A Bayesian Approach to Longevity Derivative Pricing
under Stochastic Interest Rates with a Two-factor Lee-Carter Model

Takahiro Fushimi and Atsuyuki Kogure *
Faculty of Policy Management, Keio University, Japan
July 2014

Abstract

Mortality-linked securities such as longevity bonds or survivor swaps typically depend on not only mortality risk but also interest rate risk. However, in the existing pricing methodologies, it is often the case that only the mortality risk is modeled to change in a stochastic manner and the interest rate is kept fixed at a pre-specified level. In order to develop large and liquid longevity markets, it is essential to incorporate the interest rate risk into pricing mortality-linked securities. In this paper we tackle the issue by considering the pricing of longevity derivatives under stochastic interest rates following the CIR model (Cox, Ingersoll and Ross, 1985). As for the mortality modeling, we use a two-factor extension of the Lee-Carter model by noting the recent studies which point out the inconsistencies of the original Lee-Carter model with observed mortality rates due to its single factor structure. With the issue of parameter uncertainty, we propose using a Bayesian methodology both to estimate the models and to price longevity derivatives in line with Kogure and Kurachi (2010) and Kogure, Li and Kamiya (2013). We investigate the actual effects of the incorporation of the interest rate risk by applying our methodology to price a longevity bond and a longevity cap with Japanese data. The results show significant differences according to whether the CIR model is used or the interest rate is kept fixed at a certain level.

Keywords: Longevity derivatives, Lee-Carter Methodology, Interest Rate Risk, CIR Model, Bayesian Pricing, Maximum Entropy Principle.

1 Introduction

Mortality-linked securities such as longevity bonds or survivor swaps typically depend on not only mortality risk but also interest rate risk. However, in the existing pricing methodologies, it is often the case that only the mortality risk is modeled to change in a stochastic manner and the interest rate is kept fixed at a pre-specified level. In this paper we tackle the issue by considering the pricing of longevity derivatives under stochastic interest rates following the CIR model (Cox, Ingersoll and Ross, 1985). As for the mortality modeling we use a two-factor extension of the Lee-Carter model by noting the recent studies which point out the inconsistencies of the original Lee-Carter model with observed mortality rates due to its single factor structure; see Lazer and Denuit (2009) and Renshaw and Haberman (2005).

As argued in Blake, Cairns and Dowd (2006b), in addition to the process risk due to the stochastic nature of the risk given a statistical model, the parameter uncertainty should be reflected into the pricing as well. In the same line with Kogure and Kurachi (2010) and Kogure, Li and Kamiya (2013), we propose using a Bayesian approach to deal with both the process risk and the parameter uncertainty. We

*Corresponding author Email: kogure@sfc.keio.ac.jp
employ a Bayesian CIR model formulated by Feng and Xie (2012) and a Bayesian two-factor Lee-Carter model presented by Kogure, Fushimi and Takamatsu (2012) to derive the predictive distribution for the present value of the longevity bond. We then apply the risk neutral pricing approach proposed by Stutzer (1996) to convert the predictive distribution into the risk-neutral version based on the maximum entropy principle. We investigate the actual effects of the incorporation of the interest rate risk by applying our pricing methodology to male mortality rates in Japan and interest rates of Japanese Government Bonds.

There have been several works on the valuation of mortality-linked securities under stochastic mortality rates and stochastic interest rates in the literature. Recent examples include Jalen and Ramon (2009) and Shen and Siu (2013). We believe that the present work is one of the first attempts to the valuation with mortality and interest rate risks from a Bayesian perspective. The Bayesian CIR model we use is based on an approximation to the original continuous-time CIR model discretized at the observation interval, which causes some difficulty when the magnitude of the observation interval is large. To alleviate this problem a Bayesian technique is applied to introduce augmented data between each pair of observations.

The remainder of this paper is organized as follows. Section 2 introduces longevity derivatives and a general strategy of how to evaluate them. Section 3 discusses about mortality modeling and Bayesian Lee-Carter models implemented in this paper. Section 4 discusses the Bayesian CIR model estimated in this article. Section 5 considers the Bayesian pricing of longevity derivatives and applies it Japanese data. Section 6 closes the paper.

2 Longevity Derivatives

2.1 EIB/BNP longevity bond

In November 2004, European Investment Bank (EIB) and BNP Paribas announced the first long-term longevity bond designed for pension plans and other annuity providers. Actually, this bond was not issued due to the unattractiveness to investors though, but we examine and price this security because it is considered one of the most practicable mortality linked instruments among hypothetical products. This security was to be issued by EIB, with BNP Paribas as the designer and originator, and Partner Re as the longevity risk reinsurer. The face value of the bond was 540 million ponds and it had a 25-year maturity. The security was a bond whose coupon payments are linked to the cohort survivor index based on the mortality rate of English and Welsh males aged 65 in 2002. In the absence of credit risk, survivor index at time $t$, $S(t)$ is given by

\[
S(0) = 1 \\
S(1) = S(0) \times t_{p65}(2002) \\
\vdots \\
S(t) = S(0) \times t_{p65}(2002)
\]

where $t_{p65}(2002)$ is the survivor probability that a male aged 65 at 2002 will be alive at least $t$ more years. Using this survivor index $S(t)$, the floating coupon payment at time $t$ is calculated as

\[
C_t = 50,000,000 \times S(t) = 50,000,000 \times S(0) \times t_{p65}(2002).
\]

From the investors' perspective, they pay for an initial payment of around 540 million pond which is the issue price and receive an annual mortality dependent payment of $C_t$ in each year $t$ for 25 years.

In reality, this attempt was eventually failed. Blake, Cairns, and Dowd (2006a) discussed the possible reasons for the failure.

2.2 Longevity derivatives on a survivor index

Let $S_x(t)$ denote a survival index which represents a survival rate at time $t$ of a population aged $x$ at time 0. Then a longevity derivative on $S_x(t)$ is a financial product which pays $C(S_x(t))$ for $t = 1, 2, \cdots, T$ according to a particular functional form of $C(\cdot)$. Examples include:
\[ C(S_x(t)) = S_x(t) \text{ for } t = 1, 2, \ldots, T - 1 \text{ and } C(S_x(T)) = B + S_x(T). \] Then it represents a longevity bond with principal payment \( B \) and floating coupon payments \( S_x(t) \)'s.

\[ C(S_x(t)) = S_x(t) - K_t \text{ for each } t. \] Then it represents a survivor swap which exchanges \( S_x(t) \) with \( K_t \) at time \( t \). Here \( K_t \) are swap rates predetermined at time 0.

\[ C(S_x(t)) = \max(S_x(t) - K_t, 0) \text{ for each } t. \] Then it represents a longevity cap which is a string of call options on \( S_x(t) \) with exercise price \( K_t \).

See Blake Cairns and Dowd (2006a) for the other longevity derivatives.

### 2.3 A general strategy for pricing longevity derivatives

Let \( \{r_t, t = 0, 1, \} \) represent a one-period short interest rate process. Then, theoretically, the price at time 0 of a longevity derivative that pays \( C(S_x(t)) \) at time \( t = 1, \ldots, T \), is

\[
P_0(x, T) = E^Q \left[ \sum_{t=1}^{T} \exp \left( - \sum_{u=0}^{t} r_u \right) C(S_x(t)) \mid \mathcal{F}_0 \right]
\]

(1)

where \( \mathcal{F}_0 \) represents the information about mortality and interest rate processes available at time 0 and \( E^Q[\cdot \mid \mathcal{F}_0] \) is the expectation with respect to the risk neutral measure \( Q \), conditioning on \( \mathcal{F}_0 \). Therefore, the pricing of a longevity derivative calls for the following two steps:

1. Statistical modeling of \( S_x(t) \) and \( r_t \)
2. Risk neutralization for \( S_x(t) \) and \( r_t \)

Recently, Bayesian methods have been increasingly used to deal with parameter uncertainty; see Blake, Cairns and Dowd (2006b). However, their use is mainly limited to step 1. In this paper we attempt to achieve steps 1 and 2 in a unified Bayesian framework proposed and applied in Kogure and Kurachi (2010) and Kogure, Li and Kamiya (2013).

### 3 Bayesian Lee-Carter Methodologies

#### 3.1 Mortality modeling

Let \( \mu_x(t) \) denote the force of mortality, the instantaneous death rate, for an individual aged exactly \( x \) at time \( t \). Then the probability that he or she will die within one year is given by

\[
\nu q_x = 1 - \exp \left\{ - \int_0^1 \mu_{x+s}(t + s) ds \right\}.
\]

(2)

We assume that

\[
\mu_{x+s}(t + s) = \mu_x(t)
\]

for integers \( x \) and \( t \), and for all \( 0 \leq s, u < 1 \). That is, the force of mortality is constant with squares of age and time. Then the death probability (2) for integers \( x \) and \( t \) becomes

\[
\nu q_x = 1 - \exp(-\mu_x(t)).
\]

(3)

In the Lee-Carter methodology, the force of mortality is modeled as

\[
\mu_x(t) = \exp(\alpha_x + \beta_x \kappa_t),
\]

(4)

where \( \alpha_x, \beta_x \) and \( \kappa_t \) are parameters to be estimated. Here, to ease the constraint due to its single factor structure, we consider a two-factor extension of (4):

\[
\mu_x(t) = \exp(\alpha_x + \beta_1 \kappa_{t_1} + \beta_2 \kappa_{t_2}).
\]

for which (3) becomes

\[
\nu q_x = 1 - \exp\{ - \exp(\alpha_x + \beta_1 \kappa_{t_1} + \beta_2 \kappa_{t_2}) \}.
\]

(5)
3.2 Statistical modeling

Let $m_{xt}$ denote the crude death rate

$$m_{xt} = \frac{D_{xt}}{E_{xt}}, \quad x = x_{\min}; x_{\min} + 1, \cdots, x_{\max}; \quad t = t_{\min}; t_{\min} + 1, \cdots, t_{\max},$$

where $D_{xt}$ is the number of deaths recorded at age $x$ during year $t$ and $E_{xt}$ is the number of person years from which $D_{xt}$ occurred. It is the observed mortality rate corresponding to $\mu_x(t)$ for integers $x$ and $t$ in light of the assumption. It is often the case that the mid-year population is substituted for $E_{xt}$. For each $t$, we define $y_t$ to be a vector

$$y_t = (y_{x_{\min}t}, \cdots, y_{x_{\max}t})^T$$

consisting of the log mortality rate $y_{xt} = \log m_{xt}$ across ages observed at year $t$. Let $M = x_{\max} - x_{\min} + 1$. Then the standard Lee-Carter model is given by

$$y_t = \alpha + \beta t + \varepsilon_t, \quad \varepsilon_t \sim N_M(0, \sigma^2 I_M)$$

(6) where $\alpha = (\alpha_{x_{\min}}, \cdots, \alpha_{x_{\max}})^T$, $\beta = (\beta_{x_{\min}}, \cdots, \beta_{x_{\max}})^T$, and $\varepsilon_t$ is an error vector distributed as the $M$-dimensional normal distribution with mean $0_M$ and variance $\sigma^2 I_M$. Here $0_M$ is the $M$-dimensional vector of all zeroes and $I_M$ is the identity matrix of dimension $M \times M$. The equation (6) decomposes $y_{xt} = \log m_{xt}$ into three components. The $\alpha_x$’s describe the age-pattern of mortality averaged over time, whereas the $\beta_x$’s describe the deviations from the averaged pattern when $\kappa_t$ varies. The change in the level of mortality over time is described by the univariate mortality index $\kappa_t$. The Lee-Carter model differs from "parametric" models because the dependence on age in (6) is non-parametric and is represented by the sequences of $\alpha_x$ and $\beta_x$.

The two factor version becomes

$$y_t = \alpha + B \kappa_t + \varepsilon_t, \quad \varepsilon_t \sim N_M(0, \sigma^2 I_M)$$

(7) where $\kappa_t = (\kappa_{1t}, \kappa_{2t})^T$ and $B$ is an $M \times 2$ matrix with its $i$th row being $\beta_i = (\beta_{ix_{\min}}, \cdots, \beta_{ix_{\max}})^T$. Let $L = t_{\max} - t_{\min} + 1$. Then, given $\kappa_t$’s, the likelihood function of (7) is written as

$$l(y|\alpha, B, \kappa_{t_{\min}}, \cdots, \kappa_{t_{\max}}, \sigma^2)$$

$$= \prod_{t=t_{\min}}^{t_{\max}} \prod_{x=x_{\min}}^{x_{\max}} f(y_{xt}|\alpha, \beta_{1x}, \beta_{2x}, \kappa_{1t}, \kappa_{2t}, \sigma^2)$$

$$\propto \left(\frac{1}{\sigma^2}\right)^L \exp \left\{ - \frac{1}{2\sigma^2} \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} \sum_{x=x_{\min}}^{x_{\max}} (y_{xt} - (\alpha_x + \beta_1 \kappa_{1t} + \beta_2 \kappa_{2t}))^2 \right\}.$$  

To make the likelihood identifiable, the parameters $B$ and $\{\kappa_t\}$ are restricted such that

$$\sum_{x=x_{\min}}^{x_{\max}} \beta_{ix} = 1 \quad \text{and} \quad \sum_{t=t_{\min}}^{t_{\max}} \kappa_{it} = 0, \quad i = 1, 2,$$

which in turn forces $\alpha_x$ to be the average of the log-central death rates over calendar years for each $x$.

3.3 Bayesian dynamic factor models

Considering the nature of the $\kappa_t$ as latent variables, it is often the case that Lee-Carter methodology is set up as a state-space model. The one-factor Lee-Carter model is set up as

$$\begin{cases}
\text{observation equation:} & y_t = \alpha + \beta \kappa_t + \varepsilon_t, \quad \varepsilon_t \sim N_M(0, \sigma^2 I_M) \\
\text{state equation:} & \kappa_t = \lambda + \kappa_{t-1} + \omega_t, \quad \omega_t \sim N(0, \sigma^2) 
\end{cases}$$

(10) where $\beta = (\beta_{x_{\min}}, \cdots, \beta_{x_{\max}})$. The state equation in (10) represents stochastic process of mortality index $\kappa_t$ by modeling as a random walk with drift.
In this paper, we also consider the Bayesian analysis of the two-factor Lee-Carter model:

\[
\begin{align*}
\text{observation equation: } & \quad y_t = \alpha + \beta_1 \kappa_{1t} + \beta_2 \kappa_{2t} + \epsilon_t, \quad \epsilon_t \sim N_M(0, \sigma^2_\epsilon I_M) \\
\text{1st state equation: } & \quad \kappa_{1t} = \lambda_1 + \kappa_{1t-1} + \omega_{1t}, \quad \omega_{1t} \sim N(0, \sigma^2_{\omega_{1t}}) \\
\text{2nd state equation: } & \quad \kappa_{2t} = \phi \lambda_2 + \kappa_{2t-1} + \omega_{2t}, \quad \omega_{2t} \sim N(0, \sigma^2_{\omega_{2t}})
\end{align*}
\]

where \( \beta_1 = (\beta_{1_{\text{min}}}, \ldots, \beta_{1_{\text{max}}}) \), \( \beta_2 = (\beta_{2_{\text{min}}}, \ldots, \beta_{2_{\text{max}}}) \), and \( |\phi| < 1 \). This model is considered to be a natural extension of (10). Within the state-space modeling framework, these two models are also known as dynamic factor models (Petris, Petrone and Campagnoli, 2009). The two-factor Lee-Carter model can capture the systematic effect influencing all mortalities across ages as well as the volatility of all mortalities by the second factor \( \kappa_{2t} \). We suppose that the first factor \( \kappa_{1t} \) reflects the time trend in mortality rates and that the second factor \( \kappa_{2t} \) describes a cyclical fluctuation around the trend. Thus, we assume that \( \kappa_{1t} \) is random walk with a drift term in the same way as original model and \( \kappa_{2t} \) is an AR(1) process. By including the second factor \( \kappa_{2t} \), we can capture the variability of mortality rates which cannot be accounted for by the one-factor model with its single factor structure.

In the Bayesian analysis, we first need to specify prior distributions for the parameters involved. We then make inferences based on the posterior distribution, which is obtained by combining the likelihood (8) with the prior distributions. For many cases, the analytical form of the posterior distribution cannot be obtained in an explicit form. However, using the Markov Chain Monte Carlo (MCMC) method, it is now possible to perform a high-speed sampling from the posterior distribution. The prior distributions for \( \alpha \) in the observation equation the parameter is chosen as

\[ \alpha \sim \text{Normal}_M(0_M, \sigma^2_\alpha I_M), \]

and for \( \beta_1 \) and \( \beta_2 \) are set up as

\[ \beta_i \sim \text{Normal}_M((1/M)1_M, \sigma^2_{\beta_i} I_M), \quad i = 1, 2. \]

Where \( 1_M \) is the M-dimensional vector of all ones. Here in the rest of the paper, underlined characters denote hyperparameters, which are parameters for the prior distributions. The prior for \( \sigma^2_\epsilon \) is set up as

\[ \sigma^2_\epsilon \sim \text{Inverse Gamma}(a_\epsilon, b_\epsilon) \]

The priors for the parameters \( \lambda_1, \lambda_2 \) in the state equation are chosen as

\[ \lambda_i \sim \text{Normal}(\lambda_{0i}, \sigma^2_{\lambda_i}), \quad i = 1, 2, \]

for \( \sigma^2_{\lambda_1} \) and \( \sigma^2_{\lambda_2} \) are set up as

\[ \sigma^2_{\omega_i} \sim \text{Inverse Gamma}(a_{\omega_i}, b_{\omega_i}), \quad i = 1, 2, \]

and for \( \phi \) is chosen as

\[ \phi \sim \text{Truncated Normal}_{(-1, 1)}(0, \sigma^2_\phi). \]

Also, the prior for \( \kappa_{1_{\text{min}}} \) and \( \kappa_{2_{\text{min}}} \) are chosen as

\[ \kappa_{1_{\text{min}}} \sim \text{Normal}(c_{1}, R), \quad i = 1, 2. \]

### 3.4 Sampling from posterior distribution

Denoting by \( \theta \) the set of all parameters involved, the joint posterior density \( f(\theta | y) \) of \( \theta \), given the observed data \( y \) of \( \theta \), given the observed data \( y \), may be written as

\[ f(\theta) \propto l(y | \theta) f(\theta). \]

Here, \( l(y | \theta) \) represents the likelihood and \( f(\theta) \) is the joint prior density for \( \theta \) described in the previous section. Using the MCMC procedure, we generate samples

\[ \theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(j)} \sim f(\theta | y) \]

from the joint posterior distribution.
3.5 MCMC procedure

The steps in MCMC procedure are as follows:

Step1: Initialize all parameters.

Step2: Use Gibbs sampler to
1) update \( \alpha_x \) from \( f(\alpha_x | y, \alpha_{-x}, \beta_1, \beta_2, \kappa_1, \kappa_2, \lambda_1, \lambda_2, \sigma^2_x, \sigma^2_{\omega_1}, \sigma^2_{\omega_2}) \),
2) update \( \beta_1 \) from \( f(\beta_1 | y, \alpha, \beta_{-1}, \kappa_1, \kappa_2, \lambda_1, \lambda_2, \sigma^2_x, \sigma^2_{\omega_1}, \sigma^2_{\omega_2}) \),
3) update \( \beta_2 \) from \( f(\beta_2 | y, \alpha, \beta_{-2}, \kappa_1, \kappa_2, \lambda_1, \lambda_2, \sigma^2_x, \sigma^2_{\omega_1}, \sigma^2_{\omega_2}) \),
4) update \( \sigma^2_x \) from \( f(\sigma^2_x | y, \alpha, \beta_1, \beta_2, \kappa_1, \kappa_2, \lambda_1, \lambda_2, \sigma^2_x, \sigma^2_{\omega_1}, \sigma^2_{\omega_2}) \),
5) update \( \lambda_1 \) from \( f(\lambda_1 | y, \alpha, \beta_1, \kappa_1, \kappa_2, \lambda_2, \sigma^2_x, \sigma^2_{\omega_1}, \sigma^2_{\omega_2}) \),
6) update \( \lambda_2 \) from \( f(\lambda_2 | y, \alpha, \beta_1, \beta_2, \kappa_1, \kappa_2, \lambda_1, \sigma^2_x, \sigma^2_{\omega_1}, \sigma^2_{\omega_2}) \),
7) update \( \sigma^2_{\omega_1} \) from \( f(\sigma^2_{\omega_1} | y, \alpha, \beta_1, \beta_2, \kappa_1, \kappa_2, \lambda_1, \lambda_2, \sigma^2_x, \sigma^2_{\omega_2}) \),
8) update \( \sigma^2_{\omega_2} \) from \( f(\sigma^2_{\omega_2} | y, \alpha, \beta_1, \beta_2, \kappa_1, \kappa_2, \lambda_1, \lambda_2, \sigma^2_x, \sigma^2_{\omega_1}) \),
9) update \( \phi \) from \( f(\phi | y, \alpha_x, \beta_1, \beta_2, \kappa_1, \kappa_2, \lambda_1, \lambda_2, \sigma^2_x, \sigma^2_{\omega_1}, \sigma^2_{\omega_2}) \).

Step3: Employ Forward Filtering Backward Sampling(FFBS) to
10) update \( \kappa_1 \) from \( f(\kappa_1 | y, \alpha, \beta_1, \beta_2, \kappa_1+, \kappa_2, \lambda_1, \lambda_2, \sigma^2_x, \sigma^2_{\omega_1}, \sigma^2_{\omega_2}) \),
11) update \( \kappa_2 \) from \( f(\kappa_2 | y, \alpha, \beta_1, \beta_2, \kappa_1+, \kappa_2, \lambda_1, \lambda_2, \sigma^2_x, \sigma^2_{\omega_1}, \sigma^2_{\omega_2}) \).

Step 4: Repeat Step 2 and Step 3 until the sampling size \( N \) is reached.

<table>
<thead>
<tr>
<th>hyperparameter</th>
<th>specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_x )</td>
<td>2.1</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>( (a_x - 1)\sigma^2_x ) the sample variance of ( {\hat{\alpha}_x} )</td>
</tr>
<tr>
<td>( \sigma^2_x )</td>
<td>the sample variance of ( {\hat{\beta}_1} )</td>
</tr>
<tr>
<td>( \lambda_0 )</td>
<td>the sample mean of ( {\hat{\kappa}<em>i - \hat{\kappa}</em>{i-1}} )</td>
</tr>
<tr>
<td>( \sigma^2_{\omega_1} )</td>
<td>(the sample variance of ( {\hat{\kappa}<em>i - \hat{\kappa}</em>{i-1}} ))/L</td>
</tr>
<tr>
<td>( \gamma_{\ell} )</td>
<td>( \hat{\kappa}_{\min} ) the sample variance of ( {\hat{\kappa}_i} )</td>
</tr>
<tr>
<td>( \alpha_{\omega} )</td>
<td>2.1</td>
</tr>
<tr>
<td>( \beta_{\omega} )</td>
<td>( (a_\omega - 1)\times) the sample variance of ( {\hat{\kappa}<em>i - \hat{\kappa}</em>{i-1}} )</td>
</tr>
<tr>
<td>( \sigma^2_{\omega} )</td>
<td>the sample mean of ( {\hat{\kappa}<em>i - \hat{\kappa}</em>{i-1}} )</td>
</tr>
</tbody>
</table>

Table 1: Hyperparameters of the Bayesian Lee-Carter models

3.6 Bayesian estimation

The data set we used are the numbers of deaths and populations of Japan between ages 65 and 98 over the period from 1973 to 2008 both for the male, taken from the Population Census and the Vital Statistics of Japan. We performed 15,000 steps of the MCMC sampling for the parameters involved. For all samplings, we discarded the first 5,000 steps and used the remaining 10,000 steps to eliminate initial effects. Following Kogure and Kurachi (2010), hyperparameters are chosen as shown in Table 1. In Table 1, \( \hat{\alpha}_x \), \( \hat{\beta}_x \), and \( \hat{\kappa}_i \) denote the maximum-likelihood estimators for parameters \( \alpha_x \), \( \beta_x \), and \( \kappa_i \), respectively.

The posterior mean, posterior SD, 95% intervals of the highest posterior density (HPD), and Geweke’s convergence diagnostics \( (p \text{ value}) \) are reported in Table 2. Figure 1 represents the traces of the MCMC steps for \( \alpha_{65}, \beta_{65} \), and \( \kappa_{1973} \) of the one-factor model. Figure 2 shows the trajectories of the MCMC steps for \( \alpha_{65}, \beta_{65}, \kappa_{1973}, \kappa_{1973} \), and \( \phi \) of the two-factor model. We notice that all the trajectories become stable after the first dozen steps in all of the graphs. Figure 3 shows smoothed histograms for \( \alpha_{65}, \beta_{65} \), and \( \kappa_{1973} \) of one-factor model. Figure 4 describes smoothed histograms for \( \alpha_{65}, \beta_{65}, \kappa_{1973}, \kappa_{1973} \), and \( \phi \) of the two-factor model. The two panels of Figure 5 plot the sample means of \( \alpha_x \) and \( \beta_x \) for ages between 65 and 98 of the one-factor model. The first panel reveals the linearly increasing pattern of age specific effects, as might be expected. The second panel measures the improvement in mortality at each age due to \( \kappa_1 \). It seems that those persons around age 75 enjoy the improvement most. Figure 6
Table 2: Summary statistics of MCMC samples

<table>
<thead>
<tr>
<th></th>
<th>Posterior mean</th>
<th>Posterior SD</th>
<th>95%HPD</th>
<th>Geweke</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>One-factor</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_{65}$</td>
<td>-4.074</td>
<td>0.010</td>
<td>(-4.094 , -4.053)</td>
<td>0.914</td>
</tr>
<tr>
<td>$\beta_{65}$</td>
<td>0.031</td>
<td>0.002</td>
<td>(0.028 , 0.035)</td>
<td>0.469</td>
</tr>
<tr>
<td>$\kappa_{1973}$</td>
<td>11.778</td>
<td>0.303</td>
<td>(11.206 , 12.378)</td>
<td>0.959</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>-0.582</td>
<td>0.078</td>
<td>(-0.731 , -0.421)</td>
<td>0.834</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>0.004</td>
<td>0.000</td>
<td>(0.004 , 0.004)</td>
<td>0.660</td>
</tr>
<tr>
<td>$\sigma^2_2$</td>
<td>0.343</td>
<td>0.115</td>
<td>(0.154 , 0.571)</td>
<td>0.976</td>
</tr>
<tr>
<td><strong>Two-factor</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_{65}$</td>
<td>-4.074</td>
<td>0.011</td>
<td>(-4.094 , -4.052)</td>
<td>0.079</td>
</tr>
<tr>
<td>$\beta_{65}$</td>
<td>0.031</td>
<td>0.000</td>
<td>(0.031 , 0.032)</td>
<td>0.642</td>
</tr>
<tr>
<td>$\beta_{265}$</td>
<td>0.029</td>
<td>0.004</td>
<td>(0.022 , 0.036)</td>
<td>0.569</td>
</tr>
<tr>
<td>$\kappa_{1973}$</td>
<td>11.631</td>
<td>0.313</td>
<td>(10.996 , 12.215)</td>
<td>0.847</td>
</tr>
<tr>
<td>$\kappa_{2973}$</td>
<td>-0.004</td>
<td>0.210</td>
<td>(-0.430 , 0.409)</td>
<td>0.964</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>-0.579</td>
<td>0.075</td>
<td>(-0.725 , -0.429)</td>
<td>0.489</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.016</td>
<td>0.030</td>
<td>(-0.040 , 0.077)</td>
<td>0.726</td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.000</td>
<td>0.029</td>
<td>(-0.058 , 0.057)</td>
<td>0.402</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>0.004</td>
<td>0.000</td>
<td>(0.004 , 0.004)</td>
<td>0.230</td>
</tr>
<tr>
<td>$\sigma^2_{21}$</td>
<td>0.292</td>
<td>0.102</td>
<td>(0.123 , 0.495)</td>
<td>0.970</td>
</tr>
<tr>
<td>$\sigma^2_{22}$</td>
<td>0.049</td>
<td>0.034</td>
<td>(0.010 , 0.113)</td>
<td>0.903</td>
</tr>
</tbody>
</table>

Figure 1: MCMC steps of selected parameters (one-factor)

Figure 7 gives plots of the past pattern and the future prediction of the mean levels of the $\kappa_t$. It also gives 95% prediction intervals for the future levels which depict longevity risk. The left panel in Figure 8 gives plots of the past patterns and the future predictions of the $\kappa_1$. It represents the steadily declining pattern in mortality, found in the single-factor Lee Carter model. The right panel in Figure 8 provides the plots of the past patterns and the future forecasts of the $\kappa_2$. It describes a cyclical variation which is not captured by the one-factor model.

Figure 9 plots the expected values of $t p_x$ projected by one-factor and two-factor models. The dashed curves in each plot show 95% intervals of the distribution of $t p_x$. We can observe that the resulting intervals is initially very narrow but becomes quite wide by the 25-year time horizon. The expected values of $t p_x$ is quite similar to each model but the 95% intervals of two-factor is slightly narrower than that of one-factor from the 10-year time horizon.
Figure 2: MCMC steps of selected parameters (two-factor)

Figure 3: Posterior distributions of selected parameters (one-factor)
Figure 4: Posterior distributions of selected parameters (two-factor)

Figure 5: Sample means of $\alpha_x, \beta_x$ (one-factor)

Figure 6: Sample means of $\alpha_x, \beta_{1x}, \beta_{2x}$ (two-factor)
Figure 7: Sample means of $\kappa_t$ (one-factor)

Figure 8: Sample means of $\kappa_{1,t}, \kappa_{2,t}$ (two-factor)

Figure 9: Projected survival probabilities of Japanese population
4 Bayesian CIR model

4.1 Interest rate modeling

We define $v(t)$ to be the value of a bank account at time $t > 0$. We assume $v(0) = 1$ and that the bank account evolves according to the following differential equation:

$$dv(t) = r(t)v(t)dt, \quad v(0) = 1,$$

(12)

where $r(t)$ is a positive function of time. As a consequence,

$$v(t) = \exp \left( \int_0^t r(s)ds \right).$$

(13)

The above definition tells us that investing a unit amount at time 0 yields at time $t$ the value in (13), and $r(s)$ is the instantaneous rate at which the bank account accrues. The instantaneous rate $r(s)$ is usually referred to as a spot rate. Similarly, the stochastic discount factor $D(t, T)$ between two time instants $t$ and $T$ is the amount at time $t$ that is equivalent to one unit of currency payable at time $T$, and is given by

$$D(t, T) = \frac{v(t)}{v(T)} = \exp \left( - \int_t^T r(s)ds \right).$$

(14)

4.2 CIR model

The general equilibrium approach developed by Cox, Ingersoll and Ross(1985) led to the introduction of "square root" term in the diffusion coefficient of the instantaneous short-rate dynamics proposed by Vasicek (1977). The resulting model has been a benchmark for many years because of its analytical tractability and the fact that, contrary to the Vasicek (1977) model, the instantaneous short rate is always positive. The square root process of spot rate is defined by the following stochastic differential equation

$$dr(t) = (\alpha - \beta r(t))dt + \sigma \sqrt{r(t)}dB(t)$$

(15)

where $\{B(t), t \geq 0\}$ is a standard Brownian motion and $\alpha, \beta, \sigma \geq 0$ are parameters to be estimated.

4.3 Statistical modeling

The estimation of the continuous time model (15) from discretely observed data will be encountered with difficulty since a computationally useful expression of the likelihood is difficult to obtain. Thus, it is common to use the discretization of the continuous time model. Under the Euler-Maruyama approximation, the model (15) is discretized as

$$r_{t+\Delta} = r_t + (\alpha - \beta r_t)\Delta + \sigma \sqrt{\Delta} \epsilon_t,$$

(16)

where $\Delta$ is the time interval and $\epsilon_t \sim \text{Normal}(0, 1)$. Some difficulty occurs in the use of the discretization model when the magnitude of the observation interval $\Delta$ is large. We overcome this difficulty by using the Bayesian approach proposed by Feng and Xie (2012) by introducing augmented data between each pair of observations.

4.4 Bayesian CIR model

To derive the full conditional posterior distribution, we divide the parameters $(\alpha, \beta, \sigma)$ into $(\Psi, \sigma^2)$ where $\Psi = (\alpha, \beta)^T$. Suppose that $T$ observations and $M$ augmented data between each pair of observations are available. Let $Y = (r_1, r_2, ..., r_T)$ denote all observations and $Y^* = (r_1^*, r_2^*, ..., r_{T-1}^*)$ augmented data,
where $r_t^* = (r_{t,0}^*, r_{t,1}^*, ..., r_{t,M}^*)$ and $r_{t,0}^* = r_t$. For each $t \geq 0$, we define $\Delta^+ = \frac{A}{M+1}$ and assume that $r_{t,j}^*$ is a Markov process for $j = 0, 1, ..., M$. For $\Psi$ and $\sigma^2$, we have the likelihood

$$f(Y, Y^*|\Psi, \sigma^2) = \prod_{t=1}^{T-1} \prod_{j=0}^{M} f(r_{t,j+1}|r_{t,j}^*, \Psi, \sigma^2)$$

$$\propto \exp \left\{ \sum_{t=1}^{T-1} \sum_{j=0}^{M} \frac{-(r_{t,j+1} - [r_{t,j}^* + (\alpha - \beta r_{t,j}^*)\Delta^+])^2}{2\sigma^2\Delta^+ r_{t,j}^*} \right\}$$

and this shows that $f(Y, Y^*|\Psi, \sigma^2)$ is a bivariate normal distribution. Therefore, we gain the likelihood function of $\Psi$ and $\sigma^2$:

$$f(Y, Y^*|\Psi, \sigma^2) \propto |\Lambda_\Psi|^{-1/2} \exp \left\{ -\frac{1}{2}(\Psi - \mu_\Psi)^T \Lambda_\Psi^{-1}(\Psi - \mu_\Psi) \right\}, \tag{17}$$

where

$$\mu_\Psi = \left( \frac{a_{12}C - a_{12}D}{a_{11}a_{22} - a_{12}^2}, \frac{-a_{12}C + a_{11}D}{a_{11}a_{22} - a_{12}^2} \right)^T$$

$$\Lambda_\Psi = \begin{pmatrix} \frac{\Delta^+}{\sigma^2} A & -\frac{\Delta^+}{\sigma^2} (T-1)(M+1) \\ -\Delta^+ \frac{T-1}{T} B & \frac{\Delta^+}{\sigma^2} \end{pmatrix}$$

and

$$A = \sum_{t=1}^{T-1} \sum_{j=0}^{M} \frac{1}{\epsilon_{t,j}^*}, \quad B = \sum_{t=1}^{T-1} \sum_{j=0}^{M} r_{t,j}^*,$$

$$C = -\sum_{t=1}^{T-1} \sum_{j=0}^{M} \frac{r_{t,j}^* - r_{t,j+1}^*}{\epsilon_{t,j}^*}, \quad D = \sum_{t=1}^{T-1} \sum_{j=0}^{M} (r_{t,j}^* - r_{t,j+1}^*),$$

$$a_{11} = \frac{\Delta^+}{\sigma^2} A, \quad a_{22} = \frac{\Delta^+}{\sigma^2} B, \quad a_{12} = -\frac{\Delta^+}{\sigma^2} (T-1)(M+1).$$

In Bayesian analysis, the prior distributions of parameters are required to set up as we discussed in the previous section. To make sure that $\alpha, \beta \geq 0$, the prior distribution of $\Psi$ is given as

$$\Psi \sim T trunc\{N(0, \infty), (\mu, \Sigma)^{-1} \}.$$ 

Similarly, the prior distribution of $\sigma^2$ is chosen as

$$\sigma^2 \sim T inv\{\mu, \nu\}.$$ 

### 4.5 Forecasting

Given the MCMC samples of each parameter $\{\alpha^{(i)}, \beta^{(i)}\}$ and $\sigma^{(i)}, i = 1 \cdots N$ and initial value $r_0$, the future interest rate $\{r_1^{(i)}, r_2^{(i)}, \cdots, r_T^{(i)}\}, i = 1 \cdots N$ are recursively calculated as

$$r_1^{(i)} = r_0 + (\alpha^{(i)} - \beta^{(i)} r_0) \Delta + \sigma^{(i)} \sqrt{\Delta} \sqrt{\epsilon_1^{(i)}}$$

$$r_2^{(i)} = r_1^{(i)} + (\alpha^{(i)} - \beta^{(i)} r_1^{(i)}) \Delta + \sigma^{(i)} \sqrt{\Delta} \sqrt{r_1^{(i)} \epsilon_2^{(i)}}$$

$$\vdots$$

$$r_T^{(i)} = r_{T-1}^{(i)} + (\alpha^{(i)} - \beta^{(i)} r_{T-1}^{(i)}) \Delta + \sigma^{(i)} \sqrt{\Delta} \sqrt{r_{T-1}^{(i)} \epsilon_T^{(i)}}$$

where $\epsilon_t^{(i)} \sim N(0, 1), i = 1 \cdots N, t = 1 \cdots T$. Repeating the above recursion (18) by using the MCMC sampling, we obtain $N$ simulated $T$ period ahead forecasts.
4.6 MCMC procedure
The steps in the MCMC procedure are as follows:

Step 1: Initialize $r_{1,0}, \Psi, \sigma^2$.

Step 2: Use Data augmentation to generate samplings of $r_{1}^*, r_{2}^*, ..., r_{T}^*_{T-1}$.

Step 3: Use Gibbs sampler to
- update $\Psi$ from $f(\Psi|Y, Y^*, \sigma^2)$,
- update $\sigma^2$ from $f(\sigma^2|\Psi, Y, Y^*)$.

Step 4: Update $r_{1}^*, r_{2}^*, ..., r_{T}^*_{T-1}$ from $f(r_t^*|\Psi, \sigma^2)$.

Step 5: Repeat Step 2, Step 3 and Step 4 until the sampling size $N$ is reached.

4.7 Bayesian estimation
We use the interest rates of Japanese Government Bonds with 1 year maturity. The data consists of monthly interest rate from January, 1998 through December, 2008. Figure 11 shows the term structure of Japanese Government Bonds with 1, 2, 3, 4 and 5 year maturities. We estimated Bayesian CIR model to the interest rate with 1 year maturity using MCMC method. We run 15,000 steps of MCMC samplings on the condition that $M = 30$ and $\Delta = 1/250$. We then discarded the first 5,000 steps and used the remaining 10,000 steps. The hyper parameters are set up as in Table 3 and $\hat{\alpha}$ and $\hat{\beta}$ are the maximum-likelihood estimators for each parameter.

<table>
<thead>
<tr>
<th>hyperparameter</th>
<th>specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>$\hat{\alpha}$</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>$\hat{\beta}$</td>
</tr>
<tr>
<td>A diagonal element of $\Sigma$</td>
<td>10</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.1</td>
</tr>
<tr>
<td>$\nu$</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table 3: Hyperparameters of the Bayesian CIR model

The MCMC results for $\alpha, \beta$ and $\sigma^2$ are summarized in Table 4, giving the posterior mean, the posterior SD and the 95% highest posterior density (HPD) intervals for each of $\alpha, \beta$, and $\sigma^2$, respectively. The sixth column of Table 4 lists the $p$ values of the Geweke convergence, which all fairly large and seem to support the null hypothesis of convergence. Figure 12 displays the time series plots of MCMC sequences of $\alpha, \beta$ and $\sigma^2$. It may confirm stable convergences of each sequence. The posterior distributions for $\alpha, \beta$ and $\sigma^2$ depicted in Figure 13 all look fairly normal.

Figure 12 illustrates the past pattern and the future predictions of 1-year-maturity interest rate. In CIR model, the long term mean of interest rate is modeled as $\alpha/\beta$ and the estimated value is turned out to be 0.2355%, which coincides with that of historical data. Since the estimates satisfy the non-negative condition $2\alpha > \sigma^2$, the forecast 95% intervals cannot be negative value.
5 Pricing of Longevity Derivatives

5.1 Risk-neutral distribution

We need the risk-neutral probability measure $Q$ to obtain the theoretical Price (1). A change of the real-world probability measure $P$ is commonly used to obtain such a risk-neutral probability. One of the most popular approach is to use the Wang transform proposed by Wang (2000), which transforms the original distribution $P$ of the security into its distortion distribution. In this article, to be consistent with

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Mean</th>
<th>SD</th>
<th>95% HPD</th>
<th>Geweke</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 year</td>
<td>$\alpha$</td>
<td>0.6744</td>
<td>0.299</td>
<td>(0.0777,1.2363)</td>
</tr>
<tr>
<td></td>
<td>$\beta$</td>
<td>2.8635</td>
<td>0.3274</td>
<td>(2.2318,3.504)</td>
</tr>
<tr>
<td></td>
<td>$\sigma^2$</td>
<td>0.0004</td>
<td>1.181x10^{-5}</td>
<td>(0.0004,0.0004)</td>
</tr>
</tbody>
</table>

Table 4: Summary statistics of MCMC samples
It is well known that the solution to this constrained minimization problem is given as

Then, the maximum entropy principle stipulates that the risk-neutral distribution

\[ S \]

with respect to

where

Let \( \pi \) denote the empirical distribution of the \( N \) paths of the MCMC sampling, each of which distributed as the process \( \{(p_x, r_t), t = 1, 2, \ldots, T\} \), and denote them by

\[ \left\{(p_x^{(1)}, r_t^{(1)}), (p_x^{(2)}, r_t^{(2)}), \ldots, (p_x^{(N)}, r_t^{(N)}), i = 1, 2, \ldots, N\right\} \]

Let \( \pi \) denote the empirical distribution of the \( N \) paths of the MCMC sampling which puts an equal mass of \( 1/N \) on each path. Now assume that for \( t \), the present market value \( p_x^\text{market} \) of the longevity index \( S_x(t) = p_x \) is available. Then, We convert \( \pi \) into the risk-neutral version \( \pi^* \) by imposing

\[ \sum_{i=1}^{N} \exp \left( - \sum_{u=1}^{t} r_u^{(i)} \right) p_x^{(i)} \pi^*_i = p_x^\text{market}, \text{ for } t = 1, \ldots, T. \]

Then, the maximum entropy principle stipulates that the risk-neutral distribution \( \pi^* = \{\pi_i\}_{i=1}^{N} \) should minimize the Kullback-Leibler information divergence

\[ \sum_{i=1}^{N} \pi^*_i \log \left( \frac{\pi^*_i}{\pi_i} \right) \]

with

\[ \pi^*_i > 0, \text{ for } i = 1, \ldots, N, \sum_{i=1}^{N} \pi^*_i = 1. \]

It is well known that the solution to this constrained minimization problem is given as

\[ \hat{\pi}_i = \frac{\pi_i \exp \left\{ \sum_{t=1}^{T} \exp \left( - \sum_{u=1}^{t} r_u^{(i)} \right) p_x^{(i)} \lambda_t \right\}}{\sum_{i=1}^{N} \pi_i \exp \left\{ \sum_{t=1}^{T} \exp \left( - \sum_{u=1}^{t} r_u^{(i)} \right) p_x^{(i)} \lambda_t \right\}} \text{ for } i = 1, \ldots, N, \]

where \( \lambda_t \)'s are Lagrange multipliers. Putting \( \pi_j = 1/N (j = 1, 2, \ldots, N) \), we have,

\[ p_x^\text{market} = \frac{\sum_{i=1}^{N} \exp \left\{ \sum_{t=1}^{T} \exp \left( - \sum_{u=1}^{t} r_u^{(i)} \right) p_x^{(i)} \lambda_t \right\}}{\sum_{i=1}^{N} \exp \left\{ \sum_{t=1}^{T} \exp \left( - \sum_{u=1}^{t} r_u^{(i)} \right) p_x^{(i)} \lambda_t \right\}} \text{ for } t = 1, \ldots, T, \]

and the Lagrange multipliers are found by minimizing

\[ \sum_{i=1}^{N} \exp \left\{ \sum_{t=1}^{T} \lambda_t \left\{ \exp \left( - \sum_{u=1}^{t} r_u^{(i)} \right) p_x^{(i)} - p_x^\text{market} \right\} \right\}. \]

with respect to \( \lambda_t \)'s. Finally, the derivative whose payments are \( S_x(t) \) at \( t \) is evaluated under the risk-neutral distribution as

\[ P_0(x, T) = \mathbb{E}^Q \left[ \sum_{t=1}^{T} \exp \left( - \sum_{u=1}^{t} r_u \right) S_x(t) \right| \mathcal{F}_0 \approx \sum_{i=1}^{N} \sum_{t=1}^{T} \exp \left( - \sum_{u=1}^{t} r_u^{(i)} \right) S_x(t)^{(i)} \pi^*_i. \]
5.3 Applications to Japanese data

We priced a longevity bond as well as a longevity cap whose payments are linked with Japanese male survival rate by using the entropy maximization principle on the basis of our empirical results illustrated in the previous sections. We computed the present value of longevity derivatives under the following conditions:

- maturity: 25 years
- age of the cohort: $x = 65$
- survival index $S_x(t)$: the survival probability projected by two-factor Lee-Carter model
- risk-free interest rate $r_t$: the risk free rate forecasted by CIR model.

The pricing procedure is set up as follows:

1. Forecast $t_p^x = S_x(t)$ and $r_t$ by employing the Bayesian models.
2. Risk-neutralize them under the entropy maximization principle.
3. Take the expectation of the payments of a longevity derivative with respect to the risk neutral probability.

To determine the market price $t_p^x_{\text{market}}$, we used death probabilities for annuity products tabulated in "the standard life table 2007", which is constructed by the Institute of Actuaries of Japan because most Japanese life insurance companies price their products using this table.

5.3.1 Longevity bond

We considered a longevity bond with payments

$$C(S_{65}(t)) = S_{65}(t), \ t = 1, 2, \ldots, 25$$

and derived the predictive distributions of the present value of the longevity bond for the risk-neutral and the physical (not risk-adjusted) versions. We compared the result for the CIR model with those for the cases in which the interest rate is fixed at 1%, 2% and 3% for both one-factor and two-factor Lee-Carter models. The results are shown in Figure 14 (one-factor) and Figure 15 (two-factor). They showed that for all the cases the risk-neutral distribution (solid line) is located on the right side of the physical one (dashed line), properly reflecting the risk adjustment for the mortality and the interest rate. The basic statistics of the present value of longevity bond are summarized in Table 5 (one-factor) and Table 6 (two-factor).

The mean of the risk-neutral predictive distribution is the theoretical price $P_{2008}(x = 65, T = 25)$ of the longevity bond. We found that the longevity bond under the CIR model is priced lowest with the two-factor Lee-Carter model and is priced second lowest with the one-factor Lee-Carter model. Note also that, as indicated by the values of the skewness, the risk-neutral distribution of the CIR model is more skewed to the left than those of the fixed interest rate cases both for one-factor and two-factor Lee-Carter models. This seems a direct result of the incorporation of the interest rate risk.
Table 5: Summary statistics of the present value of the longevity bond (one-factor)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean</th>
<th>SD</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIR</td>
<td>18.495</td>
<td>0.212</td>
<td>-1.561</td>
<td>3.846</td>
</tr>
<tr>
<td></td>
<td>17.595</td>
<td>0.366</td>
<td>-0.076</td>
<td>3.157</td>
</tr>
<tr>
<td>1%</td>
<td>18.841</td>
<td>0.274</td>
<td>-0.731</td>
<td>2.982</td>
</tr>
<tr>
<td></td>
<td>17.898</td>
<td>0.346</td>
<td>-0.078</td>
<td>3.245</td>
</tr>
<tr>
<td>2%</td>
<td>18.707</td>
<td>0.246</td>
<td>-1.125</td>
<td>4.773</td>
</tr>
<tr>
<td></td>
<td>17.668</td>
<td>0.406</td>
<td>-0.010</td>
<td>2.949</td>
</tr>
<tr>
<td>3%</td>
<td>18.262</td>
<td>0.410</td>
<td>-0.600</td>
<td>2.606</td>
</tr>
<tr>
<td></td>
<td>17.443</td>
<td>0.489</td>
<td>0.032</td>
<td>2.639</td>
</tr>
</tbody>
</table>

Table 6: Summary statistics of the present value of the longevity bond (two-factor)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean</th>
<th>SD</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIR</td>
<td>18.377</td>
<td>0.332</td>
<td>-1.865</td>
<td>2.610</td>
</tr>
<tr>
<td></td>
<td>17.578</td>
<td>0.350</td>
<td>-0.114</td>
<td>3.130</td>
</tr>
<tr>
<td>1%</td>
<td>19.141</td>
<td>0.294</td>
<td>-1.205</td>
<td>3.387</td>
</tr>
<tr>
<td></td>
<td>17.880</td>
<td>0.332</td>
<td>-0.116</td>
<td>3.163</td>
</tr>
<tr>
<td>2%</td>
<td>19.102</td>
<td>0.328</td>
<td>-0.154</td>
<td>5.245</td>
</tr>
<tr>
<td></td>
<td>17.650</td>
<td>0.395</td>
<td>-0.054</td>
<td>2.904</td>
</tr>
<tr>
<td>3%</td>
<td>18.460</td>
<td>0.610</td>
<td>-0.213</td>
<td>1.802</td>
</tr>
<tr>
<td></td>
<td>17.425</td>
<td>0.481</td>
<td>-0.008</td>
<td>2.602</td>
</tr>
</tbody>
</table>

Figure 14: Distributions of the present value of the longevity bond (one-factor)
Figure 15: Distributions of the present value of the longevity bond (two-factor)
5.3.2 Longevity cap

We also considered a longevity cap with payments

\[ C(S_{65}(t)) = \max(S_{65}(t) - K_t, 0), \quad t = 1, 2, \ldots, 25 \]

and derived the predictive distributions of the present value of the longevity cap for the risk-neutral and the physical (not risk-adjusted) versions. Here we set the cap rate \( K_t \) to the observed survival rate \( \hat{\nu}_{65}^{\text{ref}}(2008) \) in year 2008. We compared the result for the CIR model with those for the cases in which the interest rate is fixed at 1%, 2% and 3% for both one-factor and two-factor Lee-Carter models. The results are shown in Figure 16 (one-factor) and 17 (two-factor). As in the case of the longevity bond, the risk-neutral distribution (solid line) is located on the right side of the physical one (dashed line) for all the cases.

The basic statistics of the present value of the longevity cap are summarized in Table 7 (one-factor) and Table 8 (two-factor). The mean of the risk-neutral predictive distribution is the theoretical price \( P_{2008}(x = 65, T = 25) \) of the longevity cap. We found that the longevity cap is priced lowest under the CIR model.

![Figure 16: Distribution of the present value of the longevity cap (one-factor)](image)
Figure 17: Distributions of the present value of the longevity cap (two-factor)

Table 7: Summary statistics of the present value of the longevity cap (one-factor)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Risk Neutral</th>
<th>Physical</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIR</td>
<td>1.528</td>
<td>0.556</td>
</tr>
<tr>
<td>1%</td>
<td>2.025</td>
<td>0.915</td>
</tr>
<tr>
<td>2%</td>
<td>1.960</td>
<td>0.904</td>
</tr>
<tr>
<td>3%</td>
<td>1.750</td>
<td>0.892</td>
</tr>
<tr>
<td>4%</td>
<td>0.911</td>
<td>0.579</td>
</tr>
</tbody>
</table>

Table 8: Summary statistics of the present value of the longevity cap (two-factor)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Risk Neutral</th>
<th>Physical</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIR</td>
<td>1.220</td>
<td>0.542</td>
</tr>
<tr>
<td>1%</td>
<td>1.938</td>
<td>0.899</td>
</tr>
<tr>
<td>2%</td>
<td>1.689</td>
<td>0.887</td>
</tr>
<tr>
<td>3%</td>
<td>1.583</td>
<td>0.876</td>
</tr>
<tr>
<td>4%</td>
<td>0.746</td>
<td>0.569</td>
</tr>
</tbody>
</table>
6 Conclusions

We examined the pricing of longevity derivatives with mortality and interest rate risks in a Bayesian framework in this paper. We estimated the Bayesian two-factor Lee-Carter model and the Bayesian CIR model to obtain the predictive distributions of future mortality and interest rates by employing the MCMC method. We then derived the physical and risk-neutral predictive distributions of the present values of the longevity bond and cap under the maximum entropy principle.

We compared the results for the CIR model with those for the cases in which the interest rate is fixed at 1%, 2% and 3% for both one-factor and two-factor Lee-Carter models. Our results showed that for all the cases the risk-neutral distribution is located on the right side of the physical one. We found that the longevity bond under the CIR model is priced lowest for the one-factor Lee-Carter model and second lowest for the two-factor Lee-Carter model. We also found that the longevity cap is priced lowest under the CIR model both for one-factor and two-factor Lee-Carter models.

One implication of these findings is that the longevity derivative may be overpriced considerably when the interest rate is not modeled to change and is kept at a low level, which suggests the importance of incorporating the interest rate risk into the pricing especially for countries facing low interest rates.

These results, however, are limited in some respects. For one thing, the CIR model used in this paper is a one-factor model as it describes movements in interest rates for all maturities as driven by only a single stochastic factor and thus may not fully capture the dynamics of the term structure. We also implicitly assume that mortality rates and interest rates are independent, which is not entirely justifiable as argued in Granados (2008) among others.

References


