Demography (2013) 50:1615–1640
DOI 10.1007/s13524-013-0223-3

Perturbation Analysis of Indices of Lifespan Variability

Alyson A. van Raalte · Hal Caswell

Published online: 17 September 2013

Abstract A number of indices exist to calculate lifespan variation, each with different underlying properties. Here, we present new formulae for the response of seven of these indices to changes in the underlying mortality schedule (life disparity, Gini coefficient, standard deviation, variance, Theil's index, mean logarithmic deviation, and interquartile range). We derive each of these indices from an absorbing Markov chain formulation of the life table, and use matrix calculus to obtain the sensitivity and the elasticity (i.e., the proportional sensitivity) to changes in age-specific mortality. Using empirical French and Russian male data, we compare the underlying sensitivities to mortality change under different mortality regimes to determine the conditions under which the indices might differ in their conclusions about the magnitude of lifespan variation. Finally, we demonstrate how the sensitivities can be used to decompose temporal changes in the indices into contributions of age-specific mortality changes. The result is an easily computable method for calculating the properties of this important class of longevity indices.

Keywords Inequality · Lifespan variation · Mortality compression · Age patterns of mortality · Decomposition

Electronic supplementary material The online version of this article (doi:10.1007/s13524-013-0223-3) contains supplementary material, which is available to authorized users.

A. A. van Raalte (✉) · H. Caswell
Max Planck Institute for Demographic Research, Konrad-Zuse Str. 1, 18057 Rostock, Germany
e-mail: vanRaalte@demogr.mpg.de

H. Caswell
Biology Department MS-34, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA

Introduction

In this article, we present a detailed comparative study of the sensitivity and elasticity of indices of variation in longevity. The longevity experience of a cohort has long been summarized by life expectancy, but attention has recently expanded to focus on variation in longevity as a natural complement to describing the average length of life. Indices of variability have been compared across populations to measure the rectangularity of the survival curve or degree of mortality compression for both human and and nonhuman populations (see, e.g., Eakin and Witten 1995; Edwards and Tuljapurkar 2005; Smits and Monden 2009; van Raalte et al. 2011; Vaupel et al. 2011). They have also been employed above the modal age at death to examine whether old-age mortality is being compressed, or whether these deaths are shifting to higher ages (Brown et al. 2012; Cheung et al. 2005; Cheung and Robine 2007; Kannisto 2000, 2001; Ouellette and Bourbeau 2011; Thatcher et al. 2010). Several authors have compared the various indices of lifespan variation (Anand et al. 2001; Cheung et al. 2005; Kannisto 2000; Shkolnikov et al. 2003; Vaupel et al. 2011; Wilmoth and Horiuchi 1999), finding such high correlations among indices as to make them apparently interchangeable.

Our concern here is with a neglected aspect of these indices: their response to perturbations of the underlying mortality schedule. In general, perturbation analysis calculates the change in some quantity x attributable to changes in one or more parameters θ from which that quantity is calculated. This can be expressed as a derivative, the sensitivity $dx/d\theta$, which gives the response of x to an additive perturbation in θ, or as the proportional derivative, or elasticity $\frac{x}{\theta}\frac{dx}{d\theta}$, which gives the proportional change in x resulting from a proportional change in θ.

Sensitivity analysis has a long history in formal demography. Examples include the sensitivity of the intrinsic growth rate r or λ to changes in mortality and fertility, first derived by Hamilton (1966), then by Demetrius (1969) and Keyfitz (1971), and then in general form by Caswell (1978). The sensitivity of life expectancy to changes in mortality was derived by Keyfitz (1971), investigated in more detail by Pollard (1982, 1988), and more recently studied by Caswell (2006) and Wrycza and Baudisch (2012). Keyfitz introduced an elasticity of life expectancy in terms of life table entropy (see Keyfitz and Caswell 2005: chap. 4), which was further developed by Goldman and Lord (1986), Vaupel (1986), and Vaupel and Canudas Romo (2003). The sensitivity of the stable population structure to changes in mortality and fertility was investigated numerically by Coale (1972) and analytically by Caswell (2001, 2008). The sensitivity of short-term projections was analyzed by Caswell (2007). All of these sensitivity analyses quantify the effect of parameter changes, permit rigorous comparison of different perturbations, and provide the tools for decomposing changes in an index into components arising from changes in each parameter (Caswell 1989, 2001). Until its perturbation analysis is available, our understanding of any index is incomplete. Indices of variation in longevity are no exception, but no general perturbation theory is available.

The patterns of sensitivity and elasticity of each index will reveal whether the correlations among the values of the various indices extend to their response to changes in mortality. This includes analytically determining the age ranges over which mortality reductions lead to increases or decreases in variation. These age ranges are separated by a "threshold age," at which the derivative is zero. Perturbation results can also be used to target age-specific mortality reductions to ages that best reduce lifespan variation. Moreover, because the sensitivity results are age-specific, they can be related to normative concepts of inequality or social preference for actions designed to reduce mortality at particular ages (Anand et al. 2001; Asada 2007; Gakidou et al. 2000). This has rarely been done; a notable exception to this is the WHO attempt to quantify inequality among individuals as part of World Health Report 2000, using an index similar to a Gini coefficient, modified by expert opinion (Gakidou et al. 2000; Gakidou and King 2002; WHO 2000).

In this article, we provide the first readily computable formulae for the sensitivities and elasticities of seven of the most commonly used indices to changes in age-specific mortality and, by extension, to any parameter(s) affecting age-specific mortality. Our approach is to reformulate the problem of lifespan variability in terms of an absorbing Markov chain, and to use methods from matrix calculus (Caswell 2008, 2009, 2010). We will show how to use the results to decompose differences among populations or over time into contributions from changes in age-specific mortality, using life table response experiment (LTRE) methods (Caswell 2001). We apply these methods to data from France and Russia, and illustrate instances in which these different sensitivities cause indices to predict changes in lifespan variation that differ in magnitude or even direction.

Indices of Lifespan Variation

We compare seven indices of variability: life disparity, Gini coefficient, Theil's index, mean logarithmic deviation, standard deviation, variance, and interquartile range.

> Life disparity ($e^{\dagger}$) is a life table–based index defined as the average remaining life expectancy at death, or alternatively the average years of life lost in a population attributable to death. The elasticity of life expectancy with respect to mortality change, also known as Keyfitz' H (Keyfitz 1977), is $e^{\dagger}$ divided by the life expectancy at birth (Goldman and Lord 1986; Vaupel 1986; Vaupel and Canudas Romo 2003).
>
> The Gini coefficient (G) is often used in economic inequality research. It ranges from 0 to 1, with higher numbers signaling greater inequality. It is the mean of the absolute value of the interindividual differences in age at death, divided by the life expectancy (Shkolnikov et al. 2003).
>
> Both Theil's index (T) and the mean logarithmic deviation (MLD) are based on the entropy of the distribution of age at death, developed from information theory by Henry Theil in the 1960s (Theil 1967). The entropy of a distribution measures the amount of information needed to specify the result of sampling;

if everyone died at the same age, no information would be needed and the entropy-based measures would be zero.

Standard deviation (S), variance (V), and interquartile range (IQR) are standard statistical measures of variability applied to the distribution of age at death.

Different research objectives often call for the use of one index over another because of the underlying formal properties of each index. V, T, and MLD are all additively decomposable into between- and within-group variation (Shorrocks 1980; van Raalte et al. 2012). This decomposition can be used to study the contribution of between-group differences to the total level of lifespan variation. G can also be decomposed in this way, but it contains an overlap term (Lambert and Aronson 1993). The MLD index can also be additively decomposed over time to account for compositional change to the between- and within-group variation components (Mookherjee and Shorrocks 1982). Between-population differences in V can further be decomposed by cause into spread, allocation, and timing effects (Nau and Firebaugh 2012). $e^{\dagger}$ has interesting connections to the perturbation theory of life expectancy: the product of $e^{\dagger}$ and the average rate of progress in reducing age-specific death rates is equal to the rate of change in life expectancy (Vaupel and Canudas Romo 2003).

Indices also differ in whether they measure absolute inequality (the level of variation would be unaffected by additive gains to everyone's lifespan) or relative inequality (the level of variation would be unaffected by proportional gains to everyone's lifespan). Additive indices are more easily interpretable because they are normally expressed in years.

The sensitivity of indices to changes in mortality at different ages is an important and poorly understood property of the indices. In some circumstances, society might consider variability in ages at death caused by high levels of premature mortality to be more detrimental than variability caused by differences in old-age mortality. In such a case, use of an index with a high sensitivity to early death would emphasize the response to the mortality viewed as most important. As Paul Allison (1978:865) noted, "The choice of an inequality measure is properly regarded as a choice among alternative definitions of inequality rather than a choice among alternative ways of measuring a single theoretical construct."

The seven indices we examine here are highly correlated across countries and times. Some of these correlations have been reported by Vaupel et al. (2011) and Wilmoth and Horiuchi (1999). We present the correlations among all seven indices, from birth and from age 10 (Table 1), calculated over all female and male life tables[1] currently in the Human Mortality Database (HMD 2012). Thus, we expect that all of them will pick up most of the general patterns in lifespan variation in inter-population comparisons. Our focus is on the details of the response of the indices to changes in mortality.

[1]We excluded populations that were double-counted; that is, we excluded civilian population in favor of total national population life tables and excluded nationally aggregated life tables in favor of including regional or ethnicity-based life tables.

Table 1 Pearson correlation coefficients between pairs of indices, calculated from birth (ages 0–110+) in the top panel and calculated conditional on survival to age 10 (ages 10–110+) in the bottom panel, for all female and male life tables in the Human Mortality Database (7,516 in total)

	$e^\dagger$	G	T	MLD	S	V	IQR
$e^\dagger$	1.000						
G	.978	1.000					
T	.947	.991	1.000				
MLD	.965	.991	.992	1.000			
S	.981	.933	.893	.930	1.000		
V	.987	.945	.911	.944	.996	1.000	
IQR	.967	.966	.948	.956	.920	.944	1.000
	$e^\dagger_{10}$	G_{10}	T_{10}	MLD_{10}	S_{10}	V_{10}	IQR_{10}
$e^\dagger_{10}$	1.000						
G_{10}	.986	1.000					
T_{10}	.978	.995	1.000				
MLD_{10}	.979	.990	.995	1.000			
S_{10}	.986	.962	.961	.973	1.000		
V_{10}	.984	.964	.971	.980	.998	1.000	
IQR_{10}	.981	.978	.977	.976	.958	.966	1.000

Markov Chain Formulations of Longevity

To develop perturbation results, we formulate the mortality schedule as a finite-state absorbing Markov chain (Caswell 2001, 2006, 2009, 2010; Feichtinger 1973). This formulation lets us express the various indices in matrix notation and then apply matrix calculus to obtain the sensitivity and elasticity of each index to changes in parameters (e.g., Caswell 2006, 2009, 2011). Because this study focuses on human demography, we focus on the age-classified model. Nevertheless, these results could be generalized to apply to stage-classified populations.

Notation

We use matrix notation in deriving the sensitivities. Matrices are denoted by uppercase bold-faced symbols (e.g., $\mathbf{X}$), and vectors are denoted by lowercase bold-faced symbols ($\mathbf{x}$); vectors are column vectors by default, and $\mathbf{x}^\top$ denotes the transpose of $\mathbf{x}$. The symbol diag($\mathbf{x}$) denotes the matrix with the vector $\mathbf{x}$ on the diagonal and zeros elsewhere. The vector $\mathbf{e}$ is a vector of ones, and the vector $\mathbf{e}_i$ is the ith unit vector—that is, the vector with a 1 in the ith location and zeros elsewhere. The Hadamard product (or element-by-element product) is denoted by $\circ$, and the Kronecker product is denoted by $\otimes$. The vec operator (e.g., vec$\mathbf{X}$) stacks the columns of a matrix into a column vector. See Caswell (2001) or Keyfitz and Caswell (2005) for an account of matrix models in demography.

We consider s age classes. Let $\mathbf{U}$ be a matrix ($s \times s$) with survival probabilities on the subdiagonal and zeros elsewhere; that is,

$$u_{i+1,i} = 1 - q_{i-1} \qquad i = 1, \ldots, s-1 \tag{1}$$

where q_i is the probability of death between ages i and $i+1$ from the life table. Note that matrix entries are indexed starting at 1, but age in the life table (conventionally) starts at 0. Thus, for example, the probability of surviving from the first to the second age state is the complement of the probability of death between age 0 and age 1, or $u_{2,1} = 1 - q_0$.

The matrix $\mathbf{U}$ describes transitions among the transient states in the Markov chain. Death is an absorbing state; we classify deaths by the age class at death with a diagonal matrix $\mathbf{M}$ ($s \times s$), where

$$m_{i,i} = 1 - q_{i-1} \qquad i = 1, \ldots, s \tag{2}$$

The transition matrix[2] for the Markov chain is

$$\mathbf{P} = \left(\begin{array}{c|c} \mathbf{U} & 0 \\ \hline \mathbf{M} & \mathbf{I} \end{array}\right), \tag{3}$$

where 0 is a matrix of zeros and $\mathbf{I}$ is an ($s \times s$) identity matrix.

In this Markov chain, absorption corresponds to death, and the time to absorption corresponds to longevity. The statistical properties of longevity can be directly calculated from $\mathbf{P}$. The mean time spent in age class i, conditional on starting in age class j is given by the (i, j) entry of the fundamental matrix

$$\mathbf{N} = (\mathbf{I} - \mathbf{U})^{-1}. \tag{4}$$

Because absorption corresponds to death, the time to absorption can be treated as a measure of longevity (Caswell 2001, 2006, 2009). The mean time to absorption is given by the column sums of $\mathbf{N}$. Let $\tilde{\boldsymbol{\eta}}$ denote the vector whose ith entry is the expected time to absorption for an individual in age class i; it is given by

$$\tilde{\boldsymbol{\eta}}^\top = \mathbf{e}^\top \mathbf{N}, \tag{5}$$

where $\mathbf{e}$ is a vector of ones. However, it can be shown that this exceeds by 0.5 years the life expectancy calculated by the usual life table formulations; accordingly, we use

$$\boldsymbol{\eta} = \tilde{\boldsymbol{\eta}} - 0.5\mathbf{e} \tag{6}$$

to represent life expectancy. The subtraction of the constant 0.5 does not affect the calculations of sensitivities.

The vector of variances in longevity satisfies

$$\mathbf{v}^\top = \mathbf{e}^\top \mathbf{N}(2\mathbf{N} - \mathbf{I}) - \boldsymbol{\eta}^\top \circ \boldsymbol{\eta}^\top, \tag{7}$$

where $\circ$ denotes the Hadamard (or element-by-element) product.

The complete distribution of age at death, conditional on starting in age class j, is given by column j of the matrix

$$\mathbf{B} = \mathbf{MN}. \tag{8}$$

[2]Note that $\mathbf{P}$ is column-stochastic and operates on column vectors, in line with the orientation of population projection matrices (e.g., Caswell 2001; Keyfitz and Caswell 2005).

The distribution of age at death for an individual in the first age class is given by the first column of **B**;

$$\mathbf{f} = \mathbf{B}\mathbf{e}_1. \tag{9}$$

The survivorship function ℓ, beginning at age 1 and with a radix $\ell(0) = 1$, is given by

$$\ell = \mathbf{e} - \mathbf{C}\mathbf{f}, \tag{10}$$

where

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & 0 \end{pmatrix} \tag{11}$$

takes cumulative sums of the vector **f**.

The vector **x** contains the average age at death in the age interval; for French males in 2005, it is {0.06, 1.5, 2.5, ..., 109.5, 111.32}.

In Table 2 we present the conventional life table notation alongside the less familiar matrix notation for each index. In conventional notation, ℓ_y is survivorship, d_y is the death density, and e_y is remaining life expectancy for the age interval y to $y+1$. We further denote a_y as the length of the age interval lived by those who died. An overbar, as in $\bar{e}_y$, is used when adjustments to the variable are necessary to account for the portion of the age interval lived by those who died; that is,

$$\bar{e}_y = e_y + a_y(e_{y+1} + e_y). \tag{12}$$

By this same logic, $\bar{x}_y$ is the average age at death over the interval. Generally it is the age halfway in between the two age intervals, but in the first year of life, $\bar{x}_0 = a_0$. The highest age interval is denoted by ω.

Finally in the *IQR* formula, $\hat{x}_1$ and $\hat{x}_3$ are, respectively, the interpolated first and third age quartiles, at which 25 % and 75 % of the total deaths have occurred.

Table 2 Formulae for calculating indices in conventional life table formulation (discrete, assuming ℓ_0 of 1) and their equivalent formulation in matrix notation

	Conventional Life Table Notation	Matrix Notation
$e^\dagger$	$\sum_{y=0}^{\omega} d_y \bar{e}_y$	$\mathbf{f}^\top \boldsymbol{\eta}$
G	$1 - \frac{1}{e_0}\sum_{y=0}^{\omega} \ell_{y+1}^2$	$1 - \frac{1}{\eta_1}\mathbf{e}^\top[(\mathbf{e}-\mathbf{C}\mathbf{f}) \circ (\mathbf{e}-\mathbf{C}\mathbf{f})]$
T	$\sum_{y=0}^{\omega} d_y \left(\frac{\bar{x}_y}{e_0} ln \frac{\bar{x}_y}{e_0}\right)$	$\mathbf{f}^\top \left[\left(\frac{\mathbf{x}}{\eta_1}\right) \circ \left(\log \frac{\mathbf{x}}{\eta_1}\right)\right]$
MLD	$\sum_{y=0}^{\omega} d_y \left(ln \frac{e_0}{\bar{x}_y}\right)$	$\mathbf{f}^\top \left[\log(\eta_1)\,\mathbf{e} - \log \mathbf{x}\right]$
V	$\sum_{y=0}^{\omega} d_y \left(\bar{x}_y - e_0\right)^2$	$\left[\mathbf{e}^\top \mathbf{N}(2\mathbf{N}-\mathbf{I}) - \boldsymbol{\eta}^\top \circ \boldsymbol{\eta}^\top\right]^\top$
S	$\sqrt{V}$	$\sqrt{V}$
IQR	$\hat{x}_3 - \hat{x}_1$	$\hat{x}_3 - \hat{x}_1$

Sensitivity and Elasticity Analysis

Expressing longevity in terms of an absorbing Markov chain and applying matrix calculus has greatly expanded the possibilities for perturbation analysis (Caswell 2001, 2008, 2009, 2010; Willekens 1977). To assess the absolute and proportional effects on the indices of changes in the underlying mortality rates, we needed the analytic expressions for the sensitivity and elasticity of the seven indices of lifespan variability with respect to mortality. The sensitivity of $e^{\dagger}$ was first derived by Zhang and Vaupel (2009) in an age-classified model, which was further developed by Wagner (2010). This was later generalized to an age- and stage-classified model by Caswell (2010), who also derived expressions for the sensitivity and elasticity of the variance and the standard deviation (Caswell 2009). The other expressions were newly derived for this article.

As described in detail in the Appendix, the sensitivity of a $n \times 1$ vector $\mathbf{y}$ to a $m \times 1$ vector of parameters θ is given by the $n \times m$ matrix

$$\frac{d\mathbf{y}}{d\theta^{\top}} = \left(\frac{dy_i}{d\theta_j}\right), \tag{13}$$

whose (i, j) entry is the derivative of y_i with respect to θ_j. The elasticity of $\mathbf{y}$ to θ is

$$\frac{\epsilon\mathbf{y}}{\epsilon\theta^{\top}} = \text{diag}(\mathbf{y})^{-1}\frac{d\mathbf{y}}{d\theta^{\top}}\text{diag}(\theta). \tag{14}$$

The sensitivities and elasticities of each index to age-specific mortality were derived using matrix calculus (Magnus and Neudecker 1988). These techniques have been given extensive treatment in recent publications by Caswell (2008, 2009, 2010), using most of the same notation that we have here. The derivation of the sensitivities of all the indices to mortality can be found in the Appendix. We performed all numerical calculations in MATLAB (version 7.3.0), and include both MATLAB and R code in Online Resources 1 and 2. The formulae resulting from the matrix calculus may appear complicated, but the complication arises from, and accounts for, the network of interactions among the variables, and they are easily computed.

We now turn to the demographic applications, especially in comparing the sensitivities of these indices, examining how they have changed over time as we have moved from high- to low-mortality regimes, and using the sensitivities as a decomposition method.

A Comparison of Sensitivities: France and Russia

We used French male data to broadly illustrate the underlying sensitivities and elasticities of each index. We calculated the indices under four very different mortality regimes: high mortality (1888), medium mortality (1948), low mortality (2005), and war/epidemic year (1918). The last regime is interesting: the distribution of age at death has a second mode around young adulthood and a long right tail instead of a long left tail. To help visualize these differences, we plot all four distributions in Fig. 1.

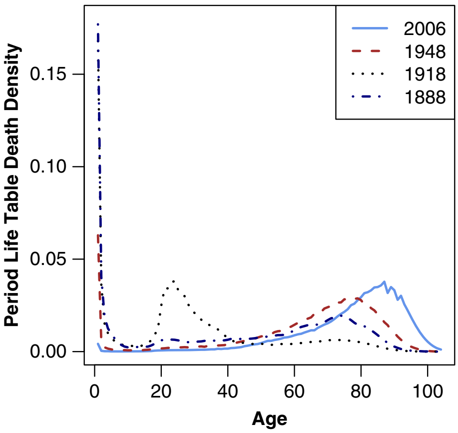

Fig. 1 The French male death densities for which the sensitivities and elasticities of the indices are compared (color figure online)

All indices are highly sensitive to changes in infant mortality. For this reason we calculated the sensitivities and elasticities both at birth (Figs. 2 and 3) and at age 10 (see the Appendix). Given the different units for each index, the elasticities are perhaps intuitively easier to interpret. The y-axis measures the percentage change in the index from a 1 % change in mortality at each age on the x-axis.

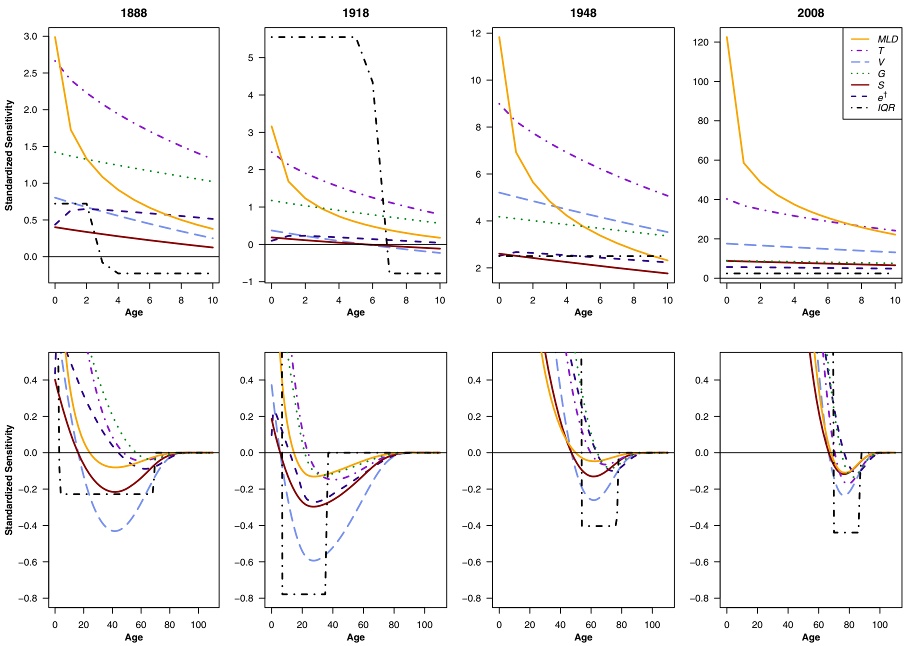

Fig. 2 The sensitivity of each index with respect to mortality change at different ages. The sensitivities were standardized to the value of each index (i.e., $y_0 \frac{dy}{d\theta}$) to make them comparable. Note the difference in scale between the top and bottom panels, plotted separately to more clearly delineate behavior of the indices at early and later ages. French males, period life table data from the Human Mortality Database (color figure online)

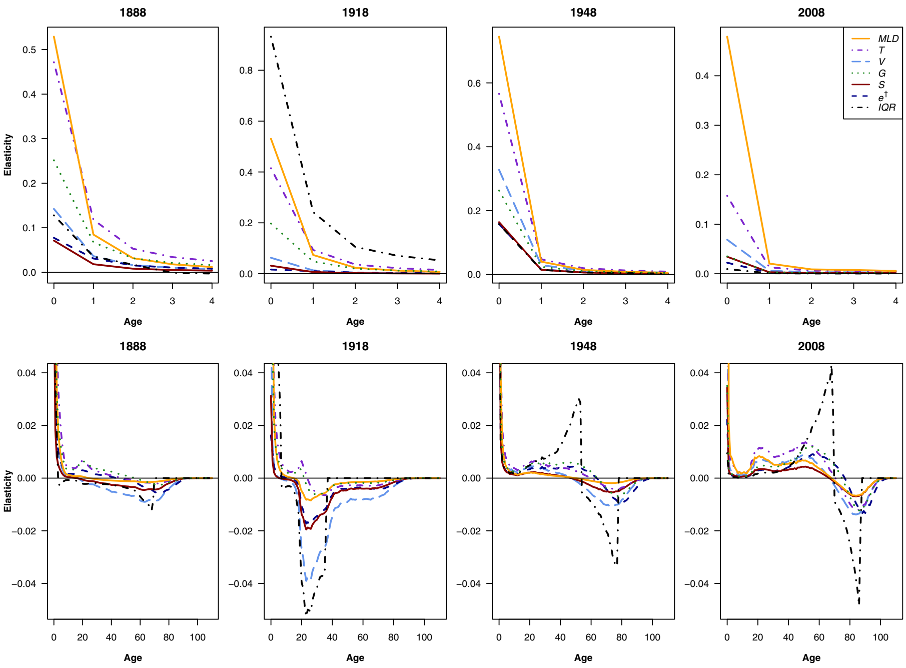

Fig. 3 The proportional change in the index from a 1 % change in mortality at each age on the x-axis. The first five ages were plotted separately in the top panel to more clearly delineate behavior of the indices at early and later ages. French males, period life table data from the Human Mortality Database (color figure online)

As we would expect from the high correlations between indices, the sensitivities follow similar general age patterns. The primary differences are in the sensitivity to infant mortality, the slope of the decline from birth to late adulthood, and in the age at which the sensitivities cross the x-axis (this age is the same for S and V). This age has been termed "the threshold age," or $a^\dagger$, because of its original derivation for the $e^\dagger$ index (Zhang and Vaupel 2009). Reductions in mortality below the threshold age reduce lifespan variation, while reductions above the threshold age increase the variation. Because it is the age at which the derivative is zero, it can be calculated using either the sensitivity or elasticity formulae in the Appendix. The threshold age has increased over time, and the differences between threshold ages of the indices have considerably diminished.

In general, conditioning on survival to age 10 resulted in only minor changes to the pattern of the sensitivity of each index to age-specific mortality, although it did remove some of the differences between indices found when examined from birth (see Fig. 6 in the Appendix). This was particular the case for the *MLD* and *T* indices, which are so highly sensitive to changes at birth that changes at other ages are largely masked. The *IQR* index produces the most unique sensitivity patterns. It is sensitive to transfers between quartiles but not to transfers within quartiles. Transfers, of course, are an awkward concept in mortality research, particularly because there are no finite life years that need to be distributed within the population. In practice, however, the idea of age rationing in health care (i.e., sacrificing facilities and medicine for older individuals to save younger individuals) comes close.

We imagined a scenario of targeted interventions leading to mortality reduction in order to examine how the different sensitivity profiles of the various indices could affect our assessment of whether a population is becoming more egalitarian in its ages at death. Using French male data from 1888 and 2008, we calculated the threshold age and the percentage change in the index from a 10 % decrease in death rates over selected age ranges. This was done for the indices calculated at birth and conditional on survival to age 10 (Table 3). The largest differences between the indices occurred for mortality change at the youngest age ranges, particularly for the 1888 age-at-death distribution, in which early death was more common. In the modern distribution, differences between indices were also large over the middle adult age range of 60–80. Indices with younger threshold ages such as *MLD*, *S*, and *V* found

Table 3 The threshold age (columns 2 and 3) and the percentage increase in each index resulting from a 10 % reduction in mortality over the given age ranges, for indices calculated from birth (top panel) and conditional upon survival to age 10 (bottom panel). Results were obtained by reducing mortality by 10 % at each age, using French male period life table data from Human Mortality Database, and then recomputing the indices

	Threshold Age		Percentage Increase in Index From Mortality Reduction at Ages:							
			0–5		20–40		60–80		85+	
Index	1888	2008	1888	2008	1888	2008	1888	2008	1888	2008
$e^\dagger$	43.4	76.2	−2.7	−0.3	−0.3	−0.6	1.0	−1.0	0.2	2.1
G	56.9	74.9	−4.6	−0.4	−0.8	−0.9	0.4	−0.9	0.0	1.2
IQR	2.8	69.7	−3.8	−0.1	0.4	−0.5	1.1	−2.7	0.0	1.6
MLD	24.8	68.7	−6.2	−5.1	0.1	−1.2	0.3	0.2	0.0	0.9
S	16.6	67.1	−1.4	−0.6	0.3	−0.8	0.9	0.3	0.0	0.9
T	46.3	69.7	−8.0	−1.9	−0.6	−1.9	0.4	−0.2	0.0	1.2
V	16.6	67.1	−2.8	−1.1	0.6	−1.6	1.9	0.6	0.1	1.9
Range	54.1	9.1	6.6	5.0	1.4	1.4	1.6	3.4	0.2	1.2
	Threshold Age		Percentage Increase in Index From Mortality Reduction at Ages:							
			10–15		20–40		60–80		85+	
Index	1888	2008	1888	2008	1888	2008	1888	2008	1888	2008
$e^\dagger_{10}$	57.2	76.7	−0.4	0.0	−1.8	−0.7	1.5	−1.0	0.3	2.2
G_{10}	58.8	75.1	−0.6	−0.1	−2.7	−1.0	0.9	−1.1	0.1	1.3
IQR_{10}	43.1	69.8	−0.5	0.0	−4.2	−0.4	2.8	−2.8	0.0	1.7
MLD_{10}	45.5	68.6	−1.6	−0.3	−3.3	−2.7	1.7	0.2	0.1	2.0
S_{10}	42.5	67.2	−0.4	−0.1	−1.2	−1.0	1.5	0.3	0.1	1.0
T_{10}	50.5	70.5	−1.3	−0.2	−3.8	−2.4	1.7	−0.2	0.1	1.5
V_{10}	42.5	67.2	−0.9	−0.2	−2.5	−1.9	3.1	0.5	0.2	2.1
Range	16.3	9.5	1.3	0.3	3.0	2.3	2.2	3.4	0.3	1.1

that mortality reduction over these ages increased lifespan variation, whereas the other indices all measured a decrease in variability. In the historic population, there were some differences between the responses of indices calculated from birth and from age 10. The initial age made relatively little difference in calculations for the modern population.

Decomposition of Temporal Trends in Variability

Our perturbation results make it possible to decompose differences or changes in an index into contributions from differences or changes in any of the parameters. The approach, known in population biology as life table response experiment (LTRE) analysis, has been widely used (see the review in Caswell 2001: chap. 10). It applies to any demographic statistic for which the sensitivity to the underlying vital rates can be calculated. The version used here is closely related to the decomposition method independently derived and described by Horiuchi et al. (2008).

Here, we use LTRE analysis to decompose temporal changes in the indices of lifespan variability into contributions from changes in age-specific mortality rates for Russian males from 1958 to 2006.

Let y be an index, let θ be a vector of parameters (mortality rates in our application), and let t denote time. The decomposition proceeds from noting that, to first order,

$$y(t + \Delta t) \approx y(t) + \frac{dy}{d\theta^{\top}} \frac{d\theta}{dt} \Delta t. \tag{15}$$

The product of the two derivatives in (15) gives the overall change in y attributable to the changes in all the parameters over the interval Δt. Thus, the contributions to that change are given by the entries of the vector

$$\mathbf{c}(t) = \left(\frac{dy}{d\theta^{\top}} \right)^{\top} \circ \left(\frac{d\theta}{dt} \right). \tag{16}$$

These contributions can be integrated over time to obtain the contributions to the change in y from t_0 to t_1,

$$y(t_1) \approx y(t_0) + \sum_{i=t_0}^{t_1} \mathbf{c}(i). \tag{17}$$

We computed the rate of change in the parameters, $d\theta/dt$, using the MATLAB function `gradient`, which uses a central difference algorithm to compute the derivatives.

The sequence of age-specific mortality changes experienced by Russian males (Anand et al. 2001; Shkolnikov et al. 2003) provides an interesting example of decomposition analysis. From 1958 to 2006, infant mortality declined substantially, from around 47 to 12 deaths per thousand live births. This decline was particularly rapid from 1958 to 1968. At the same time, adult mortality, especially between ages

40 and 50 years, fluctuated greatly. Adult mortality increased slowly but steadily until the mid-1980s, declined rapidly between 1984 and 1987 following the anti-alcohol campaigns, and then increased steeply with the mortality crisis brought on by the upheavals of transition to a market economy (Leon et al. 1997).

Figure 4 shows each index relative to its level in 1959, calculated from birth. Apart from the *IQR*, all indices show that lifespan variation decreased during the period, with large fluctuations in the interim. Among all the indices, *S* and *V* showed the least volatility, whereas *T*, *G*, and *IQR* showed the most.

Figure 5 presents the LTRE decomposition of life expectancy and of the indices, relative to their starting value, computed using Eq. (17). Reductions in infant, child, and adolescent mortality led to gains in life expectancy, but increased mortality of adults aged 20–70 tempered these gains. The positive contributions to life expectancy from reduced adult mortality during 1984–1987 are also visible.

The indices of variability show a different pattern. Their changes are a balance of strong negative contributions from infant and (to a lesser extent) child mortality, and positive contributions from changes in mortality among ages 20–50. Thus, the change in variability (no matter how it is measured) is a balance of contributions from these two age ranges.

Discussion

Summary of Results

Each of the seven indices of lifespan variation now possesses a directly computable sensitivity and elasticity to changes in age-specific mortality, derived from a

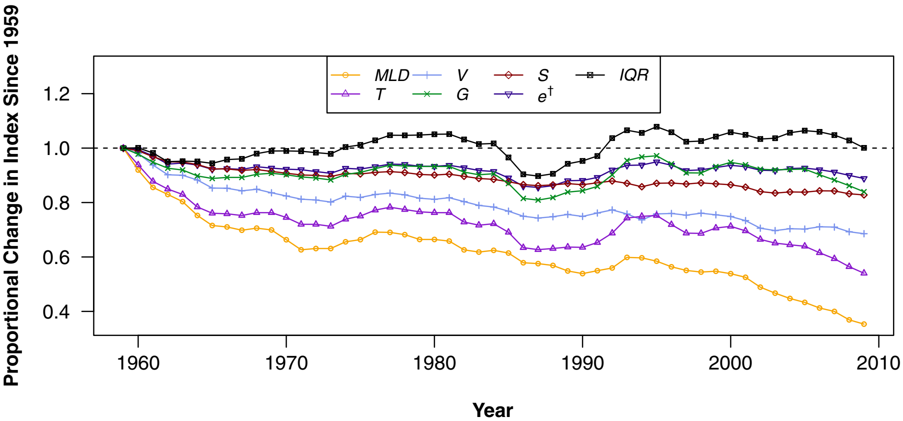

Fig. 4 The indices of lifespan variability for Russian males, measured relative to their values in 1959. All indices were calculated from period life table data, 1959–2008, from the Human Mortality Database (color figure online)

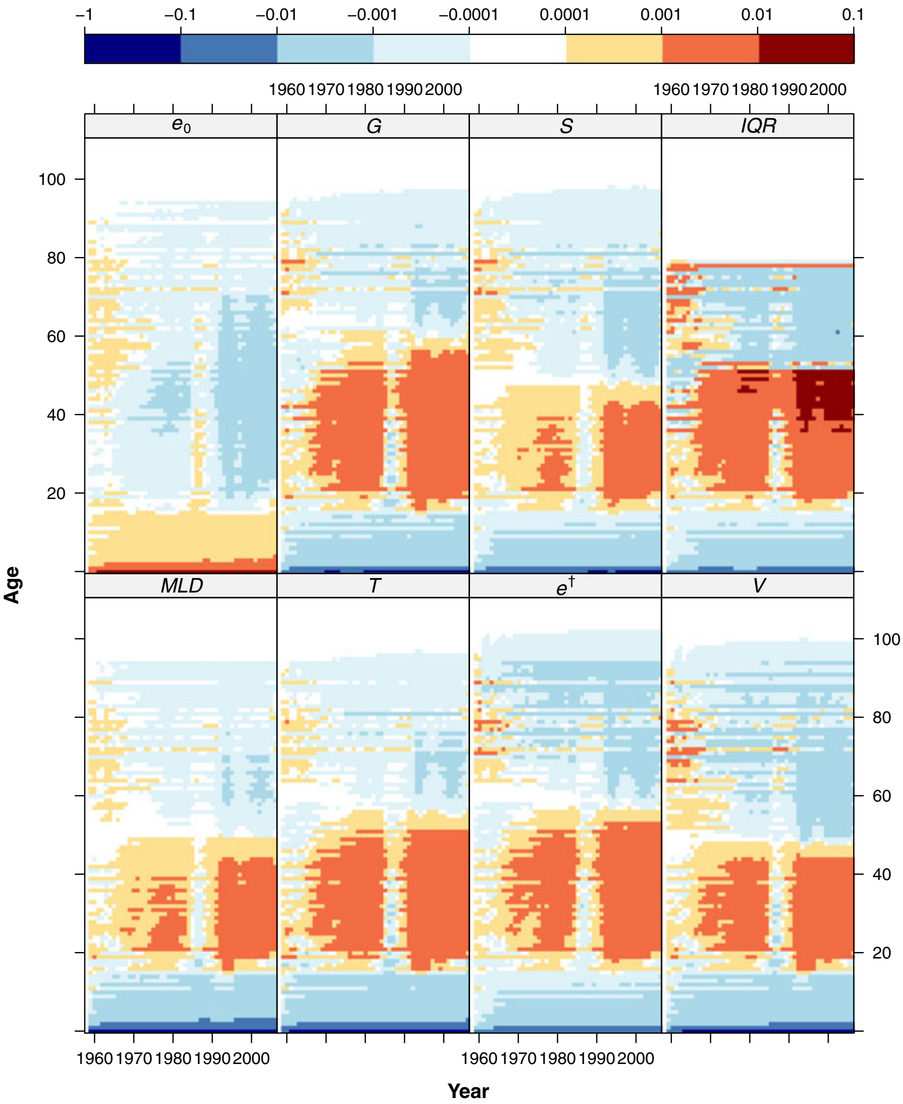

Fig. 5 The proportional contribution of changes in age-specific mortality to the changes in life expectancy, e_0, and in each index, measured relative to their values in 1959. Values were calculated by integrating the LTRE contributions calculated from Eq. (15). Note that the color scale changes by a factor of 10 in order to make contributions from all ages visible on the graphs. Calculations were based on period life table data for Russian males, 1959–2008, from the Human Mortality Database (color figure online)

consistent Markov chain formulation, with Matlab and R codes supplied in Online Resources 1 and 2 for calculation purposes. These formulae can be applied to any life table mortality schedule—period or cohort. The sensitivity patterns, applied to period data, showed similar broad patterns. Over time, the indices have become more sensitive to infant and child mortality, with *MLD*, *T*, and *V* being especially sensitive. Although earlier mortality regimes produced large differences in the threshold age of the indices, this age was similar for all indices over recent mortality schedules.

Indices of Variability and Distributive Justice

It is important to remember that variation in age at death in itself does not imply heterogeneity of individuals or inequality in the conditions they experience. Indeed, in the extreme case in which every individual experienced identical age-constant mortality hazards, the standard deviation of longevity and the life disparity would both equal the life expectancy. This variation, attributable to individual stochasticity (Caswell 2009), greatly exceeds the variation in longevity observed in actual populations. However, although individual stochasticity leads to variation in ages at death, it is unlikely that the current distribution is due entirely to individual stochasticity. If this were the case, we would not expect to find larger variation among lower socioeconomic groups (van Raalte et al. 2011) or large differences in variation between national populations with similar average mortality levels (Edwards and Tuljapurkar 2005; Smits and Monden 2009; Vaupel et al. 2011). Danish twin studies have estimated that genetic processes account for about a quarter of the the total variation in lifespan (Herskind et al. 1996; McGue et al. 1993). Although this proportion is substantial, it still leaves a large unexplained component.

To the extent that variation does represent the result of inequality in conditions, it may be interesting to think of it in terms of concepts of distributive justice. Distributive justice is concerned with how a society or group should allocate its scarce resources or products among individuals with competing needs or claims (Roemer 1996). Different notions of inequality arise from two different types of diversity: basic individual heterogeneity and assessment of inequality in terms of different variables (Sen 1992).

Most distributive justice frameworks can broadly fit into four categories: pro-poor, egalitarian, utilitarian, and minimum threshold frameworks (Marchand et al. 1998). An interesting line of future research would be to view changes in the age-at-death distribution through the lens of distributive justice. To this extent, perturbation analysis could be used to guide the selection of an index depending on how the sensitivity results align with the aimed-for distributive justice approach. Such delineations are dependent on the particular mortality schedule. As can be seen in Fig. 2, the interquartile range was most sensitive to infant mortality in the French male distribution in 1918 and was least sensitive in 1948 and 2008.

After a definition of inequality (and corresponding index) is chosen, perturbation results could be used further to target policy interventions to ages where reductions in mortality best reduce inequality.

Conclusion

We compared seven indices of lifespan variation, all of which largely correlated with one another in the mortality schedules found in the 7,516 unique life tables of the Human Mortality Database. Using matrix differentiation techniques, we derived the expressions for the sensitivities and elasticities of all indices. We highlighted four key uses of these perturbation results: (1) quantifying the effect of an age-specific mortality intervention on lifespan variation; (2) analytically determining the threshold age separating deaths at ages that increase or decrease variation in ages at death; (3) providing the output for a LTRE decomposition analysis that quantifies the changes in any lifespan variation index into contributions from mortality change at each age; and (4) comparing indices on the basis of their sensitivities to changes in age-specific mortality.

Acknowledgments Alyson van Raalte is supported by the Max Planck Society. Hal Caswell acknowledges financial support from the U.S. National Science Foundation (DEB-0816514 and SES-1156378), the Woods Hole Oceanographic Institution, and a Research Award from the Alexander von Humboldt Foundation, as well as the hospitality of the Max Planck Institute for Demographic Research. The International Max Planck Research School of Demography provided support for a course (IMPRSD 133) in which some of these ideas were first explored. The authors thank Jim Vaupel, Johan Mackenbach, Mikko Myrskylä, and two anonymous reviewers for helpful comments. An earlier version of this article was presented at the 2012 annual meeting of the Population Association of America in San Francisco, CA.

Appendix: Sensitivity Results and Derivations

This appendix provides details of the derivations. We begin with a summary of some basic matrix calculus techniques, and then present the derivations of the sensitivities of the indices of lifespan variation.

Matrix Calculus Preliminaries

The indices of lifespan variation in Table 3 are functions of scalars, vectors, and matrices. Matrix calculus permits differentiation of all three. The derivative of a scalar y with respect to a scalar x is the derivative $\frac{dy}{dx}$ familiar from basic calculus. The derivative of a $n \times 1$ vector $\mathbf{y}$ with respect to a scalar x is the $n \times 1$ vector

$$\frac{d\mathbf{y}}{dx} = \begin{pmatrix} \dfrac{dy_1}{dx} \\ \vdots \\ \dfrac{dy_n}{dx} \end{pmatrix}. \tag{18}$$

The derivative of a scalar y with respect to a $m \times 1$ vector $\mathbf{x}$ is the $1 \times m$ gradient vector

$$\frac{dy}{d\mathbf{x}^\top} = \left(\frac{\partial y}{dx_1} \cdots \frac{\partial y}{dx_m} \right). \tag{19}$$

The derivative of an $n \times 1$ vector $\mathbf{y}$ with respect to a $m \times 1$ vector $\mathbf{x}$ is the $n \times m$ Jacobian matrix, whose (i, j) entry is the derivative of y_i with respect to x_j:

$$\frac{d\mathbf{y}}{d\mathbf{x}^\top} = \left(\frac{dy_i}{dx_j} \right). \tag{20}$$

The derivatives of matrices are computed by transforming the matrices into column vectors using the vec operator and applying the rules for vector differentiation. Thus, the derivative of the $m \times n$ matrix $\mathbf{Y}$ with respect to the $p \times q$ matrix $\mathbf{X}$ is the $mn \times pq$ matrix

$$\frac{d\mathbf{Y}}{d\mathbf{X}} = \frac{d\text{vec}\mathbf{Y}}{d\text{vec}^\top\mathbf{X}}. \tag{21}$$

For notational simplicity, we denote $(d\text{vec}\mathbf{X})^\top$ as $d\text{vec}^\top\mathbf{X}$.

These definitions imply the chain rule for matrix calculus; if $\mathbf{Y}$ is a function of $\mathbf{X}$, and $\mathbf{X}$ is a function of $\mathbf{Z}$, then

$$\frac{d\text{vec}\mathbf{Y}}{d\text{vec}^\top\mathbf{Z}} = \frac{d\text{vec}\mathbf{Y}}{d\text{vec}^\top\mathbf{X}} \frac{d\text{vec}\mathbf{X}}{d\text{vec}^\top\mathbf{Z}}. \tag{22}$$

Matrix derivatives are constructed by forming differentials, where the differential of a matrix (or vector) is the matrix (or vector) of differentials of the elements; that is,

$$d\mathbf{X} = (dx_{ij}). \tag{23}$$

If, for some matrix $\mathbf{Q}$, it can be shown that

$$d\mathbf{y} = \mathbf{Q}d\mathbf{x}, \tag{24}$$

then according to the "first identification theorem" of Magnus and Neudecker (1985),

$$\frac{d\mathbf{y}}{d\mathbf{x}^\top} = \mathbf{Q}. \tag{25}$$

We will frequently obtain expressions of the form shown in Eq. (25) using a theorem originated by Roth (1934): if $\mathbf{Y} = \mathbf{ABC}$, then

$$\text{vec}\mathbf{Y} = (\mathbf{C}^\top \otimes \mathbf{A})\text{vec}\mathbf{B}. \tag{26}$$

We will also simplify expressions involving Kronecker products using

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD} \tag{27}$$

whenever $\mathbf{AC}$ and $\mathbf{BD}$ are defined.

More details on matrix calculus can be found in Magnus and Neudecker (1988). A good mathematical introduction can be found in Abadir and Magnus (2005), and demographic discussions appear in Caswell (2007, 2008, 2010).

Sensitivities of the Indices of Lifespan Variation

Preliminaries

Differentiating the various indices made use of the following sensitivities. The vector of life expectancies as a function of age is given by

$$\boldsymbol{\eta}^\top = \mathbf{e}^\top \mathbf{N}. \tag{28}$$

The derivative of this vector with respect to mortality is (Caswell 2006, 2009)

$$\frac{d\boldsymbol{\eta}}{d\boldsymbol{\theta}^\top} = \left(\mathbf{I} \otimes \mathbf{e}^\top\right)\left(\mathbf{N}^\top \otimes \mathbf{N}\right)\frac{d\mathrm{vec}\mathbf{U}}{d\boldsymbol{\theta}^\top}. \tag{29}$$

The life expectancy at birth is given by

$$\eta_1 = \boldsymbol{\eta}^\top \mathbf{e}_1, \tag{30}$$

and thus its derivative is

$$\frac{d\eta_1}{d\boldsymbol{\theta}^\top} = \left(\mathbf{e}_1^\top \otimes \mathbf{e}^\top\right)\frac{d\mathrm{vec}\mathbf{N}}{d\boldsymbol{\theta}^\top} \tag{31}$$

$$= \left(\mathbf{e}_1^\top \mathbf{N}^\top \otimes \mathbf{e}^\top \mathbf{N}\right)\frac{d\mathrm{vec}\mathbf{U}}{d\boldsymbol{\theta}^\top}. \tag{32}$$

The distribution of age at death is given by the vector

$$\mathbf{f} = \mathbf{M}\mathbf{N}\mathbf{e}_1. \tag{33}$$

Its derivative is given by (Caswell 2010),

$$\frac{d\mathbf{f}}{d\boldsymbol{\theta}^\top} = \left(\mathbf{e}_1^\top \mathbf{N}^\top \otimes \mathbf{I}\right)\frac{d\mathrm{vec}\mathbf{M}}{d\boldsymbol{\theta}^\top} + \left(\mathbf{e}_1^\top \mathbf{N}^\top \otimes \mathbf{B}\right)\frac{d\mathrm{vec}\mathbf{U}}{d\boldsymbol{\theta}^\top}. \tag{34}$$

The derivatives of **U** and **M** depend on the structure of the life cycle; in the age-classified case under consideration here, **M** contains the probabilities of death q_i on the diagonal, and **U** contains the probabilities of survival $1 - q_i$ on the subdiagonal.

Life Disparity, $\eta^\dagger$

The disparity can be written as

$$\eta^\dagger = \mathbf{f}^\top \boldsymbol{\eta}. \tag{35}$$

As shown in Caswell (2010, 2011),

$$d\eta^{\top} = \left(d\mathbf{f}^{\top}\right)\boldsymbol{\eta} + \mathbf{f}^{\top}\left(d\boldsymbol{\eta}\right), \tag{36}$$

and thus

$$\frac{d\eta^{\dagger}}{d\theta^{\top}} = \boldsymbol{\eta}^{\top}\frac{d\mathbf{f}}{d\theta^{\top}} + \mathbf{f}^{\top}\frac{d\boldsymbol{\eta}}{d\theta^{\top}}, \tag{37}$$

where $d\mathbf{f}/d\theta^{\top}$ is given by (34) and $d\boldsymbol{\eta}/d\theta^{\top}$ is given by (29).

Gini Coefficient

In matrix form, the Gini coefficient is given by

$$G = 1 - \frac{1}{\eta_1}\mathbf{e}^{\top}\left[\boldsymbol{\ell}\circ\boldsymbol{\ell}\right], \tag{38}$$

where the survivorship vector is

$$\boldsymbol{\ell} = \mathbf{e} - \mathbf{C}\mathbf{f}. \tag{39}$$

Differentiating (38), noting that η_1 is a scalar, gives

$$dG = \frac{1}{\eta_1^2}\mathbf{e}^{\top}\left(\boldsymbol{\ell}\circ\boldsymbol{\ell}\right)d\eta_1 - \frac{2}{\eta_1}\mathbf{e}^{\top}\left[\boldsymbol{\ell}\circ\left(d\boldsymbol{\ell}\right)\right]. \tag{40}$$

We apply the vec operator to both sides of (40) and obtain

$$dG = \frac{1}{\eta_1^2}\mathbf{e}^{\top}\left(\boldsymbol{\ell}\circ\boldsymbol{\ell}\right)d\eta_1 - \frac{2}{\eta_1}\ell^{\top}d\boldsymbol{\ell}. \tag{41}$$

Differentiating (39) gives $d\boldsymbol{\ell} = -\mathbf{C}d\mathbf{f}$; substituting this into (41) and using the chain rule gives

$$\frac{dG}{d\theta^{\top}} = \frac{1}{\eta_1^2}\mathbf{e}^{\top}\left(\boldsymbol{\ell}\circ\boldsymbol{\ell}\right)\frac{d\eta_1}{d\theta^{\top}} + \frac{2}{\eta_1}\ell^{\top}\mathbf{C}\frac{d\mathbf{f}}{d\theta^{\top}}, \tag{42}$$

where $d\eta_1/d\theta^{\top}$ is given by (32)

Mean Logarithmic Deviation

The mean logarithmic deviation in matrix notation is

$$MLD = \mathbf{f}^{\top}\left[\mathbf{e}\log\eta_1 - \log\mathbf{x}\right], \tag{43}$$

where the logarithm is applied elementwise. Differentiating (43) gives

$$dMLD = \left(d\mathbf{f}^{\top}\right)\left[\mathbf{e}\log\eta_1 - \log\mathbf{x}\right] + \mathbf{f}^{\top}\mathbf{e}\left(d\log\eta_1\right). \tag{44}$$

However, $\mathbf{f}^\top \mathbf{e} = 1$ because $\mathbf{f}$ is a probability distribution. Using this fact and also noting that $d \log \eta_1 = (1/\eta_1) d\eta_1$, we obtain

$$\frac{dMLD}{d\theta^\top} = \left[\mathbf{e}^\top \log \eta_1 - \log \mathbf{x}^\top\right] \frac{d\mathbf{f}}{d\theta^\top} + \frac{1}{\eta_1} \frac{d\eta_1}{d\theta^\top}, \tag{45}$$

where $d\eta_1/d\theta^\top$ is given by (32) and $d\mathbf{f}/d\theta^\top$ is given by (34).

Theil's Index

The expression for Theil's index in matrix notation is

$$T = \mathbf{f}^\top \left[\frac{\mathbf{x}}{\eta_1} \circ \log \frac{\mathbf{x}}{\eta 1}\right], \tag{46}$$

where the logarithm is applied elementwise. Differentiating (46) term by term yields

$$dT = \left(d\mathbf{f}^\top\right) \left[\frac{\mathbf{x}}{\eta_1} \circ \log \frac{\mathbf{x}}{\eta 1}\right] + \mathbf{f}^\top \left[d\left(\frac{\mathbf{x}}{\eta_1}\right) \circ \log \frac{\mathbf{x}}{\eta_1}\right] + \mathbf{f}^\top \left[\frac{\mathbf{x}}{\eta_1} \circ d\left(\log \frac{\mathbf{x}}{\eta_1}\right)\right]. \tag{47}$$

However,

$$d\left(\frac{\mathbf{x}}{\eta_1}\right) = -\frac{\mathbf{x}}{\eta_1^2} d\eta_1 \tag{48}$$

$$\begin{aligned} d\left(\log \frac{\mathbf{x}}{\eta_1}\right) &= d\left(\log \mathbf{x} - \mathbf{e} \log \eta_1\right) \\ &= -\frac{\mathbf{e}}{\eta_1} d\eta_1. \end{aligned} \tag{49}$$

Substituting (48) and (49) into (47), and transposing the first term, gives

$$dT = \left(\frac{\mathbf{x}^\top}{\eta_1} \circ \log \frac{\mathbf{x}^\top}{\eta_1}\right) d\mathbf{f} - \mathbf{f}^\top \left[\left(\frac{\mathbf{x}}{\eta_1^2} \circ \log \frac{\mathbf{x}}{\eta_1}\right) + \left(\frac{\mathbf{x}}{\eta_1} \circ \frac{\mathbf{e}}{\eta_1}\right)\right] d\eta_1. \tag{50}$$

Simplifying Eq. (50) and expressing the result in terms of a parameter vector θ gives

$$\frac{dT}{d\theta^\top} = \left(\frac{\mathbf{x}^\top}{\eta_1} \circ \log \frac{\mathbf{x}^\top}{\eta_1}\right) \frac{d\mathbf{f}}{d\theta^\top} - \left(\frac{T}{\eta_1} + \frac{\mathbf{f}^\top \mathbf{x}}{\eta_1^2}\right) \frac{d\eta_1}{d\theta^\top}, \tag{51}$$

where $d\eta_1/d\theta^\top$ is given by (32) and $d\mathbf{f}/d\theta^\top$ is given by (34).

The Variance and Standard Deviation of Longevity

The variance in longevity, conditional on survival to age class i, is given by the vector $\mathbf{v}$, which satisfies

$$\mathbf{v}^\top = \mathbf{e}^\top \mathbf{N} (2\mathbf{N} - \mathbf{I}) - \boldsymbol{\eta}^\top \circ \boldsymbol{\eta}^\top. \tag{52}$$

Caswell (2006, 2009, 2010) shows that

$$\frac{d\mathbf{v}}{d\theta^{\top}} = \left[2\left(\mathbf{N}^{\top}\otimes\mathbf{e}^{\top}\right)+2\left(\mathbf{I}\otimes\mathbf{e}^{\top}\mathbf{N}\right)-\left(\mathbf{I}\otimes\mathbf{e}^{\top}\right)\right]\frac{d\mathrm{vec}\mathbf{N}}{d\theta^{\top}} - 2\mathrm{diag}(\boldsymbol{\eta})\frac{d\boldsymbol{\eta}}{d\theta^{\top}}, \tag{53}$$

where $d\boldsymbol{\eta}/d\theta^{\top}$ is given by (29) and

$$\frac{d\mathrm{vec}\mathbf{N}}{d\theta^{\top}} = \left(\mathbf{N}^{\top}\otimes\mathbf{N}\right)\frac{d\mathrm{vec}\mathbf{U}}{d\theta^{\top}}. \tag{54}$$

The standard deviation of longevity is given by the vector

$$\mathbf{s} = \sqrt{\mathbf{v}}, \tag{55}$$

where the square root is taken elementwise and its sensitivity, as derived in Caswell (2010), is

$$\frac{d\mathbf{s}}{d\theta^{\top}} = \frac{1}{2}\mathrm{diag}\,(\mathbf{s})^{-1}\,\frac{d\mathbf{v}}{d\theta^{\top}}. \tag{56}$$

The Interquartile Range

The interquartile range is defined implicitly in terms of the distribution of ages at death. Let $f(x)$ be a probability density function and $F(x) = \int_{-\infty}^{x} f(s)\,ds$ be the cumulative distribution. The qth quantile is the value $\hat{x}$ satisfying

$$F\left(\hat{x}\right) = q. \tag{57}$$

Let $F\left(\hat{x}_1\right) = q_1$ and $F\left(\hat{x}_2\right) = q_2$, assuming that $q_2 > q_1$. The interquantile range is

$$R\,(q_1, q_2) = \hat{x_2} - \hat{x_1}. \tag{58}$$

The special case of the interquartile range refers to $R\,(0.25, 0.75)$.

Now we choose a set of probabilities of interest

$$\mathbf{q} = \begin{pmatrix} q_1 \\ \vdots \\ q_h \end{pmatrix}, \tag{59}$$

and let $\hat{\mathbf{x}}$ be the vector of quantiles that satisfy

$$F\left[\theta, \hat{\mathbf{x}}\,(\theta)\right] = \mathbf{q}, \tag{60}$$

where the distribution $f\,(\cdot)$ depends on a parameter vector θ of dimension $p \times 1$.

Next we differentiate Eq. (60) as follows:

$$\frac{\partial F}{\partial \boldsymbol{\theta}^{\top}} d\boldsymbol{\theta} + \frac{\partial F}{\partial \hat{\mathbf{x}}^{\top}} d\hat{\mathbf{x}} = 0, \tag{61}$$

and solve for $d\hat{\mathbf{x}}$, to obtain

$$d\hat{\mathbf{x}} = -\left(\frac{\partial F}{\partial \hat{\mathbf{x}}^{\top}}\right)^{-1} \left(\frac{\partial F}{\partial \boldsymbol{\theta}^{\top}}\right) d\boldsymbol{\theta}. \tag{62}$$

The first identification theorem implies that

$$\frac{d\hat{\mathbf{x}}}{d\boldsymbol{\theta}^{\top}} = -\left(\frac{\partial F}{\partial \hat{\mathbf{x}}^{\top}}\right)^{-1} \left(\frac{\partial F}{\partial \boldsymbol{\theta}^{\top}}\right). \tag{63}$$

The first term on the right side of of Eq. (63) is

$$\left(\frac{\partial F}{\partial \hat{\mathbf{x}}^{\top}}\right)^{-1} = \begin{pmatrix} \frac{1}{f(\hat{x}_1)} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{f(\hat{x}_h)} \end{pmatrix}, \tag{64}$$

while the second term is

$$\left(\frac{\partial F}{\partial \boldsymbol{\theta}^{\top}}\right) = \begin{pmatrix} \frac{\partial F(\hat{x}_1)}{\partial \theta_1} & \cdots & \frac{\partial F(\hat{x}_1)}{\partial \theta_p} \\ \vdots & & \vdots \\ \frac{\partial F(\hat{x}_h)}{\partial \theta_1} & \cdots & \frac{\partial F(\hat{x}_h)}{\partial \theta_p} \end{pmatrix}. \tag{65}$$

The product of Eqs. (64) and (65), following Eq. (63), gives

$$\left(\frac{d\hat{x}}{d\boldsymbol{\theta}^{\top}}\right) = -\begin{pmatrix} \frac{1}{f(\hat{x}_1)} \frac{\partial F(\hat{x}_1)}{\partial \theta_1} & \cdots & \frac{1}{f(\hat{x}_1)} \frac{\partial F(\hat{x}_1)}{\partial \theta_p} \\ \vdots & & \vdots \\ \frac{1}{f(\hat{x}_h)} \frac{\partial F(\hat{x}_h)}{\partial \theta_1} & \cdots & \frac{1}{f(\hat{x}_h)} \frac{\partial F(\hat{x}_h)}{\partial \theta_p} \end{pmatrix}. \tag{66}$$

The sensitivity of the interquantile range is the difference between row j and row i of (66):

$$\frac{dR_{(i,j)}}{d\boldsymbol{\theta}^{\top}} = \frac{d\hat{x}_j}{d\boldsymbol{\theta}^{\top}} - \frac{d\hat{x}_i}{d\boldsymbol{\theta}^{\top}}. \tag{67}$$

When $\mathbf{f}(x)$ is a discrete distribution, the quantiles will have to be interpolated. This is what we did to find the sensitivity of the *IQR* with quartiles $\hat{x}_3$ and $\hat{x}_1$.

Appendix Figures Depicting the Sensitivity and Elasticity of Indices Calculated From Age 10

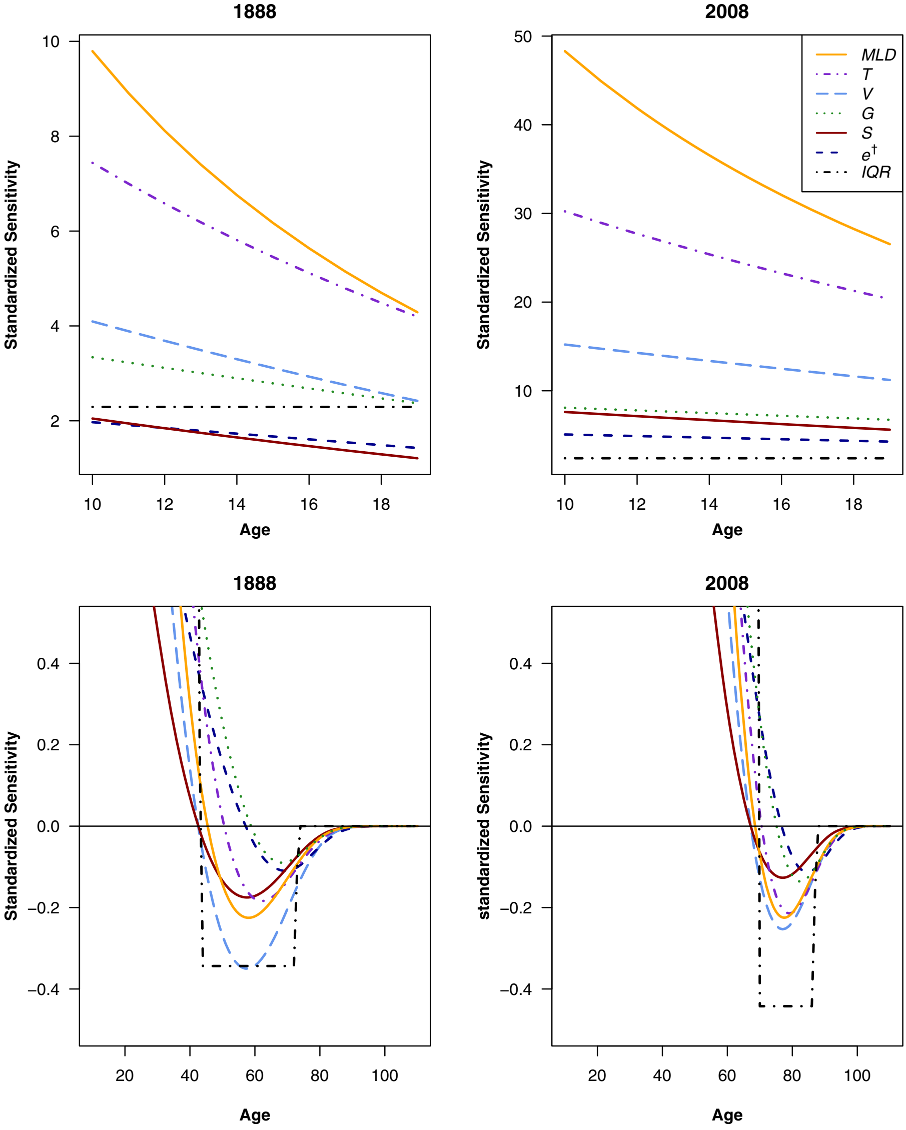

Fig. 6 The sensitivity of each index conditional on survival to age 10 with respect to mortality change at different ages. The sensitivities were standardized to the value of each index, i.e. $y_{10} \frac{dy}{d\theta}$, to make them comparable. Note the difference in scale between the top and bottom panels, plotted separately to more clearly delineate behavior of the indices at early and later ages. French males, period life table data from the Human Mortality Database (color figure online)

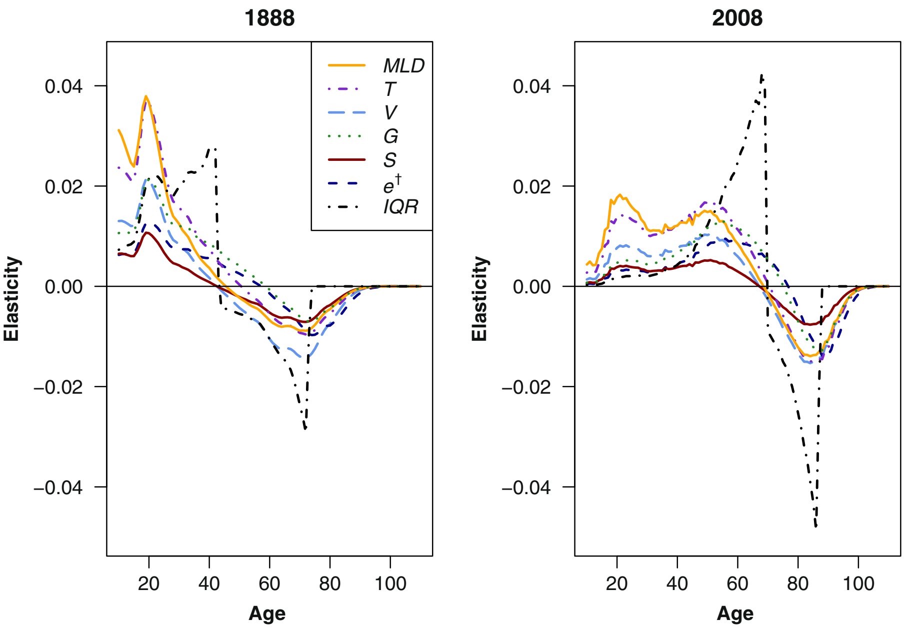

Fig. 7 The proportional change in the index calculated conditional upon survival to age 10 from a 1 % change in mortality at each age on the x-axis. French males, period life table data from the Human Mortality Database (color figure online)

References

Abadir, K. M., & Magnus, J. R. (2005). *Matrix algebra: Econometric exercises*. Cambridge, UK: Cambridge University Press.

Allison, P. D. (1978). Measures of inequality. *American Sociological Review, 43,* 865–880.

Anand, S., Diderichsen, F., Evans, T., Shkolnikov, V., Wirth, M. (2001). Measuring disparities in health: Methods and indicators. In *Challenging inequities in health: From ethics to action* (pp. 49–67). New York: Oxford University Press.

Asada, Y. (2007). *Health inequality: Morality and measurement*. Toronto, Canada: University of Toronto Press.

Brown, D. C., Hayward, M., Montez, J. K., Hummer, R. A., Chiu, C.-T., Hidajat, M. M. (2012). The significance of education for mortality compression in the United States. *Demography, 49,* 819–840.

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

Caswell, H. (1989). Analysis of life table response experiments I. Decomposition of effects on population growth rate. *Ecological Modelling, 46,* 221–237.

Caswell, H. (2001). *Matrix population models: Construction, analysis, and interpretation* (2nd ed.). Sunderland, MA: Sinauer.

Caswell, H. (2006). MAM2006: Markov anniversary meeting. In *Applications of Markov chains in demography* (pp. 319–334). Raleigh, NC: Boson Books.

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

Caswell, H. (2008). Perturbation analysis of nonlinear matrix population models. *Demographic Research, 18*(article 3), 59–116. doi:10.4054/DemRes.2008.18.3

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

Caswell, H. (2010). *Perturbation analysis of longevity using matrix calculus*. Paper presented at the annual meeting of the Population Association of America, Dallas, TX.

Caswell, H. (2011). Sensitivity analysis of discrete Markov chains via matrix calculus. *Linear Algebra and Its Applications*. doi:10.1016/j.laa.2011.07.046
Cheung, S. L., & Robine, J. M. (2007). Increase in common longevity and the compression of mortality: The case of Japan. *Population Studies, 61,* 85–97.
Cheung, S. L., Robine, J. M., Tu, E. J., Caselli, G. (2005). Three dimensions of the survival curve: Horizontalization, verticalization, and longevity extension. *Demography, 42,* 243–258.
Coale, A. J. (1972). *The growth and structure of human populations: A mathematical investigation*. Princeton, NJ: Princeton University Press.
Demetrius, L. (1969). The sensitivity of population growth rate to pertubations in the life cycle components. *Mathematical Biosciences, 4,* 129–136.
Eakin, T., & Witten, M. (1995). How square is the survival curve of a given species? *Experimental Gerontology, 30,* 33–64.
Edwards, R. D., & Tuljapurkar, S. (2005). Inequality in life spans and a new perspective on mortality convergence across industrialized countries. *Population and Development Review, 31,* 645–674.
Feichtinger, G. (1973). Markovian models for some demographic processes. *Statistical Papers, 14,* 310–334.
Gakidou, E., & King, G. (2002). Measuring total health inequality: Adding individual variation to group-level differences. *International Journal for Equity in Health, 1,* 1–12.
Gakidou, E. E., Murray, C. J. L., Frenk, J. (2000). Defining and measuring health inequality: An approach based on the distribution of health expectancy. *Bulletin of the World Health Organization, 78,* 42–54.
Goldman, N., & Lord, G. (1986). A new look at entropy and the life table. *Demography, 23,* 275–282.
Hamilton, W. (1966). The moulding of senescence by natural selection. *Journal of Theoretical Biology, 12,* 12–45.
Herskind, A., McGue, M., Holm, N., Srensen, T., Harvald, B., Vaupel, J. (1996). The heritability of human longevity: A population-based study of 2872 Danish twin pairs born 18701900. *Human Genetics, 97,* 319–323.
HMD (2012). Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Retrieved from www.mortality.org or www.humanmortality.de
Horiuchi, S., Wilmoth, J. R., Pletcher, S. D. (2008). A decomposition method based on a model of continuous change. *Demography, 45,* 785–801.
Kannisto, V. (2000). Measuring the compression of mortality. *Demographic Research, 3*(article 6), 1–24. doi:10.4054/DemRes.2000.3.6
Kannisto, V. (2001). Mode and dispersion of the length of life. *Population: An English Selection, 13,* 159–171.
Keyfitz, N. (1971). Linkages of intrinsic to age-specific rates. *Journal of the American Statistical Association, 66,* 275–281.
Keyfitz, N. (1977). *Applied mathematical demography* (Vol. 1). New York: Wiley.
Keyfitz, N., & Caswell, H. (2005). *Applied mathematical demography* (3rd ed.). New York: Springer-Verlag.
Lambert, P. J., & Aronson, J. R. (1993). Inequality decomposition analysis and the Gini coefficient revisited. *The Economic Journal, 103,* 1221–1227.
Leon, D. A., Chenet, L., Shkolnikov, V. M., Zakharov, S., Shapiro, J., Rakhmanova, G., ... McKee, M. (1997). Huge variation in Russian mortality rates 1984–94: Artefact, alcohol, or what? *The Lancet, 350,* 383–388.
Magnus, J., & Neudecker, H. (1988). *Matrix differential calculus with applications in statistics and econometrics*. Chichester, UK: Wiley.
Magnus, J. R., & Neudecker, H. (1985). Matrix differential calculus with applications to simple, Hadamard, and Kronecker products. *Journal of Mathematical Psychology, 29,* 474–492.
Marchand, S., Wikler, D., Landesman, B. (1998). Class, health, and justice. *The Milbank Quarterly, 76,* 449–467.
McGue, M., Vaupel, J. W., Holm, N. V., Harvald, B. (1993). Longevity is moderately heritable in a sample of Danish twins born 1870–1880. *The Journal of Gerontology, 48,* 237–244.
Mookherjee, D., & Shorrocks, A. (1982). A decomposition analysis of the trend in UK income inequality. *The Economic Journal, 92,* 886–902.

Nau, C., & Firebaugh, G. (2012). A new method for determining why length of life is more unequal in some populations than in others. *Demography, 49,* 1207–1230.

Ouellette, N., & Bourbeau, R. R. (2011). Changes in the age-at-death distribution in four low mortality countries: A nonparametric approach. *Demographic Research, 25*(article 19), 595–628. doi:10.4054/DemRes.2011.25.19

Pollard, J. (1982). The expectation of life and its relationship to mortality. *Journal of the Institute of Actuaries, 109,* 225–240.

Pollard, J. (1988). On the decomposition of changes in expectation of life and differentials in life expectancy. *Demography, 25,* 265–276.

Roemer, J. E. (1996). *Theories of distributive justice*. Cambridge, MA: Harvard University Press.

Roth, W. (1934). On direct product matrices. *Bulletin of the American Mathematical Society, 40,* 461–468.

Sen, A. (1992). *Inequality reexamined*. New York: Russell Sage Foundation; Cambridge, MA: Harvard University Press, 1997.

Shkolnikov, V. M., Andreev, E. E., Begun, A. Z. (2003). Gini coefficient as a life table function: Computation from discrete data, decomposition of differences and empirical examples. *Demographic Research, 8*(article 11), 305–358. doi:10.4054/DemRes.2003.8.11

Shorrocks, A. F. (1980). The class of additively decomposable inequality measures. *Econometrica, 48,* 613–625.

Smits, J., & Monden, C. (2009). Length of life inequality around the globe. *Social Science & Medicine, 68,* 1114–1123.

Thatcher, A. R., Cheung, S. L. K., Horiuchi, S., Robine, J.-M. (2010). The compression of deaths above the mode. *Demographic Research, 22*(article 17), 505–538. doi:10.4054/DemRes.2010.22.17

Theil, H. (1967). *Economics and information theory*. Amsterdam, The Netherlands: North-Holland Publishing Co.

van Raalte, A. A., Kunst, A. E., Lundberg, O., Leinsalu, M., Martikainen, P., Artnik, B., . . . Mackenbach, J. P. (2012). The contribution of educational inequalities to lifespan variation. *Population Health Metrics, 10,* 1–10.

van Raalte, A. A., Kunst, A. E., Deboosere, P., Leinsalu, M., Lundberg, O., Martikainen, P., . . . Mackenbach, J. P. (2011). More variation in lifespan in lower educated groups: Evidence from 10 European countries. *International Journal of Epidemiology, 40,* 1703–1714.

Vaupel, J. W. (1986). How change in age-specific mortality affects life expectancy. *Population Studies, 40,* 147–157.

Vaupel, J. W., & Canudas Romo, V. (2003). Decomposing change in life expectancy: A bouquet of formulas in honor of Nathan Keyfitz's 90th birthday. *Demography, 40,* 201–216.

Vaupel, J. W., Zhang, Z., van Raalte, A. A. (2011). Life expectancy and disparity. *BMJ Open, 1,* e000128. doi:10.1136/bmjopen–2011–000128

Wagner, P. (2010). Sensitivity of life disparity with respect to changes in mortality rates. *Demographic Research, 23*(3), 63–72.

WHO (2000). *The World Health Report 2000*. Geneva, Switzerland: World Health Organization.

Willekens, F. J. (1977). Sensitivity analysis in multiregional demographic models. *Environment and Planning A, 9,* 653–674.

Wilmoth, J. R., & Horiuchi, S. (1999). Rectangularization revisited: Variability of age at death within human populations. *Demography, 36,* 475–495.

Wrycza, T., & Baudisch, A. (2012). How life expectancy varies with perturbations in age-specific mortality. *Demographic Research, 27*(article 13), 365–375. doi:10.4054/DemRes.2012.27.13

Zhang, Z., & Vaupel, J. W. (2009). The age separating early deaths from late deaths. *Demographic Research, 20*(article 29), 721–730. doi:10.4054/DemRes.2009.20.29