Affine Stochastic Mortality

David F. Schrager[1,2]

This version:

October 22, 2004

Abstract

We propose a new model for stochastic mortality. The model is based on the literature on affine term structure models. It satisfies three important requirements for application in practice: analytical tractibility, clear interpretation of the factors and compatibility with financial option pricing models. We test the model fit using data on Dutch mortality rates. Furthermore we discuss the specification of a market price of mortality risk and apply the model to the pricing of a Guaranteed Annuity Option and the calculation of required Economic Capital for mortality risk.

Keywords: Stochastic Mortality, Affine Models, Mortality Laws, Longevity Risk, Market Price of Mortality Risk, Mortality Options

[1]Nationale Nederlanden, Market and Operational Risk Management, PO Box 796, 3000 AT Rotterdam, The Netherlands, Tel. +31 - (0)10 513 1262, Fax. +31 - (0)10 513 0120

[2]University of Amsterdam, Dept. of Quantitative Economics, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands, Tel. +31 - (0)20 525 4125, e-mail: d.f.schrager@uva.nl

1 Introduction

In this paper we propose a new model for mortality intensity. Our approach is based on the observation that if the mortality intensity is an affine function of a number of latent factors the survival and death probabilities are known in closed form. Most of the subsequent results are based on the literature on affine term structure models (cf. Duffie and Kan, 1996 and Duffie, Pan and Singleton, 2000) and the credit risk literature based on the subfiltration approach, see Lando (1998), Elliott, Jeanblanc and Yor (2000) and Jamshidian (2004). Our contribution consists of the application of these ideas to the modeling of the evolution of mortality rates over time. We provide in the need for a model for mortality forces which can be combined consistently with continuous time models known from the derivative pricing literature. We introduce a new setup for some well known functional dependences between age and mortality intensity (i.e. the Thiele and Makeham mortality laws). The three main advantages of the model are a rich analytical structure (inherited from the affine setup), clear interpretation of the latent factors and the aforementioned consistency with derivative pricing models. In contrast to previous work, for example Dahl (2004) and Milevsky and Promislow (2001), we consider the mortality intensity for all ages simultaneously. As opposed to e.g. Lee and Carter (1992) and Lee (2000) we do not explicitly focus on the time series properties of mortality (although the model is extremely well suited for estimation to empirical data), rather we have a pricing and risk management application in mind. Four types of mortality risk are usually distinguished: trend (i.e. longevity), level (portfolio vs. population), volatility (discrepancies between trend / level and observed mortality) and catastrophe. Our model captures three of these types and the risks are directly quantified by parameters estimates. Furthermore, we show, using historical Dutch mortality rates, the proposed Thiele and Makeham functional forms fit the data sufficiently well.

Assuming independence of financial and mortality risk one can easily combine our model with, for instance, a term structure model. One could then easily price several well studied options embedded in insurance contracts under stochastic mortality. Examples of such contracts are Guaranteed Annuity Options (GAOs) or Rate of Return Guarantees, see among others, Brennan and Schwartz (1976), Bacinello and Ortu (1993), Aase and Persson (1998), Boyle and Hardy (2003), Pelsser (2003) or Schrager and Pelsser (2004a).

The remainder of the paper is organised as follows. The general setup for stochastic mortality is discussed in section 2, together with some specific examples which will be empirically tested in section 3 on Dutch mortality data. Section 4 discusses the formulation of a market price of mortality risk. Section 5 shows how the model can be used in the pricing of embedded options, more specific in the pricing of GAOs. Section 6 discusses further applications of the specific models in section 2. Section 7 concludes.

2 Affine Mortality Intensity

The modeling of mortality in life insurance is very similar to that of default in the credit risk literature. We follow a special case of the subfiltration approach to modeling default events (see Jamshidian, 2004), namely the Cox-process approach developed by Lando (1998). In the remainder of the paper, the mortality intensity can be thought of as a hazard rate in the context of the Cox-process approach. In this setup, the time of death of a person is modeled as stopping time τ with respect to some filtration, $\mathcal{G}_t$, containing all information (both financial and actuarial). Restrict $t \in \left[0, \overline{T}\right]$ for simplicity. In our setting $\mathcal{G}_t = \mathcal{G}_t^\tau \vee \mathcal{H}_t$ where $\mathcal{G}_t^\tau$ is the filtration generated by the stopping time and $\mathcal{H}_t$ is some subfiltration containing all information except whether the person is alive or not. As a consequence of the Cox-process setup the subfiltration $\mathcal{H}_t$ is conditionally independent (Jamshidian, 2004). Consider a probability space $(\Omega, \mathcal{G}, \mathcal{G}_t, \mathbf{P})$ where the filtration $\mathcal{G}_t$ satisfies the usual conditions and $\mathbf{P}$ denotes the real world probability measure. Set $\mathcal{G}_t = \mathcal{G}_t^\tau \vee \mathcal{H}_t$ and $\mathcal{H}_t = \mathcal{M}_t \vee \mathcal{F}_t$ where $\mathcal{M}_t$ and $\mathcal{F}_t$ contain the information concerning mortality and financial markets respectively.

We propose the following general form for the mortality intensity, $\mu_x(t)$, of a person of age x at time t,

$$\mu_x(t) = g_0(x) + \sum_{i=1}^{M} Y_i(t) g_i(x) \tag{1}$$

where $g_i : x \longrightarrow \mathbf{R}_+$, some function with the positive half line as its range and $Y(t)$ are factors driving the uncertainty in the mortality intensity. The M-dimensional factor dynamics are given by the following diffusion,

$$\begin{aligned} dY(t) &= A(\theta - Y(t))\,dt + \Sigma\sqrt{V_t}dW_t^{\mathbf{P}} \\ Y(0) &= \overline{Y} \end{aligned} \tag{2}$$

where $W_t^{\mathbf{P}}$ is an M-dimensional Brownian Motion under the real world measure. Let $\mathcal{M}_t$ be the filtration generated by $W_t^{\mathbf{P}}$. Furthermore, A and Σ are $M \times M$ matrices and θ is an M vector. The matrix V_t is a diagonal matrix holding the diffusion coefficients of the factors on the diagonal, i.e.

$$V_{t,(ii)} = \alpha_i + \beta_i' Y(t) \qquad i = 1, 2, ..., M \tag{3}$$

where the β_i's are M vectors. Or directly in matrix notation, defining the matrix $\beta = \begin{bmatrix} \beta_1 & \cdots & \beta_M \end{bmatrix}'$ and the vector $\alpha = \begin{bmatrix} \alpha_1 & \cdots & \alpha_M \end{bmatrix}'$, we have[1],

$$V_t = \text{diag}(\alpha + \beta Y(t))$$

We call this an M-factor Affine Mortality Model since the mortality intensity is an affine function of the factors. Furthermore the instantaneous drift and variance of the factors are affine functions of the factors.

Notice that we explicitly mention the starting values of the factors, $Y(0)$. Mostly mortality is thought of as exhibiting a decreasing trend. Furthermore we would expect to find the variance of the factors to be small compared to its actual value. By setting the starting values of the factors, in our applications, above the (long run) mean we can model this low volatility decreasing trend using an affine diffusion. This enables us to capitalize on the analytical properties of these type of models.

It is well known that in affine term structure models bond prices are exponentially affine in the factors. In our model for the mortality intensity we have a similar result since survival probabilities are exponential affine in the factors. Thus in ATSMs we have for the time t price of a bond maturing at time T, $D(t,T) = \exp(A(t,T) - B(t,T) \cdot X_t)$. The coefficients A and B can be obtained by solving a system of ODE's, known as Riccati equations.

Survival probabilities in our model are given by the following theorem.

[1] If $\mathbf{x}$ is an M-vector then we define $diag(\mathbf{x})$ to be the $M \times M$ diagonal matrix with the elements of $\mathbf{x}$ on the diagonal.

Theorem 1 *Consider the model in (1) and (2) and define the M-vector,*

$$g(x+s) = \begin{bmatrix} g_1(x+s) & \cdots & g_M(x+s) \end{bmatrix}'$$

Then the $T-t$-year survival probability of an $x+t$ year old is given by

$$\begin{aligned} {}_{T-t}p_{x+t}(t) &= E^{\mathbf{P}}\left[\mathbf{1}_{[\tau_x > T]} | \mathcal{H}_t \cup [\tau_x > t]\right] \\ &= E^{\mathbf{P}}\left[\exp\left(-\int_t^T \mu_{x+s}(s)\, ds\right) | \mathcal{M}_t\right] \end{aligned}$$

this expectation reduces to,

$${}_{T-t}p_{x+t}(t) = \exp\left(\gamma(x,t,T) - \delta(x,t,T) Y_t\right)$$

where the coefficients $\gamma(x,t,T)$ and $\delta(x,t,T)$ are solutions of the following system of ODE[2]*,*

$$\dot{\delta}_t \equiv \frac{d\delta(x,t,T)}{dt} = -g(x+t) + A'\delta(x,t,T) + \frac{1}{2}\sum_{i=1}^{M} \left[\Sigma'\delta(x,t,T)\right]_i^2 \beta_i \tag{4}$$

$$\dot{\gamma}_t \equiv \frac{d\gamma(x,t,T)}{dt} = -g_0(x+t) - \theta' A' \delta(x,t,T) + \frac{1}{2}\sum_{i=1}^{M} \left[\Sigma'\delta(x,t,T)\right]_i^2 \alpha_i \tag{5}$$

with boundary conditions $\gamma(x,T,T) = 0$, $\delta(x,T,T) = \overrightarrow{0}$.

Proof. From the Feynman Kac theorem it follows that ${}_{T-t}p_{x+t}(t) \equiv v(t, Y_t)$ is a solution to the PDE (we abbreviate $Y_i(t)$ to Y_{ti} and $v(t, Y_t)$ to v),

$$\frac{\partial v}{\partial t} + \left(\theta' A' - Y_t{}' A'\right)\frac{\partial v}{\partial Y_t} + \frac{1}{2}\sum_{k,\ j=1}^{M} \left(\Sigma S_t \Sigma'\right)_{kj} \frac{\partial^2 v}{\partial Y_{tk} \partial Y_{tj}} - \left[g_0(x+t) + Y_t{}' g(x+t)\right] v = 0$$

to which a solution exists which is unique. Use $\exp\left(\gamma(x,t,T) - \delta(x,t,T) Y_t\right)$ as a trial solution to compute,

$$\left[\dot{\gamma}_t - Y_t{}'\dot{\delta}_t - \left(\theta' A' - Y_t{}' A'\right)\delta_t + \frac{1}{2}\sum_{k,\ j=1}^{M}\sum_{i=1}^{M}\left\{\Sigma_{ki}\left(\alpha_i + \beta_i' Y_t\right)\Sigma_{ji}\delta_{tk}\delta_{tj}\right\} - g_0(x+t) - Y_t{}' g(x+t)\right] v = 0$$

which can be simplified[3] to,

$$\dot{\gamma}_t - \dot{\delta}_t{}' Y_t - \left(\theta' A' - Y_t{}' A'\right)\delta_t + \frac{1}{2}\sum_{i=1}^{M}\left\{\left[\Sigma'\delta(x,t,T)\right]_i^2 \left(\alpha_i + Y_t{}'\beta_i\right)\right\} - g_0(x+t) - Y_t{}' g(x+t) = 0$$

then by the "affine matching principle" (Duffie and Kan, 1996) $\gamma(x,t,T)$ and $\delta(x,t,T)$ are solutions to the above system of ODE. ■

[2] We let $\left[\Sigma'\delta(x,t,T)\right]_i^2$ denote square of the i-th element of the M-vector which follows from computing $\Sigma'\delta(x,t,T)$.

[3] Note that $\sum_{k,\ j=1}^{M}\sum_{i=1}^{M}\left\{\Sigma_{ki}\left(\alpha_i + \beta_i' Y_t\right)\Sigma_{ji}\delta_{tk}\delta_{tj}\right\} = \sum_{i=1}^{M}\left\{\left(\alpha_i + \beta_i' Y_t\right)\left(\sum_{k=1}^{M}\Sigma_{ki}\delta_{tk}\right)\left(\sum_{j=1}^{M}\Sigma_{ji}\delta_{tj}\right)\right\}$ and $\left(\sum_{k=1}^{M}\Sigma_{ki}\delta_{tk}\right) = \left[\Sigma'\delta(x,t,T)\right]_i$

In equations (1) and (2) the model is formulated under the real world measure $\mathbf{P}$. When we assume the existence of a unique pricing measure, equivalent to $\mathbf{P}$, we can formulate a model for the factors under this measure. This will be the subject of section 4.

Notice that (1) enables us to model the mortality intensity of all ages simultaneously! This contrasts the approaches of Dahl (2004) and Milevsky and Promislow (2001) where only a single age at a time is considered.

The formulation of the model in (1) and (2) is very general. It already satisfies our goals of tractability and consistency with derivative pricing models. However in this form it is too general for our purposes. We next propose a parameterization of (1) and (2) which adds the required interpretation.

2.1 Special Case: Gaussian Thiele Model

In the special case of Gaussian factors one can obtain nice analytical expressions for the survival probabilities. In this section we combine Gaussian factors with the functional form for the dependence of mortality intensity on age postulated by Thiele in 1867. Thiele proposed the following functional form for the mortality intensity,

$$\mu_x = Y_1 \exp\left(-\tau_1 x\right) + Y_2 \exp\left(-\tau_2 \left(x-\eta\right)^2\right) + Y_3 \exp\left(\tau_3 x\right) \tag{6}$$

where all parameters are positive. This specification allows for modeling the behavior of seperate age groups. The third term captures the general mortality behavior (increasing with age). The first term which is decreasing with age correponds to (additional) mortality at young ages. The second term allows for (additional) hump shaped behavior of mortality of middle aged people (typically young adults). The functional form of Thiele nests the well known Makeham and Gompertz mortality laws.

We can add uncertainty to (6) if we let some or all of the parameters follow stochastic processes. We can fit (6) into the framework of (1) by choosing $g_0(x) = 0$, $g_1(x) = \exp(-\tau_1 x)$, $g_2(x) = \exp\left(-\tau_2 (x-\eta)^2\right)$ and $g_3(x) = \exp(\tau_3 x)$ and let the parameters Y_i, $i = 1, 2, 3$ follow an affine SDE as in (2),

$$\mu_x(t) = Y_1(t) \exp\left(-\tau_1 x\right) + Y_2(t) \exp\left(-\tau_2 \left(x-\eta\right)^2\right) + Y_3(t) \exp\left(\tau_3 x\right) \tag{7}$$

To obtain a Gaussian Stochastic Thiele model we restrict ourselves to Gaussian factor dynamics, i.e. we let the factors follow a multivariate Ornstein-Uhlenbeck process. Without loss of generality we have for the SDE of the factors,

$$\begin{aligned} dY(t) &= A\left(\theta - Y(t)\right) dt + \Sigma dW_t^{\mathbf{P}} \\ Y(0) &= \overline{Y} \end{aligned} \tag{8}$$

where $A = diag(a)$, and $a = [a_1; a_2; a_3]'$ is a vector in $\mathbf{R}^3_+$. Now we have the following,

Corollary 2 *Consider the model in (7) and (8). Then the $T-t$-year survival probability of an $x+t$ year old is given by*

$${}_{T-t}p_{x+t}(t) = \exp\left(C(x,t,T) - D_1(x,t,T)\, Y_1(t) - D_2(x,t,T)\, Y_2(t) - D_3(x,t,T)\, Y_3(t)\right)$$

where we can solve explicitly for D_1 and D_3

$$D_1(t,x,T) = \exp(-\beta_1[x+t])\frac{1-e^{-(\beta_1+a_1)(T-t)}}{\beta_1+a_2} \tag{9}$$

$$D_3(t,x,T) = \exp(\beta_3[x+t])\frac{1-e^{(\beta_3-a_2)(T-t)}}{a_2-\beta_3} \tag{10}$$

and D_2 and C solve the following ODE,

$$\frac{dD_2(x,t,T)}{dt} = \exp\left(-\beta_2([x+t]-\eta)^2\right)+a_2D_2(x,t,T) \tag{11}$$

$$\frac{dC(x,t,T)}{dn} = \theta' A' D(x,t,T)-\frac{1}{2}\sum_{i=1}^{M}[\Sigma' D(x,t,T)]_i^2 \tag{12}$$

with boundary conditions $C(x,T,T)=0$, $D_2(x,T,T)=0$.

Proof. This follows from directly Theorem 1 and the substitution of (7) and (8) in (1) and (2). ■

The resulting model has some desirable properties. First, if $\overline{Y} > \theta$ the factors display an exponentially decreasing trend, allowing for decreasing mortality rates over time. Second, the model is able to capture different age groups in a single equation for the relation between the factors and the mortality intensity. Finally, the model is analytically tractable which makes it very well suited for estimation and pricing.

As mentioned in the introduction our model captures three out of four types of mortality risk. Recall that the four types of mortality risk usually distinguished in practice are: trend, level, volatility and catastrophe. Our model is perfectly capable of tracking the general characteristics of the three most important types of mortality risk seperately using the parameters of the SDE of the driving factors. We capture the mortality trend using the matrix A, the mortality volatility by Σ. Furthermore the model can be estimated for both the entire population as for the population of insured with or without parameter equality restrictions linking the versions of the model. Within our specification however we are not able nor aiming to capture the effect of (mortality) catastrophe. Some jump component could be added to the specification. This is left for future research.

2.2 Special Case: Gaussian Makeham Model

A specification nested in the one by Thiele is Makeham's. Makeham proposed the following functional form for the mortality intensity,

$$\mu_x = Y_1 + Y_2 c^x \tag{13}$$

Where Y_1 and Y_2 are positive and $c > 1$. As one can see, Makeham's specification doesn't take the "middle age hump" into account. It merely adds an age related growth factor to a "base" mortality constant. Just like before we can add uncertainty to (13) if we let some or all of the parameters follow stochastic processes. We can fit (13) into the framework of (1) by choosing $g_0(x)=0$, $g_1(x)=1$ and $g_2(x)=c^x$ and let the parameters Y_i, $i=1,2$ follow an affine SDE as in (2),

$$\mu_x(t) = Y_1(t) + Y_2(t)c^x \tag{14}$$

We obtain a Gaussian Stochastic Makeham model by assuming Y_t follows an OU process like in (8). For completeness,

$$\begin{aligned} dY_i(t) &= a_i(\theta_i - Y_i(t))\,dt + \sigma_i dW_{it}^{\mathbf{P}} \quad , \; Y_i(0) = \overline{Y}_i \; , \; i = 1, 2 \\ dW_{1t}^{\mathbf{P}} dW_{2t}^{\mathbf{P}} &= \rho dt \end{aligned} \tag{15}$$

where for simplicity we assume $\theta_2 = 0$. Because of the simple structure we can obtain an analytic expression for the survival probability.

Corollary 3 *In the Gaussian Stochastic Makeham model we have the following expression for the $T-t$-year survival probability at time t of an $x+t$ year old,*

$${}_{T-t}p_{x+t}(t) = \exp\left(C(x,t,T) - D_1(x,t,T)\,Y_1(t) - D_2(x,t,T)\,Y_2(t)\right)$$

where,

$$\begin{aligned} D_1(x,t,T) &= \frac{1 - e^{-a_1(T-t)}}{a_1} \\ D_2(t,x,T) &= c^{x+t}\frac{1 - e^{(\beta - a_2)(T-t)}}{a_2 - \beta} \qquad \textit{where } \beta = \ln(c) \end{aligned}$$

furthermore we have that,

$$\begin{aligned} C(x,t,T) = & -\theta_1(T-t) - \frac{1 - e^{-a_1(T-t)}}{a_1}\theta_1 + \frac{\sigma_1^2}{2a_1^3}\left[a_1(T-t) - 2\left(1 - e^{-a_1(T-t)}\right) + \frac{1}{2}\left(1 - e^{-2a_1(T-t)}\right)\right] + \\ & \frac{\sigma_2^2 c^{2(x+t)}}{2(a_2-\beta)^3}\left[(a_2-\beta)(T-t) - 2\left(1 - e^{-(a_2-\beta)(T-t)}\right) + \frac{1}{2}\left(1 - e^{-2(a_2-\beta)(T-t)}\right)\right] + \frac{\rho\sigma_1\sigma_2 c^{x+t}}{2\left[a_1(a_2-\beta)\right]^2}\cdot \\ & [\; a_1(a_2-\beta)(T-t) - a_1\left(1 - e^{-(a_2-\beta)(T-t)}\right) - (a_2-\beta)\left(1 - e^{-a_1(T-t)}\right) + \\ & \frac{a_1(a_2-\beta)}{a_1 + (a_2-\beta)}\left(1 - e^{-(a_1 + (a_2-\beta))(T-t)}\right) \;] \end{aligned}$$

Proof. It is easily verified these expressions satisfy (4) and (5). ■

Notice that we can remove the dependence of the coefficients C and D on t using a change of variables, $n = T - t$ and $z = x + t$ and hence write ${}_np_z(t)$. We will use these expressions when we estimate the model in section 3.

2.3 Gaussian Assumption

Mortality intensity is by definition non-negative. Unfortunately when the factors are Gaussian one cannot exclude negative mortality rates. This has the same drawback as negative interest rates. In our view, given the nice analytical structure of these models, this doesn't disqualify them as good modeling tools to quantify mortality risk. We will show in section 3 that these Gaussian models do a nice job in explaining the variation of mortality rates over time. In any case, parameter values can always be chosen such that the probability of negative values is small. This is supported by our estimates in the next section which all but exclude the possibility of negative mortality intensity. As an alternative one could use a decreasing CIR processes,

$$\begin{aligned} dY(t) &= -AY(t)\,dt + \Sigma\sqrt{Y(t)}dW_i(t) \\ Y(0) &= \overline{Y} \end{aligned}$$

The factors $Y(t)$ will be absorbed at the boundary zero. In any case we have to take a pragmatic approach to obtain workable results, which is our first priority. We either have to deal with absorbing boundaries or with possible negative mortality intensities. The latter generally do not show up either in estimation or in pricing at reasonable parameter values. Absorbtion at zero will be unrealistic as well and will generally not occur very often at estimated parameter values because of the low variance of the factors.

To use an example from the literature on credit risk, Duffee (1999) allows for negative hazard rates. He defends the possibility of negative rates by saying it is necessary for the model to fit the observed term structures (both fairly steep and flat).

To make sure the hazard rate is strictly positive in the CIR model we can use, e.g. for the Thiele specification,

$$\begin{aligned} \mu_x(t) &= Y_1(t)\exp(-\beta_1 x) + Y_2(t)\exp\left(-\beta_2(x-\eta)^2\right) + [Y_3(t)+\overline{\mu}]\exp(\beta_3 x) \\ Y(t) &= -AY(t)\,dt + \Sigma\sqrt{Y(t)}dW_i(t) \qquad \text{with } Y(0)=\overline{Y} \end{aligned}$$

for some constant $\overline{\mu}$.

3 Application to Dutch Mortality Rates

In this section we test the validity of the Gaussian Thiele model as a tool for mortality risk modeling. We estimate several versions of the model using empirical data on mortality rates. The data are mortality coefficients of the Dutch male population from 1950 to 2002, age zero to 89. The data are available from the Dutch Central Bureau for Statistics. The mortality coefficients can be interpreted as one year mortality probabilities, notation q_x for age x, for the respective ages. In practice one should off course use portfolio data (at least in conjunction with population data) to estimate the relevant development of mortality. For applications, for example longevity risk in life annuities (see also the part in section 7 on Economic Capital), one can easily choose to estimate the model for ages above / below a certain treshold relevant to that application. This will improve model fit for the relevant age group. The model could be extended to allow for risk classification (e.g. smoking / non-smoking), we will come back to that in section 3.3.

We estimated several models using the following approach. Using Non Linear Least Squares we estimated several nested specifications of (6) with time varying parameters. This will help us to determine wether the specification in (7) is reasonable. This analysis is carried out in section 3.1. Next, in 3.2., we estimate (7) in combination with (8) with Maximum Likelihood by the Kalman Filter.

3.1 Specification Analysis

Before we actually estimate the model in (7) via the Kalman Filter we spend some time on model specification issues. We want to know whether fixing the parameters, τ_i, $i = 1, 2, 3$, over the estimation period is not too restrictive. To investigate this we proceed as follows. First, we estimate,

$$q_{x.t} = Y_{1.t} \exp\left(-\tau_{1.t} x\right) + Y_{2.t} \exp\left(-\tau_{2.t} \left(x - \eta_t\right)^2\right) + Y_{3.t} \exp\left(\tau_{3.t} x\right) + \varepsilon_{x.t} \tag{16}$$

for $t = 1950, ..., 2002$ by non-linear least squares (NLLS). Furthermore $\text{Var}(\varepsilon_{x.t}) = \sigma^2 q_{x.t}^2$ and $E\left(\varepsilon_{x.t-i}\varepsilon_{x.t}\right) = E\left(\varepsilon_{x.t}\varepsilon_{x-j.t}\right) = 0$, $\forall\ i, j \neq 0$. We let the standard deviation of the error increase proportional to the mortality probability so we effectively minimize the squared relative error. This model is a non-linear seemingly unrelated regression (NL-SUR) model and hence we can estimate the parameters by applying NLLS to the equation for each t. It is the most flexible model which postulates a mortality intensity of the form (6). This gives us a benchmark fit to test other specifications. In this form the model has $7 \times (2002 - 1950)$ parameters. We also estimated the following versions of the model (by NLLS),

$$q_{x.t} = Y_{1.t} \exp\left(-\tau_1 x\right) + Y_{2.t} \exp\left(-\tau_2 \left(x - \eta\right)^2\right) + Y_{3.t} \exp\left(\tau_3 x\right) + \varepsilon_{x.t} \tag{17}$$

$$q_{x.t} = A_t + B_t \exp\left(C_t\ x\right) + \varepsilon_{x.t} \tag{18}$$

$$q_{x.t} = A_t + B_t \exp\left(C\ x\right) + \varepsilon_{x.t} \tag{19}$$

Model (17) is a restricted version of (16) and points in the direction of (7). Model (18) is the Makeham version of (16), model (19) restricts the Makeham specification to have constant curvature parameter C. The results of the estimation procedure are in table 2. We present both the total sum of squared errors (TSS), the Akaike information criterion (AIC), the Schwarz criterion (SC) and the Mean Absolute Relative Error (MARE) for the four models. Based on these numbers we conclude that restricting several of the parameters in (16) to be constant over time is not too costly. The loss of MARE is only 2%, although the AIC is better for the unrestricted model, the SC is better for the restricted than for the unrestricted model. The same holds for the restricted Makeham version of the model. The total fit is not much worse than that of (16). The AIC is almost equal for the unrestricted and restricted model, whereas the SC is better for the unrestricted model. If we compare the Thiele and Makeham specification we see that adding $T + 3$ parameters makes sense, the AIC and SC of (17) are much lower than the AIC and SC of (19) respectively. The gain in MARE from (19) to (17) is 5%. We conclude that based on these results the affine specification of the Thiele model (i.e. constant β's and η) is a good candidate for a dynamic stochastic mortality model.

3.2 Kalman Filter Estimation

Based on the results in section 3.1. we now proceed with estimation of (7) and (14). This model can be formulated in state space form and estimated by the Kalman filter (cf. de Jong, 2000, Duffee and Stanton, 2004). The model equations are[4],

$$\overline{q}_x(t) \equiv -\ln(1-q_x(t)) = -A(x,t,t+1) + B(x,t,t+1)\cdot Y_t + \nu_{x,t} \tag{20}$$

$$Y_{t+1} = e^{-A}Y_t + \varepsilon_t \quad , Y_1 = \overline{Y} \tag{21}$$

$$\mathrm{Var}(\nu_{x,t}) = s^2\overline{q}_x(t)^2\,,\ \ \mathrm{Var}(\varepsilon_t) = \Omega,\ \ E(\nu_t\varepsilon_t) = 0,\ E(\varepsilon_{t-i}\varepsilon_t') = 0\ \forall\ i \neq 0 \tag{22}$$

Where Y_t is either two or three dimensional for the Makeham and Thiele versions of the model respectively. As one can see in (22), in this formulation the standard deviation of the measurement error is proportional to the measurement. This is a simple way to make sure the estimation procedure cares more about relative than absolute errors. Although there are certainly other ways, our objective in this section is to show we can easily obtain parameter estimates for both the Makeham and Thiele Affine model using the Kalman Filter. Before we can estimate the model there are two remaining problems. The first is that of starting values. We need to initialise the Kalman Filter and the estimation procedure. This can be solved relatively easy. We can use the results from the NLLS analysis to obtain starting values for the parameters and factors. Another, more difficult, problem is that of the dimension of the vector with observations. For each year we have a one year mortality probability for ages from 0 to 89. This means that in our Kalman routine we have to invert a 90 by 90 matrix which makes it slow and numerically unstable. To circumvent this problem we use an algorithm developed by Koopman and Durbin (2000) which speeds up computations considerably and makes the Kalman routine numerically robust.

For simplicity (and because theoretical motivation for correlated factors is lacking), in our estimation, we have set all correlation coefficients between factors to zero. Furthermore we have set factor means to zero as well. Estimation results for both models are in table 2. We make two important observations. First the mean reversion of all factors is generally low. This means all factors exhibit a slowly decreasing trend (towards their zero mean). Second recall that in the Makeham model the second factor (and in the Thiele model the third factor) impacts the general and old age mortality. Although the variance of all factors is generally low, the variance of the third factor is particularly so. This tells us the improvements in mortality for old ages have been more or less "deterministic" over the last half a century.

In figures 1 and 2 we plot the MARE per age for the two different models. It is clear that the Makeham model has a bad fit for the ages below 25. We can still see the "middle age hump" and the high child mortality. It is precisely these ages the additional terms in (7) aim to model. As is clear from figure 2, the Thiele model succeeds in modeling these lower ages and does a much better job overall.

A likelihood ratio test on the joint restrictions, $\beta_1 = \eta = \beta_2 = a_2 = \sigma_2 = 0$, i.e. the Makeham two factor model against the Thiele three factor model is rejected on a 99% confidence basis (test statistic of 4673.45 against a critical value of 15.09. The MARE of the Thiele model is 11%. All in all not a bad performance for a model with such simple dynamics.

[4] The parameter s denotes the standard error of the model misspecification, or measurement error in the Kalman Filter jargon.

3.3 Heterogeneity

Frailty or risk classification can easily be taken into account. In this section we take the point of view that heterogeneity at an individual level is of little use in practice. This is the case simply because individual heterogeneity cannot be identified in estimation. Therefore we would like to classify heterogeneity into a limited number of known risks. One could extend the model to,

$$\widetilde{\mu}_x(t) = \mu_x(t) + \mathbf{i} \cdot \lambda_x(t) \tag{23}$$

where λ_x is a vector of additional (mortality) intensities for each risk factor (e.g. smoking, obese). This is similar to Willemse (2004) who argues that frailty can be modeled as an individual age reduction, we model heterogeneity by adding a heterogeneous component to mortality intensity. If λ_x is driven by affine dynamics equation () leads to survival probabilities by a simple generalization of theorem 1. If the relevant data are available for each risk group the model could be estimated using techniques similar to those used in section 3.2.

4 Market Price of Mortality Risk

Up till now we have only been concerned with modeling the real world behavior of the mortality intensity. However one of our main objectives was to have a model which could easily be used for the pricing of standard insurance products *and* embedded options. Therefore we now turn attention to the behavior of the process under the pricing measure and hence the specification of the market price of mortality risk. Before doing so we should first formally give a description of an insurance contract (which can be easily extended to a portfolio of insurance contracts). The following discussion is partly based on Dahl (2004).

Note that the combined financial and insurance market is incomplete so there are many (even infinitely many) pricing measures, or equivalent martingale measures (EMM). Calculations are done here for a fixed choice of EMM. Let the financial market be driven by a vector Brownian motion $\overline{W}_t^{\mathbf{P}}$ and let $\mathcal{F}_t$ be the filtration generated by this Brownian motion, $\mathcal{F}_t = \sigma\left(\overline{W}_u^{\mathbf{P}}, u \leq t\right)$. Furthermore let $\mathcal{M}_t$ be the filtration generated by $W_t^{\mathbf{P}}$, $\mathcal{M}_t = \sigma\left(W_u^{\mathbf{P}}, u \leq t\right)$. Define the time of death of individual i aged x_i at time zero to be τ_i. Let τ_i be a $\mathcal{G}_t^{\tau_i}$ stopping time. Now let an insurance contract for this individual be described by the process $N_{x_i}^i(t)$ which starts at 1 at $t = 0$ and takes a jump to zero at time τ_i, $N_{x_i}^i(t) = \mathbf{1}_{[\tau^i \leq t]}$. Then the process $M_{x_i}^i(t)$ defined by,

$$M_{x_i}^i(t) = N_{x_i}^i(t) - \int_0^t \mathbf{1}_{[\tau^i < s]} \mu_{x_i+s}(s)\, ds \tag{24}$$

is a martingale.

Now, considering individuals $i = 1, ..., N$ we can change to an equivalent measure $\mathbf{Q}$ by defining the Radon-Nikodym kernel for a change of $\mathbf{P}$ w.r.t. $\mathbf{Q}$ by,

$$d \ln\left(\frac{d\mathbf{Q}}{d\mathbf{P}}\right)_t = -\kappa_t dW_t - \overline{\kappa}_t d\overline{W}_t + \sum_{i=1}^N \rho_t^i dM_{x_i}^i(t) \tag{25}$$

Contrary to the credit risk literature we have a specific idea about the behavior of the force of mortality over time and age. In the credit risk literature the market price of default risk is usually introduced in the form of "adjusting the default intensity", i.e. $\rho_t^i \neq 0$ and $\kappa_t = 0$, when changing to the risk neutral measure. We feel, because we have a clearcut model for the force of mortality, it is more natural to let the change of measure influence the driving process of the force of mortality, i.e. the process for Y_t instead of $\mu_x(t)$ directly. This results in the following Radon-Nikodym kernel,

$$d \ln \left(\frac{d\mathbf{Q}}{d\mathbf{P}} \right)_t = -\kappa_t dW_t - \overline{\kappa}_t d\overline{W}_t \tag{26}$$

This reasoning has as a main advantage in that we can change measure for all ages simultaneously without having to specify individual ρ_t^i. Although we could work with infinitely many ages using (26) avoids the use of an infinite dimensional Radon-Nikodym kernel. In both the Makeham and especially the Thiele specification the factors have a clear interpretation as modeling mortality risk for young / middle / old ages. Hence there's no need for us to influence the probability law of the process $N_{x_i}^i$ *explicitly*. Instead we can influence it *implicitly* by influencing the process Y_t driving $\mu_x(t)$.

As we mentioned in section 2.1 each part of equation (6) models the force of mortality for a specific age group. Furthermore in the stochastic formulation in (7) each part is driven by a latent factor. This is very convenient since this allows us to choose market prices of risk for the factors such that the force of mortality under the pricing measure is prudent compared to the real world measure. For example in the presence of longevity risk we expect $a_3^{\mathbf{Q}} < a_3$ in the Thiele model. Then we have ${}_np_x(t) <_n p_x^{\mathbf{Q}}(t)$. The real world probability of survival is lower than the same probability used for valuation. For young ages the risk is usually of different sign (if we want to be on the safe side, prudent, from the insurer's point of view) and we have $a_1^{\mathbf{Q}} > a_1$.

5 Valuation of Endowments, Annuities and Guaranteed Annuity Options

In this section we follow a credit risk approach to value insurance payments conditional upon survival of the insured. We use the obvious analogy between default intensity and force of mortality and model the "interest plus mean loss rate" to arrive at a mortality adjusted discount rate. Define the actuarial discount rate $R_x(t)$, which is both age and time dependent, as,

$$R_x(t) = r_t + \mu_x(t) \tag{27}$$

This mortality adjusted discount rate can then be used to discount payments conditional upon survival. Assume r_t is driven by a (vector) Brownian motion $\overline{W}_t^{\mathbf{P}}$. Like in section 4 let $\mathcal{F}_t$ be the filtration generated by this Brownian motion, $\mathcal{F}_t = \sigma\left(\overline{W}_u^{\mathbf{P}}, u \leq t\right)$ and $\mathcal{M}_t$ be the filtration generated by $W_t^{\mathbf{P}}$, $\mathcal{M}_t = \sigma\left(W_u^{\mathbf{P}}, u \leq t\right)$. As a consequence of our Cox process setup we can use a result in Lando (1998) which can be found in a more general form in Jamshidian (2004, proposition 3.9) and reduce the valuation of a claim conditional upon survival to that of a "non-defaultable claim". This implies we can regard the actuarial discount rate as a normal interest rate and use well known change of numeraire techniques (see e.g. Karoui and Rochet, 1989, Jamshidian, 1991, Geman et al, 1995. and Jamshidian, 1998).

The results in Lando (1998) state that the value of a pure endowment ${}_nE_x(t)$ for an x-year old is given by,

$$ {}_nE_x(t) = E_t^{\mathbf{Q}}\left[\exp\left(-\int_0^n R_{x+s}(t+s)\,ds\right)\right] \tag{28}$$

where the expectation is taken under the risk neutral measure. If we assume independence between financial markets and mortality, which is natural[5], we obtain,

$$ {}_nE_x(t) = D(t,t+n)\ \ {}_np_x^{\mathbf{Q}}(t) \tag{29}$$

Where $D(t,T)$ denotes the time t price of a zero bond with maturity T. This is analogous to the deterministic mortality case. As a first consequence of the result in Lando (1998) and Jamshidian (2004) we can regard ${}_nE_x(t)$ as a regular bond price and use it as a numeraire.

Now, an n-year annuity for an x-year old, notation $a_{x:n}(t)$, is just a sequence of endowments and can be valued as such (this is similar to the realtionship between zero and coupon bonds). To be specific,

$$\begin{aligned} a_{x:n}(t) &= \sum_{i=1}^{n} {}_iE_x(t) && (30)\\ &= \sum_{i=1}^{n} D(t,t+i)\ \ {}_ip_x^{\mathbf{Q}}(t) && (31)\end{aligned}$$

An m-year deferred annuity, $a_{x:n}^m(t)$, is an annuity with a deferred first payment date,

$$ a_{x:n}^m(t) = \sum_{i=1}^{n} {}_{m+i}E_x(t) \tag{32}$$

We could even price a Rate of Return Guarantee under stochastic mortality. Consider a contract paying out in n-years time the maximum of some investment S_t and the guaranteed amount K conditional upon survival of the insured. The time t value of the guarantee embedded in the contract for an x-year old, $G_x(t)$, is given by,

$$\begin{aligned} G_x(t) &= E_t^{\mathbf{Q}}\left[\exp\left(-\int_0^n R_{x+s}(t+s)\,ds\right)(K-S_{t+n})^+\right] && (33)\\ &= {}_np_x^{\mathbf{Q}}(t)\,D(t,t+n)\,E_t^{\mathbf{Q}^{t+n}}\left[(K-S_{t+n})^+\right] \end{aligned}$$

where $\mathbf{Q}^{t+n}$ denotes the $t+n$-forward measure. Since all the payoff in (28), (30), (32) and (33) depend linearly on mortality, because of the independence between mortality and financial markets, we are able to account for stochastic mortality in a straightforward manner.

Now consider the more interesting problem of a guaranteed annuity option (GAO). GAOs are treated in a number of papers in the literature, see Boyle and Hardy (2003) and Pelsser (2003) among others. A GAO gives

[5] Positive effects of improvements in mortality on financial markets are not considered. This would require e.g. a macro model including aggregate consumption with micro foundations. Such an approach is certainly interesting but beyond the scope of this paper.

the policyholder the right to convert an assured sum into an annuity at the maximum of the prevailing market annuity rate or a guaranteed annuity rate, say r_x^G. The market forward annuity rate at time t for an x-year old entering an n-year annuity starting at T is defined by $r_{x,T,n}(t) \equiv D(t,T) \ / \ a_{x:n}^{T-t}(t)$. Using the model of section 2 and the probabalistic setup of section 4 we can explicitly take into account the stochastic nature of mortality when pricing this option. From Pelsser (2003) we obtain for the value at time t of a GAO put option with exercise date T on an n-year annuity, which before the exercise date is a deferred annuity,

$$\begin{aligned} GAO(t,T,x,n) &= E_t^{\mathbf{Q}}\left[\exp\left(-\int_0^{T-t} R_{x+s}(t+s)\,ds\right) a_{x+T-t:n}^{T-T}(T)\left(r_x^G - r_{x+T-t,T,n}(T)\right)^+\right] \\ &= a_{x:n}^{T-t}(t)\,E_t^{\mathbf{Q}^A}\left[\left(r_x^G - r_{x+T-t,T,n}(T)\right)^+\right] \end{aligned} \tag{34}$$

where the second line follows from a change to the actuarial annuity measure[6], notation $\mathbf{Q}^A$. When we proceed along the lines of Pelsser (2003) and assume bond volatility is given by $b_T(t)$ for a bond with maturity T, where the bond price dynamics are driven by a Brownian motion $\overline{W}_t^{\mathbf{P}}$, furthermore we assume, that the volatility of a survival probability is given by $\Delta(x,t,T)$[7], then we can show the SDE for an endowment is given by,

$$d_{T-t}E_{x+t}(t) = ...dt + b(t,T)\,d\overline{W}_t^{\mathbf{P}} + \Delta(x,t,T)\,dW_t^{\mathbf{P}}$$

and hence the SDE for the forward annuity rate $r_{x,T,n}(t)$ is given by,

$$\begin{aligned} dr_{x,T,n}(t) &= -r_{x,T,n}(t)\left[\left(\sum_{i=1}^n w_i(t)\left[b(t,T+i) - b(t,T)\right]\right)d\overline{W}_t^{\mathbf{Q}^A} + \left(\sum_{i=1}^n w_i(t)\left[\Delta(x,t,T+i) - \Delta(x,t,T)\right]\right)dW_t^{\mathbf{Q}^A}\right] \\ &\overset{d}{=} -r_{x,T,n}(t)\,\sigma_{x,T,n}(t)\,dZ_t^{\mathbf{Q}^A} \end{aligned} \tag{35}$$

where $w_i(t) = {}_{T+i}\,p_x(t)\,D(t,T+i) \ / \ a_{x:n}^{T-t}(t)$, $\overline{W}_t^{\mathbf{Q}^A}$ and $W_t^{\mathbf{Q}^A}$ are uncorrelated (vector) Brownian motions under $\mathbf{Q}^A$, $\overset{d}{=}$ denotes equality in distribution and we define $\sigma_{x,T,n}(t)$ implicitly. Also define integrated annuity rate volatility by $\sigma_{x,T,n}^2 = \int_0^T \sigma_{x,T,n}(t)^2\,dt$. One can see that the volatility of the annuity rate consists of an interest and a mortality component. If we assume we have a Gaussian term structure model in combination with a Gaussian stochastic mortality model (e.g. one of the models in section 2) we can freeze the "weights", $w_i(t)$, and proceed along lines of Schrager and Pelsser (2004b)

Schrager and Pelsser (2004b) derive an analytical approximation to the price of a swaption in Gaussian term structure methods. They show the swap rate is approximately Gaussian in Gaussian term structure models. We will now use their approach to show the annuity rate is approximately Gaussian in a combined Gaussian term structure and Makeham model to derive an approximate price for a GAO. If we assume a Gaussian model for the term structure, i.e. the short rate under the risk neutral measure is given by

$$\begin{aligned} r_t &= \omega_0 + \omega_X X_t \\ dX_t &= \gamma(\phi - X_t)\,dt + \widehat{\Sigma}\,d\overline{W}_t^{\mathbf{Q}} \end{aligned}$$

[6] Define the forward annuity measure as the measure induced by taking $a_{x:n}^{T-t}(t)$ as a numeraire. This is a valid approach given the strict positivity of ${}_{T-t}p_x(t)$ implied by the Cox-process setting.

[7] Explicitly, we mean that, $d_{T-t}p_{x+t}(t) = ...dt + \Delta(x,t,T)\,dW_t^{\mathbf{P}}$.

where γ is a diagonal matrix. If we combine this with a Gaussian stochastic mortality model (more specifically we assume the Makeham model of section 2.2[8]) we have that,

$$\begin{aligned} b(t,T) &= B(t,T)\,\widehat{\Sigma} \\ \Delta(x,t,T) &= \begin{bmatrix} \sigma_1 D_1(x,t,T) & \sigma_2 D_2(x,t,T) \end{bmatrix}' \end{aligned}$$

where each element of the vector $B(t,T)$ is given by $B^{(i)}(t,T) = 1 / \kappa_{(ii)} \left(1 - e^{-\kappa_{(ii)}(T-t)}\right)$. Under this combined Gaussian model endowment volatility is deterministic.

Now observe that it is precisely the weights $w_i(t)$ which make the volatility of $r_{x,T,n}(t)$ a complicated expression. In (35) these weights are calculated by taking the value of an endowment over the value of a deferred annuity. However this annuity is also chosen to be the numeraire. This means the weights are martingales under the annuity measure, $\mathbf{Q}^A$. Schrager and Pelsser (2004b) show these type of expressions can be well approximated by their value at the time of valuation (set at zero for simplicity). If we set the value of the weights at $w_i(0)$ the approximate volatility of a forward annuity rate, $\sigma_{x,T,n}$, becomes deterministic as well,

$$\sigma_{x,T,n} = \sqrt{\left(\sum_{i=1}^{n} w_i(0) \int_0^T \left[b(s,T+i) - b(s,T)\right]^2 ds\right) + \left(\sum_{i=1}^{n} w_i(0) \int_0^T \left[\Delta(x,s,T+i) - \Delta(x,s,T)\right]^2 ds\right)}$$

which implies the forward annuity rate is approximately a Gaussian Martingale which makes (34) extremely easy to calculate. The approximate time zero price of a GAO is now given by,

$$GAO(0,T,x,n) = a_{x:n}^{T}(0)\left[\left(r_x^G - r_{x,T,n}(0)\right)\Phi\left(\frac{r_x^G - r_{x,T,n}(0)}{\sigma_{x,T,n}}\right) + \sigma_{x,T,n}\,\varphi\left(\frac{r_x^G - r_{x,T,n}(0)}{\sigma_{x,T,n}}\right)\right] \tag{36}$$

where $\varphi(\cdot)$ is the density of a Gaussian r.v. mith mean 0 and s.d. 1 and Φ is the corresponding distribution function. This formula is extremely accurate and can be calculated with means as simple as a spreadsheet.

6 Applications and Numerical Examples

To illustrate the effect of stochastic mortality on the pricing of a GAO we now calculate the value of a 5 year option on a life long annuity for a 60 year old in a combined single factor Hull-White Makeham model. We use the parameter estimates from section 3. Assume the market price of mortality risk is zero so we can price cash flows using real world mortality intensities. Assume the value of the factors is given by $Y_1 = 9.31\text{e-}5$, $Y_2 = 2.19\text{e-}5$. The Hull-White parameters are given by $\kappa = 0.05$ and $\sigma_r = 0.01$. The initial term structure is flat at 5%. Set the guaranteed rate r_x^G at 10%. The market forward annuity rate implied by our assumption is 9.54%. Approximate volatility with and without mortality risk is 10.82% and 10.59% respectively. The value of the GAO as a percentage of annuity payoff is given by 5.50% and 5.43% respectively. This implies a relative increase of the value of the GAO of 1.2% because of mortality risk. The effect of mortality risk increases with the age of the policyholder and the maturity of the option.

[8] This can be easily generalized to any Gaussian stochastic mortality model.

To further illustrate the kind of results one could easily generate using these models (e.g. in a spreadsheet) we now calculate the Economic Capital required for a stylised portfolio of pension policies. We use the following assumptions, the pension funds has 10,000 participants and a premium of 11%, the assumptions on the participants are in table 3. It can be read by row as follows, 15% of the participants is of age 25, the average wage of these participants is 20,000 (Euro or $), the percentage wage growth per year for this class is 6%, which entitles them to a pension of 205,714 at age 65.

A pension is just a deferred annuity of which we can easily calculate the value using (32). We can also calculate the value of each premium payment (assuming the premium remains constant at 11%) as an endowment using (28). The (market) value of the liabilities can be calculated as the value of the pensions net the value of the premiums.

Let's assume the pension funds uses a Gaussian Stochastic Makeham model to model mortality. Again, we assume the term structure (annual compounding) is flat at 5%. We use the Makeham parameter estimates from section 3.2. Again assume the value of the factors is given by Y_1 = 9.31e-5, Y_2 = 2.19e-5. This yields a market value of the liabilities of 2.663 billion. Now we can define the Economic Capital required for mortality risk as the 95% VaR. Since the factors are Gaussian this equals the difference in market value of the liabilities after a 2 sigma shock in each of the factors. The "new" value of the factors is given by Y_1 = 5.73e-5, Y_2 =2.11e-5. The new market value of the liabilities is 2.723 billion. This yields an Economic Capital for mortality risk of 60 million[9].

7 Conclusion

We apply techniques from the literature on term structures and credit risk to the modeling of mortality rates. We specify and estimate a model with a rich analytical structure which is well suited for a combination with continuous time option pricing models. The model can be extended to include heterogeneity in mortality. We estimate a Makeham and Thiele version of the model using data on Dutch mortality rates. Although the model has some shortcomings, in general it fits the data well and it is easily applicable to the pricing of well known embedded options, like GAOs and the calculation of risk measures adopted in the insurance industry, like Economic Capital.

Acknowledgements

The author would like to thank Antoon Pelsser, Pieter Bouwknegt, Erik Torny and KAFEE seminar participants at the University of Amsterdam for discussion and comments.

References

[9] Note that if the "best estimate" liabilities are perfectly matched by the assets the required Economic Capital for interest rate risk is zero.

Bacinello, A.R. and Ortu, F. (1993) Pricing Equity Linked Life Insurance with Endogenous Minimum Guarantee, *Insurance Mathematics and Economics*, vol. 12, pp. 245-257

Boyle, P. and Hardy, M. (2003). Guaranteed Annuity Options, *Astin Bulletin*, vol. 33, 2, pp. 125-152

Brennan M.J. and Schwartz E.S. (1976). The pricing of equity-linked life insurance policies with an asset value guarantee, *Journal of Financial Economics*, vol. 3, pp. 195-213

Chen L. and Filipovic D. (2004). Credit Derivatives in an Affine Framework, *working paper*

Dahl, M.(2004). Stochastic Mortality in Life Insurance: Market Reserves and Mortality-Linked Insurance Contracts, *Insurance Mathematics and Economics*, vol. 35, pp. 113-136

De Jong, F. (2000). Time Series and Cross Section Information in Affine Term Structure Models, *Journal of Business & Economic Statistics*, vol. 18, pp. 300-314

Duffee, G.R. (1999). Estimating the Price of Default Risk, *Review of Financial Studies*, vol. 12, pp. 197-226

Duffee G.R. and Stanton R.H. (2004). Estimation of Dynamic Term Structure Models, *working paper*

Duffie D. and Kan R. (1996). A Yield-Factor Model of Interest Rates, *Mathematical Finance*, vol. 6, pp. 379-406

Duffie D., Pan J. and Singleton K. (2000) Transform Analysis and Asset Pricing for Affine Jump Diffusions, *Econometrica*, vol. 68, pp. 1343-1376

Duffie D. and Singleton K. (1999) Modeling Term Structures of Defaultable Bonds, *Review of Financial Studies*, vol. 12, pp. 687-720

Elliott R., Jeanblanc M. and Yor M. (2000). On models of default risk, *Mathematical Finance*, vol. 10, pp. 179-195

Geman H., El Karoui N. and Rochet J-C. (1995). Changes of Numeraire, Changes of Probability Measure and Option Pricing, *Journal of Applied Probability*, vol. 32, pp 443-458.

Jamshidian, F. (1991). Contingent Claim Evaluation in the Gaussian Interest Rate Model, *Research in Finance*, vol. 9, pp. 131-170

Jamshidian, F. (1998). LIBOR and Swap Market Models and Measures, *Finance and Stochastics*, vol. 1, pp. 293-330

Jamshidian, F. (2004). Valuation of Credit Default Swaps and Swaptions, *Finance and Stochastics*, vol. 8, pp. 343-371

Karoui, N.E. and Rochet J. (1989). A Pricing Formula for Options on Coupon Bonds, *Cahier de Recherche du GREMAQ-CRES*, no. 8925

Koopman S.J. and Durbin J. (2000). Fast Filtering and Smoothing for Multivariate State Space Models, *Journal of Time Series Analysis*, vol. 21, pp.281-296.

Lando, D. (1998). On Cox Processes and Credit Risky Securities, *Review of Derivatives Research*, vol. 2, pp. 99-120

Lee, R.D. (2000). The Lee-Carter Method of Forecasting Mortality, with Various Extensions and Applications, *North American Actuarial Journal*, vol. 4, pp. 80-93

Lee, R.D. and Carter , L.R. (1992). Modelling and Forecasting US Mortality, *Journal of the American Statistical Association*, vol.87, pp. 659-671

Milevsky, M.A. and Promislow S.D. (2001). Mortality Derivatives and the Option to Annuitize, *Insurance Mathematics and Economics*, vol. 29, pp. 299-318

Pelsser A.A.J. (2003). Pricing and Hedging Guaranteed Annuity Options via Static Option Replication, *Insurance Mathematics and Economics*, vol. 33, pp. 283-296

Schrager D.F. and Pelsser A.A.J. (2004a). Pricing Rate of Return Guarantees in Regular Premium Unit Linked Insurance, *Insurance Mathematics and Economics*, vol. 35, pp. 369-398

Schrager D.F. and Pelsser A.A.J. (2004b). Pricing Swaptions in Affine Term Structure Models, *working paper*

Willemse W.J. (2004). Rational reconstruction of frailty-based mortality models by a generalisation of Gompertz' law of mortality, *working paper*

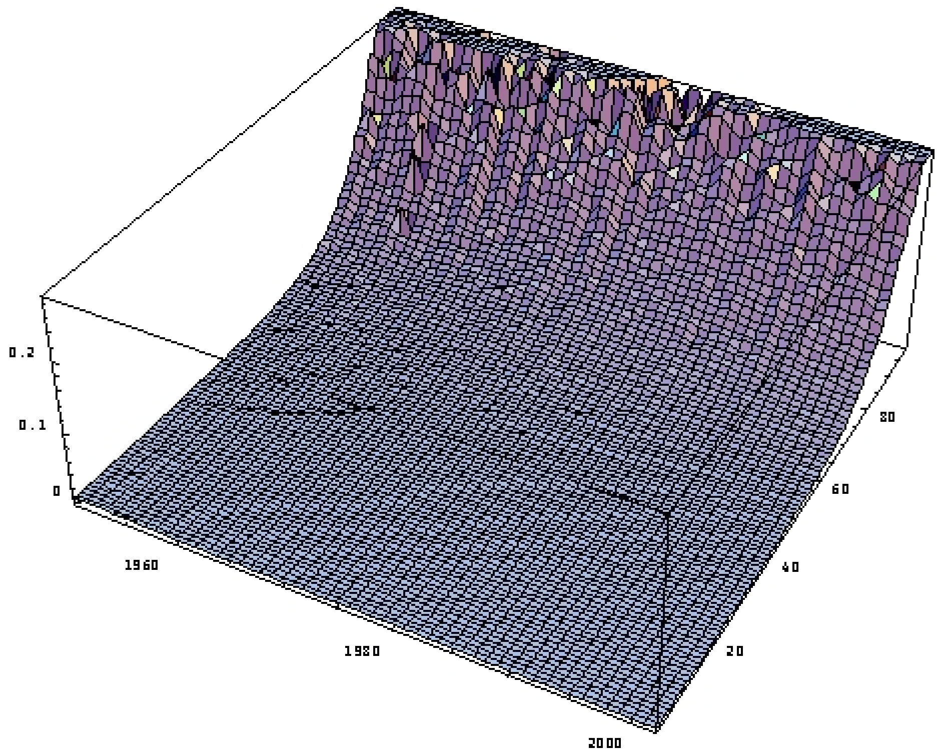

Figure 1. *Mortality coefficients of the Dutch Male population from 1950 to 2002, age zero to 98. Source: Dutch Central Bureau for Statistics.*

Perf. Measures	Thiele	Thiele_const τ_{1-3} / η	Makeham	Makeham_const C
TSS	105.15	149.97	297.04	306.06
AIC	(3.77)	(3.49)	(2.81)	(2.80)
SC	(2.70)	(3.02)	(2.35)	(2.49)
MARE	10%	12%	17%	17%

Table 1. *Performance measures for the different nested specifications of the model in (16). The Thiele specification of the mortality intensity outperforms the Makeham specification. We see that setting some convenient parameters to be constant doesn't harm model fit significantly. Based on the Schwartz criterion the Thiele model with constant τ_{1-3} and constant η is the best among the four specifications.*

Makeham		Thiele	
a_1	0.028	a_1	0.036
a_2	0.0046	a_2	0.018
σ_1 ($\times 10^5$)	1.79	a_3	0.006
σ_2 ($\times 10^7$)	3.83	σ_1 ($\times 10^5$)	7.17
c	1.11	σ_2 ($\times 10^5$)	3.69
σ_ε ($\times 10^4$)	23.90	σ_3 ($\times 10^7$)	5.70
		β_1	0.224
		β_2	0.023
		β_3	0.100
		η	21.82
		σ_ε ($\times 10^4$)	14.35
LogL / ($T \times N$)	6.2368	LogL / ($T \times N$)	6.7361

Table 2. *Kalman Filter parameter estimates and log likelihoods of the Makeham and Thiele affine mortality models.*

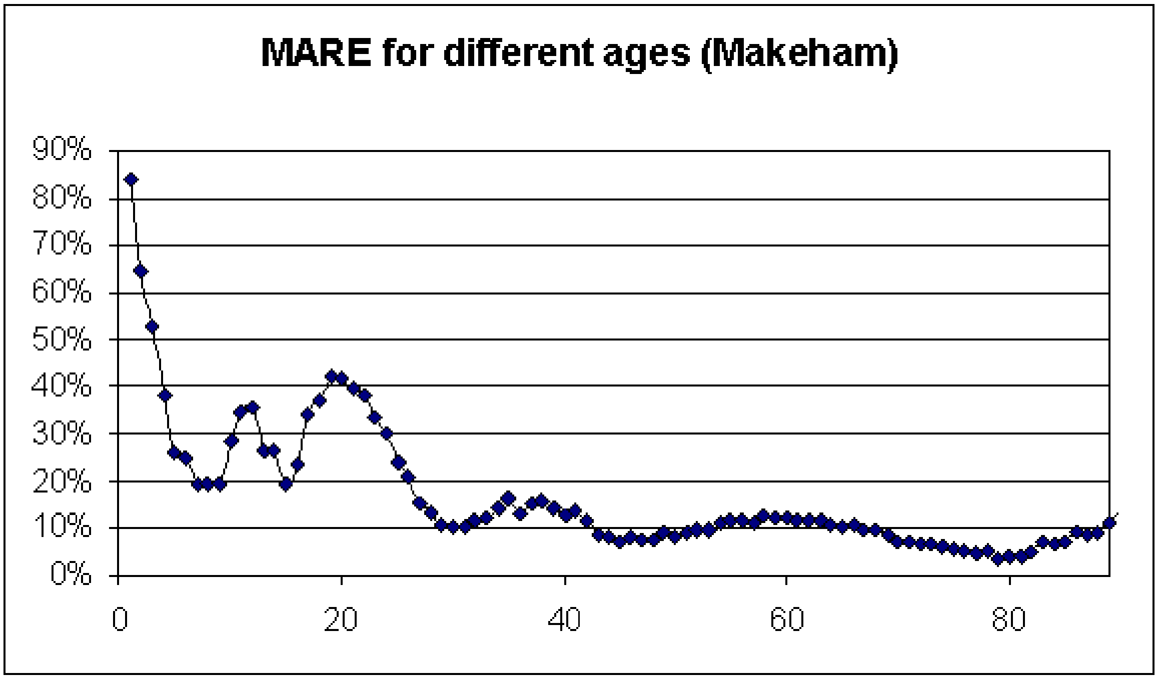

Figure 1. *Mean Absolute Relative Error (MARE) of the Stochastic Makeham model w.r.t. the data for the different ages used in estimation. The model performs well from 25 years and older. We expect the Thiele model to correct for some of the error in the Makeham specification for the lower ages.*

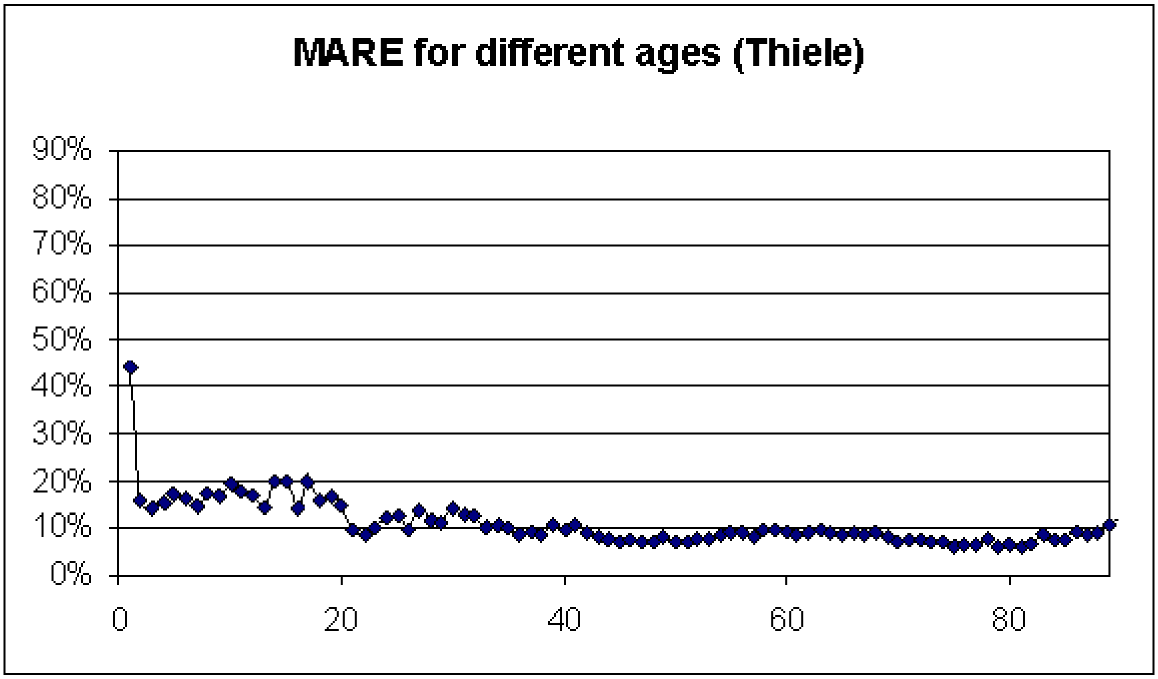

Figure 2. *Mean Absolute Relative Error (MARE) of the Stochastic Thiele model w.r.t. the data for the different ages used in estimation. The model performs well for all ages. The additional terms for mortality at young ages and the middle age "hump" correct more than half of the error of the Makeham model for those ages.*

Age	%	Wage	Wage growth	Pension
25	15	20,000	6%	205,714
30	15	25,000	5%	137,900
35	15	30,000	5%	129,658
40	15	50,000	4%	133,292
45	15	60,000	2.5%	98,317
50	15	60,000	2.5%	86,898
55	15	60,000	2.5%	76,805
60	10	60,000	2.5%	67,884

Table 3. *Data on pension fund participants.*