Perturbation analysis of longevity using matrix calculus

Population Association of America Annual Meeting
April 2010

Hal Caswell
Biology Department MS-34
Woods Hole Oceanographic Institution
Woods Hole MA 02543 USA

Max Planck Institute for Demographic Research
Rostock, Germany

March 2, 2010

Contents

1 Introduction

Studies of longevity are often concerned with *change* — change over time [41, 36, 18], differences among countries [14] , regions [18, 34], or ethnic groups, or changes in response to changing health [33] or economic conditions. Perturbation analysis is an approach to the study of the change in a dependent variable to changes in one or more variables on which it depends.

In this paper, I present a new approach to the perturbation analysis of longevity, based on the application of matrix calculus to Markov chain models. My focus is on the analytical methods, but as an example I will show results for two contrasting populations: Japanese females in 2006 (life expectancy at birth 85 years) and Indian females in 1961 (life expectancy 45 years).[1] Mortality rates in India are up to two orders of magnitude higher than in Japan, but aging proceeds more rapidly in Japan than in India (Figure 1). Life expectancy declines almost linearly with age in both countries (Figure 1). Infant mortality in India is high enough that life expectancy increases with age for the first few years of life.

Other applications of some of the results I present here will appear in a talk by Sara Zureik in Session 149, Saturday, April 17, 8:30–10:20 AM.

1.1 Measures of longevity

Longevity is the length of an individual's life. A population is characterized by the distribution of longevity. Here, I present perturbation theory for several indices:

1. Life expectancy at age.
2. Variance of longevity at age.
3. Standard deviation of longevity at age. This includes the popular standard deviation of longevity at age 10 (SD_{10}) as a special case.
4. Life lost: the average number of years of life lost due to mortality.
5. The entropy H of the distribution of age at death.
6. The dependency ratio R of the distribution of age at death.

Indices 2 – 5 are measures of the disparity or inequality implied by a distribution of age at death [16, 46, 52]. Item 6 is an example of another type of calculaton form the distribution that might be of interest.

I will present the sensitivity and elasticity of these indices with respect to changes in age-specific mortality rates.

1.2 Perturbation analysis

Perturbation analysis refers to a set of mathematical tools whose goal is to describe the *change* in one quantity resulting from a *change* in some other quantity. Because longevity depends on mortality, asking how longevity will respond to a change in mortality (e.g., [28, 29]) is an example of a perturbation analysis. Perturbation analysis is of interest precisely because change (over time

[1] Mortality data were obtained from the Human Lifetable Database.

or among populations) is of interest (see Section 3 for more discussion). The results of perturbation of a parameter θ on some dependent variable y can be described by the sensitivity $\frac{dy}{d\theta}$ or the elasticity $\frac{\theta}{y}\frac{dy}{d\theta}$ of y. The former gives the response to a small additive perturbation, the latter the proportional response to a small proportional perturbation [4].

2 Preliminaries

2.1 Markov chain formulation of the longevity problem

Perturbation analysis of longevity has been pursued mostly within the framework of age-classified life cycles [2, 28, 40, 48, 49]. This framework, however is limited in its ability to incorporate life cycle complexity. An alternative is provided by formulating the problem in terms of an absorbing Markov chain. This approach was pioneered in demography by [19, 20] and has been greatly extended in recent years [4, 5, 9, 23, 47]. A good source for the basic theory of absorbing Markov chains is [25].

Individuals move through a set of transient states in their life cycle and eventually die. Transient states may represent age classes, developmental stages, or states defined by health, employment, economic, or other kinds of status. Absorbing states may represent death, death at a particular stage, death from a specified cause, or other ways of leaving the life cycle (e.g., graduation in a model of educational states) [6, 9].

Number the stages in the life cycle so that the transient states are $1, \ldots, s$ and the absorbing states are $s+1, \ldots, s+a$. Then the transition matrix of the Markov chain is

$$\mathbf{P} = \left(\begin{array}{c|c} \mathbf{U} & 0 \\ \hline \mathbf{M} & \mathbf{I} \end{array} \right) \tag{1}$$

Here, $\mathbf{U}$ is the $s \times s$ matrix of transition probabilities among the transient states. The $a \times s$ matrix $\mathbf{M}$ gives the probabilities of absorbtion in each of the absorbing states. The columns of $\mathbf{P}$ sum to one.

I assume that the spectral radius of $\mathbf{U}$ is strictly less than one; a sufficient condition for this is that there is a non-zero probability of ultimate death for every stage.

Age-classified models are a special case, in which

$$\mathbf{U} = \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ p_1 & & & \\ & \ddots & & \\ & & p_{s-1} & p_s \end{pmatrix} \tag{2}$$

Where the age-specific survival probability is $p_i = \exp(-\mu_i)$, with μ_i a mortality rate. The (s, s) entry of $\mathbf{U}$ is an age-independent survival probability for the last age class; if it is set to zero, no one survives beyond age class s.

My results can be extended to time-varying models, but I will not do so here.

2.1.1 Longevity

Eventual absorbtion is certain. Longevity is analyzed in terms of how long it takes for absorbtion. Let ν_{ij} be the number of visits to transient state i by an individual in transient state j, prior to absorbtion. Its expectation is given by the fundamental matrix

$$\begin{aligned} \mathbf{N} &= \left(\; E(\nu_{ij}) \; \right) && (3) \\ &= (\mathbf{I} - \mathbf{U})^{-1} && (4) \end{aligned}$$

Let η_j be the number of visits to all transient states by an individual in state j, prior to absorbtion. The life expectancy of an individual in stage j is the expectation of η_j. The expectations are obtained as the column sums of $\mathbf{N}$

$$E(\boldsymbol{\eta})^{\mathsf{T}} = \mathbf{1}^{\mathsf{T}}\mathbf{N} \tag{5}$$

where $\mathbf{1}$ is a column vector of ones.[2]

The variance in longevity is [25]

$$V(\boldsymbol{\eta})^{\mathsf{T}} = \mathbf{1}^{\mathsf{T}}\mathbf{N}(2\mathbf{N} - \mathbf{I}) - E(\boldsymbol{\eta}^{\mathsf{T}}) \circ E(\boldsymbol{\eta}^{\mathsf{T}}) \tag{6}$$

The standard deviation of longevity is

$$SD(\boldsymbol{\eta}) = \sqrt{V(\boldsymbol{\eta})} \tag{7}$$

where the square root is taken element-wise. Both $V(\boldsymbol{\eta})$ and $SD(\boldsymbol{\eta})$ are zero when everyone dies at the same age; they have no clear upper limit.

Note that $V(\boldsymbol{\eta})$ and $SD(\boldsymbol{\eta})$ are vectors; the elements give the variance or standard deviation of longevity for individuals in each stage, making it easy to examine the disparity in remaining longevity as a function of age.

Suppose that the the absorbing stages are defined as the stage (or age) at death. Then

$$\mathbf{M} = \begin{pmatrix} 1-p_1 & & \\ & \ddots & \\ & & 1-p_s \end{pmatrix} \tag{8}$$

In this case, the distribution of stage at death is given by

$$\mathbf{B} = \mathbf{M}\mathbf{N} \tag{9}$$

where b_{ij} is the probability of eventual death in stage i for an individual currently in stage j. The distribution of stage at death for an individual in stage 1 is given by the vector

$$\mathbf{f} = \mathbf{B}\mathbf{e}_1 \tag{10}$$

where $\mathbf{e}_1$ is the first unit vector (i.e., column 1 of $\mathbf{I}$).

The expected life lost [49] is given by

$$\eta^{\dagger} = \mathbf{f}^{\mathsf{T}} E(\boldsymbol{\eta})^{\mathsf{T}} \tag{11}$$

In addition to being of interest in itself, $\eta^{\dagger}$ is a measure of disparity in lifespan; if everyone dies at the same age, $\eta^{\dagger} = 0$.

The entropy of the distribution of stage at death is

$$H = -\mathbf{f}^{\mathsf{T}} \log(\mathbf{f}) \tag{12}$$

where $\log(f)$is calculated element-wise. The entropy is maximized when the distribution of age at death is uniform; it is zero when everyone dies at the same age.

The dependency ratio of deaths is an example of a ratio of linear combinations of the elements of $\mathbf{f}$,

$$R = \frac{\mathbf{a}^{\mathsf{T}}\mathbf{f}}{\mathbf{b}^{\mathsf{T}}\mathbf{f}} \tag{13}$$

where $\mathbf{a}$, for example, might contain ones for ages 15–65 and zeros elsewhere, and $\mathbf{b}$ would be the one-complement of $\mathbf{a}$.

[2] I am avoiding the customary symbol e for life expectancy. For one thing, e serves to many other functions to be a good choice. For another, I want to consider not only the expectation but also the higher moments of longevity. For that, I need a symbol for longevity itself, not only its expectation. I cringe at writing $E(e)$.

2.2 Perturbation analysis

Our goal is to obtain expressions for the derivatives of $E(\boldsymbol{\eta})$, $V(\boldsymbol{\eta})$, $SD(\boldsymbol{\eta})$, $\eta^\dagger$, H, and R with respect to changes in age specific-mortality rates. This requires differentiating matrix- and vector-valued functions with respect to vector-valued arguments. To do so, we use matrix calculus.

2.3 Matrix calculus

Matrix calculus permits the consistent differentiation of scalar-, vector-, and matrix-valued functions of scalar, vector, or matrix arguments. This method has recently been introduced in ecology [5, 6, 7, 50, 8, 9], but has not seen much use in demography.[3] I give a brief summary here; more detail can be found in [6]. There exist several conventions for matrix calculus, differing in their arrangements of the matrix and vector entries. The best is that of Magnus and Neudecker [31, 32]; a useful introductory treatment is in [1].

If x and y are scalars, the derivative of y with respect to x is the familiar derivative dy/dx. If $\mathbf{y}$ is a $n \times 1$ vector and x a scalar, the derivative of $\mathbf{y}$ with respect to x is the $n \times 1$ vector

$$\frac{d\mathbf{y}}{dx} = \begin{pmatrix} \dfrac{dy_1}{dx} \\ \vdots \\ \dfrac{dy_n}{dx} \end{pmatrix}. \tag{14}$$

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

$$\frac{dy}{d\mathbf{x}^\mathsf{T}} = \begin{pmatrix} \dfrac{\partial y}{\partial x_1} & \cdots & \dfrac{\partial y}{\partial x_m} \end{pmatrix} \tag{15}$$

Note the orientation of $d\mathbf{y}/dx$ as a column vector and $dy/d\mathbf{x}^\mathsf{T}$ as a row vector.

If $\mathbf{y}$ is a $n \times 1$ vector and $\mathbf{x}$ a $m \times 1$ vector, the derivative of $\mathbf{y}$ with respect to $\mathbf{x}$ is the $n \times m$ Jacobian matrix

$$\frac{d\mathbf{y}}{d\mathbf{x}^\mathsf{T}} = \left(\frac{dy_i}{dx_j} \right). \tag{16}$$

Derivatives involving matrices are written by transforming the matrices into vectors using the vec operator (which stacks the columns of the matrix into a column vector), and then 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

$$\frac{d\text{vec}\,\mathbf{Y}}{d\text{vec}^\mathsf{T}\mathbf{X}}. \tag{17}$$

For notational convenience, I will write $\text{vec}^\mathsf{T}\mathbf{X}$ for $(\text{vec}\,\mathbf{X})^\mathsf{T}$.

These definitions (unlike some alternatives; see [31]) lead to the familiar chain rule. If $\mathbf{Y}$ is a function of $\mathbf{X}$ and $\mathbf{X}$ is a function of $\mathbf{Z}$, then

$$\frac{d\text{vec}\,\mathbf{Y}}{d\text{vec}^\mathsf{T}\mathbf{Z}} = \frac{d\text{vec}\,\mathbf{Y}}{d\text{vec}^\mathsf{T}\mathbf{X}} \frac{d\text{vec}\,\mathbf{X}}{d\text{vec}^\mathsf{T}\mathbf{Z}}. \tag{18}$$

Derivatives are constructed by forming the differentials of the expressions involving the matrices. The differential of a matrix (or vector) is the matrix (or vector) of differentials of the elements; i.e.,

$$d\mathbf{X} = \left(dx_{ij} \right). \tag{19}$$

[3] A remarkable early paper by Willekens [51] is an exception. It used a slightly different, but related, formulation of matrix calculus, but it seems to have been cited only 8 times since its publication, and only once in regard to its perturbation analytical content. See also [17].

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

$$d\mathbf{y} = \mathbf{Q} d\mathbf{x} \tag{20}$$

then

$$\frac{d\mathbf{y}}{d\mathbf{x}^{\mathsf{T}}} = \mathbf{Q}. \tag{21}$$

(the "first identification theorem" of [31]).

The combination of the chain rule and the identification theorem permits more complicated expressions involving differentials to be turned into derivatives with respect to an arbitrary vector, say $\mathbf{u}$. If

$$d\mathbf{y} = \mathbf{Q} d\mathbf{x} + \mathbf{R} d\mathbf{z} \tag{22}$$

then

$$\frac{d\mathbf{y}}{d\mathbf{u}^{\mathsf{T}}} = \mathbf{Q} \frac{d\mathbf{x}}{d\mathbf{u}^{\mathsf{T}}} + \mathbf{R} \frac{d\mathbf{z}}{d\mathbf{u}^{\mathsf{T}}} \tag{23}$$

for any $\mathbf{u}$.

We will make extensive use the Kronecker product, defined as

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}. \tag{24}$$

The vec operator and the Kronecker product are related [45]; if

$$\mathbf{Y} = \mathbf{ABC} \tag{25}$$

then

$$\operatorname{vec} \mathbf{Y} = (\mathbf{C}^{\mathsf{T}} \otimes \mathbf{A}) \operatorname{vec} \mathbf{B}. \tag{26}$$

2.4 Elasticity

It is sometimes useful to consider the effect of a proportional, rather than an absolute change in a variable. The elasticity, or proportional sensitivity, of y to x is

$$\frac{\epsilon y}{\epsilon x} = \frac{x}{y} \frac{dy}{dx} \tag{27}$$

It gives the proportional change in y caused by a small proportional change in x. In matrix calculus, it is easily calculated as

$$\frac{\epsilon \mathbf{y}}{\epsilon \mathbf{x}^{\mathsf{T}}} = \operatorname{diag}(\mathbf{y})^{-1} \frac{d\mathbf{y}}{d\mathbf{x}^{\mathsf{T}}} \operatorname{diag}(\mathbf{x}) \tag{28}$$

where $\operatorname{diag}(\mathbf{y})$ is a matrix with $\mathbf{y}$ on the diagonal and zeros elsewhere.

3 Applications and results

In the following sections, I will give the perturbation results and the results for the life tables of Japan and India. Full details of derivations will appear elsewhere; see also [10]. Results are presented in terms of an arbitrary vector $\boldsymbol{\theta}$ of parameters on which $\mathbf{U}$ and $\mathbf{M}$ depend. In the examples, $\boldsymbol{\theta}$ will be the vector $\boldsymbol{\mu}$ of age-specific mortality rates.

3.1 The fundamental matrix

The fundamental matrix $\mathbf{N}$ appears in many of these formulas. Its sensitivity was first obtained by [5] as

$$\frac{d\mathrm{vec}\,\mathbf{N}}{d\boldsymbol{\theta}^{\mathsf{T}}} = \left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{N}\right) \frac{d\mathrm{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^{\mathsf{T}}} \tag{29}$$

3.2 Life expectancy

The sensitivity of the vector of life expectancy as a function of age is obtained by differentiating (5),

$$dE(\boldsymbol{\eta})^{\mathsf{T}} = \mathbf{1}^{\mathsf{T}}(d\mathbf{N}) \tag{30}$$

Applying the vec operator and Roth's theorem (26) gives

$$\begin{aligned} dE(\boldsymbol{\eta}) &= \left(\mathbf{I} \otimes \mathbf{1}^{\mathsf{T}}\right) d\mathrm{vec}\,\mathbf{N} & (31)\\ &= \left(\mathbf{I} \otimes \mathbf{1}^{\mathsf{T}}\right)\left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{N}\right) d\mathrm{vec}\,\mathbf{U} & (32) \end{aligned}$$

Applying the chain rule and the first identification theorem gives the final result

$$\frac{dE(\boldsymbol{\eta})}{d\boldsymbol{\theta}^{\mathsf{T}}} = \left(\mathbf{I} \otimes \mathbf{1}^{\mathsf{T}}\right)\left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{N}\right) \frac{d\mathrm{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^{\mathsf{T}}} \tag{33}$$

The elasticity of $E(\boldsymbol{\eta})$ is obtained using (28);

$$\frac{\epsilon E(\boldsymbol{\eta})}{\epsilon \boldsymbol{\theta}^{\mathsf{T}}} = \mathrm{diag}\left(E(\boldsymbol{\eta})\right)^{-1}\left(\mathbf{I} \otimes \mathbf{1}^{\mathsf{T}}\right)\left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{N}\right) \left(\frac{d\mathrm{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^{\mathsf{T}}}\right) \mathrm{diag}\left(\boldsymbol{\theta}\right) \tag{34}$$

Life expectancy is more sensitive to changes in mortality in Japan than in India; the sensitivity decreases almost linearly with age in Japan, and slightly less linearly in India (Figure 3). On the other hand, life expectancy is more elastic to changes in mortality in India, and less so in Japan.

3.3 Variance and standard deviation of longevity

The variance in longevity is contained in the vector $V(\boldsymbol{\eta})$ in (6). Its sensitivity is

$$\frac{dV(\boldsymbol{\eta})}{d\boldsymbol{\theta}^{\mathsf{T}}} = \left[2\left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{1}^{\mathsf{T}}\right) + 2\left(\mathbf{I} \otimes \mathbf{1}^{\mathsf{T}}\mathbf{N}\right) - \left(\mathbf{I} \otimes \mathbf{1}^{\mathsf{T}}\right) - 2\left(\mathrm{diag}\left(E(\boldsymbol{\eta})\right) \otimes \mathbf{1}^{\mathsf{T}}\right)\right]\left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{N}\right) \frac{d\mathrm{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^{\mathsf{T}}} \tag{35}$$

The sensitivity of the standard deviation (7) is

$$\frac{dSD(\boldsymbol{\eta})}{d\boldsymbol{\theta}^{\mathsf{T}}} = \frac{1}{2}\mathrm{diag}\ \left(SD(\boldsymbol{\eta})\right)^{-1} \frac{dV(\boldsymbol{\eta})}{d\boldsymbol{\theta}^{\mathsf{T}}} \tag{36}$$

The elasticities are calculated from (28)

$$\begin{aligned} \frac{\epsilon V(\boldsymbol{\eta})}{\epsilon \boldsymbol{\theta}^{\mathsf{T}}} &= \mathrm{diag}\ \left(V(\boldsymbol{\eta})\right)^{-1} \frac{dV(\boldsymbol{\eta})}{d\boldsymbol{\theta}^{\mathsf{T}}} \mathrm{diag}\left(\boldsymbol{\theta}\right) & (37)\\ \frac{\epsilon SD(\boldsymbol{\eta})}{\epsilon \boldsymbol{\theta}^{\mathsf{T}}} &= \mathrm{diag}\ \left(SD(\boldsymbol{\eta})\right)^{-1} \frac{dSD(\boldsymbol{\eta})}{d\boldsymbol{\theta}^{\mathsf{T}}} \mathrm{diag}\left(\boldsymbol{\theta}\right) & (38) \end{aligned}$$

Figures 4 and 5 show the sensitivity and elasticity of variance and standard deviation of longevity, at birth and at age 10. Not surprisingly, the qualitative patterns are similar. Both measures of disparity are more sensitive to mortality changes in Japan than in India, and the sensitivities are highest at young ages. Both life tables have the property that sensitivities are negative over a range of ages ($\approx 20--80$ for India, $\approx 80--105$ for Japan). In this age range, reductions in mortality will increase disparity. Outside this age range, reductions in mortality will reduce disparity.

The sensitivity and elasticity patterns for $SD(eta_1)$ and $SD(\eta_1 0)$ are qualitatively similar, but conditioning on survival to age 10 does remove a large impact of changes in mortality below that age.

3.4 Distribution of age at death

The sensitivity of the distribution of age at death is given by the matrix $d\mathbf{f}/d\boldsymbol{\theta}^{\mathsf{T}}$, whose (i,j) element is the derivative $df_i/d\theta_j$. This sensitivity is given by

$$\frac{d\mathbf{f}}{d\boldsymbol{\theta}^{\mathsf{T}}} = \left(\mathbf{e}_1^{\mathsf{T}}\mathbf{N}^{\mathsf{T}} \otimes \mathbf{I}\right) \frac{d\mathrm{vec}\,\mathbf{M}}{d\boldsymbol{\theta}^{\mathsf{T}}} + \left(\mathbf{e}_1^{\mathsf{T}} \otimes \mathbf{M}\right)\left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{N}\right) \frac{d\mathrm{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^{\mathsf{T}}} \tag{39}$$

If, as is the case here, the absorbing states are defined in terms of stage at death, then $\mathbf{M} = \mathbf{I} - \mathrm{diag}\,(\mathbf{p})$, and

$$\frac{d\mathrm{vec}\,\mathbf{M}}{d\boldsymbol{\theta}^{\mathsf{T}}} = -\mathrm{diag}\,(\mathrm{vec}\,\mathbf{I})\left(\mathbf{1}^{\mathsf{T}} \otimes \mathbf{I}\right) \frac{d\mathbf{p}}{d\boldsymbol{\theta}^{\mathsf{T}}} \tag{40}$$

where $\mathbf{p}$ is the vector of survival probabilities. The resulting expression for the sensitivity of the distribution of stage at death is

$$\frac{d\mathbf{f}}{d\boldsymbol{\theta}^{\mathsf{T}}} = -\left(\mathbf{e}_1^{\mathsf{T}}\mathbf{N}^{\mathsf{T}} \otimes \mathbf{I}\right)\mathrm{diag}\,(\mathrm{vec}\,\mathbf{I})\left(\mathbf{1}^{\mathsf{T}} \otimes \mathbf{I}\right) \frac{d\mathbf{p}}{d\boldsymbol{\theta}^{\mathsf{T}}} + \left(\mathbf{e}_1^{\mathsf{T}} \otimes \mathbf{M}\right)\left(\mathbf{N}^{\mathsf{T}} \otimes \mathbf{N}\right) \frac{d\mathrm{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^{\mathsf{T}}} \tag{41}$$

3.5 Life lost due to mortality

The life lost due to the death of an individual at age class i is the life expectancy at age class i, η_i; the probability that death occurs at that age is f_i; hence the life lost due to mortality over the whole life cycle is $\eta^\dagger = E(\boldsymbol{\eta})^{\mathsf{T}}\mathbf{f}$. The values are $\eta^\dagger = 10.1$ years for Japan and $\eta^\dagger = 23.9$ years for India. The sensitivity of $\boldsymbol{\eta}^\dagger$

$$\frac{d\eta^\dagger}{d\boldsymbol{\theta}^{\mathsf{T}}} = \left[\left(\mathbf{f}^{\mathsf{T}} \otimes \mathbf{1}^{\mathsf{T}}\right) + \left(\mathbf{e}_1^{\mathsf{T}} \otimes E(\boldsymbol{\eta})^{\mathsf{T}}\mathbf{M}\right)\right] \frac{d\mathrm{vec}\,\mathbf{N}}{d\boldsymbol{\theta}^{\mathsf{T}}} + \left(\mathbf{e}_1^{\mathsf{T}}\mathbf{N}^{\mathsf{T}} \otimes E(\boldsymbol{\eta})^{\mathsf{T}}\right) \frac{d\mathrm{vec}\,\mathbf{M}}{d\boldsymbol{\theta}^{\mathsf{T}}} \tag{42}$$

Substituting the result for $d\mathbf{N}$ and simplifying gives

$$\frac{d\eta^\dagger}{d\boldsymbol{\theta}^{\mathsf{T}}} = \left[\left(\mathbf{f}^{\mathsf{T}}\mathbf{N}^{\mathsf{T}} \otimes \boldsymbol{E}(\boldsymbol{\eta})^{\mathsf{T}}\right) + \left(\mathbf{e}_1^{\mathsf{T}}\mathbf{N}^{\mathsf{T}} \otimes E(\boldsymbol{\eta})^{\mathsf{T}}\mathbf{B}\right)\right] \frac{d\mathrm{vec}\,\mathbf{U}}{d\boldsymbol{\theta}^{\mathsf{T}}} + \left(\mathbf{e}_1^{\mathsf{T}}\mathbf{N}^{\mathsf{T}} \otimes E(\boldsymbol{\eta})^{\mathsf{T}}\right) \frac{d\mathrm{vec}\,\mathbf{M}}{d\boldsymbol{\theta}^{\mathsf{T}}} \tag{43}$$

The sensitivity and elasticity of $\eta^\dagger$ show very similar patterns to those of $V(\boldsymbol{\eta})$ and $SD(\boldsymbol{\eta})$ (Figure 7), confirming that these indices are measuring similar aspects of disparity in longevity.

3.6 Dependency ratio of deaths

Consider any ratio of linear combinations of the elements of $\mathbf{f}$,

$$R = \frac{\mathbf{a}^{\mathsf{T}}\mathbf{f}}{\mathbf{b}^{\mathsf{T}}\mathbf{f}} \tag{44}$$

where **a** and **b** are weight vectors. In the case of the dependency ratio, **b** would contain the weights associated with the productive stages (e.g., ages 15–65) and **a** the weights associated with the dependent stages (e.g., younger than 15 and older than 65).

The values are $R = 14.8$ for Japan and $R = 1.35$ for India, emphasizing that deaths in the Japanese life table are overwhelmingly from dependent age classes, while most deaths in India occur in the productive age classes. It would be valuable to examine old-age dependency and young age dependency separately, but I have not done so here.

The sensitivity of such a ratio is

$$\frac{dR}{d\boldsymbol{\theta}^{\mathsf{T}}} = \left(\frac{\mathbf{b}^{\mathsf{T}}\mathbf{f}\mathbf{a}^{\mathsf{T}} - \mathbf{a}^{\mathsf{T}}\mathbf{f}\mathbf{b}^{\mathsf{T}}}{(\mathbf{b}^{t} r \mathbf{f})^2} \right) \frac{d\mathbf{f}}{d\boldsymbol{\theta}^{\mathsf{T}}} \tag{45}$$

R is by far more sensitive in Japan than in India (Figure 9), but when expressed in terms of proportional changes, the patterns are very similar.

4 Conclusion

1. Formulating the life cycle as an absorbing Markov chain makes it easy to calculate a variety of measures of longevity, including life expectancy, measures of disparity, and ratios.

2. Applying matrix calculus to these results yields formulae for the sensitivity and elasticity of these measures. The formulae are not particularly revealing, but are easily computable in software (Matlab, R) oriented toward matrix manipulations.

3. Results are presented for the sensitivity and elasticity of

   (a) the fundamental matrix **N**

   (b) the expectation of longevity $E(\boldsymbol{\eta})$

   (c) the variance and standard deviation of longevity, and the standard deviation of longevity above any specified age

   (d) the distribution of the age at death

   (e) the entropy of the distribution of age at death

   (f) the expected number of years of life lost due to mortality $\eta^{\dagger}$

   (g) the dependency ratio of deaths

4. When applied to the life tables for Japan 2006 (low mortality, low disparity) and India 1961 (high mortality and high disparity), several patterns are apparent. The qualitative patterns of sensitivity of measures of disparity are all very similar, suggesting that they are all responding in the similar ways to changes in mortality.

5 Acknowledgments

This research was supported by the National Science Foundation. I thank the students in IMPRSD 133, Perturbation Analysis of Longevity, for their patience and input, and the Max Planck Institute for Demographic Research for hospitality.

6 Literature cited

References

[1] Abadir, K.M. and J.R. Magnus. 2005. Matrix algebra. Econometric exercises 1. Cambridge University Press, Cambridge, United Kingdom.

[2] Canudas Romo, V. 2003. Decomposition methods in demography. Population Studies, Rozenberg Publishers, Amsterdam, Netherlands.

[3] Caswell, H. 1989. Matrix population models. First edition. Sinauer Associates, Sunderland, Massachusetts, USA.

[4] Caswell, H. 2001. Matrix population models: construction, analysis, and interpretation. Second edition. Sinauer Associates, Sunderland, Massachusetts, USA.

[5] Caswell, H. 2006. Applications of Markov chains in demography. pp. 319–334 in A.N. Langville and W.J. Stewart (editors), MAM2006: Markov Anniversary Meeting. Boson Books, Raleigh, North Carolina, USA.

[6] Caswell, H. 2007. Sensitivity analysis of transient population dynamics. Ecology Letters 10:1–15.

[7] Caswell, H. 2008. Perturbation analysis of nonlinear matrix population models. Demographic Research 18:59–116.

[8] Caswell, H. 2009. Sensitivity and elasticity of density-dependent population models. Journal of Difference Equations and Applications 15:349–369.

[9] Caswell, H. 2009. Stage, age, and individual stochasticity in demography. Oikos 118:1763–1782.

[10] Caswell, H. Perturbation analysis of longevity (unpublished manuscript, in prep.)

[11] Cheung, S.L.K., J.-M. Robine, E. J.-C. Tu, and G. Caselli. 2005. Three dimensions of the survival curve: horizontalization, verticalization, and longevity extension. Demography 42:243–258.

[12] Cochran, M.E. and S. Ellner. 1992. Simple methods for calculating age-based life history parameters for stage-structured populations. Ecological Monographs 62:345–364.

[13] Cruz, G.T., Y. Saito, and J.N. Natividad. 2007. Active life expectancy and functional health transition among Filipino older people. Canadian Studies in Population 34:29–47.

[14] Day, P., J. Pearce, and D. Dorling. 2008. Twelve worlds: a geo-demographic comparison of global inequalities in mortality. Journal of Epidemiology and Community Health 62:1002–1007.

[15] Dublin, L.I., A.J. Lotka, and M. Spiegelman. 1949. Length of life: a study of the life table. Revised edition. Ronald Press, New York, NY.

[16] 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.

[17] Ekamper, P. and N. Keilman. 1993. Sensitivity analysis in a multidimensional demographic projection model with a two-sex algorithm. Mathematical Population Studies 4:21–36.

[18] Ezzati, M., A.B. Friedman, S.C. Kulkarni, and C.J.L. Murray. 2008. The reversal of fortunes: trends in county mortality and cross-country mortality disparities in the United States. PLoS Medicine 5(4):e66. doi:10.1371/journal.pmed.0050066

[19] Feichtinger, G. 1971. Stochastische modelle demographischer prozesse. Lecture notes in operations research and mathematical systems. Springer-Verlag, Berlin.

[20] Feichtinger, G. 1973. Markovian models for some demographic processes. Statistiche Hefte 14:310–334.

[21] Henderson, H.V. and S.R. Searle. 1981. The vec-permutation matrix, the vec operator and Kronecker products: a review. Linear and Multilinear Algebra 9:271–288.

[22] Horiuchi, S., J.R. Wilmoth, and S.D. Pletcher. 2008. A decomposition method based on a model of continuous change. Demography 45:785–801.

[23] Horvitz, C.C. and S. Tuljapurkar. 2008. Stage dynamics, period survival, and mortality plateaus. American Naturalist 172:203–215.

[24] Hunter, C.M. and H. Caswell. 2005. The use of the vec-permutation matrix in spatial matrix population models. Ecological Modelling 188:15–21.

[25] Iosifescu, M. 1980. Finite Markov processes and their applications. Wiley, New York, USA.

[26] Jagger, C., B. Cox, S. Le Roy, and the EHEMU Team. 2006. Health expectancy calculation by the Sullivan method: a practical guide. Third edition. EHEMU Technical Report September 2006.

[27] Kemeny, J.G. and J.L. Snell.1976. Finite Markov Chains. Springer-Verlag, New York, New York, USA.

[28] Keyfitz, N. 1977. Applied mathematical demography. Wiley, New York, New York, USA.

[29] Keyfitz, N. and H. Caswell. 2005. Applied mathematical demography. Third edition. Springer-Verlag, New York, New York, USA.

[30] Land, K.C. and A. Rogers. 1982. Multidimensional Mathematical Demography. Academic Press, New York, New York, USA.

[31] Magnus, J.R. and H. Neudecker. 1985. Matrix differential calculus with applications to simple, Hadamard, and Kronecker products. Journal of Mathematical Psychology 29:474–492.

[32] Magnus, J.R. and H. Neudecker. 1988. Matrix differential calculus with applications in statistics and econometrics. John Wiley & Sons, New York, New York, USA.

[33] Murray, C.J.L. and A.D. Lopez. 1996. The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020. Harvard University School of Public Health, Cambridge, MA.

[34] Murray, C.J.L., S.C. Kulkarni, C. Michaud, N. tomijima, M.T. Bulzacchelli, t.J. Iandiorio, and M. Ezzati. 2006. Eight Americas: investigatning mortality disparities across races, counties, and race-counties in the United States. PLoS Medicine 3(9):e260. DOI: 10.1371/journal.pmed.0030260.

[35] Nel, D.G. 1980. On matrix differentiation in statistics. South African Statistical Journal 14:137–193.

[36] Oeppen, J. and J.W. Vaupel. 2002. Broken limits to life expectancy. Science 296:1029–1031.

[37] Pearl, R. 1926. Alcohol and longevity. Alfred A. Knopf, New York, NY.

[38] Pearl, R. 1928. The rate of living. Alfred A. Knopf, New York, NY.

[39] Pollard, J.H. 1982. The expectation of life and its relationship to mortality. Journal of the Institute of Actuaries 109:225–240.

[40] Pollard, J.H. 1988. On the decomposition of changes in expectation of life and differentials in life expectancy. Demography 25:265–276.

[41] Riley, J. Rising life expectancy: a global history. Cambridge University Press, Cambridge, United Kingdom.

[42] Robine, J.-M., I. Romieu, and E. Cambois. 1999. Health expectancy indicators. Bulletin of the World Health Organization 77:181–185.

[43] Robine, J.-M., C. Jagger, C.D. Mathers, E.M. Crimmins, and R.M. Suzman. 2003. Determining health expectancies. Wiley, Chichester, England.

[44] Rogers, A., R.G. Rogers, and L.G. Branch. 1989. A multistate analysis of active life expectancy. Public Health Reports 104:222–226.

[45] Roth, W.E. 1934. On direct product matrices. Bulletin of the American Mathematical Society 40:461–468.

[46] Shkolnikov, V.M., E.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.

[47] Tuljapurkar, S. and C.C. Horvitz. 2006. From stage to age in variable environments: life expectancy and survivorship. Ecology. 87:1497–1509

[48] Vaupel, J.W. 1986. How change in age-specific mortality affects life expectancy. Population Studies 40:147–157.

[49] 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.

[50] Verdy, A. and H. Caswell. 2008. Sensitivity analysis of reactive ecological dynamics. Bulletin of Mathematical Biology 70:1634–1659.

[51] Willekens, F.J. 1977. Sensitivity analysis in multiregional demographic models. Environment and Planning A9:653–674.

[52] Wilmoth, J.R. and S. Horiuchi. 1999. Rectangularization revisited: variability of age at death within human populations. Demography 36:475–495.

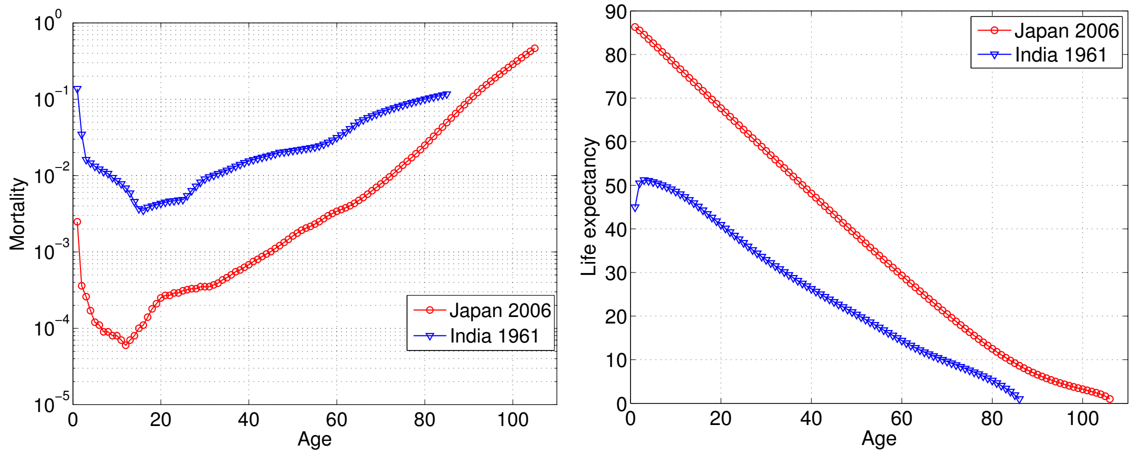
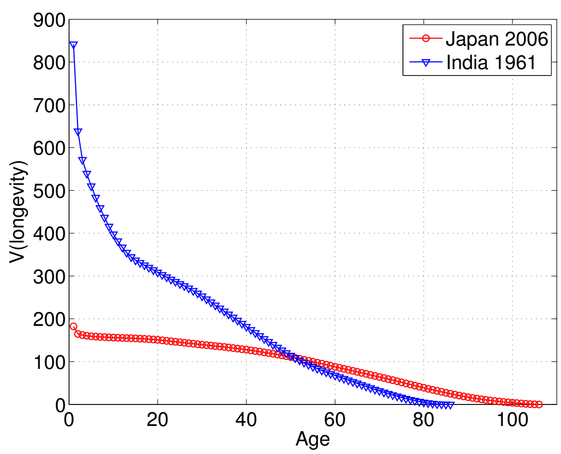


Figure 1: Mortality rate $\mu(x)$ and life expectancy $E(\boldsymbol{\eta})$ as a function of age, for Japan 2006 and India 1961 (females).

7 Figures



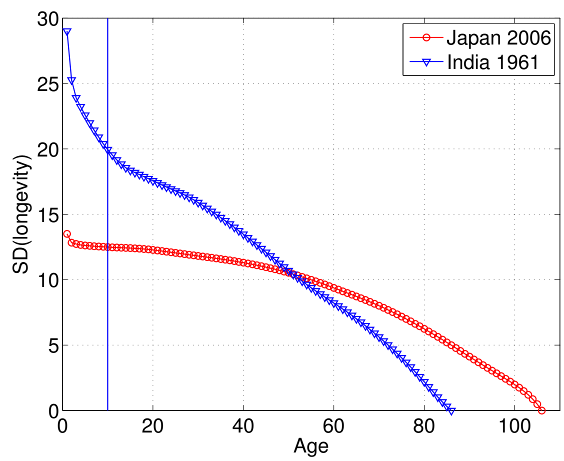


Figure 2: Variance and standard deviation of longevity as a function of age. in longevity at age.Vertical line indicates SD_{10}, sometimes used as a measure of lifespan disparity.

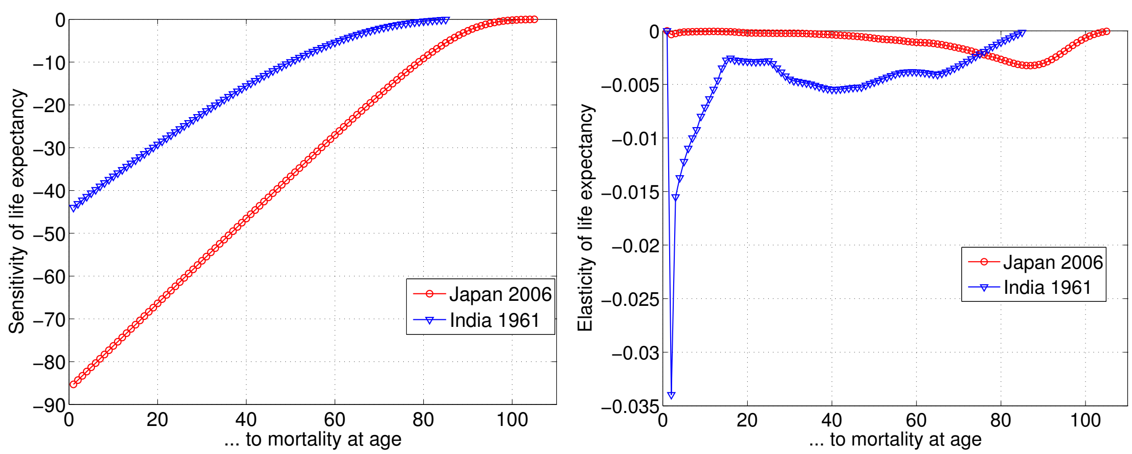


Figure 3: Sensitivity and elasticity of life expectancy at birth to changes in mortality rate at each age.

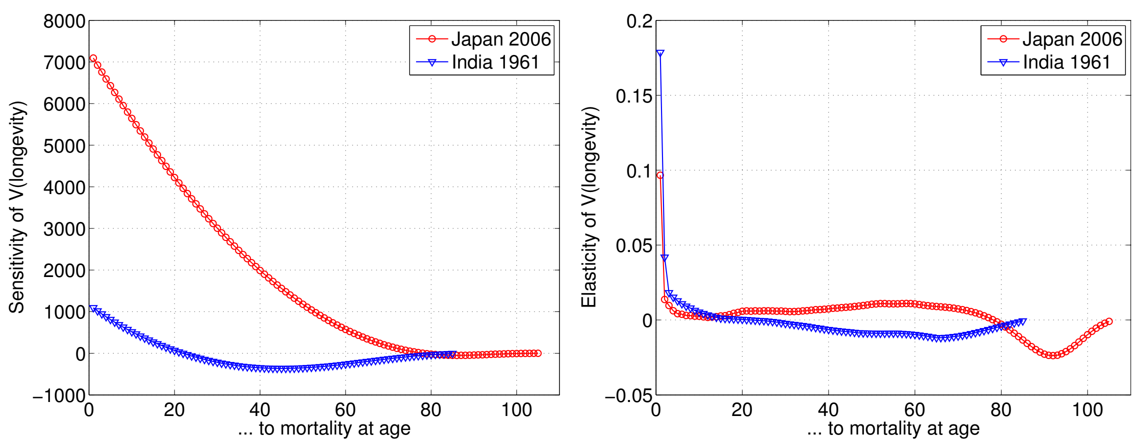


Figure 4: Sensitivity and elasticity of variance in longevity at birth to changes in mortality rate at age.

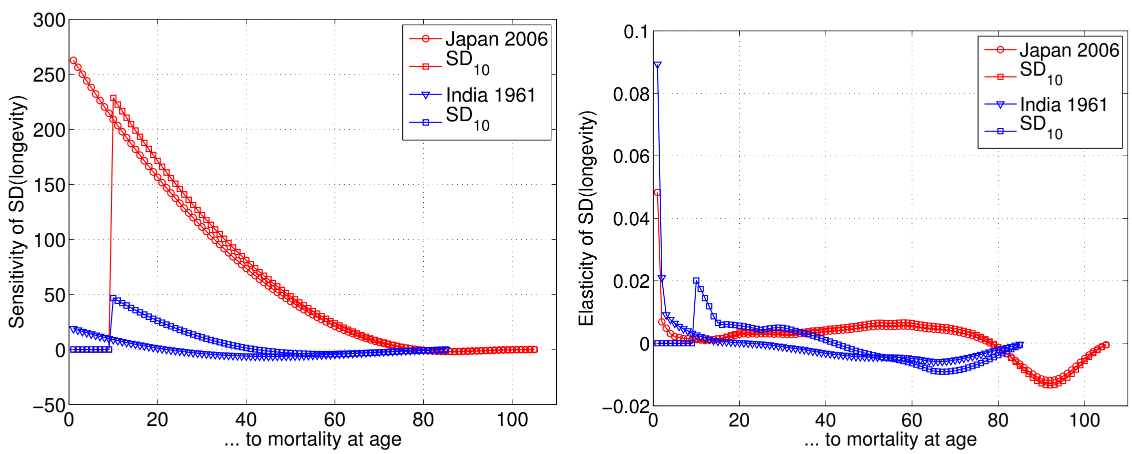


Figure 5: Sensitivity and elasticity of standard deviation at age 10, SD_{10}, to changes in mortality at age.

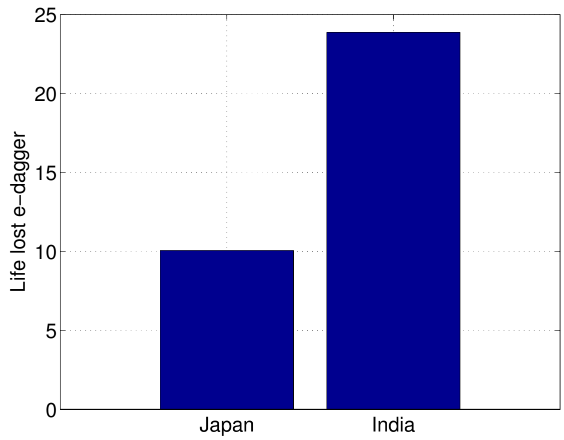


Figure 6: Life lost, $\eta^{\dagger}$.

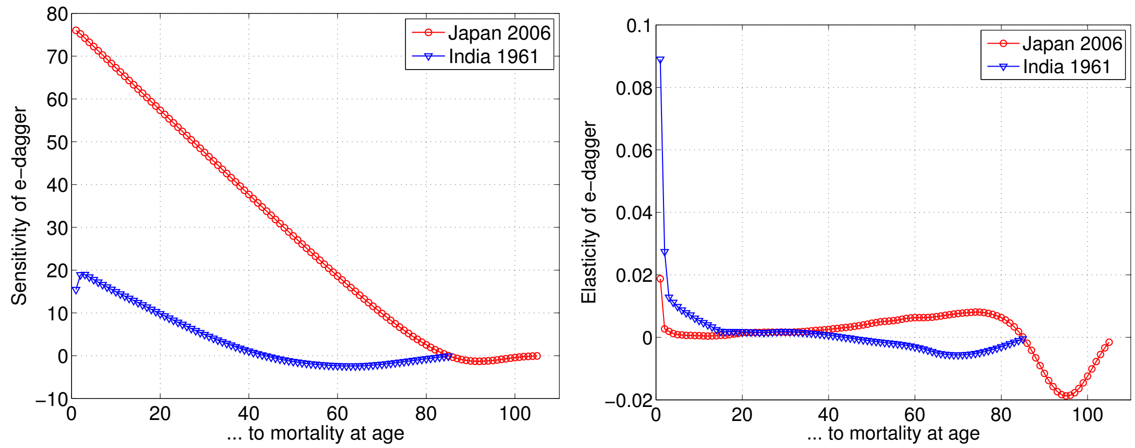


Figure 7: Sensitivity and elasticity of life lost at birth, $\eta^{\dagger}$, to mortality at specified age.

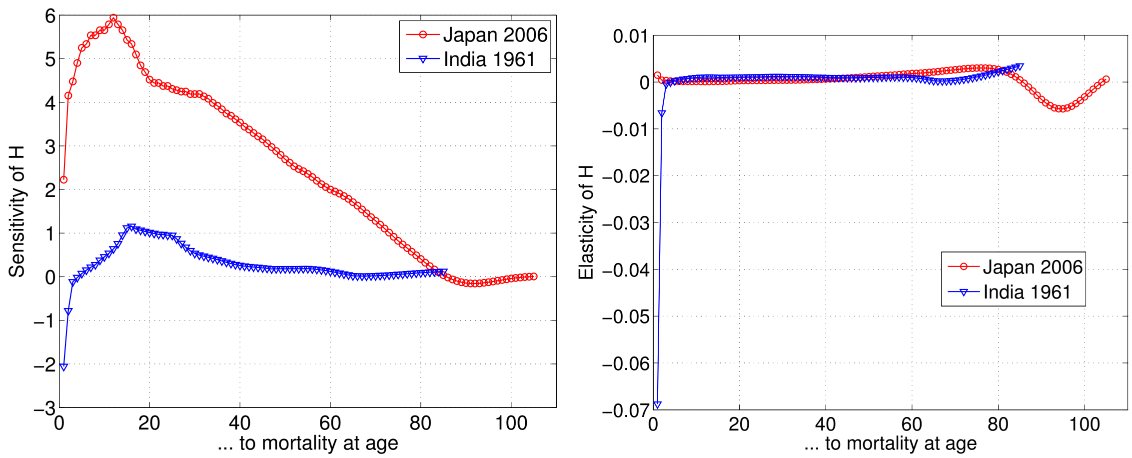


Figure 8: Sensitivity and elasticity of the entropy H of the distribution of longevity to mortality at age

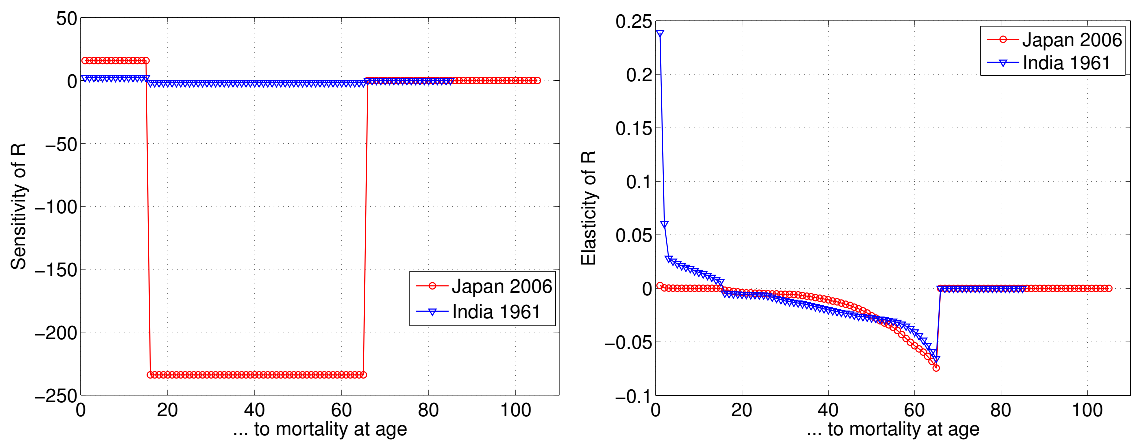


Figure 9: Sensitivity and elasicity of the dependency ratio of deaths to changes in mortality at age.