A robust approach to stochastic mortality modelling

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

Abstract

This paper proposes a general approach to stochastic mortality modelling. The logit transforms of death probabilities are modelled by linear combinations of user-specified basis functions. The resulting model is flexible yet simple, and allows for natural interpretations of the underlying risk factors. Deaths per year are assumed to be binomially distributed, which results in a strictly concave log-likelihood function when calibrating the model. This enables the use of convex optimization tools, and guarantees robustness in calibration. We fit two versions of the model into Finnish adult (18-100 years) mortality data, and present simulations for the future development of Finnish mortality.

Keywords: Mortality risk, longevity risk,stochastic modelling, convexity, basis functions

1 Introduction

General longevity has improved significantly over the 20th century, resulting in to some extent 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 behaviour. The pensions industry as well as national social security systems are now incurring the costs of unpredictably improved longevity, as they need to pay out benefits for much longer than was anticipated. 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. As a consequence, mortality-linked securities have recently attracted rapidly increasing attention (see

e.g. [3] and [2] for overviews of possible and existing instruments). Mortality-linked instruments that have already been implemented include mortality swaps [14], bonds [17] and forwards.

Stochastic mortality modelling has produced a variety of approaches for capturing the randomness in the future development of mortality rates (see [7] for a recent review). The earliest and still widely popular discrete-time model with one stochastic factor was introduced by Lee and Carter [15] in 1992. It was followed by a number of modifications (see e.g. [5, 16, 4, 13, 12]), varying the original model and addressing its shortcomings. Models with multiple stochastic factors were subsequently proposed by Renshaw and Haberman [20] and Cairns et al. [6], with extensions incorporating cohort effects by Renshaw and Haberman [21] and Cairns et al.[8]. Currie et al. [9] 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. [18, 10, 11]).

We propose a general discrete-time stochastic mortality modelling framework, which is flexible but relatively simple, enabling population-specific charasteristics as well as user preferences to be taken into consideration. An important advantageous feature of the model is the robustness of its calibration. Moreover, the framework allows for a choice of risk factors with tangible interpretations.

The logit transforms of mortality rates are modelled by linear combinations of user-specified basis functions on the cohorts, and the weights of the basis functions in the linear combinations are the stochastic factors of the model. 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 consideration when calibrating the model.

The weights of the basis functions are the stochastic factors capturing the uncertainty in the future mortality rates. An appropriate choice of basis functions ensures that the factors of the model have an easy interpretation, for instance as the logit transforms of the mortality probabilities in certain cohorts, which facilitates the assessment of the model, and enables the study of the relationships between mortality rates and economic factors. This is a central issue in the hedging of mortality-linked securities.

The chosen model is fitted into data by the maximum log-likelihood method. Deviating from the usual assumption of Poisson distributed deaths, we assume deaths to be binomially distributed. This, combined with the chosen parameterization, results in a strictly concave target function for maximization. This not only means that the problem has a unique maximum point but also enables the use of convex optimization tools, and guarantees robustness of the calibration.

As an example, we consider Finnish adult (from 18 to 100 years) mortality for males and females in the past semicentury. Basis functions are chosen to be piecewise linear, and we consider two exemplary models with two and three basis

functions. Using the resulting model, we present simulations for future development of Finnish mortality rates.

The remaining paper is organized as follows. Section 2 outlines the modelling procedure. Section 3 fits two versions of the model into Finnish mortality data, discussing the appropriateness and fit of each model. Section 4 presents some simulation results for mortality probabilities as well as the future development of certain reference cohorts. In Section 5 we give conclusions and a brief discussion on possible future refinements and applications of the model.

2 Model specification

Let $E(x,t)$ be the size of population aged x (cohort) at the beginning of year t, and define x and t to be integer-valued. Denoting the number of deaths occurring in cohort x during year t by $D(x,t)$, we have

$$D(x,t) = E(x,t) - E(x+1,t+1), \tag{1}$$

which is observable at time $t+1$, i.e. at the beginning of year $t+1$. We assume that the conditional distribution of $D(x,t)$ given $E(x,t)$ is binomial:

$$D(x,t) \sim \text{Bin}(E(x,t), q(x,t)), \tag{2}$$

where $q(x,t)$ is the probability that an individual aged x and alive at the beginning of year t dies before the end of the year. The list $q(\cdot,t)$ of numbers may thus be interpreted as the *mortality table* for year t. A stochastic mortality model is obtained by modellingthe probabilities $q(x,t)$ as stochastic processes. The future deaths $D(x,t)$ and population sizes $E(x+1,t+1)$ are then obtained by sampling from $\text{Bin}(E(x,t), q(x,t))$.

The uncertainty in the future values of $q(x,t)$ may be interpreted as the *systematic mortality risk*. As the population grows, the fraction $D(x,t)/[E(x,t)q(x,t)]$ converges in distribution to constant 1. For large populations, the population dynamics is thus well described by $D(x,t) = E(x,t)q(x,t)$ and the main uncertainty comes from unpredictable variations in the future values of $q(x,t)$.

In order to get a tractable model, we propose to model the logistic probabilities by

$$\operatorname{logit} q(x,t) = \ln\left(\frac{q(x,t)}{1-q(x,t)}\right) = \sum_{i=1}^{n} w_i(t)\phi_i(x),$$

where $\phi_i(x)$ are user-defined *basis functions* and $w_i(t)$ are stochastic risk factors that vary over time. In other words, $q(x,t) = q_{w(t)}(x)$, where $w(t) = (w_1(t), \ldots, w_n(t))$ and q_w is the parametric function defined for each $w \in \mathbb{R}$ by

$$q_w(x) = \frac{\exp\left(\sum_{i=1}^{n} w_i\phi_i(x)\right)}{1+\exp(\sum_{i=1}^{n} w_i\phi_i(x))}, \tag{3}$$

Modelling the logit transforms instead of $q(x,t)$ ensures that $q(x,t) \in (0,1)$ so that they are indeed probabilities. With appropriate choices of the basis functions $\phi_i(\cdot)$ one can guarantee certain desirable properties of $q(\cdot,t)$ in the model. For example, one may wish to construct a model where the probabilities $q(x,t)$ behave continuously or smoothly across the cohorts. This can be achieved simply by choosing continuous or smooth basis functions, respectively.

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 cohort x, then the isk factor $w_k(t)$ equals the logistic death probability in cohort x in 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 w 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.

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

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

is the zero vector $z = 0$. 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 death probabilities on A in the model.

Example 1 *In ([6]), Cairns et al. introduced a 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 age over the range of cohorts. It can also be written as

$$\operatorname{logit} q(x,t) = w_1(t)\phi_1(x) + w_2(t)\phi_2(x),$$

where $w_i = \kappa_i$ for $i = 1,2$, and the basis functions are $\phi_1(x) = 1$ and $\phi_2(x) = (x - \bar{x})$. w_1 can be interpreted as the general level of mortality, while w_2 captures how the relationship between mortalities in different age groups develops. The basis functions are linearly independent on any set of cohorts A, and also orthogonal:

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

Once the basis functions ϕ_i have been chosen, we will model the vector $w(t) = (w_1(t), \ldots, w_n(t))$ of risk factors as a multivariate stochastic process in discrete time. The simplest (nontrivial) choice would be to model w as a random walk with drift but one could also use more sophisticated models developed in the vast literaure on econometric modelling. The model specification could be based solely on the user's views about the future development of mortality or on historical data or on both. The historical values of the risk factors $w(t) = (w_1(t), \ldots, w_n(t))$ can be easily constructed by maximum likelihood estimation as follows.

Given the historical values of $E(x,t)$ and $D(x,t)$, the log-likelihood function for yearly values of $w(t)$ can be written using (2) and (3) as

$$\begin{aligned} l_t(w) &= \ln \prod_x \binom{E(x,t)}{D(x,t)} q_w(x)^{D(x,t)} \left(1 - q_w(x)\right)^{E(x,t)-D(x,t)} \qquad (4) \\ &= \sum_x \left[\ln \binom{E(x,t)}{D(x,t)} + D(x,t) \ln q_w(x) + (E(x,t) - D(x,t)) \ln(1 - q_w(x)) \right] \\ &= \sum_x \left[D(x,t) \sum_i w_i \phi_i(x) - E(x,t) \ln(1 + e^{\sum_i w_i \phi_i(x)}) + \ln \binom{E(x,t)}{D(x,t)} \right]. \end{aligned}$$

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

Proposition 2 *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 cohorts*

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

then l_t is strictly concave.

Proof. Looking at the last expression for $l_t(w)$, we see that the first term in the brackets is linear in $w = (w_1, \ldots, w_n)$ so it is concave. The last term is constant so it is trivially concave. Since $E(x,t)$ is nonnegative, the first claim follows if we can show that the functions

$$g_x(w) = \ln(1 + e^{\sum_i w_i \phi_i(x)})$$

are convex in w. Since g_x is the composition of the linear function $w \mapsto \sum_i w_i \phi_i(x)$ with the function $\varphi(z) = \ln(1 + e^z)$ it suffices to show that $\varphi : \mathbb{R} \to \mathbb{R}$ is convex; see [22, Theorem 5.7]. It is easily checked that the second derivative of φ is

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

which is strictly positive on $\mathbb{R}$. This implies that φ is strictly convex on $\mathbb{R}$ (see e.g. [23, Theorem 2.13]) proving the first claim.

As to the second claim, we first note that the log-likelihood function l_t can be written as the difference of an affine function and

$$f_t(w) = \sum_{x \in A(t)} E(x,t)\varphi\left(\sum_{i=1}^{n} \phi_i(x) w_i\right).$$

It suffices to show that f_t is strictly convex. By [23, Theorem 2.14], this holds if the Hessian matrix $\nabla^2 f_t(w)$ of f_t is positive definite for all $w \in \mathbb{R}^n$. Since

$$\frac{\partial^2 f_t(w)}{\partial w_i \partial w_j} = \sum_{x \in A(t)} E(x,t)\varphi''\left(\sum_{i=1}^{n} \phi_i(x) w_i\right) \phi_i(x)\phi_j(x),$$

we get

$$\nabla^2 f_t(w) = \Phi^T D(w) \Phi,$$

where $D(w)$ is the diagonal matrix with entries

$$\left[E(x,t)\varphi''\left(\sum_{i=1}^{n} \phi_i(x) w_i\right)\right]_{x \in A(t)}$$

and the matrix Φ has the vectors $[\phi_i(x)]_{x \in A(t)}$ as its columns $i = 1, \ldots, n$. Since the diagonal elements of $D(w)$ are strictly positive for every $w \in \mathbb{R}^n$ and since Φ has full rank by the independence assumption, the Hessian is positive definite. $\square$

The above is a definite advantage of the assumption of binomially distributed deaths over to the more common Poisson assumption. Convexity implies that local maxima of l_w are true maximum likelihood estimators. Strict convexity implies that the estimator is unique; see e.g. [23, 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. [1, 19].

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 death probability curve $q(x,t)$.

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

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

4. The future death probabilities are given by

$$q(x,t) = \left[1 + \exp\left(\sum_i w_i(t)\phi_i(x)\right)\right]^{-1}$$

5. The future deaths $D(x,t)$ are obtained by sampling from Bin $(E(x,t), q(x,t))$ or simply by $D(x,t) = E(x,t)q(x,t)$, if we are only interested in systematic mortality risk.

3 Fitting the model into Finnish mortality data

In this section we present two specifications of the stochastic mortality model described in the previous section, and fit these two models into Finnish mortality data. Both models employ piecewise linear basis functions $\phi_i(x)$. The first model consists of two stochastic factors $w = (w_1, w_2)$ and hence two basis functions, while the second one is a three-factor model with $w = (w_1, w_2, w_3)$.

We use Finnish mortality data for Finnish males and females, covering ages 18-100 and years 1900-2007. Younger ages were included in view of the fact that in the Finnish pension system the accrual of earnings-related pension starts at the age of 18. The data was obtained from the Human Mortality Database [1], and contains annual values for the numbers of deaths $D(x,t)$ during the year t of individuals aged x at the beginning of that year, and corresponding population sizes $E(x,t)$.

To produce estimates for historical values of the stochastic factors w, the expression in Equation 4 is maximized using Matlab Optimization Toolbox.

3.1 Two-parameter model

We first consider a two-parameter application of the stochastic mortality model of Section 2, which is of the form

$$\text{logit} \quad q(x,t) = w_1(t)\phi_1(x) + w_2(t)\phi_2(x). \tag{5}$$

The two linear basis functions are chosen as

$$\phi_1(x) = 1 - \frac{x-18}{82} \quad \text{and} \quad \phi_2(100) = \frac{x-18}{82}. \tag{6}$$

Consequently, the linear combination $\sum_{i=1}^{2} w_i\phi_i(x)$ is now also linear, which is illustrated in Figure 1.

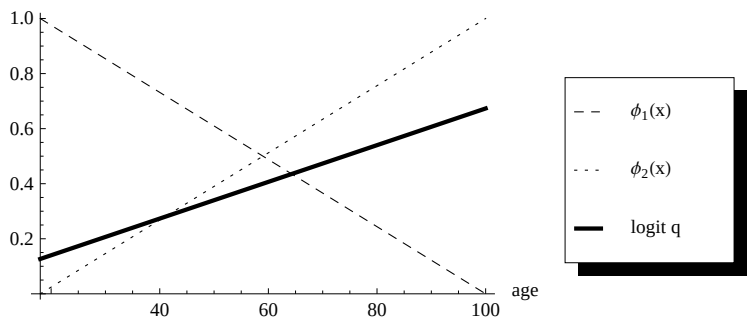

Figure 1: Linear basis functions and an exemplary linear combination

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

Note that for a given t,

$$\text{logit } \; q(18,t) = w_1(t)\phi_1(18) + w_2(t)\phi_2(18) = w_1(t), \tag{7}$$

and, similarly, $\text{logit } \; q(100,t) = w_2(t)$. In other words, the two factor values for each t are two points on the line for $\text{logit } \; q(x,t)$, and hence have a natural interpretation. This is particularly advantageous when assessing the model in the light of estimation results for historical values of w, or pricing mortality-linked financial instruments.

Estimated historical values of w for males and females are presented in Figures 2 and 3. Figures 4 and 5 show the observed death rates $\frac{D(x,t)}{E(x,t)}$ in comparison with the estimated mortality rates $q(x,t)$. $\frac{D(x,t)}{E(x,t)}$ illustrates the effect of both systematic and nonsystematic mortality, while $q(x,t)$ contains only systematic mortality.

The effect of Finnish war years (1918 and 1939-1944) can be clearly observed in the estimated values for males. However, the gaps observed in values of w_2 during the war years imply that the mortality rate $q(100,t)$ would have temporarily decreased. As it is doubtful that this is true, we conclude that this model may be too simple to capture the population dynamics in Finland during the entire 20th century. This conclusion is also supported by the fact that the observed death rates do not show the kind of sharp, temporary decline in the war years that the estimated $q(x,t)$s suggest.

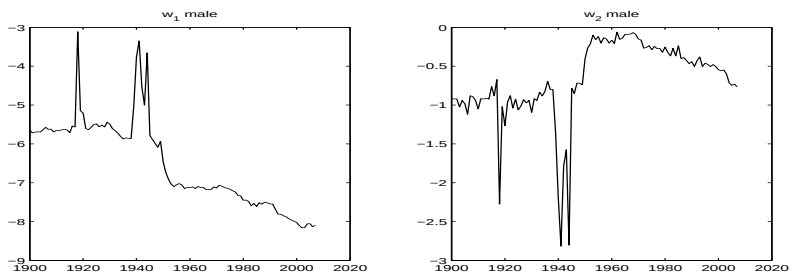

Figure 2: Estimated parameter values, males

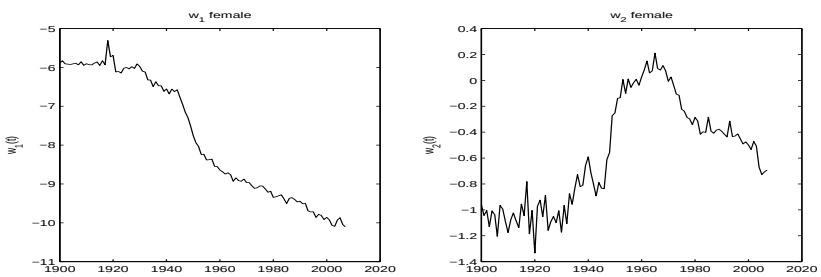

Figure 3: Estimated parameter values, females

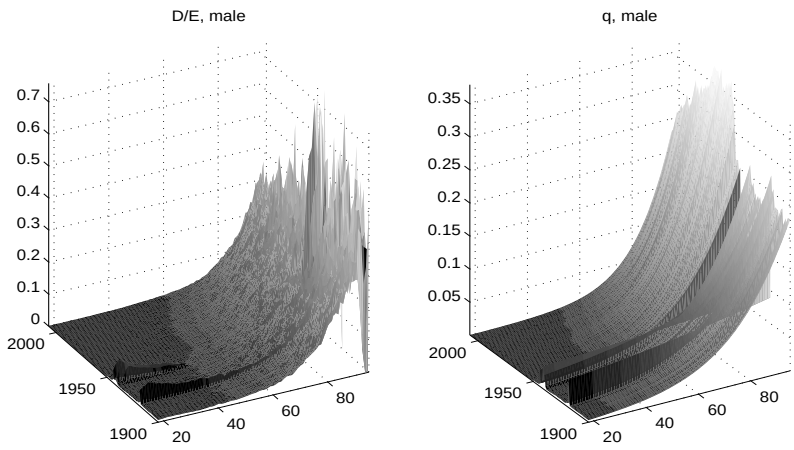

Figure 4: $\frac{D(x,t)}{E(x,t)}$ vs. estimated values of $q(x,t)$, males

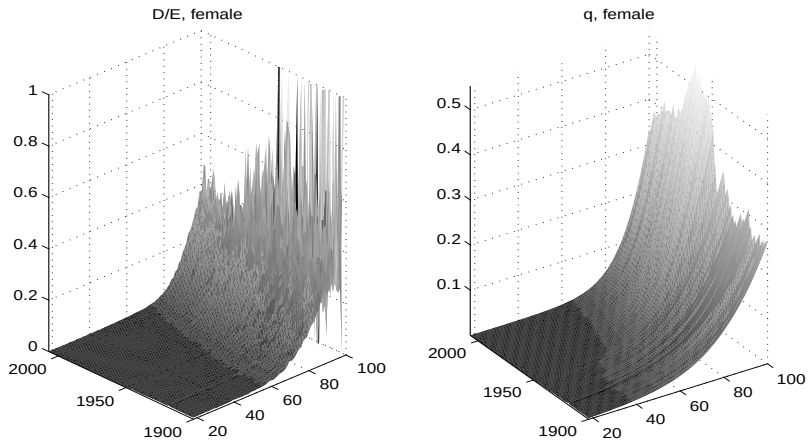

Figure 5: $\frac{D(x,t)}{E(x,t)}$ vs. estimated values of $q(x,t)$, females

3.2 Three-parameter model

The second model to consider is a three-parameter one with representation

$$\text{logit} \quad q(x,t) = w_1(t)\phi_1(x) + w_2(t)\phi_2(x) + w_3(t)\phi_3(x). \tag{8}$$

Basis functions $\phi_i(x)$ are piecewise linear, and can be written as

$$\phi_1(x) = 1 - \frac{x-18}{32} \quad , \quad \phi_2(x) = \begin{cases} \frac{1}{32} & x \leq 50 \\ 2 - \frac{x}{50} & x > 50 \end{cases} \text{ and} \tag{9}$$

$$\phi_3(x) = \frac{x}{50} + 1 \tag{10}$$

As opposed to the first model, the linear combination $\sum_{i=1}^{3} w_i \phi_i(x)$ is now piecewise linear, as depicted in Figure 6. Similarly as in the first case, the values of the factors are points in the curve for $\operatorname{logit} q$: $\operatorname{logit}\ q(18,t) = w_1(t)$, $\operatorname{logit}\ q(50,t) = w_2(t)$ and $\operatorname{logit}\ q(100,t) = w_3(t)$.

Estimation results for factor values are plotted in Figures 7 and 8. Although there are comparatively large fluctuations in the values of w_3 which now represent $\operatorname{logit} q(100,t)$ in the first half of the observation period, this appears more plausible than the gaps observed in the values for w_2 in the 2-parameter model. In both plots, a small bulge appears in the values for w_1 towards the end of the period.

The surface plots for estimated $q(x,t)$ in Figure 9 now resemble more those for the observed death rates $\frac{D(x,t)}{E(x,t)}$. The gap in Figure 4 does not appear here. In other words, the fit seems to have improved especially for the estimation results for males.

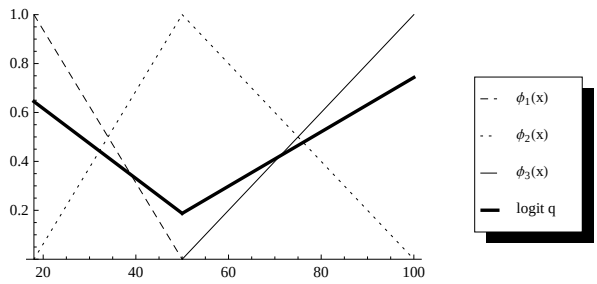

Figure 6: Three piecewise linear basis functions and their linear combination

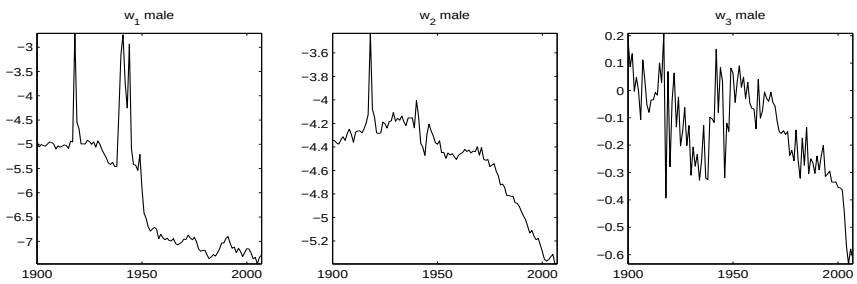

Figure 7: Estimated values of w for three-parameter model, males

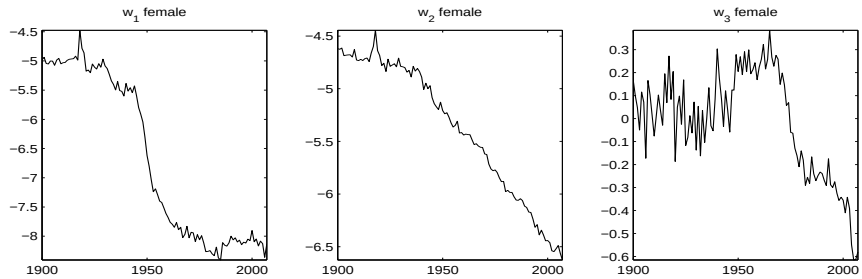

Figure 8: Estimated values of w for three-parameter model, females

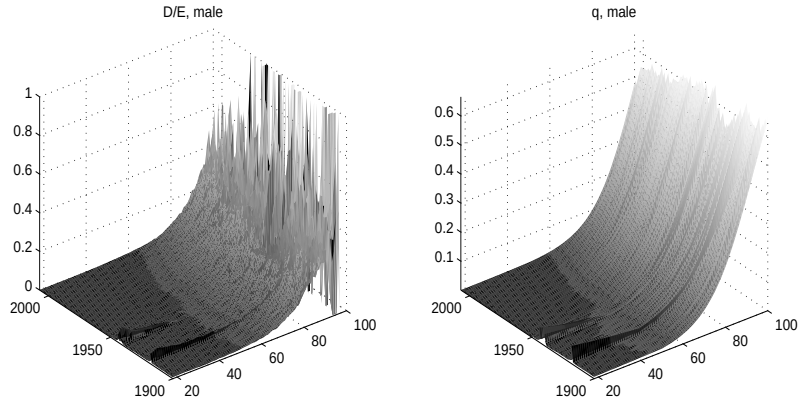

Figure 9: Estimated values of $q(x,t)$ vs. $\frac{D(x,t)}{E(x,t)}$, males

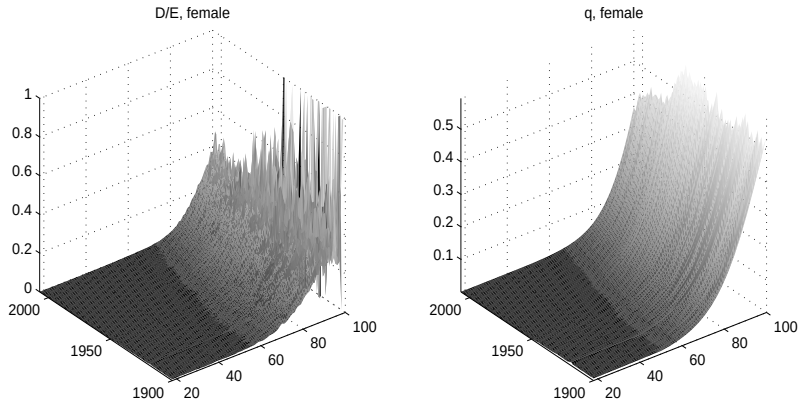

Figure 10: Estimated values of $q(x,t)$ vs. $\frac{D(x,t)}{E(x,t)}$, females

3.3 Modelling the process w

The estimated values for $w(t)$ for both models show componentwise a downward trend from about year 1960 onwards. The trends reflect a general improvement in

mortality over time. The time series also displayed varying degrees of fluctuations, the largest being observable at component w_3 of the three-parameter model for both males and females. This signals larger variance in the systematic mortality for older age groups. For simulation and forecasting purposes, we combine the male and female parameter estimates and represent $w(t) = (w_f(t), w_m(t))^T$ for both models as 4- and 6-dimensional random walks with a drift. General n-dimensional form is written as

$$w(t) - w(t-1) = \mu + CZ(t), \tag{11}$$

where $\mu \in \mathbb{R}^{n\times 1}$ is a constant drift vector, $C \in \mathbb{R}^{n\times n}$ is a constant lower diagonal matrix, and $Z(t)$ is a $n \times 1$ vector of independent standard Gaussian random variables: $Z_i(t) \sim N(0,1)$. μ is the mean vector of $\Delta w(t)$, and the matrix C is obtained from the Cholesky decomposition of the covariance matrix $V = \text{Diag}(\sigma) R\ \text{Diag}(\sigma) = CC'$ for estimated historical values of $\Delta w(t)$. The vector $\sigma \in \mathbb{R}^{n\times 1}$ is the standard deviation vector, and $R \in \mathbb{R}^{n\times n}$ is the correlation matrix. Using data for years 1960-2007 we find the results to be

$$\mu = \begin{pmatrix} -0.0311 \\ -0.0153 \\ -0.0207 \\ -0.0128 \end{pmatrix} \sigma = \begin{pmatrix} 0.0799 \\ 0.0637 \\ 0.0535 \\ 0.0626 \end{pmatrix} \tag{12}$$

$$R = \begin{pmatrix} 1.0000 & -0.6890 & 0.4051 & -0.3565 \\ -0.6890 & 1.0000 & -0.3279 & 0.6655 \\ 0.4051 & -0.3279 & 1.0000 & -0.6605 \\ -0.3565 & 0.6655 & -0.6605 & 1.0000 \end{pmatrix} \tag{13}$$

$$\mu = \begin{pmatrix} -0.0097 \\ -0.0252 \\ -0.0171 \\ -0.0075 \\ -0.0196 \\ -0.0116 \end{pmatrix} \sigma = \begin{pmatrix} 0.1149 \\ 0.0352 \\ 0.0662 \\ 0.0765 \\ 0.0372 \\ 0.0728 \end{pmatrix} \tag{14}$$

$$R = \begin{pmatrix} 1.0000 & 0.1246 & 0.1277 & 0.1341 & 0.1690 & -0.0573 \\ 0.1246 & 1.0000 & -0.3064 & -0.2564 & 0.4259 & -0.1741 \\ 0.1277 & -0.3064 & 1.0000 & 0.1067 & 0.0528 & 0.6282 \\ 0.1341 & -0.2564 & 0.1067 & 1.0000 & -0.4031 & 0.1625 \\ 0.1690 & 0.4259 & 0.0528 & -0.4031 & 1.0000 & -0.3214 \\ -0.0573 & -0.1741 & 0.6282 & 0.1625 & -0.3214 & 1.0000 \end{pmatrix}. \tag{15}$$

The observation period was chosen on the grounds that the factor values displayed fairly evenly drifting behavior during that time.

4 Simulations

In the simulation experiment, we applied the the two models presented in Chapter 3 to consider mortality probabilities and cohort sizes 30 years into the future. In both cases, the process $w(t)$ was simulated according to Equation 11, using estimates

for μ and V from the fitting of both models, based on data from years 1960-2007. The simulations were performed with the Monte Carlo method, with sample size $N = 10000$. Probabilities $q(x,t)$ were then calculated from the simulated values of $w(t)$. The number of deaths $D(x,t)$ in each cohort was approximated by its expected value $E(x,t)q(x,t)$, and cohort sizes at the end of the year were obtained by subtracting that value from $E(x,t)$, the cohort size at the beginning of the year t.

Cohorts aged 30 and 65 in the final observation year 2007 were chosen as reference cohorts. Figures 11 and 12 show the simulated means and 90% confidence intervals for $q(.,t)$ and cohort sizes $E(.,t)$ in the two-parameter case. Figures 13 and 14 give the respective results for the three-parameter model. The 30-year-old cohort displays a more notable difference between the two models than the 65-year-old one. For the younger cohort, the cohort size estimates for the two-parameter model in the final simulation year are slightly below that of the three-parameter model. For the older reference cohort the difference, albeit hardly notable, is the other way round.

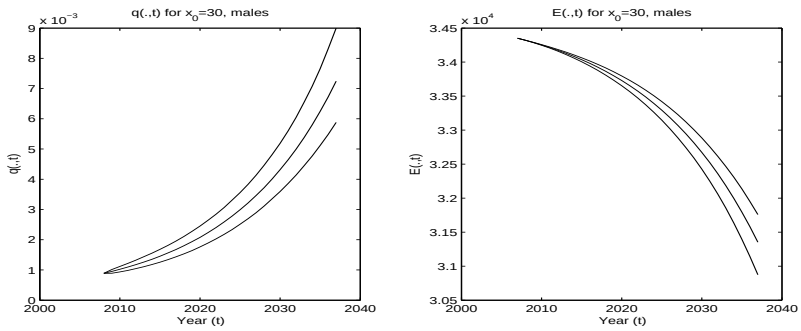

Figure 11: Medians and 90% confidence intervals for mortality rates $q(.,t)$ and cohort sizes $E(.,t)$. Cohort aged 30 in 2007, male

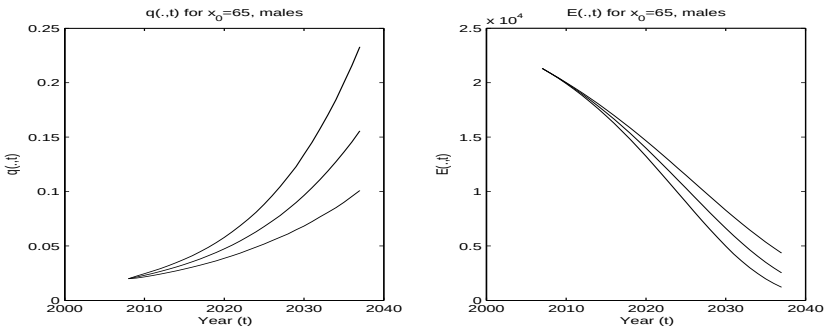

Figure 12: Medians and 90% confidence intervals for mortality rates $q(.,t)$ and cohort sizes $E(.,t)$. Cohort aged 65 in 2007, male

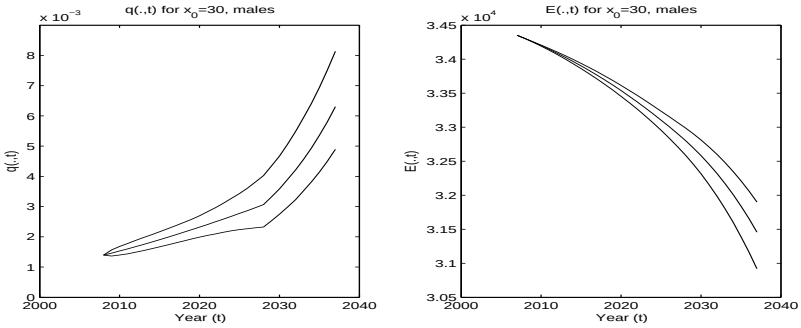

Figure 13: Medians and 90% confidence intervals for mortality rates $q(.,t)$ and cohort sizes $E(.,t)$. Cohort aged 30 in 2007, male

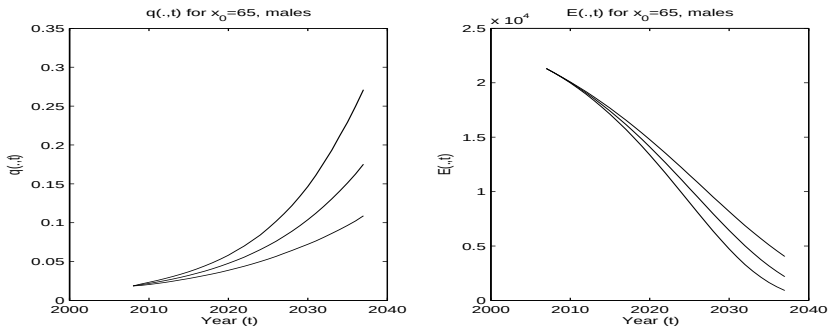

Figure 14: Medians and 90% confidence intervals for mortality rates $q(.,t)$ and cohort sizes $E(.,t)$. Cohort aged 65 in 2007, male

5 Conclusions

This paper proposed a flexible but relatively simple framework for stochastic mortality modelling. The model allows for risk factors with tangible interpretations, and features robustness in parameter estimations. Two- and three-factor versions of the model were fitted into Finnish adult 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 careful 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 uncertain future development of mortality tables, the specification need not be based completely on historical data, because that might not produce the most valid description of the future development of mortality. Instead, expert views could be incorporated into the model. Another interesting is the inter-

play between mortality and the financial instruments that are used to hedge mortality linked securities. The dependence structure is an important factor in asset and liability management of insurance companies, and in pricing and hedging of mortality-linked financial products.

References

[1] 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.

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

[3] 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.

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

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

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

[11] 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.

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

[13] 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.

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

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

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

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

[18] 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.

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

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

[21] 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.

[22] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[23] 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.