

A State-Space Estimation of the Lee-Carter Mortality Model and Implications for Annuity Pricing

M. C. Fung^a, G. W. Peters^b and P. V. Shevchenko^c

^a*Risk Analytics Group, Digital Productivity Flagship, CSIRO, Australia*

^b*Department of Statistical Science, University College London, United Kingdom*

Associate Fellow of Oxford Mann Institute, Oxford University

Associate Fellow of Systemic Risk Center, London School of Economics.

^c*Risk Analytics Group, Digital Productivity Flagship, CSIRO, Australia*

3 August 2015

Email: Simon.Fung@csiro.au

Abstract: A common feature of retirement income products is that their payouts depend on the lifetime of policyholders. A typical example is a life annuity policy which promises to provide benefits regularly as long as the retiree is alive. Consequently, insurers have to rely on “best estimate” life tables, which consist of age-specific mortality rates, in order to price these kind of products properly. Recently there is a growing concern about the accuracy of the estimation of mortality rates since it has been historically observed that life expectancy is often underestimated in the past (so-called longevity risk), thus resulting in longer benefit payments than insurers have originally anticipated. To take into account the stochastic nature of the evolution of mortality rates, Lee and Carter (1992) proposed a stochastic mortality model which primarily aims to forecast age-specific mortality rates more accurately.

The original approach to estimating the Lee-Carter model is via a singular value decomposition, which falls into the least squares framework. Researchers then point out that the Lee-Carter model can be treated as a state-space model. As a result several well-established state-space modeling techniques can be applied to not just perform estimation of the model, but to also perform forecasting as well as smoothing. Research in this area is still not yet fully explored in the actuarial literature, however. Existing relevant literature focuses mainly on mortality forecasting or pricing of longevity derivatives, while the full implications and methods of using the state-space representation of the Lee-Carter model in pricing retirement income products is yet to be examined.

The main contribution of this article is twofold. First, we provide a rigorous and detailed derivation of the posterior distributions of the parameters and the latent process of the Lee-Carter model via Gibbs sampling. Our assumption for priors is slightly more general than the current literature in this area. Moreover, we suggest a new form of identification constraint not yet utilised in the actuarial literature that proves to be a more convenient approach for estimating the model under the state-space framework. Second, by exploiting the posterior distribution of the latent process and parameters, we examine the pricing range of annuities, taking into account the stochastic nature of the dynamics of the mortality rates. In this way we aim to capture the impact of longevity risk on the pricing of annuities.

The outcome of our study demonstrates that an annuity price can be more than 4% under-valued when different assumptions are made on determining the survival curve constructed from the distribution of the forecasted mortality rates. Given that a typical annuity portfolio consists of a large number of policies with maturities which span decades, we conclude that the impact of longevity risk on the accurate pricing of annuities is a significant issue to be further researched. In addition, we find that mis-pricing is increasingly more pronounced for older ages as well as for annuity policies having a longer maturity.

Keywords: *Mortality modeling, longevity risk, Bayesian inference, Gibbs sampling, state-space models*

1 INTRODUCTION

The pricing of retirement income products depends crucially on the accuracy of the predicted death or survival probabilities. It is now widely documented that survival probability is consistently underestimated especially in the last few decades. To capture the stochastic nature of mortality trends, Lee and Carter (1992) proposed a stochastic mortality model to forecast the trend of age-specific mortality rates.

There exists a body of literature on how to estimate the Lee-Carter model. The original approach in Lee and Carter (1992) is via singular value decomposition. To overcome the unrealistic feature of homogeneity in the additive error term, Brouhns et al. (2002) recast the model as a Poisson regression model assuming Poisson random variation for the number of deaths. Estimation of the model in the Poisson regression setting under the Bayesian framework is carried out in Czado et al. (2005). Also there is a recently developed framework for modeling death counts with common risk factors via credit risk plus methodology and resultant estimation of the model via Monte Carlo Markov Chain in Hirz et al. (2015). In this paper we focus principally on the class of what has become known as the Lee-Carter models, in this regard another approach to estimating the Lee-Carter model is via state-space representation. Pedroza (2006) shows that the predictive intervals for forecasting are materially wider than using the singular value decomposition method. Kogure and Kurachi (2010) adopt the state-space modeling approach and apply it for the pricing of longevity bonds and swaps.

In this paper we aim to explore further the Bayesian state-space modeling approach and examine its implication for annuity pricing. Specifically, we provide a rigorous and detailed derivation of the posterior distributions of the static parameters and the latent process of the Lee-Carter model via Gibbs sampling. Our assumptions on the priors on the Lee-Carter model parameters are more general than Pedroza (2006) and Kogure and Kurachi (2010). Moreover, a new form of identification constraint not yet recognised in the actuarial literature is proposed which proves to be more convenient for estimating the model using an MCMC method under the state-space formulation. Using the predictive distributions of age-specific death rates, we examine the impact of longevity risk on the pricing of annuities and demonstrate that this long-term risk is indeed a significant factor when accurate pricing is required.

In Section 2 the state-space Lee-Carter model is presented together with some definitions and notation. Section 3 describes the Gibbs sampling approach to estimate the state-space Lee-Carter model. Posterior distributions of the static parameters and the latent process are derived in detail. Section 4 examines the impact of longevity risk on annuity pricing. Section 5 concludes with some remarks.

2 LEE-CARTER MODEL

2.1 Definitions and Notation

In this section we briefly recall some important definitions from actuarial literature on mortality modelling that are required to set up the Lee-Carter model below and the pricing analysis in Section 4. Let T_x be a random variable representing the remaining lifetime of a person aged x . The cumulative distribution function and survival function of T_x are written as ${}_τq_x = P(T_x \leq \tau)$ and ${}_τp_x = P(T_x > \tau)$ respectively. For a person aged x , the force of mortality at age $x + \tau$ is defined as $\mu_{x+\tau} := \lim_{h \rightarrow 0} \frac{1}{h} P(T_x < \tau + h | T_x > \tau) = \frac{1}{\tau p_x} \frac{d}{d\tau} \tau q_x = -\frac{d}{d\tau} \ln \tau p_x$ and hence ${}_τp_x = \exp(-\int_0^\tau \mu_{x+s} ds)$. The central death rate for a x -year-old, where $x \in \mathbb{N}$, is defined as $m_x := \frac{{}_1q_x}{\int_0^1 {}_s p_x ds} = \frac{\int_0^1 {}_s p_x \mu_{x+s} ds}{\int_0^1 {}_s p_x ds}$ which is a weighted-average of the force of mortality.

Under the so-called piecewise constant force of mortality assumption, that is $\mu_{x+s} = \mu_x$ where $0 \leq s < 1$ and $x \in \mathbb{N}$, we have $m_x = \mu_x$ and hence ${}_1p_x = e^{-m_x}$. Moreover, the maximum likelihood estimate of the force of mortality $\hat{\mu}_x$ (and hence \hat{m}_x) is given by $\hat{\mu}_x = D_x / E_x = \hat{m}_x$ where D_x is the number of deaths recorded at age x last birthday and the exposure-to-risk E_x is the total time lived by people aged x last birthday, during the observation year (Pitacco et al. (2009)). Note that E_x is often approximated by an estimate of the population aged x last birthday in the middle of the observation year.

2.2 The Lee-Carter State Space Model

Based on the definitions described above, we now discuss the work of Lee and Carter (1992) who proposed a stochastic mortality model specifically for forecasting age-specific central death rates m_{xt} , where $x = x_1, \dots, x_p$ and $t = 1, \dots, n$ represent age and year (time) respectively. The model assumes that the log central death rate, $y_{xt} = \ln m_{xt}$, is governed by the following equation

$$y_t = \alpha + \beta \kappa_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2 \mathbf{1}_p) \tag{1}$$

where $\mathbf{y}_t = (y_{x_1t}, \dots, y_{x_pt})'$, $\boldsymbol{\alpha} = (\alpha_{x_1}, \dots, \alpha_{x_p})'$, $\boldsymbol{\beta} = (\beta_{x_1}, \dots, \beta_{x_p})'$, $\boldsymbol{\varepsilon}_t = (\varepsilon_{x_1t}, \dots, \varepsilon_{x_pt})'$, $\mathbf{1}_p$ is the p by p identity matrix and $N(\cdot, \cdot)$ denotes the Gaussian distribution. Lee and Carter (1992) estimate the model (1) via singular value decomposition and subsequently assume that the unobserved latent time trend denoted by κ_t satisfies the following linear dynamics

$$\kappa_t = \kappa_{t-1} + \theta + \omega_t, \quad \omega_t \sim N(0, \sigma_\omega^2) \tag{2}$$

where $\boldsymbol{\varepsilon}_t$ and ω_t are independent. The parameters θ, σ_ω^2 are then estimated using standard econometric techniques. In this form the Lee-Carter model is, however, not identifiable since the model (1) is invariant up to some linear transformations of the parameters: $\mathbf{y}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}\kappa_t + \boldsymbol{\varepsilon}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}c + \frac{\boldsymbol{\beta}}{d}((\kappa_t - c)d) + \boldsymbol{\varepsilon}_t = \tilde{\boldsymbol{\alpha}} + \tilde{\boldsymbol{\beta}}\tilde{\kappa}_t + \boldsymbol{\varepsilon}_t$ where $\tilde{\boldsymbol{\alpha}} = \boldsymbol{\alpha} + \boldsymbol{\beta}c$, $\tilde{\boldsymbol{\beta}} = \frac{\boldsymbol{\beta}}{d}$ and $\tilde{\kappa}_t = (\kappa_t - c)d$. To overcome this identification issue, Lee and Carter (1992) introduced the constraints $\sum_{x=x_1}^{x_p} \beta_x = 1$ and $\sum_{t=1}^n \kappa_t = 0$ to ensure that the model becomes identifiable since, by setting $d = \sum_{x=x_1}^{x_p} \beta_x$ and $c = \sum_{t=1}^n \kappa_t$, we have $\sum_{x=x_1}^{x_p} \tilde{\beta}_x = 1$ and $\sum_{t=1}^n \tilde{\kappa}_t = 0$. Pedroza (2006) suggests that we can in fact combine the processes \mathbf{y}_t and κ_t into one dynamical system

$$\mathbf{y}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}\kappa_t + \boldsymbol{\varepsilon}_t, \quad \kappa_t = \kappa_{t-1} + \theta + \omega_t, \quad \text{where } \boldsymbol{\varepsilon}_t \sim N(0, \sigma_\varepsilon^2 \mathbf{1}_p) \text{ and } \omega_t \sim N(0, \sigma_\omega^2), \tag{3}$$

resulting in a state-space representation of the Lee-Carter model and estimate κ_t and model parameters jointly.

2.3 Lee-Carter model in ARIMA Time Series Form

We also note that, at least when one doesn't consider the identification constraints, the Lee-Carter model is a simple linear dynamic model. Hence, we also highlight that this model can be rewritten in the form of an ARIMA structure via a Local Level formulation where we denote $\boldsymbol{\eta}_t := \boldsymbol{\alpha} + \boldsymbol{\beta}\kappa_t$ and $h_t := \theta + \omega_t$. One can then rewrite the state-space form where each age x is an ARIMA(0,1,1) structure as $Z_{xt} := \nabla Y_{xt} = \nabla \boldsymbol{\eta}_{xt} + \nabla \boldsymbol{\varepsilon}_{xt}$ with a simple closed form expression for the auto-correlation function given by

$$\rho_{Z_x}(k) = \frac{\gamma(k)}{\gamma(0)} = \begin{cases} -\frac{\sigma_\varepsilon^2}{\sigma_\omega^2 + 2\sigma_\varepsilon^2}, & k = 1, \\ 0, & k \geq 1. \end{cases} \tag{4}$$

Suggesting that one can also perform estimation on the unconstrained form of the model via estimation based on the autocorrelation, though these would need to be modified subject to identification constraints. This would again complicate the estimation, suggesting the need to try to find alternative identification constraints that are more applicable to these standard estimation approaches.

3 BAYESIAN INFERENCE FOR LEE-CARTER MODEL IN STATE-SPACE FORM

Pedroza (2006) and Kogure and Kurachi (2010) both consider Bayesian formulations of the Lee-Carter model which allows the joint estimation of κ_t and model parameters. However, under their formulation they again work with the identification constraints proposed in Lee and Carter (1992) which are not obvious to use when designing efficient Monte Carlo procedures such as an Markov Chain Monte Carlo (MCMC) procedure. Such identification constraints will lead to difficulties in designing the proposal of the MCMC and difficulties in achieving suitable acceptance rates for the resultant Markov chain, resulting in high variance in estimates of mortality rates. Additionally, although these authors work in the Bayesian setting, their derivations of the posterior distributions are not fully described. In the following we derive the posterior distributions of the parameters and the state process of the Lee-Carter model under our extended Bayesian framework.

3.1 New Identification Constraints and Bayesian Formulations

We suggest an alternative new formulation of the identification constraints required which we believe is simpler and more readily applicable to most Monte Carlo based procedures such as MCMC and filtering methods such as Kalman Filter and Sequential Monte Carlo. This has the key advantage that for a given computational effort we can design efficient MCMC samplers with lower variance and therefore result in more reliable estimates of mortality rate. Our formulation of the identification constraints are given by simply setting $\alpha_{x_1} = \text{constant}$, and $\beta_{x_1} = \text{constant}$. Such a choice is a valid identification constraint since if one of the elements of each $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are known, then a non-trivial linear transformation is not allowed anymore (Section 2.2); that is, we must have $c = 0$ and $d = 1$.

Under the Bayesian approach, we aim to obtain the posterior density $\pi(\kappa_{0:n}, \boldsymbol{\Psi} | \mathbf{y}_{1:n})$ of the states ¹ $\kappa_{0:n}$ as well as the parameters, $\boldsymbol{\Psi} := (\alpha_{x_2:x_p}, \beta_{x_2:x_p}, \theta, \sigma_\varepsilon^2, \sigma_\omega^2)$, given the observations $\mathbf{y}_{1:n}$. In this paper we

¹Here $a_{1:t}$ means a_1, \dots, a_t .

explain an efficient and suitable sampling framework for actuarial applications which utilises the state-space Lee-Carter structure, in particular the fact that it is a linear Gaussian model, as well as the new constraint formulation we introduce. Under this model we develop an efficient approach involving a combined Gibbs sampling conjugate model sampler for the marginal target distributions of the static model parameters along with a forward backward Kalman filter sampler for the latent process $\kappa_{1:t}$.

A sample of the targeted density is obtained via Gibbs sampling in two steps: (1) Initialise $\Psi = \Psi^{(0)}$; (2) For $i = 1, \dots, N$, first draw $\kappa_{0:n}^{(i)}$ from $\pi(\kappa_{0:n}|\Psi^{(i-1)}, \mathbf{y}_{1:n})$, then draw $\Psi^{(i)}$ from $\pi(\Psi|\kappa_{0:n}^{(i)}, \mathbf{y}_{1:n})$.

3.2 Sampling from the full conditional density $\pi(\kappa_{0:n}|\Psi, \mathbf{y}_{1:n})$

Samples from the full conditional density $\pi(\kappa_{0:n}|\Psi, \mathbf{y}_{1:n})$ can be obtained via the so-called forward-filtering-backward sampling (FFBS) procedure (Carter and Kohn (1994)). We can write

$$\pi(\kappa_{0:n}|\Psi, \mathbf{y}_{1:n}) = \prod_{t=0}^n \pi(\kappa_t|\kappa_{t+1:n}, \Psi, \mathbf{y}_{1:n}) = \prod_{t=0}^n \pi(\kappa_t|\kappa_{t+1}, \Psi, \mathbf{y}_{1:t}) \quad (5)$$

where the last term in the product, $\pi(\kappa_n|\Psi, \mathbf{y}_{1:n})$, is distributed as $N(m_n, C_n)$ in Kalman filtering. We use the following notation

$$\kappa_{t-1}|\mathbf{y}_{1:t-1} \sim N(m_{t-1}, C_{t-1}) \quad (6)$$

$$\kappa_t|\mathbf{y}_{1:t-1} \sim N(a_t, R_t), \text{ where } a_t = m_{t-1} + \theta, R_t = C_{t-1} + \sigma_\omega^2 \quad (7)$$

$$\mathbf{y}_t|\mathbf{y}_{1:t-1} \sim N(\mathbf{f}_t, \mathbf{Q}_t), \text{ where } \mathbf{f}_t = \boldsymbol{\alpha} + \beta a_t, \mathbf{Q}_t = \beta\beta' R_t + \sigma_\varepsilon^2 \mathbf{1}_p \quad (8)$$

$$\kappa_t|\mathbf{y}_{1:t} \sim N(m_t, C_t), \text{ where } m_t = a_t + R_t\beta' \mathbf{Q}_t^{-1}(\mathbf{y}_t - \mathbf{f}_t), C_t = R_t - R_t\beta' \mathbf{Q}_t^{-1}\beta R_t \quad (9)$$

to denote the distributions involved in Kalman filtering. Once we draw a sample κ_n from $N(m_n, C_n)$, then Eq. (5) suggests that we can draw recursively and backwardly κ_t from $\pi(\kappa_t|\kappa_{t+1}, \Psi, \mathbf{y}_{1:t})$ where $t = n - 1, n - 2, \dots, 1, 0$. It can be shown that (Petris et al. (2009))

$$\pi(\kappa_t|\kappa_{t+1}, \Psi, \mathbf{y}_{1:t}) \sim N(h_t, H_t), \text{ where } h_t = m_t + C_t R_{t+1}^{-1}(\kappa_{t+1} - a_{t+1}), H_t = C_t - C_t R_{t+1}^{-1} C_t. \quad (10)$$

In summary, the FFBS algorithm consists of three steps: (1) Run Kalman filter to obtain m_n and C_n ; (2) Draw κ_n from $N(m_n, C_n)$ and (3) For $t = n - 1, \dots, 0$, draw κ_t from $N(h_t, H_t)$.

3.3 Sampling from the full conditional density $\pi(\Psi|\kappa_{0:n}, \mathbf{y}_{1:n})$

Sampling from the full conditional density $\pi(\Psi|\kappa_{0:n}, \mathbf{y}_{1:n})$ can be achieved by applying Gibbs sampling. The prior for $(\alpha_x, \beta_x, \theta, \sigma_\varepsilon^2, \sigma_\omega^2)$ are given by $\alpha_x \sim N(\tilde{\mu}_\alpha, \tilde{\sigma}_\alpha^2)$, $\beta_x \sim N(\tilde{\mu}_\beta, \tilde{\sigma}_\beta^2)$, $\sigma_\varepsilon^2 \sim \text{IG}(\tilde{a}_\varepsilon, \tilde{b}_\varepsilon)$, $\theta \sim N(\tilde{\mu}_\theta, \tilde{\sigma}_\theta^2)$, $\sigma_\omega^2 \sim \text{IG}(\tilde{a}_\omega, \tilde{b}_\omega)$ where $x \in \{x_2, \dots, x_p\}$ and $\text{IG}(\cdot, \cdot)$ denote the inverse-gamma distribution. It is assumed that the priors for all parameters are independent. In this case the posterior densities of parameters are of the same type as the prior densities, a so-called conjugate prior. In the following we derive the posterior distribution for each parameter (for ease of notation it is assumed that $\mathbf{y} = \mathbf{y}_{1:n}$, $\boldsymbol{\kappa} = \kappa_{0:n}$, family $\Psi_{-\lambda}$ means “ Ψ without the parameter λ ”):

- For α_x where $x \in \{x_2, \dots, x_p\}$, we have

$$\begin{aligned} \pi(\alpha_x|\mathbf{y}, \boldsymbol{\kappa}, \Psi_{-\alpha_x}) &\propto \pi(\mathbf{y}|\boldsymbol{\kappa}, \Psi)\pi(\boldsymbol{\kappa}|\Psi)\pi(\alpha_x|\Psi_{-\alpha_x}) \propto \prod_{t=1}^n \pi(y_{xt}|\kappa_t, \alpha_x, \beta_x, \sigma_\varepsilon^2)\pi(\alpha_x) \\ &\propto \exp\left\{-\frac{1}{2}\left(\frac{(\tilde{\sigma}_\alpha^2 n + \sigma_\varepsilon^2)\alpha_x^2 - 2(\tilde{\mu}_\alpha \sigma_\varepsilon^2 + \tilde{\sigma}_\alpha^2 \sum_t (y_{xt} - \beta_x \kappa_t))\alpha_x}{\tilde{\sigma}_\alpha^2 \sigma_\varepsilon^2}\right)\right\}. \end{aligned}$$

Hence the posterior conditional distribution of α_x is given by $N\left(\frac{\tilde{\mu}_\alpha \sigma_\varepsilon^2 + \tilde{\sigma}_\alpha^2 \sum_t (y_{xt} - \beta_x \kappa_t)}{\tilde{\sigma}_\alpha^2 n + \sigma_\varepsilon^2}, \frac{\tilde{\sigma}_\alpha^2 \sigma_\varepsilon^2}{\tilde{\sigma}_\alpha^2 n + \sigma_\varepsilon^2}\right)$.

- For β_x where $x \in \{x_2, \dots, x_p\}$, we have

$$\begin{aligned} \pi(\beta_x|\mathbf{y}, \boldsymbol{\kappa}, \Psi_{-\beta_x}) &\propto \pi(\mathbf{y}|\Psi)\pi(\boldsymbol{\kappa}|\Psi)\pi(\beta_x|\Psi_{-\beta_x}) \propto \prod_{t=1}^n \pi(y_{xt}|\kappa_t, \alpha_x, \beta_x, \sigma_\varepsilon^2)\pi(\beta_x) \\ &\propto \exp\left\{-\frac{1}{2}\left(\frac{(\tilde{\sigma}_\beta^2 \sum_t \kappa_t^2 + \sigma_\varepsilon^2)\beta_x^2 - 2(\tilde{\mu}_\beta \sigma_\varepsilon^2 + \tilde{\sigma}_\beta^2 \sum_t (y_{xt} - \alpha_x \kappa_t))\beta_x}{\tilde{\sigma}_\beta^2 \sigma_\varepsilon^2}\right)\right\}. \end{aligned}$$

Hence the posterior conditional distribution of β_x is given by $N\left(\frac{\tilde{\sigma}_\beta^2 \sum_t (y_{xt} - \alpha_x) \kappa_t + \tilde{\mu}_\beta \sigma_\varepsilon^2}{\tilde{\sigma}_\beta^2 \sum_t \kappa_t^2 + \sigma_\varepsilon^2}, \frac{\tilde{\sigma}_\beta^2 \sigma_\varepsilon^2}{\tilde{\sigma}_\beta^2 \sum_t \kappa_t^2 + \sigma_\varepsilon^2}\right)$.

- For θ , we have

$$\begin{aligned} \pi(\theta | \mathbf{y}, \boldsymbol{\kappa}, \boldsymbol{\Psi}_{-\theta}) &\propto \pi(\mathbf{y} | \boldsymbol{\kappa}, \boldsymbol{\Psi}) \pi(\boldsymbol{\kappa} | \boldsymbol{\Psi}) \pi(\theta | \boldsymbol{\Psi}_{-\theta}) \propto \prod_{t=1}^n \pi(\kappa_t | \kappa_{t-1}, \theta, \sigma_\omega^2) \pi(\theta) \\ &\propto \exp \left\{ -\frac{1}{2} \left(\frac{(\tilde{\sigma}_\theta^2 n + \sigma_\omega^2) \theta^2 - 2(\tilde{\mu}_\theta \sigma_\omega^2 + \tilde{\sigma}_\theta^2 \sum_t (\kappa_t - \kappa_{t-1})) \theta}{\tilde{\sigma}_\theta^2 \sigma_\omega^2} \right) \right\}. \end{aligned}$$

Hence the posterior conditional distribution of θ is given by $N\left(\frac{\tilde{\sigma}_\theta^2 \sum_{t=1}^n (\kappa_t - \kappa_{t-1}) + \tilde{\mu}_\theta \sigma_\omega^2}{\tilde{\sigma}_\theta^2 n + \sigma_\omega^2}, \frac{\tilde{\sigma}_\theta^2 \sigma_\omega^2}{\tilde{\sigma}_\theta^2 n + \sigma_\omega^2}\right)$.

- For σ_ε^2 , we have

$$\begin{aligned} \pi(\sigma_\varepsilon^2 | \mathbf{y}, \boldsymbol{\kappa}, \boldsymbol{\Psi}_{-\sigma_\varepsilon^2}) &\propto \pi(\mathbf{y} | \boldsymbol{\kappa}, \boldsymbol{\Psi}) \pi(\boldsymbol{\kappa} | \boldsymbol{\Psi}) \pi(\sigma_\varepsilon^2 | \boldsymbol{\Psi}_{-\sigma_\varepsilon^2}) \propto \prod_{t=1}^n \prod_{x=x_1}^{x_p} \pi(y_{xt} | \kappa_t, \alpha_x, \beta_x, \sigma_\varepsilon^2) \pi(\sigma_\varepsilon^2) \\ &\propto \frac{1}{(\sigma_\varepsilon^2)^{np/2 + \tilde{a}_\varepsilon + 1}} \exp \left\{ -\frac{1}{\sigma_\varepsilon^2} \left(\tilde{b}_\varepsilon + \frac{1}{2} \sum_t \sum_x (y_{xt} - (\alpha_x + \beta_x \kappa_t))^2 \right) \right\}. \end{aligned}$$

The posterior conditional distribution of σ_ε^2 is thus $\text{IG}\left(\tilde{a}_\varepsilon + \frac{np}{2}, \tilde{b}_\varepsilon + \frac{1}{2} \sum_{t=1}^n \sum_{x=x_1}^{x_p} (y_{xt} - (\alpha_x + \beta_x \kappa_t))^2\right)$.

- For σ_ω^2 , we have

$$\begin{aligned} \pi(\sigma_\omega^2 | \mathbf{y}, \boldsymbol{\kappa}, \boldsymbol{\Psi}) &\propto \pi(\mathbf{y} | \boldsymbol{\kappa}, \boldsymbol{\Psi}) \pi(\boldsymbol{\kappa} | \boldsymbol{\Psi}) \pi(\sigma_\omega^2 | \boldsymbol{\Psi}_{-\sigma_\omega^2}) \propto \prod_{t=1}^n \pi(\kappa_t | \kappa_{t-1}, \theta, \sigma_\omega^2) \pi(\sigma_\omega^2) \\ &\propto \frac{1}{(\sigma_\omega^2)^{n/2 + \tilde{a}_\omega + 1}} \exp \left\{ -\frac{1}{\sigma_\omega^2} \left(\tilde{b}_\omega + \frac{1}{2} \sum_t (\kappa_t - (\kappa_{t-1} + \theta))^2 \right) \right\}. \end{aligned}$$

The posterior conditional distribution of σ_ω^2 is thus $\text{IG}\left(\tilde{a}_\omega + \frac{n}{2}, \tilde{b}_\omega + \frac{1}{2} \sum_{t=1}^n (\kappa_t - (\kappa_{t-1} + \theta))^2\right)$.

3.4 Forecasting

The predictive distributions of \mathbf{y}_{n+k} , given \mathbf{y}_n , are obtained using the MCMC samples as follows. Let L be the number of samples remained after burn-in. Then for $k \geq 1$, and for $\ell = 1 \dots, L$, we sample recursively

$$\kappa_{n+k}^{(\ell)} \sim N\left(\kappa_{n+k-1}^{(\ell)} + \theta^{(\ell)}, (\sigma_\omega^2)^{(\ell)}\right), \quad \mathbf{y}_{n+k}^{(\ell)} \sim N\left(\boldsymbol{\alpha}^{(\ell)} + \boldsymbol{\beta}^{(\ell)} \kappa_{n+k}^{(\ell)}, (\sigma_\varepsilon^2)^{(\ell)} \mathbf{1}_p\right) \quad (11)$$

where the samples $\kappa_n^{(\ell)}$ are obtained from the FFBS procedure. This produces an estimate of $\pi(\mathbf{y}_{n+k} | \mathbf{y}_{1:n}) = \int \pi(\mathbf{y}_{n+k} | \kappa_{n+k}, \boldsymbol{\Psi}) \pi(\kappa_{n+k} | \kappa_{n+k-1}, \boldsymbol{\Psi}) \dots \pi(\kappa_n, \boldsymbol{\Psi} | \mathbf{y}_{1:n}) d\boldsymbol{\Psi} d\kappa_n \dots d\kappa_{n+k}$ and samples from it for forecasting.

4 IMPLICATIONS FOR ANNUITY PRICING

In this section we aim to quantify the impact of longevity risk on the pricing of annuities, using the mortality rates forecasted by the Lee-Carter model in state-space form which is estimated by the Bayesian approach described in the previous section.

4.1 Estimation using Australian mortality data

The data set consists of Australian female mortality data obtained from the Human Mortality Database (<http://www.mortality.org>). Since the application is for annuity pricing, we focus on 1-year death rates for age 60-100 from year 1975-2011. Figure 1 shows the estimation results. Here we set $\alpha_{x_1} = -5$, $\beta_{x_1} = 0.2$ and assume $m_0 = 0$, $C_0 = 100$, $\tilde{\mu}_\alpha = \tilde{\mu}_\beta = \tilde{\mu}_\theta = 0$, $\tilde{\sigma}_\alpha^2 = \tilde{\sigma}_\beta^2 = \tilde{\sigma}_\theta^2 = 100$, $\tilde{a}_\varepsilon = \tilde{a}_\omega = 2.1$ and $\tilde{b}_\varepsilon = \tilde{b}_\omega = 0.3$. Number of iterations in MCMC is 5000 and the burn-in iterations is 1000. We use very vague prior so that estimation is mainly determined by the data and the impact from prior is not material.

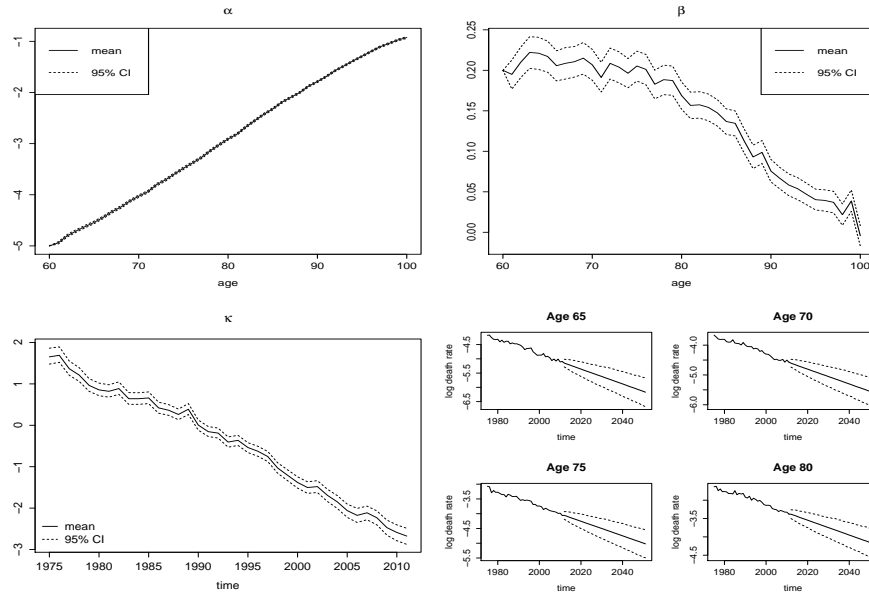


Figure 1. (Upper four panels) Posterior mean and 95% confidence interval (CI) for parameters α , β ; posterior mean and 95% CI for the latent process κ over year 1975–2011; mean and 95% CI of the predictive distributions of log central death rates ($y_{65}, y_{70}, y_{75}, y_{80}$) over 40 years forecast.

4.2 Annuity pricing

The τ year survival probability of a person aged x currently (i.e. $t = 0$ or year 2012) is determined by ${}_{\tau}p_x = \prod_{j=1}^{\tau} {}_1p_{x+j-1} = \prod_{j=1}^{\tau} e^{-m_{x+j-1,j-1}}$, which is a random variable since $m_{x+j-1,j-1}$, for $j = 1, \dots, \tau$, are random quantities forecasted by the Lee-Carter model. Assuming a large enough annuity portfolio, the price of an annuity with maturity T year, written for a x -year-old with benefit \$1 per year and conditional on the path $\mathbf{m}_{1:T}^x = (m_{x,0}, m_{x+1,1}, \dots, m_{x+T-1,T-1})$, is given by

$$a_x^T(\mathbf{m}_{1:T}^x) = \sum_{\tau=1}^T B(0, \tau) E(1_{T_x > \tau} | \mathbf{m}_{1:\tau}^x) = \sum_{\tau=1}^T B(0, \tau) {}_{\tau}p_x(\mathbf{m}_{1:\tau}^x) \tag{12}$$

where $B(0, \tau)$ is the τ -year bond price, $\mathbf{m}_{1:\tau}^x$ is the first τ elements of $\mathbf{m}_{1:T}^x$, and ${}_{\tau}p_x(\mathbf{m}_{1:\tau}^x)$ denotes the survival probability given $\mathbf{m}_{1:\tau}^x$ which is random. Denuit and Dhaene (2007) shows that some bounds of ${}_{\tau}p_x(\mathbf{m}_{1:\tau}^x)$ can be computed analytically. Biffis (2005) evaluates annuity prices allowing for longevity risk using financial theory. From an annuity provider’s perspective, what is important, however, is that the annuity price is a random quantity depending on the random paths of $\mathbf{m}_{1:T}^x$. Moreover, it is important to determine a survival curve ${}_{\tau}p_x$ (as a function of τ) in (12) that best captures the mortality experience of the portfolio for risk management purposes. In this regard, we evaluate different quantiles of the annuity price $a_x^T(\mathbf{m}_{1:T}^x)$ in Table 1 and extract the corresponding survival curves. Note that the forecasted death rate samples are used to produce sample paths $\mathbf{m}_{1:T}^{x,(\ell)}$ and hence samples of annuity prices $a_x^{T,(\ell)}(\mathbf{m}_{1:T}^x)$ where $\ell = 1, \dots, L$.

Impact of longevity risk. The possibility that the realised survival curve would be different to the survival curve assumed for pricing leads to the so-called systematic mortality risk, a.k.a. longevity risk. In Table 1 we compare the median, 0.025 quantile and 0.975 quantile of the annuity prices for different ages and maturities. We also assume a constant interest rate $r = 3\%$ and hence $B(0, \tau) = e^{-r\tau}$. Although the price difference might appear to be overall small, mis-pricing can be a significant risk when considering a large annuity portfolio. For an annuity portfolio consists of N policies where the benefit per year is B , an under-pricing of $\gamma\%$ of the “correct” annuity price will result in a shortfall of $NBa_x^T\gamma/100$ where a_x^T is the “wrong” annuity price being charged with benefit \$1 per year. For instance, $N = 10,000$ policies written to 80-year-old policyholders with maturity $\tau = 20$ years and \$20,000 benefit per year will result in a shortfall of \$67 million when the realised survival curve is the one that corresponds to the 0.975 quantile annuity price, while the survival curve corresponds to the median annuity price is assumed for pricing (here $\gamma = 4.1$ in Table 1). Moreover, as shown in Table 1, mis-pricing is increasingly more pronounced for older ages as well as for annuity policies having a longer maturity.

Table 1. Annuity price with different age and maturity (T) for female policyholder. Value in bracket () is the percentage difference compared to median annuity price. We only consider contracts with maturity so that age + maturity ≤ 100 .

| Maturity (years) | $T = 5$ | $T = 10$ | $T = 15$ | $T = 20$ | $T = 25$ | $T = 30$ |
|------------------|--------------|--------------|---------------|---------------|---------------|---------------|
| age = 65 | | | | | | |
| Median | 4.49 | 8.18 | 11.14 | 13.38 | 14.88 | 15.64 |
| 0.025 Q | 4.48 (-0.2%) | 8.13 (-0.6%) | 11.00 (-1.3%) | 13.10 (-2.1%) | 14.42 (-3.1%) | 15.03 (-3.9%) |
| 0.975 Q | 4.50 (+0.2%) | 8.22 (+0.6%) | 11.26 (+1.1%) | 13.63 (+1.9%) | 15.31 (+2.9%) | 16.22 (+3.7%) |
| age = 70 | | | | | | |
| Median | 4.42 | 7.94 | 10.57 | 12.30 | 13.15 | 13.41 |
| 0.025 Q | 4.41 (-0.4%) | 7.86 (-1.0%) | 10.37 (-1.9%) | 11.92 (-3.1%) | 12.63 (-4.0%) | 12.82 (-4.4%) |
| 0.975 Q | 4.44 (+0.4%) | 8.01 (+0.9%) | 10.76 (+1.8%) | 12.66 (+2.9%) | 13.67 (+4.0%) | 14.00 (+4.4%) |
| age = 75 | | | | | | |
| Median | 4.31 | 7.49 | 9.54 | 10.52 | 10.81 | N.A. |
| 0.025 Q | 4.29 (-0.7%) | 7.38 (-1.6%) | 9.27 (-2.8%) | 10.12 (-3.8%) | 10.35 (-4.3%) | N.A. |
| 0.975 Q | 4.34 (+0.6%) | 7.61 (+1.5%) | 9.80 (+2.8%) | 10.92 (+3.8%) | 11.28 (+4.3%) | N.A. |
| age = 80 | | | | | | |
| Median | 4.08 | 6.63 | 7.83 | 8.18 | N.A. | N.A. |
| 0.025 Q | 4.03 (-1.1%) | 6.48 (-2.4%) | 7.57 (-3.4%) | 7.86 (-3.9%) | N.A. | N.A. |
| 0.975 Q | 4.12 (+1.1%) | 6.79 (+2.3%) | 8.10 (+3.4%) | 8.51 (+4.1%) | N.A. | N.A. |

5 CONCLUSIONS

This article explores further the state-space representation of the Lee-Carter model in longevity modeling. We derive in details the posterior distributions of the static parameters and the latent process of the model under the Bayesian framework via Gibbs sampling. We suggest an identification constraint for the model that is particularly suitable for estimation under a MCMC approach. The predictive distributions of death rates are used to determine the range of annuity prices. Our results show that the assumption of survival curve has significant impact on annuity prices. Annuity written for older age policyholders is particularly vulnerable to mis-pricing caused by longevity risk.

REFERENCES

- Biffis, E. (2005). Affine processes for dynamic mortality and actuarial valuations. *Insurance: Mathematics and Economics* 37(3), 443–468.
- Brouhns, N., M. Denuit, and J. K. Vermunt (2002). A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics* 31, 373–393.
- Carter, C. K. and R. Kohn (1994). On Gibbs sampling for state space models. *Biometrika* 81(3), 541–553.
- Czado, C., A. Delwarde, and M. Denuit (2005). Bayesian Poisson log-bilinear mortality projections. *Insurance: Mathematics and Economics* 36, 260–284.
- Denuit, M. and J. Dhaene (2007). Comonotonic bounds on the survival probabilities in the Lee-Carter model for mortality projection. *Journal of Computational and Applied Mathematics* 203, 169–176.
- Hirz, J., U. Schmock, and P. V. Shevchenko (2015). Modelling annuity portfolios and longevity risk with extended creditrisk⁺. Preprint, arxiv: 1505.04757.
- Kogure, A. and Y. Kurachi (2010). A Bayesian approach to pricing longevity risk based on risk-neutral predictive distributions. *Insurance: Mathematics and Economics* 46, 162–172.
- Lee, R. D. and L. R. Carter (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association* 87, 659–675.
- Pedroza, C. (2006). A Bayesian forecasting model: predicting U.S. male mortality. *Biostatistics* 7(4), 530–550.
- Petris, G., S. Petrone, and P. Campagnoli (2009). *Dynamic Linear Models with R*. Springer.
- Pitacco, E., M. Denuit, S. Haberman, and A. Olivieri (2009). *Modelling Longevity Dynamics for Pensions and Annuity Business*. Oxford University Press.