보험금융연구 제27권 제1호(2016. 2) pp. 107-133

The Performance Evaluation on the General Procedure for Forecasting Mortality*

사망률 예측을 위한 일반적 과정 성능평가

Sang Il Lee**
이 상 일

This study investigates the forecasting ability of the general procedure(GP) using mortality data for South Korean males during 1983-2010. The GP was recently introduced to construct a stochastic mortality model by including every significant demographic feature in historical mortality data. We assess the GP via a comparison with seven existing stochastic mortality models, testing in-sample fit and out-of-sample prediction for three age groups: 1-79, 11-79, and 60-79. The results suggest that the GP consistently outperforms other models with regard to the Bayesian Information Criterion(BIC) and Mean Absolute Percentage Error (MAPE). This shows that the GP extracts optimal risk factors for the projections of age-specific mortality rates from mortality data. Furthermore, we examine predicted levels of uncertainty in forecasts at different ages and show how the risk can be hedged using q-forwards. This information is useful for pension providers or insurers to hedge future unexpected liabilities.

Key words: General Procedure, Stochastic Mortality Models, Forecasting Performance, Q-forwards, Longevity Risk
한국연구재단 분류 연구분야 코드: B051609

* We appreciate the anonymous reviewers for their valuable comments and suggestions that greatly improved the quality of the paper.
** 아주대학교 금융공학과 박사과정(silee@ajou.ac.kr)
논문 투고일: 2015. 03. 02, 논문 최종 수정일: 2015. 10. 27, 논문 게재 확정일: 2016. 02. 15

I. Introduction

Human mortality decreased significantly in the 20th century(Preston, 1993; Smith, 1993). During the first part of the 20th century, the decline in mortality primarily resulted from the reduction of infectious diseases for younger groups, and during the last decades of the 20th century, the decline resulted from the reduction in deaths owing to chronic diseases mainly for older age groups(Antolin, 2007). This increase in longevity of human life is a blessing, but creates systemic risk in pension systems and public retirement systems.

Population forecasts using static life tables would overestimate death rate because they do not consider the evolution of mortality over time. An alternative solution is to use a stochastic mortality model. Lee and Carter(1992) first proposed a stochastic mortality model for forecasting mortality in the US, which is currently being considered as the benchmark model. Subsequently, various models have been proposed to contain more detailed demographic features such as cohort effects and age-dependent differentials(Lee and Miller, 2001; Booth et al., 2002; Brouhns et al., 2002; Girosi and King, 2005; Renshaw, 2006; Cairns et al., 2006a; Currie, 2006; H′ari et al., 2008; Tulijapurkar, 2008; Plat, 2009; O'Hare, 2012).

In spite of such developments, it is ambiguous whether these models have enough flexibility to represent age-specific differentials in mortality data. Moreover, the models suffer from the limitation of universal applicability, because they were designed based on the observation of mortality data for a few developed countries, especially the UK.

These problems can be overcome by using the "General Procedure"(GP) developed by Hunt and Blake(2014). The GP provides an effective method of capturing all major age-dependent demographic features from mortality data and of incorporating them into a stochastic mortality model using non-parametric period and parametric age functions. Since age/period functions are independent of each other, we are able to

establish the age-specific mortality structure of the population under consideration and a distinctive demographic pattern by comparing estimated age/period functions across nations. We call the stochastic mortality model built by the GP as the GP model.

The GP model was originally constructed using mortality data from the UK and was found to fit well in-sample on the basis of the Bayesian Information Criterion(BIC). However, this does not necessary imply that the model is informative regarding out-of-sample predictive contents. The GP model could be useful to forecast future mortality rates, since it considers all risk factors for projections by obtaining the period functions corresponding to age functions. In this study we investigate the accuracy of projected mortality rates by the GP model. We evaluate the forecasting ability by comparing out-of-sample forecasting performances of the GP model with those of the seven popular models shown in Table 2. To obtain robustness of the out-of-sample test, we also examine the impact of parameter uncertainty using the residual bootstrapped technique(Koissi et al., 2006). This serves as a test for demonstrating computational stability of the GP model. We also present fan charts of the forecasts produced by the GP and two other models to show the impact of diverse risk sources on mortality rate forecasts. This information is valuable for pension providers and insurers to hedge unexpected liabilities.

The remainder of the paper is organized as follows. Section 2 presents a stochastic mortality model produced by the GP and assesses its goodness-of-fit. Section 3 builds a time-series forecasting model for the time-varying indexes of the GP model. Section 4 shows the in-sample fit and the out-of-sample forecasting accuracy of the GP model, along with a comparative analysis of the performance of the GP model with those of the other models. Section 5 examines the hedge effectiveness of q-forwards using sample paths generated by the GP model. Finally, Section 6 summarizes and concludes the study.

II. Constructing Stochastic Mortality Model using the General Procedure

1. Definition of Mortality and Data Source

To build stochastic mortality models, we use the crude(i.e., unsmoothed) death rate $m_{x,t}$ for age x in calendar year t:

$$m_{x,t} = \frac{Number\ of\ deaths\ during\ calendar\ year\ t\ aged\ x\ last\ birthday}{Average\ population\ during\ calendar\ year\ t\ aged\ x\ last\ birthday}, \tag{2.1}$$

where calendar year t is defined as running from t to $t+1$; and the average population is approximated by the estimate of the population aged x last birthday in the middle of the calendar year. The one-year mortality rate $q_{x,t}$ is given by

$$q_{x,t} = 1 - \exp[-m_{x,t}], \tag{2.2}$$

which is the probability that an individual aged exactly x at exact time t will die between t and $t+1$. We use the South Korean male mortality data during 1983~2010 for the age range 1~79[1)].

2. The GP Model

We construct a stochastic mortality model using the GP. For brevity, we restrict the discussion to points necessary for evaluating its predictive performance. Hunt and Blake(2014) provide more detailed information on the GP, such as the identifiability constraints and the algorithms for estimating parameters.

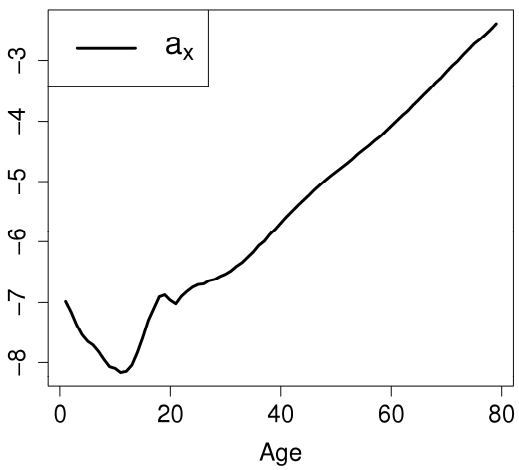
STAGE 0. The first step is to fit the mortality data to the model,

1) The data is obtained from Statistics Korea(KOSTAT). Available at www.kostat.go.kr.

$$\ln(m_{x,t}) = a_x, \tag{2.3}$$

where a_x is a nonparametric age function to be estimated, which reflects the average age-specific pattern of mortality across the full age range. To estimate the parameter a_x and the parameters in the following stages, we use Brouhns' methodology by maximizing the log-likelihood of a Poisson distribution(Brouhns et al., 2002). Figure 1 displays the estimation results, showing the age pattern of mortality in childhood, young adulthood(the accident hump), and senescence.

〈Figure 1〉 Estimated values of a_x for Stage 0



STAGE 1. For improving flexibility of the model (2.3), the next stage is to add a nonparametric age/period term $b_x^{(1)}k_t^{(1)}$ to it:

$$\ln(m_{x,t}) = a_x + b_x^{(1)}k_t^{(1)}, \tag{2.4}$$

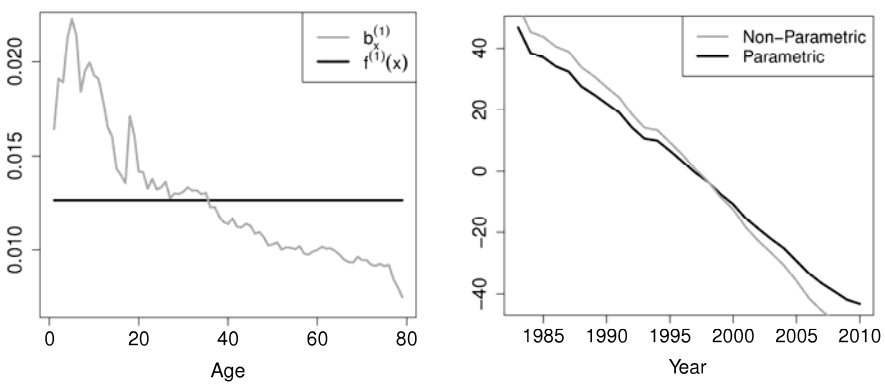
where the nonparametric age function $b_x^{(1)}$ and the nonparametric period function $k_t^{(1)}$ describe the age effect and the period effect, respectively. The two grey lines shown in the left and right panels of Figure 2 represent the fitted values of $b_x^{(1)}$ and $k_t^{(1)}$, respectively. To improve the parsimony of the model (2.4), we need to design a parametric age function $f^{(1)}(x)$ reflecting the significant demographic feature of $b_x^{(1)}$.

We select a constant function($f^{(1)}(x) \propto 1$) describing a general level of mortality for all ages. Substituting $b_x^{(1)}$ with $f^{(1)}(x)$, the updated model is expressed as

$$\ln(m_{x,t}) = a_x + f^{(1)}(x)k_t^{(1)}. \tag{2.5}$$

The two black lines shown in the left and right panels of Figure 2 represent the fitted values of $f^{(1)}(x)$ and $k_t^{(1)}$, respectively. This substitution provides a trade-off between the fit quality and the parsimony of the model.

〈Figure 2〉 Age functions(left-hand panel) and period functions(right-hand panel) for Stage 1



STAGE 2. Similarly, adding a nonparametric age/period $b_x^{(2)}k_t^{(2)}$ function to the model (2.5), we arrive at

$$\ln(m_{x,t}) = a_x + f^{(1)}(x)k_t^{(1)} + b_x^{(2)}k_t^{(2)}. \tag{2.6}$$

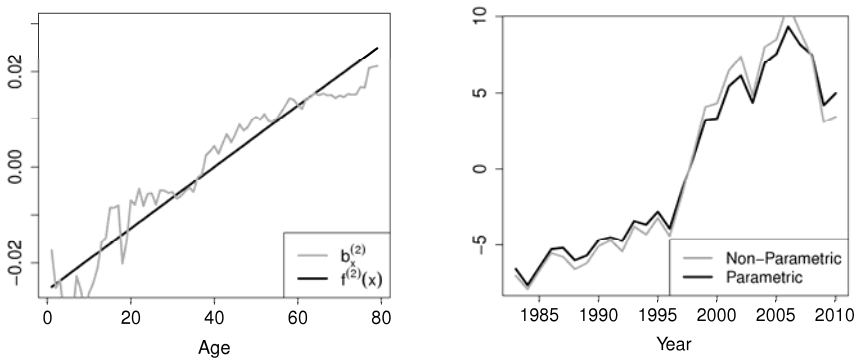
The two grey lines shown in the left and right panels of Figure 3 represent the fitted values of $b_x^{(2)}$ and $k_t^{(2)}$, respectively. We use a straight line to capture the dominant trend of $b_x^{(2)}$. Substituting $b_x^{(2)}$ with $f^{(2)}(x)(\propto x)$, we arrive at

$$\ln(m_{x,t}) = a_x + f^{(1)}(x)k_t^{(1)} + f^{(2)}(x)k_t^{(2)}. \tag{2.7}$$

The two black lines shown in Figure 3 represent the fitted values of $f^{(2)}(x)$ on the left-hand panel and $k_t^{(2)}$ on the right one.

STAGES 3–6. We repeat the procedure until the functions $f^{(3)}(x), \cdots, f^{(6)}(x)$ are obtained. Then, we arrive at

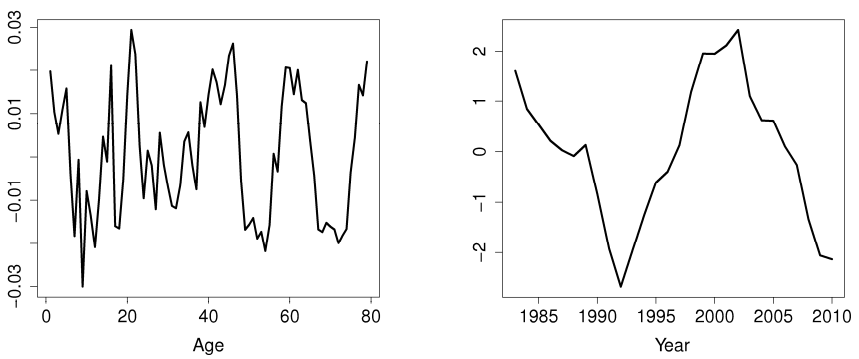
〈Figure 3〉 Age functions(left-hand panel) and period functions(right-hand panel) for Stage 2



$$\ln(m_{x,t}) = a_x + \sum_{i=1}^{6} f^{(i)}(x) k_t^{(i)}. \tag{2.8}$$

Table 1 shows the implemented six parametric functional forms and their demographic implications: the average level of mortality($i = 1$); the increase in the general level of mortality with aging($i = 2$); mortality differentials related to young adult mortality ($i = 3$); childhood mortality($i = 4$); postponement of old age mortality($i = 5$); and an accident hump($i = 6$).

〈Figure 4〉 Nonparametric age(left-hand panel) and period(right-hand panel) functions for Stage 7



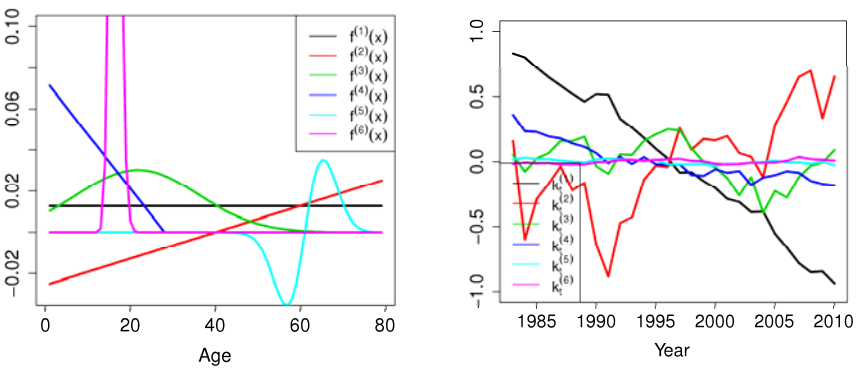
STAGE 7. Figure 4 shows the fitted values of $b_x^{(7)}$(left) and $k_t^{(7)}$(right) obtained by adding the additional age/period term $b_x^{(7)}k_t^{(7)}$ to the model (2.8). We do not observe any significant demographic features in the age function, such as distinct features superimposed on a specific age range or trends across the entire age range. Thus, we do not consider hidden parametric age functions.

STAGE 8. The final stage is to add the cohort term γ_{t-x} to estimate lifelong effects specified by different generations. Then, we arrive at the final model:

$$\ln(m_{x,t}) = a_x + \sum_{i=1}^{6} f^{(i)}(x)k_t^{(i)} + \gamma_{t-x}, \tag{2.9}$$

where $y = t - x$ (i.e., the year of birth). Figure 5 shows the estimation results obtained using the model: the left-hand panel shows parametric age functions, $f^{(i)}(x)$, and the right-hand panel the mortality indexes, $k_t^{(i)}$, scaled by deducting their means and dividing by their standard deviations for convenience sake. The period functions, $k_t^{(i)}$, represent the time-trending behavior of the corresponding age functions $f^{(i)}(x)$. The fitted cohort effects are shown as dots in Figure 7.

〈Figure 5〉 (Color online) Parametric age functions(left-hand panel) and period functions(right-hand panel) scaled by deducting their means and dividing by their standard deviations



〈Table 1〉 Parametric age function and demographic significance

Term (i)	Description	$f^{(i)}(x)\propto$	Free parameters	Demographic significance
1	Constant	1	None	General level of mortality
2	Linear	$x-\overline{x}$	None	Gompertz slope
3	Normal	$\exp(-\frac{(x-\hat{x})^2}{\sigma^2})$	$\hat{x},\sigma$	Young adult mortality
4	Put option	$(x_c-x)^+$	x_c	Childhood mortality
5	Rayleigh	$(x-\hat{x})\exp(-\rho^2(x-\hat{x})^2)$	$\hat{x},\rho$	Postponement of old age mortality
6	Log-normal	$\frac{1}{x}\exp(-\frac{(\ln(x)-\hat{x})^2}{\sigma^2})$	$\hat{x},\sigma$	Peak of accident hump

〈Table 2〉 Specifications of the seven stochastic mortality models

Model	Formula
M1	Lee and Carter(1992) $\log m_{x,t}=a_x+b_x^{(1)}k_t^{(1)}$
M2	Renshaw(2006) $\log m_{x,t}=\alpha_x+b_x^{(1)}k_t^{(1)}+b_x^{(2)}\gamma_{t-x}$
M3	Currie(2006) $\log m_{x,t}=a_x+n_a^{-1}k_t^{(1)}+n_a^{-1}\gamma_{t-x}$
M5	Cairns et al.(2006a) $\log(\frac{q_{x,t}}{1-q_{x,t}})=k_t^{(1)}+k_t^{(2)}(x-\overline{x})$
M6	Cairns et al.(2009) $\log(\frac{q_{x,t}}{1-q_{x,t}})=k_t^{(1)}+k_t^{(2)}(x-\overline{x})+\gamma_{t-x}$
M7	Cairns et al.(2009) $\log(\frac{q_{x,t}}{1-q_{x,t}})=k_t^{(1)}+k_t^{(2)}(x-\overline{x})+k_t^{(3)}((x-\overline{x})^2-\hat{\sigma}_x^2)+\gamma_{t-x}$
M8	Cairns et al.(2009) $\log(\frac{q_{x,t}}{1-q_{x,t}})=k_t^{(1)}+k_t^{(2)}(x-\overline{x})+\gamma_{t-x}(x_c-x)$

Note: $\overline{x}$ is the mean age over the range of ages being used in the analysis. n_a is the number of ages covered in the sample age range. $\hat{\sigma}_x^2$ is the mean value of $(x-\overline{x})^2$. x_c is a constant parameter to be estimated. M4 is not included in our analysis. It is the P-splines model developed in Currie(2006).

〈Table 3〉 BIC measures for GP, M1–M3 and M5–M8 models using mortality data for South Korean males aged 1–79

Model	Log–likelihood	BIC
General Procedure	-10,895.49	-12,241.69
M1	-17,138.76	-17,855.35
M2	-11,691.97	-13,121.29
M3	-15,478.96	-16,295.71
M5	-115,383.90	-115,599.60
M6	-56,615.52	-57,235.79
M7	-45,644.04	-46,368.33
M8	-33,587.82	-32,959.84

3. Assessing Model Fit

In order to assess the goodness-of-fit of the GP model and the other seven models listed in Table 2, we use the BIC measure:

$$BIC = L(\hat{\phi}) - \frac{1}{2} K \ln(P), \tag{2.10}$$

where $L(\hat{\phi})$ is the log-likelihood of the estimated parameter $\hat{\phi}$; P is the number of observations; and K is the number of parameters being estimated. It provides a trade-off between the fit quality and parsimony of the model. The best estimate is chosen based on the highest value of BIC measure. Table 3 shows the BIC scores for the eight models estimated using mortality data for South Korean males aged 1~79 during 1983~2010. The GP model has the highest value of BIC(-12,241.69), making it the best-fit model. We also observe that the models, M5~M8, do not deliver significant performance results. This is because they were solely designed for higher age groups.

III. Modeling Time-varying Indexes

To examine the future distribution of mortality rates, we build time series models for the stochastic variables(i.e., the mortality indexes and cohort effects) of the eight models.

1. Modeling GP Mortality Indexes

Chan et al.(2014) suggest a general class of vector autoregressive integrated moving average(VARIMA) model for multiple mortality indexes. However, when we applied the VARIMA model to the GP indexes, the first and second best-fit models do not pass diagnostic tests. Thus, the GP mortality indexes, $k_t^{(i)}$, are modeled using a multivariate random walk with drift(RWD) commonly used to build time-series models for mortality indexes(for example, a univariate RWD model in Lee and Carter(1992) and a multivariate RWD model in Cairns et al.(2006a, 2011)).

The multivariate RWD process for the mortality indexes $k_t = (k_t^{(1)}, \cdots, k_t^{(6)})^T$ is defined as follows:

$$k_{t+1} = k_t + \mu + CZ_{t+1}, \tag{3.1}$$

where $\mu \in R^6$ is a constant 6×1 vector; C is a constant 6×6 upper triangular matrix; and Z_t is a six-dimensional standard normal random variable. Vector μ represents the drift and matrix C the volatility of the risk factors, satisfying $Var(\triangle k_t) = CC^T$. The volatility matrix C is uniquely determined from $Var(\triangle k_t)$ based on the Cholesky-decomposition. The estimation results are

$$\hat{\mu} = (-3.413, 0.155, 0.035, -0.524, -0.040, 0.025)^T \tag{3.2}$$

and

$$Var(\nabla \hat{k}_t) = \hat{C}\hat{C}^T = \begin{pmatrix} 7.12 & -4.84 & -5.87 & -1.17 & 0.29 & -0.12 \\ -4.84 & 5.07 & 4.48 & 1.54 & -0.35 & -0.05 \\ -5.87 & 4.48 & 8.95 & 0.27 & -0.31 & 0.02 \\ -1.17 & 1.54 & 0.27 & 1.48 & -0.18 & -0.03 \\ 0.29 & -0.35 & -0.31 & -0.18 & 0.11 & 0.02 \\ -0.12 & -0.05 & 0.02 & -0.03 & 0.02 & 0.07 \end{pmatrix} \quad (3.3)$$

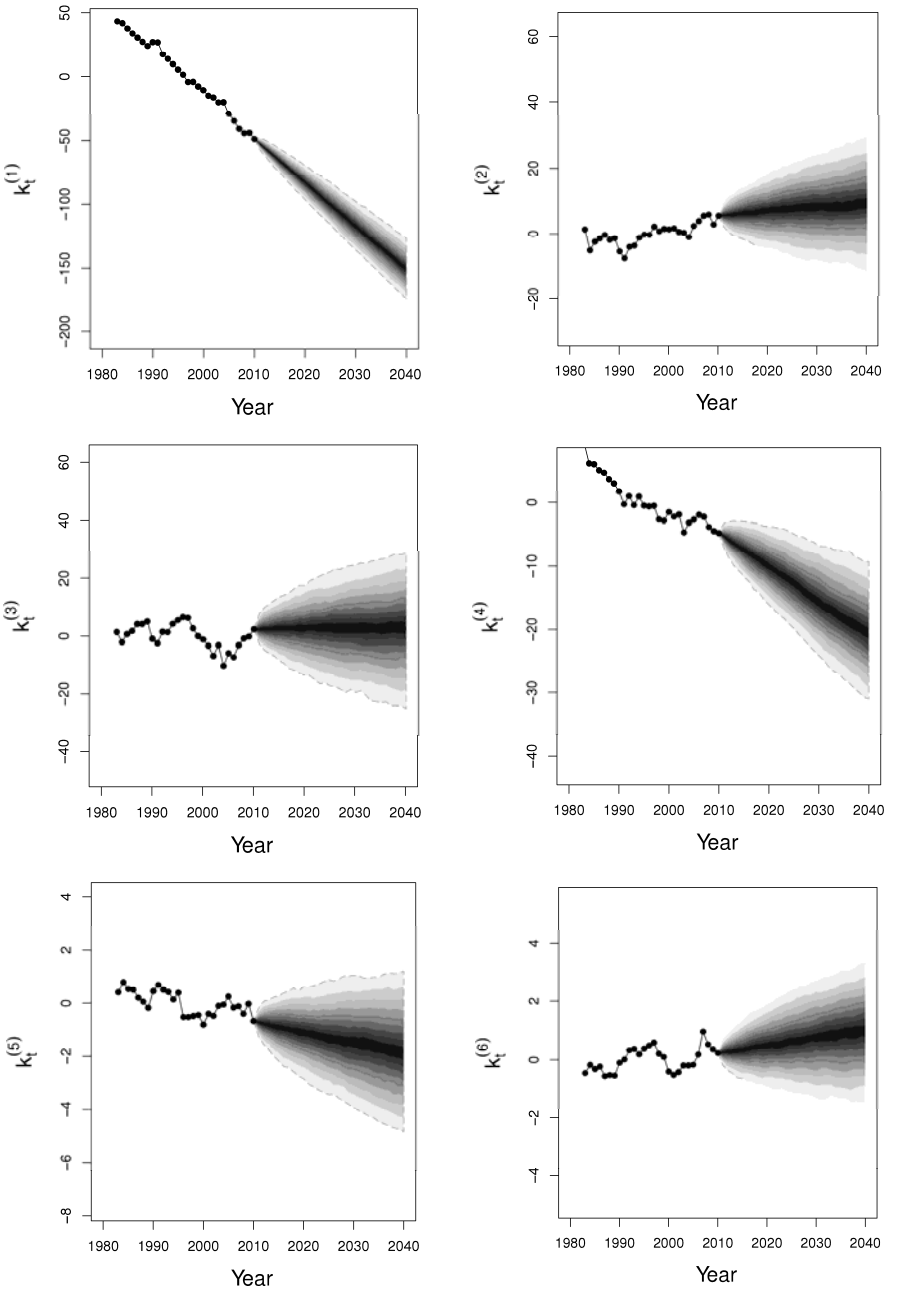
Figure 6 displays the fan charts for the GP mortality indexes obtained by simulating 1,000 paths. The dashed lines indicate 95% confidence intervals. The projections show diverse patterns of trend and level of uncertainty, which reflects own intrinsic features of specific age groups for mortality rate projections.

2. Modeling Cohort Effects

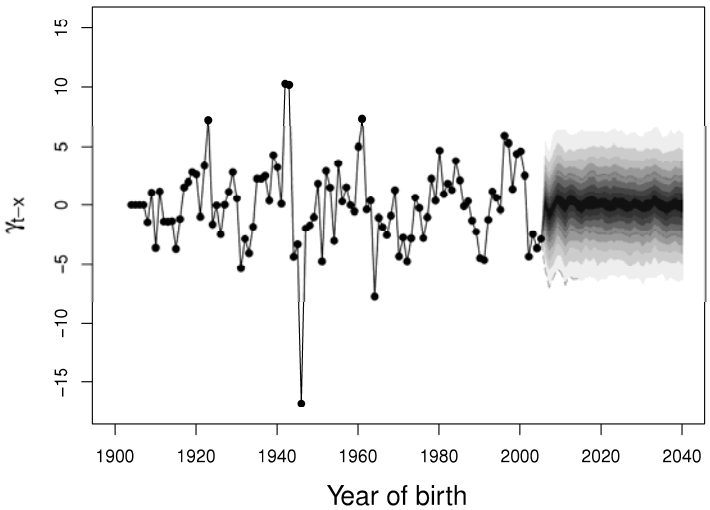
We assume that cohort effects, γ_{t-x}, are independent of $k_t^{(i)}$. For the time horizon of 1983~2010, the year of birth, $t-x$, is given from 1904 to 2009, and the cohort effects with fewer than 5 observations are excluded from the fitting procedure. Since an RWD model is unlikely to be appropriate for modeling cohort effects(Cairns et al., 2011), we use a more general autoregressive integrated moving average(ARIMA) process. The ARIMA(p,d,q) models with $d=0,1,2$ and p, $q=0,1,2,3,4,5$ are considered as candidates. Of these models, the best model is ARIMA(2,0,2) based on the BIC.

Figure 7 shows the fan chart for cohort effects obtained by simulating 1,000 paths. We observe a strong discontinuity between 1945 and 1946 relating to the end of the Second World War. A similar result is also observed in the cohort effects estimated by the GP model for the UK mortality data(Hunt and Blake 2014).

〈Figure 6〉 Mortality index fan charts, $k_t^{(i)}$. The dots display the estimates of the mortality indexes fitted to the historical data.



〈Figure 7〉 Cohort effect fan charts γ_{t-x}. The dots display the estimates of the cohort effects fitted to the historical data.



IV. Evaluating Predictive Accuracy

In this section, we examine the accuracy of projected mortality rates for the eight models. To establish the robustness of our results, we perform out-of-sample forecasting experiments for three different age groups: 11~79(younger and older age ranges), 1~79(full age range), and 60~79 years(older age range).

1. Mortality Projections for Age Range 11–79

We fit the models to mortality data over the age range 1~79 and over four different historical "look-back" windows: (1) 1983~2000; (2) 1983~2001; (3) 1983~2002; and (4) 1983~2003.

We first evaluate the fitting performances in terms of the BIC. As shown in Table 4, the best-fit model is the GP model with the highest BIC measure(marked with the

symbol *) for all look-back windows. The second-best model(marked with the symbol **) is M3 for the look-back window 2001~2010 and M2 for the other look-back windows.

Next, we evaluate the forecasting performances over the four different "look-forward" windows: (1) 2001~2010; (2) 2002~2010; (3) 2003~2010; and (4) 2004~2010. Mortality indexes over the windows are generated by using the multivariate(or univariate) RWD model fitted to the corresponding look-back windows. The models are quantitatively assessed based on the accuracy of projections using the mean absolute percentage error(MAPE) between $\bar{q}_{x,t}$, the mean forecasts of $q_{x,t}$, and $q_{x,t}$, historical data:

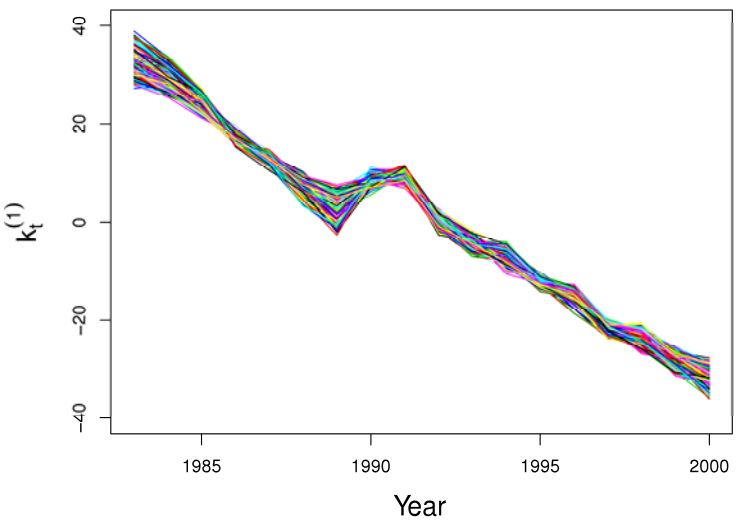
$$\text{MAPE}= \frac{1}{N}\sum_{t=t_0}\sum_{x}\left|\frac{q_{x,t}-\bar{q}_{x,t}}{q_{x,t}}\right|, \tag{4.1}$$

where t_0 is the first year of the look-forward horizons and N is the number of observations. The MAPE measure is computed only for the age range 11~79 for the convenience of not extrapolating the cohort effects γ_{t-x}. That is, the cohort effects for computing future mortality rates are obtained from the estimations over the look-back windows. As shown in Table 4, the GP model is the best-fit model with the smallest MAPE measures for all look-forward windows, and the second-best model is M1.

Further, we examine the effect of parameter uncertainty using the residual bootstrap methodology proposed by Koissi et al.(2006). Figure 8 displays the mortality index $k_t^{(1)}$ for 500 bootstrap residual matrices over a look-back window of 1983~2000. We observe that the underlying pattern of $k_t^{(1)}$ (and also the other indexes, not shown here) remain unchanged, which reflects the computational stability of the GP model. The MAPE value calculated over the period 2001~2010 is 9.31%, which is close to the value of 9.07% in the absence of parameter uncertainty. This (roughly) reflects that parameter uncertainty has little impact on ranking the forecasting performance of mortality models on the basis of MAPE.

These results show that the GP model delivers the best performance in terms of both in-sample and out-of-sample fit, regardless of the sample types. The second model, M2 or M3 in terms of BIC and M1 in terms of MAPE, is highly dependent on the sample types. We also observe that the model M5 does not deliver significance performance results, since it was designed for higher groups only.

〈Figure 8〉 (Color online) Mortality index $k_t^{(1)}$, 500 bootstraps



2. Mortality Projections for Age Range 1-79

To simulate mortality rates for the age range 1~79, we need to specify stochastic processes for modeling cohort effects. For the models with the cohort term γ_{t-x}, M2, M3, M6, M7, and M8, we consider ARIMA(p,d,q) models with $d=0,1,2$ and $p,q=0,1,2,3,4,5$ as candidates and pick the best-fit model among them using the BIC measure. The period 1983~2000 is used as the historical look-back window for examining ten-year forecasts. As shown in Table 5, the best predictor is the GP model with the smallest MAPE measure of 8.24%, and the second-best model is M3 with 11.12%.

〈Table 4〉 BIC results for age range 1–79 over a look–back window and MAPE results for age range 11–79 over a look–forward window

	GP	M1	M2	M3	M5	M6	M7	M8
(1) A look-back window of 1983-2000 and a look-forward window of 2001-2010								
BIC	-8,074.48*	-10,617.72	-8,407.68	-8,382.38**	-42,956.96	-19,859.95	-17,804.65	-19,713.98
MAPE	9.07%*	15.88%**	17.76%	18.31%	55.54%	21.60%	19.35%	20.25%
(2) A look-back window of 1983-2001 and a look-forward window of 2002-2010								
BIC	-8,511.59*	-11,354.66	-8,818.04**	-8,910.33	-44,270.16	-20,724.14	-18,505.36	-20,647.88
MAPE	8.00%*	13.54%**	22.19%	16.73%	51.00%	20.57%	17.76%	20.64%
(3) A look-back window of 1983-2002 and a look-forward window of 2003-2010								
BIC	-8,997.48*	-12,091.72	-9,217.34**	-9,440.92	-45,535.85	-21,621.59	-19,195.7	-21,644.11
MAPE	7.92%*	12.11%**	23.66%	15.97%	49.92%	20.31%	17.79%	15.67%
(4) A look-back window of 1983-2003 and a look-forward window of 2004-2010								
BIC	-9,402.29*	-12,693.66	-9,678.78**	-9,837.27	-46,691.19	-22,543.7	-19,961.73	-22,613.41
MAPE	8.94%*	11.38%**	20.24%	15.34%	52.87%	20.24%	18.76%	20.43%

〈Table 5〉 MAPE results for age range 1–79 over a look–back window of 2001–2010 and ARIMA model for cohort effects

	GP	M1	M2	M3	M5	M6	M7	M8
MAPE	8.24%*	14.52%	12.74%	11.12%**	16.44%	31.74%	19.25%	30.01%
Model	ARIMA(2,0,2)	None	ARIMA(2,0,1)	ARIMA(1,1,0)	None	ARIMA(1,1,0)	ARIMA(0,1,0)	ARIMA(2,0,1)

Note: M1 and M5 do not include a cohort term.

〈Table 6〉 MAPE results for age range 60–79 over a look–back window of 2001–2010 and ARIMA model for cohort effects

	GP	M1	M2	M3	M5	M6	M7	M8
MAPE	2.82%*	6.11%	18.06%	4.37%**	7.06%	8.68%	5.19%	7.31%
Model	ARIMA(2,0,2)	None	ARIMA(0,1,0)	ARIMA(0,1,1)	None	ARIMA(1,0,0)	ARIMA(1,0,0)	ARIMA(1,1,1)

Note: M1 and M5 do not include a cohort term.

3. Mortality Projections for Age Range 60-79

For the age range 60~79, the models M1~M3 and M5~M8 are fitted into the mortality data for a back-ward window of 1983~2000, and time-series models are also constructed. However, the GP model is fitted to the age range 1~79, because the model already has the age functions characterizing age-specific demographic features over the range. Among the parametric age functions, only the three functions, $f^{(1)}$, $f^{(2)}$, and $f^{(5)}$ effectively contribute to the projections. The others have little effect on the projections, since their values are near zero in the range as shown in the left panel of Figure 5. This might be a penalty to the GP model owing to some poor-fit arising from the difference between the fitting age range and the evaluating one.

The MAPE measures are computed over a forward window of 2001~2010. As shown in Table 6, the GP model is the best predictor with the smallest MAPE measure of 2.82%, and the second-best model is M3.

V. Applications to Mortality/Longevity-linked Derivatives

1. Comparison of Mortality Fan Charts

For the LC(M1), CBD(M5), and GP models, we look at the volatility of projections of mortality rates at younger and older ages. The top panel of Figure 9 shows the fan charts for mortality rates at age 25 for each of the LC and GP models fitted to mortality data for the age range 1~79. The fan under the GP model is significantly wider than those under the LC model, which results from the multiple risk sources of the GP model. This reflects that at the younger age, the forecasting performance of the LC model underestimates the risk associated with the forecast levels of uncertainty.

The bottom panel of Figure 9 shows the fan charts for mortality rates at age 65 for

each of the LC, CBD, and GP models. As mentioned in Subsection 4.3, the LC and CBD models are estimated using mortality data for the age range 60~79, and the GP model is estimated using mortality data for the age range 1~79. We observe that the widths of confidence intervals of the fans are broadly similar.

Information regarding forecast levels of uncertainty is valuable for pension providers or insurers to hedge future unexpected liabilities. We shall look at hedging strategy for the mortality risk using q-forwards.

2. Hedging Longevity Risk using q-forwards

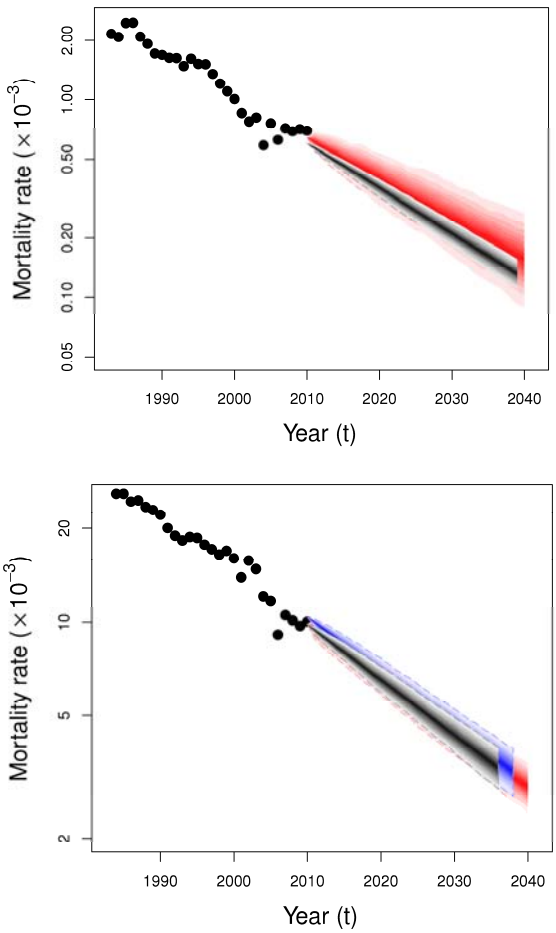
A newly emerging life market offers risk management opportunities against mortality risks such as brevity risk(i.e., the risk of premature death) and longevity risk(i.e., the risk of living too long). Mortality-linked securities and derivatives have been extensively developed by academic communities as well as industry specialists,(e.g., longevity bond(Blake 2001), k-forward(Chan et al., 2014; Tan et al., 2014), q-forward(Coughlan et al., 2007)), and theoretical frameworks for pricing them have been established(Cairns et al., 2006b; Loeys, 2007; Bauer, 2010; Barrieu, 2012). The payoff structure of such products is basically expressed as a function of current expectations for future mortality rate or indexes related to mortality rate. Thus, finding the most efficient estimator and assessing forecast levels of uncertainty in projections play key roles in mortality risk management.

As a simple example, we consider a pension provider's hedging strategy against unexpected liabilities owing to longevity risk. The pension provider uses q-forward derivatives launched by J. P. Morgan in 2007. Figure 10 illustrates the transaction between party A(e.g., a pension provider) and party B(e.g., a bank). The settlement of q-forward contract at maturity is given by

$$\begin{cases} N\times(q_f - q_r) & \text{for } party\ A \\ N\times(q_r - q_f) & \text{for } party\ B, \end{cases}$$

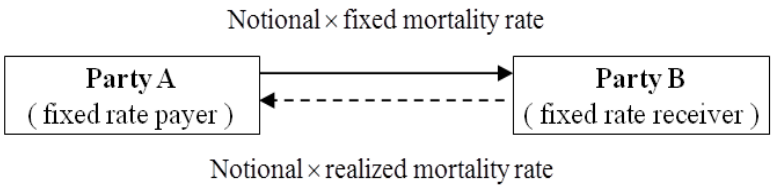
where N is the notional amount, q_f is the fixed mortality rate determined at the time of evaluation, that is, the best estimate, and q_r is the realized mortality rate at the time of maturity, T. Using q-forwards, pension providers can hedge against the risk of

〈Figure 9〉 (Color online) The top panel shows the fan charts at age 25 for the LC model(short) and the GP model(long). The bottom panel shows the fan charts at age 60 for the LC model(short), the CBD model(middle), and the GP model(long). The dots represent historical mortality rates for period 1983–2010.



decreasing mortality of plan members, and life insurers can protect themselves against significant increases in the mortality of policyholders. For example, when unexpected reduction in mortality rate arises, the pension provider can cover the loss owing to the longevity risk by receiving more funds from the bank.

〈Figure 10〉 A q-forward transaction at Maturity, T



We examine the hedging performance for a hypothetical pension plan containing one pensioner aged 65. It pays the pensioner $1 at the beginning of each year starting from a certain age until the pensioner dies or attains age 90. For simplicity, we assume that there are no other risk sources such as credit, sampling, and basis risk. Then, the present value of the unexpected cash flows, X, from the plan is given by

$$X = V_{Liability} - E(V_{Liability}), \tag{5.1}$$

where $V_{Liability}$ is the present value of the realized liability, and $E(V_{Liability})$ is the present value of the best estimated liability. The present value of the unexpected cash flow from the hedged portfolio, i.e., the liability with additional hedging portfolios, is written as

$$X^* = V_{Liability} - E(V_{Liability}) - h[V_{Hedge} - E(V_{Hedge})], \tag{5.2}$$

where V_{Hedge} is the present value of all payoffs from a hedging instrument(here, q-forward), and h is the number of units held of the hedging instrument. Hedge effectiveness is evaluated based on the amount of longevity risk reduction(LRR) defined by

where $\sigma(X)$ is the standard deviation of portfolio X. A higher value of LRR indicates better hedge effectiveness. Figure 11 shows the distribution of the two portfolios

$$\text{Longevity Risk Reduction(LRR)} = 1 - \frac{\sigma^2(X^*)}{\sigma^2(X)}, \tag{5.3}$$

obtained by using 5,000 simulation paths. The distribution of the hedged portfolio is narrower than that of the unhedged one. Here, all cash flows were discounted at a 3% interest rate. Table 7 presents the results of hedge effectiveness assessment for three different ages. $\sigma_{unhedged}$ and σ_{hedged} are the standard deviations for the unhedged and hedged portfolios, respectively. The LRRs are 90% at age 60 and 92% at ages 65 and 70. The hedge ratio h at age 60 is 6.4, implying that the optimal number of q-forward contracts is 6.4, or rounding to the nearest whole number, 6. The high LRR values reflect the high hedging effectiveness of q-forwards against mortality risk. As shown above, mortality rate predictions and their uncertainties play key roles in the management of mortality risk using q-forwards(as well as other mortality-linked products such as S-forwards, longevity swaps, and k-forwards). The GP optimally provides age-specific risk factors for the estimates.

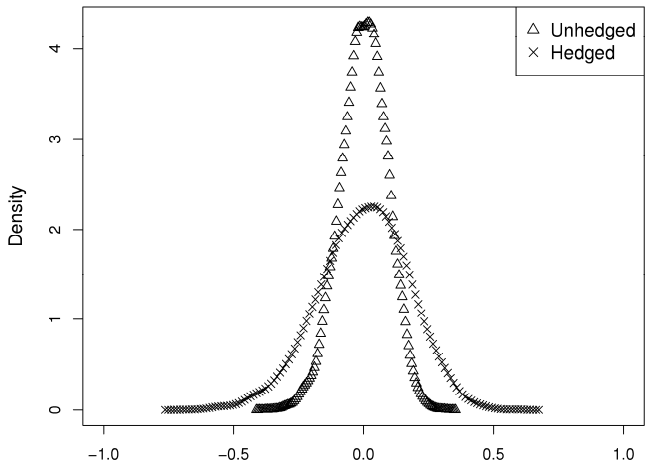
〈Table 7〉 Hedge effectiveness

Age	$\sigma_{unhedged}$	σ_{hedged}	h	LRR
60	0.0912	0.0288	6.4	90%
65	0.158	0.0433	7.2	92%
70	0.239	0.0676	8.8	92%

VI. Concluding Remarks

In this study, we constructed the GP model using mortality data for South Korean males. The six age/period functions and cohort effects are identified as the key factors for fitting the data. These factors are also shown in the GP model for the UK mortality data; however, the UK GP model has another age function for middle-age mortality between ages 55~65(Hunt and Blake, 2014), which reflects different demographic profiles between the two nations.

〈Figure 11〉 Distributions of the present values of the hedged and unhedged portfolios



The key finding of this study is that the GP model consistently outperforms the other seven models when evaluated based on both in-sample fit tests using the BIC and out-of-sample fit tests using the MAPE, achieving robustness against parameter uncertainties.

The analysis on the fluctuations of mortality indexes may be an interesting topic for future research. Hanewald(2012) and Niu and Melenberg(2013) investigate the dynamic relationships between the Lee-Carter mortality index and variables such as leading causes of death, real growth rates, and employment rates. Using the GP mortality indexes can give more detailed information on mortality dynamics, since they provide age-specific mortality indexes, unlike the Lee-Carter mortality index which provides only the overall improvement of mortality rate. Lim et al.(2014) present an analysis on the trends of leading causes of death for South Korean, which would usefully serve to investigate the research.

References

Antolin, P., "Longevity Risk and Private Pensions. OECD working papers on insurance and private pensions", *OECD publishing*, No. 3, 2007, doi:1.1787/261260613084.

Barrieu, P., Bensusan, H., El Karoui, N., Hillairet, C., Loisel, S., Ravanelli, C., and Salhi, Y., "Understanding, modelling and managing longevity risk: key issues and main challenges", *Scandinavian Actuarial Journal*, Vol. 2012, No. 3, 2012, pp. 203-231.

Bauer, D., Boerger, M., and Russ, J., "On the pricing of longevity-linked securities", *Insurance. Mathematics and Economics*, Vol. 46, No. 1, 2010, pp. 139-149.

Blake, D and Burrows, W. "Survivor Bonds: Helping to Hedge Mortality Risk", *Journal of Risk and Insurance* , Vol. 68, 2001, pp. 339-348.

Booth, H., Maindonald, J., and Smith, L., "Applying Lee-Carter under conditions of variable mortality decline", *Population Studies*, Vol. 56, 2002, pp. 325-336.

Brouhns, N., Denuit, M., and Vermunt, J. K., "A Poisson Log-Bilinear Regression Approach to the Construction of Projected Lifetables", *Insurance: Mathematics and Economics*, Vol. 31, 2002, pp. 373-393.

Cairns, A. J. G., Blake, D., and Dowd, K., "A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration", *Journal of Risk and Insurance* , Vol. 73, No. 4, 2006a, pp. 687-718.

_____, "Pricing Death: Frameworks for the Valuation and Securitization of Mortality Risk", *ASTIN Bulletin*, Vol. 36, No. 1, 2006b, pp. 79-120.

Cairns, A., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A., and Balevich, I., "A quantitative comparison of stochastic mortality models using data from England and Wales and the United States", *North American*

Actuarial Journal, Vol. 13, No. 1, 2009, pp. 1-35.

Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., and Khalaf-Allah, M., "Mortality density forecasts: an analysis of six stochastic mortality models", *Insurance: Mathematics and Economics*, Vol. 48, 2011, pp. 355-367.

Chan, W. S., Li, J. S. -H., and Li, J., "The CBD mortality indexes: Modeling and applications", *North American Actuarial Journal*, Vol. 18, No. 1, 2014, pp. 38-58.

Coughlan, G. D., Epstein, D., Shinha, A., Honig, P., "q-Forwards: Derivatives for transferring longevity and mortality risk", J.P. Morgan: London, 2007.

Currie, I. D., "Smoothing and forecasting mortality rates with P-splines", Presentation to the Institute of Actuaries, 2006.

Dowd, K., Cairns, A.J.G., Blake, D., Coughlan, G.D, Epstein, E., and Khalaf-Allah, M., "Backtesting Stochastic Mortality Models", *North American Actuarial Journal*, Vol. 14, 2010, pp. 281-298.

Girosi, F. and King, G., "A reassessment of the Lee-Carter mortality forecasting method", Working Paper. Harvard University, 2005.

Hanewald, K. "Explaining Mortality Dynamics", *North American Actuarial Journal*, Vol. 15, No. 2, 2012, pp. 290-314.

Hári, N., Waegenaere, A., Melenberg, B., and Nijman, T., "Estimating the term structure of mortality", *Insurance: Mathematics and Economics*, Vol. 42, 2008, pp. 492-504.

Hunt, A. and Blake, D., "A general procedure for constructing mortality models", Vol. 18, No. 1, 2014, pp. 116-138.

Koissi, M., A. Shapiro, and G. Hognas, "Evaluating and extending the Lee-Carter model for mortality forecasting: bootstrap confidence interval", *Insurance: Mathematics and Economics*, Vol. 38, 2006, pp. 1-20.

Lee, R. D. and Carter, L. R., "Modeling and forecasting US mortality", *Journal of the*

American Statistical Association, Vol. 87, 1992, pp. 659-675.

Lee, R. D. and Miller, T., "Evaluating the Performance of the Lee-Carter Approach to Modeling and Forecasting Mortality", *Demography* Vol. 38, No. 4, November 2001 , pp. 537-549.

Lim, D., Ha, M., and Song, I., "Trends in the Leading Causes of Death in Korea, 1983-2012", *J Korean Med Sci.*, Vol. 29, No. 12, 2014, pp. 1597-1603.

Loeys, J., Panigirtzoglou, N., and Ribeiro, R. M., "Longevity: A Market in the Making", J. P. Morgan, 2007.

Niu, G., and Melenberg, B., "Trends in mortality decrease and economic growth", DP 11/2013-071, 2013.

O'Hare C, Li, Y., "Explaining young mortality", *Insurance: Mathematics and Economics*, Vol. 50, No. 1, 2012, pp. 12-25.

Plat, R., "On stochastic mortality modeling", *Insurance: Mathematics and Economics*, Vol. 45, No. 3, 2009, pp. 393-404.

Preston, S. H., In *Forecasting the Health of Elderly Populations*, Manton K G, Singer B H, Suzman R M, editors. New York, Springer, 1993, pp. 51-77.

Renshaw, A.E. and Haberman, S., "A cohort-based extension to the Lee-Carter model for mortality reduction factors", *Insurance: Mathematics and Economics*, Vol. 38, 2006, pp. 556-570.

Smith, D. W., *Human Longevity*, Oxford Univ. Press, New York, 1993.

Tan, C. I., Li, J., Li, J. -H., Balasooriya, U., "Parametric mortality indexes: From index construction to hedging strategies", *Insurance: Mathematics and Economics*, Vol. 59, 2014, pp. 285-299.

Tulijapurkar, S., "Mortality declines, longevity risk and aging", *Asia-Pacific Journal of Risk and Insurance*, Vol. 3, No. 1, 2008, pp. 37-51.

요 약

확률적 사망모형(stochastic mortality models)은 미래 사망률 예측에 널리 사용된다. 사망률 데이터 적합(fitting)을 위해 다양한 형태의 모형들이 제시됐지만, 범용성과 최적화 측면에서 한계를 지닌다. 최근에 소개된 일반적 과정(general procedure)은 사망률 데이터로부터 연령에 따른 모든 모수적 나이 함수(parametric age functions)를 단계적으로 추출하여 모델링하므로 기존의 한계를 극복할 수 있다. 이에 본 연구는 일반적 과정의 이러한 특징이 사망률 예측에 어떻게 구현되는지 살펴보고자 한다.

1983년부터 2010년까지 한국인의 사망률 통계를 이용해 일반적 과정과 널리 사용되는 7개의 확률적 사망모형의 표본 내 적합도(in-sample fit)와 표본 외 예측력(out-of-sample forecasting)을 비교 평가했다. 표본 내 적합도 검정은 베이지안 정보 기준(Bayesian information criterion)을, 표본 외 예측력 검정은 절대 평균오차 비율(mean absolute percentage error)을 이용했다. 세 개의 연령 집단(1~79세, 11~79세, 60~79세)에 대한 검정 결과는 일반적 과정의 우수한 수행력을 보여주었다. 이는 사망률 예측을 위해 필요한 최적 변수들이 일반적 과정의 단계적 분석에 의해 검출될 수 있고, 나이 함수에 대한 모수적 접근은 과적합(overfitting) 문제를 줄여 예측 결과의 신뢰성을 높일 수 있음을 보여준다. 추정된 모수적 나이함수와 그에 대응하는 비모수적 기간함수(nonparametric period functions), 코호트 효과(cohort effects)는 사회경제적 그리고 인구통계학적 의미를 내포하므로 사망률 추세를 이해하는 데 있어 중요하다. 또한, 일반적 과정의 다중 리스크 인자(multiple risk factors)는 연금, 보험시장에서 미래 사망률에 연관된 리스크를 추정하고 관리하는 데 유용하게 사용될 수 있다. 활용사례로 연금 시장에서 장수위험(longevity risk)을 살펴봤다. 즉 연금 지급자의 불확실한 미래부채(future liabilities)를 추정하여 q-선물(q-forwards)을 이용한 헤지(hedge) 전략을 논했다.

※ 국문 색인어: 일반적 과정, 확률적 사망 모형, 표본 외 예측력, q-선물, 장수위험