

We happy few: using structured population models to identify the decisive events in the lives of exceptional individuals

Robin E. Snyder^{1,*,\dagger}

Stephen P. Ellner²

1. Case Western Reserve University, Cleveland, Ohio 44106;

2. Cornell University, Ithaca, New York 14853.

* Corresponding author; e-mail: res29@@case.edu

\dagger The authors wish to be identified to the reviewers.

Manuscript elements: Figure 1, figure 2, figure 3, figure 4, figure 5, Figure 6, figure 7, figure 8, figure 9, figure 10, figure 11, figure 12, online appendices A, B (including Table B1), C and D (including figures D1, D2, D3 D4). Figures 4,5,6,7,8 and 12 include color. We intend this to be an e-article so no figures will need to “print” in color.

Keywords: Reproductive skew, integral projection model, individual stochasticity, growth correlation, *Dacrydium elatum*, *Artemisia ordosica*, *Cedrela odorata*

Dryad doi: doi:10.5061/dryad.3b56d

Manuscript type: e-article.

Prepared using the suggested L^AT_EX template for *Am. Nat.*

Abstract

In any population, some individuals make it big: they are among the few that produce many offspring, grow to large size, etc. What distinguishes the lives of these happy few? We present three approaches for identifying what factors distinguish those “Lucky” individuals who come to dominate reproduction in a population without fixed differences between individuals (genotype, site quality, etc.): comparing life history trajectories for Lucky and Unlucky individuals, and calculating the elasticity of the probability of becoming Lucky to perturbations in demographic rates at a given size or given age. As examples we consider published size-structured integral projection models for the tropical tree *Dacrydium elatum* and semi-arid shrub *Artemisia ordosica*, and a size age-structured matrix model for the tropical tree *Cedrela ordosica*. We find that good fortune (e.g., rapid growth) when small and young matters much more than good fortune when older and larger. Becoming Lucky is primarily a matter of surviving while others die. For species with more variable growth (such as *Cedrela* and *Ordosica*), it is also a matter of growing fast. We focus on reproductive skew, but our methods are broadly applicable, and can be used to investigate how individuals come to be exceptional in any aspect.

Introduction

In many plant and animal populations, a small fraction of individuals produce the majority of all offspring, a phenomenon known as “reproductive skew” (Moran and Clark, 2012, and references therein). Even if all individuals have the same demographic rates, as in the examples we consider in this paper (e.g., the same mean and variance of growth rate as a function of size), there will often be very high variation among individuals in outcomes such as lifetime reproductive success. Who gets to join the happy few who produce many offspring? Are there windows of opportunity where good fortune at the right size or time makes it substantially more likely that an individual will become one of the few who dominate reproduction? How do these critical transition points depend on the form of the life history?

We propose here that these and similar questions about pathways through the life cycle can be addressed using methods from Markov Chain theory, for both matrix and integral projection models (IPM). Demographic studies frequently show that individuals with the same measured attributes (size, age, etc.) can have very different fates: each year some live and some die, some grow more than others, some have more offspring and some fewer. Matrix projection models and IPMs quantify this variability, so that an individual’s current state determines the relative likelihood of different fates at the next census. As a result, an individual’s transitions among age- or stage-classes in a matrix projection model are a finite-state Markov Chain with death as an absorbing state (an “absorbing” state or set of states can be entered, but not left). This perspective was introduced to ecology by Cochran and Ellner (1992) and greatly extended by Caswell, Tuljapurkar, and collaborators (Caswell, 2001, 2011*a,b*, 2012; Metcalf et al., 2009; Steiner and Tuljapurkar, 2012; Steiner et al., 2012; Tuljapurkar and Horvitz, 2006; Tuljapurkar et al., 2009). In an IPM the state space is continuous, or a mix of discrete and continuous, but an individual’s transitions are still a Markov chain and the same basic approach can be used (Ellner et al., 2015; Meyn and Tweedie, 1993).

Our work echoes Caswell (2011*a*), which calculates the moments of the distribution of life-

time reproductive output, as a way of asking how rare are the few individuals that dominate reproduction. Our work extends these calculations to ask not just how rare the happy few are, but how an individual comes to be one of them. We also show how the complete probability distribution of lifetime reproductive output can be computed.

Others have also asked how individuals come to be Lucky. Zuidema et al. (2009) analyzed an age-size structured model for tropical trees, to ask whether consistently fast-growing juveniles make a disproportionate contribution to population growth. They used loop analysis (van Groenendael et al., 1994) of an age-size structured matrix projection model to measure the importance of life-cycle pathways involving fast growth.¹ Because trees are slow-growing and long-lived, the matrix model used many size-classes, and such large models are challenging for loop analysis. First, the decomposition of the life cycle into loops is rarely unique; users must choose a preferred loop type (long, short, etc.) in a decomposition algorithm (Zuidema et al., 2009), sort groups into meaningful categories (e.g., fast growers), and check whether conclusions are robust to different preferences and groupings. Second, loop analysis measures the “importance” of different pathways by loop elasticity, which is not clearly linked to individual reproductive output or to population growth because it is generally impossible to perturb only one loop. Jansen et al. (2012) addressed the same question for an understory palm, using elasticity analysis of an age-size structured model. Transitions of individuals with above-median size for their age had 1.8 times higher total elasticity than transitions of individuals with below-median size for their age, showing that fast growers are especially important for population growth. However, this does not necessarily imply that fast growth is the only factor, or the most important factor, in becoming one of the Lucky few. Metcalf et al. (2009) used IPMs for tropical trees in a time-varying light environment to study how a new seedling’s light environment affected its chance of reaching the canopy. They computed life expectancy and the expected time to reach canopy height, as

¹Loop analysis breaks up the life-cycle graph into a set of paths along possible stage transitions (including survival without a change in stage) that begin and end at the same stage. Each such loop is assigned a “loop elasticity” for population growth rate, whose sum over all loops equals 1.

a function of the light environment at birth. By comparing those two numbers, they inferred whether individuals born into a particular light environment are more likely to reach the canopy or to die first. The methods we develop here work with the exact probability (within the model) of events such as reaching the canopy, and can make comparisons across the entire life trajectory, not just initial state.

Here we develop three approaches for identifying what events or factors distinguish those few individuals who dominate reproduction in size-structured IPMs or matrix models. The first is to compare life trajectories: how do individual growth rate as a function of size, survival as a function of age, etc., differ between Lucky and Unlucky individuals? The second is to calculate the sensitivity or elasticity of the probability of becoming Lucky to various state-dependent demographic parameters. For example, does expected survival or variability in growth have a larger influence on the probability of becoming exceptionally large? Does survival matter more at small sizes or large? Caswell (2012, 2013) recently derived some analogous sensitivity formulas for matrix models; however these are largely specific to matrix models.² The third approach is to compute the sensitivity to perturbations at different ages: if you're going to have a good year, is it better to have it when you're young, old, or at some crucial point in development? We present what needs to be calculated, and how, for each approach, and then apply the results to three examples: published IPMs for the tropical tree *Dacrydium elatum* (Zuidema et al., 2010), the semi-arid shrub *Artemisia ordosica* (Li et al., 2011), and an age-size structured matrix model for tropical tree *Cedrela odorata* (Zuidema et al., 2009). R (Team, 2015) scripts to replicate all of our calculations for these examples are provided as an online Supplement to this paper.

To introduce our approaches in the simplest setting, we consider first a purely size-structured population in which becoming Lucky is equivalent to reaching a critical size, such as a tree

²There are two reasons for this. First, the formulas for many of the quantities being perturbed involve return times or rely on Doebelin's formula for return time probabilities (Iosifescu, 1980, Chapter 2). These do not exist in IPMs because the random size-transitions imply that there is zero probability of an individual being exactly the same size in two different years. The second is that many of the sensitivity formulas and their derivations use matrix calculus, which is also not defined (at least yet) for IPMs.

reaching canopy height or a fish reaching the size at sexual maturity. We then present a more direct but less simple approach in which individuals are cross-classified by size and by total reproductive output up to the present, T , and Lucky is defined by achieving a critical level of T , such as the 90th percentile of the distribution of T at death. We derive analytic formulas for the quantities of interest, so that they can be computed without recourse to individual-based simulations, which would be enormously slower because of the sample sizes need to accurately compute small probabilities and their responses to small parameter changes.

The methods we present are broadly applicable. They can be used with both matrix models and IPMs, or with other models where individual life-trajectories are described in terms of possible transitions and their consequences. They can be applied to any individual-level state variable (e.g., size, site quality, local crowding, number of breeding attempts) or combination of variables to explore how these additional factors contribute to individuals becoming Lucky. Here we investigate reproductive skew — how do individuals come to be among the few that dominate reproduction? — but the methods can be used to study any way in which some individuals within a population come to be different from others. Also, we limit ourselves here to situations where differences between individuals arise purely by chance, what Caswell (2009) calls *individual stochasticity*, rather than fixed differences (e.g., genetic heterogeneity or other life-long differences in “quality”). It is important to analyze individual stochasticity because it provides a baseline for interpreting observed between-individual variability in life trajectories (e.g., Steiner and Tuljapurkar, 2012). And, as we outline in the Discussion, we expect that our methods can be extended to models that include fixed or varying differences, in addition to individual stochasticity.

Methods: size-structured models

The questions of whether and how an individual becomes one of the few who dominate reproduction can be studied using the mathematical theory of Markov chains with multiple absorbing

sets. Here our absorbing sets are “Lucky” — exceeded a certain size, certain number of offspring, etc., before you die — and “Unlucky” — die without ever exceeding the threshold. However, we cannot use standard results (e.g. Kemeny and Snell, 1960) “off the shelf” because Lucky is not a single state and Lucky individuals continue to change in size. In this section we explain how this situation can be analyzed in the setting of a purely size-structured model. Note: we use Lucky/Unlucky (capitalized) to refer to an individual’s final fate, and lucky/unlucky to refer to good or bad fortune at some point along the way.

Let $n_t(z)$ denote the probability of being in individual state z at time t (when z is a continuous state variable, this is to be interpreted as a shorthand way of saying that $n_t(z)$ is the probability density function for individual state at time t , evaluated at z). The state z can represent size, the number of offspring to date, size and number of offspring, etc. We let the transition kernel be denoted as $P(z', z)$, where z denotes current state and z' state at the next time step, so that

$$n_{t+1}(z') = \int_A P(z', z) n_t(z) dz. \quad (1)$$

We write this equation as $n_{t+1} = Pn_t$, analogous to the notation for matrix-vector multiplication. Similarly, for a function $f(z)$ and a kernel $G(z', z)$, we define their convolution fG as the function

$$fG(z) = \int_A f(z') G(z', z) dz'. \quad (2)$$

For example, if $f(z')$ is z' , the size next year, and $G(z', z)$ is the probability of growing from size z to size z' , then fG is a function of z whose value is the expected size next year given that the current size is z .

Since z represents size (or size and reproductive output) in our examples, we will denote the set of unlucky states by S (“small”) and the set of lucky states by B (“big”). The set of all states is A . The *transient kernel* $Q(z', z)$ is defined to be that portion of P which represents transitions from one small state to another ($z', z \in S$). Classifying individuals as Lucky if they exceed a critical value of some state variable (size, or total reproductive output) means that our state space has two absorbing sets: becoming Lucky, or death without ever exceeding the critical value. Once an individual’s state enters an absorbing set, it remains there. The set of small states is “transient”

in the sense that individuals are born into the set, but eventually exit it one way or the other. The transient kernel Q describes the transitions between one time step and the next, prior to the time at which the individual exits.

Probability of becoming lucky

Becoming lucky means reaching a state in the set B before death, starting from a state in S (by assumption, all individuals are in S at birth). The probability of becoming Lucky in the next time step, starting from state z in S , is

$$a_{\mathbb{L}}(z) = \int_B P(z', z) dz', \quad z \in S. \quad (3)$$

Here and throughout the paper we use \mathbb{L} in equations to mean “Lucky”, and L is used for the lower limit of a size interval. Following standard Markov process theory (e.g., Kemeny and Snell, 1960), the *fundamental operator* N is defined as

$$N(z', z) = (I - Q)^{-1}(z', z), \quad z, z' \in S. \quad (4)$$

The series expansion $N = I + Q + Q^2 + \dots$ shows that $N(z', z)h$ is the expected total number of times that an individual’s state is in the interval $(z', z' + h)$ in S , prior to death or becoming Lucky, if one starts in state z in S (Ellner et al., 2015, Chapter 3). Following Kemeny and Snell (1960) and Ellner et al. (2015), the probability of becoming Lucky, $q_{\mathbb{L}}$, obeys the relation $q_{\mathbb{L}} = a_{\mathbb{L}} + q_{\mathbb{L}}Q$; that is, the probability of becoming Lucky is the probability of doing so in your next transition ($a_{\mathbb{L}}$), plus the probability of moving instead to another $z \in S$ and eventually becoming Lucky from there ($q_{\mathbb{L}}Q$). This gives $q_{\mathbb{L}}(I - Q) = a_{\mathbb{L}}$ and therefore the probability of eventually becoming Lucky given that you are now in state z in S is

$$q_{\mathbb{L}}(z) = a_{\mathbb{L}}N(z) = \int_S a_{\mathbb{L}}(z')N(z', z) dz', \quad z \in S. \quad (5)$$

Eq. 5 says that the probability of becoming Lucky starting from small state z equals the probability of transitioning from z' directly to a big state, times the expected amount of time spent in state z' before becoming Lucky, totaled over all small states z' .

Comparing life trajectories

To compare life trajectories of Lucky and Unlucky individuals, we use the single-time-step transition kernels conditional on eventually becoming Lucky, and on death while still Unlucky. All individuals start in S , and we use $Q_{\mathbb{L}}$ and $Q_{\mathbb{U}}$ to denote the two conditional transition kernels on S prior to exit through becoming Lucky or dying. Again following Kemeny and Snell (1960),

$$Q_{\mathbb{L}}(z', z) = \frac{Q(z', z)q_{\mathbb{L}}(z')}{q_{\mathbb{L}}(z)}, \quad z, z' \in S. \quad (6)$$

This is a standard result for finite-state chains, and can be derived similarly for general state spaces (see Appendix A). The kernel conditional on never being Lucky takes the same form using the probability of ending up in the Unlucky absorbing state, $q_{\mathbb{U}} = 1 - q_{\mathbb{L}}$:

$$Q_{\mathbb{U}}(z', z) = \frac{Q(z', z)q_{\mathbb{U}}(z')}{q_{\mathbb{U}}(z)}, \quad z, z' \in S. \quad (7)$$

By definition, the Unlucky are destined never to become large and only leave S by dying, so $Q_{\mathbb{U}}$ governs their entire life. But the lives of the Lucky continue after they enter B and their conditional kernel must include transitions on the entire state-space A . Suppose first that individuals larger than the threshold size z_c for being Lucky cannot ever shrink back below the threshold. We can then set $S = [L, z_c)$, $B = [z_c, U]$. The conditional kernel for the Lucky is then

$$P_{\mathbb{L}}(z', z) = \frac{P(z', z)q_{\mathbb{L}}(z')}{q_{\mathbb{L}}(z)}, \quad z, z' \in A, \quad (8)$$

where $q_{\mathbb{L}}(z)$ is defined to equal 1 for all z in B .

When the initial state of newborns is given by a probability distribution $c_0(z)$, we also need the conditional distributions of initial state for Lucky and Unlucky. We can get these from (8) by imagining that all individuals start life in the same imaginary state z_{-1} at age -1 and move to a state at age 0 according to $P(z, z_{-1}) = c_0(z)$. The conditional initial state distributions are thus $c_0(z)q_{\mathbb{L}}(z)$ and $c_0(z)q_{\mathbb{U}}(z)$ normalized so that they integrate to 1, which we denote $c_{0,\mathbb{L}}$ and $c_{0,\mathbb{U}}$.

Once we have the conditional initial states and kernels, we can calculate various quantities as functions of time or state for the two groups. For example, the expected survival to age a ,

starting at size z , is

$$s_a(z) = \int_A P^a(z', z) dz', \quad (9)$$

where P^a is the a^{th} iterate of P (for example, if P is the iteration matrix for P , then the iteration matrix for P^3 is $P\%*\%P\%*\%P$ in R and $P*P*P$ in MATLAB, Octave, or Python using NumPy). To compare expected survival for Lucky and Unlucky individuals, we simply use $P_{\mathbb{L}}$ or $Q_{\mathbb{U}}$, respectively, in place of P , and integrate the result over the conditional initial distribution. (E.g., the expected survival to age a for Lucky individuals is $\int_A \int_A P_{\mathbb{L}}^a(z', z) dz' c_{0,\mathbb{L}}(z) dz$.) For Unlucky individuals, the integral in (9) only runs over the set of small sizes S because the Unlucky have no chance of reaching B .

To consider expected growth, conditional on survival, as a function of size, we calculate the expected sizes next year:

$$\mathbb{E}[z_{t+1}|z_t = z] = \frac{\int_A z' P(z', z) dz'}{\int_A P(z', z) dz'}. \quad (10)$$

Note that the denominator is the size-specific probability of surviving for one year. Again, to compare Lucky and Unlucky, we simply use either $P_{\mathbb{L}}$ or $Q_{\mathbb{U}}$ in place of P in (10). Using $P_{\mathbb{L}}^a$ or $Q_{\mathbb{U}}^a$ in place of P^a gives the expected size (conditional on survival) after a time steps.

If individuals can shrink back below z_c , then to compare life trajectories we need to define $S = [L, z_c), B = [L, U]$ considered as two separate intervals “side by side” (Figure 1). S represents individuals who have always been smaller than z_c . Sizes $z < z_c$ in B represent individuals who were larger than z_c at some time in the past, and reached z by shrinking. This is equivalent to cross-classifying individuals by a second state variable that is 0 at birth and switches permanently to 1 when an individual reaches z_c or larger. In Appendix A we explain how comparisons using conditional kernels can be done when individuals can shrink back below z_c .

Sensitivity and elasticity of $q_{\mathbb{L}}$, the probability of becoming Lucky

In this section we show how to calculate the derivative of $q_{\mathbb{L}}$ with respect to perturbations of demographic rates at a particular size z_0 , or at a particular age a . These are our main new

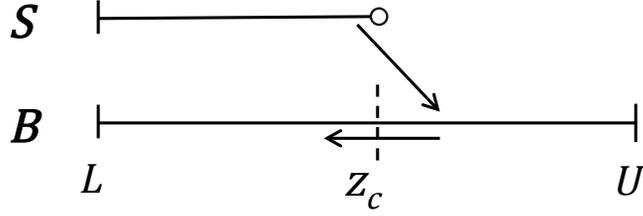


Figure 1: The expanded state space for a purely size-structured model with shrinkage below the critical size z_c (note that L and U denote the lower and upper limits of the size range, rather than the fates of being Lucky or Unlucky (\mathbb{L} and \mathbb{U})). When an individual first grows from size $z < z_c$ to size $z' \geq z_c$, it moves from S to B . When an individual in B shrinks from $z \geq z_c$ to $z < z_c$ it stays in B . Individuals who die from S are Unlucky (absorbing state 1); individuals who die from B are Lucky (absorbing state 2).

theoretical results. This section is the most technical part of the paper, and the derivations use IPM and Markov chain theory. However, readers without this background can skip from here directly to the following section *Examples: size-structured models*, where we apply the results to published models based on empirical demographic studies.

Sensitivity and elasticity to luck at a given size

Perturbations to Q at a size z_0 are characterized mathematically by $\frac{\partial Q}{\partial \phi(z_0)}$, where $\phi(z_0)$ is the attribute being perturbed, such as expected survival at state z_0 or the standard deviation of growth at state z_0 . The derivative is itself a kernel; specifically, increasing $\phi(z_0)$ by ϵ increases Q by $\epsilon \frac{\partial Q}{\partial \phi(z_0)}$ plus higher-order terms in ϵ .

From the expression for $q_{\mathbb{L}}$ (the probability of becoming a Lucky individual, eqn.(5)), and the result (familiar for matrices, but also true for invertible kernels) that $d(A^{-1}) = -A^{-1}(dA)A^{-1}$, we have that

$$\frac{\partial q_{\mathbb{L}}}{\partial \phi(z_0)} = \frac{\partial a_{\mathbb{L}}}{\partial \phi(z_0)} N + a_{\mathbb{L}} N \frac{\partial Q}{\partial \phi(z_0)} N, \quad (11)$$

where

$$\frac{\partial a_{\mathbb{L}}}{\partial \phi(z_0)} = \int_B \frac{\partial Q(z', z)}{\partial \phi(z_0)} dz'. \quad (12)$$

To calculate the elasticity, defined as $\frac{\phi(z_0)}{q_{\mathbb{L}}} \frac{\partial q_{\mathbb{L}}}{\partial \phi(z_0)}$, we replace $\frac{\partial Q}{\partial \phi(z_0)}$ with $\frac{\phi(z_0)}{q_{\mathbb{L}}} \frac{\partial Q}{\partial \phi(z_0)}$ in eqns. (11) and (12).

It is useful to consider a concrete example. Suppose that $P(z', z) = s(z)G(z', z)$, where $s(z)$ is the probability of survival at size z and $G(z', z)$ is the probability distribution of subsequent size z' conditional on current size z . Let $G(z', z)$ be a Gaussian distribution with mean $\bar{z}(z)$ and standard deviation $\sigma(z)$:

$$G(z', z) = \frac{1}{\sqrt{2\pi\sigma^2(z)}} \exp\left(\frac{-(z' - \bar{z}(z))^2}{2\sigma^2(z)}\right). \quad (13)$$

Consider the sensitivity $q_{\mathbb{L}}(z)$ to $\bar{z}(z_0)$, the expected size in the following year starting from size z_0 . A perturbation of size ϵ can be represented as changing $\bar{z}(z)$ to $\bar{z}(z) + \epsilon\delta(z - z_0)$ where δ is the Dirac δ -function.³ We then calculate the derivative of Q as

$$\begin{aligned} \frac{\partial Q(z', z)}{\partial \bar{z}(z_0)} &\equiv s(z) \lim_{\epsilon \rightarrow 0} \frac{\partial}{\partial \epsilon} \frac{1}{\sqrt{2\pi\sigma^2(z)}} \exp\left(\frac{-(z' - (\bar{z}(z) + \epsilon\delta(z - z_0)))^2}{2\sigma^2(z)}\right) \\ &= \frac{1}{\sigma^2(z)} (z' - \bar{z}(z)) \delta(z - z_0) Q(z', z). \end{aligned} \quad (14)$$

The derivative of $a_{\mathbb{L}}$ is thus

$$\frac{\partial a_{\mathbb{L}}}{\partial \bar{z}(z_0)} = \frac{1}{\sigma^2(z)} \delta(z - z_0) \int_B (z' - \bar{z}(z)) Q(z', z) dz' = \frac{1}{\sigma^2(z)} \delta(z - z_0) A_z(z), \quad (15)$$

where A_z is defined to be the integral in (15). The first term in eqn. (11) is

$$\begin{aligned} \frac{\partial a_{\mathbb{L}}}{\partial \bar{z}(z_0)} N &= \int_S \frac{1}{\sigma^2(z')} A_z(z') \delta(z' - z_0) N(z', z) dz' \\ &= \frac{1}{\sigma^2(z_0)} A_z(z_0) N(z_0, z). \end{aligned} \quad (16)$$

³If δ functions are unfamiliar: the relevant properties of $\delta(z - z_0)$ are that it is nonzero only at $z = z_0$ and that $\int \delta(z - z_0) f(z) dz = f(z_0)$ for any smooth function f . Clearly this is not a function in the usual sense. It can be thought of as representing the limit of probability density functions with mean z_0 and variance decreasing to zero: we perturb a demographic rate in a small region around z_0 , and let the size of the region shrink to zero with the total perturbation held constant.

Turning to the second term in eqn. (11),

$$\begin{aligned}\frac{\partial Q}{\partial \bar{z}(z_0)} N &= \int_S \frac{1}{\sigma^2(u)} (z' - \bar{z}(u)) \delta(u - z_0) Q(z', u) N(u, z) du \\ &= \frac{1}{\sigma^2(z_0)} (z' - \bar{z}(z_0)) Q(z', z_0) N(z_0, z).\end{aligned}\quad (17)$$

Because $a_{\mathbb{L}} N$ is a vector and $\frac{\partial Q}{\partial \bar{z}(z_0)} N$ is a kernel, their product (the second term in eqn. 11) is a vector, as it should be.

To calculate the elasticity, we would use $\frac{\bar{z}(z_0)}{q_{\mathbb{L}}} \frac{\partial Q(z', z)}{\partial \bar{z}(z_0)}$ in place of $\frac{\partial Q(z', z)}{\partial \bar{z}(z_0)}$. This produces a factor of $\frac{\bar{z}(z_0)}{q_{\mathbb{L}}}$ multiplying the expressions in eqns. (15),(16), and (17).

Sensitivity and elasticity to luck at a given age

Alternatively, we can consider the effect of a perturbation to luck at a given age. That is, we imagine that at some age an individual experiences a good year in which survival or some other vital rate is temporarily increased for all sizes equally. Mathematically we are considering $\partial a_{\mathbb{L}} / \partial \phi$, where ϕ is the vital rate of interest. As above, we can express the probability of becoming Lucky before death as the probability of becoming Lucky immediately (i.e. moving into B at the next time step), plus the probability of moving to a state z' and then becoming Lucky, plus the probability of making two transitions before becoming Lucky, etc.:

$$q_{\mathbb{L}} = a_{\mathbb{L}} + a_{\mathbb{L}} Q + a_{\mathbb{L}} Q^2 + \cdots = a_{\mathbb{L}} N. \quad (18)$$

Let us denote a perturbation to $a_{\mathbb{L}}$ or Q at age t with a subscript t . Then if the perturbation occurs at age 0 (which will affect both $a_{\mathbb{L}}$ and Q for that year of life)

$$q_{\mathbb{L}} = a_{\mathbb{L}_0} + a_{\mathbb{L}} Q_0 + a_{\mathbb{L}} Q Q_0 + a_{\mathbb{L}} Q^2 Q_0 + \cdots = a_{\mathbb{L}_0} + a_{\mathbb{L}} N Q_0. \quad (19)$$

If the perturbation occurs in year 1, we have

$$q_{\mathbb{L}} = a_{\mathbb{L}} + a_{\mathbb{L}_1} Q + a_{\mathbb{L}} Q_1 Q + a_{\mathbb{L}} Q Q_1 Q + \cdots = a_{\mathbb{L}} + a_{\mathbb{L}_1} Q + a_{\mathbb{L}} N Q_1 Q. \quad (20)$$

The general case for a perturbation at age t is

$$q_{\mathbb{L}} = a_{\mathbb{L}} \sum_{j=0}^{t-1} Q^j + a_{\mathbb{L}_t} Q^t + a_{\mathbb{L}} N Q_t Q^t. \quad (21)$$

Differentiating the last expression with respect to the perturbation ϕ in year t (i.e. only applying the derivative to $a_{\mathbb{L}_t}$ and Q_t), we have

$$\frac{\partial q_{\mathbb{L}}}{\partial \phi} = \frac{\partial a_{\mathbb{L}}}{\partial \phi} Q^t + a_{\mathbb{L}} N \frac{\partial Q}{\partial \phi} Q^t \quad (22)$$

for a perturbation to vital rate ϕ at age t .

To calculate the elasticity to ϕ , we need to consider proportional perturbations at all sizes at a given age. This is achieved by replacing $\frac{\partial Q}{\partial \phi}$ with $\frac{1}{q_{\mathbb{L}}} \Phi \circ \frac{\partial Q}{\partial \phi}$, where Φ is the kernel with values $\phi(z', z)$ and $A \circ B$ represents element-wise multiplication of A and B . For a vital rate that depends only on the current state, z , $\Phi(z', z)$ reduces to $\phi(z)$.

Again, it is useful to consider the specific example in eqn. (13) and calculate the components of sensitivity to expected growth at age t . The partial derivative of Q is

$$\frac{\partial Q(z', z)}{\partial \bar{z}} = \frac{z' - \bar{z}(z)}{\sigma^2(z)} Q(z', z) \quad (23)$$

and the partial derivative of $a_{\mathbb{L}}$ is thus

$$\frac{\partial a_{\mathbb{L}}}{\partial \bar{z}} = \frac{1}{\sigma^2(z)} \int_B (z' - \bar{z}(z)) Q(z', z) dz'. \quad (24)$$

If we were calculating elasticity, we would use

$$\left(\frac{1}{q_{\mathbb{L}}} \Phi \circ \frac{\partial Q}{\partial \bar{z}} \right) (z', z) = \frac{\bar{z}(z)}{q_{\mathbb{L}}} \frac{z' - \bar{z}(z)}{\sigma^2(z)} \cdot Q(z', z) \quad (25)$$

in place of $\frac{\partial Q}{\partial \bar{z}}$ and

$$\frac{\bar{z}(z)}{q_{\mathbb{L}}} \frac{1}{\sigma^2(z)} \int_B (z' - \bar{z}(z)) Q(z', z) dz' \quad (26)$$

in place of $\frac{\partial a_{\mathbb{L}}}{\partial \bar{z}}$.

Alternative method for calculating derivatives

If the kernel is sharply peaked, so that many grid points are required for accurate numerical integration, it may be faster to estimate the sensitivity numerically. To estimate the sensitivity of

$q_{\mathbb{L}}$ to expected growth at size z_0 , calculate a perturbed transition kernel $P_\epsilon(z', z)$. For $z = z_0$, we add ϵ to the expected growth,

$$g(z', z_0) = \frac{1}{\sqrt{2\pi\sigma^2(z_0)}} \exp\left(\frac{-(z' - (\bar{z}(z_0) + \epsilon))^2}{2\sigma^2(z_0)}\right), \quad (27)$$

and let $P_\epsilon(z', z) = P(z', z)$ for all other values of z . If $P(z', z)$ is represented by a matrix, this corresponds to perturbing one column of P . We calculate the new probability of becoming Lucky, q_{L_ϵ} , from the perturbed kernel and estimate the sensitivity as

$$\frac{\partial q_{\mathbb{L}}}{\partial \bar{z}(z_0)} \approx \frac{q_{L_\epsilon} - q_{\mathbb{L}}}{\epsilon}. \quad (28)$$

Examples: size-structured models

We now demonstrate our methods using two published size-structured IPMs for perennial plants. The first is for *Dacrydium elatum*, one of six fitted models for tropical trees in Zuidema et al. (2010). The second is for the shrub *Artemisia ordosica* at different stages of dune fixation in Mongolia (Li et al., 2011); we use the parameters for the final stage, fixed dunes with microbiotic crust. The size measure for *Artemisia* is plant height. In the *Dacrydium* model the size measure for saplings and trees ($> 1\text{cm dbh}$) is dbh, but for seedlings ($< 1\text{cm dbh}$) there is a 4-stage matrix model using height classes. All new seedlings are born into the smallest height class. One important difference between these species is that growth of *Dacrydium* is much less variable: for example, shrinkage is rare ($< 3\%$ chance at all sizes and $< 1\%$ chance above 3cm), while in *Artemisia* the chance of shrinking is $> 25\%$ for shrubs more than halfway to z_c (as set below). The choice of species and dune fixation stage for these examples was arbitrary, and our methods would work equally well for the other species and parameter sets.

We also analyzed the age-size structured matrix model for the tropical tree *Cedrela odorata* developed by Zuidema et al. (2009). We include this example not only because it was the initial motivation for our work, but also because it models a species with persisting (but not permanent) variability among individuals in growth rate: a high growth rate in one year (e.g., the year of

birth) is likely to be followed by high growth in following years. Each individual's growth is fairly deterministic, as in *Dacrydium*, but individuals of the same size may have very different growth rates, as in *Artemisia*. Persistent differences in light availability and soil moisture contribute (to differing degrees, in different locations) to between-individual differences in growth rate in *Cedrela* that persist over multiple years (Brienen et al., 2010). This example also illustrates that our methods are applicable to matrix models.

The first step is to define a threshold size for being Lucky. For *Dacrydium* we did this by looking at size-specific demography and the relationship between current size and future reproductive success (in the next section, we develop another approach in which Lucky is defined directly by lifetime reproductive output). Figure 2A shows the fitted size-specific survival and "breeding" (flowering and making seeds) probabilities for *Dacrydium* saplings and trees (> 1cm dbh). Survival is high, and nearly 100% for 10-50cm dbh trees. The model assumes that per-capita seed production of flowering adults is independent of size, so fecundity is proportional to the size-dependent breeding probability and thus rises sharply between 10 and 30cm dbh. Figure 2B shows the mean and standard deviation of the expected total number of breeding events (between "now" and death) as a function of current size (the formulas for these are derived in Ellner et al. (2015, Chapter 3)). Below about 10cm dbh, the standard deviation of future breeding events exceeds the mean: some will breed often, some not at all. Above 20cm dbh the mean and standard deviation converge onto nearly the same value. This is because the number of future breeding events is approximately the same as the time to death, which has an approximately geometric distribution with low probability of "success" (death). Looking at reproductive output at the whole-population level, small individuals are far more numerous than large ones (Fig. 2C), but most of the seeds produced each year come from individuals 20cm dbh or larger (Fig. 2D). We therefore take 20cm as the critical size for being Lucky.

Similar calculations for *Artemisia ordosica* suggest a threshold of 40cm height for being Lucky. At that size, annual fecundity begins to rise above effectively zero (Figure 3A), expectation of future reproductive success is about half that of the largest individuals observed (Fig. 3B), and

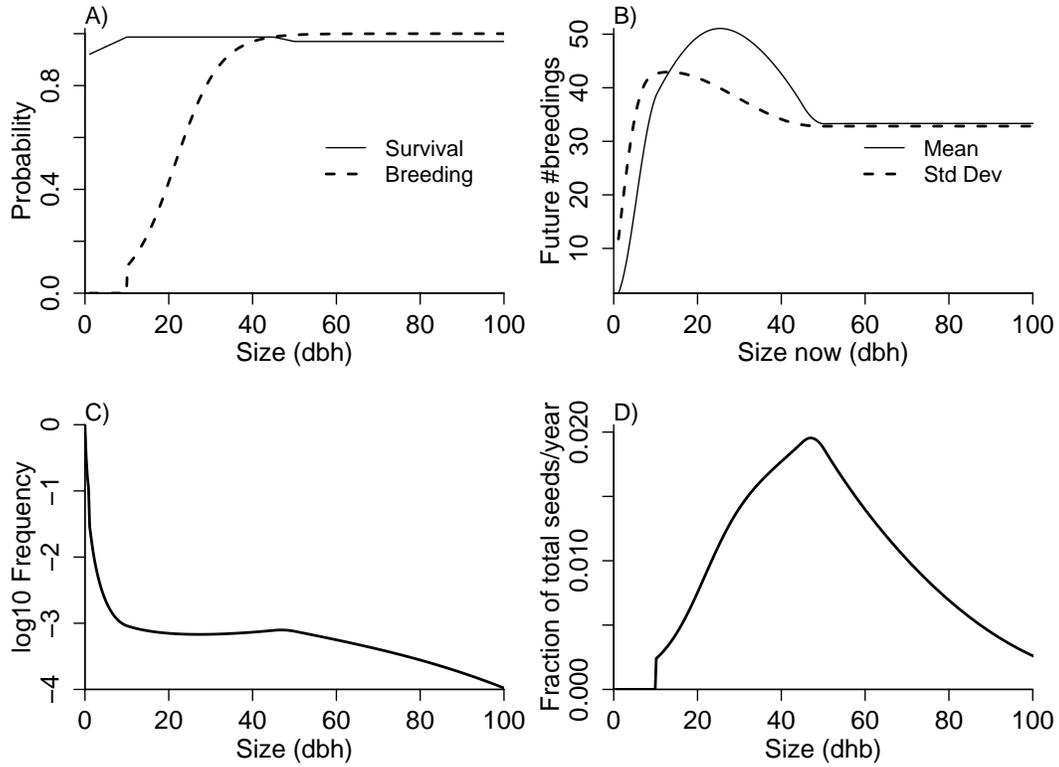


Figure 2: A) Annual probabilities of survival and "breeding" (flowering and producing seeds) as a function of size in the Zuidema et al. (2010) IPM for *Dacrydium elatum*, for saplings and trees (> 1cm dbh). B) Mean (solid) and standard deviation (dashed) of the number of breeding events from the present until death, as a function of current size. C) Stable size distribution (on log-10 scale). D) Distribution of parent size for seeds produced in a given year, for a population at stable size distribution. The plotted curve is the product of the stable size distribution and size-specific mean fecundity, normalized to a probability density function. Source file: ZuidemaBreedingFreq.R and scripts that it sources.

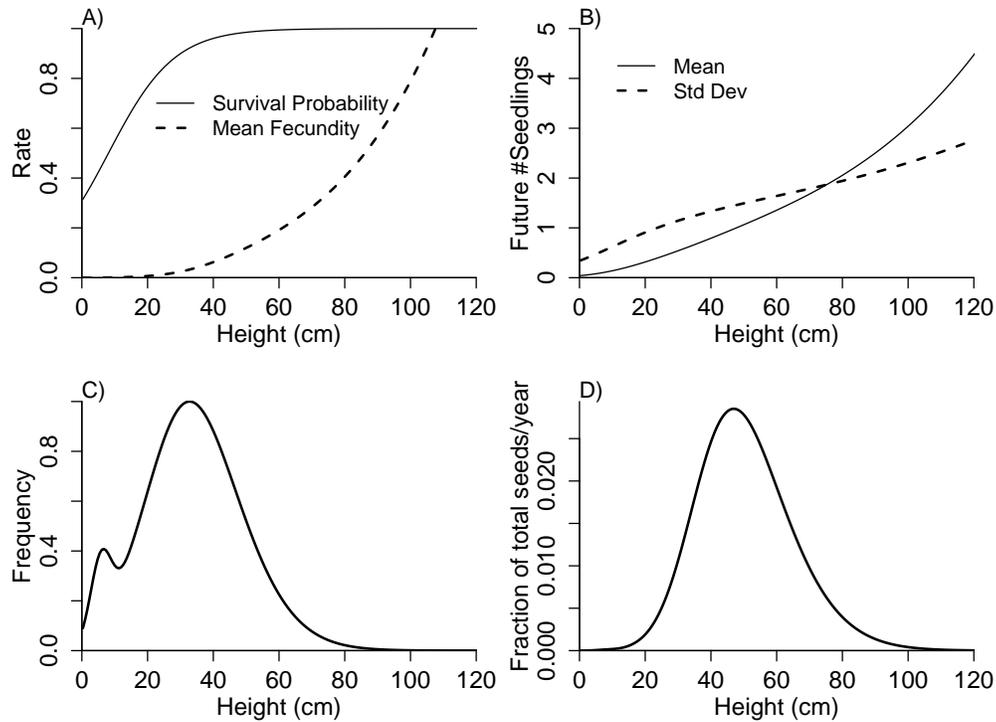


Figure 3: A) Annual survival probability and mean fecundity (number of seedlings) as a function of size in the IPM for *Artemisia ordosica* by Li et al. (2011), 2007-2008 census period on fixed dunes with microbiotic crust. B) Mean (solid) and standard deviation (dashed) of total offspring (seedlings), from the present until death, as a function of current size. The formulas for these are derived in Ellner et al. (2015, Chapter 3)). C) Stable size distribution. D) Distribution of parent size for seedlings produced in a given year, for a population at stable size distribution. The plotted curve is the product of the stable size distribution and size-specific mean fecundity, normalized to a probability density function. Source file: `ArtemisiaBreedingFreq.R` and scripts that it sources.

individuals 40cm or taller contribute most of the seedlings each year even though smaller individuals are more numerous in the stable size distribution (Fig. 3C,D). Fecundity (measured by seedlings at the next census) is low because of the very low establishment probability in dunes with a soil crust. For most individuals, the expected number of future seedlings is 1 or less.

The *Cedrela* matrix model uses 10-year age classes (0-10, 11-20 etc.) and 10cm dbh size categories, up to age 160 and size 120cm. In the model, trees in size-classes 1-5 have no fecundity, while those in size class 6 and above (50cm dbh and larger) have constant survival and fecundity (independent of age and size) and never shrink back into size classes 1-5. Thus, contributing to future population growth is equivalent to entering size class 6, and once that happens, lifetime reproductive output is strictly a matter of survival. We therefore defined *Cedrela* trees as Lucky if they reach size class 6. Additional details about the model are in Appendix C.

Comparison using conditional kernels

Dacrydium: Because the growth of *Dacrydium* saplings and trees is nearly deterministic, there is little scope for becoming Lucky through rapid growth. As a result, luck is mostly a matter of not dying young, and so the age-specific survival curves are vastly different for Lucky and Unlucky individuals (Figure 4A). The onset of visible mortality occurs at the age (≈ 100 years) at which trees typically pass the 20cm threshold (unless they die first) that defines the Lucky class. Lucky trees are also slightly larger at all ages: the trees that contribute most to population growth are fast growers, as Zuidema et al. (2009) concluded about *Cedrela* (Figure 4B). However, in *Dacrydium* the differences in growth are small, and are largely the result of Lucky trees moving faster through the seedling stages (which reduces their exposure to the high mortality of seedlings). After that, Lucky and Unlucky have very similar size-specific growth rates (Figure 4C). For several decades the Lucky have higher age-specific growth rate (Figure 4D), mainly because Lucky trees are larger than equally old Unlucky trees due to their faster initial growth as seedlings. (Note that the growth rate of Unlucky individuals eventually drops off to zero, as it must: if they kept growing, they would eventually become Lucky.)

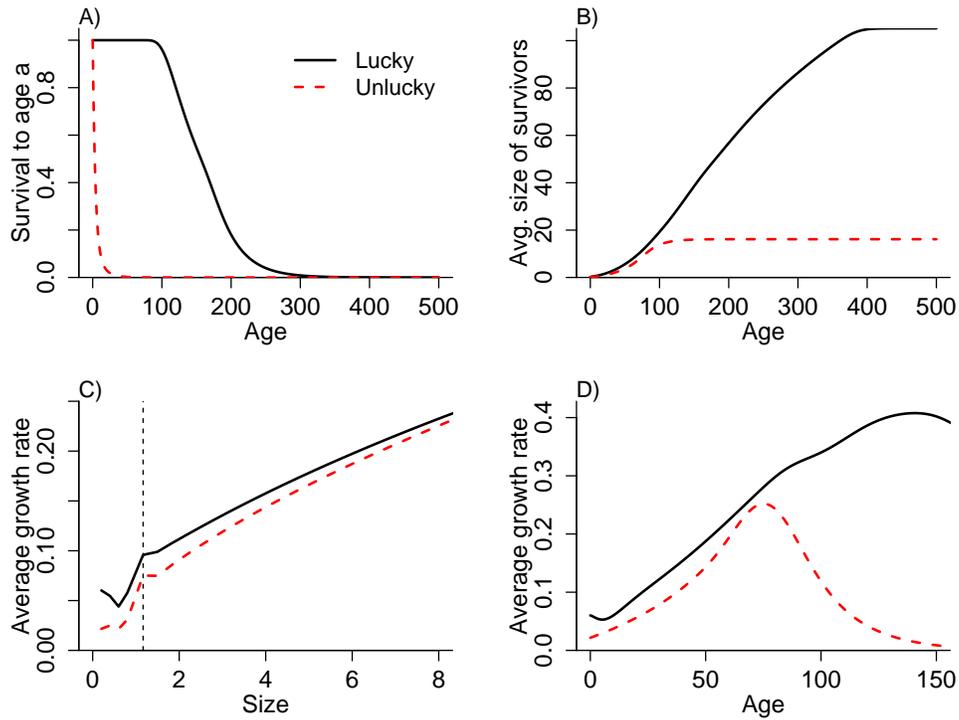


Figure 4: A) Survival to age a , $l(a)$, for Lucky (solid black curve) and Unlucky trees (dashed red curve) in the Zuidema et al. (2010) IPM for *Dacrydium*. B) Average size as a function of age. C) Average growth rate as a function of size. The vertical dashed line marks the end of the seedling stages. D) Average growth rate as a function of age. For these calculations, seedlings (0-1 cm dbh, classified into four height classes) are treated as having sizes 0.2, 0.4, 0.6, 0.8 cm dbh. Source file: ZuidemaLucky20-compareGrowth.R and scripts that it sources.

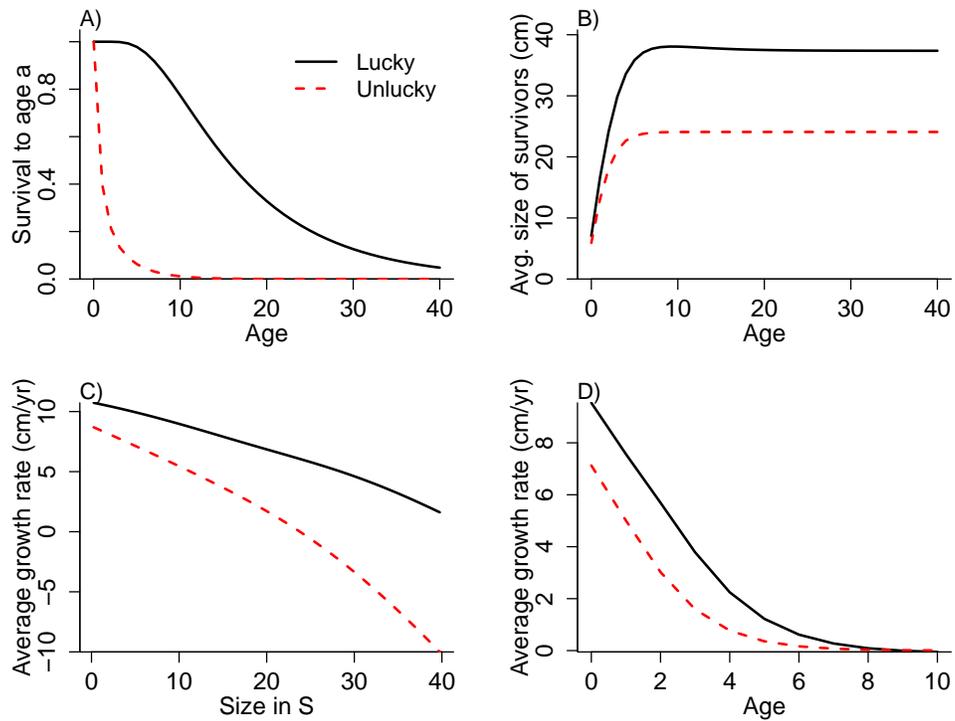


Figure 5: A) Survival to age a , $l(a)$, for Lucky (solid black curve) and Unlucky shrubs (dashed red curve) in the Li et al. (2011) IPM for *Artemisia* described in the text. B) Average size as a function of age. C) Average growth rate as a function of size, for individuals that have not yet crossed the threshold (40cm) for being Lucky. D) Average growth rate as a function of age. Source file: `ArtemisiaLucky40-compareGrowth.R` and scripts that it sources.

The faster growth of Lucky seedlings also contributes to their higher survival, but not by much. Survivorship through the seedling stages in the model is 2.2% for all individuals combined. If we give seedlings the faster stage-specific growth rate of Lucky seedlings, without changing their stage-specific survival, survival through the seedling stages only increases to 12% – whereas for truly Lucky seedlings it has to be 100%. So even in the seedling stages where the growth differential is largest, rapid growth is less important than simply surviving.

Artemisia: Because growth is less deterministic in *Artemisia*, differences in growth play a larger role in becoming Lucky. Not dying while small is still essential (Figure 5A), and mortality in the Lucky subpopulation then remains low because they have gotten past the high mortality at small sizes. However, by age 5 the Lucky are more than 50% larger than the Unlucky (Fig. 5B), and the difference in size-specific average growth rate (prior to reaching the threshold) is substantial (Fig. 5C). For example, 24cm or larger Unlucky shrubs shrink on average, while the expected growth of a 24cm Lucky shrub is 6cm/yr. The average size of Lucky shrubs never exceeds the 40cm threshold (Fig. 5B) – getting past the threshold is indeed a matter of luck, and it is typically followed by shrinking below the threshold. By age 10, surviving shrubs have settled into the stationary size distribution conditional on their fate, and the average growth rate is zero (Fig. 5D).

Cedrela: As always, there is high mortality when small, and only a tiny fraction reach reproductive maturity (size classes 6 or above, Fig. 6A), so that survival differs substantially between Lucky and Unlucky (Fig. 6B). But there is also a large difference in growth (Fig. 6C) and a consistent difference in growth rates (Fig. 6D) of Lucky and Unlucky trees. This agrees with the conclusions of Zuidema et al. (2009) about *Cedrela*: the trees that dominate the reproductive input to future generations are consistently fast-growing individuals.

Elasticity of the probability of becoming Lucky (q_L)

Dacrydium: Up to about size 8 cm dbh, expected growth has the most influence on the probability of becoming Lucky/large (Fig. 7A). This is because increased growth pushes individuals through

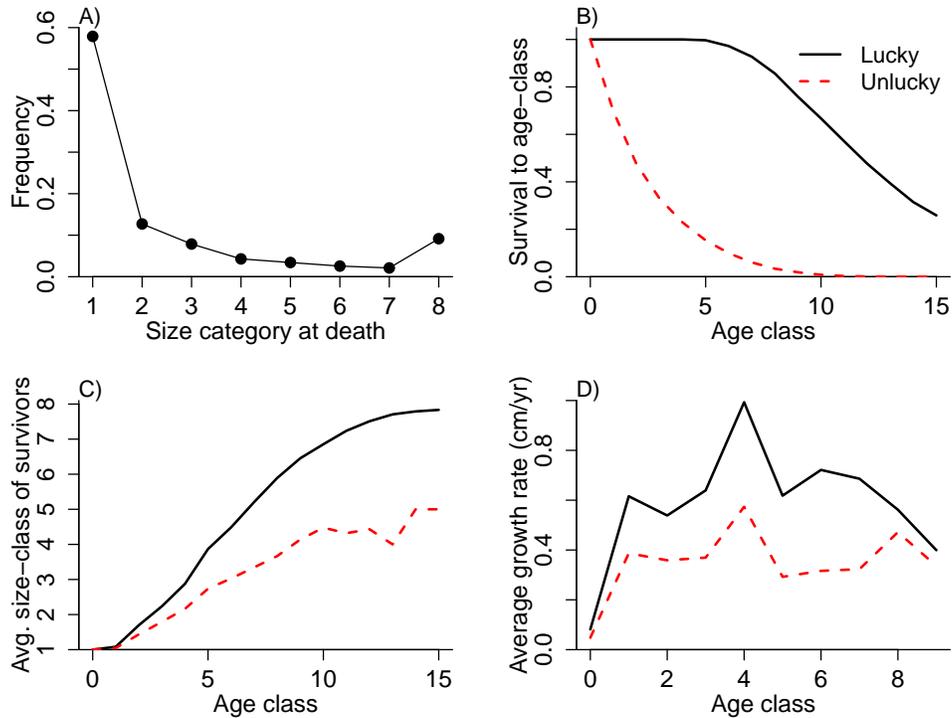


Figure 6: Comparison based on conditional kernels for the Zuidema et al. (2009) matrix model for *Cedrela odorata* with size-by-age classification and Lucky defined as reaching size-class 6. A) Frequency distribution of size-class at death. Note that size-class 8 is a “big trees” category (70cm dbh or larger). B) Survival to age-class a , $l(a)$, for Lucky (solid black curve) and Unlucky trees (dashed red curve). C) Average size-class as a function of age-class for Lucky and Unlucky. D) Average growth rate as a function of age-class for Lucky and Unlucky trees. The plotted growth rates start to level off at age-class 9 because many trees have reached the maximum possible size-class (class 8 for Lucky, class 5 for Unlucky). Source files: `CedrelaMatrices.R` and `CedrelaLucky.R`

the smaller, higher mortality sizes faster. Once they reach 10 cm dbh, they are relatively safe, with only a 1.3% chance of dying per year until they reach 45 cm dbh. This also explains the kink in the elasticity of expected growth at size 10: the mortality rate for *Dacrydium* is piecewise linear with a kink at that size. At larger sizes, survival comes to have the greatest influence on becoming Lucky. Once individuals are past the vulnerable smaller sizes, becoming Lucky is a matter of surviving long enough.

The sharp peak in elasticities at small sizes mirrors the initial sapling size distribution. (We are calculating elasticities for the sapling IPM portion of the model only, not the seedling classes, and we assume that individuals leaving the largest seedling stage have a Gaussian distribution of size as a new sapling.) Perturbations at these sizes are especially important because everyone passes through these sizes—they cannot be skipped over between one census and the next.

Examining elasticity vs. age (Fig. 7B), we see that a little luck at young ages counts more than luck at older ages. This provides another reason for why perturbations at small sizes have a large influence on the probability of becoming Lucky: small sizes correspond to young ages. Of course, since we are only perturbing sapling demographic rates, these perturbations only matter at ages where a substantial fraction of the population has reached the sapling stages. Perturbations to expected growth matter most until just past age 50, at which point perturbations to survival matter more.

Artemisia: For *Artemisia*, expected growth has the highest influence on becoming Lucky at all sizes (Fig. 8A). This is unsurprising given that surviving a long time is not sufficient to become Lucky/large for *Artemisia*—above-average growth is also required. Survival has the second-highest influence on becoming Lucky, with growth variability ($sd(\text{growth})$) a distant third. As with *Dacrydium*, there is a hump at small sizes that mirrors the birth size distribution, and as with *Dacrydium*, luck at these sizes is important because all individuals pass through these sizes. There is then a dip around size 11 cm before a second, broader peak. The dip occurs because individuals tend to jump over this size interval between the first and second year censuses. For example, if a shrub is 8 cm in its first year, its expected size the following year is 14.67 cm. The

probability of becoming Lucky also depends on size at birth, but the elasticity to size at birth is very small compared to those for expected growth and survival (results not shown; see Figure D1 in Online Appendix D, SI script ArtemisiaBirthSizeElasticity.R).

Examining elasticities versus age (Fig. 8B), luck in the first year is especially influential—it helps to be born at the right time. However, elasticity for expected growth peaks at age 2. Two-year-old shrubs are $\approx 20\text{cm}$ tall on average, and can expect to grow about 5cm in the coming year, if they survive. That is still small enough that better growth now has the added benefit of higher survivorship in the future (see Fig. 3A). Younger shrubs also get about the same absolute boost in survival, but the impact is attenuated because of mortality before age 2. After age 2, growth starts to slow down (so a 5% increase in growth is a smaller change in absolute terms) and survival begins to plateau as a function of size at the mean size of individuals, so the elasticity starts to drop. The elasticity function shows that age 2 is the “sweet spot” in the balance between a decreasing survival gain from added growth, and an increasing chance that the beneficiary will survive long enough to become Lucky.

The *Cedrela* matrix model is not parameterized in terms of growth rates. However, we can compare the effects of perturbations to survival at different ages, and at different sizes. Not surprisingly, the highest elasticities are for the survival of younger and smaller individuals (Figure 9).

Models structured by size and cumulative breeding

Defining the “Lucky few” in terms of a size threshold has the advantage of simplicity. But size is an indirect measure of reproductive success and may not always correlate well with lifetime reproductive output. We therefore extend our methods so that Lucky is instead defined directly in terms of lifetime reproductive output.

In concept, this change takes almost no effort. Instead of classifying individuals by size, cross-classify them by size x and some measure of total reproductive success to date, T . We will call this

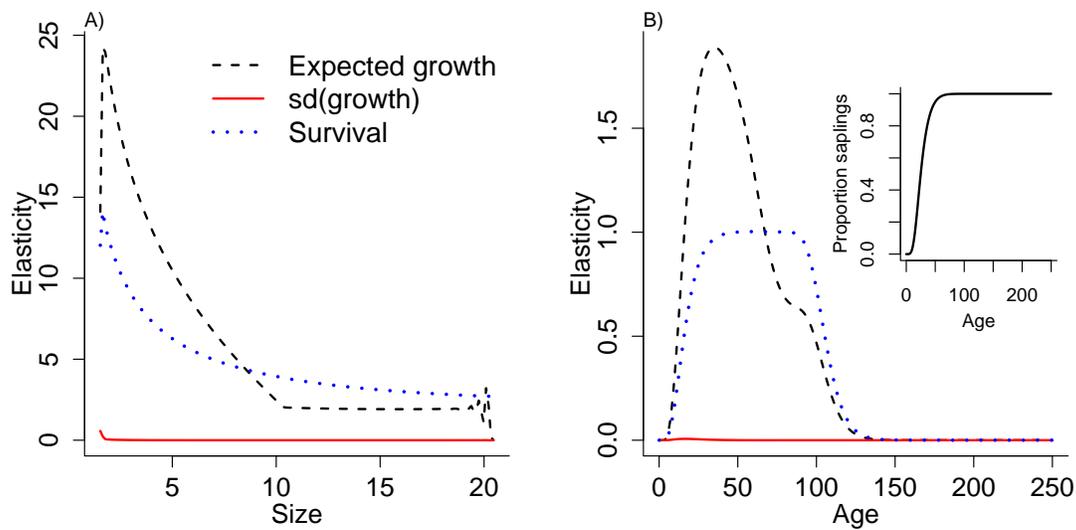


Figure 7: *Dacrydium elatum*: elasticity of the probability of a tree reaching 20 cm dbh to A) Perturbations in expected growth, the standard deviation of growth, and survival at a given size, B) The same perturbations at a given age. Based on the Zuidema et al. (2010) IPM for *Dacrydium*. Note that the IPM only covers the sapling stages — seedling dynamics are governed by a Leslie matrix — so individuals must reach the sapling stages before the perturbations can have an effect. Source file: `DacrydiumElasticities.R` and the scripts that it sources.

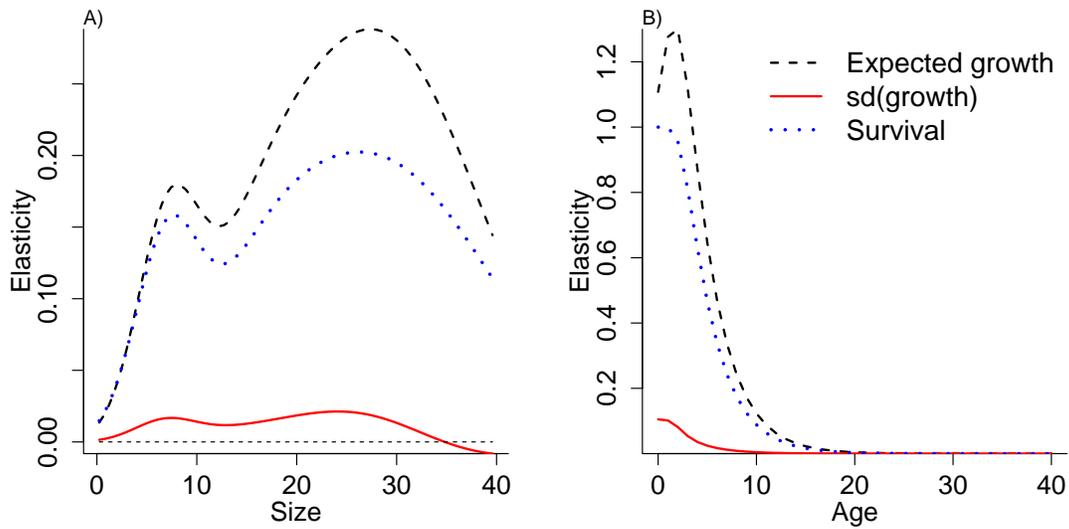


Figure 8: *Artemisia ordosica*: elasticity of the probability of reaching 40 cm height to A) Perturbations in expected growth, the standard deviation of growth, and survival at a given size, B) The same perturbations at a given age, in the Li et al. (2011) IPM for *Artemisia*. Source file: *ArtemisiaElasticities.R* and the scripts that it sources.

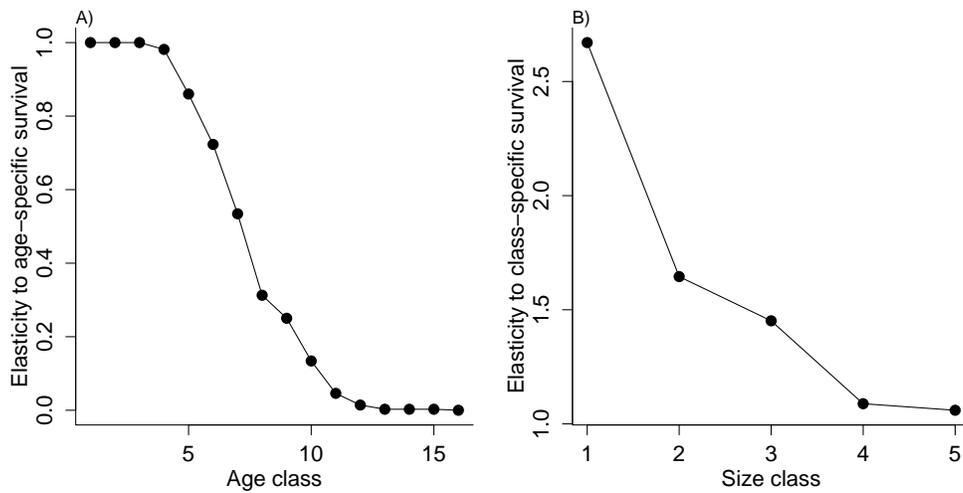


Figure 9: *Cedrela odorata*: elasticity of the probability of reaching size-class 6 to A) survival in different age-classes, and B) survival in different size-classes, for the Zuidema et al. (2009) matrix model for *Cedrela*. Source files: *CedrelaMatrices.R* and *CedrelaPerturb.R*

a *size*×*kids* model. Our general analysis still applies to $z = (x, T)$, with the formulas properly interpreted. Here we take T to be discrete (number of offspring, or the number of years in which flowering occurs), so the underlying model is a general IPM (Ellner and Rees, 2006) with continuous×discrete classification, structurally like an size-age model or a size-quality model with a discrete quality variable. The full IPM iteration is

$$n(x', T', t + 1) = \sum_{T=0}^{m_T} \int_L^U K(x', T', x, T) dx \quad (29)$$

where m_T is the (necessarily) finite upper limit on possible T values. The individual-level Markov chain that we focus on here is the same with $P(x', T', x, T)$ in place of K . When offspring count can be very large, however, a continuous measure might be more appropriate (e.g., $T = \log$ of total seed production to date). Either way, care must be taken to make sure that no errors are introduced by the finite upper limit on T . One approach is to make m_T so large that no individual ever comes close to it. Here, we instead modify the transition probabilities so that an individual with $T = m_T$ remains there. This makes m_T a fecundity “class” representing individuals with m_T or more offspring. Specifically, we set $P(x', T, x, m_T) = 0$ for $T \neq m_T$ and $P(x', m_T, x, m_T) = s(x)G(x', x)$ where s is survival and G is the growth kernel for survivors.

In a *size*×*kids* model, we can define an individual as Lucky if their value of T at death is above some threshold T_c . Because T can never go down, any set of states $\{(x, T) : T \geq T_c\}$ is absorbing; any individual entering that set will die Lucky. We therefore take $T \geq T_c$ as the definition of the “big” or Lucky states B .

A suitable value of T_c can be identified from the distribution of T at death. State at death is a random quantity described by the kernel

$$\Omega(z', z) = (1 - s(z'))N(z', z) \quad (30)$$

where N is again the fundamental operator and $s(z') = \int_A P(u, z') du$ is the survival probability (Ellner et al., 2015, Chapter 3). That is, for an individual currently at state z , $\int_X \Omega(z', z) dz'$ is the probability that the individual’s state at the last census prior to death is in the set X . So if $c_0(z)$

is the state distribution for a cohort at birth,

$$\Omega c_0 \equiv \int_A \Omega(z, z_0) c_0(z_0) dz_0 \quad (31)$$

is the distribution of state at death for the cohort. In a size \times kids model $z = (x, T)$ and Ωc_0 is a function of x and T . The distribution of total reproductive output T at death is

$$\omega(T) = \int_{\mathbb{L}}^U (\Omega c_0)(x, T) dx \quad (32)$$

where L, U are the lower and upper limits of the model's size range. $\omega(T)$ is a univariate distribution so it is straightforward to numerically compute (for example) the 95th percentile of $\omega(T)$ and take this as T_c .

Comparison based on conditional kernels

Here we use the *Dacrydium* IPM to illustrate the calculations for a size \times kids model. Very similar results for *Artemisia* are shown in online figures D2 and D3.

Tree size affects the probability of flowering but not the expected seed production of a flowering tree in the model, so we let T be the number of times that a tree has flowered to date. Counting flowering times rather than seedlings is helpful for computations because it requires a smaller iteration matrix (and “smaller” for the calculations below was a matrix with over 10^8 entries). We can also verify that it is nearly equivalent to counting seedlings, by finding the fraction of variation in lifetime reproductive output that is due to variation in the number of flowering times. To do this, we compute the standard deviation in lifetime seedling production for a new sapling (Ellner et al., 2015, Chapter 3), and then re-compute the standard deviation when each flowering event produces exactly the expected number of seedlings (the calculations are in the online Supplement script `BreedingVarCompare.R`). For the *Dacrydium* IPM, assuming that the seedling yield from a flowering event has a Poisson distribution (i.e., many seeds with a low probability of surviving to be a seedling), variation in the number of flowering times accounts for $\approx 99\%$ of the variation in lifetime reproductive output. The reason for this surprising

outcome is that the Lucky flower very many times (see below), having reached a size where they flower every year and can expect to live 40-50 more years – variability in annual seedling yield averages out over many flowering events. Similarly, in the *Cedrela* matrix model, variation in the number of breeding times accounts for $\approx 97.5\%$ of variation in total reproductive output. In the *Artemisia* IPM, where a Lucky shrub can expect to live only another 10-11 years, variation in the number of flowering times still accounts for $\approx 90\%$ of the variation in lifetime total reproductive output. For *Artemisia* we therefore define Lucky in terms of seedling production rather than flowering times. In general, for short-lived species, it will usually be necessary to count offspring rather than breeding events.

For *Dacrydium*, the marginal distributions of size and breeding times T at death are both bimodal (Figure 10). The larger mode (over 99% of seedlings) is dying young and small without ever flowering. The second mode is at $\approx 50\text{cm}$ dbh, having flowered many times; these are the Lucky (very) few. Based on the distribution of T , we define the Lucky as those who reach $T = 20$; this is somewhat arbitrary, but it picks out those who reach the upper tail.

Figure 11 shows how the probability of reaching $T = 20$ depends on current size and the number of flowering times to date. Individuals smaller than 10cm dbh cannot flower in this model, so they all have $T = 0$; after that, survivors can start to flower and T will increase. The contours for probability 0.7 or below are nearly perpendicular to the size axis, saying that the probability of becoming Lucky is mostly a function of size. This supports the simplifying approach of defining the Lucky just in terms of size. The size threshold that we used in that approach (20cm) corresponds roughly to a 70% chance of flowering 20 or more times before dying.

Figure 12 is exactly comparable to Figure 4 except that it is based on the $\text{size} \times T$ cross-classified model, with $T = 20$ flowering times as the criterion for being Lucky. The results are nearly identical, presumably because size is such a good predictor of the probability of reaching $T = 20$ (Figure 11). Individuals of size 20cm dbh typically have $T = 5$ to 10 (see Figure D4 in Appendix D) so the criterion defining the Lucky in Figure 12 is more stringent. But it is

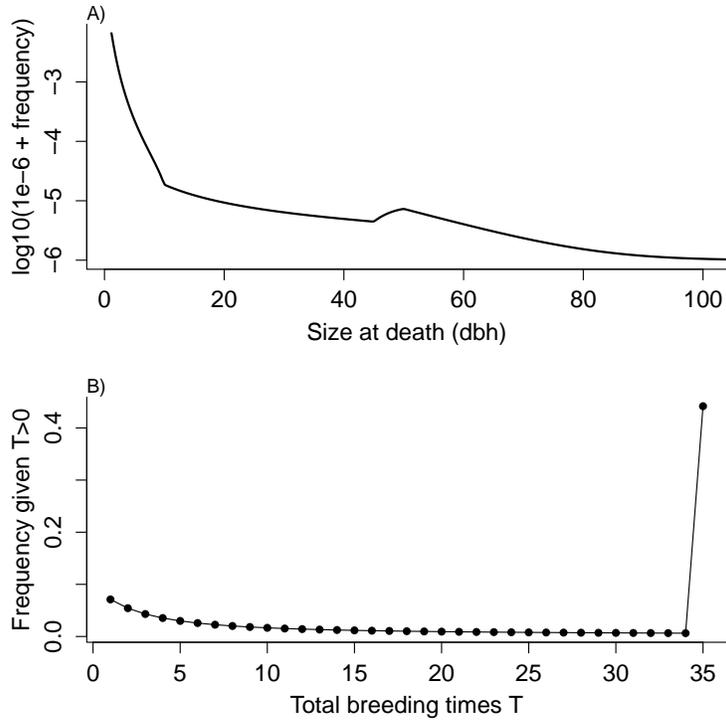


Figure 10: Marginal distributions of A) size and B) total number of flowering times T at death, in the *Dacrydium* IPM with individuals cross-classified by size and T . The marginal distributions were calculated from the bivariate death state distribution Ω for a cohort of individuals born into the smallest seedling class with $T = 0$. Note that the size distribution is plotted on a log-transformed scale so that the second mode at ≈ 50 cm dbh can be seen. The plotted distribution of T (panel B) is the conditional distribution given $T > 0$, because over 99% of individuals die before reaching the minimum size for flowering (10cm dbh); the large value at $T = 35$ is because the model was implemented with $T = 35$ as a class representing all $T \geq 35$. Source files: `ZuidemaCompareLucky-xT.R` and scripts that it sources; `MegamatrixFunctions.R`

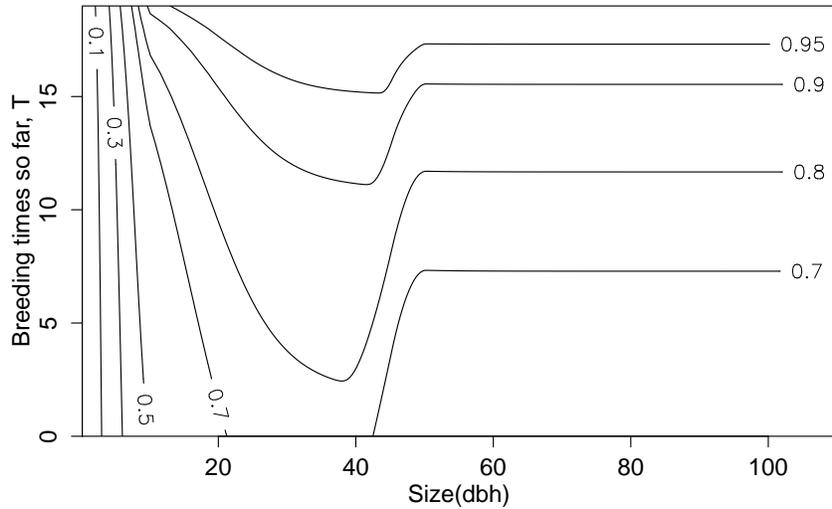


Figure 11: Contour plot of the probability of becoming Lucky ($T \geq 20$ lifetime flowering times) as a function of current size and T , on the set of small states S , in the Zuidema et al. (2010) IPM for *Dacrydium* with size \times T cross-classification and Lucky defined as reaching $T = 20$. Source files: ZuidemaCompareLucky-xT.R and scripts that it sources; MegamatrixFunctions.R

not much more stringent, because a typical 20cm dbh individual has 70-80% odds of reaching $T = 20$. Thus, being one of the Lucky few in *Dacrydium* is primarily a matter of reaching a critical size, and getting to that size is mostly a matter of living long enough, with rapid growth being somewhat important only in the earliest years when it provides an escape from high mortality. In the *Artemisia* model, because growth is more variable, current size is a less reliable predictor of future prospects, and there is more difference between results for the size- and reproduction-based definitions of Lucky. In particular, when Lucky is defined by lifetime seedling production, there is a smaller difference in growth between the Lucky and Unlucky than when Lucky is defined by a size threshold (see online Figure D3).

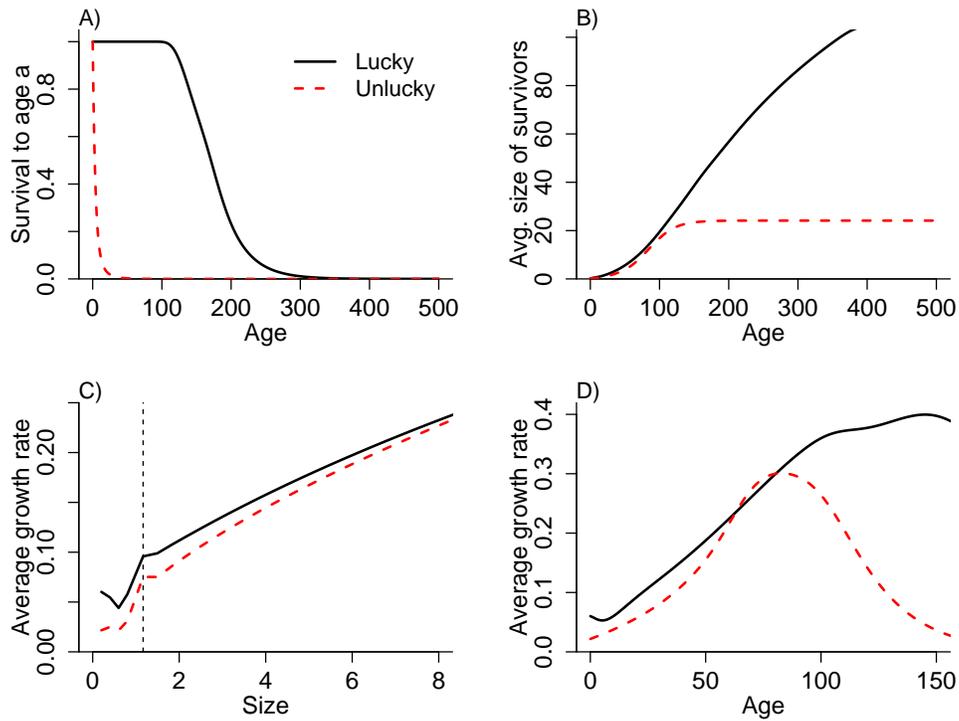


Figure 12: A) Survival to age a , $l(a)$, for Lucky (solid black curve) and Unlucky trees (dashed red curve) in the Zuidema et al. (2010) IPM for *Dacrydium* with size \times T cross-classification and Lucky defined as reaching $T = 20$. B) Average size as a function of age. C) Average growth rate as a function of size. The vertical dashed line marks the end of the seedling stages. D) Average growth rate as a function of age. For these calculations, seedlings (0-1 cm dbh, classified into four height classes) are treated as having sizes 0.2, 0.4, 0.6, 0.8 cm dbh. Source files: ZuidemaCompareLucky-xT.R and scripts that it sources.

Discussion

We have presented several ways of investigating how exceptional (“Lucky”) individuals came to be that way. The first is to compare life trajectories by calculating kernels conditional on becoming Lucky or remaining Unlucky and using these to calculate expected life history aspects for the two groups. For example, we can calculate average growth rate versus age for Lucky and Unlucky individuals in a size-structured model. In addition, we can calculate the sensitivity or elasticity of the probability of becoming Lucky to a small increase in a vital rate at a particular size (regardless of age), or an increase in a vital rate at a particular age. While we have defined Lucky individuals as those who have become exceptionally large or have had exceptionally high levels of reproduction, these methods can be used for any definition of “Lucky” that can be expressed as a set of values for individual state.

We used two IPMs as examples, for the tropical tree *Dacrydium elatum* and for the semi-arid shrub *Artemisia ordosica*, and a matrix model for the tropical tree *Cedrela odorata*. We found that in all cases, becoming exceptionally large (or breeding exceptionally often) is first and foremost a matter of living long enough. For the shrub it is also a matter of growing fast, while for *Dacrydium* the Lucky and Unlucky differ little in growth rates except at the very smallest sizes (the height-classified seedling stages). In the *Cedrela* model, individuals with fast growth initially continue to have fast growth, and as expected, this faster growth also contributes to becoming Lucky. In all cases, luck at small sizes/young ages matters more than luck at large sizes/older ages. These conclusions, of course, are contingent on the validity of the published models.

In this paper we have considered only *individual stochasticity* (Caswell, 2009): all individuals have the same transition probabilities for growth, survival, etc., so that differences between one life course and another are purely due to chance (Tuljapurkar et al. (2009) and Steiner et al. (2010) use “dynamic heterogeneity” for this kind of variation among individuals). This is true even in the *Cedrela* model, where growth rate differences persist over time, but they arise initially by chance.

However, it is also important to consider situations in which some individuals have consistently higher growth, survival, etc. than others; Tuljapurkar et al. (2009) call this “fixed heterogeneity”. Fixed differences add a deterministic component to the chance events that determine which individuals get Lucky in the models we have considered here. The differences could be due to some genetic or phenotypic difference between individuals (“intraspecific trait variation”) that persists throughout their lifetime, or a permanent feature of their habitat (e.g., a plant growing in a favorable soil type). Other differences could change gradually within an individual’s lifetime (e.g., different growth rates due to differences in light environment).

The approaches we present here can be extended to accommodate these persistent between-individual differences, by using a cross-classified individual state. Evolutionary models have cross-classified individuals by size (individual stochasticity/dynamic heterogeneity) and a heritable trait (fixed heterogeneity) that is effectively their genotype or breeding value for a quantitative trait (e.g., Bruno et al., 2011; Coulson et al., 2011; Rees et al., 2006; Rees and Ellner, 2015; Rees et al., 1999; Rose et al., 2002; Vindenes and Langangen, 2015). Metcalf et al. (2009) cross-classified trees by size and by their light environment; Gap dynamics determine transitions between light environments, while transitions between sizes depend on light environment. While the megamatrix governing transitions in a cross-classified model will often be large, if environmental transitions do not depend on the state variable, structure in the megamatrix can be exploited to simplify calculations. If differences between individuals are permanent, the transition kernel governing post-birth transitions will be block diagonal. The megamatrix for transitions in a time-varying environment (as in Metcalf et al. (2009)) can be written as the Kronecker product of the transition matrix for the environments and the transition matrix for the size transitions. Extending our approaches to these between-year and between-individual differences will be the subject of a future manuscript.

Of course, we often don’t know whether persistent individual differences are present or not. If some individuals are permanently better in ways that make them more likely to be among the few that dominate reproduction — persistently fast growth, for example — this can change

population-level quantities such as population growth rate and generation time, not to mention producing selection if the trait can be passed on to offspring. However, as Plard et al. (2015) have pointed out, if the effect of the trait can be approximated by the effect of size or some other variable that is accounted for by the IPM, then including the trait value as part of an individual's state may have only minor effects. Several studies have used models without fixed heterogeneity between individuals as null models to test for the presence of such differences (e.g. (Orzack et al., 2011; Plard et al., 2012; Steiner et al., 2010; Tuljapurkar et al., 2009)). Bonnet and Postma (2016) have found that this null model approach lacks statistical power. However, this may depend on the life history of the species, and on which observations are compared with null model predictions. The predictions developed in this paper can be used as additional, and possibly more powerful properties for testing null models, such as the mean and variance of age- or size-specific growth in individuals with low versus high lifetime total reproductive success.

The events or factors that lead to exceptional reproductive success are presumably strong targets for selection, so identifying the decisive events or factors should contribute to understanding life history evolution – when survival is what matters, we should find that organisms have evolved to maximize survival at the expense of rapid growth or early breeding. And just knowing how strongly reproduction is skewed can be important for understanding microevolutionary properties (e.g., genetic diversity, effective population size, and the effective generation time for response to selection). Our methods make it possible to partition variation in lifetime reproduction or other whole-life outcomes into contributions from different demographic processes (e.g., differences due to size at birth, due to growth at different ages or stages, etc.), and (when there are fixed differences among individuals) partitioning the total variation into effects of fixed differences and variation arising by chance.

Structured population models such as IPMs are not just population models. They are also models for the lives of individuals (Caswell, 2009; Cochran and Ellner, 1992; Metcalf et al., 2009; Tuljapurkar et al., 2009). We can extract from those models far more information than we have been, about the different paths individuals take through life, and about what events and factors

cause one individual's life to be very different from another's. The methods and results in this paper are one small step in this direction.

Acknowledgments

This project resulted from seemingly unrelated questions about IPMs from Pieter Zuidema and Hal Caswell: how can you do loop analysis (PZ), and how can you calculate the size-at-death distribution (HC)? Dylan Childs taught us essential R tools for working with cross-classified IPMs and constructing integrated bin-to-bin iteration matrices. We thank Eelke Jongejans and Hal Caswell for discussions about this work, Hal Caswell for sharing unpublished work on individual stochasticity, and Peter Adler, Collin Edwards, Giles Hooker, Hidetoshi Inamine, Irena Papst, Britta Teller and the anonymous reviewers for comments on the manuscript. Research support was provided by US NSF grants DEB-1353039 (SPE) and DEB-1354041 (RES).

Online Appendix A: Additional Derivations

Equation (6) for general state space chains

As in Ellner and Rees (2006) the state space A is most generally a union of compact metric spaces (e.g., two copies of an interval $[L, U] \subset \mathbb{R}$, one each for males and females) and the kernel is defined relative to a Borel measure μ_j on component space j . Here and in most other applications this generality isn't needed, so we write dz instead of $\mu(dz)$ or $d\mu(z)$.

$q_{\mathbb{L}}(z)$ denotes the probability of entering B (and therefore being absorbed there) before death given current state z . That is,

$$q_{\mathbb{L}}(z) = \Pr[z_t \in B \text{ or } z_{t+1} \in B \text{ or } z_{t+2} \in B \text{ or } \dots | z_t = z].$$

Note $q_{\mathbb{L}} \equiv 1$ on B . Then if the state at time t is $z_t = z \in S$, for any measurable set X

$$\Pr[z_{t+1} \in X | \text{enter } B] = \frac{\Pr[z_{t+1} \in X \text{ and enter } B]}{\Pr[\text{enter } B]} = \frac{\Pr[z_{t+1} \in X \text{ and enter } B]}{q_{\mathbb{L}}(z)}$$

By the Markov property, the probability of entering B at $t + 1$ or later, given the state at $t + 1$, depends only on the state at $t + 1$, so the numerator is $\int_X q_{\mathbb{L}}(z')P(z', z)dz'$.

$$\Pr[z_{t+1} \in X | \text{enter } B] = \int_X [P(z', z)q_{\mathbb{L}}(z') / q_{\mathbb{L}}(z)] dz'.$$

The density of z_{t+1} conditional on $z_t = z$ and absorption into B is therefore $P(z', z)q_{\mathbb{L}}(z') / q_{\mathbb{L}}(z)$.

This implies that the state conditional on absorption into B is a Markov chain with transition kernel $P(z', z)q_{\mathbb{L}}(z') / q_{\mathbb{L}}(z)$ and the transient kernel (describing transitions prior to absorption into B) is $Q(z', z)q_{\mathbb{L}}(z') / q_{\mathbb{L}}(z)$.

Comparing life trajectories in a size-structured model with shrinkage below z_c

As noted in the main text, for this case we need to create an extended state space $S \cup B$ where $S = [L, z_c)$, $B = [L, U]$ considered as two separate sets. All individuals are born into S (unless their initial size is above z_c , in which case they are already Lucky at birth). The interval $[L, z_c)$ in B represents individuals who entered $[z_c, U]$ and then shrank below z_c .

Define $\mathbf{Z}_1 = [L, z_c), \mathbf{Z}_2 = [z_c, U]$ considered as two subsets of the original state space $[L, U]$. Schematically, the transition kernel on the enlarged state space is

$$\tilde{P} = \begin{matrix} & S & & B \\ \begin{matrix} S \\ \\ B \end{matrix} & \begin{pmatrix} P(\mathbf{Z}_1, \mathbf{Z}_1) & \vdots & \mathbf{0} \\ \dots\dots\dots & \vdots & \dots\dots\dots \\ \mathbf{0} & \vdots & P(B, B) \\ P(\mathbf{Z}_2, \mathbf{Z}_1) & \vdots & \end{pmatrix} \end{matrix}. \quad (\text{A1})$$

The zero-matrix in the left column has the size of $P(\mathbf{Z}_1, \mathbf{Z}_1)$ and is there because individuals in S who move to B do so by growing to a size above z_c . The zero-matrix in the right column has the size of $P(\mathbf{Z}_1, B)$ and is there because individuals in B stay there regardless of how their size changes.

We can now apply the general machinery to \tilde{P} to compute the conditional kernels. The two absorbing states are $\mathcal{A}_1 = \text{“die from } S\text{”}$ (Unlucky) and $\mathcal{A}_2 = \text{“die from } B\text{”}$ (Lucky). The probability of immediate absorption into \mathcal{A}_1 from $z \in S$ is the mortality rate, $a_1(z) = 1 - s(z)$. Because the Unlucky never leave S we only have to compute their conditional kernel on S . From (5), the probability of absorption into \mathcal{A}_1 from $z \in S$ is

$$q_1 = a_1(I - Q)^{-1} \quad (\text{A2})$$

where Q is again P restricted to S , and $q_1 \equiv 0$ on B . The conditional kernel for the Unlucky is then

$$Q_U(z', z) = Q(z', z)q_1(z')/q_1(z), \quad z', z \in S. \quad (\text{A3})$$

This is exactly the same as when shrinkage is impossible.

Because all individuals absorb into \mathcal{A}_1 or \mathcal{A}_2 , the probability of absorption into \mathcal{A}_2 is

$$q_2(z) = \begin{cases} 1 - q_1(z), & z \in \mathbf{Z}_1 \\ 1, & z \in \mathbf{Z}_2 \end{cases}.$$

The conditional kernel is therefore

$$P_{\mathbb{L}}(z', z) = \tilde{P}(z', z)q_2(z')/q_2(z) \quad (\text{A4})$$

In evaluating the two equations above numerically, care must be taken because small sizes occur in both S and B . If size $z = 3$ is small, there are two “copies” $3_1 \in S$ and $3_2 \in B$, and $q_2(3_1) \neq q_2(3_2)$.

Online Appendix B: Megamatrix calculations and size \times kids IPMs

A “size \times kids IPM” means that individuals are cross-classified by their current size x and by T , their total number of offspring (or total number of breeding times) from birth to the present. We used bin-to-bin integration (Dawson, 2013; Ellner et al., 2015, Chapter 6) to construct the iteration matrices for numerical calculations, with 7-point Gauss-Legendre quadrature for initial state and cumulative kernel for final state (i.e., exact integration using the Gaussian cumulative distribution function as implemented in R by `pnorm`). Like midpoint rule, bin-to-bin uses meshpoints

$$x_i = L + (i - 0.5)h, \quad i = 1, 2, \dots, m_x, \quad h_x = (U - L)/m_x.$$

The numerical iteration for the individual-level Markov chain is then

$$n(x_k, T_l, t + 1) = \sum_{i=1}^{m_x} \sum_{j=1}^{m_T} P(x_k, T_l, x_i, T_j) n(x_i, T_j, t) \quad (\text{B1})$$

The iteration matrix P in the equation above includes both size transitions and fecundity, and it includes the integration with respect to final state (which is why the h_x , that would be present in a midpoint rule implementation, is absent).

In our scripts P is calculated as the product of the size-transition iteration matrix $P_x(x_k, x_i)$, with the probability of having $T_l - T_j$ offspring conditional on initial size x (or the probability of breeding $T_l - T_j$ times, in which case the only possible values of $T_l - T_j$ are 0 and 1). In the *Dacrydium* model reproduction is conditional on survival, so the offspring number distribution is the conditional distribution given that survival occurs. The published *Dacrydium* and *Artemisia* IPMs only give information on the breeding (i.e. flowering) probability and the average seedling production given that flowering occurs. We assumed a Poisson distribution of seedlings, conditional on the mean, which can be rationalized by assuming that seedlings are the result of producing many seeds with a low probability of establishment as seedlings.

It is natural to think of n as a “state matrix” where column j describes the size distribution of individuals with $T = j - 1$, and to think of the transition probabilities as a 4-dimensional array laid out like P in (B1). However, in interpreted matrix languages such as R or MATLAB calculations

are much faster when we implement them as a large 2-dimensional iteration matrix multiplying a large state vector, especially in a version of R linked to a multithreaded linear algebra library (such as CRAN R for OS X, or Revolution R Open with MKL installed). Fortunately, R makes it very easy to move back and forth between the 4-D and 2-D layouts by manipulating the `dim` attribute of a matrix or array, as follows.

The state vector is created by stacking the columns of the state matrix: column 1 first, column 2 below it, and so on. Assuming m mesh points for size and $T = 0, 1, 2, \dots$, the state vector is structured as

$$\begin{array}{l}
 n(x_1, 0) \\
 \vdots \\
 n(x_m, 0) \\
 n(x_1, 1) \\
 \vdots \\
 n(x_m, 1) \\
 n(x_1, 2) \\
 \vdots
 \end{array} \tag{B2}$$

This is the “megamatrix” layout. The first m states are all sizes and “quality”(T) level 1. The next m states are all sizes and the next “quality” level. The corresponding iteration matrix for \mathbf{n} then is structured as an matrix of “blocks” (each of size $m \times m$) in which T is constant. Diagonal blocks represent transitions in which T remains constant; off-diagonal blocks are ones where T changes.

Going from a state matrix n to a state vector \mathbf{n} is called the “vec” operation: $\mathbf{n} = \text{vec}(n)$. In R, `vec` can be done with the `matrix` function,

```
vec <- function(nmat) matrix(nmat,ncol=1)
```

Going back to the state matrix is also done with `matrix`,

```
unvec <- function(nvec,nrow=NULL,ncol=NULL){
```

```

if(is.null(nrow)) return(matrix(nvec,ncol=ncol));
if(is.null(ncol)) return(matrix(nvec,nrow=nrow));
}

```

vec works automatically, but unvec needs to “know” the number of rows or number of columns in the state matrix.

Next we need the iteration matrix that acts correctly on \mathbf{n} . Let A be a four-dimensional array where $A[i, j, k, l]$ holds the value of $h_x A(x_k, T_l, x_i, T_j)$, where $A = K$ or P . Because of the order in which R fills in matrices by default (see Ellner et al. (2015, Chapter 6) for details), converting this to the iteration matrix on \mathbf{n} is just

```
dim(A) <- rep(sqrt(length(A)),2)
```

Converting the other way, from iteration matrix to four-dimensional array, can also be done just by changing the `dims` attribute. As this involves no copying or moving of data in memory, it’s very fast. Skeptics are invited to run the code in Table B1.

The only real challenge is to construct A efficiently. Because A is four-dimensional, vectorizing is essential. In our applications it is enough to vectorize along one size dimension. The functions in `ZuidemaTreeFunctions.R` and `ShrubFunctions.R` compute $A[, l, i, j]$ as a vector and fill in A using nested loops in l, i, j . This can take a minute or two, but it only needs to be done once. If the kernel function vectorizes in all of its arguments then iteration matrices can be constructed very quickly without loops Ellner et al. (2015, Chapter 6), but that is not the case for the models we consider here.

For the comparisons based on conditional kernels, we need to construct state vectors and iteration matrices on the set S of small states defined by $T < T_c$. This is easy to do in the iteration array/state matrix format, by subsetting array and matrix dimensions corresponding to T . The subsetted arrays and matrices can then be converted back to the iteration matrix and state vector on S , using `unvec` and `unfold` with the reduced length in the T dimensions.

```

### Turn state matrix to state vector
vec <- function(nmat) matrix(nmat,ncol=1)

### Turn state vector to state matrix
unvec <- function(nvec,nrow=NULL,ncol=NULL){
  if(is.null(nrow)) return(matrix(nvec,ncol=ncol));
  if(is.null(ncol)) return(matrix(nvec,nrow=nrow));
}

### Turn 4-D iteration array to 2-D iteration matrix
flatten <- function(A4) {
  dim(A4)<-rep(sqrt(length(A4)),2)
  return(A4)
}

## Convert iteration matrix to iteration array
unfold <- function(A2,dim) {
  dim(A2) <- dim; return(A2);
}

### Test using random state matrix n2 and iteration array A4
mx=20; mq=10;
n2=matrix(rpois(mx*mq,lambda=3),mx,mq);
A4=array(rpois(mx*mq*mx*mq,lambda=3),dim=c(mx,mq,mx,mq));

## Do an iteration of one time-step the hard way
nnew.1 = matrix(0,mx,mq);
for(k in 1:mx){
  for(l in 1:mq){
    nnew.1[k,l]=sum(A4[k,l,,]*n2)
  }
}

## Do it the easy way
A2 <- flatten(A4);
nnew.2 <- matrix(A2%*%vec(n2),nrow=mx);
range(nnew.1-nnew.2); # should be (0,0)

```

Box B1: R code demonstrating conversion from state matrix to state vector, and from iteration array to iteration matrix.

Online Appendix C: The Zuidema et al. (2009) matrix model for *Cedrela odorata*

The model was based on growth trajectories (estimated from tree rings) of 60 trees, of size at least 60cm dbh, felled in a closed-canopy, nonflooded tropical moist forest in Pando Department, Bolivia. A matrix model was constructed using 10 year age-classes (0-10, 11-20 etc.) and 10cm dbh size categories, up to age 160 and size 120cm. All of the age-specific transition matrices are given in Appendix B of Zuidema et al. (2009). Zuidema et al. (2009) report that the canopy height of the forest was 25 - 35m, and that *Cedrela odorata* trees typically reach the canopy at size 30-40cm dbh.

We reduced the original model to 8 size (cm dbh) classes (0-9.9, 10-19.9, \dots , 60-69.9, 70 and above). This was possible because trees in size classes 6 and above all have the same survival and fecundity, and there is no shrinkage out of this size range. Our reduced model therefore makes exactly the same population projections as the original model, except that we lose information about the size distribution of larger trees (> 70 cm dbh). We did this reduction because the original model had some quirks in the larger, older age \times size categories, which are probably the result of the limited sample size. For example, in the published matrix P9 for age-class 9, individuals in size-class 11 have 100% mortality, contradicting the model's general assumption that mortality is size-independent above 50cm dbh. This feature would occur in a matrix model parameterized in the usual manner if at least one tree in the data set set transitions into an age \times size category, but none transitions out of it due to natural death or (depending on study methods) the study coming to an end. Quirks like this only arises in categories that few individuals enter, so they are probably unimportant for population projection, but they can have noticeable effects on life cycle analysis because transitions into black-hole categories are more likely in the conditional kernel for the Unlucky.

The fraction of variation in lifetime reproductive output due to variation in number of breeding events can be computed as follows. From the state at death distribution, 14% of individuals

reach size class 6 (> 50 cm), at which point they begin to reproduce, producing 2.03 recruits per 10 year time step of the model and having constant survival probability of 82% per time step. The number of breeding events (i.e., model time steps during which the tree breeds) is equal to the remaining lifespan (including the year of entering size class 6) which we denote here by N . Individuals who reach size class 6 remain in class 6 or larger until death with constant survivorship of 0.82 per model time step, so N has a geometric distribution with parameter $p = 0.18$. Reproductive output including variability due to randomness in the number of recruits per breeding event is the sum of N Poisson random variables with mean 2.03 (assuming a Poisson distribution of recruit numbers, resulting from many seeds with a low chance of recruitment). Removing the variability in recruits per breeding event, reproductive output is $2.03N$.

We use the following general results. First, if X_1, X_2, \dots are independent random variables and N is a non-negative random variable independent of the X 's, then the mean of $Z = X_1 + X_2 + \dots + X_N$ is $E[N]E[X]$, and the variance is

$$E[N] \text{Var}(X) + E[X]^2 \text{Var}(N).$$

Second, if a is a 0-1 random variable with probability f of getting 1, and Z is a random variable independent of a , then

$$\text{Var}(aZ) = f \text{Var}(Z) + f(1 - f)E(Z)^2.$$

We can apply these to N being geometric with parameter p , X =recruits per breeding event (either Poisson with mean 2.03, or constant with value 2.03), and $f = 0.14$ the probability that a new recruit survives to reach size class 6. Computing the standard deviations with and without variation in recruits per breeding event, we find that variation in number of breeding events accounts for $\approx 97.5\%$ of the total variation in reproductive output (the calculations are in the on-line SI script file `BreedingVarCompare.R`). Moreover, variation in the number of breeding events accounts for $\approx 95\%$ of the variation in reproductive output among the Lucky.

Online Appendix D: Additional Figures

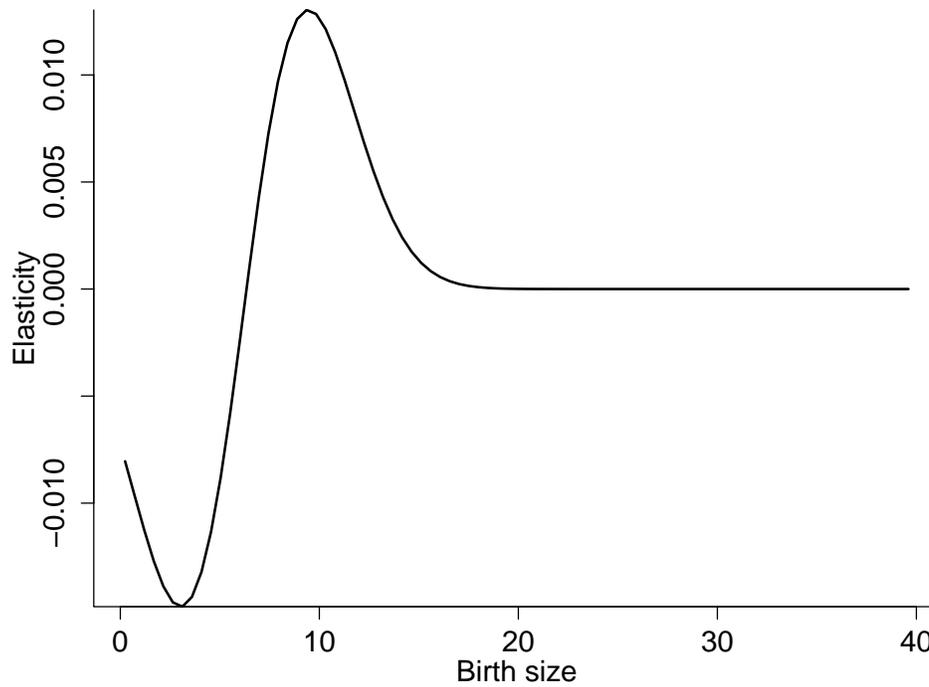


Figure D1: Elasticity of the probability of *Artemisia ordosica* reaching 40 cm height to perturbations in the probability of being born at a given size. If the birth size distribution is $c(z)$, then the elasticity $\frac{c(z_0)}{q_L} \frac{\partial q_L}{c(z_0)}$ equals $-a_L N c + (a_L N)(z_0)$. As expected, an individual is less likely to become Lucky if it is born at below-average size and more likely to become Lucky if it is born at above-average size, but the birth size elasticities are much smaller than those for expected growth and survival. Source files: `ArtemisiaBirthSizeElasticity.R` and the scripts that it sources.

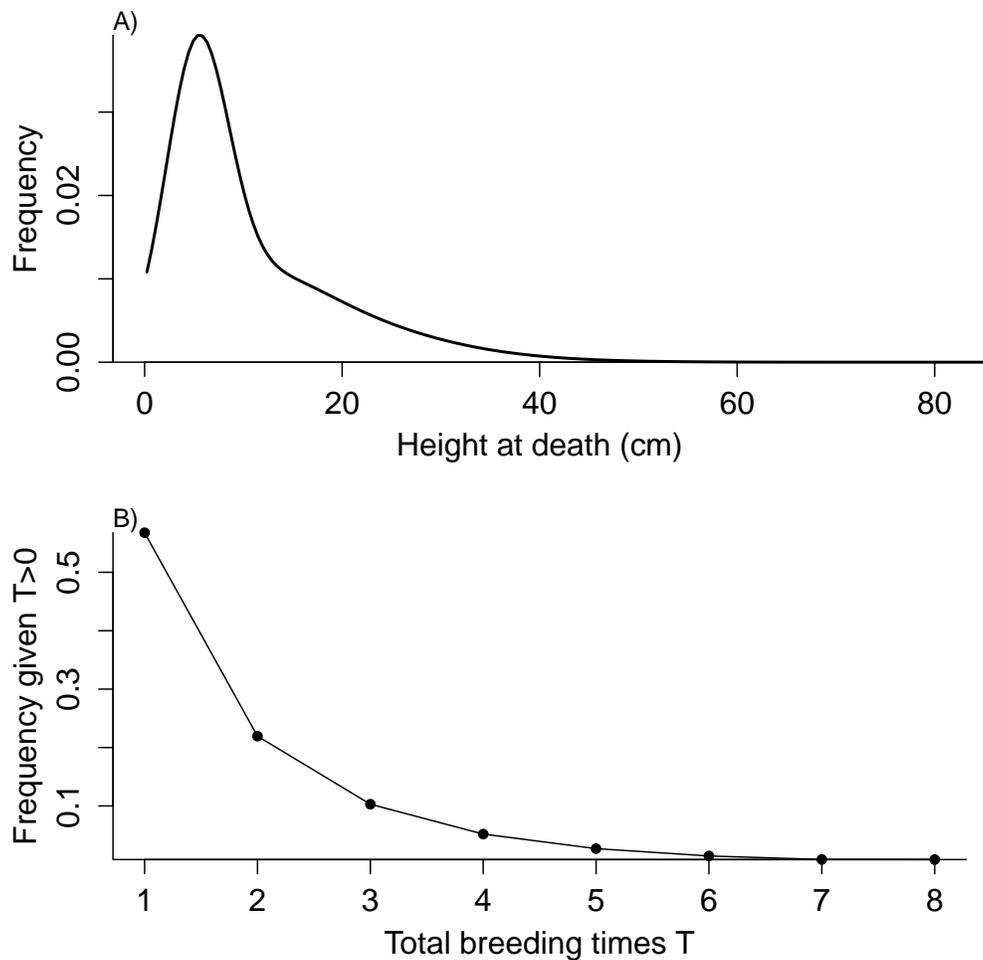


Figure D2: Marginal distributions of A) size and B) total number of seedlings T at death, in the *Artemisia* IPM with individuals cross-classified by size and T . The marginal distributions were calculated from the bivariate death state distribution Ω for a cohort of individuals born with $T = 0$. The plotted distribution of T (panel B) is the conditional distribution given $T > 0$, because over 95% of individuals die without ever producing a seedling. Source files: `ArtemisiaCompareLucky-xT.R` and scripts that it sources; `MegamatrixFunctions.R`

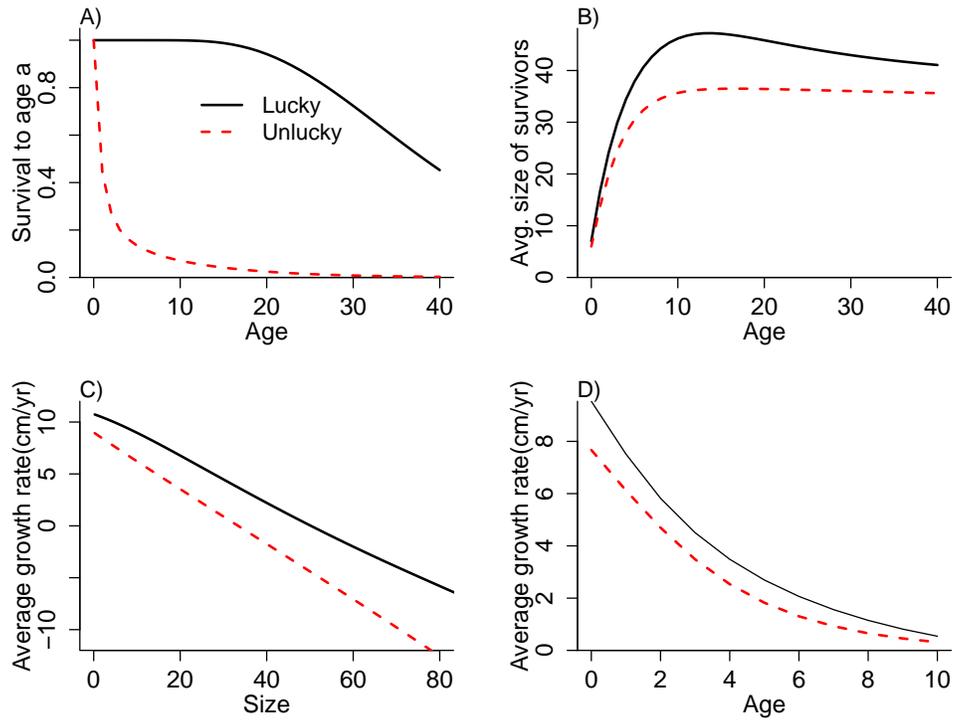


Figure D3: A) Survival to age a , $l(a)$, for Lucky (solid black curve) and Unlucky shrubs (dashed red curve) in the Li et al. (2011) IPM for *Artemisia* described in the text, with Lucky defined by producing 5 or more seedlings before death. B) Average size as a function of age. C) Average growth rate as a function of size, for individuals that have not yet crossed the threshold (40cm) for being Lucky. D) Average growth rate as a function of age. Source file: *ArtemisiaCompareLucky-xT.R* and scripts that it sources.

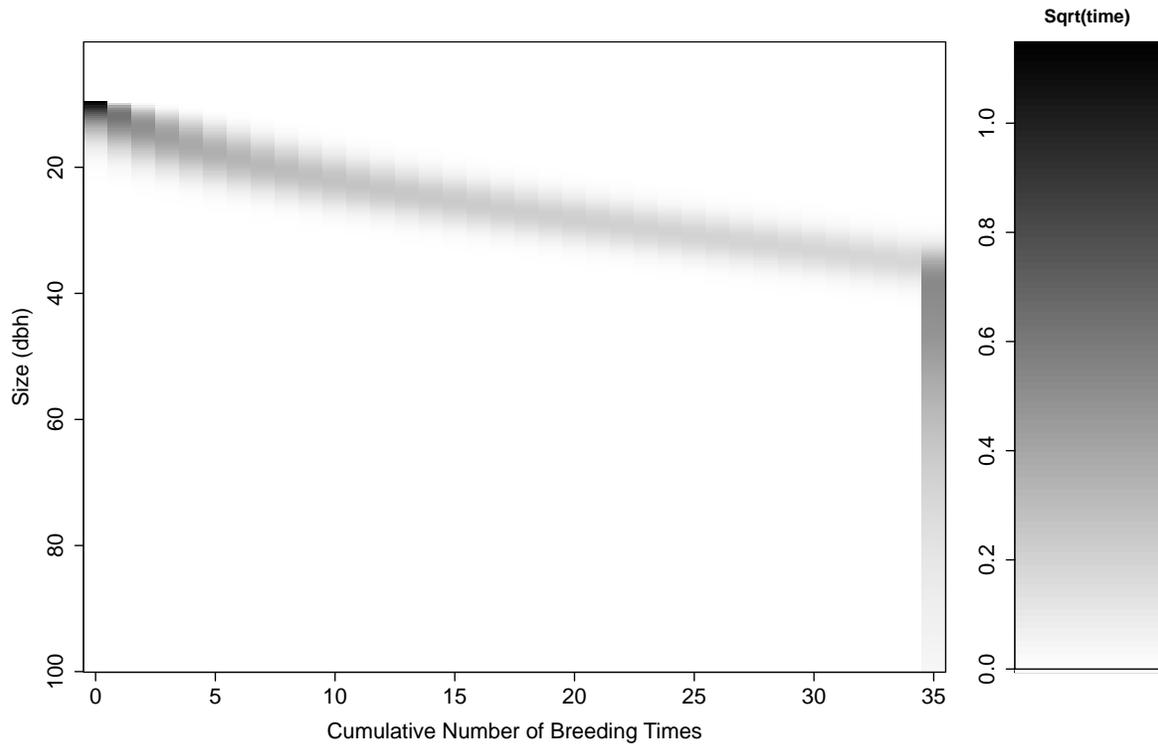


Figure D4: Image plot of the fundamental operator N in the *Dacrydium* IPM with cross-classification by size x and breeding times T . $N(z, z_0)dz$ is the expected total time in $(z, z + dz)$ from now until death, conditional on initial state in z_0 . Here we show N as a function of $z = (x, T)$ for initial state z_0 having $T = 0$ and size just below the minimum size for flowering (10cm dbh in this model). Note the plotting on square-root scale so that lower values of N are emphasized. Source files: ZuidemaCompareLucky-xT.R and scripts that it sources; MegamatrixFunctions.R

References

- Bonnet, T., and E. Postma. 2016. Successful by chance? The power of mixed models and neutral simulations for the detection of individual fixed heterogeneity in fitness components. *American Naturalist* 187:60–74.
- Brienen, R. J. W., P. A. Zuidema, and M. Martínez-Ramos. 2010. Attaining the canopy in dry and moist tropical forests: strong divergences in tree growth trajectories reflect variation in growing conditions. *Oecologia* 163:485–496.
- Bruno, J. F., S. P. Ellner, I. Vu, K. Kim, and C. D. Harvell. 2011. Impacts of aspergillosis on sea fan coral demography: modeling a moving target. *Ecological Monographs* 81:123–139.
- Caswell, H. 2001. *Matrix Population Models*, 2nd edition. Sinauer Associates, Sunderland MA.
- . 2009. Stage, age and individual stochasticity in demography. *Oikos* 118:1763–1782.
- . 2011a. Beyond R_0 : Demographic models for variability of lifetime reproductive output. *PLoS ONE* 6:e20809.
- . 2011b. Perturbation analysis of continuous-time absorbing Markov chains. *Numerical Linear Algebra With Applications* 18:901–917.
- . 2012. Matrix models and sensitivity analysis of populations classified by age and stage: a vec-permutation matrix approach. *Theoretical Ecology* 5:403–417.
- . 2013. Sensitivity analysis of discrete markov chains via matrix calculus. *Linear Algebra and its Applications* 438:1725–1745.
- Cochran, M. E., and S. Ellner. 1992. Simple methods for calculating age-specific life history parameters from stage-structured models. *Ecological Monographs* 62:345–364.
- Coulson, T., D. MacNulty, D. Stahler, B. vonHoldt, R. Wayne, and D. Smith. 2011. Modeling

- effects of environmental change on wolf population dynamics, trait evolution, and life history. *Science* 334:1275–1278.
- Dawson, A. E. 2013. *Models for Forest Growth and Mortality: Linking Demography to Competition and Climate*. Ph.D. thesis. University of Alberta, Department of Mathematical and Statistical Sciences. [Http://hdl.handle.net/10402/era.31581](http://hdl.handle.net/10402/era.31581).
- Ellner, S. P., D. Z. Childs, and M. Rees. 2015. *Data-driven Modeling of Structured Populations: A Practical Guide to the Integral Projection Model*. Springer, New York.
- Ellner, S. P., and M. Rees. 2006. Integral projection models for species with complex demography. *American Naturalist* 167:410–428.
- Iosifescu, M. 1980. *Finite Markov Processes and their Applications*. Wiley Series in Probability and Mathematical Statistics. John Wiley and Sons.
- Jansen, M., P. A. Zuidema, N. P. R. Anten, and M. Martínez-Ramos. 2012. Strong persistent growth differences govern individual performance and population dynamics in a tropical forest understorey palm. *Journal of Ecology* 100:1224–1232.
- Kemeny, J. G., and J. L. Snell. 1960. *Finite Markov Chains*. D. Van Nostrand Company, New York.
- Li, S.-L., F.-H. Yu, M. J. A. Werger, M. Dong, and P. A. Zuidema. 2011. Habitat-specific demography across dune fixation stages in a semi-arid sandland: understanding the expansion, stabilization and decline of a dominant shrub. *Journal of Ecology* 99:610–620.
- Metcalf, C. J. E., C. C. Horvitz, S. Tuljapurkar, and D. A. Clark. 2009. A time to grow and a time to die: a new way to analyze the dynamics of size, light, age, and death of tropical trees. *Ecology* 90:2766–2778.
- Meyn, S. P., and R. L. Tweedie. 1993. *Markov Chains and Stochastic Stability*. Springer-Verlag, New York.

- Moran, E. V., and J. S. Clark. 2012. Causes and consequences of unequal seedling production in forest trees: a case study in red oaks. *Ecology* 93:1082–1094.
- Orzack, S. H., U. K. Steiner, S. Tuljapurkar, and P. Thompson. 2011. Static and dynamic expression of life history traits in the northern fulmar *Fulmarus glacialis*. *Oikos* 120:369–380.
- Plard, F., C. Bonenfant, D. Delorme, and J.-M. Gaillard. 2012. Modeling reproductive trajectories of roe deer females: fixed or dynamic heterogeneity? *Theoretical Population Biology* 82:317–329.
- Plard, F., J.-M. Gaillard, T. Coulson, D. Delorme, C. Warnant, G. Michallet, S. Tuljapurkar, S. Krishnakuma, and C. Bonenfant. 2015. Quantifying the influence of measured and unmeasured individual differences on demography. *Journal of Animal Ecology* 84:1434–1445.
- Rees, M., D. Z. Childs, C. J. E. Metcalf, K. E. Rose, A. W. Sheppard, and P. J. Grubb. 2006. Seed dormancy and delayed flowering in monocarpic plants: selective interactions in a stochastic environments. *American Naturalist* 168:E53–E71.
- Rees, M., and S. P. Ellner. 2015. Evolving integral projection models: Evolutionary demography meets eco-evolutionary dynamics. *Methods in Ecology and Evolution*, *in press*. .
- Rees, M., A. Sheppard, D. Briese, and M. Mangel. 1999. Evolution of size-dependent flowering in *Onopordum illyricum*: A quantitative assessment of the role of stochastic selection pressures. *American Naturalist* 154:628–651.
- Rose, K., M. Rees, and P. Grubb. 2002. Evolution in the real world: Stochastic variation and the determinants of fitness in *Carlina vulgaris*. *Evolution* 56:1416–1430.
- Steiner, U. K., and S. Tuljapurkar. 2012. Neutral theory for life histories and individual variability in fitness components. *Proceedings of the National Academy of Sciences of the United States of America* 109:4684–4689.

- Steiner, U. K., S. Tuljapurkar, T. Coulson, and C. Horvitz. 2012. Trading stages: Life expectancies in structured populations. *Experimental Gerontology* 47:773–781.
- Steiner, U. K., S. Tuljapurkar, and S. H. Orzack. 2010. Dynamic heterogeneity and life history variability in the kittiwake. *Journal of Animal Ecology* 79:436–444.
- Team, R. C. 2015. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN3-900051-07-0.
- Tuljapurkar, S., and C. Horvitz. 2006. From stage to age in variable environments: Life expectancy and survivorship. *Ecology* 87:1497–1509.
- Tuljapurkar, S., U. K. Steiner, and S. H. Orzack. 2009. Dynamic heterogeneity in life histories. *Ecology Letters* 12:93–106.
- van Groenendael, J., H. de Kroon, S. Kalisz, and S. Tuljapurkar. 1994. Evaluating Life History Pathways in Population Projection Matrices. *Ecology* 75:2410–2415.
- Vindenes, Y., and Ø. Langangen. 2015. Individual heterogeneity in life histories and eco-evolutionary dynamics. *Ecology Letters* 18:417–432.
- Zuidema, P. A., R. J. W. Brienen, H. J. D. During, and B. Guneralp. 2009. Do persistently fast-growing juveniles contribute disproportionately to population growth? A new analysis tool for matrix models and its application to rainforest trees. *the American Naturalist* 174:709–719.
- Zuidema, P. A., E. Jongejans, P. D. Chien, H. J. During, and F. Schieving. 2010. Integral Projection Model for trees: a new parameterization method and a validation of model output. *Journal of Ecology* 98:345–355.

Figure legends

Figure 1: The expanded state space for a purely size-structured model with shrinkage below the critical size z_c . When an individual first grows from size $z < z_c$ to size $z' > z_c$, it moves from S to B . When an individual in B shrinks from $z > z_c$ to $z < z_c$ it stays in B . Individuals who die from S are Unlucky (absorbing state 1); individuals who die from B are Lucky (absorbing state 2).

Figure 2: A) Annual probabilities of survival and “breeding” (flowering and producing seeds) as a function of size in the Zuidema et al. (2010) IPM for *Dacrydium elatum*, for saplings and trees ($> 1\text{cm dbh}$). B) Mean (solid) and standard deviation (dashed) of the number of breeding events from the present until death, as a function of current size. C) Stable size distribution (on log-10 scale). D) Distribution of parent size for seeds produced in a given year, for a population at stable size distribution. The plotted curve is the product of the stable size distribution and size-specific mean fecundity, normalized to a probability density function. Source file: ZuidemaBreedingFreq.R and scripts that it sources.

Figure 3: A) Annual survival probability and mean fecundity (number of seedlings) as a function of size in the IPM for *Artemisia ordosica* by Li et al. (2011), 2007-2008 census period on fixed dunes with microbiotic crust. B) Mean (solid) and standard deviation (dashed) of total offspring (seedlings), from the present until death, as a function of current size. The formulas for these are derived in Ellner et al. (2015, Chapter 3)). C) Stable size distribution. D) Distribution of parent size for seedlings produced in a given year, for a population at stable size distribution. The plotted curve is the product of the stable size distribution and size-specific mean fecundity, normalized to a probability density function. Source file: `ArtemisiaBreedingFreq.R` and scripts that it sources.

Figure 4: A) Survival to age a , $l(a)$, for Lucky (solid black curve) and Unlucky trees (dashed red curve) in the Zuidema et al. (2010) IPM for *Dacrydium*. B) Average size as a function of age. C) Average growth rate as a function of size. The vertical dashed line marks the end of the seedling stages. D) Average growth rate as a function of age. For these calculations, seedlings (0-1 cm dbh, classified into four height classes) are treated as having sizes 0.2, 0.4, 0.6, 0.8 cm dbh. Source file: `ZuidemaLucky20-compareGrowth.R` and scripts that it sources.

Figure 5: A) Survival to age a , $l(a)$, for Lucky (solid black curve) and Unlucky shrubs (dashed red curve) in the Li et al. (2011) IPM for *Artemisia* described in the text. B) Average size as a function of age. C) Average growth rate as a function of size, for individuals that have not yet crossed the threshold (40cm) for being Lucky. D) Average growth rate as a function of age. Source file: `ArtemisiaLucky40-compareGrowth.R` and scripts that it sources.

Figure 6: Comparison based on conditional kernels for the Zuidema et al. (2009) matrix model for *Cedrela odorata* with size-by-age classification and Lucky defined as reaching size-class 6. A) Frequency distribution of size-class at death. Note that size-class 8 is a “big trees” category (70cm dbh or larger). B) Survival to age-class a , $l(a)$, for Lucky (solid black curve) and Unlucky trees (dashed red curve). C) Average size-class as a function of age-class for Lucky and Unlucky. D) Average growth rate as a function of age-class for Lucky and Unlucky trees. The plotted growth rates start to level off at age-class 9 because many trees have reached the maximum possible size-class (class 8 for Lucky, class 5 for Unlucky). Source files: `CedrelaMatrices.R` and `CedrelaLucky.R`

Figure 7: *Dacrydium elatum*: elasticity of the probability of a tree reaching 20 cm dbh to A) Perturbations in expected growth, the standard deviation of growth, and survival at a given size, B) The same perturbations at a given age. Based on the Zuidema et al. (2010) IPM for *Dacrydium*. Note that the IPM only covers the sapling stages — seedling dynamics are governed by a Leslie matrix — so individuals must reach the sapling stages before the perturbations can have an effect. Source file: `DacrydiumElasticities.R` and the scripts that it sources.

Figure 8: *Artemisia ordosica*: elasticity of the probability of reaching 40 cm height to A) Perturbations in expected growth, the standard deviation of growth, and survival at a given size, B) The same perturbations at a given age, in the Li et al. (2011) IPM for *Artemisia*. Source file: `ArtemisiaElasticities.R` and the scripts that it sources.

Figure 9: *Cedrela odorata*: elasticity of the probability of reaching size-class 6 to A) survival in different age-classes, and B) survival in different size-classes, for the Zuidema et al. (2009) matrix model for *Cedrela*. Source files: `CedrelaMatrices.R` and `CedrelaPerturbS.R`

Figure 10: Marginal distributions of A) size and B) total number of flowering times T at death, in the *Dacrydium* IPM with individuals cross-classified by size and T . The marginal distributions were calculated from the bivariate death state distribution Ω for a cohort of individuals born into the smallest seedling class with $T = 0$. Note that the size distribution is plotted on a log-transformed scale so that the second mode at ≈ 50 cm dbh can be seen. The plotted distribution of T (panel B) is the conditional distribution given $T > 0$, because over 99% of individuals die before reaching the minimum size for flowering (10cm dbh); the large value at $T = 35$ is because the model was implemented with $T = 35$ as a class representing all $T \geq 35$. Source files: `ZuidemaCompareLucky-xT.R` and scripts that it sources; `MegamatrixFunctions.R`

Figure 11: Contour plot of the probability of becoming Lucky ($T \geq 20$ at death) as a function of current size and T , on the set of small states S . Source files: `ZuidemaCompareLucky-xT.R` and scripts that it sources; `MegamatrixFunctions.R`

Figure 12: A) Survival to age a , $l(a)$, for Lucky (solid black curve) and Unlucky trees (dashed red curve) in the Zuidema et al. (2010) IPM for *Dacrydium* with size by T cross-classification, and Lucky defined as reaching $T = 20$. B) Average size as a function of age. C) Average growth rate as a function of size. The vertical dashed line marks the end of the seedling stages. D) Average growth rate as a function of age. For these calculations, seedlings (0-1 cm dbh, classified into four height classes) are treated as having sizes 0.2, 0.4, 0.6, 0.8 cm dbh. Source files: `ZuidemaCompareLucky-xT.R` and scripts that it sources.

Online figure legends

Figure D1: Elasticity of the probability of *Artemisia ordosica* reaching 40 cm height to perturbations in the probability of being born at a given size. If the birth size distribution is $c(z)$, then the elasticity $\frac{c(z_0)}{q_{\mathbb{L}}} \frac{\partial q_{\mathbb{L}}}{\partial c(z_0)}$ equals $-a_{\mathbb{L}} N c + (a_{\mathbb{L}} N)(z_0)$. As expected, an individual is less likely to become Lucky if it is born at below-average size and more likely to become Lucky if it is born at above-average size, but the birth size elasticities are much smaller than those for expected growth and survival. Source files: `ArtemisiaBirthSizeElasticity.R` and the scripts that it sources.

Figure D2: Marginal distributions of A) size and B) total number of seedlings T at death, in the *Artemisia* IPM with individuals cross-classified by size and T . The marginal distributions were calculated from the bivariate death state distribution Ω for a cohort of individuals born with $T = 0$. The plotted distribution of T (panel B) is the conditional distribution given $T > 0$, because over 95% of individuals die without ever producing a seedling. Source files: `ArtemisiaCompareLucky-xT.R` and scripts that it sources; `MegamatrixFunctions.R`

Figure D3: A) Survival to age a , $l(a)$, for Lucky (solid black curve) and Unlucky shrubs (dashed red curve) in the Li et al. (2011) IPM for *Artemisia* described in the text, with Lucky defined by producing 5 or more seedlings before death. B) Average size as a function of age. C) Average growth rate as a function of size, for individuals that have not yet crossed the threshold (40cm) for being Lucky. D) Average growth rate as a function of age. Source file: `ArtemisiaCompareLucky-xT.R` and scripts that it sources.

Figure D4: Image plot of the fundamental operator N in the *Dacrydium* IPM with cross-classification by size x and breeding times T . $N(z, z_0)dz$ is the expected total time in $(z, z + dz)$ from now until death, conditional on initial state in z_0 . Here we show N as a function of $z = (x, T)$ for initial state z_0 having $T = 0$ and size just below the minimum size for flowering (10cm dbh in this model). Note the plotting on square-root scale so that lower values of N are emphasized. Source files: `ZuidemaCompareLucky-xT.R` and scripts that it sources; `MegamatrixFunctions.R`