A user-friendly approach to stochastic mortality modelling

Helena Aro Teemu Pennanen
Department of Mathematics and Systems Analysis
Helsinki University of Technology
PL 1100, 02015 TKK
[haro,teemu]@math.hut.fi

Abstract

This paper proposes a general approach to stochastic mortality modelling, where the logit transforms of annual survival probabilities in different age groups are modelled by linear combinations of user-specified basis functions. The flexible construction allows for an easy incorporation of population-specific characteristics and user preferences into the model. Moreover, the structure enables the assignment of tangible demographic interpretations to the risk factors of the model. Survivor numbers are assumed to be binomially distributed, and, under very general assumptions, the resulting log-likelihood function in model calibration is shown to be strictly concave. This facilitates the use of convex optimization tools, and guarantees that the underlying risk factors are well-defined. We fit two versions of the model into Finnish adult (18-100 years) population and mortality data, and present simulations for the future development of Finnish life spans.

Keywords: Mortality risk, longevity risk, survival probabilities, stochastic modelling, convexity

1 Introduction

General longevity has improved significantly over the 20th century, with unexpectedly high increases in life spans. Mortality has not only been falling unpredictably in general, but there have also been considerable fluctuations in the rate of improvement over time. In addition, the changes in mortality rates across different age groups have also displayed different behavior. The pensions industry as well as national social security systems incur the costs of unpredictably improved longevity, as they need to pay out benefits for much longer than was anticipated. The effects of mortality risk on demographics and on fiscal sustainability have been studied e.g. in [2, 1].

As the effects of factors such as medical advances, environmental changes or lifestyle issues on mortality remain unpredictable, life and pensions insurance industry as well as national pensions funds have become increasingly aware of the need for longevity risk management. Consequently, several new financial instruments have been introduced for the management of mortality risk; see e.g. [5, 4, 17, 22]. The mathematical tools for pricing and hedging of such products are still rather under-developed, compared with more traditional financial instruments. The markets for these new securities would benefit from well-founded models for mortality risk management.

This paper proposes a general approach, where the logistic transforms of annual survival probabilities in different age groups are modelled by linear combinations of basis functions specified by the user. The flexible construction allows population-specific characteristics as well as user views and preferences to be incorporated into the model. The structure also enables the assignment of tangible demographic interpretations to the risk factors of the model. Survivor numbers are assumed to be binomially distributed, which, under very general assumptions about the basis functions, results in a strictly convex log likelihood function when calibrating the model. This guarantees the uniqueness of risk factors, and facilitates the use of convex optimization tools in the estimation of risk factor values by the maximum likelihood method.

Several models have been proposed for capturing the uncertainty in future development of mortality rates; see [10] for a recent review. The earliest and still widely popular discrete-time model with one stochastic factor was introduced by Lee and Carter [20] in 1992. It was followed by a number of modifications (see e.g. [7, 21, 6, 16, 15]), varying the original model and addressing its shortcomings. Models with multiple stochastic factors were subsequently proposed by Renshaw and Haberman [25] and Cairns et al. [9], with extensions incorporating cohort effects by Renshaw and Haberman [26] and Cairns et al. [11]. Currie et al. [12] have applied penalized splines in mortality modelling. In addition, although mortality data is generally published on discrete time intervals, rendering the discrete-time framework a natural choice for practical implementations, the development of mortality has also been considered in continuous time (see e.g. [23, 13, 14]).

In the modelling approach proposed in this paper, the logit transforms of survival probabilities are modelled by linear combinations of user-specified basis functions on the age groups. The weights of the basis functions are the stochastic risk factors capturing the uncertainty in the future survival probabilities. As the number of basis functions as well as their properties, such as piecewise linearity, continuity and smoothness can be chosen by the user, population specific characteristics as well as user preferences and other expert opinions can be taken into account.

An appropriate choice of basis functions ensures that the risk factors of the model have an easy interpretation, for instance as the logit transforms of the survival probabilities in certain cohorts, which facilitates the assessment of the model, and enables the study of the relationships between survival probabilities and economic factors. This is a central issue in e.g. the engineering of mortality-linked securities.

The chosen model is fitted into data by the maximum log-likelihood method, assuming deaths to be binomially distributed. This results in a strictly concave log likelihood function to be maximized. This feature not only means that the risk factors are well-defined, but also enables the use of convex optimization tools in model calibration.

As an example, we study Finnish adult (from 18 to 100 years) longevity. We choose piecewise linear basis functions, and consider two exemplary models with two and three basis functions, respectively. Using the resulting model, we present simulations for future development of Finnish survival probabilities and cohort sizes with some promising results.

The rest of this paper is organized as follows. Section 2 presents the modelling procedure. Section 3 illustrates the modeling procedure by fitting two models to Finnish adult population data. The resulting models are analyzed in terms of historical data as well as in simulations. Section 4 concludes.

2 Model specification

Let $E(x,t)$ be the number of individuals aged $[x, x+1)$ years at the beginning of year t in a given population. Our aim is to model the values of $E(x,t)$ over time $t = 0, 1, 2, \dots$ for a given set $X \subset \mathbb{N}$ of ages. We assume that the conditional distribution of $E(x+1, t+1)$ given $E(x,t)$ is binomial:

$$E(x+1, t+1) \sim \mathrm{Bin}(E(x,t), p(x,t)), \tag{1}$$

where $p(x,t)$ is the *survival probability*, the probability that an individual aged x and alive at the beginning of year t is still alive at the end of that year. Although previous literature predominantly applies the Poisson distribution to the numbers of deaths during a year, the binomial distribution is more realistic in a discrete time framework. We adopt a discrete time framework from the beginning since most available data sets are for yearly observations and it is the yearly values that are of interest in many applications. The distinction between binomial and Poisson distributions becomes important especially when $E(x,t)$ is small as is often the case in older age groups and in the small populations of countries such as Finland.

A stochastic mortality model is obtained by modelling the survival probabilities $p(x,t)$ as stochastic processes. The future values of $E(x+1, t+1)$ are then obtained by sampling from $\mathrm{Bin}(E(x,t), p(x,t))$. The uncertainty in the future values of $p(x,t)$ represents the systematic risk in future values of $E(x,t)$. Even if the 'true' survival probabilities were known, future population sizes would still be random. However, as the population grows, the fraction $E(x{+}1, t{+}1)/[E(x,t)p(x,t)]$ converges in distribution to constant 1. For large populations, the population dynamics is thus well described by $E(x+1, t+1) = E(x,t)p(x,t)$, and the main uncertainty comes from unpredictable variations in the future values of $p(x,t)$.

We propose to model the logistic probabilities by

$$\operatorname{logit} p(x,t) := \ln\Big(\frac{p(x,t)}{1-p(x,t)}\Big) = \sum_{i=1}^{n} v_i(t)\phi_i(x), \tag{2}$$

where ϕ_i are user-defined *basis functions* and v_i are stochastic *risk factors* that vary over time. In other words, in this framework $p(x,t) = p_{v(t)}(x)$, where $v(t) = (v_1(t),\ldots,v_n(t))$ and $p_v : \mathcal{X} \to (0,1)$ is the parametric function defined for each $v \in \mathbb{R}^n$ by

$$p_v(x) = \frac{\exp\left(\sum_{i=1}^{n} v_i\phi_i(x)\right)}{1+\exp(\sum_{i=1}^{n} v_i\phi_i(x))}. \tag{3}$$

Modelling the vector of risk factors $v = (v_1,\ldots,v_n)$ as a real-valued stochastic process implies $p(x,t) \in (0,1)$, guaranteeing that they are indeed probabilities.

With appropriate choices of the basis functions $\phi_i(\cdot)$ one can incorporate chosen properties of $p(\cdot,t)$ in the model. For example, one may wish to construct a model where the probabilities $p(x,t)$ behave continuously or smoothly across ages. This can be achieved simply by choosing continuous or smooth basis functions, respectively.

Another natural requirement is that the basis functions be sufficiently independent in the sense that they each contain features that cannot be represented by the other basis functions. The functions ϕ_i are *linearly independent* on a set $A \subset X$ of age groups if the only vector $v \in \mathbb{R}^n$ that satisfies

$$\sum_{i=1}^{n} v_i\phi_i(x) = 0 \quad \forall x \in A$$

is the zero vector $v = 0$. A violation of this condition would mean that the set of basis functions is redundant in the sense that we could remove at least one basis function without affecting the range of possible survival probabilities on A in the model.

The choice of the basis functions also determines the interpretation of the risk factors. If, for example, the basis functions are such that $\phi_k(x) = 1$ but $\phi_i(x) = 0$ for $i \neq k$ for a certain age x, then the risk factor $v_k(t)$ equals the logistic survival probability in x in the year t. Such concrete interpretations facilitate the modelling of future values of the risk factors. For example, one may be able to deduce dependencies between between v and certain economic factors such as investment returns. Such dependencies play a crucial role in asset and liability management of insurance companies as well as in pricing and hedging of mortality-linked securities.

Example 1 *When the survival probabilities are given by* (3) *the logistic* mortality rates $q(x,t) := 1 - p(x,t)$ *are given by*

$$\operatorname{logit} q(x,t) = -\operatorname{logit} p(x,t) = -\sum_{i=1}^{n} w_i(t)\phi_i(x).$$

In [9], Cairns et al. introduced the model

$$\operatorname{logit} q(x,t) = \kappa_1(t) + \kappa_2(t)(x - \bar{x}),$$

where κ_1 and κ_2 follow a two-dimensional random walk, and $\bar{x}$ is the mean over all ages X. This fits our framework with $n = 2$, $v_i = -\kappa_i$, $\phi_1(x) = 1$ and $\phi_2(x) = (x - \bar{x})$. The parameter v_1 can be interpreted as the general level of mortality, while v_2 describes how mortality rates change with age. In this case, the basis functions are not only linearly independent on X but also orthogonal*:*

$$\sum_{x \in X} \phi_1(x)\phi_2(x) = 0.$$

Example 2 *Following Lee and Carter [20], we could model the mortality rates by*

$$\operatorname{logit} q(x,t) = \beta_1(x) + \beta_2(x)\kappa_t,$$

where the functions β_i as well as the historical values κ_t are obtained from the singular value decomposition of the matrix of historical death rates. In our framework this corresponds to $n = 2$, $\varphi_i = \beta_i$, $v_2 = -\kappa_t$ and the first risk factor v_1 being equal to constant -1. We will, however, deviate from the calibration procedure of [20].

Once the basis functions ϕ_i have been chosen, the vector $v = (v_1, \ldots, v_n)$ of risk factors is modelled as a multivariate stochastic process in discrete time. The simplest (nontrivial) choice would be to model v as a random walk with a drift, but one could also use more sophisticated models developed in the broad range of literature on econometric modelling. The model specification could be based solely on the user's views about the future development of survival probabilities, on historical data, or on both. The historical values of the risk factors $v(t) = (v_1(t), \ldots, v_n(t))$ can be easily constructed by maximum likelihood estimation as follows.

Given the historical values of $E(x,t)$, the log-likelihood function for yearly values of $v(t)$ can be written using (1) and (3) as

$$\begin{aligned} l_t(v) &= \ln \prod_{x \in X} \binom{E(x,t)}{E(x{+}1,t{+}1)} p_v(x)^{E(x+1,t+1)}(1 - p_v(x))^{E(x,t) - E(x+1,t+1)} \\ &= \sum_{x \in X} \{E(x{+}1,t{+}1)[\ln p_v(x) - \ln(1 - p_v(x))] + E(x,t)\ln(1 - p_v(x))\} + c_t \\ &= \sum_{x \in X} \left[E(x{+}1,t{+}1) \sum_i v_i \phi_i(x) - E(x,t) \ln(1 + e^{\sum_i v_i \phi_i(x)}) \right] + c_t \quad (4) \end{aligned}$$

where

$$c_t = \sum_{x \in X} \ln \binom{E(x,t)}{E(x{+}1,t{+}1)}.$$

Maximizing $l_t(v)$ over $v \in \mathbb{R}^n$ gives an estimate of the factor vector $v(t)$ for year t. In general, the maximization requires techniques of numerical optimization, but the following result significantly facilitates the task.

Proposition 3 *The log-likelihood function $l_t : \mathbb{R}^n \to \mathbb{R}$ is concave. If the basis functions ϕ_i are linearly independent on the set of ages*

$$A(t) = \{x \,|\, E(x,t) > 0\},$$

then l_t is strictly concave.

Proof. The log-likelihood function in (4) can be written as

$$l_t(v) = f_t(\Phi v),$$

where the functions $\Phi : \mathbb{R}^n \to \mathbb{R}^{A(t)}$ and $f : \mathbb{R}^{A(t)} \to \mathbb{R}$ are defined by

$$\Phi v = \left[\sum_{i=1}^{n} v_i \phi_t(x)\right]_{x \in A(t)}$$

and

$$f_t(z) = \sum_{x \in A(t)} [E(x+1, t+1) z(x) - E(x,t)\varphi(z(x))] + c_t$$

where $\varphi(z) = \ln(1 + e^z)$. For convexity, it suffices to show that Φ is linear and f_t is convex. For strict convexity, it suffices to show that f is strictly convex and that Φ is injective.

The linearity of Φ is clear and the injectivity of Φ is equivalent to the linear independence condition. To finish the proof it suffices to show that φ is strictly convex on $\mathbb{R}$ since that implies that f_t is strictly convex on $\mathbb{R}^{A(t)}$. The second derivative of φ can be written as

$$\varphi''(z) = \frac{e^z}{(1 + e^z)^2},$$

which is strictly positive on $\mathbb{R}$. This completes the proof; see e.g. [27, Theorem 2.13]. ☐

Concavity implies that local maxima of l_v are true maximum likelihood estimators. Strict concavity implies that the estimator is unique; see e.g. [27, Theorem 2.6]. Besides guaranteeing well defined estimators, convexity facilitates the numerical maximization of l_t. There exists a wide literature on numerical techniques for convex optimization; see e.g. [3, 24].

We end this section by a brief summary of our modelling procedure.

1. Choose a set $\{\phi_i\}_{i=1}^{n}$ of basis functions that is rich enough to allow for a description of features of interest in the survival probability curve $p(x., t)$.

2. Construct historical values of $v(t)$ from data using maximum likelihood estimation.

3. Model the future development of $v(t)$ as a stochastic process, using its historical values and/or expert information.

4. The future survival probabilities are given by

$$p(x,t) = \mathrm{logit}^{-1}\big(\sum_i v_i(t)\phi_i(x))\big) = \big[1 + \exp\big(-\sum_i v_i(t)\phi_i(x)\big)\big]^{-1}$$

5. The future population sizes $E(x+1, t+1)$ are obtained by sampling from $\mathrm{Bin}(E(x,t), p(x,t))$ or simply by $E(x+1, t+1) = E(x,t)p(x,t)$, if we are only interested in the systematic risk in future values of $E(x,t)$.

3 Modelling Finnish adult mortality

As an example, we construct two stochastic mortality models using the framework presented in the previous section, and fit the models into Finnish population and mortality data. The first, simpler model comprises two linear basis functions and two stochastic risk factors, while the second one has three piecewise linear basis functions and three risk factors.

We employ population and mortality data for Finnish adults, obtained from the Human Mortality Database[1]. We use data covering age groups from 18 to 100 and years 1900 to 2007 in the fitting of our models. The population data is recorded as the annual age-specific population sizes $E(x,t)$ at the beginning of year t, separately for males and females.

The yearly values of the risk factors v were obtained by maximizing the log-likelihood functions (4) using Matlab Optimization Toolbox.

3.1 Two-parameter model

We first consider a two-factor model with two linear basis functions, which are of the form

$$\phi_1(x) = 1 - \frac{x-18}{82} \quad \text{and} \quad \phi_2(x) = \frac{x-18}{82}.$$

Consequently, their linear combination $\sum_{i=1}^{2} v_i\phi_i(x)$ is also linear, as illustrated in Figure 1. Note that the same linear form for the curve of $\mathrm{logit}\, p$ could have been obtained using any two linearly independent linear basis functions. However, this particular choice ensures that the risk factors have a certain natural interpretation. Namely, for every t,

$$\mathrm{logit}\ \ p(18,t) = v_1(t)\phi_1(18) + v_2(t)\phi_2(18) = v_1(t),$$

---

[1] University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany); www.mortality.org.

and, similarly, $\operatorname{logit} p(100,t) = v_2(t)$. Hence, the risk factors are the logit survival probabilities of ages 18 and 100.

This interpretation will be useful when assessing estimation results for historical values of v, and the general validity of the model in this particular population. In addition, expert views on the future development of mortality in these age groups may be incorporated into the model when choosing the appropriate stochastic process that governs the evolution of the risk factors. Furthermore, in the engineering of mortality-linked securities this feature facilitates, for instance, the assessment of various other instruments for hedging purposes.

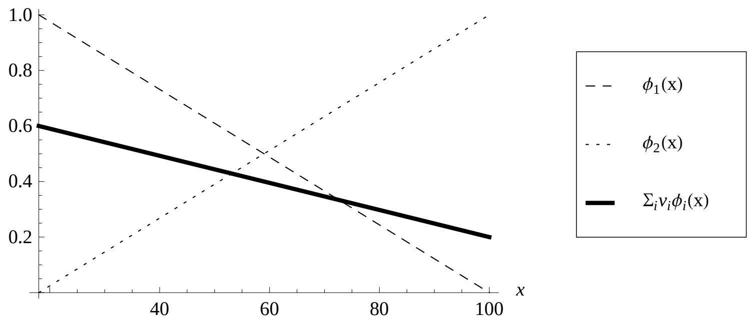

Figure 1: Linear basis functions and an exemplary linear combination $\sum_{i=1}^{2} v_i \phi_i(x)$ with $v_1 = 0.6$ and $v_2 = 0.2$.

Historical values of v_1 and v_2 for males and females are presented in Figures 2 and 3, respectively. The effect of Finnish war years (1918 and 1939-1944) can be clearly observed in the estimated risk factors for males. The dramatic drop in the wartime survival probabilities of younger age groups tilts the fitted linear logit survival probability curve resulting in the peaks in v_2. This two-dimensional model thus suggests that the survival probabilities of the oldest age groups would have temporarily increased during the war years. As this is doubtful, we conclude that the two-dimensional model is be too simple to capture the population dynamics in Finland during the entire 20th century. The female parameter estimates, in contrast, are less affected by the war years. The values of v_1 grows fairly steadily, while v_2 shows some fluctuations before settling into a growing trend around the 60s.

Figures 4 and 5 plot the survivals ratios $E(x+1,t+1)/E(x,t)$ in comparison with the estimated survival probabilities $p(x,t)$ for males and females. The values of $E(x+1,t+1)/E(x,t)$ reflect both systematic and nonsystematic risks, while $p(x,t)$ captures only the estimated systematic risk. Contrary to the behavior of $E(x+1,t+1)/E(x,t)$, the probabilities $p(x,t)$ for older males in Figure 4 display sharp ridges during the war year. This is caused by the spurious peaks in v_2 in Figure 2.

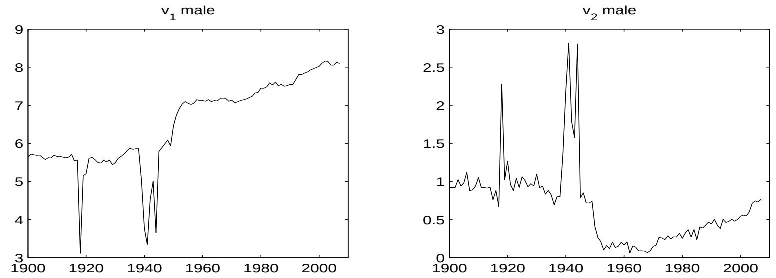

Figure 2: Estimated factor values for two-factor model, males.

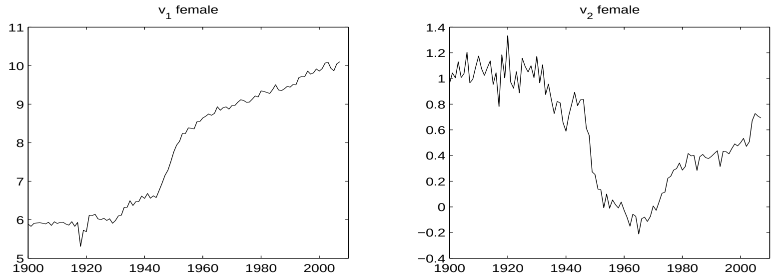

Figure 3: Estimated factor values for two-factor model, females.

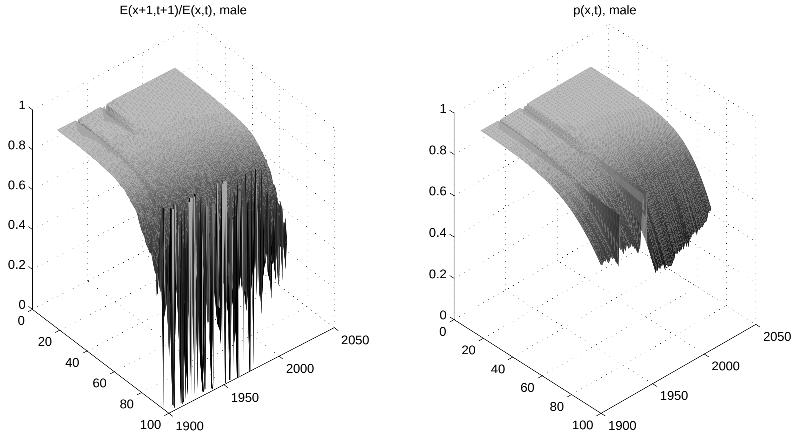

Figure 4: $\frac{E(x+1,t+1)}{E(x,t)}$ vs. estimated values of $p(x,t)$ for two-factor model, males.

3.2 Three-parameter model

To correct the somewhat unrealistic results obtained with the two dimensional model above, we will next implement a three dimensional model with three piece-

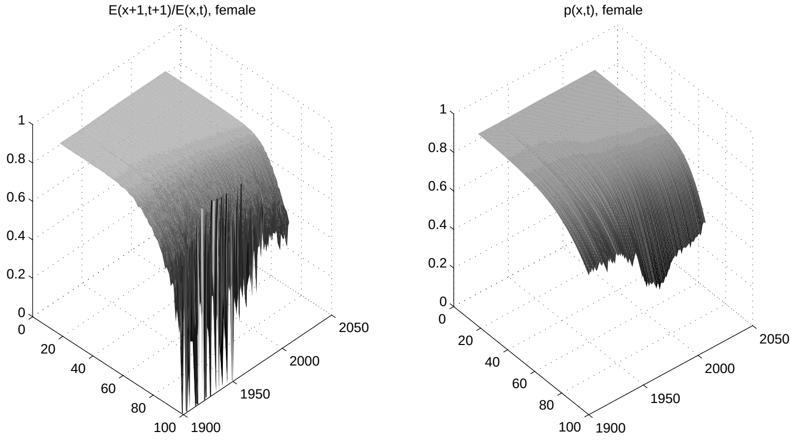

Figure 5: $\frac{E(x+1,t+1)}{E(x,t)}$ vs. estimated values of $p(x,t)$ for two-factor model, females.

wise linear basis functions given by

$$\phi_1(x) = \begin{cases} 1 - \frac{x-18}{32} & \text{for } x \leq 50 \\ 0 & \text{for } x \geq 50, \end{cases}$$

$$\phi_2(x) = \begin{cases} \frac{1}{32}(x-18) & \text{for } x \leq 50 \\ 2 - \frac{x}{50} & \text{for } x \geq 50, \end{cases}$$

$$\phi_3(x) = \begin{cases} 0 & \text{for } x \leq 50 \\ \frac{x}{50} - 1 & \text{for } x \geq 50. \end{cases}$$

The linear combination $\sum_{i=1}^{3} v_i \phi_i(x)$ is now piecewise linear and continuous; see Figure 6. Like in the two-dimensional model, the values of the factors are points on logit survival probability: $\operatorname{logit} p(18,t) = v_1(t)$, $\operatorname{logit} p(50,t) = v_2(t)$ and $\operatorname{logit} p(100,t) = v_3(t)$.

Risk factor estimates for males and females are plotted in Figures 7 and 8. For males, the effect of war years shows as sharp declines in v_1 and v_2, i.e. the logit survival probabilities for 18- and 50-year-olds, respectively. As opposed to the two-dimensional model in Section 3.1, the survival probabilities for the 100-year-olds behave more realistically and the sharp peaks during the war years are now gone. This is further illustrated in the surface plots of the estimated $p(x,t)$ in Figures 9 and 10. They now correspond better to the survival probabilities $\frac{E(x+1,t+1)}{E(x,t)}$ than in the two dimensional case. The unrealistic ridges in the survival probabilities of older males in the two-parameter model (Figure 4) do not appear here.

Another notable difference compared to the two-dimensional model is that now all the risk factors achieve their highest values during the most recent years. This is quite natural since the risk factors represent logistic survival probabilities which are generally viewed as upward drifting processes. In the two-dimensional model, the strong improvements in the larger young age groups tilt the linear logistic probability curve in the estimation at the expense of a worse fit of the smaller old age groups.

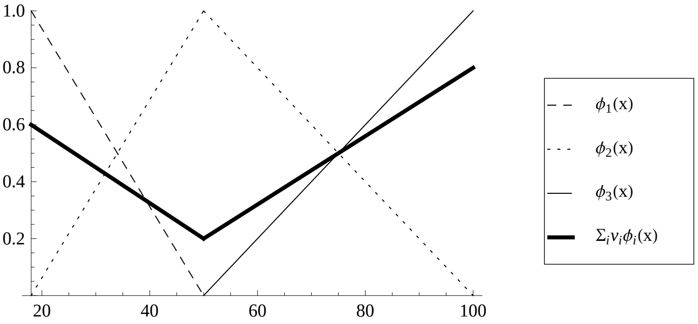

Figure 6: Three piecewise linear basis functions and their linear combination $\sum_{i=1}^{2} v_i\phi_i(x)$ with $v_1 = 0.6$, $v_2 = 0.2$ and $v_3 = 0.8$.

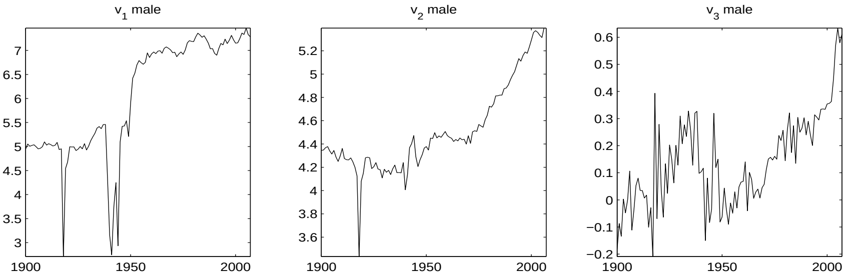

Figure 7: Estimated factor values of v for three-factor model, males.

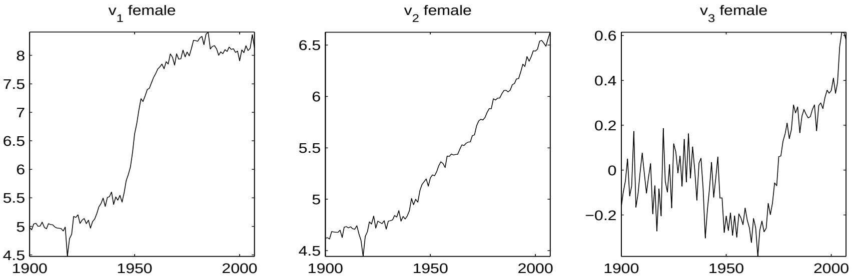

Figure 8: Estimated factor values of v for three-factor model, females.

As a more quantitative method of comparison between the two- and three-factor models, we calculated the Bayesian Information Criterion (BIC) for both;

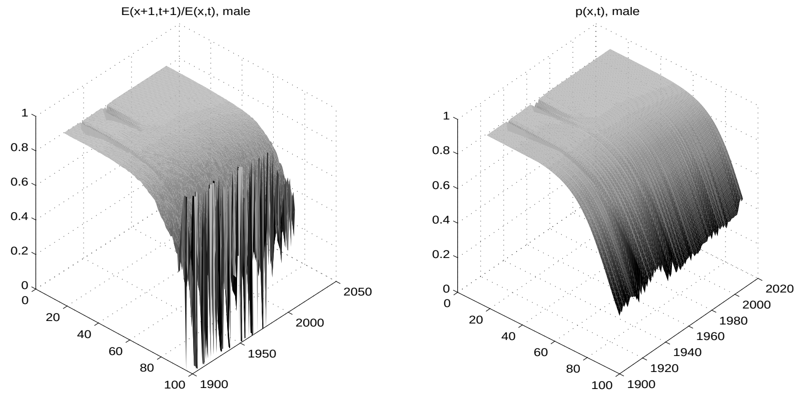

Figure 9: Estimated values of $\frac{E(x+1,t+1)}{E(x,t)}$ vs. $p(x,t)$ for three-factor model, males.

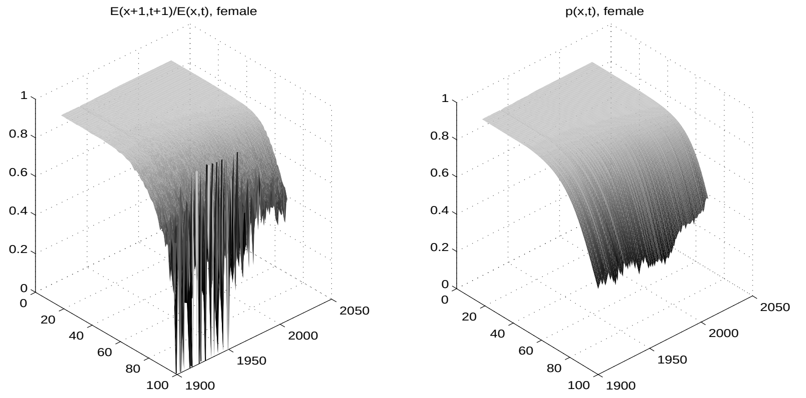

Figure 10: Estimated values of $\frac{E(x+1,t+1)}{E(x,t)}$ vs. $p(x,t)$ for three-factor model, females.

see [28]. Although some of the independence assumptions may be violated by the data, BIC gives a simple means to quantify the model fit; see also [8]. The criterion is defined as $BIC = l - \frac{1}{2}k\log(N)$, where l is the maximum log-likelihood, k the number of parameters, and N the number of observations. In our case, the yearly log-likelihood function is given by (4), $N = 100 - 18 = 82$ and k is either 2 or 3. The annual BIC values for the two models are compared in Figures 11 and 12. The BIC prefers the three-factor model to the two-factor one for both sexes and all years in the data.

3.3 Modelling the risk factors

To obtain a stochastic mortality model, we model the vector v of risk factors as a multivariate stochastic process. In order to capture the dependencies between male

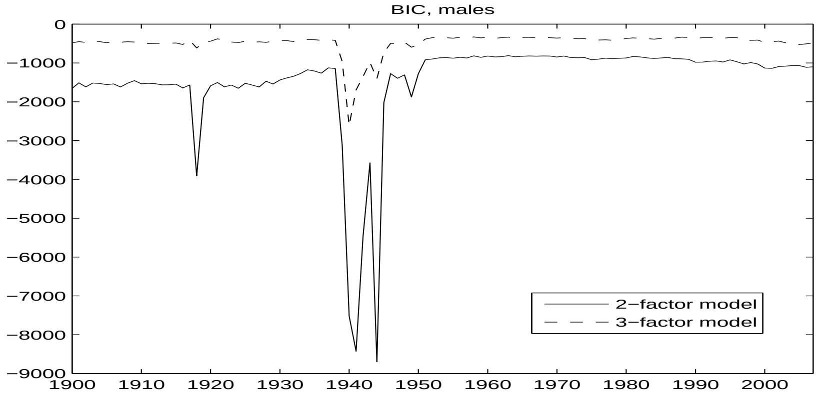

Figure 11: Annual BIC values, males.

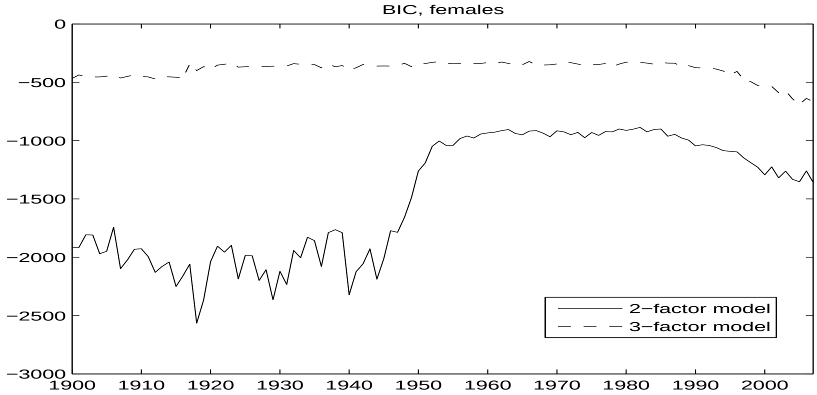

Figure 12: Annual BIC values, females.

and female mortality, we model the joint behavior of all the risk factors as a single multivariate process.

In this study, the risk factors are modelled as a simple multivariate Brownian motion (random walk) with drift. The combined vector of female and male risk factors $v(t)$ satisfies

$$\Delta v(t) = \mu + CZ(t), \tag{5}$$

where $Z(t)$ are n-dimensional independent standard Gaussian random variables, and vector $\mu \in \mathbb{R}^n$ and matrix $C \in \mathbb{R}^{n\times n}$ are parameters of the model. Here n is the total number of the risk factors for females and males. Vector μ gives the *drift* and matrix C the *volatility* of the risk factors, i.e. $E(\Delta v(t)) = \mu$ and $\mathrm{Var}(\Delta v(t)) = CC^T$. The volatility matrix C can be chosen e.g. as the Cholesky factor of the covariance matrix $\mathrm{Var}(\Delta v(t))$.

We choose to model Finnish mortality in relatively stable conditions, and therefore use historical values of the risk factors only from the period 1960-2007 in calibration of the processes. In the two-parameter model, the estimated risk factor values for both males and females show an upward trend from year 1960 onwards. In the case of the two-factor model we obtain, for the combined vector of female (f) and male (m) risk factors $v = (v_1^f, v_2^f, v_1^m, v_2^m)$

$$\mu = \begin{pmatrix} 0.0311 \\ 0.0153 \\ 0.0207 \\ 0.0128 \end{pmatrix}, \quad \sigma = \begin{pmatrix} 0.0799 \\ 0.0637 \\ 0.0535 \\ 0.0626 \end{pmatrix}$$

and

$$R = \begin{pmatrix} 1.0000 & -0.6890 & 0.4051 & -0.3565 \\ -0.6890 & 1.0000 & -0.3279 & 0.6655 \\ 0.4051 & -0.3279 & 1.0000 & -0.6605 \\ -0.3565 & 0.6655 & -0.6605 & 1.0000 \end{pmatrix},$$

where σ and R give the standard deviations and correlations of Δv so that

$$\mathrm{Var}(\Delta v) = \mathrm{diag}(\sigma) R \, \mathrm{diag}(\sigma).$$

For the three-parameter model with the vector $v = (v_1^f, v_2^f, v_3^f, v_1^m, v_2^m, v_3^m)$ of combined male and female risk factors, we obtain

$$\mu = \begin{pmatrix} 0.0097 \\ 0.0252 \\ 0.0171 \\ 0.0075 \\ 0.0196 \\ 0.0116 \end{pmatrix}, \quad \sigma = \begin{pmatrix} 0.1149 \\ 0.0352 \\ 0.0662 \\ 0.0765 \\ 0.0372 \\ 0.0728 \end{pmatrix}$$

and

$$R = \begin{pmatrix} 1.0000 & 0.1246 & 0.1277 & 0.1341 & 0.1690 & -0.0573 \\ 0.1246 & 1.0000 & -0.3064 & -0.2564 & 0.4259 & -0.1741 \\ 0.1277 & -0.3064 & 1.0000 & 0.1067 & 0.0528 & 0.6282 \\ 0.1341 & -0.2564 & 0.1067 & 1.0000 & -0.4031 & 0.1625 \\ 0.1690 & 0.4259 & 0.0528 & -0.4031 & 1.0000 & -0.3214 \\ -0.0573 & -0.1741 & 0.6282 & 0.1625 & -0.3214 & 1.0000 \end{pmatrix}.$$

3.4 Simulations

We applied the the two models presented in Section 3.2 to simulate male survival probabilities and cohort sizes 30 years into the future. The process $v(t)$ was modelled as a random walk with a drift as described in the previous section. A sample of 10000 scenarios for v was generated using Latin hypercube sampling, and the probabilities $p(x,t)$ were then calculated from each simulated path of $v(t)$. The number of survivors $E(x+1,t+1)$ in each cohort was approximated by its expected value $E(x,t)p(x,t)$.

Cohorts aged 30 and 65 in the final observation year 2007 were chosen as reference cohorts. Figures 13 and 14 plot the development of the medians and 90% confidence intervals for $p(x_0+t,t)$ and cohort sizes $E(x_0+t,t)$. Figures 15 and 16 give the corresponding results for the three-factor model. In the three factor model, the younger reference cohort displays a notable kink in the survival probability curves, which results from the cohort shifting from one part of the piecewise linear logit probability curve to another during the simulation period. For the younger cohort, the cohort size estimates for the two-factor model in the final simulation year are slightly below that of the three-factor model. For the older reference cohort the difference, albeit hardly notable, is the other way round.

Figure 17 plots the density function of the survival probability of the younger reference cohort for the three-factor model at the end of the simulation period i.e. 60-year-old males in year 2037.

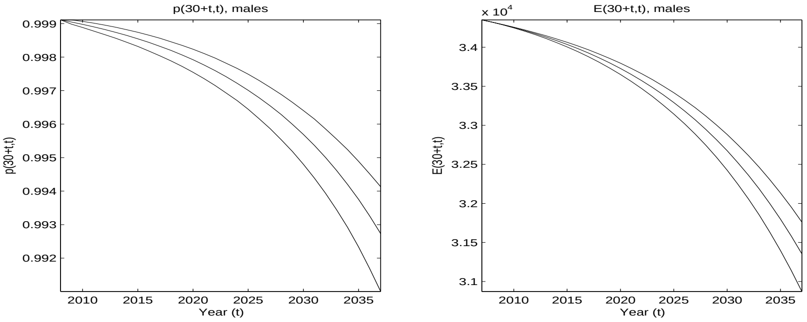

Figure 13: Medians and 90% confidence intervals for survival probabilities $p(x_0,t)$ and cohort sizes $E(x_0,t)$. Cohort aged 30 in 2007, male. Two-factor model.

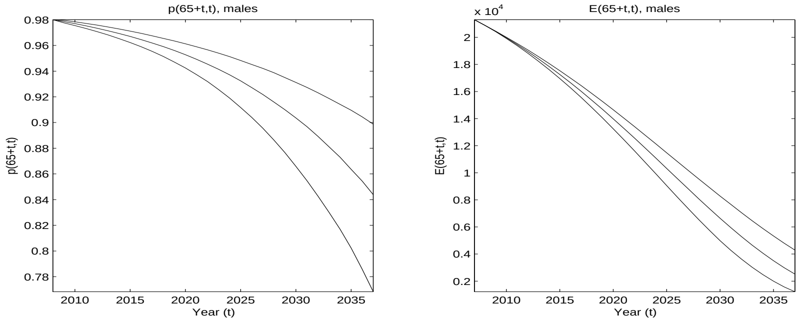

Figure 14: Medians and 90% confidence intervals for survival probabilities $p(x_0,t)$ and cohort sizes $E(x_0,t)$. Cohort aged 65 in 2007, male. Two-factor model.

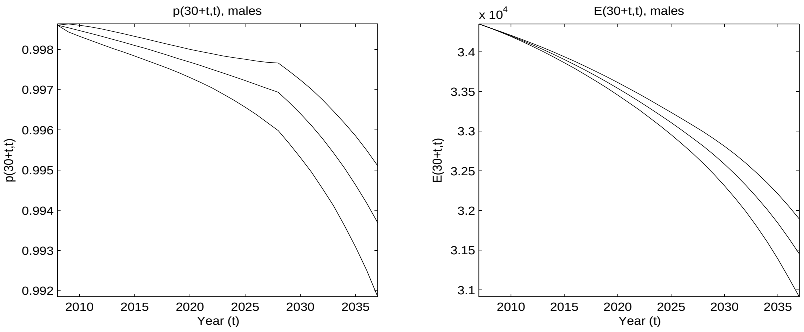

Figure 15: Medians and 90% confidence intervals for survival probabilities $p(x_0 + t, t)$ and cohort sizes $E(x_0 + t, t)$. Cohort aged 30 in 2007, male. Three-factor model.

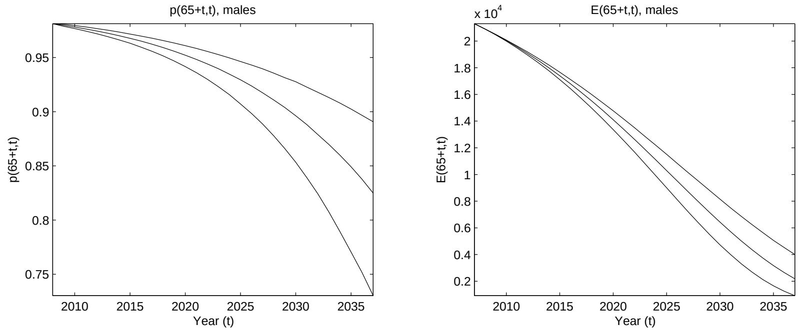

Figure 16: Medians and 90% confidence intervals for survival probabilities $p(x_0 + t, t)$ and cohort sizes $E(x_0 + t, t)$. Cohort aged 65 in 2007, male. Three-factor model.

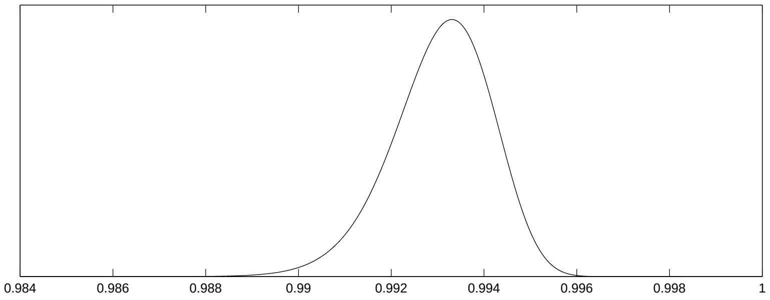

Figure 17: Distribution of the survival probability of 60-year-olds in year 2037.

4 Conclusions

This paper proposed a flexible but simple framework for stochastic mortality modelling. The framework allows for risk factors with tangible interpretations, and guarantees that the parameter estimates are well-defined. Two- and three-factor versions of the model were fitted into Finnish adult population and mortality data, and the factors were modelled as a simple random walk process with a drift. Using the resulting models, future values of death probabilities and cohort sizes were simulated with plausible results.

In real-life applications, modelling of the risk factors should receive more attention. A straightforward extension would be to use more flexible econometric models. For example, by allowing heavy tailed distributions one might be able to better capture the effects of epidemics or natural disasters on mortality. When describing the uncertainty in future development of longevity, the model need not be based completely on historical data, as that might not produce the best description of the future development of mortality. Instead, expert views of future mortality and survival probabilities could be incorporated into the model.

Another interesting aspect is the interplay between mortality and the financial instruments that are used to hedge mortality linked securities. The dependence structures therein are instrumental in asset and liability management of insurance companies, and in pricing and hedging of mortality-linked financial products; see e.g. [19, 18].

References

[1] J. M. Alho, S. Hougaard Jensen, and J. Lassila, editors. *Uncertain Demographics and Fiscal Sustainability*. Cambridge University Press, New York, NY, 2008.

[2] J. M. Alho and B. D. Spencer. *Statistical Demography and Forecasting*. Springer Series in Statistics. Springer, New York, NY, 2005.

[3] A. Ben-Tal and A. Nemirovski. *Lectures on modern convex optimization*. MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001. Analysis, algorithms, and engineering applications.

[4] E. Biffis and D. Blake. Mortality-linked securities and derivatives. *Pensions Institute Discussion Paper*, PI 0901, 2009.

[5] D. Blake, A. J. G. Cairns, and K. Dowd. Living with mortality: Longevity bonds and other mortality-linked securities. *British Actuarial Journal*, 12:153–197(45), 2006.

[6] H. Booth, J. Maindonald, and L. Smith. Applying Lee-Carter under conditions of variable mortality decline. *Population Studies*, 56(3):325–336, 2002.

[7] N. Brouhns, M. Denuit, and J. Vermunt. A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics*, 31(3):373–393, December 2002.

[8] A. Cairns. A discussion of parameter and model uncertainty in insurance. *Insurance: Mathematics and Economics*, 27:313–330, 2000.

[9] A.J. G. Cairns, D. Blake, and K. Dowd. A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *Journal of Risk & Insurance*, 73(4):687–718, 2006.

[10] A.J.G. Cairns, D. Blake, and K. Dowd. Modelling and management of mortality risk: a review. *Scandinavian Actuarial Journal*, 2-3:79–113, 2008.

[11] A.J.G. Cairns, D. Blake, K. Dowd, G. D. Coughlan, D. Epstein, A. Ong, and I Balevich. A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *Pensions Institute Discussion Paper*, PI 0701, 2007.

[12] I. D. Currie, M. Durban, and P. H. C. Eilers. Smoothing and forecasting mortality rates. *Statistical Modelling: An International Journal*, 4(4):279 – 298, 2004.

[13] M. Dahl. Stochastic mortality in life insurance: market reserves and mortality-linked insurance contracts. *Insurance: Mathematics and Economics*, 35(1):113 – 136, 2004.

[14] M. Dahl and T. Moller. Valuation and hedging of life insurance liabilities with systematic mortality risk. *Insurance: Mathematics and Economics*, 39(2):193–217, October 2006.

[15] P. De Jong and L Tickle. Extending the Lee-Carter model of mortality projection. *Mathematical Population Studies*, 13:1–18, 2006.

[16] A. Delwarde, M. Denuit, and P. Eilers. Smoothing the Lee-Carter and Poisson log-bilinear models for mortality forecasting: A penalised log-likelihood approach. *Statistical Modelling*, 7:29–48, 2007.

[17] K. Dowd, D. Blake, A.J.G Cairns, and P. Dawson. Survivor swaps. *The Journal of Risk and Insurance*, 73(1):1–17, 2006.

[18] P. Hilli, M. Koivu, and T. Pennanen. Cash-flow based valuation of pension liabilities. *Submitted*.

[19] P. Hilli, M. Koivu, and T. Pennanen. Optimal construction of a fund of funds. *Submitted*.

[20] R.. Lee and L. Carter. Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association*, 87(419):659–671, 1992.

[21] R. Lee and T. Miller. Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography*, 38(4):537–549, 2001.

[22] Y. Lin and S. H. Cox. Securitization of mortality risks in life annuities. *The Journal of Risk and Insurance*, 72(2):227–252, 2005.

[23] M.A. Milevsky and S.D. Promislow. Mortality derivatives and the option to annuitise. *Insurance: Mathematics and Economics*, 29(3):299 – 318, 2001. Papers presented at the 4th IME Conference, Universitat de Barcelona, Barcelona, 24-26 July 2001.

[24] Yu. Nesterov. *Introductory lectures on convex optimization*, volume 87 of *Applied Optimization*. Kluwer Academic Publishers, Boston, MA, 2004. A basic course.

[25] A. E. Renshaw and S. Haberman. Lee-Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics*, 33(2):255–272, October 2003.

[26] A.E. Renshaw and S. Haberman. A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, 38(3):556 – 570, 2006.

[27] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.

[28] G. Schwartz. Estimating the dimension of a model. *Annals of Statistics*, 6(2):461–464, 1978.