ARC Centre of Excellence in Population Ageing Research

Working Paper 2019/12

The Application of Affine Processes in Cohort Mortality Risk Models

Zhiping Huang, Michael Sherris, Andrés M Villegas and Jonathan Ziveyi

---

The Application of Affine Processes in Cohort Mortality Risk Models

Zhiping Huang [*], Michael Sherris[†], Andrés M Villegas[‡] and Jonathan Ziveyi[§]

3rd September 2019

Abstract

This paper assesses and compares multi-factor continuous time affine mortality models applied to age-cohort mortality curves that are well suited for theoretical and practical application in finance and insurance. Models based on Gaussian distributed mortality rates, as well as the Cox-Ingersoll-Ross (CIR) process allowing for Gamma distributed mortality rates, are compared, also quantifying the probability of negative rates in the Gaussian models. In particular, we introduce the Gaussian Arbitrage-Free Nelson-Siegel (AFNS) mortality model incorporating level, slope and curvature factors. The models have appealing features including efficient estimation and computation. We estimate models using age-cohort data to capture cohort effects more effectively and in order to explain the variability in cohort mortality curves in the continuous time framework. The models allow for Poisson variation in the model estimation using the Kalman filter. The affine mortality models facilitate the derivation of closed-form survivor curves allowing for efficient valuation of mortality-linked claims. The models can also incorporate factor dependence allowing for age-dependence in the mortality curves. Importantly we show that the Gaussian independent factor AFNS model performs very well in explaining and forecasting cohort mortality.

Paper submitted for Society of Actuaries Living to 100 Symposium VII January 13-15, 2020.

Keywords: mortality models, continuous time, cohort curve, affine rates, Kalman filter

JEL Classification Numbers: G22, C13, C22, C52, J11

[*]School of Risk & Actuarial Studies, Australian Research Council Centre of Excellence in Population Ageing Research (CEPAR), UNSW Australia, email: zhipinghuang0204@hotmail.com

[†]School of Risk & Actuarial Studies, Australian Research Council Centre of Excellence in Population Ageing Research (CEPAR), UNSW Australia, email: m.sherris@unsw.edu.au

[‡]School of Risk & Actuarial Studies, Australian Research Council Centre of Excellence in Population Ageing Research (CEPAR), UNSW Australia, email: a.villegas@unsw.edu.au

[§]School of Risk & Actuarial Studies, Australian Research Council Centre of Excellence in Population Ageing Research (CEPAR), UNSW Australia, email: j.ziveyi@unsw.edu.au

1 Introduction

Longevity risk is the, now well recognised, risk that the overall survival probability of a reference population is higher than expected (Cairns et al., 2006a). The mortality improvements experienced over past decades have highlighted its significance. Life insurance companies and pension funds, as the holders of substantial longevity risk, are often reluctant to provide financial products for post-retirement wealth planning (Blake et al., 2014). The Continuous Mortality Investigation (2018) highlights the potential for mortality experience to improve, and the underestimation of longevity risk has been estimated as having a significant financial impact requiring an extra $450 billion in payments per year (The Joint Forum, 2013). This also results in a larger capital requirement for insurers (Barrieu et al., 2012). The quantification of longevity risk and the design of financial and insurance products to manage this risk are fundamental to life insurance companies and pension funds.

Many methods have been proposed for insurers and pension funds to manage the risk of mortality improvement including transfering longevity risk to counterparties through longevity swaps and transferring the risk to capital markets through securitization of insurance assets and liabilities (The Joint Forum, 2013). Capital markets are expected to become increasingly significant with potential for an expanding global market, arising from efficiency and effectiveness of longevity management (Blake et al., 2018). Longevity-linked securities have been proposed including longevity bonds (Blake and Burrows, 2001), longevity swaps (Dowd et al., 2006), and q-forwards (Coughlan et al., 2007). As with all financial market innovations, risk quantification and fair pricing are major concerns. The Life and Longevity Markets Association (2010) acknowledge that expected mortality improvement is a key input in pricing longevity risk, but forecasting such improvement remains challenging.

Many different stochastic mortality models have been proposed to describe the stochastic evolution of mortality rates at an aggregate population level. These include the Lee-Carter model (Lee and Carter, 1992) and its extensions under a discrete-time modeling framework. The Lee-Carter model is considered as a cornerstone and is widely used. Numerous extensions and improvements have been proposed. Cairns et al. (2009) provides a comprehensive summary and comparison of these extensions. The Cairns-Blake-Dowd model (Cairns et al., 2006b) and the age-period-cohort model (Renshaw and Haberman, 2006) have also become popular models. Discrete-time mortality models describe the mortality intensity as a discrete time series and usually require simulation to implement with no closed-form solutions for survival curves.

Continuous-time affine mortality models have also been proposed based on mathematical finance techniques for interest rate and credit risk modelling. These mortality models reflect the similarities between mortality rates and interest rates, as discussed in Milevsky and Promislow (2001), Dahl (2004), and Cairns et al. (2006a). Continuous-time mortality models using diffusion processes have many advantages over discrete-time models. Continuous-time models are readily incorporated into the modelling of longevity risk where both mortality models and financial models are required, as in the case of longevity-linked securities (Jevtic et al., 2013).

Affine mortality models are based on the application of Affine Term Structure Models (ATSMs) for interest rate modeling, as in Duffie and Kan (1996) and Dai and Singleton (2000), to mortality modelling. Affine mortality models, because of the similar structure to interest rate models, allow for an integrated pricing framework (Barrieu et al., 2012) and closed-form solutions to survival probabilities (Dahl, 2004). Affine processes are well suited for continuous-time mortality models because of their flexibility and analytical tractability. Affine mortality models that satisfy a consistency requirement (Björk and Christensen, 1999) have stable parameters and ensure consistency between the dynamics of mortality rates and the functional form for the survival curve. This ensures consistency in projected survival curves as discussed in De Rossi (2004) and Blackburn and Sherris (2013). The models use continuous-time dynamics for risk

factors that drive changes in the cohort survival curve through time with factor loadings that quantify how the risk factors impact the different ages in the survival curve.

Affine mortality models have been predominantly considered for capturing the mortality dynamics of a single cohort as in Dahl and Møller (2006), Biffis (2005) and Luciano et al. (2008). These models are usually calibrated to age-period mortality data for the reference population as in Schrager (2006) and Blackburn and Sherris (2013). The models do not fit as well at older ages and do not capture the effect of heterogeneity (Pitacco, 2016). Alai et al. (2019) find that the Gamma distribution fits mortality intensities well, which is consistent with mortality heterogeneity.

Cohort effects have been observed in age-period data for many countries, for example Willets (2004), Cairns et al. (2009) and Gallop (2008). The pricing of longevity-linked cash-flows requires the survival rates of a cohort in the reference population (Xu et al., 2015). Recently Jevtic et al. (2013), Xu et al. (2015) and Chang and Sherris (2018) have proposed affine age-cohort mortality models to capture the dynamics of survival probabilities of multiple cohorts.

This paper assesses a range of continuous-time affine cohort mortality models. We propose an affine mortality model based on the Arbitrage-Free Nelson-Siegel (AFNS) model (Christensen et al., 2011) with identifiable factors of level, slope and curvature of the mortality curve. We investigate the impact of incorporating factor dependence to capture correlations for the models. To improve the model fit we use a Gamma distribution based on the Cox-Ingersoll-Ross (CIR) model (Cox et al., 1985) for the affine mortality models. We capture cohort effects directly using age-cohort data to calibrate and assess the model survival curve fit and forecasting performance. Age-period models, that are fitted to age-period data, require additional factors to capture cohort effects. Fitting age-cohort models using age-cohort data captures cohort effects more naturally than age-period models.

This paper is structured as follows. Section 2 provides a general model framework for affine mortality models and specifies the structure of the continuous-time cohort mortality models. Section 3 describes the US mortality data for calibration. We use US data since this is a large developed country with mortality experience typical of such economies. Section 4 outlines the estimation methodology using the Kalman filter and provides an analysis of the estimation results and model comparison. In Section 5, out-of-sample expected survival probabilities are estimated for the latest cohort with full mortality data and the out-of-sample forecasting ability is assessed. Section 6 concludes the paper with a summary and major findings.

2 Affine Mortality Models

We outline the continuous-time model framework of affine mortality models, then introduce the AFNS and the CIR model applied to mortality modeling for multiple cohorts. Finally we discuss the incorporation of factor dependence in the affine models. We initially derive the mortality models in a financial modelling setting with risk-neutral pricing measures. We then give the link between the risk-neutral dynamics and the real world dynamics which allows us to calibrate the models to historical mortality data.

2.1 General Model Framework

The models are formally defined based on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, where Ω is the set of possible states of nature and $\mathbb{F} = \{\mathcal{F}_t\}_{0 \leq t \leq T}$, where $\mathcal{F}_t = \mathcal{H}_t V \mathcal{M}$ is the combined filtration for both the term structure of interest rates and mortality, assumed to satisfy the conditions of right continuity, with $\mathcal{H}_t$ the filtration generated by the term structure of interest rates up to time t, and $\mathcal{M}_t$ the filtration containing all the information generated by the evolution of the survival curves for mortality up to time t.

There is an incomplete market for longevity risk so that there exists no unique risk-neutral measure Q for pricing mortality linked cash flows (Xu et al., 2015). In pricing longevity risk and longevity-linked financial products, the risk-neutral measure Q is defined in terms of the zero-coupon longevity bond (Cairns et al., 2006a; Bauer et al., 2008; Blackburn and Sherris, 2013). The real-world measure P reflects the best estimate of mortality, which is related to historical mortality data (Bauer et al., 2008).

We use $\widetilde{S}(t,T,x)$ to denote the survival probability of an individual aged x at time t surviving to time T. The price of a zero coupon bond paying \$1 at time T is denoted as $P(t,T)$. The price of longevity bond at time t that pays the amount $\widetilde{S}(t,T,x)$ at time T is denoted by $\widetilde{P}(t,T,x)$. If there exists a measure Q, equivalent to the real-world measure P for all t, T, and x, such that

$$P(t,T) = E^Q\left[\exp\left(-\int_t^T r_u du\right)\middle|\mathcal{H}_t\right], \tag{1}$$

$$\widetilde{P}(t,T,x) = E^Q\left[\exp\left(-\int_t^T r_u du\right)\widetilde{S}(t,T,x)\middle|\mathcal{F}_t\right], \tag{2}$$

then the dynamics of the combined financial market are arbitrage-free.

Assuming independence between interests rates and mortality, the price of a zero coupon longevity bond can be written as:

$$\widetilde{P}(t,T,x) = E^Q\left[\exp\left(-\int_t^T r_u du\right)\middle|\mathcal{H}_t\right] E^Q\left[\widetilde{S}(t,T,x)\middle|\mathcal{M}_t\right] = P(t,T)\,S(t,T,x), \tag{3}$$

where $S(t,T,x)$ is the risk-neutral survival probabilty, which is a martingale under the measure Q.

The instantaneous mortality intensity for individuals aged x at time t of a given cohort i is defined as:

$$\mu_x^i(t) = \rho_1' X_t, \tag{4}$$

where $\rho_1 \in \mathbb{R}^n$, and $X_t \in \mathbb{R}^n$ is a vector of n latent factors that are assumed to drive the mortality intensity.

In what follows the subscript x indicating ages and i indicating cohorts will be dropped since we are considering a single cohort. Individuals age as time increases in the cohort.

The dynamics of the latent factors X_t are given by the following system of stochastic differential equations (SDEs) under the risk-neutral measure Q (Duffie and Kan, 1996; Christensen et al., 2011):

$$dX_t = K^Q\left[\theta^Q - X_t\right]dt + \Sigma D(X_t,t)\,dW_t^Q, \tag{5}$$

where $K^Q \in \mathbb{R}^{n\times n}$ is the mean reversion matrix, $\theta^Q \in \mathbb{R}^n$ is the long-term mean, $\Sigma \in \mathbb{R}^{n\times n}$ is the volatility matrix, $W_t^Q \in \mathbb{R}^n$ is a standard Brownian motion, and $D(X_t,t)$ is a diagonal matrix with the ith diagonal entry as $\sqrt{\alpha^i(t) + \beta_1^i(t)\,x_t^1 + \ldots + \beta_n^i(t)\,x_t^n}$. α and β are bounded continuous functions.

We will consider 3-factor affine models since this has been found to be satisfactory in capturing mortality variations at older ages (Blackburn and Sherris, 2013).

Under these dynamics the risk-neutral survival probabilities for any age x from time t to time T can be represented as (Blackburn and Sherris, 2013):

$$S(t,T) = E^Q\left[\exp\left(-\int_t^T \mu(s)\,ds\right)\right] = \exp\left(B(t,T)'\,X_t + A(t,T)\right), \tag{6}$$

where $B(t,T)$ and $A(t,T)$ are the solutions to the following system of ordinary differential equations (ODEs):

$$\frac{dB(t,T)}{dt} = \rho_1 + \left(K^Q\right)' B(t,T), \tag{7}$$

$$\frac{dA(t,T)}{dt} = -B(t,T)' K^Q \theta^Q - \frac{1}{2}\sum_{j=1}^{3}\left(\Sigma' B(t,T) B(t,T)' \Sigma\right)_{j,j}, \tag{8}$$

with boundary conditions $B(T,T) = A(T,T) = 0$.

The survival probability can be written also in terms of the average force of mortality over the duration $(T-t)$ for any age x. The average force of mortality is affine in the latent factors and is defined as (Blackburn and Sherris (2013); Xu et al. (2015)):

$$\bar{\mu}(t,T) = -\frac{1}{T-t}\log\left[S(t,T)\right] = -\frac{B(t,T)'}{T-t}X_t - \frac{A(t,T)}{T-t}. \tag{9}$$

This allows us to construct the survival curve given the factors and the factor loadings.

2.2 Multi-Factor Affine Cohort Mortality Models

We present the dynamics of the latent factors in 3-factor affine models. For the different models we specify the dynamics for Equations (4) and (5) and give the solutions to the ODEs for Equations (7) and (8).

2.2.1 Independent Factor Models with Gaussian Processes

We first consider 3-factor affine mortality models with factors following Gaussian processes. We consider the 3-factor independent model in Blackburn and Sherris (2013) (hereafter Blackburn-Sherris model). We also consider an affine mortality model based on the AFNS interest rate term structure model with the survival curve driven by factors for level (L_t), slope (S_t), and curvature (C_t).

Table 1 summarises the assumptions for these models. The independent factor models have an identity matrix for $D(X_t,t)$ in Equation (5). The long-term mean θ^Q in the models is assumed to be a vector of zeros for both models, as explained in Blackburn and Sherris (2013).

The model dynamics are given as follows:

- The independent Blackburn-Sherris model has instantaneous mortality rate given by

$$\mu(t) = X_t^1 + X_t^2 + X_t^3, \tag{10}$$

  with $\rho_1 = (1,1,1)^T$ and $X_t = \left(X_t^1, X_t^2, X_t^3\right)$ in Equation (4).

  The dynamics of the state variables X_t have the following form under the risk-neutral measure Q

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = -\begin{pmatrix} \delta_{11} & 0 & 0 \\ 0 & \delta_{22} & 0 \\ 0 & 0 & \delta_{33} \end{pmatrix}\begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix}\begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix}. \tag{11}$$

- The independent AFNS mortality model has instantaneous mortality rate given by

$$\mu(t) = L_t + S_t, \tag{12}$$

  with $\rho_1 = (1,1,0)^T$ and $X_t = (L_t, S_t, C_t)$ in Equation (4).

The dynamics of the factors under the Q-measure are given by:

$$\begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} = -\begin{pmatrix} 0 & 0 & 0 \\ 0 & \delta & -\delta \\ 0 & 0 & \delta \end{pmatrix} \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix}. \quad (13)$$

Table 1: Affine Mortality Models - Independent Factor Model Specifications

Model	Factors X_t	ρ_1	K^Q	K^P	Σ
Blackburn-Sherris Model	$\begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} \delta_1 & 0 & 0 \\ 0 & \delta_2 & 0 \\ 0 & 0 & \delta_3 \end{pmatrix}$	$\begin{pmatrix} k_1^P & 0 & 0 \\ 0 & k_2^P & 0 \\ 0 & 0 & k_3^P \end{pmatrix}$	$\begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix}$
AFNS Model	$\begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & \delta & -\delta \\ 0 & 0 & \delta \end{pmatrix}$	$\begin{pmatrix} k_1^P & 0 & 0 \\ 0 & k_2^P & 0 \\ 0 & 0 & k_3^P \end{pmatrix}$	

The solutions for the survival curve require $B(t,T)$, the factor loadings, and $A(t,T)$ from Equations (7) and (8) which can be explicitly solved. The results are:

- The independent Blackburn-Sherris model (Blackburn and Sherris, 2013)

$$B^j(t,T) = -\frac{1-e^{-\delta_{jj}(T-t)}}{\delta_{jj}}, \quad j=1,2,3, \quad (14)$$

$$A(t,T) = \frac{1}{2}\sum_{j=1}^{3}\frac{\sigma_{jj}^2}{\delta_{jj}^3}\left[\frac{1}{2}\left(1-e^{-2\delta_{jj}(T-t)}\right) - 2\left(1-e^{-\delta_{jj}(T-t)}\right) + \delta_{jj}(T-t)\right]. \quad (15)$$

- The independent AFNS model (Christensen et al., 2011)

$$\begin{aligned} B^1(t,T) &= -(T-t), \quad B^2(t,T) = -\frac{1-e^{-\delta(T-t)}}{\delta}, \\ B^3(t,T) &= (T-t)e^{-\delta(T-t)} - \frac{1-e^{-\delta(T-t)}}{\delta}, \end{aligned} \quad (16)$$

$$\begin{aligned} \frac{A(t,T)}{T-t} = {} & \sigma_{11}^2\frac{(T-t)}{6} + \sigma_{22}^2\left[\frac{1}{2\delta^2} - \frac{1}{\delta^3}\frac{1-e^{-\delta(T-t)}}{T-t} + \frac{1}{4\delta^3}\frac{1-e^{-2\delta(T-t)}}{T-t}\right] + \\ & \sigma_{33}^2\left[\frac{1}{2\delta^2} + \frac{1}{\delta^2}e^{-\delta(T-t)} - \frac{1}{4\delta}(T-t)e^{-2\delta(T-t)} - \frac{3}{4\delta^2}e^{-2\delta(T-t)}\right. \\ & \left. -\frac{2}{\delta^3}\frac{1-e^{-\delta(T-t)}}{T-t} + \frac{5}{8\delta^3}\frac{1-e^{-2\delta(T-t)}}{T-t}\right]. \end{aligned} \quad (17)$$

The factor loadings for the independent Blackburn-Sherris model have the same functional form and differ because of the value of the fitted δ_{jj} values which results in different impacts of the factors across ages for the cohort curve. For the independent AFNS model the factor loadings have direct interpretation. $B^1(t,T)$ is level so that the level factor impacts all ages in the cohort curve the same. $B^2(t,T)$ is increasing so that the slope factor impacts older ages more than younger ages. $B^3(t,T)$ is decreasing and produces curvature in the survival curve through time.

The selection of ρ_1 and the specific structure of the mean reversion matrix K^Q in Equation (13) ensure that the factor loadings $-\frac{B(t,T)}{T-t}$ (Equation (9)) of the AFNS model maintain the exact Nelson-Siegel structure such that the latent factors can be interpreted as level, slope and

curvature factors driving the mortality curves (Diebold and Li, 2006; Diebold and Rudebusch, 2013). Moreover, since the AFNS model is derived from the ATSM in Duffie and Kan (1996), this model maintains the arbitrage-free affine structure, which is suitable for financial and pricing applications (Christensen et al., 2011).

Although Björk and Christensen (1999) argues that the Nelson-Siegel model does not satisfy the consistency requirement (proposed by Björk and Christensen, 1999), Diebold and Rudebusch (2013) explain that failing to meet the consistency requirement is due to the failure of the Nelson-Siegel model to link the parameters in the state transition equation to the parameters in the measurement equation in the state space form. The AFNS model establishes this link in the yield-adjustment term, which is $-\frac{A(t,T)}{T-t}$.

2.2.2 Dependent Factor Models with Gaussian Processes

In the independent factor models all ages in the cohort survival curve are impacted by the common factors to a greater or lesser extent as determined by the factor loadings. Factor dependence is incorporated to capture the correlation between different ages.

We give the factor dynamics and solutions to the dependent factor models with Gaussian processes. For both of the Blackburn-Sherris model and the AFNS model, we assume the volatility matrix Σ to be lower-triangular which allows correlated shocks in the models. Although correlation can also be incorporated through K^Q, for the AFNS model the structure of K^Q has to be the same as for the independent factor model in order to preserve the Nelson-Siegel structure for the survival curve factors.

The risk-neutral dynamics of the dependent factor models are specified as:

- The dependent Blackburn-Sherris model

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = - \begin{pmatrix} \delta_{11} & 0 & 0 \\ \delta_{21} & \delta_{22} & 0 \\ \delta_{31} & \delta_{32} & \delta_{33} \end{pmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix}. \quad (18)$$

- The dependent AFNS model

$$\begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} = - \begin{pmatrix} 0 & 0 & 0 \\ 0 & \delta & -\delta \\ 0 & 0 & \delta \end{pmatrix} \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix}. \quad (19)$$

For completeness, the factor loadings $B(t,T)$ and $A(t,T)$ of the dependent Blackburn-Sherris model are given in Appendix A. For the dependent AFNS model the factor loadings $B(t,T)$ are the same as for the independent factor model and an explicit expression for $A(t,T)$ is given in Christensen et al. (2011).

2.2.3 The Cox-Ingersoll-Ross Mortality Model

The Gaussian models considered so far allow mortality rates to become negative, although as we will show later, this is empirically small for most of the models we consider. To avoid negative mortality rates we consider a multi-factor affine mortality model with each factor following a square-root process, based on the Cox-Ingersoll-Ross model (CIR) (Cox et al., 1985) frequently used as an affine term structure models for interest rates and credit risk.

Another benefit of the Cox-Ingersoll-Ross model (CIR) model is that it can capture the effect of mortality heterogeneity. Under the CIR mortality model, mortality rates follow a non-central Chi-square distribution and are asymptotically Gamma distributed (Cox et al., 1985).

Following Chen and Scott (2003) and Geyer and Pichler (1999) for interest rates, we define the instantaneous mortality intensity as before to be affine with:

$$\mu_x^i(t) = \rho_1' X_t = X_t^1 + X_t^2 + X_t^3, \tag{20}$$

where $X_t = \left(X_t^1, X_t^2, X_t^3\right)$ are the state variables that are driving the mortality intensity and ρ_1 is assumed to be $(1,1,1)^T$.

The factor dynamics driving the mortality survival curve are then described by the following system of SDEs under the risk-neutral measure Q:

$$\begin{aligned}\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = &- \begin{pmatrix} \delta_{11} & 0 & 0 \\ 0 & \delta_{22} & 0 \\ 0 & 0 & \delta_{33} \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \\ \theta_3^Q \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt \\ &+ \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{X_t^2} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix}.\end{aligned} \tag{21}$$

With this structure, each factor follows a single-factor CIR process, a square-root process. The matrix $D(X_t, t)$ in Equation (5) is defined as a diagonal matrix with the j-th element on the diagonal as $\sqrt{X_t^j}$ $(j = 1, 2, 3)$.

The explicit expressions for $B(t,T)$ and $A(t,T)$ are

$$B^j(t,T) = -\frac{2\left(e^{\gamma_j(T-t)} - 1\right)}{\left(\delta_{jj} + \gamma_j\right)\left(e^{\gamma_j(T-t)} - 1\right) + 2\gamma_j}, \quad j = 1, 2, 3, \tag{22}$$

$$A(t,T) = \sum_{j=1}^{3} \frac{2\delta_{jj}\theta_j^Q}{\sigma_{jj}^2} \ln\left[\frac{2\gamma_j \exp\left(\frac{(\delta_{jj}+\gamma_j)(T-t)}{2}\right)}{\left(\delta_{jj} + \gamma_j\right)\left(e^{\gamma_j(T-t)} - 1\right) + 2\gamma_j}\right], \tag{23}$$

with $\gamma_j = \sqrt{\delta_{jj}^2 + 2\sigma_{jj}^2}$, $\quad j = 1, 2, 3$ (Duan and Simonato, 1999; Chen and Scott, 2003; Geyer and Pichler, 1999).

2.3 Real World Dynamics and Change of Measure

The affine mortality models are specified under the risk-neutral measure Q. Since we fit the models using historical data, this measure has to be changed to the real-world measure P. From Girsanov′s theorem, the relationship between the dynamics under the measure P and under the measure Q is given by:

$$dW_t^Q = dW_t^P + \Lambda_t dt, \tag{24}$$

where Λ_t is the risk premium.

To specify the structure of the risk premium of longevity risk, we adopt the essentially affine model proposed by Duffee (2002). The essentially affine model removes the strong link between the factor loadings $-\frac{B(t,T)}{T-t}$ and the drift term under the real-world measure (Blackburn and Sherris, 2013) and preserves the affine dynamics under the P-measure (Christensen et al., 2011).

The form of the risk premium is (Duffee, 2002):

$$\Lambda_t = \begin{cases} \lambda^0 + \lambda^1 X_t, & \text{for models with Gaussian processes;} \\ D(X_t, t)\,\lambda^0, & \text{for the CIR model.} \end{cases} \tag{25}$$

where $\Lambda_t \in \mathbb{R}^{n\times 1}$, $\lambda^0 \in \mathbb{R}^{n\times 1}$ and $\lambda^1 \in \mathbb{R}^{n\times n}$.

With these assumptions, the SDEs for factors under the measure P can be written as:

$$dX_t = \begin{cases} K^P\left[\theta^P - X_t\right]dt + \Sigma dW_t^P, & \text{for models with Gaussian processes;} \\ K^P\left[\theta^P - X_t\right]dt + \Sigma D\left(X_t, t\right)dW_t^P, & \text{for the CIR model.} \end{cases} \tag{26}$$

The form of K^P and θ^P are derived in Appendix B. In the essentially affine model we are free to choose the mean reversion matrix K^P and the mean vector θ^P.

Under the real-world measure P, the dynamics of the factors in each model that we will estimate from historical mortality data are:

- The independent Blackburn-Sherris model (Blackburn and Sherris, 2013)

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = -\begin{pmatrix} k_{11}^P & 0 & 0 \\ 0 & k_{22}^P & 0 \\ 0 & 0 & k_{33}^P \end{pmatrix}\begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix}dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix}\begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}. \tag{27}$$

- The independent AFNS model (Christensen et al., 2011)

$$\begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} = -\begin{pmatrix} k_{11}^P & 0 & 0 \\ 0 & k_{22}^P & 0 \\ 0 & 0 & k_{33}^P \end{pmatrix}\begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix}dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix}\begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}. \tag{28}$$

- The dependent Blackburn-Sherris model

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = -\begin{pmatrix} k_{11}^P & 0 & 0 \\ 0 & k_{22}^P & 0 \\ 0 & 0 & k_{33}^P \end{pmatrix}\begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix}dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}\begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}. \tag{29}$$

- The dependent AFNS model

$$\begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} = -\begin{pmatrix} k_{11}^P & 0 & 0 \\ 0 & k_{22}^P & 0 \\ 0 & 0 & k_{33}^P \end{pmatrix}\begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix}dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}\begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}, \tag{30}$$

- The CIR model

$$\begin{aligned} \begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = &-\begin{pmatrix} k_{11}^P & 0 & 0 \\ 0 & k_{22}^P & 0 \\ 0 & 0 & k_{33}^P \end{pmatrix}\left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix}\right]dt \\ &+ \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix}\begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{X_t^2} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix}\begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}. \end{aligned} \tag{31}$$

3 Mortality Data

The US mortality data from the Human Mortality Database (2017) (HMD) is used to calibrate and compare the mortality models. We use age-cohort data as opposed to age-period data since

we are estimating cohort models (Blackburn and Sherris, 2013; Xu et al., 2015; Chang and Sherris, 2018).

We extract the mortality data of males from ages 50 to 100 for the cohorts born from 1883 to 1915. We are interested in the older ages for post-retirement applications. This gives us complete mortality rate data for each complete cohort. The cohort death rates are obtained by reading the age-period life table diagonally.

We then determine the historical survival probability, $S^i(x;t,T)$, and the historical average forces of mortality $\bar{\mu}^i(x;t,T)$ over the period $\tau = T-t$ for each cohort i aged x at time t from the data, using:

$$S^i(x;t,T) = \prod_{s=1}^{T-t} \left[1 - q^i(x+s-1, t+s-1)\right], \tag{32}$$

$$\bar{\mu}^i(x;t,T) = -\frac{1}{T-t} \log\left[S^i(x;t,T)\right], \tag{33}$$

where $q^i(x,t)$ is the one year death probability for an individual aged x at time t in cohort i.

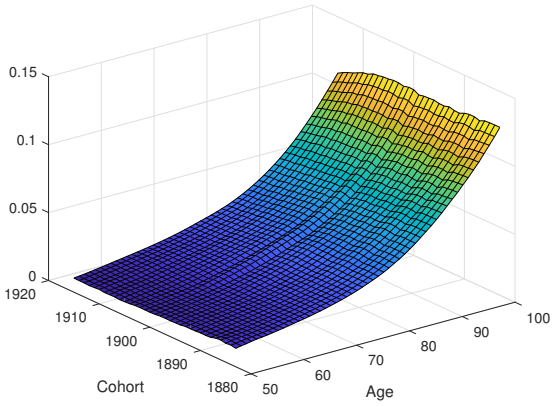


Figure 1: Average Force of Mortality for Males Born from 1883 to 1915

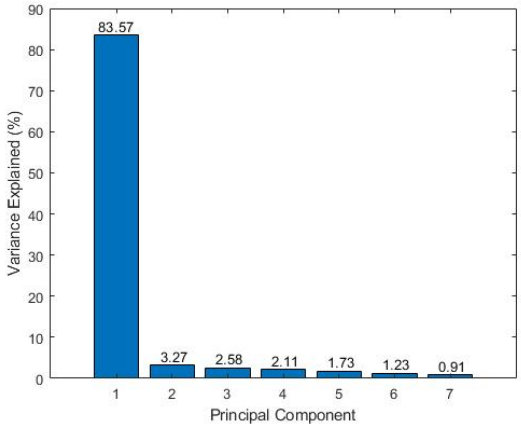


Figure 2: Fractions (%) of Variance Explained by Each of the First 7 Principal Components

The average force of mortality for cohorts born between 1883 and 1915, aged 50 to 100, is shown in Figure 1. Mortality improvement across cohorts is seen from the downward trend of the average force of mortality at each age. The rate of mortality improvement differs by age. The average force of mortality of each cohort grows exponentially in each cohort.

Figure 2 shows the principal component analysis (PCA) for the change of mortality intensity for all cohorts as they age. The first three principal components are able to explain approximately 90% of the total variance which supports our choice of 3-factor affine mortality models while maintaining parsimony. This is consistent with the analysis in Blackburn and Sherris (2013).

4 Model Assessment and Comparison

We estimate the model parameters for all the models using the Kalman filter. The fitted models are then compared using a number of model selection criteria.

4.1 Parameter Estimation

The Kalman filter (Kalman, 1960) with maximum likelihood estimation is used to estimate the parameters in the affine mortality models, following Christensen et al. (2011) and Blackburn

and Sherris (2013). Since our models only capture the volatility of the mortality rates we need to account for the exponentially increasing Poisson variation in the historical data reflecting the size of the population at each age. We do this in the measurement equation for the Kalman filter (Xu et al., 2018).

The estimation process is as follows:

1. Represent the affine mortality models in the state space form which consists of two components, the measurement equation and the state transition equation (Xu et al., 2015; Shumway and Stoffer, 2017).

   The measurement equation describes the affine relationship between the average force of mortality and the state variables (Xu et al., 2015; Durbin and Koopman, 2012). Based on Blackburn and Sherris (2013) and Xu et al. (2015), the measurement equation in terms of the average forces of mortality is:

$$\bar{\mu}(t,T) = -\frac{B(t,T)'}{T-t}X_t - \frac{A(t,T)}{T-t} + \varepsilon_t, \quad \varepsilon_t \sim N(0,H), \tag{34}$$

   where the measurement error ε_t is independently and identically distributed noise with the covariance matrix of the measurement error, H, being diagonal.

   To capture the increasing nature of the Poisson variation, the parameteric form assumed for the diagonal of the covariance matrix H is

$$H(t,T) = \frac{1}{T-t}\sum_{i=1}^{T-t}\left[r_c + r_1 e^{r_2 i}\right], \tag{35}$$

   where the values of r_c, r_1 and r_2 are estimated from the data.

   The state transition equation represents the unobserved dynamics of the state variables (Xu et al., 2015; Durbin and Koopman, 2012) and is given by:

$$X_t = \exp\left(-K^P\right)X_{t-1} + \eta_t, \quad \eta_t \sim N(0,R), \tag{36}$$

   where η_t is the transition error vector with diagonal matrix R the covariance matrix of the transition error.

   The matrix R has the following structure:

$$R = \int_{t-1}^{t} e^{-K^P(t-s)}\Sigma\Sigma' e^{-\left(K^P\right)'(t-s)}ds. \tag{37}$$

2. Use the Kalman filter to evaluate the likelihood function of affine mortality models and to extract the values of the state variables. The information available at time t is denoted by $Y_t = (y_1, \ldots, y_t)$ and the model parameters given by ψ.

   In the forecasting step, using the state update X_{t-1} and its mean square error Σ_{t-1} obtained at $t-1$,

$$X_{t|t-1} = E[X_t|Y_{t-1}] = \Phi(\psi)X_{t-1}, \tag{38}$$

$$\Sigma_{t|t-1} = \Phi(\psi)\Sigma_{t-1}\Phi(\psi)' + R(\psi), \tag{39}$$

   where $\Phi = \exp\left(-K^P\right)$.

   In the update step, the information at time t, Y_t, is used to update the forecasts $X_{t|t-1}$

and we obtain:

$$X_t = E\left[X_t|Y_t\right] = X_{t|t-1} + \Sigma_{t|t-1}B\left(\psi\right)^{'} F_t^{-1}\nu_t, \tag{40}$$

$$\Sigma_t = \Sigma_{t|t-1} - \Sigma_{t|t-1}B\left(\psi\right)^{'} F_t^{-1}B\left(\psi\right)\Sigma_{t|t-1}, \tag{41}$$

where

$$\nu_t = y_t - E\left[y_t|Y_{t-1}\right] = y_t - A\left(\psi\right) - B\left(\psi\right)X_{t|t-1}, \tag{42}$$

$$F_t = cov\left(\nu_t\right) = B\left(\psi\right)\Sigma_{t|t-1}B\left(\psi\right)^{'} + H\left(\psi\right). \tag{43}$$

3. Evaluate the following log-likelihood function with the values obtained in the previous step:

$$\log L\left(y_1, \ldots, y_t; \psi\right) = \sum_{t=1}^{T}\left(-\frac{N}{2}\log\left(2\pi\right) - \frac{1}{2}\log\left|F_t\right| - \frac{1}{2}\nu_t^{'}F_t\nu_t\right), \tag{44}$$

where N is the number of observed average forces of mortality.

The log-likelihood function is maximized with respect to ψ to obtain the optimal parameter set. For the CIR mortality model we use quasi-maximum likelihood estimation.

4.2 Model Parameter Estimation Results

Table 2 summarizes the parameter estimates for each model, along with the standard errors. The standard errors determine the significance of factor loadings, which further indicates whether the corresponding factor has a significant influence on mortality rates. The risk neutral parameter values are reported as well as the real world values.

The parameters in K^Q ($\delta^{'}$s) show the impact of each of the factors on changes in the factors and summarise the significance of the factor loadings at the different ages. The diagonal components are largely negative. The $\delta^{'}$s in the dependent-factor AFNS model are smaller in absolute value than those in the independent-factor AFNS model. For the two AFNS models, $\delta^{'}$s are both negative which results in the sensitivity at older ages for the slope factor.

The mean reversion K^P parameters give the speed to revert to the long-term mean. These values vary across the models reflecting different rates of mean reversion in the real world measure. There are negative correlations between factors in the two dependent-factor models, which impacts the adjustment term $A\left(t,T\right)$.

All the long-term mean θ parameters in the CIR model are positive which ensures positive factors and hence positive mortality rates in the model. The second factor, X^2, has the largest mean reversion speed, k_{22}^P, and largest volatility, σ_{22}, so this factor is more related to the short term (Geyer and Pichler (1999)). The mean reversion rate k_{11}^P and the volatility σ_{11} of X^1 are lowest, compared with the other two factors. The first factor has less impact on the mortality dynamics and is less volatile.

Because of the structure of the matrix H in Equation (35), the measurement errors are age-dependent and exponentially increasing with age. By comparing the r_2, the scalar in the exponential function in matrix H, the independent-factor Blackburn-Sherris model has the largest r_2, so that there is a larger measurement error volatility estimated for this model. Values of all parameters in matrix H of the CIR model are negligible, indicating smaller measurement errors and reflecting a better in-sample model fit.

4.3 Assessing Model Goodness-of-Fit

Table 3 shows, for each model, the Root Mean Square Error (RMSE), the Akaike information criterion (AIC), and the Bayesian information criterion (BIC). Since the models with Gaussian

Table 2: Estimated Parameters

	The Blackburn-Sherris Model		The AFNS Model		The CIR Model
	Independent-Factor	Dependent-Factor	Independent-Factor	Dependent-Factor	
δ_{11} (AFNS: δ)	-0.01106 (0.00123)	-0.20183 (2.944e-05)	-0.08348 (1.580e-04)	-0.04725 (8.096e-05)	-0.09652 (2.513e-04)
δ_{21}	-	0.56206 (4.091e-05)	-	-	-
δ_{22}	0.07484 (0.00432)	-0.07092 (1.766e-05)	-	-	0.12627 (2.412e-03)
δ_{31}	-	0.24075 (1.555e-05)	-	-	-
δ_{32}	-	0.80809 (4.102e-05)	-	-	-
δ_{33}	-0.06883 (2.452e-04)	0.77825 (1.461e-05)	-	-	-0.11153 (3.060e-04)
k_{11}^P	0.38753	-0.04248	0.18793	0.01810	0.00077
k_{22}^P	0.13910	0.01869	0.01361	0.02002	0.59402
k_{33}^P	0.00718	0.01827	0.02701	0.04972	0.06842
σ_{11}	0.00782	7.557e-11	9.593e-04	0.00400	0.00265 (6.894e-05)
σ_{21}	-	0.01110	-	-0.00387	-
σ_{22}	0.00125	3.370e-11	1.120e-04	0.00091	0.02848 (1.250e-03)
σ_{31}	-	-0.01190	-	-0.00183	-
σ_{32}	-	0.00047	-	0.00123	-
σ_{33}	5.409e-04	0.00029	3.549e-05	0.00023	0.01360 (9.494e-05)
r_1	1.071e-11	4.337e-08	1.422e-10	6.272e-08	5.498e-10
r_2	0.37797	0.11375	0.17784	0.10742	6.646e-07
r_c	4.360e-08	5.705e-08	4.963e-07	4.636e-13	3.410e-07
The CIR Model					
θ_1^Q	0.00080	θ_2^Q	0.01010	θ_3^Q	0.00137
θ_1^P	0.00697	θ_2^P	0.00415	θ_3^P	0.00356

processes allow for negative mortality, we show the probabilities of negative mortality for these models.

Table 3: Comparison of Affine Mortality Models

	The Blackburn-Sherris Model		The AFNS Model		The CIR Model
	Independent-Factor	Dependent-Factor	Independent-Factor	Dependent-Factor	
Log Likelihood	9896.419	9938.696	9665.801	9887.878	10045.70
RMSE	0.00250	7.601e-04	6.856e-04	9.160e-04	5.227e-04
No. of Parameteres	12	18	10	13	18
AIC	-19570.837	-19643.392	-19113.602	-19551.757	-19857.40
BIC	-18968.292	-19008.277	-18521.914	-18943.783	-19222.29
Probability of Negative Mortality	0.02700	1.011e-32	1.722e-31	4.34e-14	-

In terms of model selection, the CIR model has the highest log-likelihood and the smallest RMSE. The AIC and the BIC of the CIR model show this to be a better model even though it has more parameters than most of the other models. The CIR model, by construction, precludes the probability of negative mortality.

The Gaussian models perform well, particularly the dependent-factor models. The dependent factor Blackburn-Sherris model and the AFNS models all have low probabilities of negative mortality rates. The dependent-factor Blackburn-Sherris model has the largest log-likelihood and better AIC and BIC than the other Gaussian models.

4.4 Factors and Factor Loadings

We consider the factors and the factor loadings for the cohort survival curves for the independent AFNS mortality model, where the factors for level, slope and curvature have a direct interpretation, and the CIR mortality model which is the best performing model.

Figures 3 and 4 show the fitted values of the factors and factor loadings of the independent AFNS model. For the factor loading B_1 the impact of the level factor L is constant across all ages. There is an increase in the factor level for all ages for the cohorts born around 1900 onwards.

The factor loading B_2 increases exponentially, so it impacts older ages more than younger ages and mortality rates at older ages are therefore more sensitive to the slope factor S. For cohorts born after 1900, corresponding to the rise in the level factor L, there is a decline in the slope factor S.

The factor loading B_3 is negative and decreasing across all ages and, as a result, the convexity of the survival curve at older ages decreases faster than at younger ages. For cohorts born after 1900, the decline in C results in mortality rate curves that are less convex across age. This corresponds to mortality improvement at older ages being larger than for the younger ages.

The adjustment term A in the survival curve is negative and decreasing.

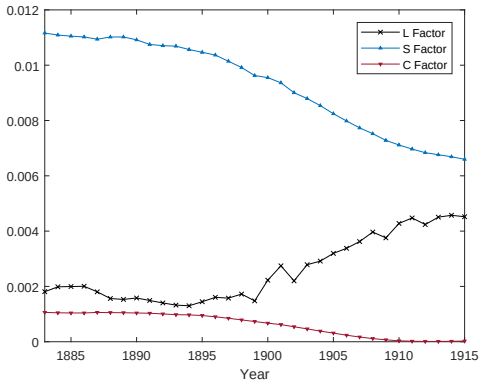
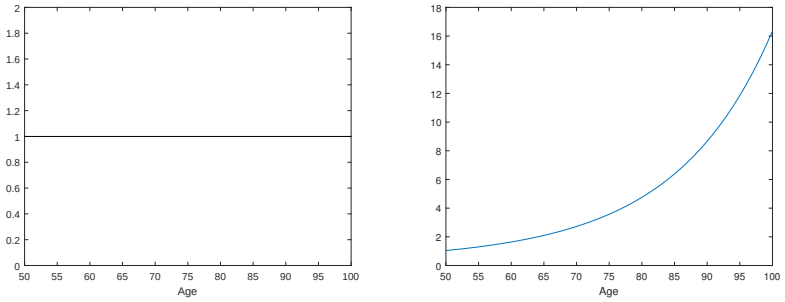


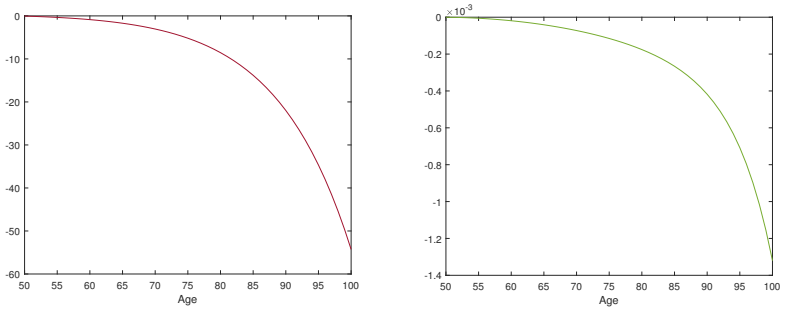
Figure 3: Factors in the Independent AFNS Model

Figures 5 and 6 show the estimated latent factors and factor loadings for the CIR mortality model. The factors and factor loadings are quite different to those of the independent AFNS mortality model. The first factor X^1 is relatively stable for cohorts born before 1900, then increases, followed by a moderate decline for cohorts born around 1910 and after. The factor loading, B_1, is positive and increases with age, so that the first factor X^1 impacts older ages more than younger ages.

The second factor X^2 shows a general downward trend across time. The factor loading B_2 is



(a) $B_1 = -\frac{B^1(t,T)}{T-t}$ (b) $B_2 = -\frac{B^2(t,T)}{T-t}$



(c) $B_3 = -\frac{B^3(t,T)}{T-t}$ (d) $A = -\frac{A(t,T)}{T-t}$

Figure 4: Factors Loadings in the Independent AFNS Model

decreasing with age and smaller than B_1, so that younger ages show mortality improvement, but the size of improvement from this factor is smaller.

The third factor X^3 is relatively constant for cohorts born before 1900 and and decreases afterwards, with a moderation in the rate of decrease for cohorts born after around 1910. The factor loading B_3 is positive and impacts older ages more than younger ages, with a convex shape, which also results in curvature changes in the survival curve.

The adjustment term A in the survival curve is negative and, following a small increase for younger ages, is then decreasing.

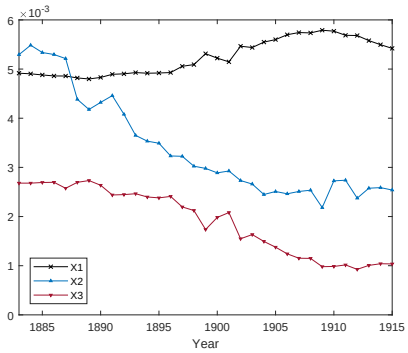
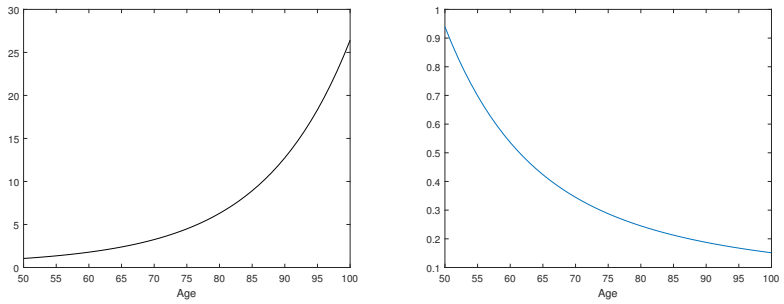


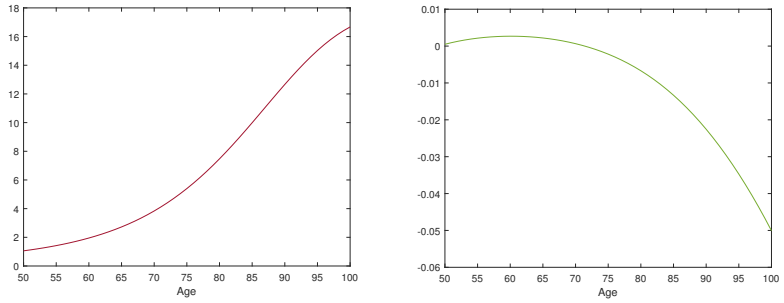
Figure 5: Factors in the CIR Model

4.5 Residual Analysis

Figure 7 plots the residuals of the mortality models. The residuals are the differences between the average force of mortality from the historical mortality data and those determined from the fitted mortality models. Plots are on the same scale on the z-axis except for the independent



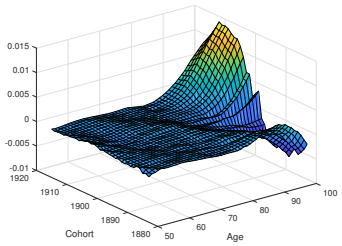
(a) $B_1 = -\frac{B^1(t,T)}{T-t}$ (b) $B_2 = -\frac{B^2(t,T)}{T-t}$



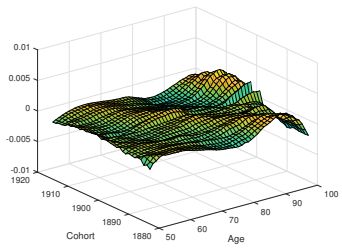
(c) $B_3 = -\frac{B^3(t,T)}{T-t}$ (d) $A = -\frac{A(t,T)}{T-t}$

Figure 6: Factors Loadings in the CIR Model

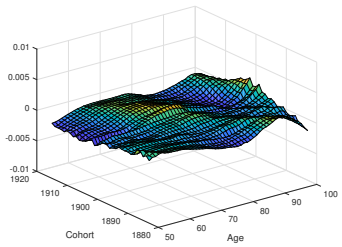
Blackburn-Sherris model which has large residuals at older ages reflecting a poorer fit at these ages.



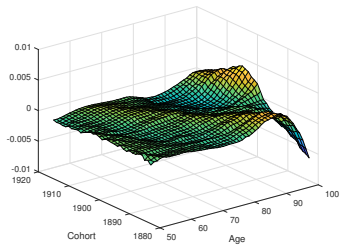
(a) The Independent Blackburn-Sherris Model



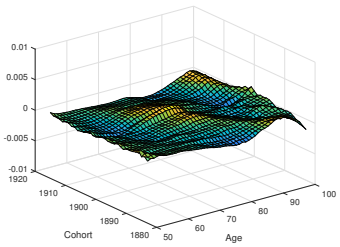
(b) The Dependent Blackburn-Sherris Model



(c) The Independent AFNS Model



(d) The Dependent AFNS Model



(e) The CIR Model

Figure 7: Residuals of Affine Mortality Models

Apart from the independent Blackburn-Sherris model, all of the models have similar residual plots. The independent AFNS model (Figure 7(c)) with three factors for level, slope and curvature has a better fit than the independent Blackburn-Sherris model. The AFNS model

reduces the magnitude of residuals at the older ages in the independent Blackburn-Sherris model without adding additional parameters. The use of level, slope and curvature factors better capture the variation in mortality curves, especially at older ages.

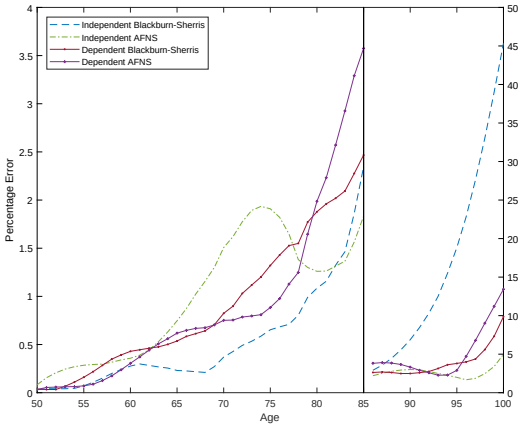
Introducing factor dependence in the Blackburn-Sherris model also reduces the size of residuals and better accounts for the mortality variation at older ages. This can be seen by comparing the dependent Blackburn-Sherris model (Figure 7(b)) with the independent model (Figure 7(a)). This is not the case for the independent AFNS model in Figure 7(c). The size of the residuals in the dependent AFNS model (Figure 7(d)) is larger, particularly at older ages. The independent AFNS model is able to capture the mortality variability better than the dependent AFNS model.

The CIR model has lower residuals at older ages and an overall flatter residual surface similar to the residuals of the independent AFNS model in Figure 7(c). The CIR model provides slightly smaller residuals at older ages and ages younger than 60, in terms of absolute values compared with the other models.

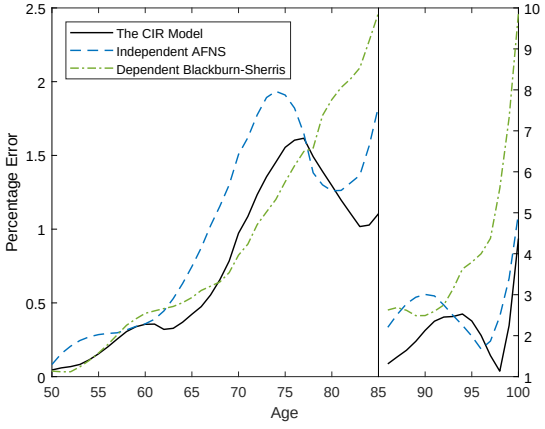
In all of the residual plots there is a hump shape running diagonally across the cohorts. This reflects a period mortality factor that has impacted all of the cohorts around the year 1970 when period mortality improvement trends experienced a significant change to a higher level of improvement. For the later cohorts at the older ages residuals are also higher reflecting the recent slowing of mortality improvement rates.

4.6 In-Sample Analysis

We also use an in-sample model performance analysis by comparing the estimated cohort survival probabilities from the fitted mortality models with the cohort survival probabilities from the historical data. Figure 8 summarizes the in-sample model fit results using the Mean Absolute Percentage Error (MAPE) for each age, across all cohorts. To reflect the difference in values, the scale of the percentage error is different above and below age 85.



(a) The Models with Gaussian Processes



(b) The CIR Model, the Dependent Blackburn-Sherris Model and the Independent AFNS Model

Figure 8: MAPE of Affine Mortality Models

Figure 8(a) shows the MAPE for the affine mortality models with Gaussian processes. Below age 85, all models have similar performance and the differences between the percentage errors of the different mortality models are relatively small. Above the age of 85, the independent Blackburn-Sherris model produces significantly larger percentage errors, while the dependent Blackburn-Sherris model and the independent AFNS model provide similar and improved model fit.

In Figure 8(b), the Gaussian mortality models with the better performance are compared with the CIR mortality model. The CIR mortality model and the independent AFNS mortality model are very similar, with the latter having only slightly larger percentage errors at most ages. The dependent Blackburn-Sherris model is similar to the CIR model below age 75, but the percentage errors of the Blackburn-Sherris model increase quickly after age 75 with values as high as 10%.

Both the independent AFNS mortality model and the CIR mortality model provide satisfactory performance based on MAPE for the historical age-cohort data for complete cohorts.

5 Forecasts of Survival Probabilities

To compare the predictive performance of the affine mortality models we use an out-of-sample forecast with the fitted parameter values estimated from the cohorts born 1883 to 1915 for ages 50 to 100 to forecast the survival curve of the cohort born in 1916. For this cohort we have full historical mortality data.

Following Christensen et al. (2011), who uses the optimal forecasts for predicting yields to maturity, we use the optimal forecasts, also referred to as the best-estimate forecasts, to project average forces of mortality and survival probabilities.

At time t, the average force of mortality over $\tau = (T+1) - (t+1)$ periods at time $t+1$ for cohort i, $\bar{\mu}^i(t+1, T+1)$, is

$$\bar{\mu}^i(t+1, T+1) = -\frac{B(t,T)'}{T-t} E[X_{t+1}|X_t] - \frac{A(t,T)}{T-t}, \tag{45}$$

where $B(t,T)$ and $A(t,T)$ only depend on $\tau = T - t$.

The forcasts of survival probabilities are then:

$$S(t+1, T+1) = \exp\left(B(t,T)' E[X_{t+1}|X_t] + A(t,T)\right). \tag{46}$$

Since the factor dynamics under measure P in the independent Blackburn-Sherris model and the 3-factor independent AFNS model, are the same, the conditional expectation of state variables for these two models are as follows:

$$E\left[X^1_{t+1}|X^1_t\right] = e^{-k^P_{11}} X^1_t, \quad E\left[X^2_{t+1}|X^2_t\right] = e^{-k^P_{22}} X^2_t, \quad E\left[X^3_{t+1}|X^3_t\right] = e^{-k^P_{33}} X^3_t. \tag{47}$$

For the independent AFNS model, the conditional mean has the same structure but with $X_t = (L_t, S_t, C_t)$.

The SDEs describing the P-dynamics of the dependent Blackburn-Sherris model and the dependent AFNS model are the same as for the independent factor model in Equation (47).

The conditional mean of the CIR model is given in Geyer and Pichler (1999) so that for the mortality model:

$$\begin{gathered} E\left[X^1_{t+1}|X^1_t\right] = e^{-k^P_{11}} X^1_t + \theta^P_1\left(1 - e^{-k^P_{11}}\right), \quad E\left[X^2_{t+1}|X^2_t\right] = e^{-k^P_{22}} X^2_t + \theta^P_2\left(1 - e^{-k^P_{22}}\right), \\ E\left[X^3_{t+1}|X^3_t\right] = e^{-k^P_{33}} X^3_t + \theta^P_3\left(1 - e^{-k^P_{33}}\right). \end{gathered} \tag{48}$$

The RMSE for projecting the 1916 cohort survival curve under each mortality model are shown in Table 4. The indepedent AFNS mortality model performs best. The dependent AFNS mortality model and the dependent Blackburn-Sherris models perform similarly. The independent

Blackburn-Sherris mortality model shows the poorest performance. The CIR mortality model has reasonable RMSE but is outperformed by the AFNS mortalty models.

Table 4: RMSE by Comparing the Actual and Best-Estimate Survival Probabilities of the 1916 Cohort

	The Blackburn-Sherris Model		The AFNS Model		The CIR Model
	Independent	Dependent	Independent	Dependent	
RMSE	0.03197	0.00726	0.00668	0.00754	0.01835

To understand these differences, Figure 9 shows the survival probabilities for the different mortality models using the best estimate forecasts compared to the actual survival probabilities from the historical mortality data. Most models produce reasonable survival curve fits except the independent Blackburn-Sherris mortality model and the CIR mortality model that both underestimate the survival rates of the 1916 cohort.

Figure 10 shows the absolute percentage errors across ages for all the mortality models confirming the superior forecasting performance of the independent AFNS mortality model. The dependent Blackburn-Sherris mortality models performs better than the independent Blackburn-Sherris mortality model, showing the benefit of including correlations between the factors in this model. The level, slope and curvature structure of the factors in the AFNS mortality model capture the impact of dependence in the factors in the Blackburn-Sherris mortality model.

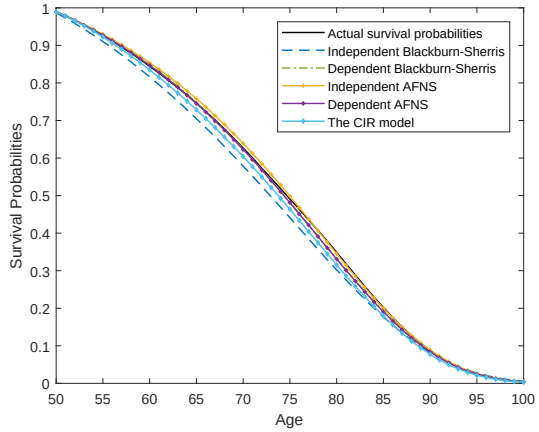


Figure 9: Actual and Best-Estimate Survival Probabilities of the 1916 Cohort

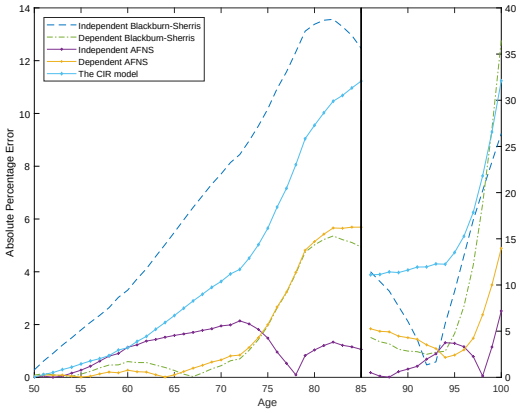


Figure 10: Absolute Percentage Errors between Actual and Best-Estimate Survival Probabilities

6 Conclusion

This paper introduces continuous time affine mortality models applied to cohort survival curve data. We introduce an AFNS mortality model with interpretable latent stochastic factors for level, slope and curvature of the survival curve. We outline, compare and assess a number of independent-factor and dependent-factor affine mortality models with Gaussian processes including the Blackburn-Sherris mortality model (Blackburn and Sherris, 2013; Christensen et al., 2011) as well as an affine mortality model with square-root processes (the CIR mortality model). The CIR mortality model precludes negative mortality rates that can occur in the Gaussian models. The CIR latent factors and the mortality intensity have non-central Chi-square distributions which can capture mortality heterogeneity.

Affine mortality models produce survival rates consistent with the dynamics of the latent

stochastic factors. The structure is similar to interest rate models providing a closed-form solution for survival probabilities and are suitable for financial and insurance applications involving longevity risk management. We use US historical age-cohort data to fit and assess the mortality models.

Although incorporating dependence in the Blackburn-Sherris mortality model improves in-sample model fit and out-of-sample forecasting performance, we find that that independent-factor AFNS mortality model performs well. It can better capture the variation in cohort mortality rates in US data and produces a better fit at older ages than the independent-factor Blackburn-Sherris model. For the 1916 cohort the independent-factor AFNS mortality model has better predictive performance compared to the other models. Negative mortality rates have very low probability in the AFNS mortality models.

The CIR mortality model has the best in-sample model fit including model residuals. The superior in-sample performance of the CIR mortality model reflects the more realistic assumption of Gamma-distributed mortality rates. The benefits of the independent-factor AFNS in modelling cohort mortality survival curves lead us to favour this model over the CIR model.

Based on our assessment of affine mortality models, the independent AFNS model provides satisfactory model fit and satisfactory predictive performance. The model is parsimonious and can be readily estimated using the Kalman filter, allowing for Poisson mortality variation in the measurement equation. The model allows for intuitive factor interpretation in terms of the dynamics of the mortality survival curve and is well suited for financial and insurance applications including pricing and longevity risk management.

7 Acknowledgement

The authors acknowledge financial support from the Society of Actuaries Center of Actuarial Excellence Research Grant 2017–2020: Longevity Risk: Actuarial and Predictive Models, Retirement Product Innovation, and Risk Management Strategies as well as support from CEPAR Australian Research Council Centre of Excellence in Population Ageing Research project number CE170100005.

Appendices

A Solutions to the Ordinary Differential Equations

We have

$$\frac{d}{dt}\left[e^{\left(K^Q\right)'(T-t)}B\left(t,T\right)\right]=e^{\left(K^Q\right)'(T-t)}\frac{dB\left(t,T\right)}{dt}-\left(K^Q\right)'e^{\left(K^Q\right)'(T-t)}B\left(t,T\right), \tag{49}$$

and substituting equation (7), we simplify to obtain

$$\int_t^T\frac{d}{ds}\left[e^{\left(K^Q\right)'(T-s)}B\left(s,T\right)\right]ds=\int_t^T e^{\left(K^Q\right)'(T-s)}\rho_1 ds, \tag{50}$$

with has the solution, after including the boundary conditions,

$$B\left(t,T\right)=-e^{\left(-K^Q\right)'(T-t)}\int_t^T e^{\left(K^Q\right)'(T-s)}\rho_1 ds. \tag{51}$$

With K^Q in equation (18) and $\rho_1=(1,1,1)^T$ in the Blackburn-Sherris model,

$$e^{\left(-K^Q\right)'(T-s)}\rho_1=\begin{pmatrix}a_{11}&a_{21}&a_{31}\\0&a_{22}&a_{32}\\0&0&a_{33}\end{pmatrix}\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}a_{11}+a_{21}+a_{31}\\a_{22}+a_{32}\\a_{33}\end{pmatrix} \tag{52}$$

where

$$\begin{aligned}
&a_{11}=e^{\delta_{11}(T-s)},\quad a_{21}=D_1\left(e^{\delta_{11}(T-s)}-e^{\delta_{22}(T-s)}\right),\\
&a_{31}=\left(D_4+D_1D_5\right)e^{\delta_{11}(T-s)}-D_1D_2e^{\delta_{22}(T-s)}+\left(D_2D_3-D_4\right)e^{\delta_{33}(T-s)},\\
&a_{22}=e^{\delta_{22}(T-s)},\quad a_{32}=D_2\left(e^{\delta_{22}(T-s)}-e^{\delta_{33}(T-s)}\right),\quad a_{33}=e^{\delta_{33}(T-s)},
\end{aligned} \tag{53}$$

and

$$D_1=\frac{\delta_{21}}{\delta_{11}-\delta_{22}},\quad D_2=\frac{\delta_{32}}{\delta_{22}-\delta_{33}},\quad D_3=\frac{\delta_{21}}{\delta_{11}-\delta_{33}},\quad D_4=\frac{\delta_{31}}{\delta_{11}-\delta_{33}},\quad D_5=\frac{\delta_{32}}{\delta_{11}-\delta_{33}}.$$

Integrating each element in Equation (52),

$$\begin{aligned}
b_1&=\int_t^T\left(a_{11}+a_{21}+a_{31}\right)ds\\
&=\left(1+D_1+D_1D_5+D_4\right)\frac{1-e^{-\delta_{11}(T-t)}}{-\delta_{11}}-D_1\left(1+D_2\right)\frac{1-e^{-\delta_{22}(T-t)}}{-\delta_{22}}\\
&\quad+\left(D_2D_3-D_4\right)\frac{1-e^{-\delta_{33}(T-t)}}{-\delta_{33}},\\
b_2&=\int_t^T\left(a_{22}+a_{32}\right)ds=\left(1+D_2\right)\frac{1-e^{-\delta_{22}(T-t)}}{-\delta_{22}}-D_2\frac{1-e^{-\delta_{33}(T-t)}}{-\delta_{33}},\\
b_3&=\int_t^T a_{33}ds=\frac{1-e^{-\delta_{33}(T-t)}}{-\delta_{33}}.
\end{aligned}$$

Let

$$e^{\left(-K^{Q}\right)'(T-t)} = \begin{pmatrix} c_{11} & c_{21} & c_{31} \\ 0 & c_{22} & c_{32} \\ 0 & 0 & c_{33} \end{pmatrix}, \tag{54}$$

where

$$\begin{aligned} c_{11} &= e^{-\delta_{11}(T-t)}, \quad c_{21} = D_1\left(e^{-\delta_{11}(T-t)} - e^{-\delta_{22}(T-t)}\right), \\ c_{31} &= \left(D_4 + D_1 D_5\right) e^{-\delta_{11}(T-t)} - D_1 D_2 e^{-\delta_{22}(T-t)} + \left(D_2 D_3 - D_4\right) e^{-\delta_{33}(T-t)}, \\ c_{22} &= e^{-\delta_{22}(T-t)}, \quad c_{32} = D_2\left(e^{-\delta_{22}(T-t)} - e^{-\delta_{33}(T-t)}\right), \quad c_{33} = e^{-\delta_{33}(T-t)}. \end{aligned} \tag{55}$$

Equation (51) now can be written as:

$$\begin{pmatrix} B^1(t,T) \\ B^2(t,T) \\ B^3(t,T) \end{pmatrix} = -\begin{pmatrix} c_{11} & c_{21} & c_{31} \\ 0 & c_{22} & c_{32} \\ 0 & 0 & c_{33} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}. \tag{56}$$

Therefore, the solutions of $B(t,T)$ are

$$\begin{aligned} B^1(t,T) &= -E_1 \frac{1-e^{-\delta_{11}(T-t)}}{\delta_{11}} + E_2 \frac{1-e^{-\delta_{22}(T-t)}}{\delta_{22}} - E_3 \frac{1-e^{-\delta_{33}(T-t)}}{\delta_{33}}, \\ B^2(t,T) &= -\left(1+D_2\right) \frac{1-e^{-\delta_{22}(T-t)}}{\delta_{22}} + D_2 \frac{1-e^{-\delta_{33}(T-t)}}{\delta_{33}}, \\ B^3(t,T) &= -\frac{1-e^{-\delta_{33}(T-t)}}{\delta_{33}}, \end{aligned} \tag{57}$$

where $E_1 = 1 + D_1 + D_1 D_5 + D_4, \quad E_2 = D_1\left(1+D_2\right), \quad E_3 = D_2 D_3 - D_4$.

From Equation (8) and the boundary condition,

$$\begin{aligned} A(t,T) &= \frac{1}{2}\int_t^T \sum_{j=1}^{3} \left(\Sigma' B(s,T) B(s,T)' \Sigma\right)_{j,j} ds \\ &= \frac{1}{2}\int_t^T \Big[\sigma_{11}^2 B^1(s,T)^2 + \left(\sigma_{21}^2 + \sigma_{22}^2\right) B^2(s,T)^2 + \left(\sigma_{31}^2 + \sigma_{32}^2 + \sigma_{33}^2\right) B^3(s,T)^2 \\ &\quad + 2\sigma_{11}\sigma_{21} B^1(s,T) B^2(s,T) + 2\sigma_{11}\sigma_{31} B^1(s,T) B^3(s,T) \\ &\quad + 2\left(\sigma_{21}\sigma_{31} + \sigma_{22}\sigma_{32}\right) B^2(s,T) B^3(s,T)\Big] ds. \end{aligned}$$

Terms with $B(s,T)$ are expanded:

$$\begin{aligned} B^1(s,T)^2 &= \frac{E_1^2}{\delta_{11}^2}\left(1-e^{-\delta_{11}(T-s)}\right)^2 + \frac{E_2^2}{\delta_{22}^2}\left(1-e^{-\delta_{22}(T-s)}\right)^2 + \frac{E_3^2}{\delta_{33}^2}\left(1-e^{-\delta_{33}(T-s)}\right)^2 \\ &\quad - \frac{2}{\delta_{11}\delta_{22}} E_1 E_2 \left(1-e^{-\delta_{11}(T-s)}\right)\left(1-e^{-\delta_{22}(T-s)}\right) \\ &\quad + \frac{2}{\delta_{11}\delta_{33}} E_1 E_3 \left(1-e^{-\delta_{11}(T-s)}\right)\left(1-e^{-\delta_{33}(T-s)}\right) \\ &\quad - \frac{2}{\delta_{22}\delta_{33}} E_2 E_3 \left(1-e^{-\delta_{22}(T-s)}\right)\left(1-e^{-\delta_{33}(T-s)}\right), \end{aligned}$$

$$
\begin{aligned}
B^2(s,T)^2 &= \frac{(1+D_2)^2}{\delta_{22}^2}\left(1-e^{-\delta_{22}(T-s)}\right)^2 + \frac{D_2^2}{\delta_{33}^2}\left(1-e^{-\delta_{33}(T-s)}\right)^2 \\
&\quad - \frac{2}{\delta_{22}\delta_{33}} D_2(1+D_2)\left(1-e^{-\delta_{22}(T-s)}\right)\left(1-e^{-\delta_{33}(T-s)}\right), \\
B^3(s,T)^2 &= \frac{1}{\delta_{33}^2}\left(1-e^{-\delta_{33}(T-s)}\right)^2, \\
B^1(s,T)B^2(s,T) &= \frac{E_1(1+D_2)}{\delta_{11}\delta_{22}}\left(1-e^{-\delta_{11}(T-s)}\right)\left(1-e^{-\delta_{22}(T-s)}\right) \\
&\quad - \frac{E_1 D_2}{\delta_{11}\delta_{33}}\left(1-e^{-\delta_{11}(T-s)}\right)\left(1-e^{-\delta_{33}(T-s)}\right) \\
&\quad + \frac{1}{\delta_{22}\delta_{33}}\left(E_2 D_2 + E_3(1+D_2)\right)\left(1-e^{-\delta_{22}(T-s)}\right)\left(1-e^{-\delta_{33}(T-s)}\right) \\
&\quad - \frac{E_2(1+D_2)}{\delta_{22}^2}\left(1-e^{-\delta_{22}(T-s)}\right)^2 - \frac{E_3 D_2}{\delta_{33}^2}\left(1-e^{-\delta_{33}(T-s)}\right)^2, \\
B^1(s,T)B^3(s,T) &= \frac{E_1}{\delta_{11}\delta_{33}}\left(1-e^{-\delta_{11}(T-s)}\right)\left(1-e^{-\delta_{33}(T-s)}\right) + \frac{E_3}{\delta_{33}^2}\left(1-e^{-\delta_{33}(T-s)}\right)^2 \\
&\quad - \frac{E_2}{\delta_{22}\delta_{33}}\left(1-e^{-\delta_{22}(T-s)}\right)\left(1-e^{-\delta_{33}(T-s)}\right), \\
B^2(s,T)B^3(s,T) &= \frac{1+D_2}{\delta_{22}\delta_{33}}\left(1-e^{-\delta_{22}(T-s)}\right)\left(1-e^{-\delta_{33}(T-s)}\right) - \frac{D_2}{\delta_{33}^2}\left(1-e^{-\delta_{33}(T-s)}\right)^2.
\end{aligned}
$$

Collecting terms with $\left(1-e^{-\delta_{jj}(T-s)}\right)^2$ and $\left(1-e^{-\delta_{ii}(T-s)}\right)\left(1-e^{-\delta_{jj}(T-s)}\right)$ $(i,j=1,2,3$ and $i\neq j)$ and integrating,

$$
\begin{aligned}
A(t,T) = \frac{1}{2}\Bigg[& \frac{F_1}{\delta_{11}^3}\left(\frac{1}{2}\left(1-e^{-2\delta_{11}(T-t)}\right) - 2\left(1-e^{-\delta_{11}(T-t)}\right) + \delta_{11}(T-t)\right) \\
& + \frac{F_2}{\delta_{22}^3}\left(\frac{1}{2}\left(1-e^{-2\delta_{22}(T-t)}\right) - 2\left(1-e^{-\delta_{22}(T-t)}\right) + \delta_{22}(T-t)\right) \\
& + \frac{F_3}{\delta_{33}^3}\left(\frac{1}{2}\left(1-e^{-2\delta_{33}(T-t)}\right) - 2\left(1-e^{-\delta_{33}(T-t)}\right) + \delta_{33}(T-t)\right) \\
& + \frac{F_4}{\delta_{11}\delta_{22}}\left((T-t) - \frac{1-e^{-\delta_{11}(T-t)}}{\delta_{11}} - \frac{1-e^{-\delta_{22}(T-t)}}{\delta_{22}} - \frac{1-e^{-(\delta_{11}+\delta_{22})(T-t)}}{\delta_{11}+\delta_{22}}\right) \\
& + \frac{F_5}{\delta_{11}\delta_{33}}\left((T-t) - \frac{1-e^{-\delta_{11}(T-t)}}{\delta_{11}} - \frac{1-e^{-\delta_{33}(T-t)}}{\delta_{33}} - \frac{1-e^{-(\delta_{11}+\delta_{33})(T-t)}}{\delta_{11}+\delta_{33}}\right) \\
& + \frac{F_6}{\delta_{22}\delta_{33}}\left((T-t) - \frac{1-e^{-\delta_{22}(T-t)}}{\delta_{22}} - \frac{1-e^{-\delta_{33}(T-t)}}{\delta_{33}} - \frac{1-e^{-(\delta_{22}+\delta_{33})(T-t)}}{\delta_{22}+\delta_{33}}\right)\Bigg],
\end{aligned}
\tag{58}
$$

where

$$
\begin{aligned}
F_1 &= \sigma_{11}^2 E_1^2, \\
F_2 &= \sigma_{11}^2 E_2^2 - 2\sigma_{11}\sigma_{21} E_2 (1+D_2) + \left(\sigma_{21}^2 + \sigma_{22}^2\right)(1+D_2)^2, \\
F_3 &= \sigma_{11}^2 E_3^2 - 2\sigma_{11}\sigma_{21} E_3 D_2 + \left(\sigma_{21}^2 + \sigma_{22}^2\right) D_2^2 - 2\left(\sigma_{21}\sigma_{31} + \sigma_{22}\sigma_{32}\right) D_2 + 2\sigma_{11}\sigma_{31} E_3 \\
&\quad + \left(\sigma_{31}^2 + \sigma_{32}^2 + \sigma_{33}^2\right), \\
F_4 &= -2\sigma_{11}^2 E_1 E_2 + 2\sigma_{11}\sigma_{21} E_1 (1+D_2), \\
F_5 &= 2\sigma_{11}^2 E_1 E_3 - 2\sigma_{11}\sigma_{21} E_1 D_2 + 2\sigma_{11}\sigma_{31} E_1, \\
F_6 &= -2\sigma_{11}^2 E_2 E_3 + 2\sigma_{11}\sigma_{21}\left[E_3(1+D_2) + E_2 D_2\right] - 2\sigma_{11}\sigma_{31} E_2 + 2\left(\sigma_{21}\sigma_{31} + \sigma_{22}\sigma_{32}\right)(1+D_2) \\
&\quad - 2\left(\sigma_{21}^2 + \sigma_{22}^2\right) D_2 (1+D_2).
\end{aligned}
$$

B Real World Dynamics and Change of Measure

B.1 Models with Gaussian Processes

The market price of risk has the following form:

$$\Lambda_t = \lambda^0 + \lambda^1 X_t, \tag{59}$$

where $\Lambda_t \in \mathbb{R}^{n\times 1}$, $\lambda^0 \in \mathbb{R}^{n\times 1}$ and $\lambda^1 \in \mathbb{R}^{n\times n}$.

With the above specification, the SDEs of the state variables X_t under the real-world measure P are derived as following:

$$
\begin{aligned}
dX_t &= K^Q\left[\theta^Q - X_t\right] dt + \Sigma\left[dW_t^P + \Lambda_t dt\right] \\
&= K^Q\left[\theta^Q - X_t\right] dt + \Sigma\left[\lambda^0 dt + \lambda^1 X_t dt + dW_t^P dt\right] \\
&= \left[K^Q\theta^Q + \Sigma\lambda^0\right] dt - \left[K^Q - \Sigma\lambda^1\right] X_t dt + \Sigma dW_t^P \\
&= \left(K^Q - \Sigma\lambda^1\right)\left[\frac{K^Q\theta^Q + \Sigma\lambda^0}{K^Q - \Sigma\lambda^1} - X_t\right] dt + \Sigma dW_t^P \\
&= K^P\left[\theta^P - X_t\right] dt + \Sigma dW_t^P,
\end{aligned}
\tag{60}
$$

where

$$K^P = K^Q - \Sigma\lambda^1, \quad \theta^P = \frac{K^Q\theta^Q + \Sigma\lambda^0}{K^Q}. \tag{61}$$

B.2 The CIR Model

Following the essentially affine model structure in Duffee (2002), the market price of longevity risk for the multi-factor CIR model is specified as:

$$\Lambda_t = D(X_t, t)\,\lambda^0 = \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{X_t^2} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} \lambda_1^0 \\ \lambda_2^0 \\ \lambda_3^0 \end{pmatrix}, \tag{62}$$

where $\Lambda_t \in \mathbb{R}^{3\times 1}$ represents risk premium and $\lambda^0 \in \mathbb{R}^{3\times 1}$, and let

$$\begin{aligned} D^2\left(X_t,t\right)\lambda^0 &= \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{X_t^2} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{X_t^2} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} \lambda_1^0 \\ \lambda_2^0 \\ \lambda_3^0 \end{pmatrix} \\ &= \begin{pmatrix} X_t^1 & 0 & 0 \\ 0 & X_t^2 & 0 \\ 0 & 0 & X_t^3 \end{pmatrix} \begin{pmatrix} \lambda_1^0 \\ \lambda_2^0 \\ \lambda_3^0 \end{pmatrix} = \begin{pmatrix} \lambda_1^0 X_t^1 \\ \lambda_2^0 X_t^2 \\ \lambda_3^0 X_t^3 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1^0 & 0 & 0 \\ 0 & \lambda_2^0 & 0 \\ 0 & 0 & \lambda_3^0 \end{pmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} = \Lambda_0 X_t. \end{aligned} \tag{63}$$

The SDEs of the state variables X_t under the real-world measure P are derived as following

$$\begin{aligned} dX_t &= K^Q\left[\theta^Q - X_t\right]dt + \Sigma D\left(X_t,t\right)\left[dW_t^P + \Lambda_t dt\right] \\ &= K^Q\left[\theta^Q - X_t\right]dt + \Sigma D\left(X_t,t\right)\left[D\left(X_t,t\right)\lambda^0 dt + dW_t^P dt\right] \\ &= \left[K^Q\theta^Q - K^Q X_t + \Sigma D^2\left(X_t,t\right)\lambda^0\right]dt + \Sigma D\left(X_t,t\right)dW_t^P \\ &= \left[K^Q\theta^Q - \left(K^Q - \Sigma\Lambda_0\right)X_t\right]dt + \Sigma D\left(X_t,t\right)dW_t^P \\ &= \left(K^Q - \Sigma\Lambda_0\right)\left[\frac{K^Q\theta^Q}{K^Q - \Sigma\Lambda_0} - X_t\right]dt + \Sigma D\left(X_t,t\right)dW_t^P \\ &= K^P\left[\theta^P - X_t\right]dt + \Sigma D\left(X_t,t\right)dW_t^P, \end{aligned} \tag{64}$$

where

$$K^P = K^Q - \Sigma\Lambda_0, \quad \theta^P = \frac{K^Q\theta^Q}{K^Q - \Sigma\Lambda_0}. \tag{65}$$

References

Alai, D., Ignatieva, K., and Sherris, M. (2019). The investigation of a forward-rate mortality framework. *Risks*, 7(2).

Barrieu, P., Bensusan, H., Karoui, N. E., Hillairet, C., Loisel, S., Ravanelli, C., and Salhi, Y. (2012). Understanding, modelling and managing longevity risk: key issues and main challenges. *Scandinavian Actuarial Journal*, 2012(3):203–231.

Bauer, D., Börger, M., Ruß, J., and Zwiesler, H.-J. (2008). The volatility of mortality. *Asia-Pacific Journal of Risk and Insurance*, 3(1).

Biffis, E. (2005). Affine processes for dynamic mortality and actuarial valuations. *Insurance: Mathematics and Economics*, 37(3):443–468.

Björk, T. and Christensen, B. (1999). Interest rate dynamics and consistent forward rate curves. *Mathematical Finance*, 9(4):323–348.

Blackburn, C. and Sherris, M. (2013). Consistent dynamic affine mortality models for longevity risk applications. *Insurance: Mathematics and Economics*, 53:64–73.

Blake, D., Boardman, T., and Cairns, A. (2014). Sharing longevity risk: Why governments should issue longevity bonds. *North American Actuarial Journal*, 18(1):258–277.

Blake, D. and Burrows, W. (2001). Survivor bonds: Helping to hedge mortality risk. *Journal of Risk and Insurance*, 68(2):339–348.

Blake, D., Karoui, N. E., Loisel, S., and MacMinn, R. (2018). Longevity risk and capital markets: The 2015-16 update. *Insurance: Mathematics and Economics*, 78:157 – 173.

Cairns, A. J., Blake, D., and Dowd, K. (2006a). Pricing death: Frameworks for the valuation and securitization of mortality risk. *ASTIN Bulletin*, 36(1):79 – 120.

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

Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A., and Balevich, I. (2009). A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *North American Actuarial Journal*, 13(1):1–35.

Chang, Y. and Sherris, M. (2018). Longevity Risk Management and the Development of a Value-Based Longevity Index. *Risks*, 6(1):1–20.

Chen, R.-R. and Scott, L. (2003). Multi-factor Cox-Ingersoll-Ross models of the term structure: Estimates and tests from a Kalman filter model. *The Journal of Real Estate Finance and Economics*, 27(2):143–172.

Christensen, J. H., Diebold, F. X., and Rudebusch, G. D. (2011). The affine arbitrage-free class of Nelson–Siegel term structure models. *Journal of Econometrics*, 164(1):4–20.

Continuous Mortality Investigation (2018). The CMI mortality projections model, CMI2017. Working paper, the Institute and Faculty of Actuaries.

Coughlan, G., Epstein, D., Sinha, A., and Honig, P. (2007). q-forwards: Derivatives for transferring longevity and mortality risks. *JPMorgan Pension Advisory Group, London, July*, 2.

Cox, J. C., Ingersoll, J. E., and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407.

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

Dahl, M. and Møller, T. (2006). Valuation and hedging of life insurance liabilities with systematic mortality risk. *Insurance: mathematics and economics*, 39(2):193–217.

Dai, Q. and Singleton, K. J. (2000). Specification analysis of affine term structure models. *The Journal of Finance*, 55(5):1943–1978.

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(2):277 – 308.

Diebold, F. X. and Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics*, 130(2):337–364.

Diebold, F. X. and Rudebusch, G. D. (2013). *Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach*. Princeton University Press.

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

Duan, J.-C. and Simonato, J.-G. (1999). Estimating and testing exponential-affine term structure models by Kalman filter. *Review of Quantitative Finance and Accounting*, 13(2):111–135.

Duffee, G. R. (2002). Term premia and interest rate forecasts in affine models. *The Journal of Finance*, 57(1):405–443.

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

Durbin, J. and Koopman, S. J. (2012). *Time series analysis by state space methods*, volume 38. Oxford University Press.

Gallop, A. (2008). Mortality projections in the United Kingdom. In *Society of Actuaries Living to 100 Symposium*. Citeseer.

Geyer, A. L. and Pichler, S. (1999). A state-space approach to estimate and test multifactor Cox-Ingersoll-Ross models of the term structure. *Journal of Financial Research*, 22(1):107–130.

Human Mortality Database (2017). U.S.A. Life Tables. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at `http://www.mortality.org/cgi-bin/hmd/country.php?cntr=USA&level=1`, data downloaded on March 2018 and October 2018.

Jevtic, P., Luciano, E., and Vigna, E. (2013). Mortality surface by means of continuous time cohort models. *Insurance: Mathematics and Economics*, 53:122–133.

Kalman, R. (1960). A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45.

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

Life and Longevity Markets Association (2010). Longevity pricing framework: framework for pricing longevity exposures developed by the LLMA (life and longevity markets association). Technical report. Available at `https://llmablog.files.wordpress.com/2017/08/longevity_pricing_framework_final.pdf`.

Luciano, E., Spreeuw, J., and Vigna, E. (2008). Modelling stochastic mortality for dependent lives. *Insurance: Mathematics and Economics*, 43(2):234–244.

Milevsky, M. A. and Promislow, S. D. (2001). Mortality derivatives and the option to annuitise. *Insurance: Mathematics and Economics*, 29(3):299–318.

Pitacco, E. (2016). High age mortality and frailty. some remarks and hints for actuarial modeling. Available at `http://www.cepar.edu.au/working-papers/working-papers-2016.aspx`.

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

Schrager, D. F. (2006). Affine stochastic mortality. *Insurance: Mathematics and Economics*, 38(1):81–97.

Shumway, R. H. and Stoffer, D. S. (2017). *Time series analysis and its applications*. Springer.

The Joint Forum (2013). Longevity risk transfer markets: Market structure, growth drivers and impediments, and potential risks. Technical report. Available at `https://www.bis.org/publ/joint34.htm`.

Willets, R. C. (2004). The cohort effect: Insights and explanations. *British Actuarial Journal*, 10(4):833877.

Xu, Y., Sherris, M., and Ziveyi, J. (2015). The application of affine processes in multi-cohort mortality models. Working paper, UNSW Business School Research Paper.

Xu, Y., Sherris, M., and Ziveyi, J. (2018). Market price of longevity risk for a multi-cohort mortality model with application to longevity bond option pricing. *forthcoming Journal of Risk and Insurance*.