ARC Centre of Excellence in Population Ageing Research

Working Paper 2011/7

Consistent Dynamic Affine Mortality Models for Longevity Risk Applications

Craig Blackburn and Michael Sherris*

* Blackburn is a doctoral student in the School of Risk and Actuarial at the University of New South Wales (UNSW) and the ARC Centre of Excellence in Population Ageing Research (CEPAR). Sherris is Professor of Actuarial Studies at UNSW and a CEPAR Chief Investigator.

Consistent Dynamic Affine Mortality Models for Longevity Risk Applications

Craig Blackburn
School of Actuarial Studies
University of New South Wales
c.blackburn@unsw.edu.au

Michael Sherris
School of Actuarial Studies and CEPAR
University of New South Wales
m.sherris@unsw.edu.au

May 5, 2011

Abstract

This paper proposes and assesses consistent multi-factor dynamic affine mortality models for longevity risk applications. The dynamics of the model produce closed-form expressions for survival curves. The framework includes an arbitrage-free model specification. Importantly, the mortality model provides consistent future survival curves with the same parametric form as the initial curve. There are multiple risk factors allowing applications to hedging and pricing mortality and longevity bonds, mortality derivatives and more general risk management problems. A state-space representation is used to estimate parameters for the model with the Kalman filter. The state-space form provides a separate measurement and transition system of equations. A measurement error variance is included for each age to capture the effect of sample population size. The transition system dynamics capture the stochastic properties of the underlying mortality rate. Swedish mortality data is used to assess 2- and 3-factor implementations of the model. A 3-factor model specification is shown to provide a good fit to the observed survival curves especially for older ages, and performs better than the 2-factor models. Bootstrapping is used to derive model parameter estimate distributions. Residual analysis is used to confirm model fit. Consistent models are shown to improve model performance and stability.

Keywords: Mortality model, longevity risk, multi-factor, affine, arbitrage-free, consistent, Kalman filter, Swedish mortality

JEL Classifications: G12, G22, G23, C13, C51, C52, J11

1 Introduction

This paper proposes and assesses consistent multi-factor dynamic affine mortality models for modelling longevity risk based on the Affine Term Structure Model (ATSM), Duffie and Kan (1996), that is used extensively in interest rate modelling. The mortality survival curve age structure is modelled in a similar manner to the term structure of interest rates. Other authors have proposed and applied mortality models with an affine structure. Dahl (2004) presents an affine mortality structure for a cohort of lives of the same age. Schrager (2006) develops an affine mortality intensity model for the Thiele and Makeham mortality laws and considers all ages simultaneously. They do do not consider the consistency property of the models.

The parametric survival curves proposed in this paper are consistent ensuring the admissible set of forecast survival curves have the same parametric form as the original fitted parametric curve. Models will need re-parameterizing at future dates if they do not have the consistency property. If the model is rich enough to capture the dynamics of mortality rates then this will be unnecessary. As a result the model provides a more reliable basis for hedging and pricing longevity risk transactions. Bjork and Christensen (1999) establishes conditions for consistency. De Rossi (2004) applies these conditions to derive a consistent 2-factor ATSM for interest rate modelling.

The model framework used has an arbitrage-free formulation allowing a calibration of market prices of risk to relevant and available market data. The model's factor dynamics are defined under a risk-neutral measure for the risk-adjusted survival curve. Instead of assuming a constant price of mortality risk for each factor, as in the "completely affine" model of Duffie and Kan (1996), an "essentially affine" model introduced by Duffee (2002) is adopted.

The modelled survival curve is exponentially affine in the stochastic factors and factor loadings. The survival curve factor dynamics assume a mean-reverting process. Use of mean-reverting dynamics in fitting and forecasting survival curves has been questioned, Luciano and Vigna (2005). Historical data for mortality rates shows evidence of mean reversion, Njenga and Sherris (2009). The mean-reverting assumption is assessed when the models are fitted to historical data.

Estimation of the model uses a state-space form, with measurement and transition systems. The model parameters are estimated using the Kalman filter, Kalman (1960), to obtain a likelihood for a given set of parameters based on historical data. The state-space representation has a measurement system for the parametric model of the observed survival curve. An explicit allowance is included for an age-dependent measurement error covariance matrix. Measurement errors are assumed independent between ages. A parametric function of age is used for the errors. The transition system for the latent stochastic factors driving the survival curve dynamics over time is a multi-factor Gaussian process. Although this allows the possibility of negative mortality rates, which is also the case for interest rate models, the fitted parameter values limit this happening.

The model framework allows closed-form expressions for risk-adjusted survival curves. This is an advantage over popular econometric mortality pricing models. The specification of the price of mortality risk allows different drift terms under the real-world, $\mathbb{P}$-measure, and risk-neutral, $\mathbb{Q}$-measure, dynamics. Market traded longevity pricing information is not currently available, so the price of mortality risk structure is not readily calibrated. As mortality swaps and other mortality based pricing information becomes available, the model is readily calibrated to market data. The number of traded instruments required is the same as the number of latent factors.

Swedish data is used to calibrate and assess the model fit. The exponential survival curve assumption is shown to be robust to the data fitting period. A state-space resampling method, STOFFER and WALL (2004), is used to produce bootstrap distribution confidence intervals for the parameter estimates. Under the model assumptions the bootstrapped distribution for each parameter has an asymptotically normal distribution.

Different model assumptions are assessed including 2- and 3-factor models with and without factor dependence. It is shown that 2-factor models are able to fit the data up to the age of 85, while the 3-factor models are able to capture the majority of data variation for the whole data age range from 50 to 99. The 3-factor model improves the fit over the 2-factor models. Consistent dynamic mortality models are shown to perform better than a model based on a modification of the Nelson-Siegel curve used in interest rate modelling which is not consistent.

Section 2 outlines the survival curve modelled and summarizes the Swedish data used to assess the models. In Section 3, the concept of consistency is introduced and illustrated with the popular Nelson-Siegel curve in interest rate yield curve models. Section 4 presents an outline of the derivation of the models with a 2-factor model used for illustration. The models that are considered in the paper are then summarized in Section 5. Section 6 outlines the Kalman filter estimation of the model parameters. Section 7 provides an analysis of the results from fitting the models with Swedish mortality data. Section 8 concludes with a summary of the paper.

2 Mortality survival curves

The model for survival curves used has a functional form that is a weighted sum of exponentials with parameters that can be time varying. This is motivated by the popular use of the Nelson-Siegel formula used to fit yield curves for interest rates. The mortality equivalent of a yield curve of zero coupon bond yields for different maturities is the curve of the average of the survival probabilities for different survival times.

The survival probability under the risk neutral probability for an individual aged x at time t to age $x+T$, hence surviving another $\tau = T - t$ years, in the affine framework has the form

$$S_{x,t}(\tau) = E_t^Q \left[e^{-\int_0^\tau \mu_{x+u,t} du} \middle| \mathbb{F}_t \right] = e^{-B(\tau)' Z_t + C(\tau)} \tag{1}$$

where $\mu_{x+u,t}$ is the force of mortality for an individual aged $x+u$ at time t and Z_t is a vector of random factors. For an n-factor model, $B(\tau)$ is a matrix of the n-factor loadings and $C(\tau)$ is a constant. The log of the survival probability is affine in the random factors driving changes in mortality through time.

The maturity curve of survival probabilities equivalent to zero coupon yields to maturity in interest rate models is given by

$$M_{x,t}(\tau) = -\frac{1}{\tau} ln[S_{x,t}(\tau)]$$

which is the average force of mortality for survival for τ years for a person aged x at time t. The $M_{x,t}(\tau)$ is used for model fitting.

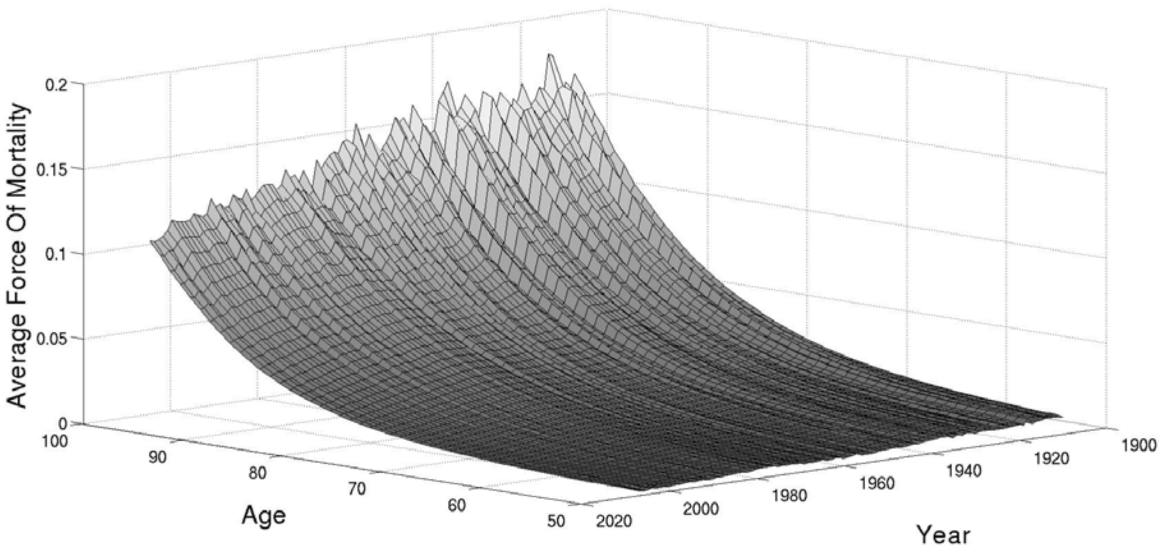


Figure 1: Age-Period average force of mortality for Swedish males ages 50-99 for 1910 to 2007

The survival probability determined from mortality data is

$$S_{x,t}(\tau) = \prod_{s=0}^{\tau-1}(1 - q_{x+s,t})$$
$$q_{x,t} = 1 - e^{-m_{x,t}}$$

where the one year death probability is $q_{x,t}$, for a person aged x in year t with and $m_{x,t}$ is the central rate of mortality. Which is calculated from mortality data as

$$m_{x,t} = \frac{D_{x,t}}{E_{x,t}} = \frac{\text{\# of deaths aged x in year t}}{\text{Exposure aged x in year t}}.$$

The models are fitted and assessed using raw deaths and population estimates for Swedish mortality. Population exposures and number of deaths data is taken from the Human Mortality Database for Sweden for the years 1910 to 2007. Sweden has been recording deaths by individual ages since 1861 and cohort information since 1901. Although Swedish deaths data is generally reliable, the Human Mortality Database has a method's protocol, Wilmoth *et al.* (2007), that describes the various smoothing techniques used for the life-table data. Deaths data is reliable whereas population exposures by age are mostly estimates. In the first half of the 20th century, accurate population estimates were only available for census years. For Sweden, this was every ten years. Since the 1970s a registry system has been used to track population estimates on a yearly basis. Population estimates between census years are estimates. As a result the mortality rates include smoothing of the population estimates.

Figure 1 plots the average force of mortality for different survival times for Swedish males between the ages of 50 and 99 for the years 1910 to 2007. This is the mortality curve data used for our model fitting. The Figure shows clearly the improvement in

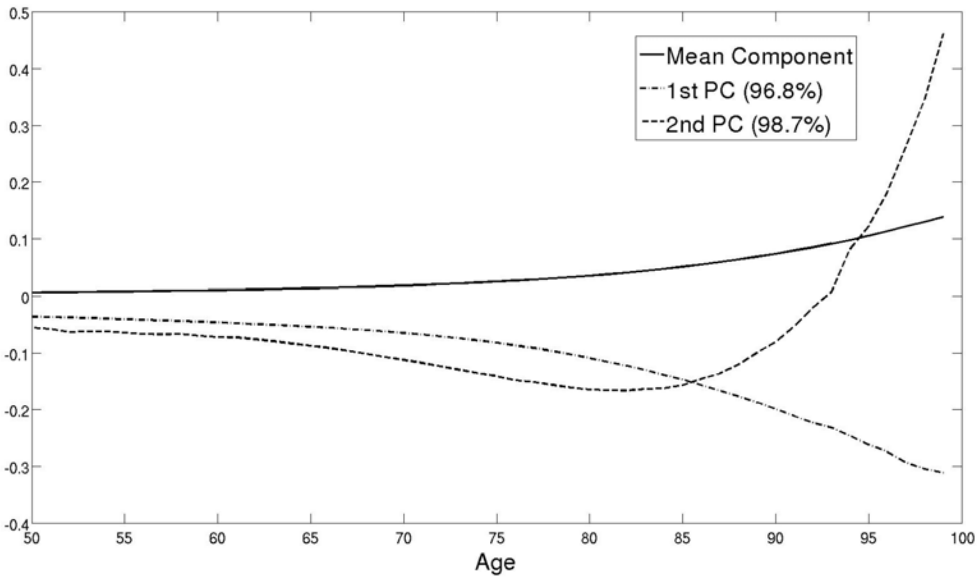


Figure 2: Principal Component Analysis for Swedish males ages 50-99 for 1910 to 2007

mortality over the course of the 20th century and an exponential shape of the curve. Mortality improvements have occurred at differing rates for different ages.

Figure 2 gives the mean and first two principal components of the average force of mortality survival curve for the Swedish data. The mean and first principal component explains 96.8% of the variation in the mortality curve data. This suggests that 2- and 3-factor mortality models should parsimoniously capture variations in observed mortality survival curves. The principal component assumptions are that the mean is constant and the principal components are orthogonal. The framework in this paper assumes a mean equivalent that is a dynamic factor and allows interaction between factors under the risk neutral, or risk adjusted, measure dynamics.

3 Consistent Survival Curve Dynamics

A main aim of this paper is to propose and assess consistent models for the survival probability curve. These maintain the same functional form of the survival curve across time. By assessing different model assumptions we show that these models perform well across time. Mortality model consistency is defined in a similar way to Bjork and Christensen (1999) for interest rate models. The dynamics of the underlying mortality rate are determined to ensure consistency in the model.

Because of its popularity in interest rate modelling, a version of the Nelson-Siegel model in Christensen *et al.* (2009) is used as one of the models for mortality survival curves. We use this model in this section to demonstrate the concept of consistency. Because the Nelson-Siegel model is not consistent it will provide a contrast with our consistent 2- and 3-factor models.

The risk factor dynamics under the risk-neutral $\mathbb{Q}$-measure of the force of mortality curve is

$$d\mu_{x,t}(\tau) = \nu_{x,t}(\tau)dt + \Sigma_{x,t}(\tau)dW^Q$$

where $\nu_{x,t}(\tau)$ is the drift vector, $\Sigma_{x,t}(\tau)$ is the vector of volatilities and dW^Q is a vector of Brownian motions.

Under these dynamics the risk-neutral price for \$1 at time t paid on survival to time T for an individual currently aged x is

$$S_{x,t}(\tau) = e^{-B(\tau)' Z_t + C(\tau)}$$

where $B(\tau)$ and $C(\tau)$ are the solutions to a set of ordinary differential equations and Z_t are the latent risk factors at time t driving the dynamics of the mortality curve.

The stochastic differential equations for the risk factors of the 3-factor Nelson-Siegel model in Christensen *et al.* (2009) under the $\mathbb{Q}$ measure are

$$\begin{pmatrix} dZ_t^1 \\ dZ_t^2 \\ dZ_t^3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[\begin{pmatrix} \mu_1^Q \\ \mu_2^Q \\ \mu_3^Q \end{pmatrix} - \begin{pmatrix} Z_t^1 \\ Z_t^2 \\ Z_t^3 \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix}$$

where $\lambda, \mu_1^Q, \mu_2^Q, \mu_3^Q, \sigma_1, \sigma_2, \sigma_3$ are parameters determining the drift and volatility.

The force of mortality dynamics can be written as

$$d\mu_{x,t}(\tau) = \nu(Z_t^1, Z_t^2, Z_t^3, t)dt + \sigma_1 dW_t^1 + \sigma_2 e^{-\lambda(T-t)} dW_t^2 + \sigma_3 \lambda e^{-\lambda(T-t)} dW_t^3.$$

The solution for the factor loadings are exponential functions given by

$$\begin{aligned} B^1(\tau) &= \tau \\ B^2(\tau) &= \frac{1 - e^{-\lambda\tau}}{\lambda} \\ B^3(\tau) &= \frac{1 - e^{-\lambda\tau}}{\lambda} - \tau e^{-\lambda\tau} \end{aligned}$$

The solution for $C(\tau)$ is found in Christensen *et al.* (2009).

The force of mortality of the model is then given by

$$\mu_{x,t}(\tau) = -\frac{\partial \left[\log S_{x,t}(\tau)\right]}{\partial T}$$

Theorem 4.1 of Bjork and Christensen (1999) is directly applied to the Fréchet derivatives of $\mu_{x,t}(\tau)$ to determine consistency. These derivatives are used to verify drift and volatility conditions for the underlying risk factor dynamics that ensure consistency. The conditions are based on manifolds for future survival curves and Fréchet derivatives of the mortality rate dynamics and full details are found in Bjork and Christensen (1999).

The consistent drift condition requires the following for the drift:

$$\mu_\tau(\tau, z) + \sigma(\tau) \left(\int_0^\tau \sigma(y) dy \right)^T \in Im[\mu_z(\tau, z)]$$

where $\mu_\tau(\tau, z)$ and $\mu_z(\tau, z)$ are Fréchet derivatives with respect to τ and z respectively, where the x, t subscript has been dropped for convenience of notation. The volatility term in the drift can be written as

$$\sigma(\tau)\left(\int_0^\tau \sigma(y)dy\right)^T$$
$$= \begin{bmatrix}\sigma_1 & \sigma_2 e^{-\lambda\tau} & +\sigma_3\lambda e^{-\lambda\tau}\end{bmatrix}\begin{bmatrix}\sigma_1\tau & \sigma_2\frac{(1-e^{-\lambda\tau})}{\lambda} & +\sigma_3\left(\frac{(1-e^{-\lambda\tau})}{\lambda}-\tau e^{-\lambda\tau}\right)\end{bmatrix}^T$$
$$= \sigma_1^2\tau + \sigma_2^2\frac{(e^{-\lambda\tau}-e^{-2\lambda\tau})}{\lambda} + \sigma_3^2\left(\tau\frac{(e^{-\lambda\tau}-e^{-2\lambda\tau})}{\lambda} - \tau^2\lambda e^{-2\lambda\tau}\right)$$

The force of mortality curve can be simplified in terms of time varying parameters denoted by z and exponential terms in τ as

$$\mu_{x,t}(\tau,z) = z_1 + z_2 e^{-\lambda\tau} + z_3\tau e^{-\lambda\tau} + z_4 e^{-2\lambda\tau} + z_5\tau e^{-2\lambda\tau} + z_6\tau^2 e^{-2\lambda\tau} + z_7\tau^2$$

The Fréchet derivatives of $\mu_{x,t}(\tau,z)$ are

$$\mu_z(\tau,z) = \begin{bmatrix}1 & e^{-\lambda\tau} & \tau e^{-\lambda\tau} & e^{-2\lambda\tau} & xe^{-2\lambda\tau} & \tau^2 e^{-2\lambda\tau} & \tau^2\end{bmatrix}$$
$$\mu_\tau(\tau,z) = (z_3-\lambda z_2)e^{-\lambda\tau} - \lambda z_3\tau e^{-\lambda\tau} + (z_5-2\lambda z_4)e^{-2\lambda\tau} + (2z_6-2\lambda z_5)\tau e^{-2\lambda\tau} - 2\lambda z_6\tau^2 e^{-2\lambda\tau} + z_7 2\tau$$

The consistent drift requirement can be written as

$$\mu_\tau(\tau,z) + \gamma_1 e^{-\lambda\tau} + \gamma_2\tau e^{-\lambda\tau} + \gamma_3 e^{-2\lambda\tau} + \gamma_4\tau e^{-2\lambda\tau} + \gamma_5\tau^2 e^{-2\lambda\tau} + \gamma_6\tau \in Im[\mu_z(\tau,z)] \tag{2}$$

where γ_i are constants after collecting terms in τ.

The vector of volatilities for $d\mu_{x,t}(\tau,z)$ is

$$\sigma(\tau) = \begin{bmatrix}\sigma_1 & \sigma_2 e^{-\lambda\tau} & \sigma_3\lambda e^{-\lambda\tau}\end{bmatrix}$$

It then follows from (2) that since $\gamma_6\tau \notin Im[\mu_z(\tau,z)]$ the Nelson Siegel survival curves are not consistent.

4 Model Derivation

The model derivation for the survival curve and mortality rate dynamics follow the standard affine term structure model defined by Duffie and Kan (1996). We summarise the 2-factor model derivation for completeness. Mortality rates, and the survival curve, are driven by unobserved (latent) risk factors following n-factor stochastic differential equations under the risk-neutral $\mathbb{Q}$-measure

$$dZ_t = \Delta^Q\left[\Theta^Q - Z_t\right]dt + \Sigma\sqrt{v(Z)}dW_t^Q$$

where

$$v(Z) = \begin{pmatrix}\alpha_1+\beta_1 Z_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \alpha_n+\beta_n Z_n\end{pmatrix}$$

with $\Delta^Q \in \mathbb{R}^{n\times n}$, $\Theta^Q \in \mathbb{R}^{n\times 1}$, $\Sigma \in \mathbb{R}^{n\times n}$.

The risk-adjusted survivor index is given by

$$S_{x,t}(\tau) = E_t^Q\left[e^{-\int_0^\tau \mu_{x+u,t}du}\middle|\mathbb{F}_t\right] = e^{-B(\tau)'Z_t+C(\tau)} \tag{3}$$

where $\tau = T - t$. This is the survival probability for an individual aged x at time t to age $x + T$ under the risk neutral probability or the price of \$1 paid on survival to age $x + T$. For an n-factor model, $B(\tau)$ is a matrix of the n-factor loadings and $C(\tau)$ is a constant.

In order to use Equation (3) to determine $B(\tau)$ and $C(\tau)$ either the Martingale approach or the Feynman-Kac theorem can be applied to derive Partial Differential Equations (PDEs). The derivation here is based on the Martingale method under the $\mathbb{Q}$-measure. The "essentially affine" price of risk structure is used to transform the model dynamics to the $\mathbb{P}$-measure.

Define a $\mathbb{Q}$-Martingale Y by

$$Y_{x,t}(\tau) := E_t^Q\left[e^{-\int_0^T \mu_{x+u,t}du}\middle|\mathcal{F}_t\right] = e^{-\int_0^t \mu_{x+u,t}du}S_{x,t}(\tau). \tag{4}$$

Using the multi-dimensional Ito's lemma, with independent Brownian Motions, the dynamics of Y are

$$\begin{aligned} dY_{x,t}(\tau) &= \frac{\partial Y_{x,t}(\tau)}{\partial t}dt + \sum_{i=1}^n \frac{\partial Y_{x,t}(\tau)}{\partial Z_i}dZ_i + \frac{1}{2}\sum_{j=1}^n \frac{\partial^2 Y_{x,t}(\tau)}{\partial^2 Z_j}d[Z_jZ_j] \\ &= \left(\frac{\partial Y_{x,t}(\tau)}{\partial t} + \sum_{i=1}^n \frac{\partial Y_{x,t}(\tau)}{\partial Z_i}\Delta^Q\left[\Theta^Q - Z_i\right] + \frac{1}{2}\sum_{j=1}^n \frac{\partial^2 Y_{x,t}(\tau)}{\partial^2 Z_j}\sigma_{jj}^2 v_j(Z_j)\right)dt \\ &\quad + \frac{\partial Y_{x,t}(\tau)}{\partial Z}\Sigma\sqrt{v(Z)}dW^Q \end{aligned} \tag{5}$$

Since Y is a Martingale, the drift component of equation (5) must equal zero. The partial derivatives from equation (4) are

$$\begin{aligned} \frac{\partial Y_{x,t}(\tau)}{\partial t} &= -\mu_{x,t}e^{-\int_0^t \mu_{x+u,t}du}S_{x,t}(\tau) + e^{-\int_0^t \mu_{x+u,t}du}\frac{\partial S_{x,t}(\tau)}{\partial t} \\ \frac{\partial Y_{x,t}(\tau)}{\partial Z_i} &= e^{-\int_0^t \mu_{x+u,t}du}\frac{\partial S_{x,t}(\tau)}{\partial Z_i} \\ \frac{\partial^2 Y_{x,t}(\tau)}{\partial^2 Z_j} &= e^{-\int_0^t \mu_{x+u,t}du}\frac{\partial^2 S_{x,t}(\tau)}{\partial Z_i\partial Z_i} \end{aligned}$$

Upon substitution in equation (5) and equating the drift term to zero, the partial differential equation can be written

$$\left(\frac{\partial S_{x,t}(\tau)}{\partial t} + \sum_{i=1}^n \frac{\partial S_{x,t}(\tau)}{\partial Z_i}\Delta^Q\left[\Theta^Q - Z_i\right] + \frac{1}{2}\sum_{j=1}^n \frac{\partial^2 S_{x,t}(\tau)}{\partial^2 Z_j\sigma_{jj}^2}v_j(Z_j)\right) - \mu_{x,t}S_{x,t}(\tau) = 0 \tag{6}$$

The risk-adjusted survivor curve partial derivatives are

$$\begin{aligned}
S_{x,t}(\tau) &= e^{-B(\tau)'Z_t+C(\tau)} \\
\frac{\partial S_{x,t}(\tau)}{\partial t} &= \left(-\dot{C}(\tau) + \dot{B}(\tau)'Z\right) S_{x,t}(\tau) \\
\frac{\partial S_{x,t}(\tau)}{\partial Z} &= -B(\tau) S_{x,t}(\tau) \\
\frac{\partial^2 S_{x,t}(\tau)}{\partial^2 Z} &= B(\tau)'B(\tau)S_{x,t}(\tau)
\end{aligned}$$

where $\dot{C} = \frac{\partial C}{\partial t}$ and $B(\tau)'$ is the transpose of $B(\tau)$.

Substituting these into equation (6) gives the general form of the PDE.

$$\left(-\dot{C}(\tau) + B(\tau)'Z\right) - \sum_{i=1}^{n} B(\tau)'\Delta^Q \left[\Theta^Q - Z_i\right] + \frac{1}{2}\sum_{j=1}^{n} \left(\Sigma' B(\tau) v(Z) B(\tau)'\Sigma\right)_{jj} - \mu_{x,t}(0) = 0$$

with the instantaneous force of mortality

$$\mu_{x,t}(0) = \rho_0 + \rho Z = \lim_{\tau \to 0} \mu_{x,t}(\tau) \tag{7}$$

The PDE can be rearranged into two ordinary differential equations in terms of $\dot{B}(\tau)$ and $\dot{C}(\tau)$,

$$\begin{aligned}
\dot{B}(\tau) &= \rho + \Delta^Q B(\tau) - \frac{1}{2}\sum_{j=1}^{n} \Sigma' B(\tau) B(\tau)'\Sigma\beta_j, \quad B(0) = 0 \\
\dot{C}(\tau) &= \rho_0 - B(\tau)'\Delta^Q\Theta^Q - \frac{1}{2}\sum_{j=1}^{n} \Sigma' B(\tau) B(\tau)'\Sigma\alpha_j, \quad C(0) = 0
\end{aligned}$$

Christensen *et al.* (2009) shows, by an appropriate selection of ρ and Δ^Q, how an arbitrage-free version of the Nelson-Siegel model is derived. Schrager (2006) specified ρ so that $\mu_{x,t}$ was a Gaussian stochastic version of the Thiele mortality model under the $\mathbb{P}$-measure.

We develop the model dynamics in the $\mathbb{Q}$-measure and solve the ordinary differential equations with ρ a constant in our derivation.

So far the derivation has been under the $\mathbb{Q}$-measure, but mortality rates are observed under the $\mathbb{P}$-measure. We adopt the "essentially affine" model of Duffee (2002) which allows us to estimate drift terms for the real-world stochastic processes and modify the drift to obtain the risk-neutral dynamics.

The risk-neutral stochastic process for the model factors are

$$dZ_t = \Delta^Q \left[\Theta^Q - Z_t\right] dt + \Sigma V(Z_t) dW_t^Q$$

In the "essentially affine" model the market price of mortality risk is

$$\Lambda_t = V(Z_t)\lambda_0 + V^-(Z_t)\lambda Z_t$$

where $\lambda_0 \in \mathbb{R}^{n\times 1}$ and $\lambda \in \mathbb{R}^{n\times n}$. $V(Z_t)^-$ is a matrix of the same order as $V(Z_t)$ and given by

$$V^-(Z_t) = \begin{pmatrix} (\alpha_1 + \beta_1 Z_t)^{-1/2} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (\alpha_n + \beta_n Z_t)^{-1/2} \end{pmatrix}$$

The change of measure is

$$dW_t^Q = dW_t^P + \Lambda_t dt$$

and the stochastic process under the $\mathbb{P}$-measure is

$$dZ_t = \left[\Delta^Q \Theta^Q - \Delta^Q Z_t\right] dt + \Sigma \left[V^2(Z_t)\lambda_0 + I^- \lambda Z_t\right] dt + \Sigma V(Z_t) dW_t^P$$

where I^- is a diagonal matrix that equals one for indices where $V^-(Z_t) \neq 0$. This simplifies to

$$\begin{aligned} dZ_t &= \left[\Delta^Q \Theta^Q + \Sigma \alpha \lambda_0\right] dt - \left[\Delta^Q - \Sigma \beta \lambda_0 - \Sigma I^- \lambda\right] Z_t dt + \Sigma V(Z_t) dW_t^P \\ &= \Delta^P \left[\Theta^P - Z_t\right] dt + \Sigma V(Z_t) dW_t^P \end{aligned}$$

where

$$\alpha = \begin{pmatrix} \alpha_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha_n \end{pmatrix} \qquad \beta = \begin{pmatrix} \beta_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \beta_n \end{pmatrix}$$

Under this change of measure the structure of the price of mortality risk is included. The flexibility in the model is that λ_0 only influences the long term mean while the λ term only influences the speed of mean reversion when $\beta = 0$.

The solutions to the ordinary differential equations are derived under the $\mathbb{Q}$-measure and we estimate the drift term under the $\mathbb{P}$-measure. The $\Sigma V(Z_t)$ terms are the same under both probability measures.

5 Models for Longevity Risk Dynamics

This section summarizes the survival curve models to be assessed. Each model assumes different structure for the dynamics of the latent risk factors driving the survival curve. Dynamics include 2- and 3-factor models both with and without dependence between the factors. A 3-factor version of the Nelson-Siegel dynamics for the survival curve is included for comparison since this is known not to be a consistent model. The other models are consistent.

5.1 Dependent 2-Factor Model

This is a 2-factor affine survival curve model with constant volatility term for each factor and $\alpha_i = 1$ and $\beta_i = 0$. The risk neutral dynamics are

$$\begin{pmatrix} dZ_{1,t} \\ dZ_{2,t} \end{pmatrix} = \begin{pmatrix} \delta_{11} & 0 \\ \delta_{21} & \delta_{22} \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \end{pmatrix} - \begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \begin{pmatrix} dW_{1,t}^Q \\ dW_{2,y}^Q \end{pmatrix}$$

There is dependence in the drift and the volatility. The sign of δ_{11} determines if the factor loading for $Z_{1,t}$ is exponentially increasing or decreasing. The dependence produces a factor loading that can be an exponentially increasing upward sloping, downward sloping or curved shape, similar to the Nelson-Siegel specification. This dependence allows more flexibility in the survival curve fitting data at the older ages.

The instantaneous force of mortality is

$$\mu_{x,t}(0) = \rho_1 Z_{1,t} + \rho_2 Z_{2,t}$$

with $\rho_0 = 0$. The set of ordinary differential equations are

$$\begin{aligned}\frac{dB(\tau)}{dt} &= \rho + \Delta^Q B(\tau), \quad B(0) = 0 \\ \frac{dC(\tau)}{dt} &= -B(\tau)'\Delta^Q\Theta^Q - \frac{1}{2}\sum_{j=1}^{n}\left(\Sigma' B(\tau)B(\tau)'\Sigma\right)_{j,j}, \quad C(0) = 0\end{aligned}$$

In this section of the paper we explicitly include the time index t and write $B(\tau)$ as $B(t,T)$ for clarity of presentation. To solve the differential equation for $B(t,T)$ note that

$$\begin{aligned}\frac{d}{dt}\left[e^{\Delta^Q(T-t)}B(t,T)\right] &= e^{\Delta^Q(T-t)}\frac{dB(t,T)}{dt} - \Delta^Q e^{\Delta^Q(T-t)}B(t,T) \\ &= e^{\Delta^Q(T-t)}\frac{dB(t,T)}{dt} - \left[\frac{dB(t,T)}{dt} - \rho\right]e^{\Delta^Q(T-t)} \\ &= e^{\Delta^Q(T-t)}\rho.\end{aligned}$$

Integrating both sides from t to T gives

$$\int_t^T \frac{d}{ds}\left[e^{\Delta^Q(T-s)}B(t,T)\right] = \int_t^T e^{\Delta^Q(T-s)}\rho ds$$

and using the boundary conditions we have

$$B(t,T) = -e^{-\Delta^Q(T-s)}\int_t^T e^{\Delta^Q(T-s)}\rho ds. \tag{8}$$

For this model assumptions, the matrix exponentials in equation (8) are as follows

$$\begin{aligned}e^{\Delta^Q(T-t)} &= \begin{pmatrix} e^{\delta_{11}(T-t)} & 0 \\ \frac{\delta_{21}}{\delta_{11}-\delta_{22}}\left(e^{\delta_{11}(T-t)} - e^{\delta_{22}(T-t)}\right) & e^{\delta_{22}(T-t)}\end{pmatrix} \\ e^{-\Delta^Q(T-t)} &= \begin{pmatrix} e^{-\delta_{11}(T-t)} & 0 \\ \frac{\delta_{21}}{\delta_{11}-\delta_{22}}\left(e^{-\delta_{11}(T-t)} - e^{-\delta_{22}(T-t)}\right) & e^{-\delta_{22}(T-t)}\end{pmatrix}\end{aligned}$$

Substituting these into equation (8) gives

$$\begin{aligned}B(t,T) = &-\begin{pmatrix} e^{-\delta_{11}(T-t)} & 0 \\ \frac{\delta_{21}}{\delta_{11}-\delta_{22}}\left(e^{-\delta_{11}(T-t)} - e^{-\delta_{22}(T-t)}\right) & e^{-\delta_{22}(T-t)}\end{pmatrix} \times \\ &\int_t^T \begin{pmatrix} e^{\delta_{11}(T-s)} & 0 \\ \frac{\delta_{21}}{\delta_{11}-\delta_{22}}\left(e^{\delta_{11}(T-s)} - e^{\delta_{22}(T-s)}\right) & e^{\delta_{22}(T-s)}\end{pmatrix}\begin{pmatrix}\rho_1 \\ \rho_2\end{pmatrix} ds \\ = &-\begin{pmatrix} e^{-\delta_{11}(T-t)} & 0 \\ \frac{\delta_{21}}{\delta_{11}-\delta_{22}}\left(e^{-\delta_{11}(T-t)} - e^{-\delta_{22}(T-t)}\right) & e^{-\delta_{22}(T-t)}\end{pmatrix} \times \\ &\int_t^T \begin{pmatrix} \rho_1 e^{\delta_{11}(T-s)} \\ \frac{\rho_1\delta_{21}}{\delta_{11}-\delta_{22}}\left(e^{\delta_{11}(T-s)} - e^{\delta_{22}(T-s)}\right) + \rho_2 e^{\delta_{22}(T-s)}\end{pmatrix} ds\end{aligned}$$

Finally simplifying the integral gives

$$B(t,T) = -\begin{pmatrix} e^{-\delta_{11}(T-t)} & 0 \\ \frac{\delta_{21}}{\delta_{11}-\delta_{22}}\left(e^{-\delta_{11}(T-t)} - e^{-\delta_{22}(T-t)}\right) & e^{-\delta_{22}(T-t)} \end{pmatrix} \times$$
$$\begin{pmatrix} -\rho_1 \frac{(1-e^{\delta_{11}(T-t)})}{\delta_{11}} \\ \rho_1 \frac{\delta_{21}}{\delta_{11}-\delta_{22}}\left(\frac{1-e^{\delta_{11}(T-t)}}{-\delta_{11}} - \frac{1-e^{\delta_{22}(T-t)}}{-\delta_{22}}\right) + \rho_2 \frac{1-e^{\delta_{22}(T-t)}}{-\delta_{22}} \end{pmatrix}$$

so that

$$B(t,T) = -\begin{pmatrix} \rho_1 \frac{(1-e^{-\delta_{11}(T-t)})}{\delta_{11}} \\ \rho_1 \frac{\delta_{21}}{\delta_{11}-\delta_{22}}\left(\frac{1-e^{-\delta_{11}(T-t)}}{\delta_{11}} - \frac{1-e^{-\delta_{22}(T-t)}}{\delta_{22}}\right) + \rho_2 \frac{1-e^{-\delta_{22}(T-t)}}{\delta_{22}} \end{pmatrix} \tag{9}$$

For the constant term, $C(t,T)$, assuming that the long term mean parameter Θ^Q is equal to zero we have

$$\begin{aligned} C(t,T) &= -\frac{1}{2}\int_t^T \sum_{j=1}^{n} \left(\Sigma' B(s,T) B(s,T)' \Sigma\right)_{jj} ds \\ &= -\frac{1}{2}\int_t^T \Big[(\sigma_{11}^2 + \sigma_{12}^2) B_1(s,T)^2 + (2\sigma_{11}\sigma_{21} + 2\sigma_{12}\sigma_{22}) B_1(s,T) B_2(s,T) \\ &\quad + (\sigma_{21}^2 + \sigma_{22}^2) B_2(s,T)\Big] ds \end{aligned} \tag{10}$$

There are only 3 integrals in terms of $B(t,T)$ with 4 unknown Σ parameters. This means one of the Σ parameters will not be identifiable. Setting $\sigma_{12} = 0$ Equation (10) becomes

$$\begin{aligned} C(t,T) &= -\frac{1}{2}\int_t^T \Big[\sigma_{11}^2 B_1(s,T)^2 + 2\sigma_{11}\sigma_{21} B_1(s,T) B_2(s,T) + (\sigma_{21}^2 + \sigma_{22}^2) B_2(s,T)^2\Big] ds \\ &= -\frac{1}{2}\int_t^T \Bigg[\sigma_{11}^2 \Big(\rho_1 \frac{(1 - e^{-\delta_{11}(T-s)})}{\delta_{11}}\Big)^2 + 2\sigma_{11}\sigma_{21}\Big(\rho_1 \frac{(1 - e^{-\delta_{11}(T-s)})}{\delta_{11}}\Big) \times \\ &\Bigg(\rho_1 \frac{\delta_{21}}{\delta_{11} - \delta_{22}}\Bigg(\frac{1 - e^{-\delta_{11}(T-s)}}{\delta_{11}} - \frac{1 - e^{-\delta_{22}(T-s)}}{\delta 22}\Bigg) + \rho_2 \frac{1 - e^{-\delta_{22}(T-s)}}{\delta_{22}}\Bigg) + \\ &(\sigma_{21}^2 + \sigma_{22}^2)\Bigg(\rho_1 \frac{\delta_{21}}{\delta_{11} - \delta_{22}}\Bigg(\frac{1 - e^{-\delta_{11}(T-s)}}{\delta_{11}} - \frac{1 - e^{-\delta_{22}(T-s)}}{\delta 22}\Bigg) + \rho_2 \frac{1 - e^{-\delta_{22}(T-s)}}{\delta_{22}}\Bigg)^2\Bigg] ds \end{aligned}$$

Simplifying the form of $C(t,T)$ in terms of $(T-t)$ with other terms combined into constants, ξ_i we have

$$\begin{aligned} C(t,T) = \xi_0 + \xi_1(T-t) + \xi_2 \frac{1 - e^{-\delta_{11}(T-t)}}{\delta_{11}} + \xi_3 \frac{1 - e^{-2\delta_{11}(T-t)}}{2\delta_{11}} + \xi_4 \frac{1 - e^{-\delta_{22}(T-t)}}{\delta_{22}} + \\ \xi_5 \frac{1 - e^{-2\delta_{22}(T-t)}}{2\delta_{22}} + \xi_6 \frac{1 - e^{-(\delta_{11}+\delta_{22})(T-t)}}{\delta_{11} + \delta_{22}} \end{aligned} \tag{11}$$

Applying the consistent drift and volatility conditions shows that the model is consistent.

5.2 Independent 2-Factor Model

A 2-factor model with no factor dependence and no interaction of the Brownian motions is also considered. The model dynamics are

$$\begin{pmatrix} dZ_{1,t} \\ dZ_{2,t} \end{pmatrix} = \begin{pmatrix} \delta_{11} & 0 \\ 0 & \delta_{22} \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \end{pmatrix} - \begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{pmatrix} \begin{pmatrix} dW_{1,t}^Q \\ dW_{2,y}^Q \end{pmatrix}$$

with the instantaneous force of mortality as

$$\mu_{x,t}(0) = Z_{1,t} + Z_{2,t}$$

This model has 4 fewer parameters compared to the dependent model. The solutions for the factor loadings for this independent model are

$$B_1(t,T) = -\frac{1 - e^{-\delta_{11}(T-t)}}{\delta_{11}} \qquad B_2(t,T) = -\frac{1 - e^{-\delta_{22}(T-t)}}{\delta_{22}} \tag{12}$$

and

$$\begin{aligned} C(t,T) &= -\frac{1}{2} \int_t^T \sum_{j=1}^n \left(\Sigma' B(s,T) B(s,T)' \Sigma \right)_{j,j} ds \\ &= -\frac{1}{2} \int_t^T \left[(\sigma_{11}^2) B_1(s,T)^2 + (\sigma_{22}^2) B_2(s,T)^2 \right] ds \\ &= \frac{1}{2} \left[\frac{\sigma_{11}^2}{\delta_{11}^3} \left(2e^{-\delta_{11}(T-t)} - \frac{1}{2} e^{-2\delta_{11}(T-t)} + \delta_{11}(T-t) - \frac{3}{2} \right) \right. \\ &\qquad \left. \frac{\sigma_{22}^2}{\delta_{22}^3} \left(2e^{-\delta_{22}(T-t)} - \frac{1}{2} e^{-2\delta_{22}(T-t)} + \delta_{22}(T-t) - \frac{3}{2} \right) \right] \end{aligned} \tag{13}$$

This is similar in form to equation (11), with the ξ_6 term zero since there is no interaction between the factors. The model dynamics are also consistent.

5.3 2-Factor Model - Real World Measure

Estimation uses historical data under the $\mathbb{P}$-measure. The 2-factor model under the real-world $\mathbb{P}$-measure is based on the dynamics

$$dZ_t = K^P \left[\mu^P - Z_t \right] dt + \Sigma V(Z_t) dW_t^P \tag{14}$$

which in the 2-factor case is

$$\begin{pmatrix} dZ_{1,t} \\ dZ_{2,t} \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & 0 \\ 0 & \kappa_{22}^P \end{pmatrix} \left[\begin{pmatrix} \mu_1^P \\ \mu_2^P \end{pmatrix} - \begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \begin{pmatrix} dW_{1,t}^P \\ dW_{2,y}^P \end{pmatrix}$$

This differs from the risk-neutral $\mathbb{Q}$-measure dynamics. Under the $\mathbb{P}$-measure interaction between the factors is not included and the "essentially affine" specification of the price of mortality risk allows different drift terms in both measures. In the independent model the σ_{21} parameter is zero.

The survival curve estimated under the real-world measure is

$$M_{x,t}(\tau) = -\frac{1}{\tau} ln[S_{x,t}(\tau)] = -\frac{B_1(\tau)}{\tau} Z_{1,t} - \frac{B_2(\tau)}{\tau} Z_{2,t} + \frac{C(\tau)}{\tau}$$

which for the dependent model is

$$M_{x,t}(\tau) = \left[\rho_1 \frac{1-e^{-\delta_{11}(T-t)}}{\delta_{11}}\right] \frac{Z_{1,t}}{\tau} + \left[\rho_1 \frac{\delta_{21}}{\delta_{11}-\delta_{22}} \left(\frac{1-e^{-\delta_{11}(T-t)}}{\delta_{11}} - \frac{1-e^{-\delta_{22}(T-t)}}{\delta 22}\right) + \rho_2 \frac{1-e^{-\delta_{22}(T-t)}}{\delta_{22}}\right] \frac{Z_{2,t}}{\tau} - \frac{C(\tau)}{\tau}$$

and for the independent model is

$$M_{x,t}(\tau) = \left[\frac{1-e^{-\delta_{11}(T-t)}}{\delta_{11}}\right] \frac{Z_{1,t}}{\tau} + \left[\frac{1-e^{-\delta_{22}(T-t)}}{\delta_{22}}\right] \frac{Z_{2,t}}{\tau} - \frac{C(\tau)}{\tau}$$

The constant term, $C(\tau)$, is given by equation (10) and equation (13) for the dependent and independent models respectively.

5.4 3-Factor Model

For the 3-factor models we assess; a consistent dependent factor model, a consistent independent factor model and a Nelson-Siegel model, which is known not to be consistent. The factor dynamics of each of the models under the $\mathbb{Q}$-measure are summarized here.

- Dependent 3-factor model. This model adds an additional independent factor to the 2-factor dependent model. The dynamics are

$$\begin{pmatrix} dZ_{1,t} \\ dZ_{2,t} \\ dZ_{3,t} \end{pmatrix} = \begin{pmatrix} \delta_{11}^Q & 0 & 0 \\ 0 & \delta_{22}^Q & 0 \\ 0 & \delta_{32}^Q & \delta_{33}^Q \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \\ \theta_3^Q \end{pmatrix} - \begin{pmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \end{pmatrix}\right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} dW_{1,t}^Q \\ dW_{2,y}^Q \\ dW_{3,y}^Q \end{pmatrix} \tag{15}$$

- Independent 3-factor model. Each of the factors are independent and there is no correlation of the Brownian motions.

$$\begin{pmatrix} dZ_{1,t} \\ dZ_{2,t} \\ dZ_{3,t} \end{pmatrix} = \begin{pmatrix} \delta_1^Q & 0 & 0 \\ 0 & \delta_2^Q & 0 \\ 0 & 0 & \delta_3^Q \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \\ \theta_3^Q \end{pmatrix} - \begin{pmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \end{pmatrix}\right] dt + \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} \begin{pmatrix} dW_{1,t}^Q \\ dW_{2,t}^Q \\ dW_{3,t}^Q \end{pmatrix} \tag{16}$$

- Nelson-Siegel model. This is a restricted form of the 3-factor dependent model,

$$\begin{pmatrix} dZ_{1,t} \\ dZ_{2,t} \\ dZ_{3,t} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \delta^Q & 0 \\ 0 & -\delta^Q & \delta^Q \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \\ \theta_3^Q \end{pmatrix} - \begin{pmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \end{pmatrix}\right] dt + \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} \begin{pmatrix} dW_{1,t}^Q \\ dW_{2,y}^Q \\ dW_{3,y}^Q \end{pmatrix} \tag{17}$$

5.5 3-Factor Dependent Model

For this model the instantaneous force of mortality is

$$\mu_{x,t}(0) = \rho_1 Z_{1,t} + \rho_2 Z_{2,t} + \rho_3 Z_{3,t}$$

The solution for $B(t,T)$ is

$$B(t,T) = -\begin{pmatrix} \rho_1 \frac{(1-e^{-\delta_{11}(T-t)})}{\delta_{11}} \\ \rho_2 \frac{(1-e^{-\delta_{22}(T-t)})}{\delta_{22}} \\ \rho_2 \frac{\delta_{32}}{\delta_{22}-\delta_{33}} \left(\frac{1-e^{-\delta_{22}(T-t)}}{\delta_{22}} - \frac{1-e^{-\delta_{33}(T-t)}}{\delta 33} \right) + \rho_3 \frac{1-e^{-\delta_{33}(T-t)}}{\delta_{33}} \end{pmatrix} \tag{18}$$

The constant adjustment term $C(t,T)$ follows as an extension of the 2-factor dependent model,

$$\begin{aligned} C(t,T) = &-\frac{1}{2}\int_t^T \Big[\sigma_{11}^2 B_1(s,T)^2 + \sigma_{22}^2 B_2(s,T)^2 + 2\sigma_{22}\sigma_{32} B_2(s,T) B_3(s,T) + \\ &(\sigma_{32}^2 + \sigma_{33}^2) B_3(s,T)^2\Big] ds \\ = &-\frac{1}{2}\int_t^T \Bigg[\sigma_{11}^2 \Big(\rho_1 \frac{(1-e^{-\delta_{11}(T-s)})}{\delta_{11}}\Big)^2 + \sigma_{22}^2 \Big(\rho_2 \frac{(1-e^{-\delta_{22}(T-s)})}{\delta_{22}}\Big)^2 + \\ &2\sigma_{22}\sigma_{32}\Big(\rho_2 \frac{(1-e^{-\delta_{22}(T-s)})}{\delta_{22}}\Big) \times \left(\rho_2 \frac{\delta_{32}}{\delta_{22}-\delta_{33}} \left(\frac{1-e^{-\delta_{22}(T-s)}}{\delta_{22}} - \frac{1-e^{-\delta_{33}(T-s)}}{\delta 33} \right) + \right. \\ &\left. \rho_3 \frac{1-e^{-\delta_{33}(T-s)}}{\delta_{33}} \right) + (\sigma_{32}^2 + \sigma_{33}^2) \left(\rho_2 \frac{\delta_{32}}{\delta_{22}-\delta_{33}} \left(\frac{1-e^{-\delta_{22}(T-s)}}{\delta_{22}} - \frac{1-e^{-\delta_{33}(T-s)}}{\delta 33} \right) + \right. \\ &\left. \rho_3 \frac{1-e^{-\delta_{33}(T-s)}}{\delta_{33}} \right)^2 \Bigg] ds \end{aligned} \tag{19}$$

5.6 3-Factor Independent Model

The 3-factor independent dynamics under the $\mathbb{Q}$-measure from equation (16) has 5 fewer parameters than the dependent model. The instantaneous force of mortality is

$$\mu_{x,t}(0) = Z_{1,t} + Z_{2,t} + Z_{3,t}$$

The solution to the ordinary differential equations is

$$B_i(t,T) = \frac{1-e^{-\delta_i(T-t)}}{\delta_i} \tag{20}$$

where $i = 1, 2, 3$ for each factor and

$$C(t,T) = \frac{1}{2}\int_t^T \sum_{j=1}^{3} \left(\Sigma^T B(s,T) B(s,T)^T \Sigma\right)_{jj} ds$$

The constant term $C(t, T)$ becomes

$$
\begin{aligned}
C(t,T) = \frac{1}{2}\Bigg[& \frac{\sigma_1^2}{\delta_1^3}\left(2e^{-\delta_1(T-t)} - \frac{1}{2}e^{-2\delta_1(T-t)} + \delta_1(T-t) - \frac{3}{2}\right) \\
& \frac{\sigma_2^2}{\delta_2^3}\left(2e^{-\delta_2(T-t)} - \frac{1}{2}e^{-2\delta_2(T-t)} + \delta_2(T-t) - \frac{3}{2}\right) \\
& \frac{\sigma_3^2}{\delta_3^3}\left(2e^{-\delta_3(T-t)} - \frac{1}{2}e^{-2\delta_3(T-t)} + \delta_3(T-t) - \frac{3}{2}\right)\Bigg]
\end{aligned} \tag{21}
$$

5.7 Nelson-Siegel Model

Christensen *et al.* (2009) present a Nelson-Siegel parametric model for the yield curve with a 3-factor Gaussian model for the dynamics. This model is not consistent but provides a contrast in assessing the models. This is a restricted version of the 3-factor dependent model. The $\mathbb{Q}$-measure dynamics of the survival curve are given by equation (17). The solution is

$$
\begin{aligned}
B_1(\tau) &= -\tau \\
B_2(\tau) &= -\frac{1-e^{-\delta\tau}}{\delta} \\
B_3(\tau) &= -\Big[\frac{1-e^{-\delta\tau}}{\delta} - \tau e^{-\delta\tau}\Big]
\end{aligned} \tag{22}
$$

The solution for $C(\tau)$ is given in Christensen *et al.* (2009).

5.8 Three-Factor Consistency Test

For the 3-factor models the survival curve, or more specifically the curve of the average force of mortality, is

$$
M_{x,t}(\tau) = -\frac{B_1(\tau)}{\tau}Z_{1,t} - \frac{B_2(\tau)}{\tau}Z_{2,t} - \frac{B_3(\tau)}{\tau}Z_{3,t} + \frac{C(\tau)}{\tau}
$$

For the 3-factor dependent model this becomes

$$
\begin{aligned}
M_{x,t}(\tau) = \frac{1}{\tau}\Bigg[& \xi_0 + \xi_1(T-t) + \xi_2(t)\frac{1-e^{-\delta_{11}(T-t)}}{\delta_{11}} + \xi_3\frac{1-e^{-2\delta_{11}(T-t)}}{2\delta_{11}} + \\
& \xi_4(t)\frac{1-e^{-\delta_{22}(T-t)}}{\delta_{22}} + \xi_5\frac{1-e^{-2\delta_{22}(T-t)}}{2\delta_{22}} + \xi_6(t)\frac{1-e^{-\delta_{33}(T-t)}}{\delta_{33}} + \\
& \xi_7\frac{1-e^{-2\delta_{33}(T-t)}}{2\delta_{33}} + \xi_8\frac{1-e^{-(\delta_{22}+\delta_{33})(T-t)}}{\delta_{22}+\delta_{33}}\Bigg]
\end{aligned} \tag{23}
$$

The force of mortality is

$$
\begin{aligned}
\mu_{x,t}(\tau) = \xi_1 + & \xi_2 e^{-\delta_{11}(T-t)} + \xi_3 e^{-2\delta_{11}(T-t)} + \xi_4 e^{-\delta_{22}(T-t)} + \xi_5 e^{-2\delta_{22}(T-t)} \\
& + \xi_6 e^{-\delta_{33}(T-t)} + \xi_7 e^{-2\delta_{33}(T-t)} + \xi_8 e^{-(\delta_{22}+\delta_{33})(T-t)}
\end{aligned}
$$

This shows that the addition of a third independent factor to the 2-factor dependent model does not change the model consistency.

For the Nelson-Siegel model the average force of mortality is

$$M_{x,t}(\tau) = Z_{1,t} + \frac{1-e^{-\delta\tau}}{\delta\tau} Z_{2,t} + \left[\frac{1-e^{-\delta\tau}}{\delta\tau} - e^{-\delta\tau}\right] Z_{3,t} - \frac{C(\tau)}{\tau} \tag{24}$$

5.8.1 3-Factor Model - Real World Measure

The factor dynamics are estimated based on the independent factor model under the real world measure

$$\begin{pmatrix} dZ_{1,t} \\ dZ_{2,t} \\ dZ_{3,t} \end{pmatrix} = \begin{pmatrix} \kappa_1 & 0 & 0 \\ 0 & \kappa_2 & 0 \\ 0 & 0 & \kappa_3 \end{pmatrix} \left[\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} - \begin{pmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} dW^P_{1,t} \\ dW^P_{2,t} \\ dW^P_{3,t}. \end{pmatrix}$$

The drift term under the real-world measure is the same for all the 3-factor models. Under the independent and Nelson-Siegel models the σ_{32} term is equal to zero.

Under the $\mathbb{P}$-measure the average force of mortality curve becomes

$$M_{x,t}(\tau) = -\frac{B_1(\tau)}{\tau} Z_{1,t} - \frac{B_2(\tau)}{\tau} Z_{2,t} - \frac{B_3(\tau)}{\tau} Z_{3,t} + \frac{C(\tau)}{\tau}$$

The average force of mortality for the dependent model is

$$M_{x,t}(\tau) = \left[\rho_1 \frac{1-e^{-\delta_{11}(T-t)}}{\delta_{11}}\right] \frac{Z_{1,t}}{\tau} + \left[\rho_2 \frac{1-e^{-\delta_{22}(T-t)}}{\delta_{22}}\right] \frac{Z_{2,t}}{\tau} + \left[\rho_2 \frac{\delta_{32}}{\delta_{22}-\delta_{33}} \left(\frac{1-e^{-\delta_{22}(T-t)}}{\delta_{22}} - \frac{1-e^{-\delta_{33}(T-t)}}{\delta 33}\right) + \rho_3 \frac{1-e^{-\delta_{33}(T-t)}}{\delta_{33}}\right] \frac{Z_{3,t}}{\tau} - \frac{C(\tau)}{\tau}$$

The average force of mortality for the independent model is

$$M_{x,t}(\tau) = \left[\frac{1-e^{-\delta_{11}(T-t)}}{\delta_{11}}\right] \frac{Z_{1,t}}{\tau} + \left[\frac{1-e^{-\delta_{22}(T-t)}}{\delta_{22}}\right] \frac{Z_{2,t}}{\tau} + \left[\frac{1-e^{-\delta_{33}(T-t)}}{\delta_{33}}\right] \frac{Z_{3,t}}{\tau} - \frac{C(\tau)}{\tau}$$

The constant adjustment terms, $C(\tau)$, is given in equations (19) and (21) for the dependent and independent 3-factor models respectively. The Nelson-Siegel average force of mortality curve is given by equation (24), with the constant adjustment term $C(\tau)$ given in Christensen *et al.* (2009).

6 Model Estimation

6.1 Kalman Filter

The Kalman filter is an optimal linear estimator, Kalman (1960). The Kalman filter provides a recursive solution to the discrete linear filtering problem, see Welch and Bishop (1995). The filter is optimal since it minimises the error covariance. The Kalman filter has been increasingly used in financial applications including estimating affine term structure models, BABBS and NOWMAN (1999) and ANDERSEN and LUND (1997). The Kalman filter is based on a set of state-space equations with a linear stochastic difference equation for the unobservable factors of the system, Z, where Z has n-dimensions. This is the transition system and is given in discrete form by

$$Z_t = A \times Z_{t-1} + B + \eta_{t-1}$$

The matrix $A \in \mathbb{R}^{n \times n}$, assumed constant, gives the relationship between the previous state of the factor, Z_{t-1}, and the current state, Z_t. The matrix $B \in \mathbb{R}^n$ is an external control for the factors.

The second equation of the state-space representation is the measurement system relating the unobserved factors to observed data, M, where M has m-dimensions.

$$M_t = H \times Z_t + C + \epsilon_t$$

The matrix $H \in \mathbb{R}^{m \times n}$, assumed constant, gives the relationship between the current factor of the process, Z_t, and the measurement M_t. A constant $C \in \mathbb{R}^m$ provides external control to the measurement equation.

The random variables $\eta_t \in \mathbb{R}^{n \times n}$ and $\epsilon_t \in \mathbb{R}^{m \times m}$ are the transition and measurement noise respectively. They are assumed to be independent and normally distributed,

$$\eta_t \sim N(0, Q)$$
$$\epsilon_t \sim N(0, R)$$

The measurement system noise covariance, $R \in \mathbb{R}^{m \times m}$, is assumed constant. The transition system noise covariance, $Q \in \mathbb{R}^{n \times n}$, is determined by the transition system dynamics. We assume measurement noise is independent, so that R is a diagonal matrix.

The Kalman filter estimates the unobserved states, Z_t, at each time interval and uses feedback from the measurements, M_t to update estimates. The Kalman filter has time update equations and measurement update equations. The time update equations project the current state and error covariance to obtain an estimate for the next time step. The measurement update equations incorporate new information to improve the estimate from the time update equations.

The time update equations are,

$$\hat{Z}_t^- = A \times \hat{Z}_{t-1} + B \times u_t$$
$$\hat{P}_t^- = A \times \hat{P}_{t-1} \times A' + Q.$$

When initialising the filter, the starting values for the recursion are the unconditional mean, Z_0, and variance, P_0, of the transition system. These are calculated from historical data given by

$$\hat{Z}_t^- = A \times Z_0 + B \times u_t$$
$$\hat{P}_t^- = A \times P_0 \times A' + Q.$$

The Kalman Gain, K, minimizes the covariance error of the system. It weights the difference between the actual measurement, M_t, and the predicted measurement, $H\hat{Z}_t^- + C$. As the measurement error covariance approaches zero, the actual measurement is weighted more heavily.

The error term and the residual variance-covariance matrix is

$$\epsilon_t = M_t - C - H \times \hat{Z}^-$$
$$S_t = H \times \hat{P}_t^- \times H' + R.$$

These two quantities determine the standardised residuals,

$$e_t = \frac{\epsilon_t}{\sqrt{S_t}}$$

The Kalman Gain is then calculated,

$$K_t = \frac{\hat{P}_t^- \times H'}{S_t}$$

The measurement update equations are,

$$\hat{Z}_t = \hat{Z}_t^- + K_t \left(M_t - H \times \hat{Z}_t^- \right)$$
$$\hat{P}_t = (I - K_t \times H) \times \hat{P}_t^-.$$

Under the assumptions of the state-space model, where all the prediction errors are Gaussian, the Kalman filter maximizes a multi-variate normal likelihood. The likelihood is computed for each recursion step of the filter with

$$l(\theta) = -\frac{1}{2} \sum_{t=1}^{N} \left[n\ln(2\pi) + \ln|S_t| + \epsilon_t' S_t^{-1} \epsilon_t \right].$$

Since we require the optimal set of parameters, $\hat{\theta}$, that maximizes the likelihood, this is computed from

$$\hat{\theta} = \text{argmax}_\theta l(\theta).$$

The Kalman filter is used to compute the likelihood for a given parameter set. To find the optimal parameter set, a non-linear optimiser recursively calls the Kalman filter to solve for parameter estimates.

6.2 Estimation Method

The complete state-space formulation of the survival curve model, in matrix form, is given by

$$M_t = B' Z_t - C + \epsilon_t \quad \epsilon_t \sim N(0, R)$$
$$dZ_t = K \left[\mu - Z_t \right] dt + \Sigma dW$$

For the Kalman filter, the system is implemented in a discrete form as

$$\Delta Z_t = \left(\mathbf{I} - e^{-K} \right) (\mu - Z_t) + \eta_t$$
$$\eta_t \sim N(0, Q)$$
$$Q = \int_0^1 e^{-Ks} \sigma \sigma' e^{-Ks} ds$$

Details for Q are given in Christensen *et al.* (2009).
The estimation of the parameters is computed using

$$M_t = B'Z_t - C + \epsilon_t \quad \epsilon \sim N(0, R)$$
$$Z_t = \left(I - e^{-K}\right)\mu + e^{-K}Z_{t-1} + \eta_t \quad \epsilon \sim N(0, Q)$$

The structure of the measurement covariance matrix, R, is diagonal and allows for the measurement error variance to increase for older ages. A constant measurement error is assumed up to a specific age. After this age the measurement error variance is assumed to increase exponentially with age based on the data. This models the measurement error variance across all ages with three parameters. We could use a separate measurement error for each age, but this would require 50 parameters for our data set. This is simplified with the measurement error as a parametric curve. We reduce the number of parameters to estimate from 50 down to 3. We specify the measurement error of the average force of mortality as

$$\epsilon_t \sim N(0, R)$$

With the covariance matrix, R, constant over time and independent between ages. The parametric form used for the diagonal of the covariance matrix, with all other entries zero, is,

$$R(\tau) = \frac{1}{\tau}\sum_{i=1}^{\tau}\left[r_c + r_1 e^{r_2 \times i}\right] \tag{25}$$

where the values of r_c, r_1 and r_2 are estimated as part of the parameter set, $\hat{\theta}$. This specification enables the maximum likelihood and mean squared error to converge at the optimal parameter set.

6.3 Bootstrapping State-Space models

The parameter estimates of the real-world drift terms are found to have wide confidence intervals. Because of the flat likelihood at the optimal parameter estimate values and the potential non-linearities in the model a bootstrap method is used to re-sample the observed data with the standardised residuals from the optimal Kalman filter parameter estimates. This produces a distribution of parameter estimates that is asymptotically Gaussian.

The re-sample method is detailed in Stoffer and Wall (2004). The process used is as follows:

1. Determine the standardized residuals from the optimal parameter set

$$e_t = \frac{\epsilon_t}{\sqrt{S_t}} \tag{26}$$

2. Re-sample the standardized residuals with replacement to generate a bootstrap set of residuals, e_t^* excluding the first four residuals because of Kalman filter start up irregularities.

3. Recompute the transition and measurement equations given the bootstrap residuals, e_t^*.

$$Z_t^* = \left(I - e^{-K}\right)\mu + e^{-K} Z_{t-1}^* + \frac{e_t^*}{\sqrt{S_t}}$$
$$M_t^* = B' Z_t^* - C + K_t \frac{e_t^*}{\sqrt{S_t}}$$

4. Use the bootstrap data set, M_t^*, to estimate a new set of parameters that maximize the likelihood,

$$\hat{\theta}^* = \text{argmax}_\theta^* l(\theta^*)$$

where θ^* is the parameter set determined from the bootstrapped data set M_t^*.

5. Repeat steps 2-4 300 times to obtain a sample distribution for each of the parameter estimates.

7 Estimation Results

7.1 Model Goodness of Fit

Results from the estimation and assessment of the five models are presented and discussed in this section. Table (1) shows the maximum Log Likelihood and the Root Mean Squared Error (RMSE) for each of the models. The 2- and 3-factor models all have similar log Likelihood and RMSE. The 3-factor dependent model estimates 20 parameters and 294 latent factors (3 factor estimates for each of the 98 years of data we tested).

Models were compared with a standard Likelihood Ratio (LR) test, since all models are a subset of the 3-factor dependent model. The null hypothesis is that an increase in the number of parameters does not significantly improve the fit. For all the 3-factor models the null hypothesis is not rejected. Extra parameters do not improve the fit of the 3-factor models. The 2-factor models are rejected in favour of the 3-factor models.

	3-Factor Dependent	3-Factor Independent	3-Factor Nelson-Siegel	2-Factor Dependent	2-Factor Independent
Log Likelihood	31672	31805	31707	29222	29276
RMSE	0.00088	0.00090	0.00094	0.00195	0.00221
No. of Model Parameters	20	15	13	15	11
No. of Factors Estimated	294	294	294	196	196
Parameter Restrictions	-	5	7	203	207
Likelihood Ratio Test	-	<0	<0	4900	4792
ΔAICb	-	<0	<0	5355	4983

Table 1: Comparison of Log Likelihood, RMSE and the number of parameters estimated for each model. A Likelihood Ratio test and AICb is also performed.

A modified Akaike Information Criteria (AIC) test for state-space models is also performed, Cavanaugh and Shumway (1997). The AICb test uses the bootstrapped Log Likelihood results to estimate the penalty term. There is no improvement in using

ML Estimates	2-Factor Dependent	95% Confidence Lower	Upper	2-Factor Independent	95% Confidence Lower	Upper
δ_{11}	-0.09224	-0.09686	-0.09180	-0.09072	-0.09577	-0.08920
δ_{21}	0.23533	0.21922	0.27258	-	-	-
δ_{22}	0.01208	-0.00242	0.01893	0.04031	0.03612	0.05131
ρ_1	0.02938	0.02552	0.03470	-	-	-
ρ_2	0.56477	0.49948	0.58019	-	-	-
κ_{11}	0.04766	0.05600	0.38681	0.00180	0.00010	0.00210
κ_{22}	0.01092	0.00011	0.01103	0.01266	0.00014	0.01622
μ_1	0.39532	0.13192	0.26123	0.01197	-0.04011	0.09040
μ_2	-0.01392	-0.06527	0.03513	-0.19773	-0.19999	0.34400
σ_{11}	6.347e-03	3.074e-03	7.371e-03	4.851e-04	2.810e-04	5.519e-04
σ_{21}	1.231e-02	9.373e-03	1.258e-02	-	-	-
σ_{22}	1.790e-03	8.286e-04	1.901e-03	1.131e-02	7.682e-03	1.163e-02
r_1	2.237e-09	6.352e-10	7.569e-09	7.555e-10	2.316e-10	1.427e-09
r_2	0.21817	0.19642	0.26022	0.25390	0.23649	0.29648
r_c	4.323e-08	1.950e-08	1.338e-07	3.407e-08	1.769e-08	6.595e-08

Table 2: 2-Factor Dependent and Independent Model Parameter Estimates.

the 3-factor dependent model over the other 3-factor models, although the 2-factor models were rejected. The LR and AICb tests are calculated as

$$\text{LR} = 2\left[\log \text{L}(\hat{\theta})_{3f} - \log \text{L}(\hat{\theta})_{\text{nest}}\right] \sim \chi^2(q) \tag{27}$$

$$\text{AICb} = -2\log \text{L}(\hat{\theta}) + 2\left[\frac{1}{N}\sum_{i=1}^{N} -2\log \text{L}(\hat{\theta}_i^{\,*}) - (-2\log \text{L}(\hat{\theta}))\right] \tag{28}$$

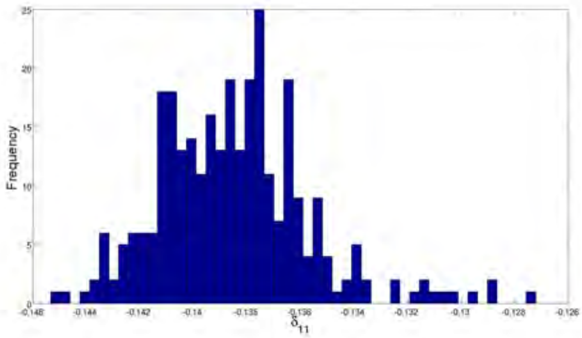
where q is the number of parameter restrictions, N is the number of bootstrap distributions and $\text{L}(\hat{\theta}_i^{\,*})$ is the maximum likelihood for ith bootstrap distribution. The number of bootstrap distributions used in all the model is 300.

7.2 Parameter Estimates

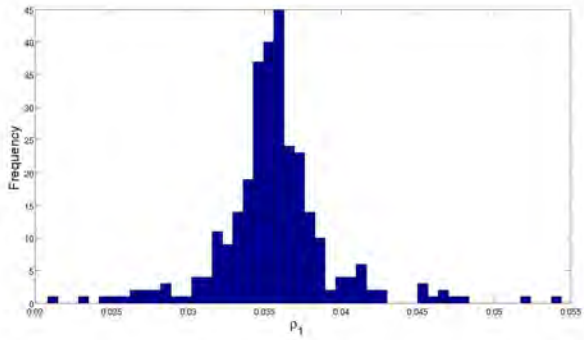
Table (2) gives the estimated parameters for the 2-factor dependent and independent models along with 95% confidence intervals from the state-space bootstrapping method described in section (6.3). All the δ parameters of the independent model are significant. The confidence interval for the δ_{22} parameter of the dependent model includes zero, indicating this parameter is not significant and this drift of the second factor, Z_2, is dependent on Z_1 under the $\mathbb{Q}$-measure.

The parameters ρ_1 and ρ_2, that scale $B_1(\tau)$ and $B_2(\tau)$ in equation (9) are significant and different from 1. The larger value of ρ_2 compared to ρ_1 implies that B_2 has more impact for younger ages.

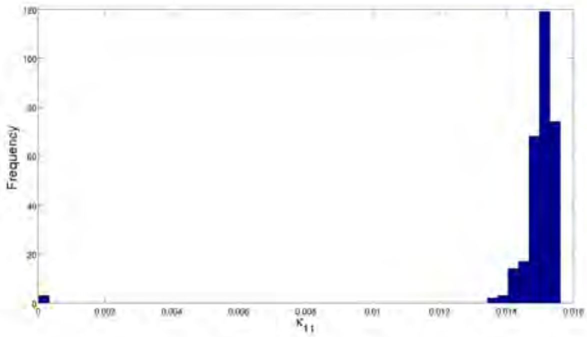
The κ parameters for the real-world drift process have wide bootstrap confidence intervals, with lower bounds close to zero, indicating a random walk, and their corresponding μ parameter estimates are unreliable. κ_{11} has a wide confidence interval, with the MLE estimate outside this range, this parameter is also unreliable.



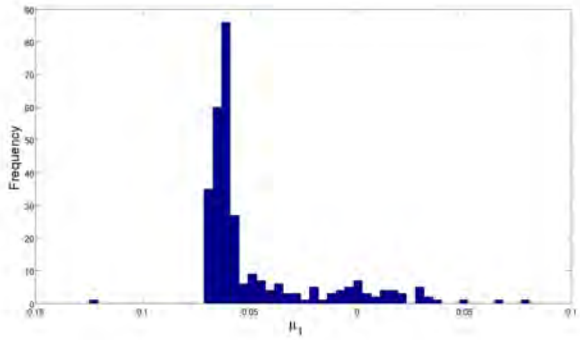
(a) δ_{11} Estimation



(b) ρ_1 Estimation



(c) κ_{11} Estimation



(d) μ_1 Estimation

Figure 3: 3-Factor Dependent Model Bootstrap Distribution

The volatility matrix, Σ and the measurement variance parameters r_1, r_2 and r_c for both models are significant. The σ_{21} parameter estimate suggests correlation between the Brownian motions of the unobserved stochastic process.

Table (3) shows the estimated parameters of the 3-factor dependent, independent and the 3-factor Nelson-Siegel models. Results are similar to the 2-factor models. Factor loading parameters are significant except for δ_{33} in the dependent model.

None of the real-world drift parameters for the 3-factor independent and Nelson-Siegel model's parameters are identifiable, they all have very wide confidence intervals. The 3-factor dependent model has significant drift parameters, although the distribution of these parameters are skewed and have a very high kurtosis due to a number of outliers. The scale parameters, ρ, are significantly different from 1. There are 5 more parameters being estimated in the dependent model over the independent model, and these parameters are significant, although there is not a significant improvement in model fit in terms of maximum likelihood or RMSE. Again, the Σ and measurement error variance parameters were able to be identified.

Figure (3) displays the bootstrap distributions and table (4) shows the metrics for four parameters from the 3-factor dependent model. We can see some problems in assuming Gaussian distribution for all the parameters. The factor loading and scaling parameters are approximately Gaussian, given the 300 bootstrap simulation that were performed. We can see that δ_{11} is skewed and ρ_1 has a number of outliers. The real-world drift parameter κ_{11} is highly skewed (and high kurtosis) due to a small number of outliers close to zero.

ML Estimates	3-Factor Dependent	95% Confidence Lower	95% Confidence Upper	3-Factor Independent	95% Confidence Lower	95% Confidence Upper	3-Factor Nelson-Siegel	95% Confidence Lower	95% Confidence Upper
δ_{11} (NS: δ)	-0.14062	-0.14269	-0.13376	-0.10107	-0.10377	-0.09755	-0.11652	-0.11634	-0.10657
δ_{22}	-0.10194	-0.10427	-0.09364	0.01072	0.00039	0.01439	-	-	-
δ_{32}	0.01775	0.01507	0.04502	-	-	-	-	-	-
δ_{33}	0.00884	-0.00386	0.01468	-0.15334	-0.15564	-0.14419	-	-	-
ρ_1	0.03305	0.02928	0.04150	1	-	-	1	-	-
ρ_2	0.05222	0.04962	0.08890	1	-	-	1	-	-
ρ_3	0.09371	0.06886	0.09658	1	-	-	0	-	-
κ_{11}	0.01536	0.01423	0.01551	0.00013	0.00010	0.02551	0.00747	0.00005	0.59999
κ_{22}	0.09560	0.08641	0.11151	0.29723	0.22008	0.99759	0.01286	0.00004	0.05607
κ_{33}	0.00772	0.00171	0.01836	0.04343	0.00011	0.13402	0.02012	0.00005	0.07054
μ_1	-6.647e-002	-6.887e-002	1.745e-002	-3.435e-003	-1.104e-002	1.878e-002	1.162e-002	-9.316e-002	5.265e-002
μ_2	9.075e-002	5.621e-002	9.836e-002	1.725e-003	-4.280e-003	4.900e-003	4.285e-003	1.488e-004	8.346e-003
μ_3	7.979e-002	1.218e-002	8.009e-002	8.957e-005	-4.057e-004	6.338e-004	3.478e-004	-2.488e-005	6.016e-004
σ_{11}	2.849e-003	2.755e-003	3.688e-003	3.966e-004	1.887e-004	4.014e-004	4.650e-003	2.592e-003	5.303e-003
σ_{22}	1.038e-007	1.000e-007	3.740e-003	4.156e-003	2.787e-003	5.719e-003	1.499e-004	1.492e-004	2.531e-004
σ_{32}	5.043e-004	2.065e-004	6.631e-004	-	-	-	-	-	-
σ_{33}	6.045e-003	3.617e-003	9.750e-003	3.998e-005	3.582e-005	7.255e-005	3.421e-005	3.434e-005	7.188e-005
r_1	8.462e-010	5.855e-010	9.252e-010	3.462e-010	1.270e-010	3.490e-010	3.172e-010	1.100e-010	3.158e-010
r_2	0.21481	0.21167	0.22686	0.23727	0.23474	0.26493	0.24162	0.23908	0.26993
r_c	1.651e-008	1.348e-008	1.729e-008	1.930e-008	1.214e-008	2.018e-008	1.363e-008	1.207e-008	1.719e-008

Table 3: 3-Factor Dependent, Independent and Nelson-Siegel Model fit results. Brackets indicate 95% bootstrap confidence intervals.

	δ_{11}	ρ_1	κ_{11}	μ_1
Kalman MLE	-0.1406	0.0331	0.0154	-0.0665
Mean	-0.1385	0.0356	0.0149	-0.0482
Median	-0.1386	0.0356	0.0151	-0.0606
Standard Deviation	0.0028	0.0038	0.0015	0.0289
Skewness	0.7795	0.5735	-9.0787	1.7826
Excess Kurtosis	1.5687	4.8238	84.9490	2.8314

Table 4: Metrics of Bootstrap Distributions (3-Factor Dependent Model).

7.3 In-Sample Analysis

Figure (4) shows the percentage error of the survival curve estimates for each of the models. The percentage errors are shown for years 1910, 1940, 1970 and 2000. The scale of each error plot is different for above and below the age of 85, allowing more detail to be shown below 85 where the percentage errors are lower. For the early years, 1910 and 1940, there are little differences between the models below the age of 85. In the years 1970 and 2000 the 3-factor models are significantly better. The 2-factor models do not fit the survivor curve well over the age of 85, and this becomes worse in later observation years.

Figure (5) shows the Mean Absolute Relative Error (MARE) of the fitted survivor curve for all time periods. The 3-factor models provide a better fit, but the percentage error of all models is low under age 85. Over age 85 the 3-factor models are required to capture the variation in the survival curve. There are limited differences between the 3-factor models.

7.4 Residual Analysis

Figure (6) shows the standardised force of mortality residuals, $\hat{m}$, for the 2-factor dependent and independent models. These residuals are re-constructed from the fitted average force of mortality residuals, $\hat{e}$. The relationship is given in equation (29)

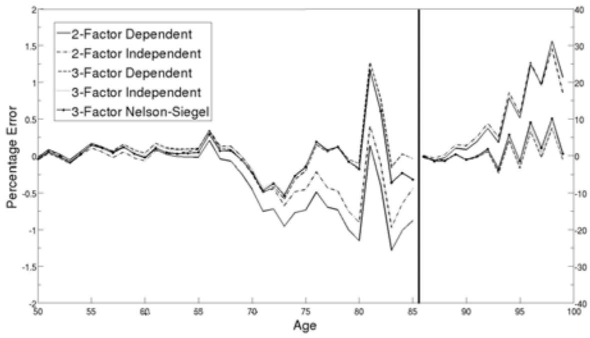
$$\hat{e}_{x,t}(\tau) = \frac{1}{\tau}\sum_{i=0}^{\tau-1} \hat{m}_{x+i,t} \tag{29}$$

Figure (6) shows the positive and negative residuals for the models. They show a good fit before 1960. After 1960 the 2-factor models residuals are less random than the 3-factor models. It is of interest to note that the 1960's was the period when the reporting of the Swedish population estimates changed.

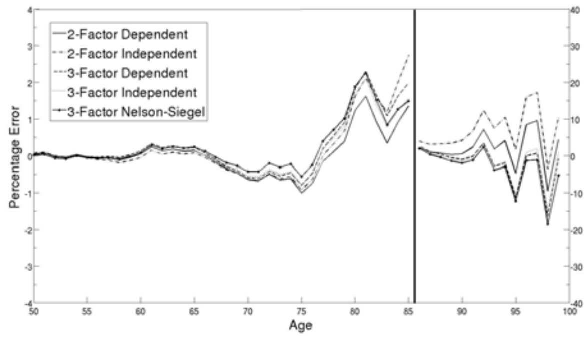
Figure (7) show the standardised residuals for the 3-factor dependent, independent and Nelson-Siegel models. The additional factor captures the structural change in mortality rates after 1960. All the 3-factor model residual patterns are similar, but there some noticeable 'cohort effects' that cannot be captured by the model.

7.5 Estimated Factor and Factor Loadings

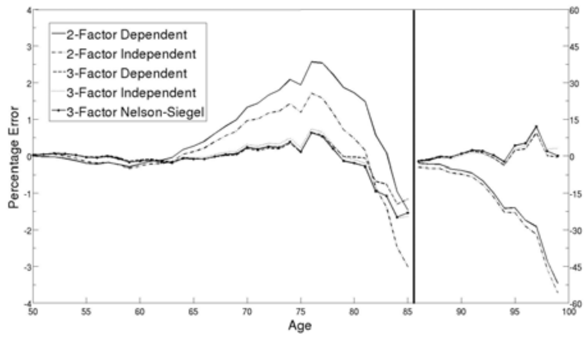
The latent factors and the factor loadings for the 2-factor dependent model and independent models are shown in Figures (8) and (9). The first factor, Z_1, is similar in both



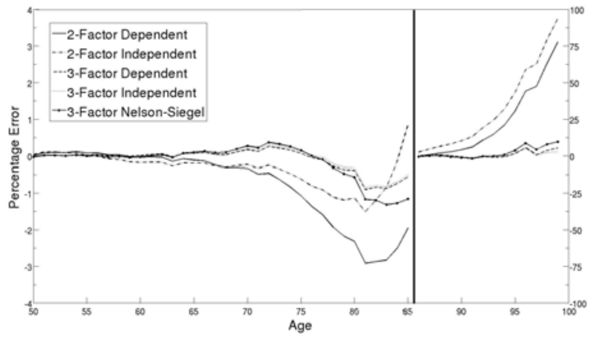
(a) Percentage Error year 1910



(b) Percentage Error year 1940



(c) Percentage Error year 1970



(d) Percentage Error year 2000

Figure 4: 2 and 3-Factor Survival Curve Percentage Errors

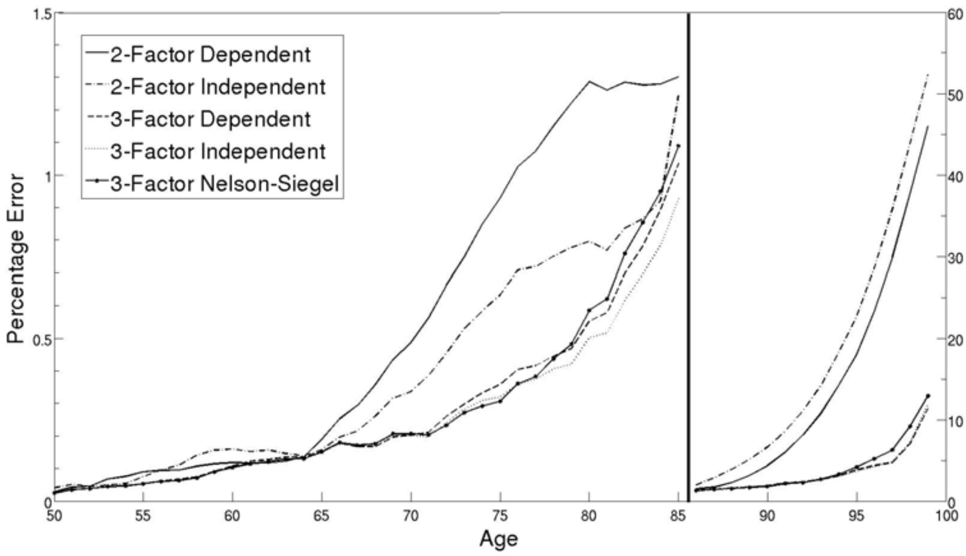
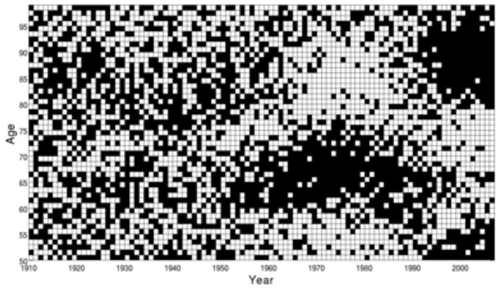
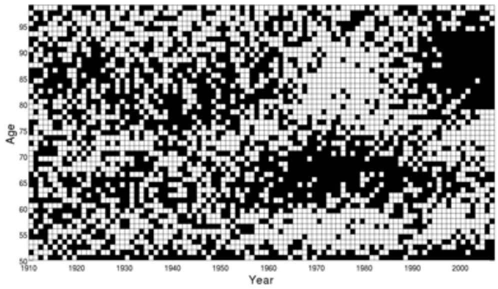


Figure 5: 2 and 3-Factor Survival Curve MARE

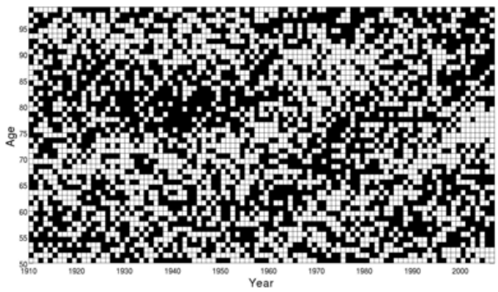


(a) Dependent Model Residuals

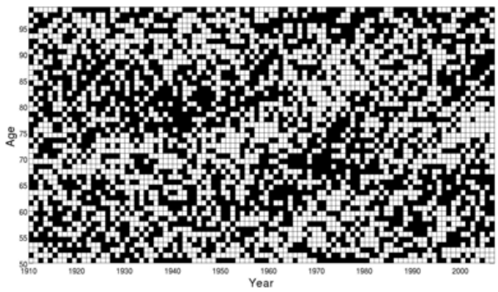


(b) Independent Model Residuals

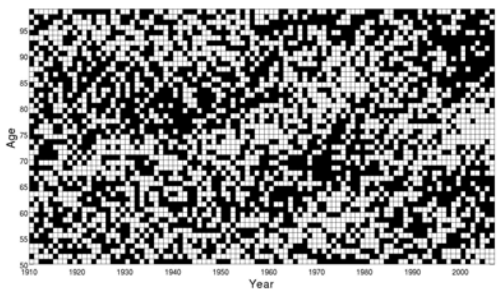
Figure 6: 2-Factor Model Residuals



(a) Dependent Model Residuals

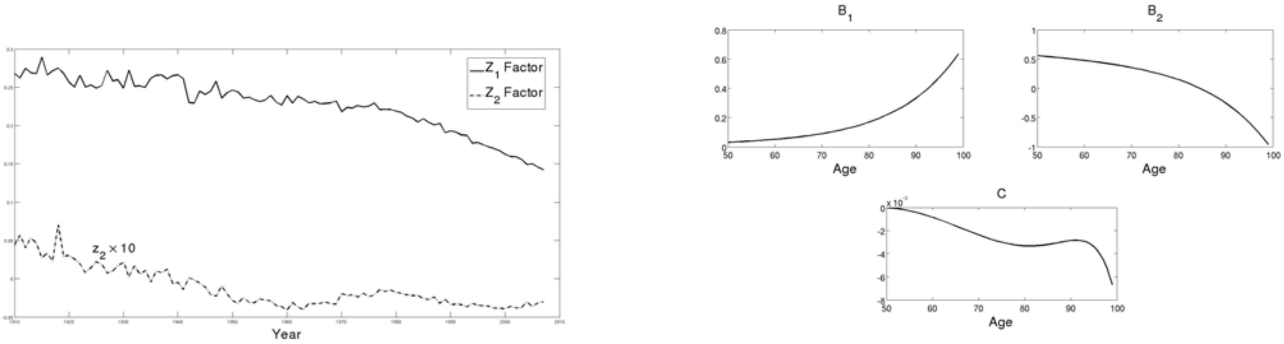


(b) Independent Model Residuals



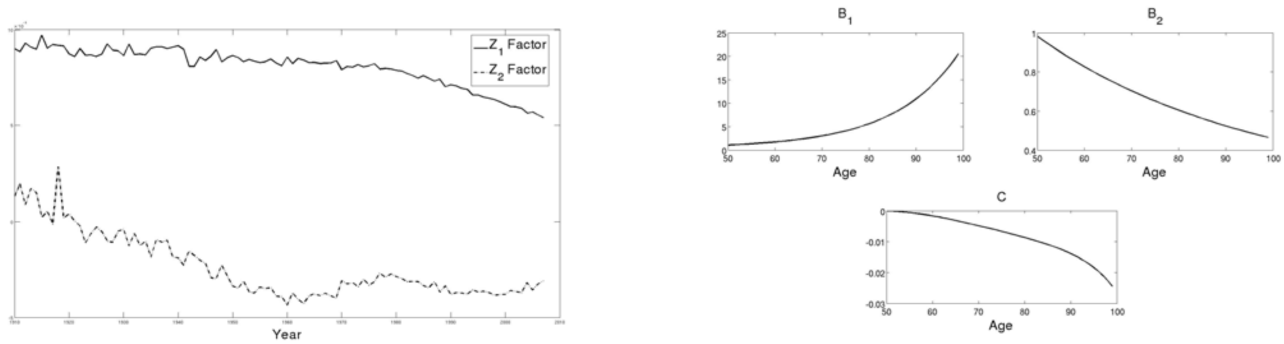
(c) Nelson-Siegel Model Residuals

Figure 7: 3-Factor Model Residuals



(a) Factors (b) Factor Loadings

Figure 8: 2-Factor Dependent Model



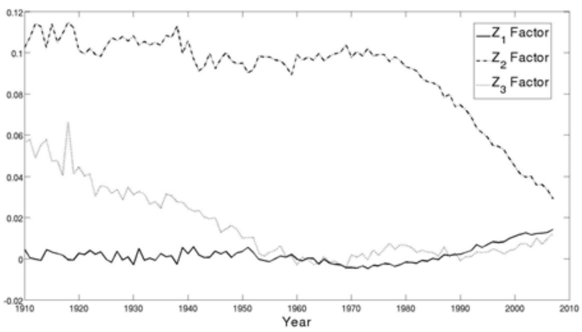
(a) Factors (b) Factor Loadings

Figure 9: 2-Factor Independent Model

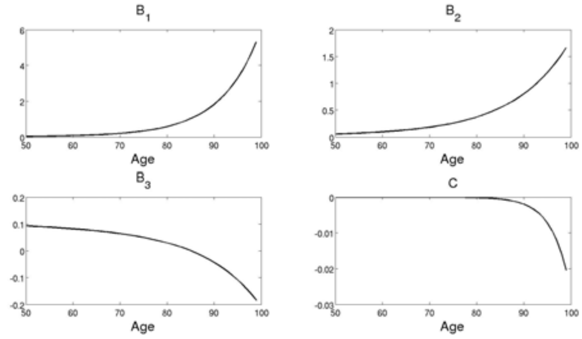
models and shows the general mortality trend. This factor, and mortality rates, were slowly improving between 1910 and 1970 with a fairly large variation between years. Since the 1970's the improvement trend has increased with a lower volatility between years. The second factor, Z_2, also similar in both models, decreases between 1910 and 1960. Since 1960, the volatility is lower and the downward trend has stopped. The 2-factor models do not fit the change in mortality rates after this period, as seen in the residuals. This change in trend from the 1960's has had more of an effect on the population over 85. This is seen in figures (4) and (5) where the 2-factor model fit becomes substantially worse in observation years 1970 and 2000 for people over the age of 85.

The first factor loading, B_1, for both models shows the general shape of the average force of mortality curve. The different shape of factor loading B_2 in the dependent and independent models is due to different risk-neutral drift specifications of the models. The dependent model allows for a negative but increasing exponential curve. This has a definite benefit in the fitting performance over the independent model for people over the age of 85, as shown in figure (5).

The factors and factor loadings for the 3-factor dependent model are shown in figure (10). The factors Z_2 and Z_3 are very similar to the 2-factor results. Z_3 has the same downward slope as Z_2 in the 2-factor models. This trend also stops around 1960. The slope of Z_2 is greater then Z_1 in the 2-factor models after 1970. The general trend of mortality rates has been improving faster than can be fitted with the 2-factor models.

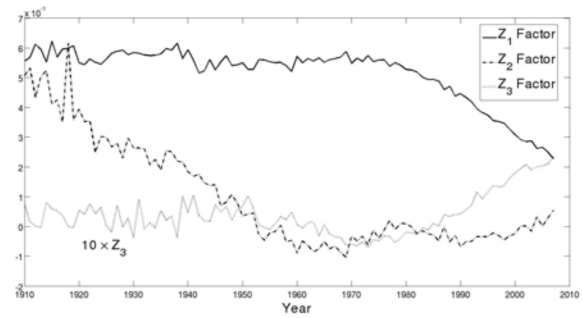


(a) Factors

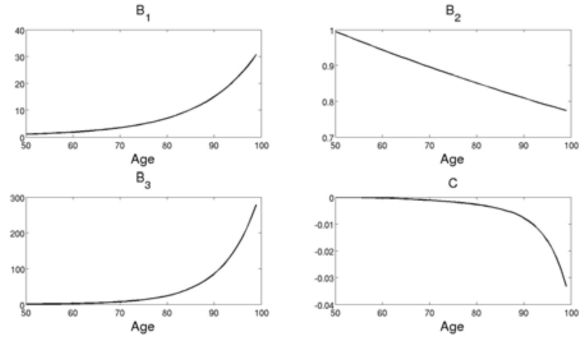


(b) Factor Loadings

Figure 10: 3-Factor Dependent Model



(a) Factors



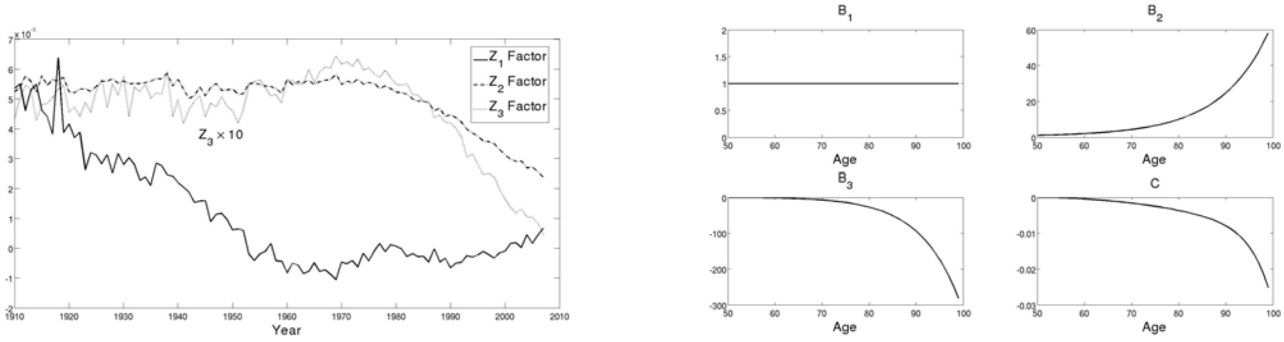
(b) Factor Loadings

Figure 11: 3-Factor Independent Model

The additional factor, in this case Z_1, captures the mortality trend after 1960. This factor is level until the 1960s, with some volatility between periods, then increasing at a constant rate.

The 3-factor dependent model also allows for a negative but increasing factor loading, B_3. Z_3 represents the trend before 1970 and Z_1 after, so that the effect of these trends on mortality is given by the factor loadings B_3 and B_1 respectively. Mortality rates in the older population (over 85) have been improving faster than the rest of the population until 1970, but since then the rate of improvement has been decreasing compared to the rest of the population. A similar interpretation based on figure (11) applies for the 3-factor independent model.

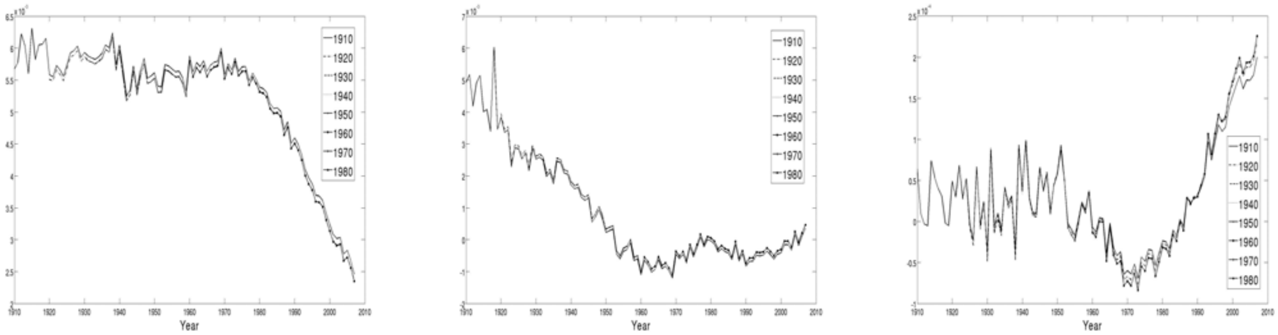
Although there are fewer parameters, the Nelson-Siegel model does not perform significantly worse then the consistent models. The Nelson-Siegel factors and factor loadings are shown in figure (12). The factor load B_1 is constant forcing the factor Z_1 to weight all ages the same. This loss of flexibility gives a worse fit in the older ages. The factors Z_2 and Z_3 appear very similar with each factor having opposite effects on the mortality rate, as indicated by their factor loadings.



(a) Factors

(b) Factor Loadings

Figure 12: 3-Factor Nelson-Siegel Model



(a) Factor Z_1

(b) Factor Z_2

(c) Factor Z_3

Figure 13: 3-factor Independent Model Robustness - Factors

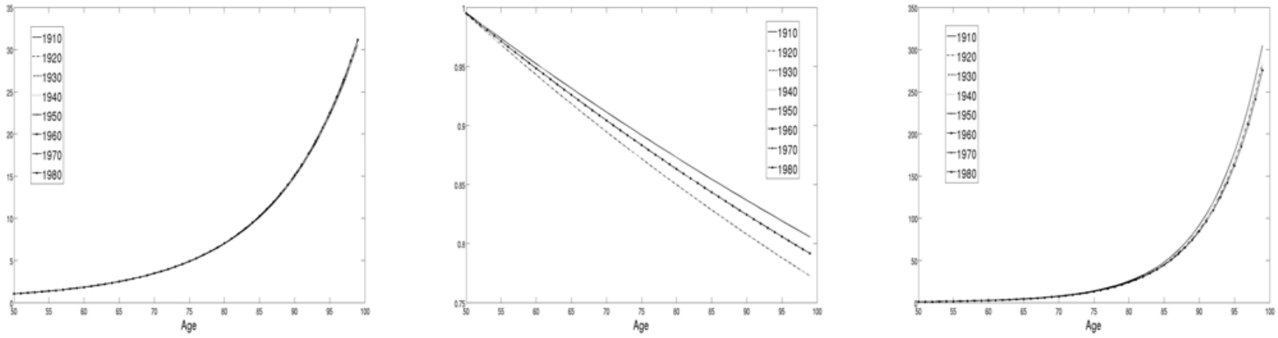
7.6 Model Robustness

Figures (13) and (14) show the robustness of the estimated factors and factor loadings to starting year in the data set. The 3-factor independent model was re-run a number of times with different starting years. Every 10 years from 1910 to 1980 was assessed. The data set finishes in the year 2007. We can see very similar results for all the factors and factor loadings for each starting year.

The factor loadings B_1 and B_3 estimates are both very stable for each starting year. The factor loading B_2 shows some variation, although the only deviation is in the 1910 and 1920 starting years. For the two identified periods of different trends, pre- and post-1970, this demonstrates the robustness of the consistent models across time.

8 Conclusion

This paper has developed and assessed a range of models for survival curves based on the Affine Term Structure Models used in interest rate modelling. We have derived closed-form expressions for the risk-adjusted survival curve and estimated model parameters under an “essentially affine” model. Models have included 2- and 3-factor models in the consistent framework. The goal of this framework is to produce consistent forecasts that can be used in pricing and risk management applications.



(a) Factor B_1 (b) Factor B_2 (c) Factor B_3

Figure 14: 3-factor Independent Model Robustness - Factor Loadings

The data set used to estimate and compare the models was Swedish male mortality for ages from 50 to 99 over the years 1910 to 2007. The 3-factor models performed better than 2-factor models. Of the 3-factor models, an independent model is recommended that is both parsimonious and consistent. The models are robust to the starting year used for estimation and the Mean Absolute Relative Error (MARE) of the fitted survival curves for all the 3-factor models are very low.

The approach assessed here has ready applications to risk management and pricing of longevity risk. Traded instruments can be used to calibrate the price of risk in the model. The consistent framework shows that across time the models do not require re-calibration and can be effectively used for assessing capital requirements for risk management as well as pricing and hedging longevity risk for insurers and pension funds.

9 Acknowledgement

The authors acknowledge the financial support of ARC Linkage Grant Project LP0883398 Managing Risk with Insurance and Superannuation as Individuals Age with industry partners PwC and APRA.

References

Andersen, T., and J. Lund, 1997 Estimating continuous-time stochastic volatility models of the short-term interest rate. Journal of Econometrics **77**: 343–377.

Babbs, S. H., and K. B. Nowman, 1999 Kalman Filtering of Generalized Vasicek Term Structure Models. The Journal of Financial and Quantitative Analysis **34**: 115.

Bjork, T., and B. J. Christensen, 1999 Interest Rate Dynamics and Consistent Forward Rate Curves. Mathematical Finance **9**: 323–348.

Cavanaugh, J. E., and R. H. Shumway, 1997 A Bootstrap Variant Of AIC For Staet-Space Model Selection. Statistica Sinica **7**: 473–496.

Christensen, J. H. E., F. X. Diebold, and G. D. Rudebusch, 2009 An arbitrage-free generalized NelsonâSiegel term structure model. Econometrics Journal **12**: C33–C64.

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

De Rossi, G., 2004 Kalman filtering of consistent forward rate curves: a tool to estimate and model dynamically the term structure. Journal of Empirical Finance **11**: 277–308.

Duffee, G. R., 2002 Term Premia and Interest Rate Forecasts in Affine Models. The Journal of Finance **57**: 405–443.

Duffie, D., and R. Kan, 1996 A yield-factor model of interest rates. Mathematical finance **6**: 379–406.

Kalman, R., 1960 A new approach to linear filtering and prediction problems. Journal of basic Engineering **82**: 35–45.

Luciano, E., and E. Vigna, 2005 Non mean reverting affine processes for stochastic mortality. International Centre for Economic Research **4/2005**.

Njenga, C. N., and M. Sherris, 2009 Longevity Risk and the Econometric Analysis of Mortality Trends and Volatility Trends and Volatility. SSRN Electronic Journal .

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

Stoffer, D. S., and K. D. Wall, 2004 Resampling in State Space Models. State Space and Unobserved Component Models Theory and Applications : 227–258.

Welch, G., and G. Bishop, 1995 An introduction to the Kalman filter. University of North Carolina at Chapel Hill, Chapel Hill, NC : 1–16.

Wilmoth, J., K. Andreev, D. Jdanov, and D. Glei, 2007 Methods protocol for the human mortality database. University of California, Berkeley, and Max Planck Institute for Demographic Research, Rostock. URL: http://mortality. org [version 31/05/2007] **2007**.