Gaussian Forward Mortality Factor Models: Specification, Calibration, and Application*

Nan Zhu
Department of Risk Management and Insurance, Georgia State University
35 Broad Street, 11th Floor; Atlanta, GA 30303; USA
Email: nzhu1@student.gsu.edu

Daniel Bauer†
Department of Risk Management and Insurance, Georgia State University
35 Broad Street, 11th Floor; Atlanta, GA 30303; USA
Email: dbauer@gsu.edu

September 2010

Abstract

Two of the most important challenges for the application of stochastic mortality models in life insurance practice are their complexity and the apparent incompatibility with classical life contingencies theory, which provides the backbone of insurers' EDP systems.

One model class that overcomes these challenges are so-called *Gaussian Forward Mortality Factor Models*. This paper determines suitable specifications based on a Principal Component Analysis of generation mortality data, addresses their calibration, and discusses practically important example applications. In particular, we derive the Economic Capital for a stylized life insurance company. Our preliminary analyses indicate that the presence of mortality risk has a drastic impact on the results.

Keywords: Stochastic Mortality Models. Forward Mortality Models. Principal Component Analysis. Economic Capital.

1 Introduction

Two of the most important challenges for the application of stochastic mortality models in life insurance practice are the apparent incompatibility of most stochastic methods with classical life contingencies theory, which presents the backbone of insurers' EDP systems, and the complexity of many of the proposed approaches. These obstacles have not only led to an increasing discrepancy between life insurance research and actuarial practice and education, but the reluctance of practitioners to rely on stochastic mortality models may also be a primary reason for the sluggish development of the longevity-linked capital market. In

*Preliminary version prepared for the Sixth International Longevity Risk and Capital Markets Solutions Conference. Parts of this paper are taken from the second author's doctoral dissertation (Bauer (2009)). We are grateful to Matthias Börger for the supply of the pre-compiled generation data. D. Bauer gratefully acknowledges financial support from the Willis Research Network.

†Corresponding author. Phone: +1-(404)-413-7490; Fax: +1-(404)-413-7499.

particular, stochastic methods are necessary to assess a company's capital relief when hedging part of their longevity risk exposure, which should be one of the key drivers for the demand of longevity-linked securities.

One model class that overcomes these problems are so-called *forward mortality models*, which infer dynamics on the entire age/term-structure of mortality. As already pointed out by Milevsky and Promislow (2001), the "traditional rates used by actuaries" really are *forward rates* so that such an approach can be viewed as the natural extension of traditional actuarial theory. In particular, the actuarial present values for traditional insurance products such as term-life insurance, endowment insurance, or life annuity contracts are of the same form as in classical actuarial theory, where the "survival probabilities" now are to be interpreted as expected values of realized survival probabilities (cf. Bauer (2009)). Hence, the inclusion of such models in the operations of a life insurer or a pension fund will not require alterations in the management of traditional product lines, but nonetheless present a coherent way to take mortality risk into account when necessary. Examples of such situations include the calculation of economic capital based on internal models or the pricing and risk management of mortality-linked guarantees in life insurance or pension products.

However, only few forward mortality models have been proposed so far, and most authors have relied on "qualitative" insights and/or modeling convenience for determining suitable specifications (cf. Bauer et al. (2008a), Dawson et al. (2010), or Schrager (2006)). Moreover, some of the presented models entail a high degree of complexity, which may lead to problems in their calibration (see e.g. Bauer et al. (2008a)).

The current paper overcomes these shortcomings by relying on *Gaussian forward mortality factor models*, the (necessary) explicit functional form of which has been identified by Bauer et al. (2010a). More specifically, using data for the term structure of mortality rates which either has been compiled based on a fixed assumption of how to furnish deterministic mortality forecasts or is obtained from available generational mortality tables, we use a Principle Component Analysis to determine a suitable number of stochastic factors and their functional form. These specifications are then (re)calibrated based on Maximum Likelihood Estimation.

To demonstrate the advantages in applications of forward mortality models and Gaussian factor models in particular, we consider two representative and practically important examples. The first application concerns the calculation of Economic Capital in life insurance companies. After providing a general framework for this problem, we determine the economic capital for a stylized life insurance company offering traditional products both without and with taking systematic mortality risk into account. Our preliminary analyses indicate that the presence of mortality risk changes the results drastically. As a second application, we discuss the valuation of annuitization options. Specifically, we derive a valuation formula for simple Guaranteed Annuitization Options within traditional endowment policies and describe the Monte-Carlo valuation of Guaranteed Minimum Income Benefits within Variable Annuities based on forward mortality models.

The remainder of the paper is organized as follows: In Section 2, we introduce the relevant definitions and results on forward mortality models from Bauer et al. (2010a); in Section 3, we describe our Principle Component Analyses and present their results; Section 4 discusses the calibration approach and the corresponding results; and our two example applications are put forward in Section 5 and 6, respectively. Finally, Section 7 concludes with a summary of our main results and an outlook on future research.

2 The Forward Mortality Framework

In a best estimate generation life table at time $t \geq 0$, forward survival probabilities[1]

$$ {}_{\tau}p_x(t, t+\tau) := \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\mathbb{P}}\left[\mathbf{1}_{\{\Upsilon_{x-t} > t+\tau\}}\middle| \mathcal{F}_{t+\tau} \vee \{\Upsilon_{x-t} > t\}\right]\middle| \mathcal{F}_t\right],\ x \geq t \geq 0,\ \tau \geq 0, $$

are listed for a (large) collection of ages x and terms τ, where Υ_{x_0} is the (random) time of death or future lifetime of an x_0-year old at time zero. For convenient modeling, we introduce the so-called *forward force of mortality*

$$ \mu_t(\tau, x) = -\frac{\partial}{\partial \tau} \log\left\{{}_{\tau}p_x(t, t+\tau)\right\} $$

and consider time-homogenous diffusion-driven models of the form

$$ d\mu_t = (A\,\mu_t + \alpha)\ dt + \sigma\, dW_t,\ \mu_0(\cdot, \cdot) > 0, \tag{1} $$

where α and $\sigma = (\sigma^{(1)}, \ldots, \sigma^{(d)})$ are suitable continuous functions $\alpha, \sigma^{(i)} : [0, \infty)^2 \to \mathbb{R}$, $1 \leq i \leq d$, $A = \frac{\partial}{\partial \tau} - \frac{\partial}{\partial x}$, and $W = (W^{(1)}, \ldots, W^{(d)})'$ is a d-dimensional standard Brownian motion. In particular, if the dynamics (1) are formulated under $\mathbb{P}$, we have the drift condition (cf. Cor. 3.1 in Bauer et al. (2010a))

$$ \alpha(\tau, x) = \sigma(\tau, x) \times \int_0^{\tau} \sigma'(s, x)\, ds. \tag{2} $$

Note that the forward force of mortality is just a convenient representation of the forward survival probabilities constituting the generation life table at time t:

$$ {}_{\tau}p_x(t, t+\tau) = \exp\left\{-\int_0^{\tau} \mu_t(s, x)\, ds\right\}. $$

Hence, Equation (1) implicitly defines a dynamic model for generation life tables.

A shortcoming of the above formulation is the inherent complexity associated with the infinite-dimensional stochastic differential equation (1): For practical purposes, we would like to have a factor model, i.e. a *finite dimensional realization*

$$ \mu_t(\tau, x) = G\left(\tau, x; Z_t\right), $$

where G is a known deterministic function and Z_t is *some convenient* finite-dimensional random variable (so that $(Z_t)_{t \geq 0}$ is *some convenient* stochastic process). Here "convenient" primarily concerns the econometrical tractability of Z_t and the properties of the resulting model. In particular, if Z_t is Normal distributed, we can use a Principle Component Analysis to determine suitable specifications. This is akin to the fixed income

[1]As usually in this context, underlying our considerations is a filtered probability space $\left(\Omega, \mathcal{H}, \mathbf{F} = (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}\right)$. Here, the filtration $\mathbf{F}$ satisfies the usual conditions and models the information flow of aggregate population dynamics, whereas the sigma algebra $\mathcal{H}$ also contains information about individuals within the population.

literature (see e.g. Litterman and Scheinkman (1991) or Rebonato (1998)), and several authors have taken a similar approach to the analysis of period mortality data (see e.g. Lee and Carter (1992) or Njenga and Sherris (2009)). However, thus far, there has been no attempt to analyze generational mortality data in order to identify the drivers of the entire age/term-structure of mortality.

Proposition 4.1 in Bauer et al. (2010a) shows that for *Gaussian realizations*, that is if $G(T, x; \cdot)$ is affine and Z_t is Normal distributed, the volatility structure must necessarily be of the form

$$\sigma(\tau, x) = C(x + \tau) \times \exp\{M\,\tau\} \times N, \tag{3}$$

where $N \in \mathbb{R}^{m \times d}$, $M \in \mathbb{R}^{m \times m}$, and $C' \in C^1([0, \infty), \mathbb{R}^m)$; a finite-dimensional realization is then given by

$$\mu_t(\tau, x) \;=\; \mu_0(\tau + t, x - t) + \int_0^t \alpha(\tau + t - s, x - t + s)\, ds + C(x + \tau)\, \exp\{M\,\tau\} \underbrace{\int_0^t \exp\{M\,(t - s)\}\; N\, dW_s}_{=Z_t}.$$

This representation can be conveniently employed to determine both the suitable number of factors for such Gaussian forward mortality factor models and their functional form in a Principle Component Analysis. However, before diving into this problem deeper in the next section, we note the following result that will prove to be convenient in what follows.[2]

Proposition 2.1. *Let $\sigma(\tau, x) = (\sigma_1(\tau, x), \ldots, \sigma_d(\tau, x))$, where each function $\sigma_i(\tau, x)$ is of the form*

$$\sigma_i(\tau, x) = C_i(x + \tau) \times \exp\{M_i\,\tau\} \times N_i, \tag{4}$$

$C_i'(x) \in \mathbb{R}^{m_i}$, $M_i \in \mathbb{R}^{m_i \times m_i}$, $N = \mathbb{R}^{m_i \times 1}$, $m_i \in \mathbb{N}$, $i = \{1, 2, \ldots, d\}$. *Then $\sigma(\tau, x)$ is also of the form (3), i.e. the model implied by $\sigma(\tau, x)$ allows for a Gaussian realization, where $C(x) = [C_1(x), \ldots, C_d(x)]$, $M =$ diag $\{M_1, \ldots, M_d\}$, and $N =$ diag $\{N_1, \ldots, N_d\}$.*

Proof.

$$\begin{aligned}
& [\sigma_1(\tau, x), \ldots, \sigma_d(\tau, x)] \\
= \; & [C_1(x + \tau) \times \exp\{M_1\,\tau\} \times N_1, \ldots, C_d(x + \tau) \times \exp\{M_d\,\tau\} \times N_d] \\
= \; & \underbrace{[C_1(x + \tau), \ldots, C_d(x + \tau)]}_{C(x+\tau)} \times \begin{pmatrix} \exp\{M_1\,\tau\} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \exp\{M_2\,\tau\} & & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \exp\{M_d\,\tau\} \end{pmatrix} \times \underbrace{\begin{pmatrix} N_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & N_2 & & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & N_d \end{pmatrix}}_{N} \\
= \; & C(x + \tau) \times \underbrace{\exp\{\text{diag}\{M_1, \ldots, M_d\}\;\tau\}}_{\exp\{M\tau\}} \times N.
\end{aligned}$$

□

[2] A similar result for forward interest rate models is provided in Angelini and Herzel (2005).

3 Principal Component Analyses of Generational Mortality Data

3.1 Approach

Assume that we are given generational mortality data in the form of (forward) survival probabilities $({}_\tau p_x(t_j, t_j+\tau))_{(\tau,x)\in\mathcal{C}}$, for different evaluation dates t_j, $j=1,\ldots,N$, where $\mathcal{C}$ denotes a (large) collection of term/age combinations. Let l denote a lag time, and choose a sub-collection $\tilde{\mathcal{C}} \subset \mathcal{C}$ such that $(\tau+l, x)$, $(\tau+t_{j+1}-t_j, x-t_{j+1}+t_j)$, $(\tau+l+t_{j+1}-t_j, x-t_{j+1}+t_j) \in \mathcal{C}\,, \forall j \in \{1,2,\ldots,N-1\}$, for $(\tau,x)\in\tilde{\mathcal{C}}$, $|\tilde{\mathcal{C}}|=K$.

For each $(\tau,x)\in\tilde{\mathcal{C}}$, $j\in\{1,2,\ldots,N-1\}$, define

$$\begin{aligned} F_l\left(t_j, t_{j+1}, (\tau,x)\right) &= -\log\left\{\frac{\frac{{}_{\tau+l}p_x(t_{j+1},t_{j+1}+\tau+l)}{{}_{\tau+l+t_{j+1}-t_j}p_{x-t_{j+1}+t_j}(t_j,t_{j+1}+\tau+l)}}{\frac{{}_\tau p_x(t_{j+1},t_{j+1}+\tau)}{{}_{\tau+t_{j+1}-t_j}p_{x-t_{j+1}+t_j}(t_j,t_{j+1}+\tau)}}\right\} \\ &= -\log\left\{\frac{{}_{\tau+l}p_x(t_{j+1},t_{j+1}+\tau+l)}{{}_\tau p_x(t_{j+1},t_{j+1}+\tau)}\Big/\frac{{}_{\tau+l+t_{j+1}-t_j}p_{x-t_{j+1}+t_j}(t_j,t_{j+1}+\tau+l)}{{}_{\tau+t_{j+1}-t_j}p_{x-t_{j+1}+t_j}(t_j,t_{j+1}+\tau)}\right\}. \end{aligned}$$

Now with some basic manipulations, we obtain

$$\begin{aligned} F_l\left(t_j, t_{j+1}, (\tau,x)\right) &\stackrel{d}{=} \int_0^{t_{j+1}-t_j}\int_\tau^{\tau+l} \alpha(v+t_{j+1}-t_j-s, x-t_{j+1}+t_j+s)\,dv\,ds \\ &\quad + \int_0^{t_{j+1}-t_j}\int_\tau^{\tau+l} C(x+v)\,\exp\left\{M\left(v+t_{j+1}-t_j-s\right)\right\}\,N\,dv\,dW_s \\ &\stackrel{d}{=} \int_0^{t_{j+1}-t_j}\int_\tau^{\tau+l} \alpha(v+t_{j+1}-t_j-s, x-t_{j+1}+t_j+s)\,dv\,ds \\ &\quad + \underbrace{\int_\tau^{\tau+l} C(x+v)\,\exp\left\{Mv\right\}\,dv}_{=O(\tau,x)} \times \underbrace{\int_0^{t_{j+1}-t_j} \exp\left\{M\left(t_{j+1}-t_j-s\right)\right\}\,N\,dW_s}_{=Z_{t_{j+1}-t_j}} \end{aligned} \tag{5}$$

is Normal distributed. In particular, if the data is equidistant, that is if $t_{j+1}-t_j=\Delta$ is constant, the vectors

$$\begin{aligned} \bar{F}_l(t_j,t_{j+1}) &= \left(\frac{F_l(t_j,t_{j+1},(\tau,x))}{\sqrt{t_{j+1}-t_j}}\right)_{(\tau,x)\in\tilde{\mathcal{C}}} \\ &= \left(\frac{F_l(t_j,t_{j+1},(\tau_1,x_1))}{\sqrt{t_{j+1}-t_j}}, \frac{F_l(t_j,t_{j+1},(\tau_2,x_2))}{\sqrt{t_{j+1}-t_j}}, \ldots, \frac{F_l(t_j,t_{j+1},(\tau_K,x_K))}{\sqrt{t_{j+1}-t_j}}\right) \end{aligned}$$

$j=1,2,\ldots,N-1$, are independent identically distributed, and we may compute the empirical covariance matrix $\hat{\Sigma}$.[3] Then the procedure is standard: We compute $\hat{\Sigma}$ and decompose it as

$$\hat{\Sigma} = U \times \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_K \end{pmatrix} \times U' = \sum_{\nu=1}^{K} \lambda_\nu \, u_\nu \, u'_\nu,$$

where $U = (u_1, u_2, \cdots, u_K)$ is an (orthogonal) matrix consisting of the eigenvectors of $\hat{\Sigma}$ and λ_ν, $\nu = 1, 2, \ldots, K$, are the corresponding (ordered) eigenvalues. We then pick the d greatest eigenvalues that explain the majority of the variation in the data, i.e. we choose d such that

$$\frac{\sum_{\nu=1}^{d} \lambda_\nu}{\sum_{\nu=1}^{K} \lambda_\nu} \geq \text{THRSH},$$

where THRSH is a given threshold.

Notice that the resulting approximative covariance matrix is

$$\begin{aligned} & (u_1, \cdots, u_d, 0, \cdots, 0) \times \begin{pmatrix} \lambda_1 & 0 & \cdots & & & 0 \\ 0 & \lambda_2 & & & & 0 \\ \cdots & & \ddots & & & \vdots \\ 0 & & & \lambda_d & & \\ \vdots & & & & 0 & \\ & & & & & \ddots \\ 0 & & & & & 0 \end{pmatrix} \begin{pmatrix} u'_1 \\ u'_2 \\ \vdots \\ u'_d \\ 0 \\ \vdots \\ 0 \end{pmatrix} \\ = & \sum_{\nu=1}^{d} \lambda_\nu \, u_\nu \times u'_\nu = \text{Cov}\left(\sum_{\nu=1}^{d} u_\nu \sqrt{\lambda_\nu} Z_{\nu,j} \right), \end{aligned}$$

where the $Z_{\nu,j}$ are i.i.d. (scalar) standard Normal random variables, $1 \leq j \leq N-1$. Hence, isolating the d relevant eigenvalues suggest the approximation

$$\bar{F}_l(t_j, t_{j+1}) \approx \mathbb{E}\left[\bar{F}_l(t_j, t_{j+1})\right] + \sum_{\nu=1}^{d} u_\nu \sqrt{\lambda_\nu} Z_{\nu,j}.$$

From Proposition 2.1 and Equation (5), on the other hand, for the model with volatility structure $\sigma(\tau, x) =$

[3]By scaling the datapoints by $\frac{1}{\sqrt{t_{j+1}-t_j}}$, we ascertain that the vectors $\bar{F}_p(t_j, t_{j+1})$, $j = 1, 2, \ldots, N-1$, are also approximately i.i.d. when relying on non-equidistant data.

$(\sigma_1(\tau,x),\ldots,\sigma_d(\tau,x))$ as in (4), we obtain

$$\begin{aligned}\bar{F}_l(t_j,t_{j+1}) \;\; &\overset{d}{\approx} \;\; \mathbb{E}\left[\bar{F}_l(t_j,t_{j+1})\right] \\ &+\sum_{\nu=1}^{d} \left(O_\nu(\tau_i,x_i)\right)_{1\le i\le K} \frac{1}{\sqrt{t_{j+1}-t_j}} \int_0^{t_{j+1}-t_j} \exp\{M_\nu(t_{j+1}-t_j-s)\}N_\nu\, dW_s^{(\nu)},\end{aligned}$$

where $O_\nu(\tau_i,x_i)=\int_\tau^{\tau+l} C_\nu(x+s)\,\exp\left\{M_\nu\, s\right\}\,ds$ and $\frac{1}{\sqrt{t_{j+1}-t_j}}\int_0^{t_{j+1}-t_j}\exp\{M_\nu(t_{j+1}-t_j-s)\}N_\nu\,dW_s^{(\nu)}$ is an m_ν-dimensional vector of Normal random variables, $1\le\nu\le d$. While this vector may not necessarily consist of perfectly correlated random variables, they are all driven by the same (scalar) Brownian motion and will thus be strongly related. In particular, for a short time step $(t_{j+1}-t_j)$, the standard Euler scheme yields an approximation by a perfectly correlated random vector

$$\begin{aligned}&\frac{1}{\sqrt{t_{j+1}-t_j}}\int_0^{t_{j+1}-t_j}\exp\{M_\nu(t_{j+1}-t_j-s)\}N_\nu\,dW_s^{(\nu)} \\ &\quad= \frac{1}{\sqrt{t_{j+1}-t_j}}\exp\{M_\nu(t_{j+1}-t_j)\}N_\nu\,\underbrace{(W^{(\nu)}_{t_{j+1}-t_j}-W^{(\nu)}_0)}_{\overset{d}{=}\sqrt{t_{j+1}-t_j}Z_{\nu,j}} \\ &\quad= \underbrace{\exp\{M_\nu(t_{j+1}-t_j)\}N_\nu}_{\equiv\tilde{N}_{\nu,j}\in\mathbb{R}^{m_\nu}}\times Z_{\nu,j},\end{aligned} \tag{6}$$

where $Z_{\nu,j}$ is a standard Normal random variable and independent for different $\nu,j\in\{1,\ldots,d\}\times\{1,\ldots,N-1\}$. Hence,

$$\left(O_\nu(\tau_i,x_i)\right)_{1\le i\le K}\times\tilde{N}_\nu\approx u_\nu\sqrt{\lambda_\nu},\; 1\le\nu\le d, \tag{7}$$

and the task is now to find parameters $C_\nu(x),\; M_\nu$, and $\tilde{N}_\nu$, $1\le\nu\le d$, that fit the data optimally.

3.2 Data

We rely on two different data sources:

1. **Population Data**: Generation life tables complied based on the Lee and Carter (1992) methodology from England and Wales male mortality data as available from the *Human Mortality Database*.[4] More specifically, we have generation mortality tables compiled for thirty consecutive years (1978-2008) each using the mortality experience of the past thirty years with a two year lag (that is, for the 1978 table, data from 1947-1976 is employed; for the 1979 table, data from 1948-1977 is employed; etc.)

2. **Pensioners Data**: We rely on UK life tables and projections for pension annuities as published on the website of the Institute of Actuaries and the Faculty of Actuaries:[5]

   - Period table PA(90)m projected backward (until 1968) and forward using the projection applied

[4] Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (downloaded 12/04/2009).

[5] `http://www.actuaries.org.uk`.

in constructing this table;[6]

- Basic table PMA80 and projection as published in the "80" Series of mortality tables;
- Basic table PMA92 and projection as published in the "92" Series of mortality tables;
- Basic table PMA92 and the medium cohort projection published in 2002;
- Basic table PNMA00 from the "00" Series of mortality tables and the Lee-Carter projection based on data up to 2003 (part of the Library of mortality projections);
- Basic table PNMA00 from the "00" Series of mortality tables and the Lee-Carter projection based on data up to 2004 (part of the Library of mortality projections);
- Basic table PNMA00 from the "00" Series of mortality tables and the Lee-Carter projection based on data up to 2005 (part of the Library of mortality projections).

3.3 Analyses for Both Data Sets

For each set of population and pensioners data, we conduct two principle component analyses with $l = 1$ and $l = 5$, respectively, where we choose $\tilde{C}$ as large as possible. For the population data, mortality tables cover ages from 20 to 120, while for the pensioners data ages range from 20 to 112.[7] Therefore, we obtain $K = 4,656$ for the population data and $K = 3,916$ for the pensioners data.

For the population data, when $l = 1$, the six greatest eigenvalues are shown in Table 1(a). We find that λ_1 explains more than 98% of the total variation, while λ_2 and λ_3 explain 84% and 8.5% of the remaining variation, respectively. Since eigenvalue four only amounts to approximately one third of λ_3 and is comparable in size to the subsequent eigenvalues, it appears sufficient to choose $d = 3$. The eigenvectors for these three largest eigenvalues as functions of τ and x are plotted in Figure 1. With $l = 5$, the six greatest eigenvalues are shown in Table 1(b). The relative weights in explaining the variance are similar as in the case when $l = 1$, and the eigenvectors for λ_1, λ_2, and λ_3 are plotted in Figure 2.

(a) $l = 1$

	Value	Percentage
λ_1	1.282189301257414	98.33%
λ_2	0.018239303117445	1.40%
λ_3	0.001836674163609	0.14%
λ_4	0.000673258163755	
λ_5	0.000537909570434	
λ_6	0.000450538704874	

(b) $l = 5$

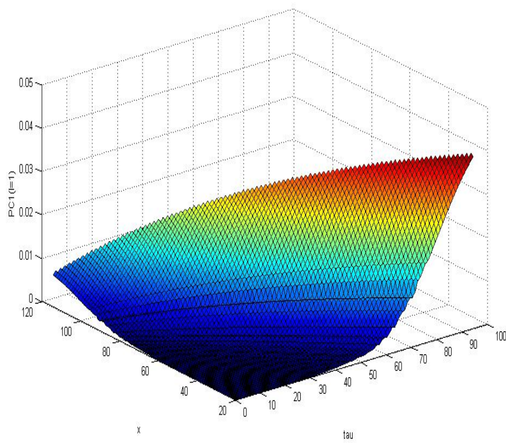
	Value	Percentage
λ_1	37.872071313908357	98.63%
λ_2	0.456871181188469	1.19%
λ_3	0.041244039023286	0.11%
λ_4	0.010970675366440	
λ_5	0.008854799063682	
λ_6	0.007036983099771	

Table 1: The six largest eigenvalues for the population data

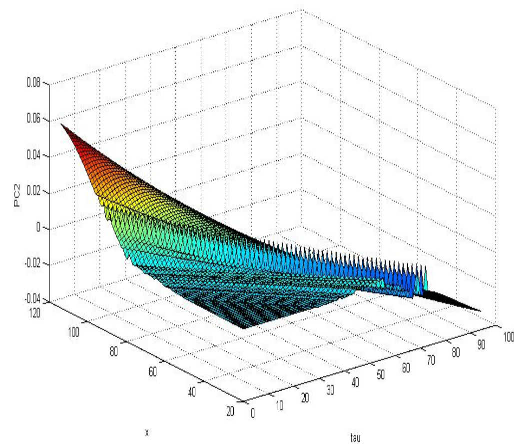
Similarly, for the pensioners data, when $l = 1$, the six largest eigenvalues are shown in Table 2(a) and the corresponding eigenvectors for λ_1, λ_2, and λ_3 are plotted in Figure 3. Eigenvalues for $l = 5$ are shown in Table 2(b) and the corresponding eigenvectors for λ_1, λ_2, and λ_3 are plotted in Figure 4.

[6]The PA(90)m table is a projected version of a period table from 1968. Hence, the best estimate mortality forecast in 1968 in form of a generation table is based on the period table for 1968 and the aforecited mortality projection.

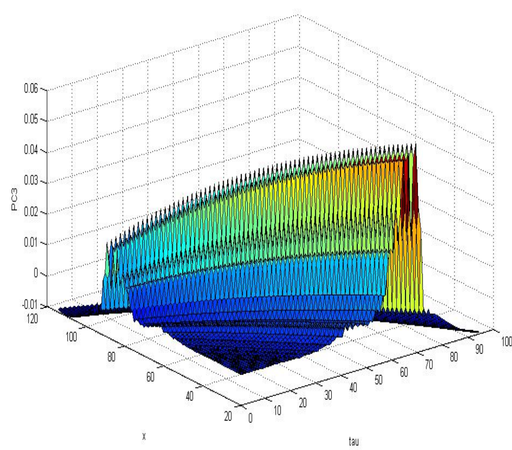
[7]This is the overlapping range of ages for all tables within each data set.



(a) PC1

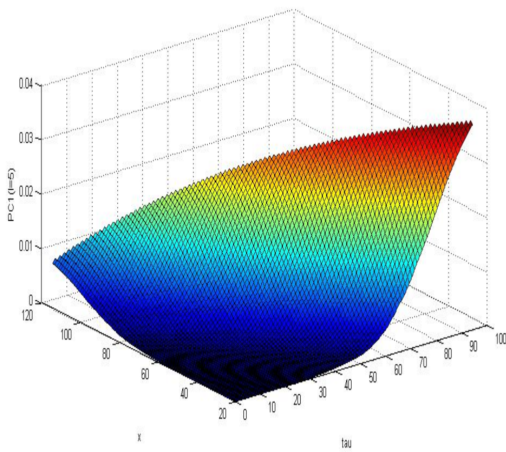


(b) PC2

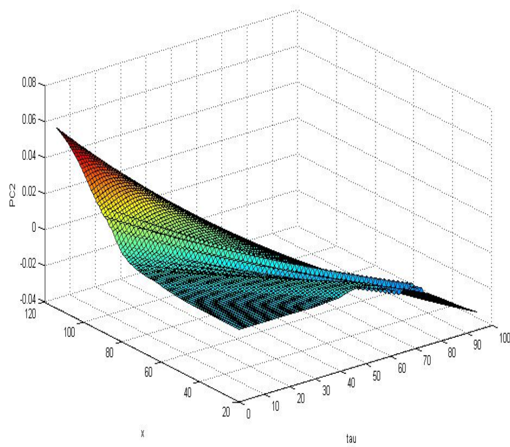


(c) PC3

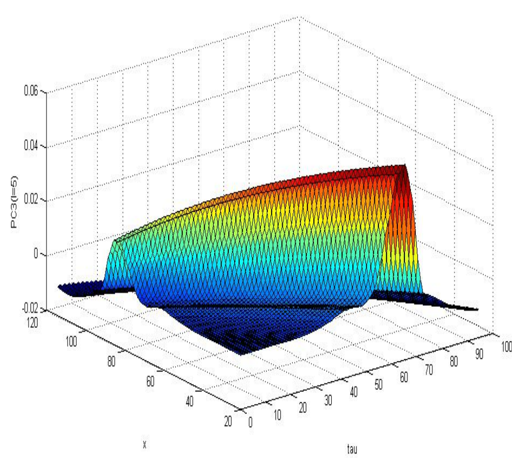
Figure 1: Eigenvectors for the population data, $l = 1$



(a) PC1

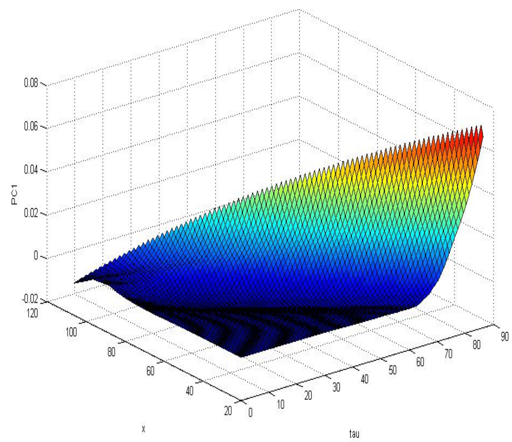


(b) PC2

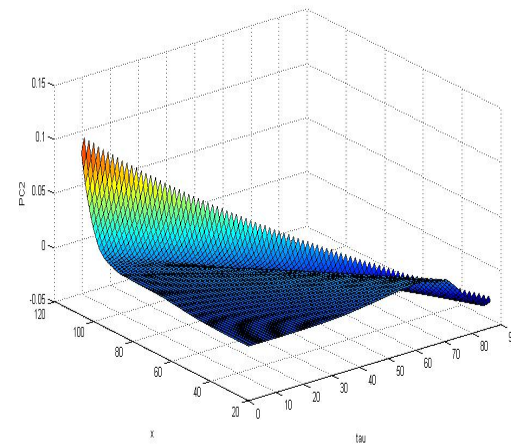


(c) PC3

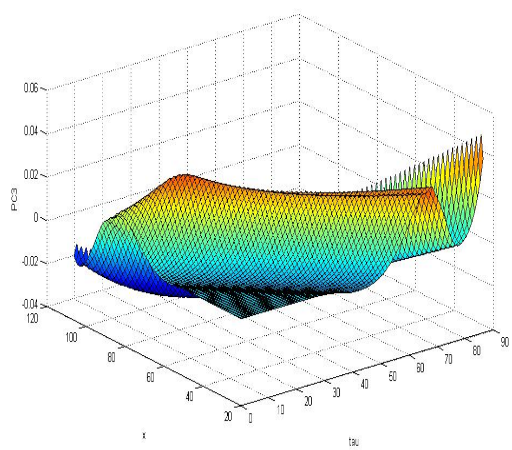
Figure 2: Eigenvectors for the population data, $l = 5$



(a) PC1

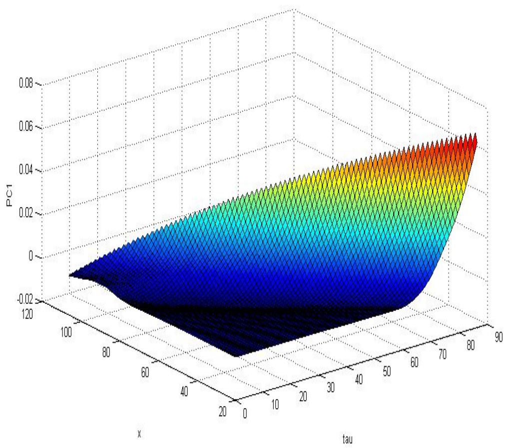


(b) PC2

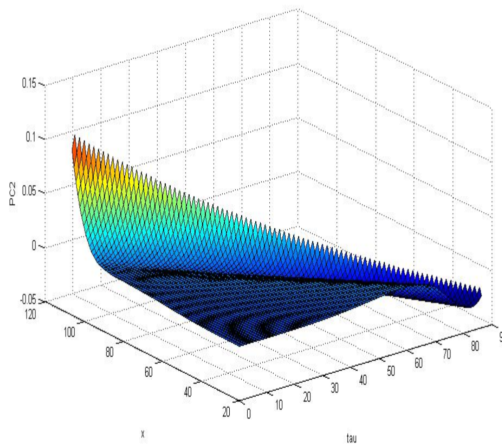


(c) PC3

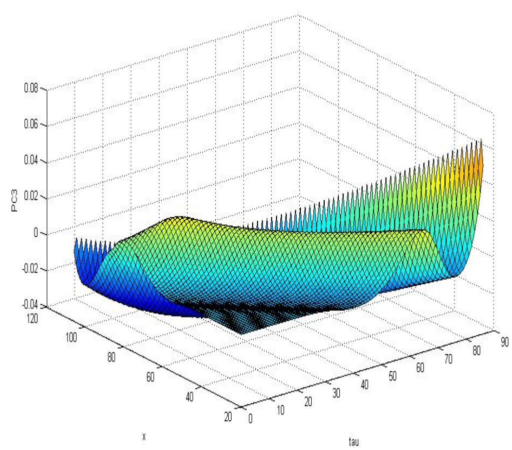
Figure 3: Eigenvectors for the pensioners data, $l = 1$



(a) PC1



(b) PC2



(c) PC3

Figure 4: Eigenvectors for the pensioners data, $l = 5$

(a) $l = 1$

	Value	Percentage
λ_1	0.750649573259028	94.61%
λ_2	0.029016064213762	3.66%
λ_3	0.010287991627984	1.30%
λ_4	0.002884535223968	
λ_5	0.000567832692818	
λ_6	0.000000000000000	

(b) $l = 5$

	Value	Percentage
λ_1	28.424093862578264	94.38%
λ_2	1.265600259549415	4.20%
λ_3	0.301118397288314	1.00%
λ_4	0.114438538260296	
λ_5	0.012348103607240	
λ_6	0.007036983099771	

Table 2: The six largest eigenvalues for pensioners data

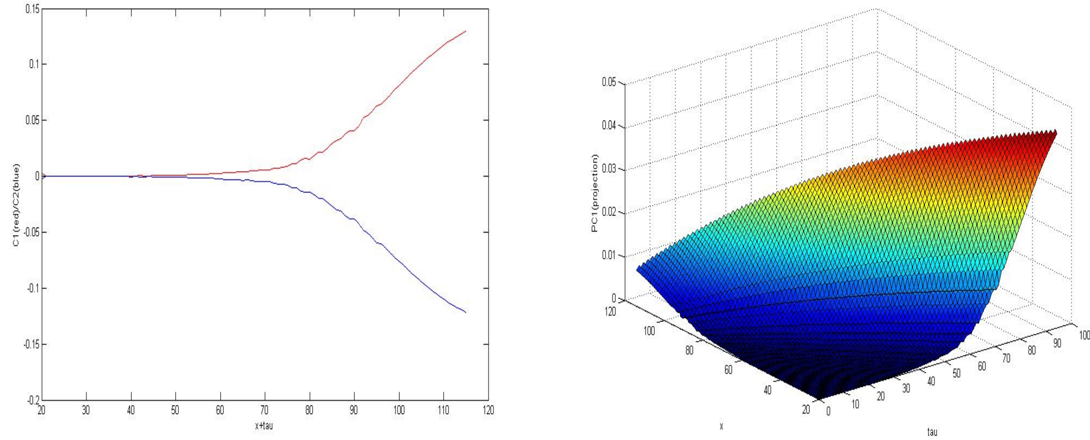
When comparing the eigenvectors between the two data sets, we notice that both exhibit very similar shapes at least for the leading two principle components. In particular, we notice that the structure of the first principle component is primarily governed by a simple age effect, which may be the key reason for its dominant role in explaining the variation of the mortality data. However, it is worth noting that the force of mortality for high ages in the far future appears to be more volatile than in the close future, which may be a consequence of the linear extrapolation prescribed by the Lee-Carter model. As for the second principle component, the data suggests an over time decreasing influence that even generates an inverse relationship for high ages in the close and the far future. Hence, this factor may be interpreted as the sensitivity of the fixed age effects in the Lee-Carter model. The third principle component also shows similarities between the two data sets. In particular, the mortality of the "old" is more affected than the mortality of the "oldest-old" (cf. Suzman et al. (1992)).

3.4 Regression Without a Functional Assumption on $C(x)$

Proposition 2.1 implies that we may examine each component $\sigma_\nu(\tau, x)$ and, hence, each eigenvector, separately, $\nu \in \{1, \ldots, d\}$. To simplify notation, we assume $(t_{j+1} - t_j) = \Delta$, although similar relationships hold for non-equidistant data. From Equations (6) and (7), we obtain for small l (we use $l = 1$ in what follows)

$$\begin{aligned} u_\nu \sqrt{\lambda_\nu} &\approx (O_\nu(\tau_i, x_i))_{1 \le i \le K} \times \tilde{N}_{\nu,j} = \left(\int_{\tau_i}^{\tau_i + l} C_\nu(x_i + s) \exp\{M_\nu \, s\} \, ds \right)_{1 \le i \le K} \times \tilde{N}_{\nu,j} \\ &\approx (C_\nu(x_i + \tau_i + l/2) \times \exp\{M_\nu(\tau_i + l/2)\} \cdot l)_{1 \le i \le K} \times \exp\{M_\nu \Delta\} N_\nu \\ &= (C_\nu(x_i + \tau_i + 1/2) \times \exp\{M_\nu(\tau_i + 1/2 + \Delta)\} \times N_\nu)_{1 \le i \le K} . \end{aligned} \tag{8}$$

Based on Equation (8), we are now able to estimate $C_\nu(x)$, M_ν, and N_ν via regression. Note, however, that in doing so, we are only utilizing the variance part of $\bar{F}_l(t_j, t_{j+1})$, with all information on $\mathbb{E}\left[\bar{F}_l(t_j, t_{j+1})\right]$ being neglected (cf. Equation (2)). Furthermore, the underlying approximations may lead to a slight bias in our estimation of the parameters. Therefore, we are not going to finalize the estimation of the parameter values here, but rather rely on the gained insights in order to determine a suitable functional assumption for $C(x)$, M, and N. The actual calibration of the model is based on Maximum Likelihood estimation described in Section 4.



(a) Estimated $C_i(x)$, $i = 1, 2$ (b) Projection

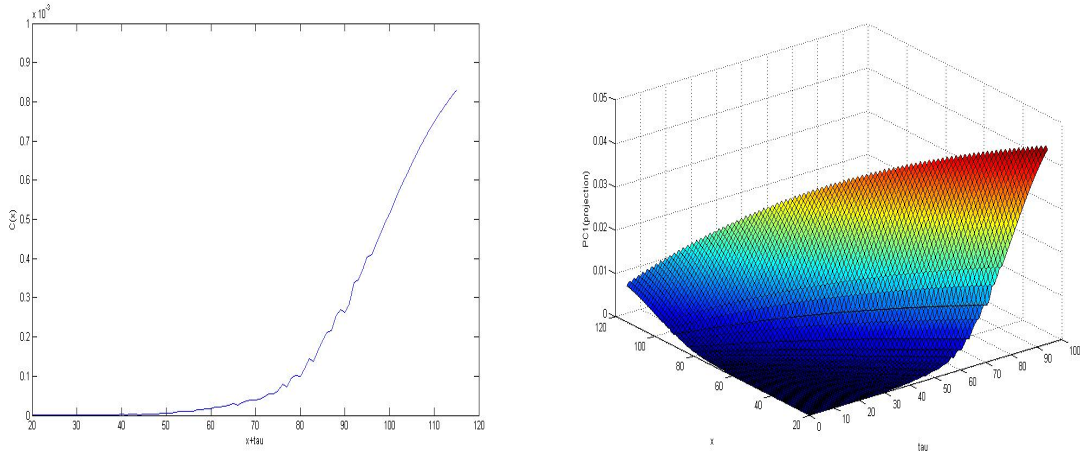
Figure 5: First principle component, method 1, population data

For M and N, two methods are used for each data set. The first method makes no specific assumption on M and N at all. The second method, on the other hand, mimics the observed shapes based on some well-known examples from interest rate modeling (see, e.g. Björk and Gombani (1999)), which significantly reduces the number of parameters to obtain accurate results. Since $m_\nu = 1$ does not allow for a pertinent fit for any of the principle components whereas $m_\nu = 3$ leads to identification problems, we set $m_\nu = 2$ for all $\mu \in \{1, 2, 3\}$. In what follows, we only show our results for the population data, although similar results were found for the pensioners data.

3.4.1 The First Principle Component

For the population data, with $m_1 = 2$, in total there are $96 \times 2 + 4 + 1 = 197$ number of parameters for the first method.[8] However, we encounter identification problems when considering four free parameters for M: For different starting values, we obtain different estimated parameters for $C(x)$ and M with the same sum of squared errors, which suggests that a reduced number of parameters for M is sufficient. Our calculations indicate that it is sufficient to assume $M = \text{diag}\,\{M_{11}, M_{22}\}$, which also simplifies the calculation of the matrix exponential. Based on this assumption, M_{11} and M_{22} are estimated as -0.0037 and -0.0104, respectively, with the estimated sum of squared errors SSE $= 5.26597e - 006$. $C_1(x)$, $C_2(x)$, and the projection based on the estimated parameters are plotted in Figure 5.

[8]The second component of N can be incorporated in C here.



(a) Estimated $f(x)$ (b) Projection

Figure 6: First principle component, method 2, population data

The second method imposes a specific assumption on the parameters. By setting

$$\begin{aligned} C(x+\tau) &= f(x+\tau) \times \begin{bmatrix} 0 & 1 \end{bmatrix}, \\ M &= \begin{bmatrix} -2b & -b^2 \\ 1 & 0 \end{bmatrix}, \text{ and} \\ N &= \begin{bmatrix} -ab-1 \\ a \end{bmatrix}, \end{aligned}$$

which is a slight modification of Example 6.2 in Björk and Gombani (1999), we obtain

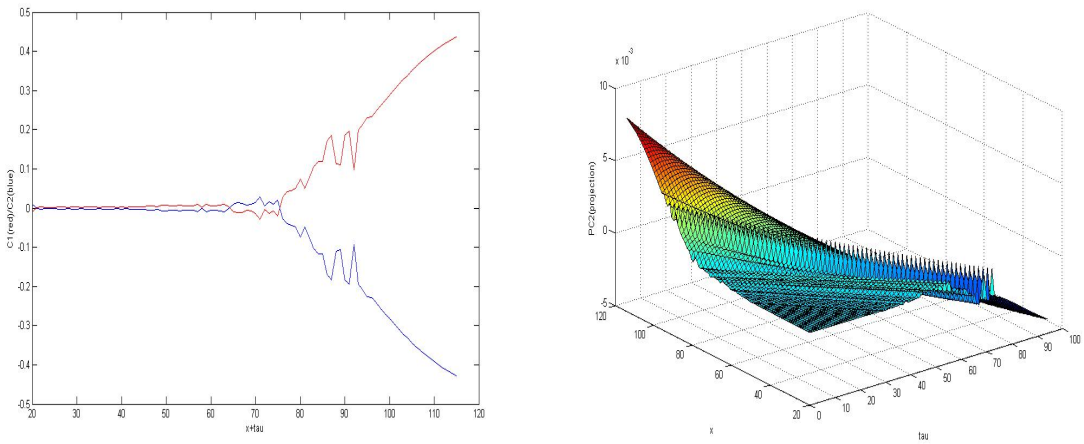
$$\begin{aligned} \sigma_1(\tau, x) &= C(x+\tau)\exp(M\tau)N \\ &= f(x+\tau)(a+\tau)\exp(-b\tau). \end{aligned}$$

Hence, this functional form is chosen to mimic the increasing, concave shape of the “diagonal curves” cd_{x_0}: $k \mapsto f(x_0)(a+k)\exp\{-bk\}$, $x_0 \in \mathbb{N}$.

With above assumptions, similar to Equation (6), we can approximate

$$\begin{aligned} u_1\sqrt{\lambda_1} &= \int_{\tau}^{\tau+l} \sigma_1\left(s+\Delta, x-\Delta\right) ds \approx \sigma_1\left(\tau+\Delta+\frac{l}{2}, x-\Delta\right) \\ &= f\left(x+\tau+\frac{l}{2}\right)\left(a+\tau+\Delta+\frac{l}{2}\right)\exp\left\{-b\left(\tau+\Delta+\frac{l}{2}\right)\right\}. \end{aligned}$$

Notice that even though $m = 2$, $f(x+\tau)$ is one-dimensional and there is only one single driving Wiener process. This reduces the number of parameters to 98. The estimation results are robust: a and b are estimated as 9.7690 and 0.0068, respectively, with an estimated sum of squared errors SSE $= 1.91691e-005$. $f(x)$ and the projection based on the estimated parameters are plotted in Figure 6.



(a) Estimated $C_i(x),\ i = 1, 2$ (b) Projection

Figure 7: Second principle component, method 1, population data

3.4.2 The Second Principle Component

Similarly to the first principle component, "too many" free parameters in M lead to problems with the identification of the model while not considerably improving the fit as measured by the SSE. Hence, again we choose a diagonal form yielding estimates $M_{11} = -0.0079$ and $M_{22} = -0.0075$, with an estimated sum of squared errors as $1.59817e - 007$. $C_1(x)$, $C_2(x)$, and the projection based on the estimated parameters are plotted in Figure 7.

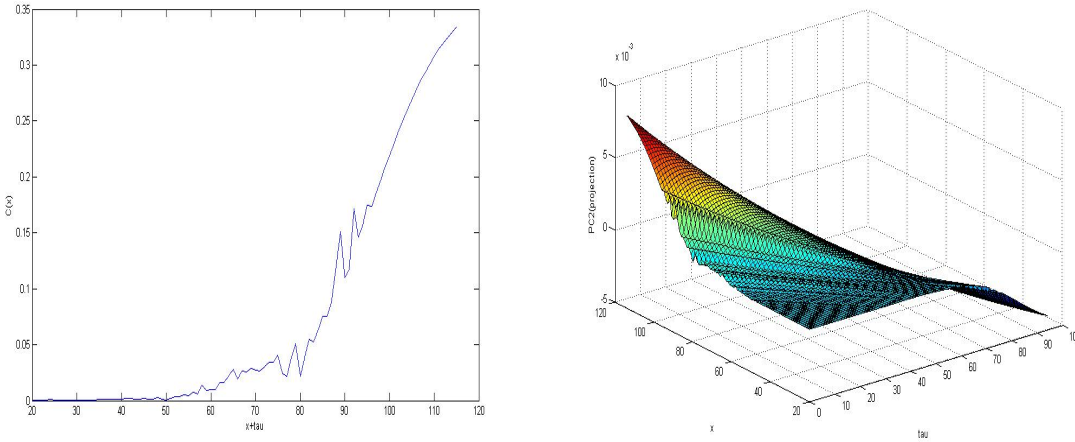
For method 2, we similarly set $M = \text{diag}\{M_{11}, M_{22}\}$ and $N = [1, N_2]$, since a difference of exponential function seems suitable. However, in contrast to method 1, we restrict the two components of C to be identical, leading to a volatility structure

$$\sigma_2(\tau, x) = C(x + \tau)\left[\exp\{M_{11}\tau\} + N_2 \cdot \exp\{M_{22}\tau\}\right].$$

In doing so, again the number of free parameters decreases considerably. The resulting estimated are $M_{11} = -0.0077$, $M_{22} = -0.0072$, and $N = -0.9744$ with an estimated SSE $= 0.000322727$. $C(x)$ and the projection based on the estimated parameters are plotted in Figure 8.

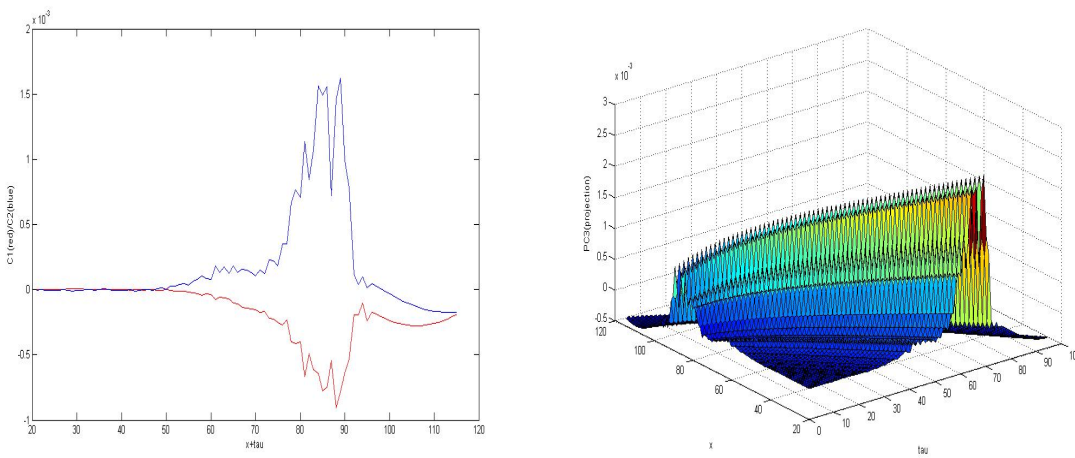
3.4.3 The Third Principle Component

Again relying on a diagonal form for M, method 1 yields $M_{11} = 0.0069$ and $M_{22} = -0.0608$ with estimated SSE $= 1.20027e - 006$. $C_1(x)$, $C_2(x)$, and the projection based on the estimated parameters are plotted in Figure 9. For method 2, the same approach as for the first principle component $\sigma_3(\tau, x) = f(x + \tau)(a + \tau)\exp\{-b\tau\}$ is taken due to the similar shape of the diagonal curves cd_{x_0}. The parameters a and b are estimated to be 17.2581 and 0.0081, respectively, with the estimated sum of squared error as $SSE = 1.51628e - 005$. $C(x)$ and the projection based on the estimated parameters are plotted in Figure 10.



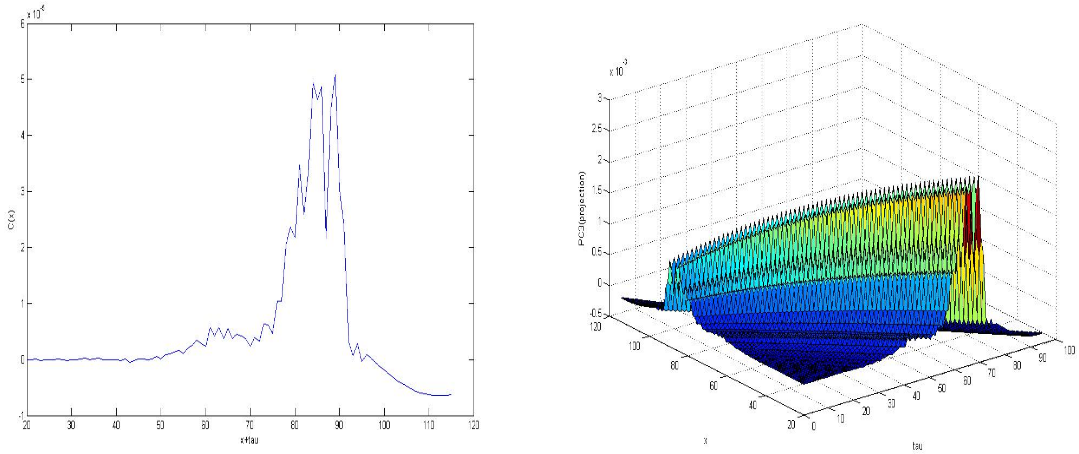
(a) Estimated $C(x)$ (b) Projection

Figure 8: Second principle component, method 2, population data



(a) Estimated $C_i(x),\ i = 1, 2$ (b) Projection

Figure 9: Third principle component, method 1, population data



(a) Estimated $C(x)$

(b) Projection

Figure 10: Third principle component, method 2, population data

3.5 Functional Assumptions for $C(x)$

In order to further reduce the number of parameters to make the calibration procedure tractable, as the next step we determine parametric assumptions for each of the functions $C_\nu(x)$, $1 \leq \nu \leq d$.

3.5.1 The First Principle Component

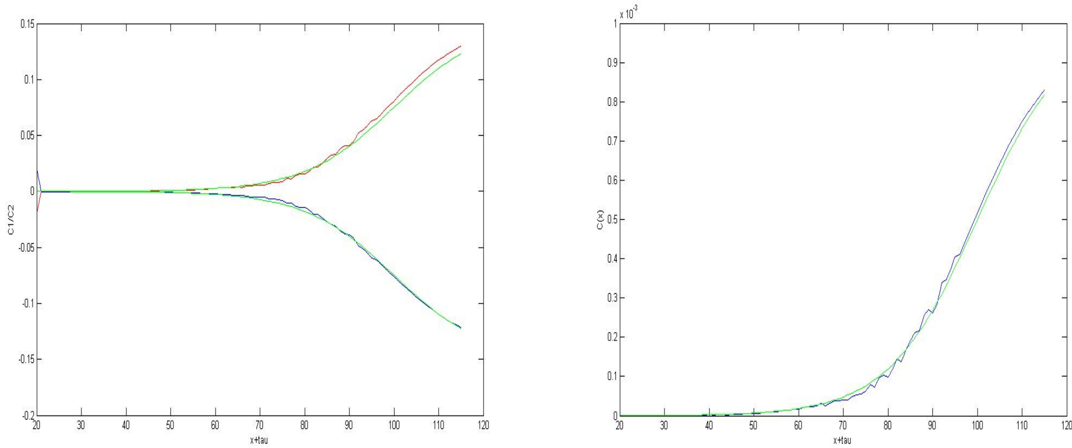
For the first principle component, the functional form depicted in Figure 5(a) suggests that a logistic-Gompertz function, $k \times \exp(cx + d)/(1 + \exp(cx + d))$, is a suitable choice. For both methods, one set of approximations with logistic-Gompertz functions (green lines) are shown in Figure 11.

3.5.2 The Second Principle Component

For the second principle component, similarly to the first one, we use a logistic-Gompertz form. Figure 12 shows the fitted curves with the corresponding set of parameter values.

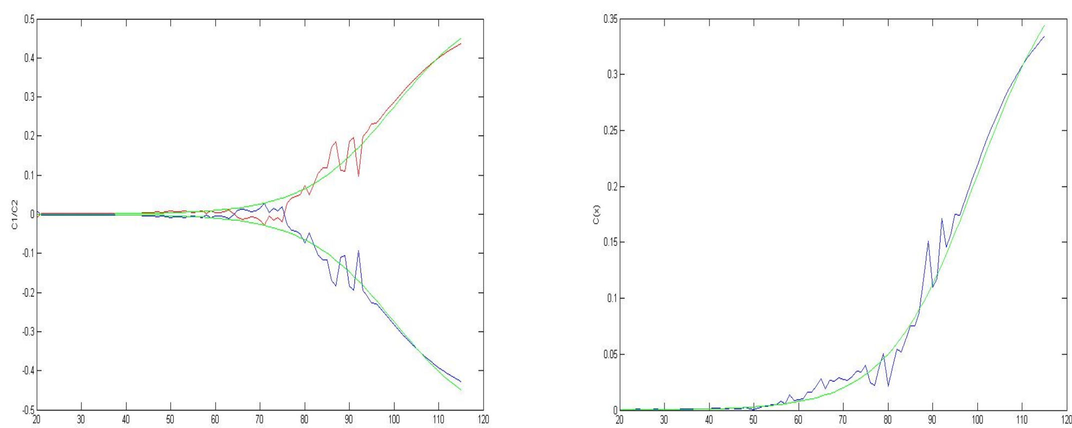
3.5.3 The Third Principle Component

For the third principle component, it is obvious that the logistic-Gompertz form alone cannot replicate the functional form of $C(x)$. However, when using the difference of two logistic-Gompertz functions, $k\frac{\exp(c_1 x+d_1)}{1+\exp(c_1 x+d_1)} - k\frac{\exp(c_2 x+d_2)}{1+\exp(c_2 x+d_2)}$, we can replicate the main features of the curves. Figure 13 shows the fitted curves with the corresponding set of parameter values.



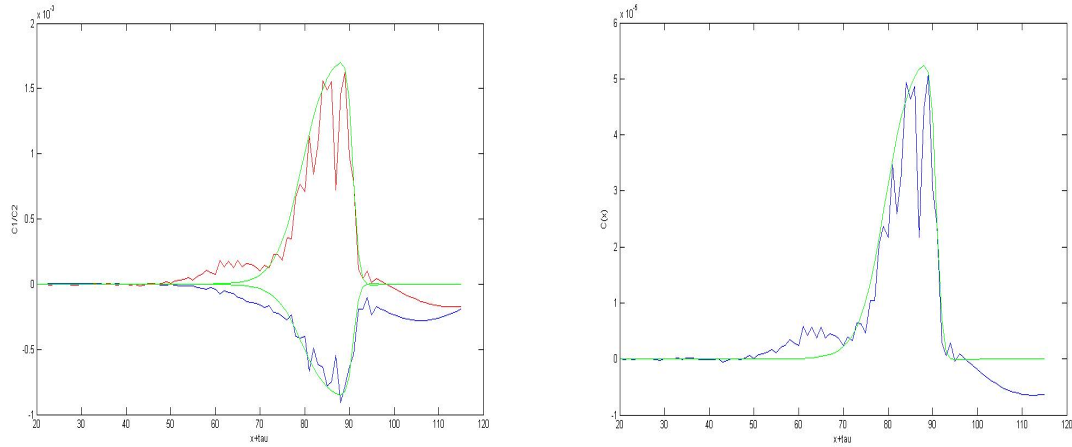
(a) Method 1, $c = 0.1$, $d = -10$, and $k = \pm 0.15$. (b) Method 2, $c = 0.1$, $d = -10$, and $k = 0.001$.

Figure 11: Fit $C_1(x)$ with Logistic-Gompertz function



(a) Method 1, $c = 0.1$, $d = -10$, and $k = \pm 0.55$. (b) Method 2, $c = 0.1$, $d = -10$, and $k = 0.42$.

Figure 12: Fit $C_2(x)$ with Logistic-Gompertz function



(a) Method 1, $c_1 = 0.3422$, $d_1 = -27.14$, $c_2 = 1.6209$, $d_2 = -147.32$, while $k = 1.8e-3$ or $k = 0.9e-3$.

(b) Method 2, $c_1 = 0.3422$, $d_1 = -27.14$, $c_2 = 1.6209$, $d_2 = -147.32$, and $k = 5.55e-5$.

Figure 13: Fit $C_3(x)$ with Logistic-Gompertz function

4 Calibration

4.1 Approach

Similarly to Bauer et al. (2008a) and Bauer (2009), we rely on the quantities $F_l\left(t_j, t_{j+1}, (\tau, x)\right)$, $(\tau, x) \in \tilde{\mathcal{C}}$ introduced above for our consideration. We can express Equation (5) as:

$$\begin{aligned} F_l\left(t_j, t_{j+1}, (\tau, x)\right) &= -\log\left\{ \frac{{}_{\tau+l}p_x(t_{j+1}, t_{j+1}+\tau+l)}{{}_{\tau}p_x(t_{j+1}, t_{j+1}+\tau)} \Big/ \frac{{}_{\tau+l+t_{j+1}-t_j}p_{x-t_{j+1}+t_j}(t_j, t_{j+1}+\tau+l)}{{}_{\tau+t_{j+1}-t_j}p_{x-t_{j+1}+t_j}(t_j, t_{j+1}+\tau)} \right\} \\ &= \int_{t_j}^{t_{j+1}} \int_0^l \alpha(v+\tau+t_{j+1}-s, x-t_{j+1}+s)\, dv\, ds \\ &\quad + \int_{t_j}^{t_{j+1}} \int_0^l \sigma(v+\tau+t_{j+1}-s, x-t_{j+1}+s)\, dv\, dW_s. \end{aligned}$$

Therefore, with Equation (5), $F_l\left(t_j, t_{j+1}, (\tau, x)\right)$ is Normal distributed with expected value

$$\mathbb{E}\left[F_l\left(t_j, t_{j+1}, (\tau, x)\right)\right] = \int_{t_j}^{t_{j+1}} \int_0^l \sigma(v+\tau+t_{j+1}-s, x-t_{j+1}-s) \int_0^{v+\tau+t_{j+1}-s} \sigma'(u, x-t_{j+1}-s)\, du\, dv\, ds$$

and covariance structure

$$\begin{aligned} &\mathrm{Cov}\left[F_l\left(t_j, t_{j+1}, (\tau_1, x_1)\right), F_l\left(t_k, t_{k+1}, (\tau_2, x_2)\right)\right] = \\ &\delta_{jk} \times \int_{t_j}^{t_{j+1}} \int_0^l \sigma(v+\tau_1+t_{j+1}-s, x_1-t_{j+1}+s)\, dv \int_0^l \sigma'(v+\tau_2+t_{j+1}-s, x_2-t_{j+1}+s)\, dv\, ds \end{aligned}$$

by a simple application of Itô's product formula. In particular, for $t_{j+1} - t_j = \Delta$ is constant,[9] the vectors $(F_L(t_j, t_{j+1}, (\tau, x)_i))_{1 \le i \le K}$ are i.i.d. Normal with expected values

$$\bar{\mu} = \left(\int_0^\Delta \left\{ \frac{1}{2} \int_{\tau+\Delta-s}^{l+\tau+\Delta-s} \sigma(u, x-\Delta+s)\, du \int_{\tau+\Delta-s}^{l+\tau+\Delta-s} \sigma'(u, x-\Delta+s)\, du \right.\right.$$
$$\left.\left. + \int_{\tau+\Delta-s}^{l+\tau+\Delta-s} \sigma(u, x-\Delta+s)\, du \int_0^{\tau+\Delta-s} \sigma'(u, x-\Delta+s)\, du \right\} ds \right)_{(\tau,x)\in\tilde{C}}$$

and covariance matrix $\Sigma = (\Sigma_{ij})_{1 \le i,j \le K}$, where

$$\Sigma_{ij} = \int_0^\Delta \int_0^l \sigma(v + \tau_i + \Delta - s, x_i - \Delta + s)\, dv \int_0^l \sigma'(v + \tau_j + \Delta - s, x_j - \Delta + s)\, dv\, ds.$$

These ideas were applied in Bauer et al. (2008a) and Bauer (2009) for their maximum-likelihood calibration algorithms. However, as pointed out in their contributions, their approach only allows for the consideration of a (very) small number of term/age combinations (τ_i, x_i) (i.e. a small value of K) since (non-systematic) deviations are not admissible. In order to overcome this problem, similarly to measurement equations in *state-spare models*, we allow for non-systematic deviations in the "observed" vectors $\bar{F}^{obs}(t_j, t_{j+1})$, from our model-endogenous vectors $\bar{F}^{mod}(t_j, t_{j+1})$. More specifically, we assume

$$\bar{F}^{obs}(t_j, t_{j+1}) = \bar{F}^{mod}(t_j, t_{j+1}) + \underline{\epsilon}_j, \tag{9}$$

where $\underline{\epsilon}_j$ are mutually independent and independent of $\bar{F}^{mod}(t_j, t_{j+1})$, $\underline{\epsilon}_j \sim N(0, \sigma_e \cdot I)$, $j = 1, \ldots, N-1$, and I is the $(K \times K)$ identity matrix. Intuitively, the $\underline{\epsilon}_j$ pick up the variation not accounted for by the considered eigenvalues. Thus, we obtain

$$\bar{F}^{obs}(t_j, t_{j+1}) \sim N(\bar{\mu}, \tilde{\Sigma}), \tag{10}$$

where $\tilde{\Sigma} = \Sigma + \sigma_e \cdot I$ and the log-likelihood function is of the form

$$\begin{aligned} & L(\bar{F}^{obs}(t_1, t_2), \ldots, \bar{F}^{obs}(t_{N-1}, t_N); C, M, N, \sigma_e) \\ &= \log \left\{ \prod_{j=2}^{N} \frac{1}{\sqrt{(2\pi)^K \det(\tilde{\Sigma})}} \exp\left\{ -\frac{1}{2} (\bar{F}^{obs}(t_{j-1}, t_j) - \bar{\mu}) \tilde{\Sigma}^{-1} (\bar{F}^{obs}(t_{j-1}, t_j) - \bar{\mu})' \right\} \right\} \\ &= \frac{1}{2} \left[\underbrace{\sum_{j=2}^{N} - \log\left\{\det(\tilde{\Sigma})\right\} - (\bar{F}^{obs}(t_{j-1}, t_j) - \bar{\mu}) \tilde{\Sigma}^{-1} (\bar{F}^{obs}(t_{j-1}, t_j) - \bar{\mu})'}_{\equiv \tilde{L}} \right] + \text{const.} \end{aligned} \tag{11}$$

To determine Maximum-likelihood estimates for our model parameters, it now suffices to determine the maximum values for $\tilde{L}$, which can be carried out numerically.

[9] We consider this special case for notational convenience. Analogous results apply for non-equidistant data.

4.2 Results

Maximum-likelihood estimations are conducted for both methods discussed in Section 3.4. For each method, we use two different values of σ_e, $\sigma_{e,1} = 0.0001$ and $\sigma_{e,2} = 0.00001$, to test the sensitivity of the estimated parameters to σ_e. Due to numerical problems when inverting the covariance matrix for a "too large" collection of age/term combinations, we limit the algorithm to a subset of the data. More precisely, the considered subset consists of ages $x \in \{20, 25, \dots, 65\}$ and terms $\tau \in \{0, 5, \dots, 35\}$.

For method 1, there are in total 24 free parameters. The starting values together with the resulting estimates under $\sigma_{e,1}$ and $\sigma_{e,2}$ are shown in Table 3.

Parameters	Starting Value	Estimation: $\sigma_{e,1}$	Estimation: $\sigma_{e,2}$
c_{11}	0.1	0.1000	0.1024
d_{11}	-10	-9.8787	-10.0959
c_{12}	0.1	0.1021	0.0998
d_{12}	-10	-10.1537	-10.2455
$M_{11;1}$	-0.0037	-0.0036	-0.0034
$M_{22;1}$	-0.0104	-0.0102	-0.0118
$N_{1;1}$	0.15	0.1548	0.1620
$N_{2;1}$	-0.15	-0.1412	-0.1452
c_{21}	0.1	0.1021	0.1041
d_{21}	-10	-10.2235	-10.6675
c_{22}	0.1	0.0996	0.0942
d_{22}	-10	-10.2491	-9.6441
$M_{11;2}$	-0.0079	-0.0080	-0.0086
$M_{22;2}$	-0.0075	-0.0076	-0.0082
$N_{1;2}$	0.55	0.5489	0.5827
$N_{2;2}$	-0.55	-0.5386	-0.6241
c_{31}	0.3422	0.3336	0.3422
d_{31}	-27.14	-27.9204	-26.1938
c_{32}	1.6209	1.6040	1.1563
d_{32}	-147.32	-148.5822	-131.5073
$M_{11;3}$	0.0069	0.0069	0.0070
$M_{22;3}$	-0.0608	-0.0605	-0.0592
$N_{1;3}$	$0.001.8$	0.0018	0.0017
$N_{2;3}$	0.0009	0.0009	0.0010
$\tilde{L}$	$\sigma_{e,1}$: 29569.7 $\sigma_{e,2}$: 35786.0	29641.3	36473.4

Table 3: MLE: method 1

For method 2, there is a total of 18 parameters to estimate. The starting values together with estimations under $\sigma_{e,1}$ and $\sigma_{e,2}$ are shown in Table 4. Comparing the adjusted log-likelihood $\tilde{L}$ between the two methods, we find that the they achieve almost the same values $\tilde{L}$. Therefore, method 2 should be preferred since it has a smaller number of parameters. Furthermore, by comparing $\tilde{L}$ with different σ_e, it is higher for smaller σ_e, while the estimation results are generally similar between the two different values. Therefore, in what follows, we use the estimated parameters from method 2 with $\sigma_{e,2}$.

Parameters	Starting Value	Estimation: $\sigma_{e,1}$	Estimation: $\sigma_{e,2}$
k_1	0.001	0.001	0.0013
c_1	0.1	0.1002	0.0988
d_1	-10	-9.9652	-9.9224
a_1	9.7690	13.8602	13.7281
b_1	0.0068	0.0060	0.0056
k_2	0.42	0.4308	0.4162
c_2	0.1	0.1611	0.1365
d_2	-10	-3.5210	-10.9224
M_{11}	-0.0077	-0.0056	-0.0153
M_{22}	-0.0072	-0.0069	-0.0142
N_2	-0.9744	-1.0001	-0.8995
k_3	$5.55e-5$	0.0001	0.0001
c_{31}	0.3422	0.3205	0.1723
d_{31}	-27.14	-24.8102	-6.2499
c_{32}	1.6209	1.4615	1.5163
d_{32}	-147.32	-139.5818	-116.5989
a_3	17.2581	15.1627	7.6889
b_3	0.0081	0.0066	0.0001
$\tilde{L}$	$\sigma_{e,1}$: 29638.5 $\sigma_{e,2}$: 36410.9	29641.8	36473.4

Table 4: MLE: method 2

5 Application I: Economic Capital in Internal Models

As indicated in the introduction, one potential reason for the sluggish development of the longevity-linked capital market may be the inability of insurers to assess their capital relief when hedging part of their longevity exposure. In this regard, due to their compatibility with classical actuarial methods, forward mortality models may present an instrumental tool in incorporating mortality into insurer's EC calculations. To demonstrate this, we first introduce a mathematical framework similar to that from Bauer et al. (2010b) for determining economic capital in a one-year mark-to-market approach as required by the new Solvency II regulation (see Börger (2010) for similar considerations). Subsequently, we demonstrate the effects of incorporating mortality based on example calculations for a stylized insurance company only offering "traditional" (life) policies.

5.1 Model Framework

Assume the uncertainty with respect to a life insurer's future profits arises from the uncertain development of a number of financial/economic and demographic factors, which are modeled with the help of the d-dimensional, sufficiently regular Markov process $Y = (Y_t)_{t\geq 0} = (Y_t^{(1)}, \ldots, Y_t^{(d)})_{t\geq 0}$, the so-called *state process*. More specifically, we assume that the prices of all risky assets in the market can be expressed in terms of Y_t, and that there exists a locally risk-free asset (bank account) $B = (B_t)_{t\geq 0}$ with $B_t = \exp\{\int_0^t r_s\, ds\}$, where $r_t = r(Y_t)$ is the instantaneous risk-free rate at time t. Similarly, we assume $\mu_t(x) =$

$\mu_t(0,x) = \mu(x,Y_t)$ for the (instantaneous) force of mortality at time t. In this market environment, we take for granted the existence of a risk-neutral probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ under which payment streams can be valued via their expected discounted values with respect to the numéraire B.[10]

Based on this economic/demographic environment, we assume that there exists a cash flow projection model, i.e. there exist functionals $f_1,\ldots,f_T$ that derive the cash flows at time t from the state process up to time t, where T is the time horizon. Hence, the random variable $X_t = f_t(Y_s,\, s\in[0,t])$ corresponds to the benefits paid minus the premiums earned at time t, $t=1,\ldots,T$, and the value of the liabilities can be determined via[11]

$$V_0 = \mathbb{E}^{\mathbb{Q}}\left[\sum_{k=0}^{T}\frac{1}{B_k}\,X_k\right].$$

Thus, the *Available Capital* at time zero can be derived from V_0 and the value of assets $A_0 = A(Y_0)$ as

$$AC_0 = A_0 - V_0.$$

However, within the one-year mark-to-market approach for calculating economic capital, it is not sufficient to determine the available capital at time zero, but it is also necessary to assess the available capital at time 1, $AC_1 = A_1 - V_1$, where

$$V_1 = X_1 + B_1\,\mathbb{E}^{\mathbb{Q}}\left[\sum_{k=2}^{T}\frac{1}{B_k}X_k\,\middle|\,Y_s, 0\le s\le 1\right].$$

More specifically, one defines the one-year loss as the $\mathcal{F}_1$-measurable random variable

$$L = AC_0 - \frac{AC_1}{1+s(0,1)},$$

where $s(0,1)$ is the one year risk-free interest rate at time zero. The economical capital is then defined with the help of a *monetary risk measure* ρ: $L^2(\Omega,\mathcal{F},\mathbb{P})\to\mathbb{R}$ as $\rho(L)$. For instance, if the economic capital is defined based on the Value-at-Risk such as the *Solvency Capital Requirement* (SCR) within the Solvency II framework, we have

$$\text{SCR} = \text{VaR}_\alpha(L) = \arg\min_x\{\mathbb{P}(L>x)\le 1-\alpha\},$$

where α is a given threshold (99.5% within Solvency II).

5.2 A Stylized Life Insurance Company

Consider now a (stylized) newly founded life insurance company only selling traditional business to male individuals, who form a representative sample of the England and Wales population. More specifically, let us assume that the company's portfolio of policies consists of $n_{x,i}^{\text{term}}$ i-year term-life policies for individuals aged x with face value B_{term}, $n_{x,i}^{\text{end}}$ i-year endowment policies for individuals aged x with face value B_{end},

[10] According to the *Fundamental Theorem of Asset Pricing*, this assumption is essentially equivalent with the absence of arbitrage in the market.

[11] Note that we implicitly adopted the *direct method* for determining insurance liabilities, see e.g. Girard (2002).

and n_x^{ann} single premium life annuities with an annual benefit of B_{ann} paid in arrears, $x \in \mathcal{X}, i \in \mathcal{I}$. Furthermore, we assume that the (for term and endowment policies annual) premium is calculated by the Equivalence principle profits and disregarding expenses using the concurrent yield curve and the concurrent best-estimate generation table.[12] Hence, the insurer's available capital at time zero AC_0 amounts to its equity capital E. For the available capital at time 1, on the other hand, we have $AC_1 = A_1 - V_1$, where

$$\begin{aligned} A_1 &= \left(E + B_{\text{ann}} \sum_{x\in\mathcal{X}} a_x(0) n_x^{\text{ann}} + B_{\text{term}} \sum_{x\in\mathcal{X}, i\in\mathcal{I}} \frac{A_{\overset{1}{x}:\overline{i}|}(0)}{\ddot{a}_{x:\overline{i}|}(0)} n_{x,i}^{\text{term}} + B_{\text{end}} \sum_{x\in\mathcal{X}, i\in\mathcal{I}} \frac{A_{x:\overline{i}|}(0)}{\ddot{a}_{x:\overline{i}|}(0)} n_{x,i}^{\text{end}} \right) \times R_1, \\ V_1 &= B_{\text{ann}} \sum_{x\in\mathcal{X}} \ddot{a}_{x+1}(1) \left(n_x^{\text{ann}} - \mathfrak{D}_x^{\text{ann}}(0,1) \right) + B_{\text{term}} \sum_{x\in\mathcal{X}, i\in\mathcal{I}} \mathfrak{D}_{x,i}^{\text{term}}(0,1) + B_{\text{end}} \sum_{x\in\mathcal{X}, i\in\mathcal{I}} \mathfrak{D}_{x,i}^{\text{end}}(0,1) \\ &\quad + B_{\text{term}} \sum_{x\in\mathcal{X}, i\in\mathcal{I}} \left[A_{\overset{1}{x+1}:\overline{i-1}|}(1) - \frac{A_{\overset{1}{x}:\overline{i}|}(0)}{\ddot{a}_{x:\overline{i}|}(0)} \ddot{a}_{x+1:\overline{i-1}|}(1) \right] \times \left(n_{x,i}^{\text{term}} - \mathfrak{D}_{x,i}^{\text{term}}(0,1) \right) \\ &\quad + B_{\text{end}} \sum_{x\in\mathcal{X}, i\in\mathcal{I}} \left[A_{x+1:\overline{i-1}|}(1) - \frac{A_{x:\overline{i}|}(0)}{\ddot{a}_{x:\overline{i}|}(0)} \ddot{a}_{x+1:\overline{i-1}|}(1) \right] \times \left(n_{x,i}^{\text{end}} - \mathfrak{D}_{x,i}^{\text{end}}(0,1) \right). \end{aligned}$$

Here, R_1 is the return on the insurer's asset portfolio, $\mathfrak{D}_{x,i}^{\text{con}}(0,1)$ is the number of deaths in the cohort of x-year old insured with policies of term i of type con $\in \{\text{ann}, \text{term}, \text{end}\}$, and $\ddot{a}_x(t)$, $A_{x:\overline{i}|}(t)$, etc. are the values of the contracts corresponding to the actuarial symbols calculated at time t based on the yield curve and the generation table at time t. For instance,

$$\ddot{a}_x(t) = \sum_{k=0}^{\infty} {}_k p_x(t; t+k)\, p(t,k),$$

where $p(t,\tau)$ denotes the price of a zero-coupon bond at time t with time to maturity τ.

The economic capital of this insurer can then be determined as

$$EC = \rho\left(E - (A_1 - V_1) p(0,1)\right),$$

where $\rho(\cdot)$ is a monetary risk measure as described above.

5.3 Implementation and Results

5.3.1 Setup

We assume that our UK insurer only invests in 1, 3, 5, and 10-year government bonds as well as an equity index $S = (S_t)_{t\geq 0}$ (FTSE) at equal proportions. For the evolution of these assets, we assume a generalized

[12] Here, we implicitly assume that the insurer is risk-neutral with respect to mortality risk, i.e. that his valuation measure is the product measure of the risk-neutral measure for financial and the physical measure for (independent) biometric events. This is without much loss of generality since, under the assumption of a deterministic market price of systematic mortality risk, a risk-adjusted generation table can be derived from the best estimate generation table via a deterministic transformation (see Bauer et al. (2010b) and Bauer et al. (2010c) for details).

Black-Scholes model with stochastic interest rates, that is, under $\mathbb{P}$

$$\begin{aligned} dS_t &= S_t(\mu dt + \rho\sigma_A dW_t + \sqrt{1-\rho^2}\sigma_A dZ_t),\ S_0 > 0, \\ dr_t &= \kappa(\gamma - r_t)dt + \sigma_r dW_t,\ r_0 > 0, \end{aligned}$$

where μ, σ_A, κ, γ, $\sigma_r > 0$ and $\rho \in [0,1]$. Moreover, we assume that the market price of interest rate risk is constant and denote it by λ, i.e. we have $\mu = r_t$ and $\gamma = \gamma - \frac{\lambda\sigma_r}{\kappa}$ for the evolution under the risk-neutral measure $\mathbb{Q}$.

We estimate the parameters based on UK data from June 1998 to June 2008 using a Kalman filter. More precisely, we use monthly data for the FTSE 100 index,[13] and government bonds with maturities of 3 months, 1 year, 3 years, 5 years, and 10 years.[14] The resulting parameter estimates are displayed in Table 5.

Parameter	μ	σ_A	ρ	κ	γ	σ_r	λ	r_0 (06/2008)
	0.0571	0.1407	0.3145	0.2826	0.0472	0.0075	0.0275	0.04844

Table 5: Estimated Parameters

Hence, based on realizations of the asset process and the interest rate at time 1, S_1 and r_1, respectively, R_1 can be determined as

$$R_1 = 20\%\frac{S_1}{S_0} + 20\%\frac{1}{p(0,1)} + 20\%\frac{p(1,2)}{p(0,3)} + 20\%\frac{p(1,4)}{p(0,5)} + 20\%\frac{p(1,9)}{p(0,10)}.$$

The parameters for our company are displayed in Table 6.

	x	i	$n_{x,i}^{\text{term/end/ann}}$	Benefit $B_{\text{term/end/ann}}$
Term Life	30	20	250	$100,000$
	35	15	250	
	40	10	250	
	45	5	250	
Endowment	40	20	500	$50,000$
	45	15	500	
	50	10	500	
Annuities	60	(45)	250	$18,000$
	70	(35)	250	
$E = 2,000,000$				

Table 6: Parameters for the company

5.3.2 Simulation

S_1 and r_1 are simulated from a joint normal distribution via a standard procedure (see e.g. Zaglauer and Bauer (2008)) using $25,000$ simulations. Furthemore, $p(t,\tau)$ can be written as $\exp(-A(\tau)r_t + C(\tau))$,

[13]Downloaded on 09/01/2010 from Yahoo! Finance UK&Ireland. `http://uk.finance.yahoo.com`.

[14]Downloaded on 09/01/2010 from the Bank of England's website, `http://bankofengland.uk/statistics/yieldcurve/archive.htm`.

where

$$A(\tau) = \frac{1}{\kappa}(1-e^{-\kappa\tau}),$$
$$C(\tau) = (\gamma - \frac{\lambda\sigma_r}{\kappa} - \frac{\sigma_r^2}{2\kappa^2})(\frac{1-\exp(-\kappa\tau)}{\kappa} - \tau) - \frac{\sigma_r^2(1-\exp(-\kappa\tau))^2}{4\kappa^3}.$$

Therefore, for r_0 and each simulated r_1, we can calculate $p(0,\tau)$ and $p(1,\tau)$, and finally R_1.

With respect to mortality risk, two different approaches are taken. First, we assume that mortality evolves deterministically, that is, we use the latest (2008) Lee-Carter generation life table for England and Wales males and r_0, r_1 to determine values of $\ddot{a}_x(k)_{k=0,1}$, $A_{\overset{1}{x}:\overline{i}|}(k)_{k=0,1}$, etc.;[15] the realized deaths $\mathfrak{D}_{x,i}^{\text{term}}$,... are simulated from Binomial distribution with also $25,000$ simulations, $x \in \mathcal{X}, i \in \mathcal{I}$. Based on the simulated A_1 and V_1, we can then calculate the desired quantile of $E-(A_1-V_1)p(0,1)$ from the resulting empirical distribution function.

For the second approach, we use the same method to determine values of actuarial symbols at time 0 and the same simulation method for $\mathfrak{D}_{x,i}$, $x \in \mathcal{X}, i \in \mathcal{I}$. However, for $\ddot{a}_x(1)$, $A_{\overset{1}{x}:\overline{i}|}(1)$, etc., the values are calculated from simulations of future generation tables as described in Section 3, with parameters calibrated from Section 4. More specifically, we have

$${}_{\tau}p_x(1;1+\tau) = \frac{{}_{\tau+1}p_{x-1}(0;1+\tau)}{{}_1p_{x-1}(0;1)} \times \exp\left\{-\int_0^1\int_0^\tau \alpha(v+1-s,x-1+s)dvds - \int_0^\tau C(x+v)e^{Mv}dv \cdot Z_1\right\}.$$

Therefore, in addition to the uncertainty of future financial/economic market, the second approach also takes into account the uncertainty of future mortality evolution. With already calibrated parameter values for $C(x)$, M, and N, we generate $25,000$ simulations for Z_1. Again, we calculate the desired quantile of $E-(A_1-V_1)p(0,1)$ based on the resulting empirical distribution function.

5.3.3 Preliminary Results

For the first approach, i.e. when not taking systematic mortality risk into account, the empirical cumulative distribution function of the portfolio loss L is displayed in Figure 14. In particular, for the 90th percentile, that is the Value-at-Risk at the 90% level, we obtain $\text{VaR}_{90\%} = 4.7206e+006$.

When taking mortality risk into account, the situation changes drastically. The corresponding empirical distribution function is displayed in Figure 15, and in this case we obtain a 90th percentile of $\text{VaR}_{90\%} = 8.9171e+006$. Hence, the economic capital as measured by the Value-at-Risk at the 90% level almost doubles when taking mortality risk into account.

While this result has to be considered with care as we neither take expenses nor profits into account and since we assume that there is no mortality risk premium associated with systematic mortality risk, the result demonstrates that mortality risk is a significant risk factor and should be taken into account in risk-based capital calculations.

[15] A maximum age of $\omega = 105$ is used instead of ∞.

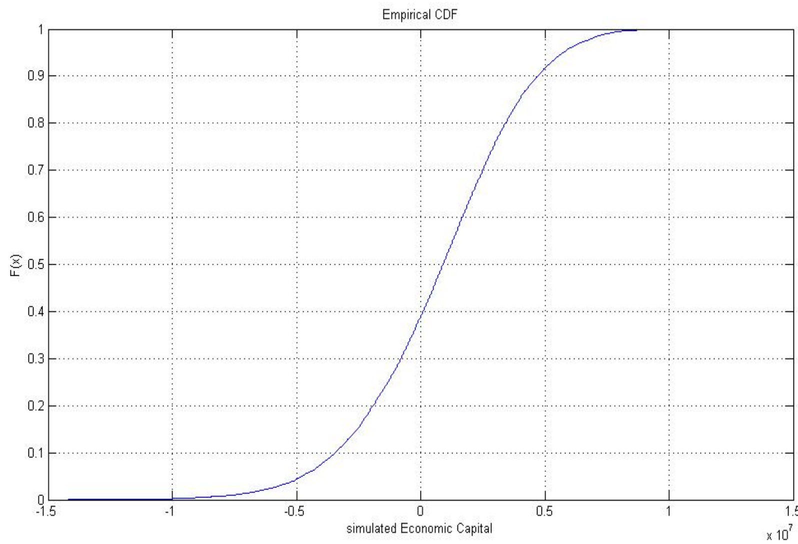


Figure 14: Empirical CDF of L, Approach 1

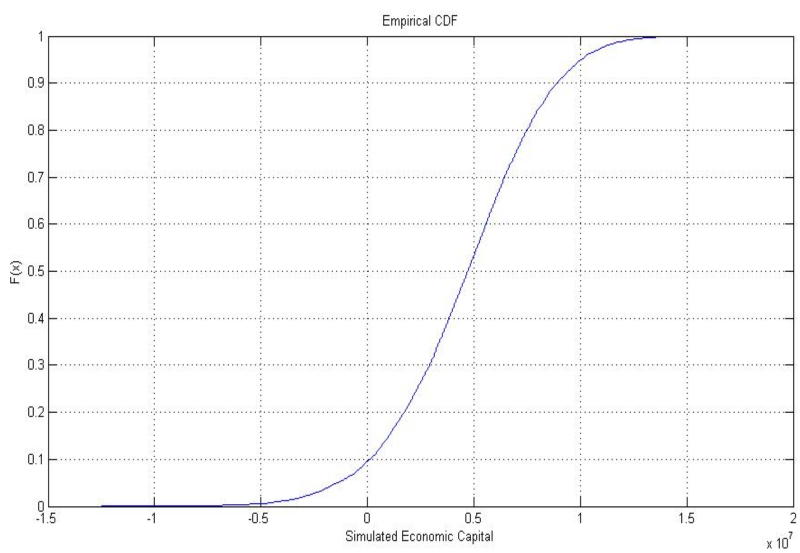


Figure 15: Empirical CDF of L, Approach 2

6 Application II: Valuation of Annuitization Options

As demonstrated in the previous section, under the assumption of independence between the demographic and economic environment, *traditional* life insurance and pension products such as whole, term, or endowment insurances or life annuities may be evaluated directly using the "concurrent" (time t) mortality surface $\mu_t(\tau, x)$. For more complex life insurance and annuity contracts such as mortality-contingent options, however, the stochasticity has to be taken into account. Common types of longevity-contingent options are so-called *Guaranteed Annuity Options* (GAOs) within traditional or participating life insurance contracts and *Guaranteed Minimum Income Benefits* (GMIBs) within *Variable Annuities*. A GAO provides the policyholder with the option to choose, at retirement, between a lump-sum payment or a life-long immediate annuities which is calculated based on a guaranteed annuity rate. In contrast, a GMIB gives the insured the possibility to annuitize a guaranteed amount at a pre-specified rate.

6.1 Valuation of Guaranteed Annuity Options

Several authors have studied the problem of evaluating GAOs without taking mortality risk into account (see Boyle and Hardy (2003), Pelsser (2003), and references therein). In contrast, Milevsky and Promislow (2001) also take the stochasticity of mortality rates into account. They provide a discrete- and continuous-time pricing framework for simple annuitization options. Similarly, Ballotta and Haberman (2006) present a pricing approach for GAOs which accounts for both interest rate and mortality risk. Since the solution for the price is not in closed form, they rely on Monte Carlo simulations for the derivation of their numerical results. Using affine processes for modeling the financial market as well as the (spot) mortality evolution, Biffis and Millossovich (2006) also present a pricing framework for GAOs. Under some structural assumptions, they derive analytical solutions up to the computation of Fourier transforms and/or numerical integrals for various contract designs.

Following the ideas of Cairns et al. (2006) in their *annuity market model*, we consider the valuation of simple GAOs in the forward mortality modeling framework. Specifically, we regard a contract providing payoffs of the following form at time T if the policyholder is alive:

$$V_T^{\mathrm{GAO}} = \max\left\{1, g_{x_0,T}^{\mathrm{GAO}}\, \ddot{a}_{x_0+T}(T)\right\}. \tag{12}$$

Here, $g_{x_0,T}^{\mathrm{GAO}}$ is the guaranteed annuity rate under the GAO for annuitization at T contracted at time 0 for a, then, x_0-year old. Thus, the contract may be interpreted as a T-year pure endowment policy with the additional option to annuitize at the fixed rate $g_{x_0,T}^{\mathrm{GAO}}$ at maturity.[16]

We have

$$V_T^{\mathrm{GAO}} = 1 + \underbrace{\ddot{a}_{x_0+T}(T)\ \max\left\{g_{x_0,T}^{\mathrm{GAO}} - \frac{1}{\ddot{a}_{x_0+T}(T)}, 0\right\}}_{=:C_T^{\mathrm{GAO}}}$$

[16] Of course, this pure endowment part may be combined with a term insurance contract to obtain an endowment contract including a GAO.

and thus,[17]

$$V_0^{\text{GAO}} = \mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{\{\Upsilon_{x_0}>T\}} e^{-\int_0^T r_s\,ds} V_T^{\text{GAO}}\right] = p(0,T)\,{}_T p_{x_0}(0,T) + \underbrace{\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^T \mu_s(0,x_0+s)\,ds} e^{-\int_0^T r_s\,ds} C_T^{\text{GAO}}\right]}_{=C_0^{\text{GAO}}}.$$

Now define $X = (X_t)_{t\geq 0}$ via

$$X_t = \sum_{k=T}^{\infty} p(t,k-t)\,{}_k p_{x_0}(t,k).$$

Then, for the value of the GAO

$$\begin{aligned} C_0^{\text{GAO}} &= \mathbb{E}^{\mathbb{Q}}\left[\frac{e^{-\int_0^T \mu_s(0,x_0+s)\,ds}}{B_T} \overbrace{\sum_{k=T}^{\infty} \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_T^k r_s\,ds}\mathbf{1}_{\{\Upsilon_{x_0}>k\}}\middle|\mathcal{F}_T \vee \{\Upsilon_{x_0} > T\}\right]}^{\ddot{a}_{x_0+T}(T)} \left(g_{x_0,T}^{\text{GAO}} - \frac{e^{-\int_0^T \mu_s(0,x_0+s)\,ds}}{e^{-\int_0^T \mu_s(0,x_0+s)\,ds}\,\ddot{a}_{x_0+T}(T)}\right)^+\right] \\ &= \mathbb{E}^{\mathbb{Q}}\left[\frac{X_T}{B_T}\left(g_{x_0,T}^{\text{GAO}} - \frac{e^{-\int_0^T \mu_s(0,x_0+s)\,ds}}{X_T}\right)^+\right] = X_0 \mathbb{E}^{\mathbb{Q}_X}\left[\left(g_{x_0,T}^{\text{GAO}} - \frac{{}_T p_{x_0}(T,T)}{X_T}\right)^+\right], \end{aligned} \tag{13}$$

where $\mathbb{Q}_X$ is the EMM associated with the numéraire process (X_t). In particular, the price process of a security with payoff ${}_T p_{x_0}(T,T)$ at time T discounted by X will be a martingale under $\mathbb{Q}_X$. Hence, in order to evaluate the expectation in (13), we need to solely assess the volatility term of $\left(\frac{p(t,T-t)\,{}_T p_{x_0}(t,T)}{X_t}\right)$ and, thus, the volatility term of (X_t). From Proposition 2.1 from Bauer et al. (2010a) and Proposition 20.5 in Björk (1999), we obtain under $\mathbb{Q}$

$$dX_t = \sum_{k=T}^{\infty} p(t,k-t)\,{}_k p_{x_0}(t,k)\left(S(t,k,x_0) + v(t,k-t)\right) dW_t,$$

where $v(t,k)$ is the time t volatility of a zero coupon bond maturing at time $t+k$ and

$$S(t,k,x_0) = -\int_0^{k-t} \sigma(s,x_0+t)\,ds.$$

Therefore, the volatility of X is given by

$$\begin{aligned} &\sum_{k=T}^{\infty} p(t,k-t)\,{}_k p_{x_0}(t,k)\left(S(t,k,x_0) + v(t,k-t)\right) \\ = \quad & X_t \sum_{k=T}^{\infty} \underbrace{\frac{p(t,k-t)\,{}_k p_{x_0}(t,k)}{\sum_{k=T}^{\infty} p(t,k-t)\,{}_k p_{x_0}(t,k)}}_{=:w_t(k)} \left(S(t,k,x_0) + v(t,k-t)\right) \end{aligned}$$

[17] Again, similarly to the previous section, for simplicity and without much loss of generality, we assume that the insurer is risk-neutral with respect to mortality risk.

By an application of Itô's Lemma, the volatility of $\left(\frac{p(t,T-t)\,{}_T p_{x_0}(t,T)}{X_t}\right)$ is then given by

$$
\begin{aligned}
&\left(\frac{p(t,T-t)\,{}_T p_{x_0}(t,T)}{X_t}\right)\left((S(t,T,x_0)+v(t,T-t))-\sum_{k=T}^{\infty} w_t(k)(S(t,k,x_0)+v(t,k-t))\right)\\
=\;&\left(\frac{p(t,T-t)\,{}_T p_{x_0}(t,T)}{X_t}\right)\underbrace{\left(\sum_{k=T}^{\infty} w_t(k)\left((S(t,T,x_0)+v(t,T-t))-(S(t,k,x_0)+v(t,k-t))\right)\right)}_{\gamma(t,T,x_0)}.
\end{aligned}
$$

If now $\gamma(\cdot,\cdot,\cdot)$ were deterministic, we would be able to derive an analytical expression for (13) via a Black-type formula (this is basically the idea of Cairns et al. (2006), who propose to directly modeling the *forward annuity rate* $\left(\frac{p(t,T-t)\,{}_T p_{x_0}(t,T)}{X_t}\right)$). However, in our more general framework, the weights $w_t(k)$ are in fact stochastic. Nevertheless, such an approach may be understood as an *approximation*, and Pelsser (2003) points out that for the deterministic mortality case, one may infer $\gamma(t,T,x_0)$ "*by "freezing" the stochastic weights at their current values*". This is similar to the approximation used by Brace et al. (1997) in Theorem 3.2. In particular, for a deterministic choice of $v(\cdot,\cdot)$, in our problem we may fix

$$
\gamma(t,T,x_0)\approx\sum_{k=T}^{\infty} w_0(k)\left((S(t,T,x_0)+v(t,T-t))-(S(t,k,x_0)+v(t,k-t))\right), \tag{14}
$$

which then yields:

$$
\begin{aligned}
C_0^{\mathrm{GAO}} &= X_0\left(g_{x_0,T}^{\mathrm{GAO}}\Phi\left(-d_2^{\mathrm{GAO}}\right)-\frac{p(0,T)\,{}_T p_{x_0}(0,T)}{X_0}\Phi\left(-d_1^{\mathrm{GAO}}\right)\right), \qquad (15)\\
\text{where } d_1^{\mathrm{GAO}} &= \frac{\log\left\{\frac{p(0,T)\,{}_T p_{x_0}(0,T)}{X_0\, g_{x_0,T}^{\mathrm{GAO}}}\right\}+\frac{1}{2}\sigma_{\mathrm{GAO}}^2}{\sigma_{\mathrm{GAO}}},\\
d_2^{\mathrm{GAO}} &= d_1^{\mathrm{GAO}}-\sigma_{\mathrm{GAO}},\\
\sigma_{\mathrm{GAO}}^2 &= \int_0^T \|\gamma(u,T,x_0)\|^2\,du,
\end{aligned}
$$

and $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution.

However, here we only allow for constant payoffs. Particularly, this means that the GAO considered here provides a constant annuity guarantee, and thus, in this special case it coincides with a GMIB as introduced above. In practice, GAOs are often attached to unit-linked or participating policies, and GMIBs are usually granted within VA contracts; in these cases, the lump-sum payment will be stochastic, and hence, formula (15) does not apply. Ballotta and Haberman (2003) provide pricing formulas for GAOs within unit-linked policies for a deterministic mortality evolution by applying the ideas of Jamshidian (1989). However, their results cannot be easily carried over to the stochastic mortality environment unless interest rates and the mortality evolution are driven by the same one-dimensional Brownian motion W, which seems very unrealistic. Thus, one may have to resort to numerical methods for the valuation.

6.2 Valuation of Garanteed Minimum Income Benefits

For example, for the time zero value of a VA contract including a GMIB with guaranteed annuity payment $g_{x_0,T}^{\text{GMIB}}$, for an insured of age x_0 at 0 when annuitizing at time T and no death benefit guarantee, i.e. in the case of death only the current account value is paid out, we have:

$$\begin{aligned} V_0^{\text{GMIB}} &= \mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{\{\Upsilon_{x_0}>T\}} e^{-\int_0^T r_s\,ds} \max\left\{g_{x_0,T}^{\text{GMIB}}\,\ddot{a}_{x_0+T}(T), A_T\right\}\right] \\ &\quad + \sum_{k=0}^{T-1} \mathbb{E}^{\mathbb{Q}}\left[\mathbf{1}_{\{\Upsilon_{x_0}\in[k,k+1)\}} e^{-\int_0^{k+1} r_s\,ds} A_{k+1}\,dt\right] \\ &= \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^T r_s+\mu_s(0,x_0+s)\,ds} \max\left\{g_{x_0,T}^{\text{GMIB}} \sum_{k=T}^{\infty} p(T,k-T)_k p_{x_0+T}(T,T+k), A_T\right\}\right] \\ &\quad + A_0\,e^{-\phi T}\left(1 - {}_T p_{x_0}(0,T)\right), \end{aligned}$$

where A_t denotes the insured's account value at time t, which – in a Black-Scholes framework under $\mathbb{Q}$ – is assumed to evolve according to the stochastic differential equation (cf. Bauer et al. (2008b))

$$dA_t = A_t\left((r_t - \phi)\,dt + \sigma_A dW_t\right), A_0 > 0.$$

Here ϕ denotes the continuously deducted option fee. Now, methods similar to those proposed in Bauer et al. (2008b), where a deterministic evolution of mortality is assumed, can be employed to determine the value. The latter formula illustrates why Cairns (2007) considers forward mortality models to be "*ideal for pricing contracts with embedded options*": Within other mortality models, $\ddot{a}_{x_0+T}(T)$ is typically not available, but each "*simulation*" will require a "*further bundle of simulations from time T to evaluate forward survival probabilities. In contrast, forward survival probabilities are a standard part of the output* [of forward force models] *at time T*".

7 Conclusion

Due to their tractability and their compatibility with classical actuarial theory, *Gaussian Forward Mortality Factor Models* present a convenient way of introducing systematic mortality risk in actuarial practice. This should not only improve the accuracy of common actuarial calculations, but should help to provide a more coherent "risk picture" of a life insurance company's operations.

The current paper addresses the specification, the calibration, and the application of such models. More precisely, we use data on the age/term structure of mortality to determine suitable specifications based on a Principal Component Analysis. The resulting models are then be calibrated based on a Maximum Likelihood estimation. The consider example applications are practically relevant and illustrate the merits of this model class.

References

Angelini, F., Herzel, S., 2005. Consistent Calibration of HJM Models to CAP Implied Volatilities. Journal of Futures Markets 25: 1093-1120.

Ballotta, L., Haberman, S., 2003. Valuation of guaranteed annuity conversion options. Insurance: Mathematics and Economics, 33: 87-108.

Ballotta, L., Haberman, S., 2006. The fair valuation problem of guaranteed annuity options: The stochastic mortality environment case. Insurance: Mathematics and Economics, 38: 195-214.

Bauer, D., 2009. Stochastic Mortality Modeling and Securitization of Mortality Risk. ifa-Verlag, Ulm, Germany.

Bauer, D., Benth, F.E., Kiesel, R., 2010a. Modeling the Forward Surface of Mortality. Working Paper, Georgia State University.

Bauer, D., Bergmann, D., Reuss, A., 2010b. On the Calculation of the Solvency Capital Requirement based on Internal Models. Working Paper, Georgia State University.

Bauer, D., Börger, M., Ruß, J., 2010c. On the Pricing of Longevity-Linked Securities. Insurance: Mathematics and Economics, 46: 139-149.

Bauer, D., Börger, M., Ruß, J., Zwiesler, H.J., 2008a. The Volatility of Mortality. Asia-Pacific Journal of Risk and Insurance, 3: 172-199.

Bauer, D., Kling, A., Russ, J., 2008b. A Universal Pricing Framework for Guaranteed Minimum Benefits in Variable Annuities. Astin Bulletin, 38: 621-651.

Björk, T., 1999. Arbitrage Theory in Continuous Time. Oxford University Press, Oxford.

Boyle, P., Hardy, M., 2003. Guaranteed Annuity Options. Astin Bulletin, 33: 125-152.

Brace, A., Gatarek, D., Musiela, M., 1997. The Market Model of Interest Rate Dynamics. Mathematical Finance, 7: 127-147.

Biffis, E., Millossovich, P., 2006. The Fair Value of Guaranteed Annuity Options. Scandinavian Actuarial Journal, 2006: 23-41.

Björk, T., Gombani, A., 1999. Minimal realizations of interest rates. Finance and Stochastics, 3: 413-432. Working paper version: `http://swopec.hhs.se/hastef/papers/hastef0182.pdf`.

Börger, M., 2010. Deterministic Shock vs. Stochastic Value-at-Risk – An Analysis of the Solvency II Standard Model Approach to Longevity Risk. Forthcoming in the DGVFM Blaetter.

Cairns, A.J., 2007. A Multifactor Generalisation of the Olivier-Smith Model for Stochastic Mortality. Working Paper, Heriot-Watt University.

Cairns, A.J., Blake, D., Dowd, K., 2006. Pricing Death: Frameworks for the Valuation and Securitization of Mortality Risk. Astin Bulletin, 36: 79-120.

Dawson, P., Dowd, K., Cairns, A.J., Blake, D., 2010. Survivor Derivatives: A Consistent Pricing Framework. The Journal of Risk and Insurance, 77: 579-596.

Girard, L., 2002. An Approach to Fair Valuation of Insurance Liabilities Using the Firm's Cost of Capital. North American Actuarial Journal, 6: 18-41.

Jamshidian, F., 1989. An exact bond option formula. Journal of Finance, 44: 205-209.

Lee, R.D., Carter, L.R., 1992. Modeling and Forecasting U. S. Mortality. Journal of the American Statistical Association, 87: 659-675.

Litterman, R., Scheinkman, J.A., 1991. Common Factors Affecting Bond Returns. Journal of Fixed Income, 1: 62-74.

Milevsky, M.A., Promislow, S.D., 2001. Mortality Derivatives and the Option to Annuitize. Insurance: Mathematics and Economics, 29: 299-318.

Njenga, C.N., Sherris, M., 2009. Longevity Risk and the Econometric Analysis of Mortality Trends and Volatility. Working Paper, University of New South Wales.

Pelsser, A., 2003. Pricing and Hedging Guaranteed Annuity Options via Static Option Replication. Insurance: Mathematics and Economics, 33: 283-296.

Rebonato, R., 1998. Interest Rate Option Models, Second Edition. Wiley, Chichester.

Schrager, D., 2006. Affine stochastic mortality. Insurance: Mathematics and Economics, 38: 81-97.

Suzman, R.M., Willis, D.P., Manton, K.G., 1992. The Oldest Old. Oxford University Press, New York.

Zaglauer, K., Bauer, D., 2008. Risk-Neutral Valuation of Participating Life Insurance Contracts in a Stochastic Interest Rate Environment. Insurance: Mathematics and Economics, 43: 29-40.