Perturbation analysis of indices of lifespan variability

Alyson A. van Raalte[1,2], Hal Caswell[3,1]

[1]Max Planck Institute for Demographic Research, Rostock, Germany

[2]Department of Public Health, Erasmus MC, Rotterdam, the Netherlands

[3]Woods Hole Oceanographic Institution, Woods Hole MA, USA

January 13, 2012

Abstract

A number of indices have been used in recent years to calculate lifespan variation, each with different underlying properties. Although these indices are assumed to be interchangeable, little research has been conducted to show under which conditions this assumption is appropriate, or how to compare their responses to the underlying mortality schedule. We compare seven indices of lifespan variation: life disparity, the Gini coefficient, the standard deviation, the variance, Theil's index, the mean logarithmic deviation, and the inter-quartile range. We derive the sensitivity and elasticity of each index by applying Markov chain theory and matrix calculus. Using empirical French and Russian male data we compare the underlying sensitivities to mortality change under different mortality regimes in order to test under which conditions the indices might differ in their conclusions about the magnitude of lifespan variation. Finally we demonstrate how integrating these sensitivities can be used as a method of age decomposition. The result is an easily computable method for calculating the properties of this important class of longevity indices.

Contents

1 Introduction

The longevity experience of a cohort has long been summarized by its expectation, but recently attention has 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 non-human populations (Demetrius 1978; Eakin and Witten 1995; Edwards and Tuljapurkar 2005; Edwards 2011; Engelman, Canudas-Romo and Agree 2010; Go et al. 1995; Hill 1993; Kibele 2010; Le Grand 1987; Myers and Manton 1984; Nusselder and Mackenbach 1996; Paccaud et al. 1998; Robine 2001; Shkolnikov, Andreev and Begun 2003; Shkolnikov et al. 2010; Smits and Monden 2009; van Raalte et al. 2011; Vaupel, Zhang and van Raalte 2011; Wilmoth and Horiuchi 1999; Zureick 2010). 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. in press; Cheung et al. 2005; Cheung and Robine 2007; Kannisto 2000; 2001; Ouellette and Bourbeau 2011; Thatcher et al. 2010).

There are many indices of lifespan variation, which have been compared by Anand et al. (2001); Cheung et al. (2005); Kannisto (2000); Shkolnikov et al. (2003); Vaupel et al. (2011) and Wilmoth and Horiuchi (1999). These authors have found such high correlations among indices as to make them apparently interchangeable. Less attention has been paid to differences among the indices in their responses to perturbations in mortality schedules, or to understanding when indices can be expected to disagree. Rarely are reasons for choosing one index over another tied in to any normative concept of inequality or to social preference for the weights placed on deaths at different ages (Asada 2007; Anand et al. 2001; Gakidou, Murray and Frenk 2000). A notable exception to this is the WHO attempt to quantify inequality over 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).

Our aim in this paper is to derive and compare the sensitivities and elasticities to mortality changes of seven of the most commonly used indices of lifespan variability. We do this by reformulating the problem of lifespan variability in terms of an absorbing Markov chain, and using methods derived from matrix

calculus (Caswell 2007; 2008; 2009; 2010). These sensitivity results may permit researchers to better tie their choice of index to their research aims. They also make possible the decomposition of 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 where these different sensitivities cause indices to disagree on the magnitude or even direction of changes in lifespan variation.

2 Indices of lifespan variation

We make comparisons of the following indices of variability:

1. Life disparity $e^\dagger$

   Life disparity 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 due 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).

2. Gini coefficient G

   The Gini coefficient 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 inter-individual differences in age at death, divided by the life expectancy (Shkolnikov et al. 2003).

3. Theil's index T

4. Mean logarithmic deviation MLD

   Both T and 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 is needed and the entropy-based measures are zero.

5. Standard deviation of the distribution of age at death, S

6. Variance in the distribution of age at death, V

7. Inter-quartile range of the distribution of age at death, IQR

   The indices S, V and 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 due to their underlying formal properties. The variance V, Theil index T and the mean logarithmic deviation MLD are all additively decomposable into between- and within-group variation (Shorrocks 1980; van Raalte 2011). This decomposition can used to study the contribution of between-group differences to the total level of lifespan variation. The Gini coefficient G can also be decomposed in this way, but contains an overlap term (Lambert and Aronson 1993). The MLD index can additionally be additively decomposed over time, to account for compositional change to the between- and within-group variation components (Mookherjee and Shorrocks 1982; van Raalte 2011). The life disparity $e^{\dagger}$ index 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, as they are normally expressed in years.

Finally the sensitivity of indices to changes in mortality at different ages is perhaps the most important and least 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, usage of an index with a high sensitivity to early death would be appropriate. As economist Paul Allison (1978) 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 currently in the HMD (2011). 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.

	$e^\dagger$	G	T	MLD	S	V	IQR
$e^\dagger$	1.000						
G	0.977	1.000					
T	0.945	0.991	1.000				
MLD	0.964	0.991	0.992	1.000			
S	0.981	0.931	0.890	0.928	1.000		
V	0.987	0.943	0.907	0.941	0.996	1.000	
IQR	0.968	0.965	0.946	0.955	0.921	0.944	1.000

	$e^\dagger_{10}$	G_{10}	T_{10}	MLD_{10}	S_{10}	V_{10}	IQR_{10}
$e^\dagger_{10}$	1.000						
G_{10}	0.984	1.000					
T_{10}	0.979	0.995	1.000				
MLD_{10}	0.967	0.986	0.995	1.000			
S_{10}	0.986	0.958	0.961	0.952	1.000		
V_{10}	0.985	0.960	0.967	0.960	0.998	1.000	
IQR_{10}	0.981	0.978	0.978	0.969	0.958	0.965	1.000

Table 1: Pearson correlation coefficients between pairs of indices, calculated from birth (ages 0-110+) in the top panel and calculated conditional upon survival to age 10 (ages 10-110+) in the bottom panel, for all female and male life tables in the Human Mortality Database (6860 in total).

3 Markov chain formulations of longevity

To analyze these indices, we use a flexible and powerful formulation of 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). Since this study focuses on human demography, we focus on the age-classified model. Nevertheless, these models could be generalized to apply to stage-classified populations.

Notation. We use matrix notation in deriving the sensitivities. Matrices are denoted by upper case bold faced symbols (e.g., $\mathbf{X}$) and vectors by lower case bold faced symbols ($\mathbf{x}$); vectors are column vectors by default. The superscript $\intercal$ denotes the transpose. The symbol $\mathrm{diag}\,(\mathbf{x})$ 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; i.e., the vector with a 1 in the ith location and zeros elsewhere. The Hadamard, or element-by-element product is denoted by $\circ$ and the Kronecker product by $\otimes$. The vec operator (e.g., $\mathrm{vec}\,\mathbf{X}$) stacks the columns of a matrix into a column vector.

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

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

where q_i is the probability of death between ages i and $i+1$ from the life table.

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[1] for the Markov chain is

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

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

$$\tilde{\boldsymbol{\eta}}^{\mathsf{T}} = \mathbf{e}^{\mathsf{T}}\mathbf{N} \tag{5}$$

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

$$\mathbf{v}^{\mathsf{T}} = \mathbf{e}^{\mathsf{T}}\mathbf{N}\left(2\mathbf{N} - \mathbf{I}\right) - \boldsymbol{\eta}^{\mathsf{T}} \circ \boldsymbol{\eta}^{\mathsf{T}} \tag{7}$$

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

$$\mathbf{B} = \mathbf{M}\mathbf{N}. \tag{8}$$

[1]Note that $\mathbf{P}$ is column-stochastic and operates on column vectors, to agree 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 $\mathbf{B}$;

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

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 $\mathbf{f}$.

The vector $\mathbf{x}$ contains the average age at death in the age interval (i.e. for French males in 2005 it is $\{0.06, 1.5, 2.5, ..., 109.5, 111.32\}$).

In Table 2 we present the conventional lifetable notation alongside the less familiar matrix notation for each index. In conventional notation ℓ_y is survivorship, d_y the death density, and e_y 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, for example $\bar{e}_y$, is used when adjustments to the variable are necessary to account for the portion of the age interval lived by those who died, i.e.

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

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 the interpolated first and third age quartiles, at which 25 and 75 percent of the total deaths have occurred.

	conventional LT notation	**matrix notation**
$e^{\dagger}$	$\sum_{y=0}^{\omega} d_y \bar{e}_y$	$\mathbf{f}^{\mathsf{T}}\boldsymbol{\eta}$
G	$1 - \frac{1}{e_0}\sum_{y=0}^{\omega} \ell_{y+1}^2$	$1 - \frac{1}{\eta_1}\mathbf{e}^{\mathsf{T}}\left[(\mathbf{e}-\mathbf{Cf})\circ(\mathbf{e}-\mathbf{Cf})\right]$
T	$\sum_{y=0}^{\omega} d_y \left(\frac{\bar{x}_y}{e_0} ln \frac{\bar{x}_y}{e_0}\right)$	$\mathbf{f}^{\mathsf{T}}\left[\left(\frac{\mathbf{x}}{\eta_1}\right)\circ\left(\log\frac{\mathbf{x}}{\eta_1}\right)\right]$
MLD	$\sum_{y=0}^{\omega} d_y \left(ln \frac{e_0}{\bar{x}_y}\right)$	$\mathbf{f}^{\mathsf{T}}\left[\log\left(\eta_1\right)\mathbf{e} - \log\mathbf{x}\right]$
V	$\sum_{y=0}^{\omega} d_y \left(\bar{x}_y - e_0\right)^2$	$\left[\mathbf{e}^{\mathsf{T}}\mathbf{N}\left(2\mathbf{N}-\mathbf{I}\right) - \boldsymbol{\eta}^{\mathsf{T}}\circ\boldsymbol{\eta}^{\mathsf{T}}\right]^{\mathsf{T}}$
S	$\sqrt{V}$	$\sqrt{V}$
IQR	$\hat{x}_3 - \hat{x}_1$	$\hat{x}_3 - \hat{x}_1$

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

4 Sensitivity and elasticity analysis

Perturbation analysis was first introduced to demography in the 1960s and 1970s in assessing the sensitivity of population growth rates and life expectancy to changes in the underlying mortality rates (Caswell 1978; Demetrius 1969; Hamilton 1966; Keyfitz 1971). In recent years this work has been extended and further life table relationships have been derived (Goldman and Lord 1986; Pollard 1982; 1988; Vaupel 1986; Vaupel and Canudas Romo 2003). Widespread usage of perturbation analysis in demography, however, was somewhat limited by the complexity in deriving the analytic expressions for the derivatives of different indices and in its ability to handle complexities in life history. Expressing the problem in terms of an absorbing Markov chain and applying matrix calculus has greatly expanded the possibilities (Caswell 2001; 2008; 2009; 2010; Willekens 1977).

To assess respectively the absolute and proportional effects on the indices from 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. 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 paper.

As described in detail in Section 8.1, the sensitivity of a $n \times 1$ vector $\mathbf{y}$ to a $m \times 1$ vector of parameters $\boldsymbol{\theta}$ is given by the $n \times m$ vector

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

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

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

The formulas 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 calculated. We performed all calculations in MATLAB 7.3.0 and have attached the code as an appendix. The

sensitivities and elasticities of each index to age-specific mortality were derived using matrix calculus techniques (Magnus and Neudecker 1988). These techniques are also given extensive treatment in recent papers by Caswell, using most of the same notation that we have here (Caswell 2009; 2010). The derivation of the sensitivities of G, MLD, T, and IQR to mortality can be found in the appendix.

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.

5 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 latter distribution is interesting as the second mode is around young adulthood, and the distribution has a long right tail instead of the long left tail. To help visualize these differences, all four distributions are plotted in Figure 1.

All indices are highly sensitive to changes in infant mortality. For this reason we compared the sensitivities and elasticities at birth (Figures 2 and 3) and age 10 (see appendix). Given the different units for each index, the elasticities are perhaps intuitively easier to interpret. The y axis measures the proportional change in the index from a one percent change in mortality at each age on the x axis.

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). Improvements in mortality below this age reduce lifespan variation, while improvements after this age increase the variation. The age itself has been termed the threshold age or $a^\dagger$ due to

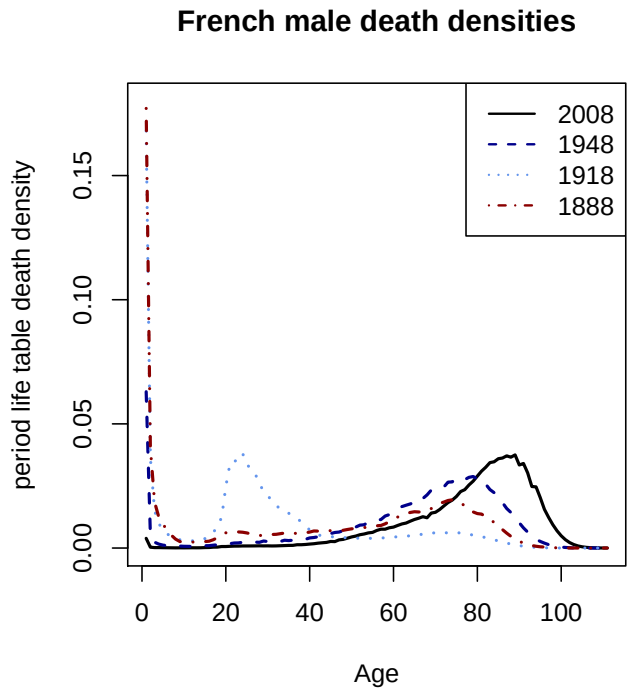

Figure 1: The French male death densities for which the sensitivities and elasticities of the indices are compared

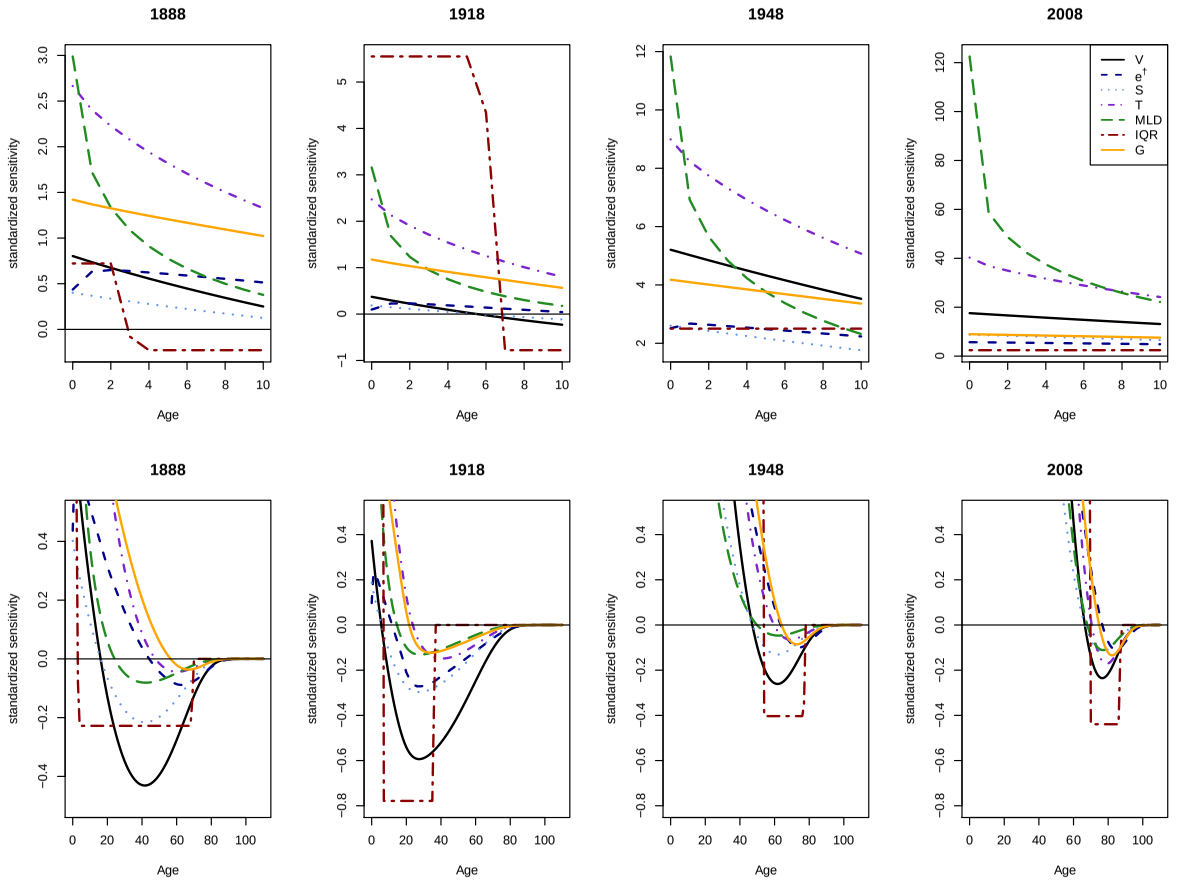

Figure 2: The sensitivity of each index with respect to mortality change at different ages. The sensitivities were standardized to the value of each index, i.e. $y_0 \frac{dy}{d\theta}$, to make them comparable. Note the difference in scale between the top and bottom panels, plotted separately to more clearly delineate behaviour of the indices at early and later ages. French males, period lifetable data from the HMD.

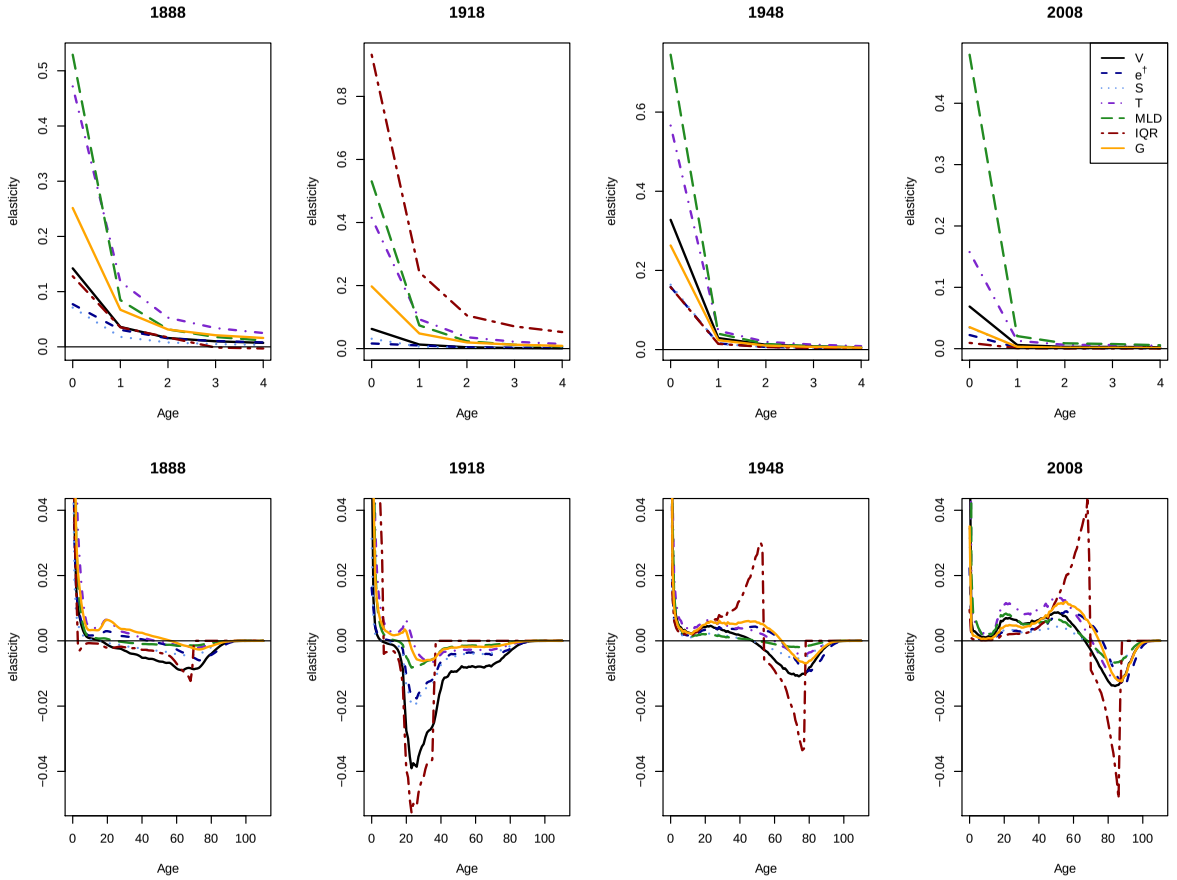

Figure 3: The proportional change in the index from a one percent change in mortality at each age on the x axis. The first five ages were plotted separately in the top panel to more clearly delineate behaviour of the indices at early and later ages. French males, period lifetable data from the HMD.

its original derivation for the $e^{\dagger}$ index (Zhang and Vaupel 2009). This age has pushed itself out to later and later ages with time, and the differences between threshold ages of the indices have considerably diminished.

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

To see how the different threshold ages and sensitivity profiles of the various indices could affect our assessment of whether a population is becoming more egalitarian in its ages at death, we imagined a scenario of targeted interventions leading to mortality reduction. Using French male data from 1888 and 2008, we calculated the threshold age and the percentage change in the index from a 10 percent decrease in death rates over selected age ranges. This was done for the indices calculated at birth and conditional upon 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 where early death was more common. In the modern distribution, differences between indices were also large over the middle adult age range 60-80. Indices with younger threshold ages such as MLD,S, and V found that mortality reduction over these ages increased lifespan variation, while the other indices all measured a decrease in variability. The different responses to mortality change over adult ages between using an index calculated at birth or at age 10 was large for the historic population, but made relatively little difference for the modern population.

	Threshold Age		Percent 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		Percent 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*

Table 3: The threshold age (columns 2 and 3) and the percentage increase in each index resulting from a 10 percent reduction in mortality over the given age ranges, for indices calculated from birth (top panel) and conditional upon survival to age 10 (bottom panel). French male period lifetable data from Human Mortality Database.

6 Decomposition of temporal trends in variability

Our perturbation results make it possible to decompose differences or changes in any of the indices into contributions from differences or changes in any of the parameters. The approach is known in population biology as Life Table Response Experiment (LTRE) analysis, and has been widely used (see review in Caswell 2001, Chapter 10). It applies to any demographic statistic for which the sensitivity to the underlying vital rates can be calculated (e.g., Caswell 2011). 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, $\boldsymbol{\theta}$ a vector of parameters (mortality rates in our application) and let t denote time. The decomposition proceeds from noting that, to first order,

$$y(t + \Delta t) \approx y(t) + \frac{dy}{d\boldsymbol{\theta}^\mathsf{T}} \frac{d\boldsymbol{\theta}}{dt} \Delta t \tag{15}$$

The product of the two derivatives in (15) sums the contributions of the changes in all the parameters to the change in y. Thus the contributions to that change are given by the entries of the vector

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

These contributions can be integrated 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, $d\boldsymbol{\theta}/dt$, 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 makes them an interesting example to examine how indices differ in their sensitivity (Anand et al. 2001; Shkolnikov et al. 2003). From 1958 to 2006, infant mortality declined substantially, from around 47 to 12 deaths per thousand

live births. This decline was particularly rapid from 1958–1968. At the same time, adult mortality, especially between ages 40 and 50 years, fluctuated a great deal. Mortality over these ages increased slowly but steadily until the mid-1980s, then 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 (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. The S and V showed less volatility than the other indices while T, G and IQR showed the most.

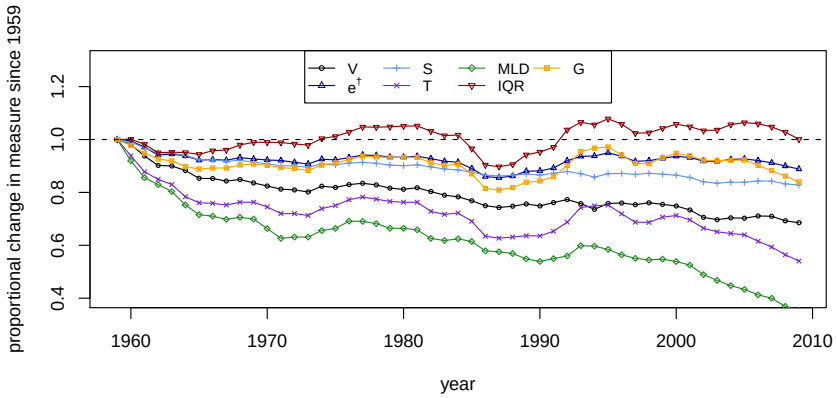

Figure 4: The indices of lifespan variability for Russian males, measured relative to their values in 1959. All indices were calculated from period lifetable data, 1959-2008, from the Human Mortality Database.

Figure 5 presents the LTRE decomposition of life expectancy and of the indices, relative to their starting value, computed using equation (17). Reductions in infant, child, and adolescent mortality led to gains in life expectancy, but increased mortality of adults age 20-70 tempered these gains. The positive contributions 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 mortality in ages 20-50. Thus the change in variability (no matter how it is measured) is a balance of contributions from these two age ranges.

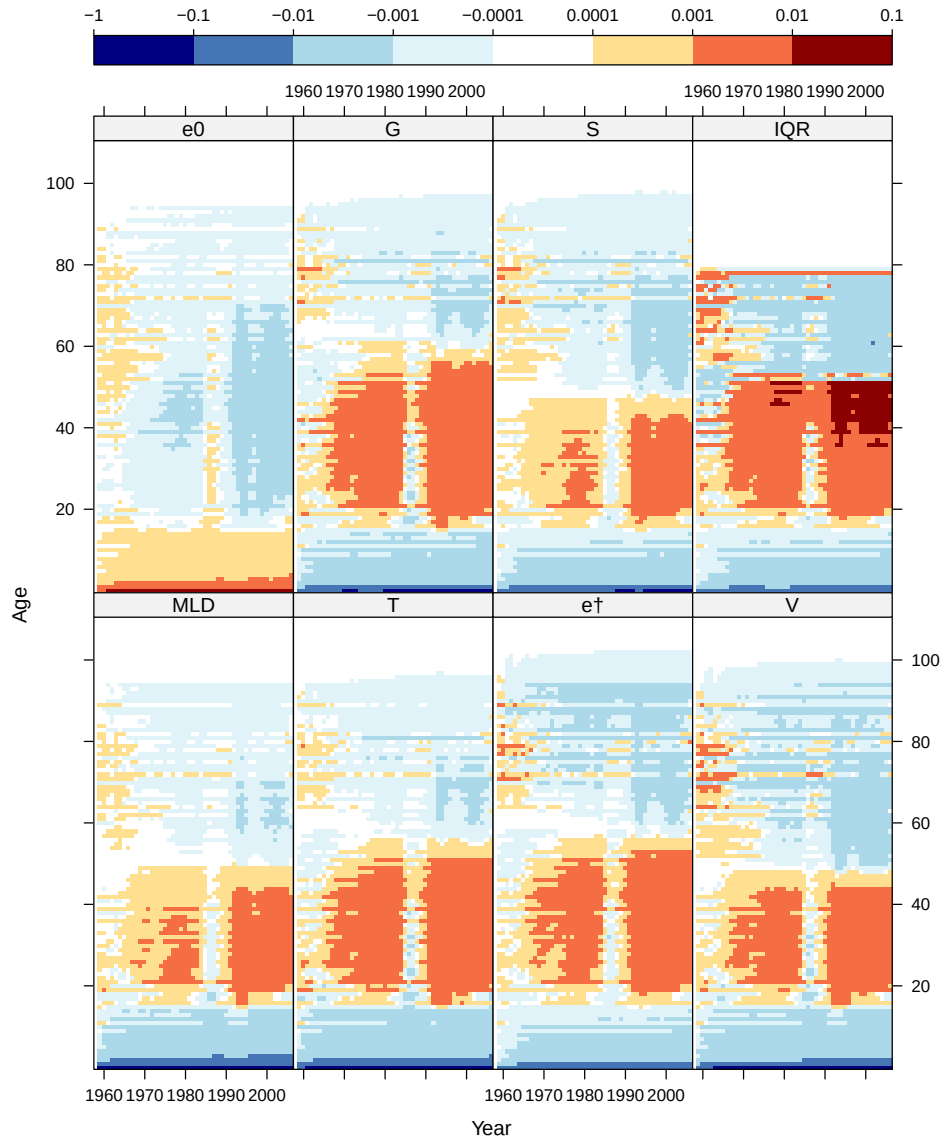

Figure 5: The proportional contribution of changes in age-specific mortality to the changes in life expectancy, e_0 and in each index, measured relative to their values in 1959. Values were calculated by integrating the LTRE contributions calculated from (15). Note that the color scale changes by a factor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

7 Conclusion

We compared seven indices of lifespan variation, all of which largely correlated with one another over the mortality schedules found in the 6860 lifetables of the Human Mortality Database. Using matrix differentiation techniques we derived the expressions for the sensitivities and elasticities of all indices. We compared these sensitivities under different age-at-death profiles and related changes over time in the indices to the underlying sensitivities through a LTRE decomposition.

The aim of this paper was not to come out in favor of any one method of measuring lifespan variation but rather to make explicit the differences in the underlying sensitivities of each index to age-specific mortality. This is essential for formulating any larger normative concept of inequality or variation. It is also clear from this analysis that some indices are better suited to certain tasks than others. The MLD, T and V indices are so sensitive to infant mortality that they are not ideal candidates for studies over the entire age range, if adult mortality is also of interest. In comparing distributions above childhood, however, they become more suitable indices, particularly if there is a strong aversion to death at younger versus older ages. The IQR differs the most from the other indices. Although it has great intuitive appeal, it can be expected to deviate from the other six indices of variation the most often. Unless a clearly defined concept of variation is specified outright, we would recommend using two or more indices with different sensitivity patterns before coming to any strong conclusions about the magnitude or direction of change in lifespan variability.

References

Abadir, K. M. and J. R. Magnus (2005). *Matrix Algebra: Econometric Exercises*. Cambridge University Press.

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

Anand, S., F. Diderichsen, T. Evans, V. Shkolnikov and M. Wirth (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: University of Toronto Press.

Brown, D. C., M. Hayward, J. K. Montez, R. A. Hummer, C.-T. Chiu and M. M. Hidajat (in press). "The significance of education for mortality compression in the United States." *Demography* pp. 1–45.

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.

——— (2001). *Matrix population models: Construction, analysis, and interpretation.* 2nd ed., Sunderland, Massachusetts, USA.: Sinauer.

——— (2006). *MAM2006: Markov Anniversary Meeting*, chap. Applications of Markov chains in demography, pp. 319–334. Raleigh, North Carolina, USA,: Boson Books.

——— (2007). "Sensitivity analysis of transient population dynamics." *Ecology Letters* 10:1–15.

——— (2008). "Perturbation analysis of nonlinear matrix population models." *Demographic Research* 18:59–116.

——— (2009). "Stage, age and individual stochasticity in demography." *Oikos* 118:1763–1782.

——— (2010). "Perturbation analysis of longevity using matrix calculus." *Paper presented at the 75th Annual Population Association of America meeting, in session 141: "Mathematical Aspects of Mortality and Longevity". Dallas, Texas* .

——— (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. and J. M. Robine (2007). "Increase in common longevity and the compression of mortality: the case of Japan." *Population Studies* 61:85–97.

Cheung, S. L., J. M. Robine, E. J. Tu and G. Caselli (2005). "Three dimensions of the survival curve: horizontalization, verticalization, and longevity extension." *Demography* 42:243–58.

Demetrius, L. (1969). "The sensitivity of population growth rate to pertubations in the life cycle components." *Mathematical Biosciences* 4:129 – 136.

——— (1978). "Adaptive value, entropy and survivorship curves." *Nature* 275:213–214.

Eakin, T. and M. Witten (1995). "How square is the survival curve of a given species?" *Experimental Gerontology* 30:33–64.

Edwards, R. D. (2011). "Changes in world inequality in length of life: 1970-2000." *Population and Development Review* 37:499–528.

Edwards, R. D. and S. Tuljapurkar (2005). "Inequality in life spans and a new perspective on mortality convergence across industrialized countries." *Population and Development Review* 31:645–674.

Engelman, M., V. Canudas-Romo and E. M. Agree (2010). "The implications of increased survivorship for mortality variation in aging populations." *Population and Development Review* 36:511–539.

Feichtinger, G. (1973). "Markovian models for some demographic processes." *Statistical Papers* 14:310–334, 10.1007/BF02923066.

Gakidou, E. and G. King (2002). "Measuring total health inequality: Adding individual variation to group-level differences." *International Journal for Equity in Health* 1:1–12.

Gakidou, E. E., C. J. L. Murray and J. Frenk (2000). "Defining and measuring health inequality: An approach based on the distribution of health expectancy." *Bulletin of the World Health Organization* 78:42–54.

Go, C. G., J. E. Brustrom, M. F. Lynch and C. M. Aldwin (1995). "Ethnic trends in survival curves and mortality." *Gerontologist* 35:318–326.

Goldman, N. and G. Lord (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.

Hill, G. (1993). "The entropy of the survival curve: An alternative measure." *Canadian Studies of Population* 20:43–57.

HMD (2011). "Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded on16/10/2011)."

Kannisto, V. (2000). "Measuring the compression of mortality." *Demographic Research* 3:1–24.

——— (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.

——— (1977). *Applied Mathematical Demography*, vol. 1st. New York: Wiley.

Kibele, E. (2010). *Regional mortality differentials in Germany: Spatial patterns, temporal trends, contextual and individual-level determinants*. Ph.D. Thesis, University of Rostock.

Lambert, P. J. and J. R. Aronson (1993). "Inequality decomposition analysis and the Gini coefficient revisited." *The Economic Journal* 103:1221–1227.

Le Grand, J. (1987). "Inequalities in health: Some international comparisons." *European Economic Review* 31:182–191.

Leon, D. A., L. Chenet, V. M. Shkolnikov, S. Zakharov, J. Shapiro, G. Rakhmanova, S. Vassin and M. McKee (1997). "Huge variation in Russian mortality rates 1984-94: Artefact, alcohol, or what?" *The Lancet* 350:383–388.

Magnus, J. and H. Neudecker (1988). *Matrix differential calculus with applications in statistics and econometrics*. John Wiley & Sons.

Magnus, J. R. and H. Neudecker (1985). "Matrix differential calculus with applications to simple, Hadamard, and Kronecker products." *J. Math. Psychol.* 29:474–492.

Mookherjee, D. and A. Shorrocks (1982). "A decomposition analysis of the trend in UK income inequality." *The Economic Journal* 92:886–902.

Myers, G. and K. Manton (1984). "Compression of mortality: myth or reality?" *Gerontologist* 24:346–53.

Nusselder, W. J. and J. P. Mackenbach (1996). "Rectangularization of the survival curve in The Netherlands, 1950-1992." *Gerontologist* 36:773–782.

Ouellette, N. and R. R. Bourbeau (2011). "Changes in the age-at-death distribution in four low mortality countries: A nonparametric approach." *Demographic Research* 25:595–628.

Paccaud, F., C. Sidoti Pinto, A. Marazzi and J. Mili (1998). "Age at death and rectangularisation of the survival curve: Trends in Switzerland, 1969-1994." *Journal of Epidemiology and Community Health* 52:412–415.

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

——— (1988). "On the decomposition of changes in expectation of life and differentials in life expectancy." *Demography* 25:265–276.

Robine, J.-M. (2001). "Redefining the stages of the epidemiological transition by a study of the dispersion of life spans: The case of France." *Population: An English Selection* 13(1):173–193.

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

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

Shkolnikov, V., E. M. Andreev, Z. Zhang, J. Oeppen and J. W. Vaupel (2010). "Losses of expected lifetime in the US and other developed countries: Methods and empirical analysis." *Demography* 47:1–50.

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

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

Thatcher, A. R., S. L. K. Cheung, S. Horiuchi and J.-M. Robine (2010). "The compression of deaths above the mode." *Demographic Research* 22:505–538.

Theil, H. (1967). Amsterdam: North-Holland Publishing Co.

van Raalte, A. A. (2011). *Lifespan variation: Methods, trends and the role of socioeconomic inequality*. Ph.D. thesis, Erasmus University.

van Raalte, A. A., A. E. Kunst, P. Deboosere, M. Leinsalu, O. Lundberg, P. Martikainen, B. H. Strand, B. Artnik, B. Wojtyniak and J. P. Mackenbach (2011). "More variation in lifespan in lower educated groups: evidence from 10 European countries." *International Journal of Epidemiology* doi: 10.1093/ije/dyr146.

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

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

WHO (2000). "The World Health Report 2000."

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

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

Zhang, Z. and J. W. Vaupel (2009). "The age separating early deaths from late deaths." *Demographic Research* 20:721–730.

Zureick, S. (2010). "Certainty in timing of death: A new analysis of shifting mortality and life span disparity." *Ph.D. thesis, University of California at Berkeley* .

8 Appendix A - 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.

8.1 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

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

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

$$\frac{dy}{d\mathbf{x}^\intercal} = \begin{pmatrix} \frac{\partial y}{dx_1} \cdots \frac{\partial y}{dx_m} \end{pmatrix}. \tag{19}$$

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:

$$\frac{d\mathbf{y}}{d\mathbf{x}^\intercal} = \begin{pmatrix} \frac{dy_i}{dx_j} \end{pmatrix}. \tag{20}$$

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

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

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

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

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

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; i.e.

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

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 (Magnus and Neudecker 1985)

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

We will frequently obtain expressions of the form (25) using a theorem originally due to Roth (Roth 1934), that if $\mathbf{Y} = \mathbf{ABC}$ then

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

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 is in Abadir and Magnus (2005), and demographic discussions appear in Caswell (2007; 2008; 2010).

8.2 Sensitivities of the indices of lifespan variation

8.2.1 Preliminaries

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

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

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

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

The life expectancy at birth is given by

$$\eta_1 = \boldsymbol{\eta}^{\mathsf{T}} \mathbf{e}_1 \tag{30}$$

and thus its derivative is

$$\begin{aligned} \frac{d\eta_1}{d\boldsymbol{\theta}^{\mathsf{T}}} &= (\mathbf{e}_1^{\mathsf{T}} \otimes \mathbf{e}^{\mathsf{T}}) \frac{d\mathrm{vec}\,\mathbf{N}}{d\boldsymbol{\theta}^{\mathsf{T}}} & (31) \\ &= (\mathbf{e}_1^{\mathsf{T}} \mathbf{N}^{\mathsf{T}} \otimes \mathbf{e}^{\mathsf{T}} \mathbf{N}) \frac{d\mathrm{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^{\mathsf{T}}} & (32) \end{aligned}$$

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),

$$\frac{d\mathbf{f}}{d\boldsymbol{\theta}^{\mathsf{T}}} = (\mathbf{e}_1^{\mathsf{T}} \mathbf{N}^{\mathsf{T}} \otimes \mathbf{I}) \frac{d\mathrm{vec}\,\mathbf{M}}{d\boldsymbol{\theta}^{\mathsf{T}}} + (\mathbf{e}_1^{\mathsf{T}} \mathbf{N}^{\mathsf{T}} \otimes \mathbf{B}) \frac{d\mathrm{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^{\mathsf{T}}} \tag{34}$$

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

8.2.2 Life disparity $\eta^\dagger$

The disparity can be written

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

As shown in Caswell (2010, 2011),

$$d\eta^{\mathsf{T}} = (d\mathbf{f}^{\mathsf{T}})\, \boldsymbol{\eta} + \mathbf{f}^{\mathsf{T}} (d\boldsymbol{\eta})\,, \tag{36}$$

and thus

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

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

8.2.3 Gini coefficient

In matrix form, the Gini coefficient is given by

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

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

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

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

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

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

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

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

8.2.4 Mean Logarithmic Deviation

The mean logarithmic deviation in matrix notation is

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

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

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

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

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

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

8.2.5 Theil's index

The expression for Theil's index in matrix notation is

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

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

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

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

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

Simplifying equation (50) and expressing the result in terms of a parameter vector $\boldsymbol{\theta}$ gives

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

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

8.2.6 The variance and standard deviation of longevity

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

$$\mathbf{v}^{\mathsf{T}} = \mathbf{e}^{\mathsf{T}} \mathbf{N} \left(2\mathbf{N} - \mathbf{I} \right) - \boldsymbol{\eta}^{\mathsf{T}} \circ \boldsymbol{\eta}^{\mathsf{T}} \tag{52}$$

Caswell (2006, 2009, 2010) shows that

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

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

$$\frac{d\text{vec}\,\mathbf{N}}{d\boldsymbol{\theta}^{\intercal}} = (\mathbf{N}^{\intercal} \otimes \mathbf{N}) \frac{d\text{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^{\intercal}}. \tag{54}$$

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 was derived in Caswell (2010),

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

8.2.7 The inter-quartile range

The inter-quartile 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

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

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

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

The special case of the inter-quartile 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

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

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

Next we differentiate equation (60)

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

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 hand side of of equation (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

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

The product of equations (64) and (65, following equation (63) gives

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

The sensitivity of the inter-quantile 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 $\hat{x}_1$.

9 Appendix B - Matlab code

9.1 Function needed in your working directory, titled `vec.m`

```
function v=vec(x)

    v=x(:);
```

9.2 Main code

```
% Code developed by Alyson van Raalte and Hal Caswell, November 2011
% NEED TO HAVE FILE 'vec.m' in working directory
% INPUT DATA are "q" a vector of death probabilities
% (we used ages 0 to 110+ from HMD),
% and "x", a vector of the average age at death within each age interval,
% (i.e. {a0, 1.5, 2.5, ..., 110.5, 110 + a(omega)} - total length = q+1)


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Preliminaries
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% defining survival probabilities
    p=1-q;

% # of transient states + 1, because in construction of U first row is zero
    s=length(p)+1;

% # transient states only
    s2=length(p);

% other things we will need later
    I = speye(s);   % identity matrix
    e = ones(s,1);  % column vector of ones for summations
    e1 = [1,linspace(0,0,s2)]';

% C matrix is for calculating cumulative sums
    for i=1:s
        C(:,i) = [zeros(i,1);ones(s-i,1)];
```

```
    end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Markov chain formulation of longevity
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% U matrix describes transient states in the Markov chain
    U=sparse(diag(p,-1));
% N matrix, where (i,j) entry is the mean time spent in each age class i,

% conditional upon starting in age class j
    N=inv(I-U);

% M matrix has probability of death at each age on diagonal
    M=sparse(diag(1-p));
    M(s,s)=1;

% The distribution of ages at death
    B=M*N; % the complete distribution of ages at death
    f = B*e1; % the distribution of ages at death from birth (or first age class)

% survivorship (alternatively ell=N*e1)
    ell = e - C*f;

% remaining life expectancy at each age
    mean_eta = sum(N)' - 0.5;

% life expectancy at birth (or first age class)
    eta = e'*N*e1 - 0.5;
% NB: in Markov chain formulations, the life expectancy at birth is always
% 0.5 years higher than that found by conventional life table methods,
% which is why we subtract 0.5 years


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Indices of lifespan variation
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% variance in lifespan
    V=(sum(N)*(2*N-I) - mean_eta'.*mean_eta')';

% standard deviation in lifespan
    S=sqrt(V);

% e  measure
```

```
    dagger_eta=mean_eta'*B;

% Theil's index
    T = f'*((x*eta^(-1)).*(log(x*eta^(-1))));

% Mean Log Deviation
    MLD = f'*(e*log(eta)-log(x));

% Gini coefficient
    G=1-(1/eta)*e'*(ell.*ell);

% IQR calculations
    F=[cumsum(f)]; % cumulative deaths
    age=[0:s-1]';
    xhat1=interp1(F(1:s-10),age(1:s-10),0.25);

% the 's-10' is in the code to handle zero deaths at oldest ages
    xhat2=interp1(F(1:s-10),age(1:s-10),0.75);
    IQR=xhat2-xhat1;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% SENSITIVITIES
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% derivatives of U with respect to mortality change
    for i=1:s-1
        dU_dmu=sparse(zeros(s,s));
        dU_dmu(i+1,i)=-p(i);
        dvecU_dmu(:,i)=vec(dU_dmu);
    end

%derivatives of M with respect to mortality change
    for i=1:s-1
        dM_dmu=sparse(zeros(s,s));
        dM_dmu(i,i)=p(i);
        dvecM_dmu(:,i)=vec(dM_dmu);
    end

%derivative of f with respect to mortality change  %should it be B or N
    df_dmu=kron(e1'*N',I)*dvecM_dmu + kron(e1'*N',B)*dvecU_dmu;

%sensitivity of expected longevity with respect to mortality change
    deta_dmu=kron(e1'*N',e'*N)*dvecU_dmu;

%sensitivity of variance in longevity with respect to mortality change
```

```
    dV_dmu  = (2*kron((N')^2,mean_eta')...
            + 2*kron(N',mean_eta'*N)...
            - (I + 2*sparse(diag(mean_eta)))*kron(N',mean_eta'))...
            * dvecU_dmu;

%sensitivity of standard deviation with respect to mortality change
    dS_dmu=0.5*diag(1./S)*dV_dmu;

%sensitivity of eta-dagger with respect to mortality change
    deta_dagger_dmu =(kron(B(:,1)'*N',mean_eta') ...
                    + kron(N(:,1)',mean_eta'*B))*dvecU_dmu ...
                    + kron(N(:,1)',mean_eta')*dvecM_dmu;

%sensitivity of Theil's with respect to mortality change
    dT_dmu  = (x'*eta^(-1)).*log(x'*eta^(-1))*df_dmu...
            - (T*eta^(-1)+f'*x*eta^(-2))*deta_dmu;

%sensitivity of MLD with respect to mortality change
    dMLD_dmu = (e'*log(eta)-log(x'))*df_dmu+eta^(-1)*deta_dmu;

%sensitivity of Gini with respect to mortality change
    dG_dmu = eta^(-2)*e'*(ell.*ell)*deta_dmu + 2/eta*e'*diag(ell)*C*df_dmu;

% sensitivity of IQR with respect to mortality change
    % sensitivity of cumulative deaths with respect to mortality change
        dF_dmu=[zeros(1,s2);cumsum(df_dmu)];

    % the death density at the quantiles by interpolation
        fxhat1=interp1(age,f,xhat1);
        fxhat2=interp1(age,f,xhat2);
        dFxhat1_dmuT=interp1([0:s]',dF_dmu(:,:),xhat1);
        dFxhat2_dmuT=interp1([0:s]',dF_dmu(:,:),xhat2);

    % finally taking sensitivities
        dxhat_dmu1 = -fxhat1^(-1)*dFxhat1_dmuT;
        dxhat_dmu2 = -fxhat2^(-1)*dFxhat2_dmuT;
        dIQR_dmu=dxhat_dmu2-dxhat_dmu1;


%%%%%%%%%%%%%%%%%%%%%%%%
%ELASTICITIES
%%%%%%%%%%%%%%%%%%%%%%%%

% elasticity of standard deviation
    ES = (1./S(1))*dS_dmu(1,:)*diag(1-p);
```

```
% elasticity of variance
    EV = (1./V(1))*dV_dmu(1,:)*diag(1-p);
% elasticity of edagger
    Eedag = (1./dagger_eta(1))'*deta_dagger_dmu*diag(1-p);
% elasticity of Gini
    EG = (1./G(1))*dG_dmu*diag(1-p);
% elasticity of Theil's
    ET = (1./T(1))*dT_dmu*diag(1-p);
% elasticity of MLD
    EMLD = (1./MLD(1))*dMLD_dmu*diag(1-p);
% elasticity of IQR
    EIQR = (1./IQR(1))*dIQR_dmu*diag(1-p);


%%%%%%%%%%%%%%%%%%%%%%%%
%SAVING THE RESULTS
%%%%%%%%%%%%%%%%%%%%%%%%

    res = [stV stedag stS stT stMLD stIQR stG EV Eedag ES ET EMLD EIQR EG];
    res=full(res); % MATLAB can't save sparse arrays
    save res.txt res -ascii;
```

10 Appendix C - Figures depicting the sensitivity and elasticity of indices calculated from age 10

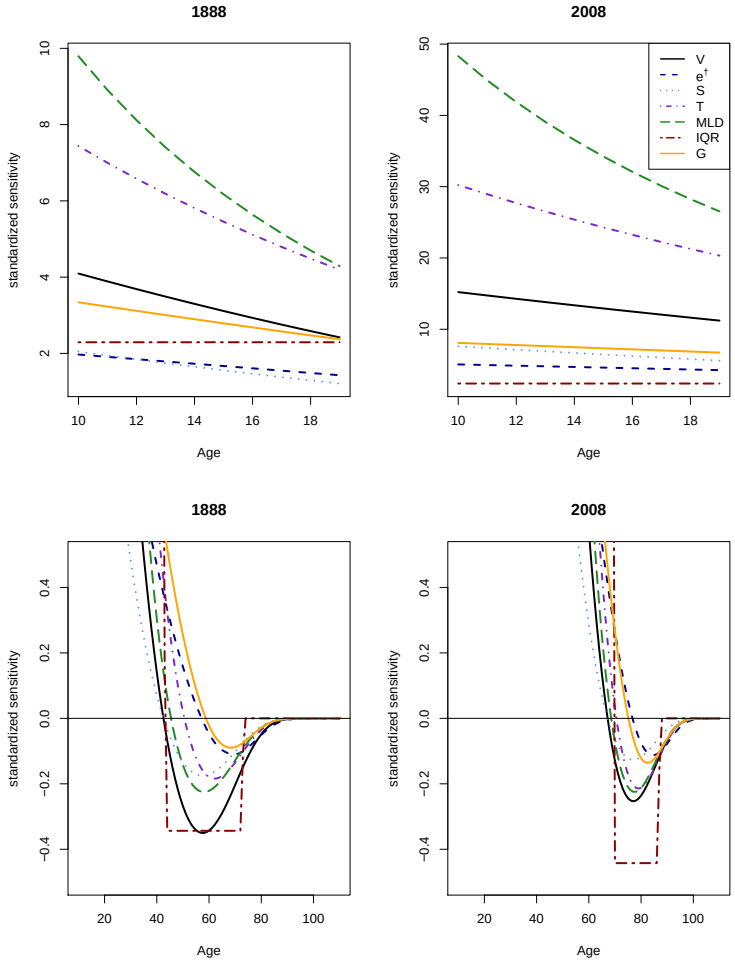

Figure 6: The sensitivity of each index conditional upon survival to age 10 with respect to mortality change at different ages. The sensitivities were standardized to the value of each index, i.e. $y_{10}\frac{dy}{d\theta}$, to make them comparable. Note the difference in scale between the top and bottom panels, plotted separately to more clearly delineate behaviour of the indices at early and later ages. French males, period lifetable data from the HMD.

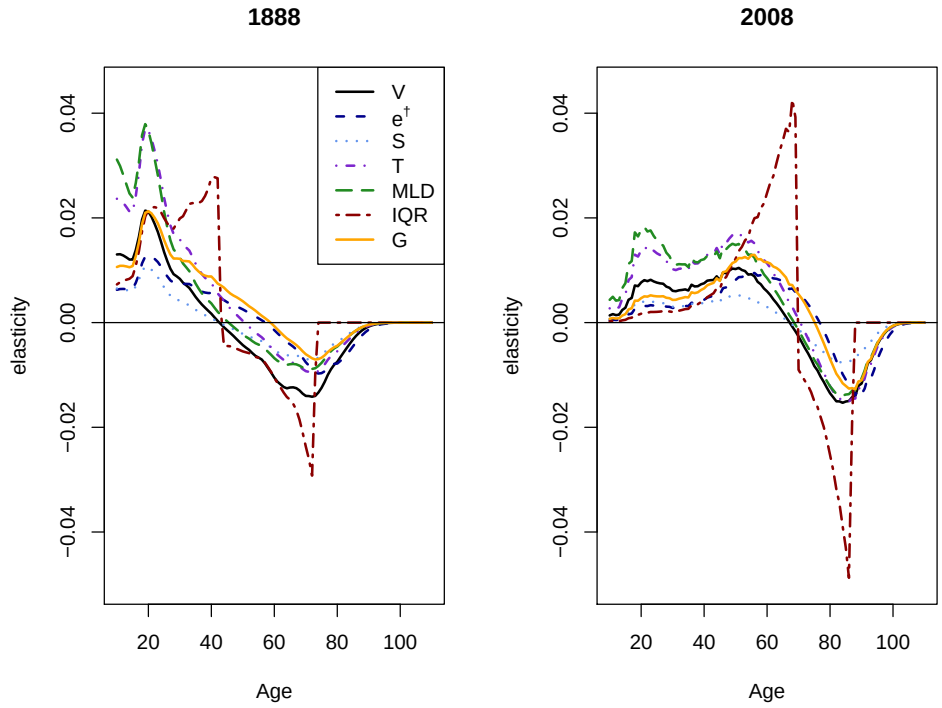

Figure 7: The proportional change in the index calculated conditional upon survival to age 10 from a one percent change in mortality at each age on the x axis. French males, period lifetable data from the HMD.