Communications for Statistical Applications and Methods
2022, Vol. 29, No. 6, 679–694

https://doi.org/10.29220/CSAM.2022.29.6.679
Print ISSN 2287-7843 / Online ISSN 2383-4757

Sustainability of pensions in Asian countries

Hyunoo Shim[a], Siok Kim[b], Yang Ho Choi[1,a]

[a]Department of Actuarial Science, Hanyang University, Korea;
[b]Department of Finance and Insurance, Hanyang University, Korea

Abstract

Mortality risk is a significant threat to individual life, and quantifying the risk is necessary for making a national population plan and is a traditionally fundamental task in the insurance and annuity businesses. Like other advanced countries, the sustainability of life pensions and the management of longevity risks are becoming important in Asian countries entering the era of aging society. In this study, mortality and pension value sustainability trends are compared and analyzed based on national population and mortality data, focusing on four Asian countries from 1990 to 2017. The result of analyzing the robustness and accuracy of generalized linear/nonlinear models reveals that the Cairns-Blake-Dowd model, the nonparametric Renshaw-Haberman model, and the Plat model show low stability. The Currie, CBD M5, M7, and M8 models have high stability against data periods. The M7 and M8 models demonstrate high accuracy. The longevity risk is found to be high in the order of Taiwan, Hong Kong, Korea, and Japan, which is in general inversely related to the population size.

Keywords: life annuity, stochastic mortality models, cross-country comparison, Lee-Carter, CBD

Improving mortality trends is a global phenomenon including Korea (Kim, 2012). Several countries have already entered the aging stage, and the longevity risk owing to the increasing life expectancy is on the rise. Longevity risk ocuurs when longevity increases owing to a decrease in mortality.

Excessive longevity risk negatively affects the sustainability of insurers in multiple coverage areas. For example, an well-known insurance that contains longevity risk is private and public pensions. Tian and Zhao (2016); Le *et al.* (2021) analyzed the financial sustainability and risk of basic public pensions via stochastic mortality focusing on China. Won (2010) found that as volatility in mortality trends increases, the risk that the funding ratio of public pensions will fall below a certain level increases. Kim (2013) quantitatively measured the longevity risk using the shock factor, VaR, and stress trend method.

Prediction models for such mortality have been used in diverse fields. Stefani and Kwon (2021) models the longevity risk in a pension by a multi-state model, and Kim (2006) forecasted mortality rates and predicted the population of Korea using the Lee-Carter model. Chung (2007) used the Lee-Carter model to predict long-term medical expenses of the national health insurance. Kang *et al.* (2015) presented a premium rating model for long-term care insurance suited for in an aging society. Specifically, stochastic models for mortality are required for insurers, as they must retain solvency capital against longevity risk. Korean-Insurance Capital Standards (K-ICS) require an insurer to have sufficient economic capital not to go bankrupt with a 99.5% confidence level for one year. In either

[1] Corresponding author: Department of Actuarial Science, Hanyang University-ERICA Campus, 55 Hanyangdaehak-ro, Sangnok-gu, Ansan, Gyeonggi-do 15588, Korea. E-mail: ychoi@hanyang.ac.kr

Published 30 November 2022 / journal homepage: http://csam.or.kr

Solvency II or K-ICS, longevity risk measurement is divided into a separate submodule and probabilistic mortality models is the most promising methodology to measure the risk. For example, Jho (2020) measured longevity risk by considering stochastic interest rate and stochastic mortality models.

Among a variety of mortality models, we considered the following models: Lee and Carter (1992) introduced to the actuarial discipline a stochastic mortality model characterizing mortality trends differed by ages. Renshaw and Haberman (2003) analyzed improved mortality model, which added a cohort effect to the aforementioned Lee-Carter model. Cairns *et al.* (2009) approached modeling mortality by introducing a trend term that is linearly dependent on ages and developed a model, called Cairns-Blake-Dowd model.

It is necessary to perform a comparative analysis of stochastic mortality models to select the most precise model. Previous studies, for example Kim (2012); Bozikas and Pitselis (2018), have comparatively analyzed these stochastic mortality models. Le and Kwon (2021) analyzed the suitability of stochastic mortality models limited to Korean mortality data. Hwang *et al.* (2018) conducted a cross-country comparative study on the trend of male mortality using the Lee-Carter model and revealed that the speed of improvement in male mortality rates was the highest in Korea against other countries considered in the study. However, this study has the following limitations: first, it is based on the single stochastic model, which cannot explain the cohort effect. Second, variations in one-year age-specific trends are smoothed out, since the mortality data are grouped into 5-year age groups.

We recognized these limitations; hence, this study attempted to find more detailed similarities and differences through comparative analysis among East-Asian countries for the most well-known models of stochastic mortality, all of which belong to the framework of generalized linear/nonlinear models. This paper is organized as follows: Section 1 explains all stochastic mortality models covered in this study. Section 2 fit the models and compares characteristics and parameters between countries for each mortality model. Section 3 compares mortality models' predictive power. Section 4 analyzes sustainability of life annuities based on the mortality models. Section 5 summarizes the results of this study and suggests future research directions.

1. Stochastic mortality models

The Lee-Carter model with the age-period factors constructed by Lee and Carter (1992) and its modified model, the Lee-Miller model (Lee and Miller, 2001), were pioneering works that constructed a statistical model for mortality rates. In this section, we discuss such models and their extended models in a generalized age-period-cohort framework.

Let $E^0_{x,t}$ be the initial population at age x exposed to the risk of dying in year t, $E^c_{x,t}$ be a central population at age x exposed to the risk of dying in year t, $D_{x,t}$ be the random variable representing the number of deaths among people at age x in year t, $\mu_{x,Lett}$ be the force of mortality, $d_{x,t}$ be the number of observed deaths, $\hat{q}_{x,t}$ be the estimate of mortality rate for the initial exposures, $\hat{m}_{x,t}$ be the estimate of the central mortality rate for central exposures. In this case, $\hat{q}_{x,t} = d_{x,t}/E^0_{x,t}$ and $\hat{m}_{x,t} = d_{x,t}/E^c_{x,t}$. The distribution of error is determined according to $D_{x,t}$. Currie (2016) argued that many common models of mortality can be fit into generalized linear / non-linear models. Likewise, we build our model upon the generalized age-period cohort (GAPC) stochastic mortality model framework.

Among the probability distributions of $D_{x,t}$, the most easily considerable distribution is the Poisson or binomial distributions. Brouhns *et al.* (2005) presents fitting methods for the Lee-Carter model assuming a Poisson distribution for the number of deaths. Previous studies have mostly dealt with the Poisson or binomial distribution. For example, Coffie (2015) compares the Poisson and the binomial

Table 1: Structure of stochastic mortality models in the generalized age-period-cohort framework

Model	Equation for $\eta_{x,t}$	Included effect
LC	$\alpha_x + \beta_x^{(1)}\kappa_t^{(1)}$	age, period
RH	$\alpha_x + \beta_x^{(1)}\kappa_t^{(1)} + \gamma_{t-x}$	age, period, cohort
Currie	$\alpha_x + \kappa_t^{(1)} + \gamma_{t-x}$	age, period, cohort
CBD M5	$\kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)}$	period (2)
CBD M6	$\kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + \gamma_{t-x}$	period (2), cohort
CBD M7	$\kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + \left((x - \bar{x})^2 - \hat{\sigma}_x^2\right)\kappa_t^{(3)} + \gamma_{t-x}$	period (3), cohort
CBD M8	$\kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + (x_c - x)\gamma_{t-x}$	period (2), cohort
Plat	$\alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + (\bar{x} - x)^+\kappa_t^{(3)} + \beta_x^{(0)}\gamma_{t-x}$	age, period (3), cohort

LC is the Lee and Carter model in 1992 (Lee and Carter, 1992); RH is the Renshaw and Haberman model in 2006 (Renshaw and Haberman, 2006); Currie is the Currie model in 2006 (Currie, 2006); CBD M5 is the Cairns-Blake-Dowd M5 model in 2006 (Cairns *et al.*, 2006); CBD M6 is the Cairns-Blake-Dowd M6 model in 2009 (Cairns *et al.*, 2009); CBD M7 is the Cairns-Blake-Dowd M7 model in 2009 (Cairns *et al.*, 2009); CBD M8 is the Cairns-Blake-Dowd M8 model in 2009 (Cairns *et al.*, 2009); Plat is the Plat model in 2009 (Plat, 2009).
$\eta_{x,t}$ is defined in Equation (1.3), α_x is a static age function, $\kappa_t^{(i)}$ is a period index for the i^{th} mortality trend, $(x - \bar{x})$, $(\bar{x} - x)$, $(\bar{x} - x)^+$, $(x_c - x)$, $\{(x - \bar{x})^2 - \hat{\sigma}^2\}$ are parametric functions of age controling the trend effect across ages. γ_{t-x} is a cohort effect term with $\beta_x^{(0)}$ modulating it over ages.
The number in parentheses means the number of different period terms included in the regression equation.

distribution's settings for the Lee-Carter model. For the Poisson distribution of $D_{x,t}$,

$$D_{x,t} \sim \text{Poisson}\left(E_{x,t}^c \mu_{x,t}\right) \tag{1.1}$$

with $\mathbb{E}(D_{x,t}/E_{x,t}) = \mu_{x,t}$. For the binomial distribution, it is

$$D_{x,t} \sim \text{Binomial}\left(E_{x,t}^0, q_{x,t}\right) \tag{1.2}$$

with $\mathbb{E}(D_{x,t}/E_{x,t}) = q_{x,t}$. Generally, Neves *et al.* (2017) extended the Lee-Carter model with five different probability models by utilizing generalized autoregressive score models. Hence, the number of deaths among exposures, that is, the death rate, can be predicted by the predictor $\eta_{x,t}$ through the link function $g(\cdot)$ as follows:

$$g\left(\mathbb{E}\left(\frac{D_{x,t}}{E_{x,t}}\right)\right) = \eta_{x,t}, \tag{1.3}$$

where x is age, t is year, $c = t - x$ is year-of-birth (cohort), g is a link function. The $\eta_{x,t}$ function characterizes the stochastic mortality model. α_x is an age function or factor, describing mortality by age. Table 1 summarizes the function and effect of all the models considered in this study.

The κ_t^i and γ_{t-x} follow stochastic processes, which we will discuss in Section 2.2. Even though a link function, $g(\cdot)$ can be arbitrarily chosen, the canonical link function for the Poisson distribution is the log function. Thus, we choose this function for $g(\cdot)$. The above parameters are not completely determined unless some conditions constrain them. Hunt and Blake (2015) carefully discussed the identifiability issues of general mortality models.

2. Robustness of models

2.1. Data

The study is based on the total population and the number of deaths of selected Asian countries including Korea, Japan, Hong Kong, Taiwan from the Human Mortality Database (HMD) (University of

California, Berkeley and Max Planck Institute for Demographic Research, 2021). The HMD database is divided by year, age, and gender. We selected common central population and deaths among those countries for men and women aged 0 to 100 years-old from 2003 to 2017. It provides detailed information on population and death counts in one-year age intervals. The short period limitation of the sample source limits the maximal fit of the stochastic mortality models.

2.2. Model fit

In this Section, we estimate the parameters of the models. We fit the selection of models by maximizing the following log-likelihood under the assumption of a Poisson distribution:

$$\mathcal{L}\left(d_{x,t}, \hat{d}_{x,t}\right) = \sum_{x} \sum_{t} \omega_{x,t} \left\{ d_{x,t} \log \hat{d}_{x,t} - \hat{d}_{x,t} - \log d_{x,t}! \right\}, \tag{2.1}$$

where $\omega_{x,t}$ are weights whose values are 0 if the observed data of (x, t) are absent, and 1 if the data are present. BIC is defined as follows:

$$\text{BIC}_m = L\left(\hat{\phi}_m\right) - \frac{1}{2}\nu_m \log N, \tag{2.2}$$

where ϕ_m is the parameter set for model m, ν_m is the effective number of parameters (degree of freedom) for model m, $\hat{\phi}_m$ is the maximum likelihood estimate of the parameter set, $L(\hat{\phi}_m)$ is the maximum log-likelihood, and N is the number of observations.

Lee and Carter (1992) fitted their models by singular value decomposition (SVD), and Bell and Monsell (1991) used the principal component analysis, and Girosi and King (2008) suggested a Bayesian approach. Currie (2006) showed that most, if not all, stochastic mortality models are generalized linear or non-linear models. As in Millossovich *et al.* (2018), we estimate the models' parameters by fitting them to generalized non-linear models that have $D_{x,t}/E_{x,t}$ as a response variable and α_x, $\kappa_t^{(1)}$, $(\overline{x} - x)\kappa_t^{(2)}$, $[(x_l - x)^+ + b(x_l - x)^+)^2]\kappa_t^{(3)}$, $\beta_x^{(0)}\gamma_{t-x}$ as explanatory variables. We used the StMoMo package in R (Villegas *et al.*, 2018) to fit the parameters and predict and simulate future mortalities.

2.3. Robustness of models

Understanding how robust the model is against a selection of train data is important for verifying the validity of the model itself and for trusting the predictive power of the model. In Asian countries, the history of collecting mortality data has not been long; therefore, it is necessary to check whether the models' parameters are robust for various sample data.

The robustness of our stochastic mortality models for E&W, UK, and the US can be tested by fitting the models with different train data set. For testing the robustness, we prepared two data set of numbers of deaths and population: the base data group ranging from the year 1990 to 2017 (Set1), and the other test data group from 2003 to 2017 (Set2). For this work, we dropped the data set from Korea, since the mortality data of Korea were not consistently collected before the year 2003. From Figure 1 to Figure 5 we show parameter estimates of the models considered in our study by countries.

β_x is a sensitivity to mortality trends at age x. Figure 1 shows that LC and RH models are not sufficiently robust to changes in data sets, though the age patterns of the two data sets coincide. For example, in the LC model, the patterns of the base β_x and the test β_x decrease when age goes from 0 to 40 years old, increase until 75 years old, and decrease until 100 years old. In contrast, the raw values of the base β_x (large data set) are more stable than that of test β_x (small data set). Cairns *et al.* (2009, 2011); Currie (2016) finds that the model is not robust in selecting data sets. As the RH model

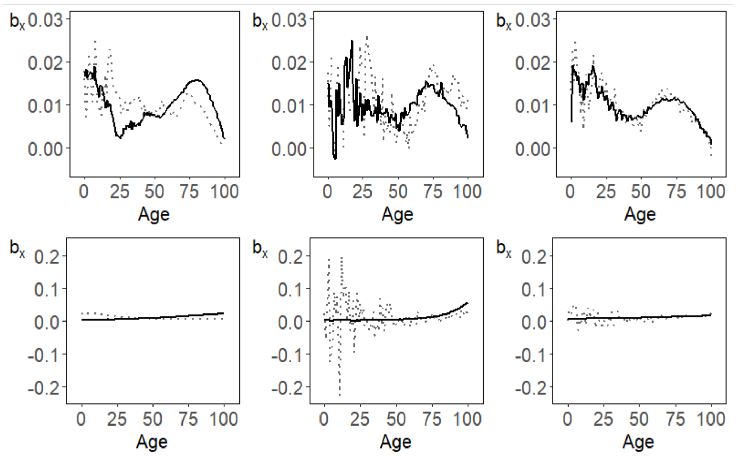

Figure 1: *Estimates of* β_x *by LC (top row), RH model (bottom row) and by countries (Japan, Hong Kong, Taiwan, left-to-right) using the base data set from the year 1990 to 2017 (solid line) or the test data set from the year 2003 to 2017 (dotted line). Each horizontal axis is age (unit: years old).*

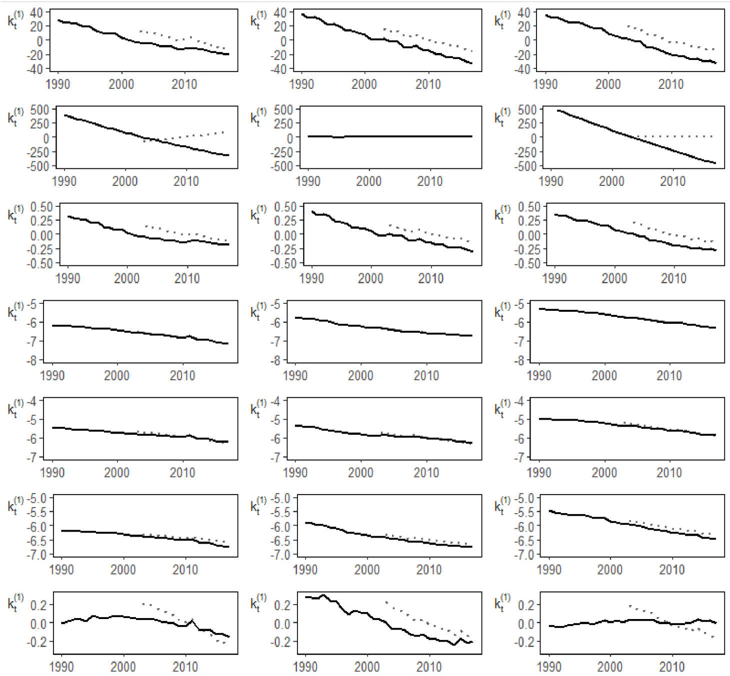

Figure 2: *Estimates of* $\kappa_t^{(1)}$ *by LC, RH, Currie, M5, M7, M8, Plat model (top-to-bottom) and by countries (Japan, Hong Kong, Taiwan, left-to-right) using the base data set from the year 1990 to 2017 (solid line) or the test data set from the year 2003 to 2017 (dotted line). Each horizontal axis represents the year.*

contains more parameters, it can be inferred to be less robust than the LC model. Figure 1 shows that the instability owing to the small size of samples (dotted line in the figure) is more prominent in the

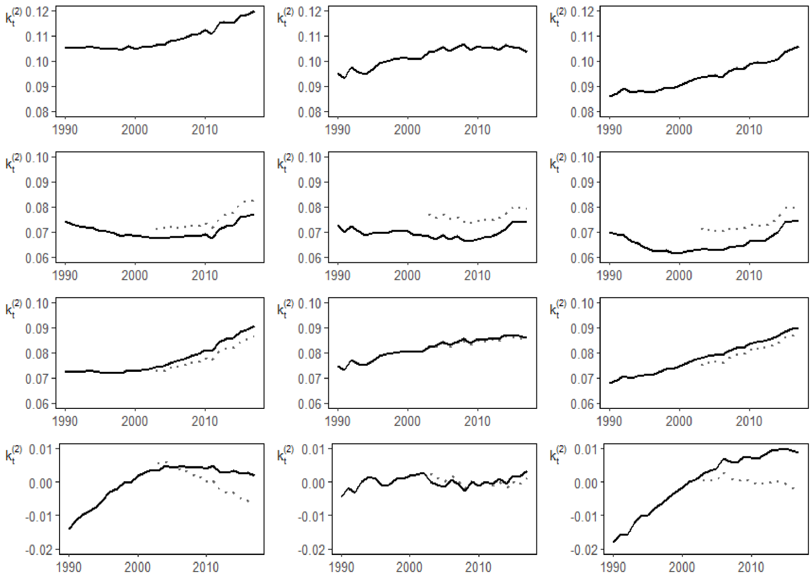

Figure 3: *Estimates of* $\kappa_t^{(2)}$ *by M5, M7, M8, Plat model (top-to-bottom) and by countries (Japan, Hong Kong, Taiwan, left-to-right) using the base data set from the year 1990 to 2017 (solid line) or the test data set from the year 2003 to 2017 (dotted line). Each horizontal axis represents the year.*

RH model, for example, for Hong Kong (the center plot in the bottom). Moreover, the positions of peaks and troughs are a little bit offset between data sets.

$\kappa_t^{(i)}$ is a period index for the i^{th} mortality trend. Figure 2 shows the first mortality trend for each model. There are two things to note in interpreting the results. First, the meaning of $\kappa_t^{(1)}$ differs by models: $\kappa_t^{(1)}$ in the LC and RH model is a sole mortality trend modulated by non-parametric age factors, $\kappa_t^{(1)}$ in the other models have no additional control parameters. Thus, the models are considered more robust than other non-parametric models such as LC and RH. Cairns *et al.* (2009, 2011); Currie (2016) argued that their parametric models with no age-modulating coefficients are more robust than non-parametric models of LC and RH. Second, M5, M7, and M8 models do not have identifiability issues, whereas constraints are imposed to fix the parameters $\kappa_t^{(i)}$ in LC, RH, Currie, and Plat model such that their sums are equal to zero: $\sum_t \kappa_t^{(i)} = 0$. However, they are centered on the yearly midpoint of the data. So the slopes of $\kappa_t^{(i)}$ do matter, and they are needed to be put under scrutiny rather than their actual values.

In Figure 2, we notice that the LC, Currie, M5, M7, and M8 model have almost numerically stable values of $\kappa_t^{(1)}$, whereas the $\kappa_t^{(1)}$ in the RH and Plat model have numerical instabilities: for example, they fail to converge to a significant value (in the RH model for Hong Kong and Taiwan), or Set1 and Set2 have slightly different trends in the Plat model. The data set (Set2) from 2003 to 2017 has 15 years, almost half of Set1. It seems 15 years are not sufficient for robust estimation of the model parameters. Furthermore, the Plat model produces distinguishable patterns between different countries. Furthermore, $\kappa_t^{(1)}$s keep decreasing and the patterns of their yearly changes differ by country.

$\kappa_t^{(2)}$s are shown in Figure 3. The M5, M7, and M8 models are numerically stable even with the small data set (Set2), whereas the Plat model may produce different estimates depending on sample

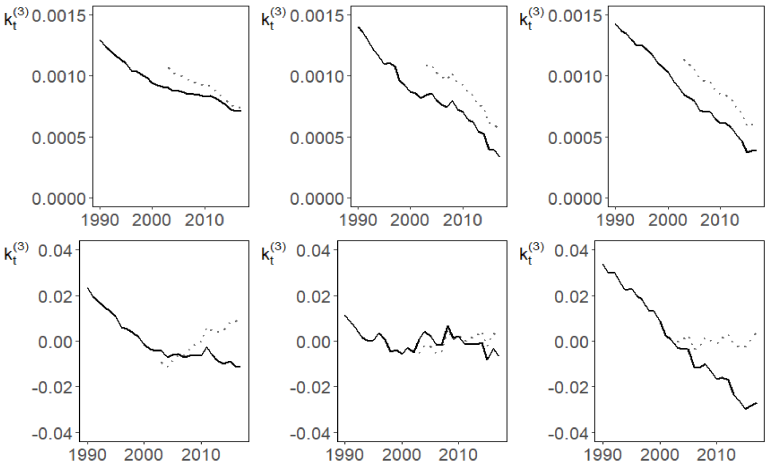

Figure 4: *Estimates of* $\kappa_t^{(3)}$ *by M7 (top row), Plat model (bottom row) and by countries (Japan, Hong Kong, Taiwan, left-to-right) using the base data set from the year 1990 to 2017 (solid line) or the test data set from the year 2003 to 2017 (dotted line). Each horizontal axis represents the year.*

periods or sample data. When it comes to their trends, $\kappa_t^{(2)}$ slowly increases by the year, and the patterns of their yearly changes are almost similar for all models and all countries.

The above properties are still valid for $\kappa_t^{(3)}$ as shown in Figure 4. Hereby, it is necessary to note that $\kappa_t^{(2)}$ represents mortality trends modulated by age difference from mean age, whereas the meaning of $\kappa_t^{(3)}$ differs by models. The quadratic age term affecting mortality trends, $\kappa_t^{(3)}$ in the M7 model decreases by the year and the patterns of their yearly changes are almost similar for all countries. In contrast, the old age term affecting mortality trends, $\kappa_t^{(3)}$ in the Plat model decreases by the year and the patterns of their yearly changes differ by country.

γ_{t-x} is shown in Figure 5. Hunt and Villegas (2015) assesses the impact of the existence of an additional identifiability issue when a model has a cohort effect term, to the robustness of the model, and they improve the robustness and convergence rate of the models. Interpreting the numerical stabilities of γ_{t-x} needs more care: they seem to have numerical stability and no dependence on the data volume for all models except the RH and Plat model. Also, although the γ_{t-x} in the Currie model has similar patterns for all countries, it zitters in some dubious ways at the end of data periods. When it comes to trends, they change in time differently over models, since competing terms differ by models.

Overall, the LC model is simple and robust but not quite exact compared to other models. The RH and Plat model lacks in sufficient stability. Currie, M5, M7, and M8 models are all parametric, they require fewer paramter estimations than other models, and they are numerically more stable. However, they are not superior to other models. It might have resulted from an artifact of parametric properties of the models.

3. Forecast and predictive power

3.1. Forecast

Like other stochastic mortality models, the dynamics of our model are driven by the stochastic processes of period indices κ_t^i and Cohort index γ_{t-x}. A first option is a multivariate ARIMA processes (for example, Lee and Carter (1992), a functional data model in Hyundman and Ullah (2007)), a sec-

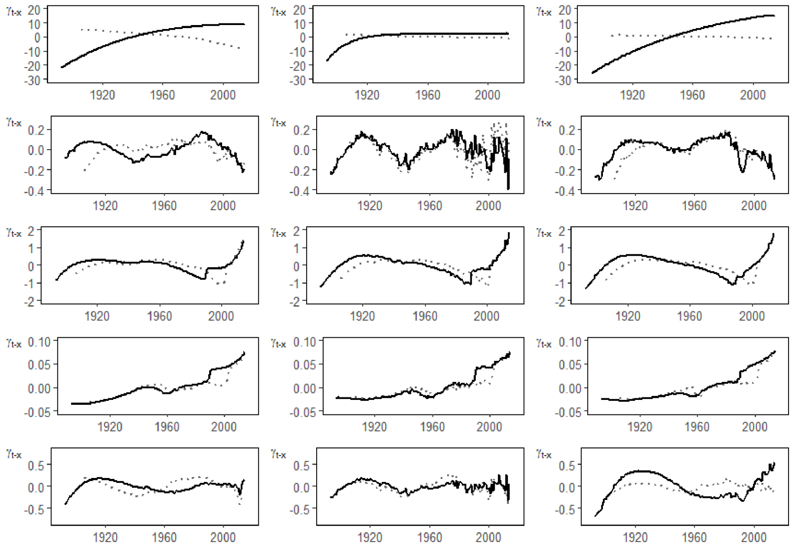

Figure 5: *Estimates of γ_{t-x} by RH, Currie, M7, M8, and Plat model (top-to-bottom) and by countries (Japan, Hong Kong, Taiwan, left-to-right) using the base data set from the year 1990 to 2017 (solid line) or the test data set from the year 2003 to 2017 (dotted line). Each horizontal axis is $t-x$.*

ond option is a set of complicated univariate ARIMA process (for example, an AR-ARCH model in Giacometti *et al.* (2012), exponential smoothing state space model in Hyndman *et al.* (2002)). Feng and Shi (2018) argues that the accuracy of multivariate ARIMA does not outperform the univariate ARIMA processes, and the exponential smoothing state-space model is the best in forecasting accuracy. In our model, we assume that each period index follows a univariate ARIMA(p_i, d_i, q_i) model for each period index using the notation in Millossovich *et al.* (2018):

$$\Delta^{d_i}\kappa_t^{(i)} = \delta_0^{(i)} + \phi_1^{(i)}\Delta^{d_i}\kappa_{t-1}^{(i)} + \cdots + \phi_{p_i}^{(i)}\Delta^{d_i}\kappa_{t-p_i}^{(i)} + \xi_t^{(i)} + \delta_1^{(i)}\xi_{t-1}^{(i)} + \cdots + \delta_{q_i}^{(i)}\xi_{t-q_i}^{(i)}, \tag{3.1}$$

where Δ is the difference operator, $\delta_0^{(i)}$ is the drift term, p_i is the order of the autoregressive model, d_i is the degree of differencing, q_i is the order of the moving-average model, $\phi_1, \ldots, \phi_{p_i}$ are the autoregressive coefficients with $\phi_{p_i} \neq 0$, $\delta_1, \ldots, \delta_{q_i}$ are the moving average coefficients with $\delta_{q_i} \neq 0$. $\xi_t^{(i)}, \ldots, \xi_{t-q_i}^{(i)}$ follow the white noise process with the normal distribution, $\mathcal{N}(0, (\sigma_\xi^{(i)})^2)$.

We assume that the cohort index is independent of the dynamics of other period indices, and it also follows a univariate ARIMA process as commonly adopted in previous research (Cairns *et al.*, 2011).

$$\Delta^d\gamma_c = \delta_0 + \phi_1\Delta^d\kappa_{c-1} + \cdots + \phi_p\Delta^d\kappa_{c-p} + \xi_c + \delta_1\xi_{c-1} + \cdots + \delta_q\xi_{c-q}, \tag{3.2}$$

where $c = t-x$ and $\xi_t, \ldots, \xi_{c-q}$ follow the white noise process with the normal distribution, $\mathcal{N}(0, (\sigma_\xi)^2)$. All other notations are similar to the above.

The error in (3.2) comes only from the error while forecasting period indices. There exist, however, uncertainties in estimating the parameters of mortality models. We estimate the parameter uncertainty by the bootstrap method. Specifically, we utilized the bootstrap algorithm already implemented in the package StMoMo (Millossovich *et al.*, 2018), wherein the semi-parametric bootstrap method

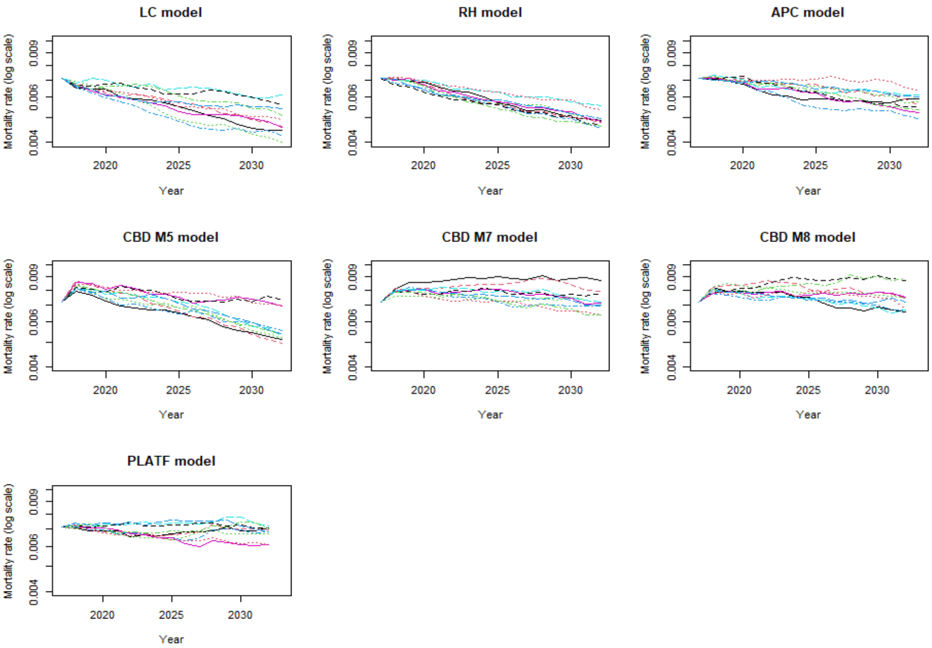

Figure 6: *Mortality rate $q_{x,t}$ of Japan at aged 70 by LC, RH, Currie, M5, M7, M8, Plat model (top-left to bottom-right) using the base data set from the year 2003 to 2017 to 2017. Each horizontal axis represents the year.*

Table 2: BIC of stochastic mortality models in the generalized age-period-cohort framework

Model	Korea	Japan	Hong Kong	Taiwan	Mean rank
LC	7782 (6)	8850 (4)	6063 (7)	6531 (4)	5.25 (6)
RH	7081 (5)	8247 (3)	5991 (5)	6647 (6)	4.75 (5)
Currie (APC)	7055 (4)	8144 (2)	5828 (3)	6457 (3)	3.00 (3)
CBD M5	15857 (7)	54825 (7)	6033 (6)	6957 (7)	6.75 (7)
CBD M7	6817 (2)	9295 (5)	5729 (2)	6412 (1)	2.50 (1)
CBD M8	6807 (1)	10968 (6)	5646 (1)	6430 (2)	2.50 (1)
Plat	6996 (3)	8066 (1)	5929 (4)	6561 (5)	3.25 (4)

The numbers in parentheses indicate the ranking of each model by country.
A mean rank means the average of rankings over countries.

was originally proposed by Brouhns *et al.* (2005), and the residual bootstrap method was discussed by Koissi *et al.* (2006) in the actuarial literature.

Figure 6 shows sample scenarios of future mortality rates generated by 7 models by year for a person aged 70 in 2017 in Japan. Future mortality rates at other ages or in other countries are similar.

3.2. Predictive power

This section verifies the predictive power of our proposed model. Accuracy measures how close an estimate of a fitted model is to the actual value for train data. In contrast, predictive accuracy or predictive power indicates how well a model fitted to train data predicts actual values for test data. Here we estimate predictive power for each model. We divided the data set into train data and test data. The period of the train data ranges from the year 2003 to 2014. That of the test data from the

Table 3: Root mean squared error (RMSE) of stochastic mortality models in the generalized age-period-cohort framework (unit: 10^{-3})

Model	Korea	Japan	Hong Kong	Taiwan	Mean rank
LC	4.450 (3)	2.191 (2)	7.514 (7)	3.713 (2)	3.50 (3)
RH	4.260 (2)	2.963 (3)	5.276 (5)	3.644 (1)	2.75 (2)
Currie (APC)	5.208 (4)	8.489 (6)	3.524 (3)	6.736 (5)	4.50 (5)
CBD M5	15.234 (7)	3.439 (4)	6.192 (6)	9.278 (7)	6.00 (7)
CBD M7	6.728 (5)	8.484 (5)	2.988 (1)	5.630 (4)	3.75 (4)
CBD M8	9.183 (6)	15.932 (7)	3.479 (2)	8.490 (6)	5.25 (6)
PLAT	3.796 (1)	1.087 (1)	4.149 (4)	4.186 (3)	2.25 (1)

The numbers in parentheses indicate the ranking of each model by country.
A mean rank means the average of rankings over countries.

Table 4: Symmetric mean absolute percentage error (SMAPE) of stochastic mortality models in the generalized age-period-cohort framework (unit: percentage)

Model	Korea	Japan	Hong Kong	Taiwan	Mean rank
LC	3.931 (2)	5.481 (3)	10.536 (7)	4.041 (2)	3.50 (4)
RH	4.163 (3)	3.456 (2)	6.837 (3)	3.344 (1)	2.25 (2)
Currie (APC)	4.509 (4)	8.491 (4)	5.456 (1)	6.484 (4)	3.25 (3)
CBD M5	10.238 (6)	11.570 (5)	10.158 (6)	6.790 (5)	5.50 (6)
CBD M7	7.327 (5)	25.938 (6)	6.860 (4)	10.496 (6)	5.25 (5)
CBD M8	14.327 (7)	35.805 (7)	7.477 (5)	14.462 (7)	6.50 (7)
PLAT	3.894 (1)	2.831 (1)	6.509 (2)	4.881 (3)	1.75 (1)

The numbers in parentheses indicate the ranking of each model by country.
A mean rank means the average of rankings over countries.

year 2015 to 2017.

We adopt two different measures for comparison of prediction errors. Root mean square error (RMSE) and symmetric mean absolute percentage errors (SMAPE). They are defined as the following:

$$\text{RMSE} = \sqrt{\frac{1}{h}\sum_{T+1}^{T+h}(\hat{y}_t - y_t)^2}, \tag{3.3}$$

$$\text{SMAPE} = \frac{1}{h}\sum_{T+1}^{T+h}\frac{|\hat{y}_t - y_t|}{(|y_t| + |\hat{y}_t|)/2}, \tag{3.4}$$

where $y_t = q_{x,t}$ is the observed mortality rate, $\hat{y}_t = \hat{q}_{x,t}$ is the forecasted mortality rate.

Tables 2–4 show the BICs, root mean squared error (RMSE), and symmetric maximum absolute percentage error (SMAPE) of the fitted models for each country. Although the model fit results are different for each country, a certain trend can be found for the most suitable and the most inappropriate models. LC and Currie have the moderate ranks in all indices; Plat and RH are the best and the second-best model in RMSE and SMAPE, respectively. However, they have a moderate rank in BIC. M7 and M8 are the best in BIC. However, they have a moderate rank in RMSE and SMAPE. M5 is the worst model for all indices. Nevertheless, those findings do not necessarily mean that the M5 model is always unsuitable. This model has the advantage of having a relatively small number of parameters and a simple form.

4. Sustainability of life annuities

Longevity poses risks from various perspectives. Solvency II suggests including at least level, volatility, and trend risk when measuring the longevity risk. Additionally, model risk, idiosyncratic risk (Buckham *et al.*, 2010), mis-estimation risk (Richards, 2016) can be considered as risks associated with longevity. In the quantitative measurement of longevity risk, the stressed-trend approach, shock approach, and value-at-risk approach exist among others (Richards *et al.*, 2014). The stressed-trend approach finds percentiles from the probability distribution of pension values, which are generated by first stressing the mortality trends and then simulating forecasted stochastic mortalities. This is also called a run-off approach because it considers all future cash flows over an extended period. The shock approach is a method of calculating risk capital by first multiplying the same shock factor by the mortality rates for each age, recalculating the pension value, and finally obtaining offset amount offset from the pre-shock pension value. Finally, the value-at-risk approach uses a stochastic mortality model similarly to the stressed-trend approach, taking into account the requirements of Solvency II, which requires risk capital for a one-year window. It, however, differs from the stressed-trend approach in the following processes. In the first step, stress such as volatility risk is applied for one year from the time of measurement to create a new set of samples. Then, the same statistical model is fitted to this new data to predict new mortality rates. A value-at-risk is measured in the statistical distribution of new pension values obtained by repeating these simulations n times.

To measure the longevity risk, we follow a new approach, a scenario approach similar to the stressed-trend approach. In Solvency II, which requires measurement of risk in a one-year window, the value-at-risk framework might be the most appropriate. However, this study aims to identify long-term risks by paying attention to the sustainability of capital invested in pensions. The stressed-trend approach applies the same constant shock to all future mortalities, but in the scenario approach, different random variable of shocks are applied to different years. This is called a scenario since it simulates forecasted mortalities. A Value-at-Risk value can be calculated from the distribution of these simulated mortalities.

We explain this scenario approach in detail. Values of annuities are calculated with discount rates of r as follows:

$$a^{Z=0}_{x:\overline{\omega-x}|} = \sum_{t=1}^{\omega-x} E\left[e^{-rt}{}_tp_x^{Z=0}\right] = \sum_{t=1}^{\omega-x} E\left[e^{-rt}\left(1-{}_tq_x^{Z=0}\right)\right], \tag{4.1}$$

where ${}_tp_x$ is a probability of survival at time t of a policyholder at age x, ω is the ultimate age which is 100 years-old in our work and r is a discount rate. We assume the discount rate is constant. In contrast, each pension value of simulation path i is as follows:

$$a^{Z_i}_{x:\overline{\omega-x}|} = \sum_{t=1}^{\omega-x} E\left[e^{-rt}{}_tp_x^{Z_i}\right] = \sum_{t=1}^{\omega-x} E\left[e^{-rt}\left(1-{}_tq_x^{Z_i}\right)\right], \tag{4.2}$$

where Z_i is not constant in time, but a scenario of the random variable Z following the standard normal distributions. We define a random variable of pension value at outset age x with a policy period of $\omega - x$, $a_{x:\overline{\omega-x}|}$. The ith realization of the random variable is $a^{Z_i}_{x:\overline{\omega-x}|}$. Then, the value-at-risk at the significance level of p by the scenario approach can be empirically estimated as follows:

$$\mathrm{VaR}_p\left(a_{x:\overline{\omega-x}|}\right) = \inf\left\{a_{x:\overline{\omega-x}|} \middle| F_p\left(a_{x:\overline{\omega-x}|}\right) \geq p\right\}. \tag{4.3}$$

Table 5: Expected values of annuities of an insured at age 60

Model	Korea	Japan	Hong Kong	Taiwan	Mean
LC	7.647	8.157	8.535	7.145	7.871
RH	7.456	8.087	10.012	8.138	8.423
Currie (APC)	7.991	8.303	8.556	7.259	8.027
CBD M5	7.790	8.100	8.560	7.161	7.903
CBD M7	7.882	8.327	8.580	7.183	7.993
CBD M8	8.863	9.143	8.995	7.847	8.712
Plat	7.720	8.224	8.536	7.103	7.896
Mean	7.907	8.334	8.825	7.405	8.118

The main difference between Equation (4.1) and Equation (4.3) is that Equation (4.1) is the pension value obtained by the mortality rate with the random variable Z equal to 0, the average value of the standard normal distribution. In this scenario, the mortality curve follows the central line of simulated stochastic mortalities in the future. In contrast, Equation (4.3) is the percentile of the pension value distributions obtained from the simulated stochastic mortality rates, each with different scenarios of random variable Z. This is closer to reality than just setting Z to 0 or a stress factor. This is because it is generally true to assume annual mortality rates as independent of each other, except for cases where mortality rates of consecutive years are affected by pandemic diseases over the years.

This analysis considers a male policyholder at age 60 at the inception of an annuity contract that pays 1 at the end of every policy year. We assumed that the discount rate of r is 4% for all years. For the time-series process of $\kappa_t^{(i)}$, the best-fit parameters were applied for each model and each country. γ_c is assumed to follow the ARIMA (1, 1, 0) process when necessary. In previous studies, the parameters of time series processes on the cohort effecte range from $p; q = 0, 1, 2, 3$, and $d = 0, 1, 2$. The parameter set of $p = 1, q = 1, d = 0$ is the most commonly used in the original studies (Renshaw and Haberman, 2006) and review papers (Villegas *et al.*, 2018; Lovász, 2011). Assuming that the time series process of the cohort effect is independent of the dynamics of period indices, it can be considered that the time series process of the cohort will be similar regardless of the models or countries. The annuity period is 40 years, p is 0.995%, and the number of simulations is 10,000 times.

Table 5 shows the expected value of the annuity for the policyholder. According to the CIA's 2017 estimates (Central Intelligence Agency, 2022), life expectancy at birth in Hong Kong, Japan, Korea, and Taiwan is ranked 7, 2, 12, and 41st worldwide. However, the estimate is based on the assumption that the current mortality rate remains the same. Table 5 reflects the tendency of mortality rates to decrease differently from country to country according to the stochastic mortality scenarios, showing that the mortality rate is low and the annuity value is high in the order of Hong Kong, Japan, Korea, and Taiwan. In general, if the life expectancy is high (if the expected mortality rate is low), the expected value of the annuity is high. In contrast, from Table 5, we find a counterexample to such a relation, which is that Hong Kong's current life expectancy is lower than that of Japan. However, Hong Kong's annuity value is higher than that of Japan. This suggests that Hong Kong's mortality rate is expected to decline faster in the future than that of Japan. Bartlett and Phillips (1995) had already predicted in 1995 that Hong Kong's aging rate would overtake Japan.

On the other hand, the expected values in Table 5 show extraordinary values for some models, so care must be taken when interpreting the results. For example, the expected values of the RH model and the Plat model are higher than those of the other models, which seems to be due to the low stability of the two models. Thus, it is not recommended to accept the results without further analysis on, for example, stability of the model.

Table 6: $\mathrm{VaR}_{0.995}(a_{x:\overline{\omega-x}|})/a^{Z=0}_{x:\overline{\omega-x}|} - 1$, ratio of VaR to expected value of annuities for an insured at age $x = 60$ (unit: %)

Model	Korea	Japan	Hong Kong	Taiwan	Mean
LC	2.381	2.371	4.797	3.537	3.272
RH	4.596	2.947	3.318	6.390	4.313
Currie (APC)	3.376	2.671	4.587	4.984	3.905
CBD M5	4.184	2.958	5.253	6.632	4.757
CBD M7	4.612	3.056	5.241	6.842	4.938
CBD M8	4.076	2.819	4.218	6.169	4.321
Plat	4.342	2.850	5.261	6.777	4.808
Mean	3.938	2.911	4.668	5.904	4.331

Table 6 presents the ratio of VaR to the expected value of an annuity for the policyholder, $\mathrm{VaR}_{0.995}(a_{x:\overline{\omega-x}|})/a^{Z=0}_{x:\overline{\omega-x}|} - 1$. The range of the VaR ratio in Table 2 is 2 to 6% is similar to the Richards *et al.* (2014)'s 3 to 6% of the risk capital calculated by the 1−year VaR approach for the 55 to 90 years-old policyholders. It is found that countries with a relatively high risk of longevity compared to the expected annuity values are in the order of Taiwan, Hong Kong, Korea, and Japan. Taiwan and Hong Kong, ranked No.1 and No.2 respectively, have a small population among comparative countries; therefore, there is a greater risk of longevity in Taiwan owing to the baseline effect (the denominator of VaR ratio), which is their low annuity value compared to Hong Kong. In the case of Japan, the uncertainty of annuity payments is the lowest owing to their large population and high annuity value. Summarily, the stability and sustainability of annuity payments of small countries by population are low, and vice versa.

Richards *et al.* (2014) showed that the capital requirement for annuities depends on models, policyholder's outset age, and discount rates. According to the study, the ratio of the capital requirement to the value of annuities may depend on models and it may vary by two or three times depending on age. This study also shows that VaR varies depending on the model. It should also be noted that the level of VaR may vary depending on the age and the discount rate is not considered in this study. However, the sensitivity analysis as an exercise can show that when the discount rate increases (decreases), the value of annuities decreases (increases) and the longevity risk increases (decreases).

5. Conclusion

The progress trend toward an aging society worldwide has increased concerns about the sustainability of the pension business. The sustainability of pensions can be predicted and analyzed through stochastic process models of mortality. Additionally, the comparison of stochastic mortality trends between countries is important by comparison itself. However, it is of greater importance in that it makes it possible to understand the globaltrend and differences between countries regarding the stability of future pension payments.

This study compared the robustness and predictive power of the stochastic mortality model, the Lee-Carter model (LC) and its generalized models including the Renshaw-Haberman (RH), Currie, Cairns-Black-Dowd M5, M7, M8, and Plat model focusing on four Asian countries. The nonparametric models, RH and Plat models have low stability, and the other models, Currie, M5, M7, and M8 models have relatively high stability. Among the latter, the M7 and M8 models show high predictive power.

The pension value is found to be high in the order of Hong Kong, Japan, Korea, and Taiwan. This order generally coincides with the order of life expectancy of countries, as the pension value has a negative correlation with the expected mortality rate. In contrast, the relative instability of pension value against changes in survival rate, that is, the relative longevity risk ranges from 2 to 6%. It is in the order of Taiwan, Hong Kong, Korea, and Japan, generally in the opposite order of population size. The smaller the population, the less sustainable the pension is.

This study has limitations in that it focuses on four Asian countries with reliable long-term mortality data, and it is necessary to expand the study by comparing it with different countries in the future. It is also necessary to investigate the sustainability of health insurance, which is as important as longevity risk in an aging society, or general whole-life insurance with mortality risk, the opposite of longevity risk.

References

Bartlett HP and Phillips DR (1995). Aging trends-Hong Kong, *Journal of Cross-Cultural Gerontology*, **10**, 257–265.

Bell W and Monsell B (1991). Using principal components in time series modeling and forecasting of age-specific mortality rates. In *Proceedings of the Social Statistics Section*, 154–159.

Bozikas A and Pitselis G (2018). An empirical study on stochastic mortality modelling under the age-period-cohort framework: The case of greece with applications to insurance pricing, *Risks*, **6**, 44.

Brouhns N, Denuit M, and Keilegom IV (2005). Bootstrapping the Poisson log-bilinear model for mortality forecasting, *Scandinavian Actuarial Journal*, **2005**, 212–224.

Buckham D, Wahl J, and Rose S (2010). *Executive's Guide to Solvency II*, John Wiley & Sons, New Jersey.

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

Cairns AJG, Blake D, Dowd K, Coughlan GD, 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**, 1–35.

Cairns AJG, Blake D, Dowd K, Coughlan GD, Epstein D, and Khalaf-Allah M (2011). Mortality density forecasts: An analysis of six stochastic mortality models, *Insurance: Mathematics And Economics*, **48**, 355–367.

Central Intelligence Agency (2022). *The World Factbook*. Available from: https://www.cia.gov

Chung WJ (2007). Stochastic forecasting health expenditure with the application to the Korea's national health insurance system, *Korean Social Security Studies*, **23**, 249–270.

Coffie E (2015). *A Comparison of Poisson or Negative Binomial Regression and Lee-Carter Models of Forecasting Norwegian Male Mortality*, Master's Thesis, University of Oslo.

Currie I (2006). *Smoothing and Forecasting Mortality Rates with p-Splines*, Presentation, London.

Currie ID (2016). On fitting generalized linear and non-linear models of mortality, *Scandinavian Actuarial Journal*, **2016**, 356–383.

Feng L and Shi Y (2018). Forecasting mortality rates: Multivariate or univariate models?, *Journal of Population Research*, **35**, 289–318.

Giacometti R, Bertocchi M, Rachev ST, and Fabozzi FJ (2012). A comparison of the Lee–Carter model and AR–ARCH model for forecasting mortality rates, *Insurance: Mathematics and Economics*, **50**, 85–93.

Girosi F and King G (2008). *Demographic Forecasting*, Princeton University Press, New Jersey.

Hunt A and Blake D (2015). Identifiability in age/period/cohort mortality models, *Annals of Actuarial Science*, **14**, 500–536.

Hunt A and Villegas AM (2015). Robustness and convergence in the Lee–Carter model with cohort effects, *Insurance: Mathematics And Economics*, **64**, 186–202.

Hwang J-Y, Thao BT, and Ko B (2018). A Bayesian comparative study on mortality improvements across nations using the Lee-Carter model, *The Journal of Risk Management*, **29**, 89–112.

Hyndman RJ, Koehler AB, Snyder RD, and Grose S (2002). A state space framework for automatic forecasting using exponential smoothing methods, *International Journal of Forecasting*, **18**, 439–454.

Hyndman RJ and Ullah S (2007). Robust forecasting of mortality and fertility rates: A functional data approach, *Computational Statistics & Data Analysis*, **51**, 4942–4956.

Jho JH (2020). Applications of stochastic interest and mortality in insurance risk calculation, *Korean Insurance Journal*, **122**, 1–34.

Kang H, Han ST, and Lee SK (2015). A Study on assessment index for long-term care insurance assessment committee using logistic regression model, *Journal of the Korean Data Analysis Society*, **17**, 3015–3023.

Kim TH (2006). Mortality forecasting for population projection, *Korea Journal of Population Studies*, **29**, 27–51.

Kim S (2012). A comparison study on the stochastic mortality models for measuring longevity risk, *Korean Insurance Journal*, **93**, 213–236.

Kim S (2013). A comparison study on methods of assessing longevity risk, *Journal of Insurance and Finance*, **24**, 93–121.

Koissi MC, Shapiro AF, and Högnäs G (2006). Evaluating and extending the Lee–Carter model for mortality forecasting: Bootstrap confidence interval, *Insurance: Mathematics and Economics*, **38**, 1–20.

Le M, Xiao X, Pamučar D, and Liang Q (2021). A study on fiscal risk of China's employees basic pension system under longevity risk, *Sustainability*, **13**, 5526.

Le TTN and Kwon HS (2021). Suitability of stochastic models for mortality projection in Korea: A follow-up discussion, *Communications for Statistical Applications and Methods*, **28**, 171–188.

Lee R and Miller T (2001). Evaluating the performance of the Lee-Carter method for forecasting mortality, *Demography*, **38**, 537–549.

Lee RD and Carter L (1992). Modeling and forecasting US mortality, *Journal of the American Statistical Association*, **87**, 659–671.

Lovász E (2011). Analysis of Finnish and Swedish mortality data with stochastic mortality models, *European Actuarial Journal*, **1**, 259–289.

Millossovich P, Villegas AM, and Kaishev VK (2018). StMoMo: An R package for stochastic mortality modelling, *Journal of Statistical Software*, **84**, 1–38.

Neves C, Fernandes C, and Hoeltgebaum H (2017). Five different distributions for the Lee–Carter model of mortality forecasting: A comparison using GAS models, *Insurance: Mathematics and Economics*, **75**, 48–57.

Plat R (2009). On stochastic mortality modeling, *Insurance: Mathematics and Economics*, **45**, 393–404.

Renshaw AE and Haberman S (2003). Lee–Carter mortality forecasting with age-specific enhancement, *Insurance: Mathematics and Economics*, **33**, 255–272.

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

Richards SJ (2016). Mis-estimation risk: Measurement and impact, *British Actuarial Journal*, **21**, 429–457.

Richards SJ, Currie ID, and Ritchie GP (2014). A value-at-risk framework for longevity trend risk, *British Actuarial Journal*, **19**, 116–139.

Stefani A and Kwon HS (2021). A multi-state model approach for risk analysis of pensions for married couples with consideration of mortality difference by marital status, *Communications for Statistical Applications and Methods*, **28**, 611–626.

Tian Y and Zhao X (2016). Stochastic forecast of the financial sustainability of basic pension in China, *Sustainability*, **8**, 46.

University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany) (2021). Human Mortality Database. Available from: https://www.mortality.org

Villegas AM, Kaishev VK, and Millossovich P (2018). StMoMo: An R package for stochastic mortality modeling, *Journal of Statistical Software*, **84**, 1–38.

Won CH (2010). Estimation of longevity risk and funding ratio of public pension, *The Korean Journal of Financial Management*, **27**, 1–26.

Received May 05, 2022; Revised July 23, 2022; Accepted August 01, 2022