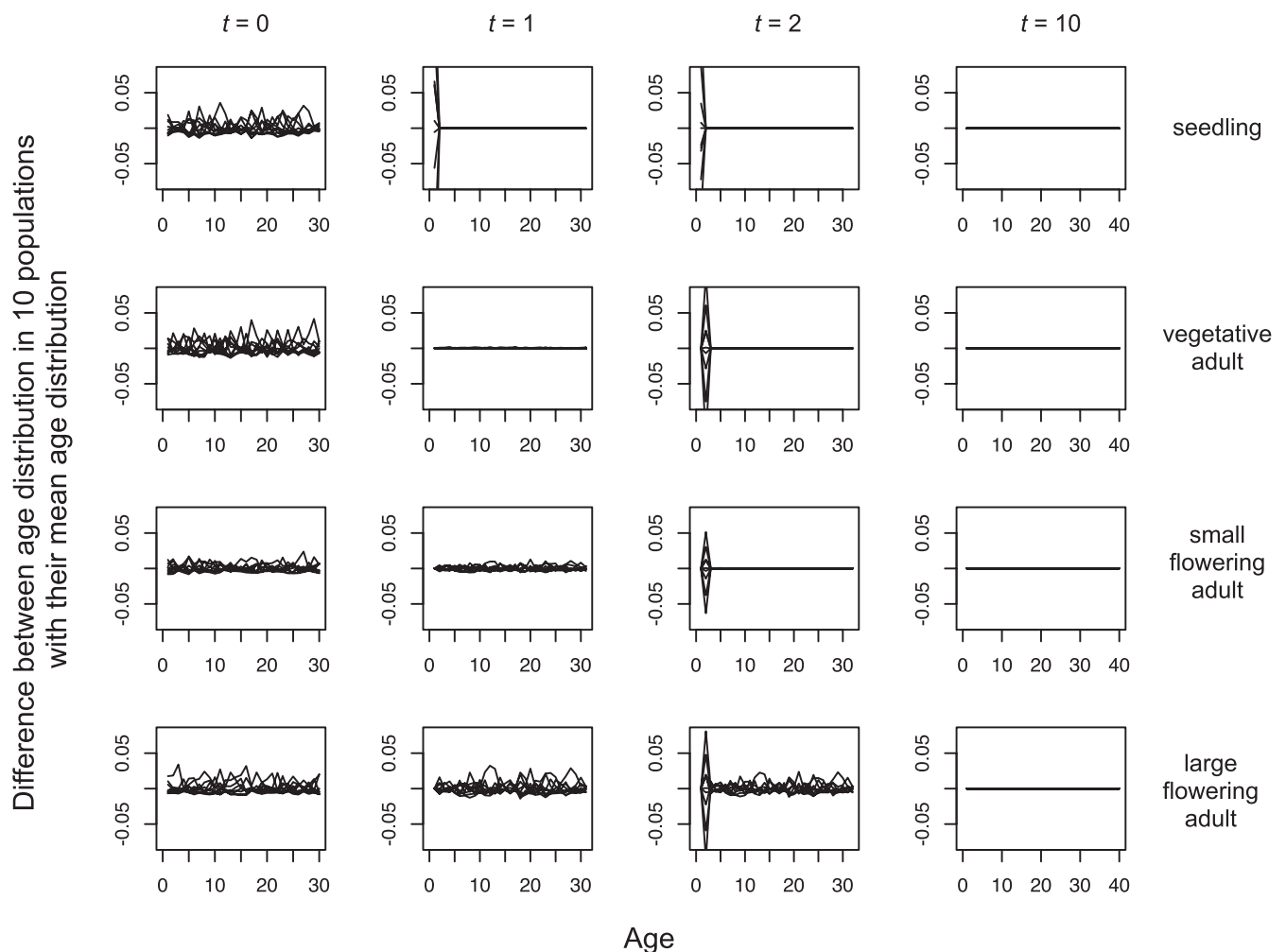

FIGURE B1 Convergence to a common age distribution. We used demographic data for population C of kidney vetch in Davison et al. (2010). We initiated ten populations. To each, we assigned a random age and stage distribution with maximum age 30. We updated the populations for several time steps by subjecting the populations to the same randomly generated sequence of projection matrices. Here, we report for different time points (columns) and within each stage (row), the differences between the age distribution of each population and the arithmetic mean of the ten age distributions. Different lines correspond to different populations. As time increases, differences vanish and lines overlap so that, eventually, only a single horizontal line at level 0 is visible. The Supporting Information (Giaimo & Traulsen, 2022a) contains code to generate this figure.