# A Synthesis Mortality Model for the Elderly

Karen C. Su[★]
Department of Insurance and Finance
National Taichung University of Science and Technology

Jack C. Yue
Department of Statistics
National Chengchi University

[★]Corresponding author: Karen C. Su, e-mail: chsu@nutc.edu.tw, Assistant Professor of Department of Insurance and Finance, National Taichung University of Science and Technology, No. 129, Sec. 3, San-Min Rd., Taichung 40401, Taiwan (R.O.C), TEL: (886)4-22196067

# A Synthesis Mortality Model for the Elderly


## Abstract

Mortality improvement has been a common phenomenon since the 20th century and the human longevity continues to prolong. Post-retirement life receives a lot of attention and the need for modelling mortality rates of the elderly (ages 65 and beyond) is essential because life expectancy has reached the highest level in history. Mortality models can be divided into two groups: relational and stochastic models, but there is no consensus which model is better in modelling the elderly's mortality rates. In this study, instead of choosing either relational or stochastic models, we propose a synthesis model, selecting and modifying appropriate models from both groups, which not only has satisfactory estimation result but also can be used for mortality projection. We use the data from U.S., U.K., Japan, and Taiwan to evaluate the proposed approach (Data source: Human Mortality Database). We found that the proposed model performs well and is a possible choice for modelling the elderly's mortality rates.



Keywords: Longevity Risk, Relational Models, Stochastic Models, Mortality Projection, Mortality Improvement

# A Synthesis Mortality Model for the Elderly

## 1. Introduction

Mortality improvement has been well recognized by life insurers, pension providers, and government officials. As the number of elder population increases, more related financial and insurance products are created, and suitable tools are needed to price these products. There is a growing interest in modelling and projecting mortality rates of the elderly (ages 65 and beyond), among all solutions in better managing longevity risk (Bohk-Ewald and Rau, 2017; Li et al., 2011; Preston and Stokes, 2012; Thatcher et al., 1999). However, the process of modelling mortality rates of elderly population didn't go so well, because of lacking good quality and enough quantity of elderly data. For example, Taiwan has fine population registration system but the highest attained age is 100+ (ages 100 and beyond) in 1992. This indicates that only 25 years of data are available if we want to study the mortality rates of oldest-old (85+) population.

The inadequate data availability of the elderly population increases the difficulty of mortality modelling. There is no consensus about the future trend of elderly mortality and life expectancy is with/without a limit is one of the points in dispute (Carnes et al., 2003; Yue, 2012). Researchers and institutions also noticed that mortality improvement patterns of different age groups may not be the same (Bohk-Ewald and Rau, 2017; Börger and Schupp, 2018; Kannisto, 1994; Kannisto et al., 1994). Figure 1 displays the trends of mortality rates of age 30 from the U.S., U.K., Japan, and Taiwan for the period 1970-2005, whereas Figure 2 shows those of ages 65 and 80 at the same period. As mortality improvements of younger populations began to stabilize, there are still obvious decreasing trends for the elderly's mortality rates. As a result, applying stochastic mortality models, such as the Lee-Carter model (Lee and Carter, 1992), assuming that the changing rates with respect to time of all ages are similar, may not obtain required accuracy for mortality fitting or projection. We need to modify the mortality models in order to incorporate with various rates of mortality improvement at different ages.

(Please insert Figure 1 about here)

(Please insert Figure 2 about here)

Mortality models can be separated into two categories: the traditional relational models and the modern stochastic models. Traditional relational models, such as the Gompertz model, the Makeham model, and the Weibull model, often use single-year mortality data and assume that the age-specific mortality rates satisfy certain function form. The Coale-Kisker model (Coale & Kisker, 1990) and the logistic model are two other popular relational models which receive lots of attention in modelling elderly mortality. Most relational models usually provide sound mortality estimation but they are not good in prediction over time. Many researchers proposed modifications to increase the predicting accuracy of relational models. Tsai and Yang (2014) proposed a linear regression approach to relate one mortality sequence to the other of equal length. Cadena and Denuit (2016) applied the accelerated failure time model and propose a semi-parametric accelerated hazard relational model.

Stochastic models, on the other hand, use multiple-year data and focus on the time/cohort trend of mortality rates. Most of them are good in mortality prediction but may not be used for data extrapolation, i.e., predicting the mortality rates beyond highest attained ages. The LC model (Lee and Carter, 1992) probably is the most widely used stochastic model in mortality predictions and applications. The RH model (Renshaw and Haberman, 2006), generalized the LC model by including a cohort effect. The CBD model (Cairns et al., 2006), has been commonly used for modeling the mortality rates of higher ages (Cairns et al., 2009). Among these three models, only the CBD model can extrapolate the mortality rates of ages without mortality records.

More studies are on the stochastic models in recent years mainly because of the ability for their performance in modelling mortality improvement, and many modifications of the LC and CBD models are proposed, including, to name a few, Chen and Cox (2009), Haberman and Renshaw (2009), Li et al. (2009), Cox et al. (2010), Cairns et al. (2009), and Mitchell et al. (2013). To improve predictability, Brouhns et al. (2002) substituted a log-bilinear Poisson regression model for single vector decomposition (SVD) in the LC model's parameter estimation. Ahcan et al. (2013) suggested forecasting mortality for small populations by mixing appropriately the mortality data obtained from other populations. Tsai and Lin (2017) incorporated Bühlmann credibility into mortality models to improve forecasting performances. However, as mentioned previously, most of the analysis results of stochastic models are limited to the age range of the historical data. For ages beyond the given sample age range, stochastic models may not provide results for mortality estimation and projection.

The main goal of this study is to propose a mortality model for post-retirement populations, ages 65 and over. Instead of setting up a new model, we aim to combine existing mortality models with accurate and stable results in mortality estimation and/or prediction. In particular, we first choose a stochastic model as the basis and modify it with a relational model. We hope that the proposed synthesis model can be used to extrapolate mortality rates of higher ages, as well as possessing good performance in estimation and prediction. In this study, we choose LC and CBD models for the group of stochastic model and use the Gompertz, Coale-Kisker, and logistic models for the group of relational model.

The concept of combining two different models is used quite often in constructing life tables (or mortality graduation), particularly for graduating the mortality rates for the elderly. For example, it is believed that the elderly data quality is more reliable from the Medicare program, U.S. Social Security Administration, and the mortality rates for ages 85 and over were produced based on the Medicare data. In constructing the 1979-81 U.S. Life Table, the mortality rates of ages 85 and over were decided by the following formula (Brown, 1993),

$$q_x = \frac{1}{11}\left[(95 - x)q_x^C + (x - 84)q_x^M\right]. \quad (1)$$

$q_x^C$ and $q_x^M$ are the estimated mortality rates of age $x\,(85 \le x \le 94)$ from the population census and Medicare, respectively, and more weights are on the Medicare mortality rates for higher ages. This simple but useful weighted formula provides smooth and reliable estimates between two models, provided that they produce smooth and accurate estimates and the weights are properly chosen. This formula also motivates the proposed synthesis model.

There are other methods for handling the case of populations with limited data, including survivor ratio method and partial standard mortality ratio (SMR) method (Lee, 2003; Thatcher et al., 2002; Wang et al., 2018). Terblanche (2016) proposed retrospective tests based on a number of extrapolative methods to forecast mortality rates of elderly populations in Australia. For a population with one million or more, graduation methods such as Whittaker and LC model can produce stable mortality estimates at ages other than the elderly (Wang et al., 2018). We often apply relational models to acquire mortality estimates for the elderly. For example, Taiwan government uses a two-stage estimation to construct official life tables.[1] At the first stage, Whittaker method is used to graduate mortality rates for ages 0-79 and the Gompertz model is applied to acquire mortality rates for ages 65 and over. At the second stage, the mortality rates at ages 65-79 are determined by the linear combination of Whittaker and Gompertz estimates, larger weights are on the Gompertz estimates for higher ages.

This paper is organized as follows. Section 2 provides a brief review of both relational models and stochastic model applied in the paper. Section 3 describes the process of constructing the synthesis models and compares the proposed model with existing models. In section 4, we apply the forecasted mortality rates to calculate the premium of a pure annuity. Final conclusion and discussions are given in Section 5.

[1] Source: Ministry of Interior, Department of Statistics http://www.moi.gov.tw/stat

## 2. Review of Mortality Models

We will first give a brief introduction of mortality models used in this study (Table 1), following by the description of proposed approach. For the group of stochastic models, we choose the LC and the CBD models. Both are popular and well-known for providing good mortality estimation and prediction. For the group of relational models, we select the Gompertz, the Coale-Kisker, and the logistic model. These models are known as candidates of providing sound fitting results for the elderly's mortality rates. We introduce the stochastic models first.

(Please insert Table 1 about here)

Among all stochastic mortality models, the simplicity and accuracy of the LC model makes it the most popular and widely used. Lee and Carter (1992) assumed that the central death rate of an individual aged-*x* at year *t* follows the form:

$$\ln m_{x,t} = \alpha_x + \beta_x \kappa_t + \varepsilon_{x,t} \,, \tag{2}$$

where

$\alpha_x$ describes the average age pattern of mortality over time,

$\beta_x$ is the deviations from the average pattern,

$\kappa_t$ describes the variation in the level of mortality over time, and

$\varepsilon_{x,t}$ is the error term.

The parameters are subject to constraints

$$\sum \kappa_t = 0 \text{ and } \sum \beta_x = 1, \tag{3}$$

to ensure model identification. When forecasting mortality rates, it is assumed that $\alpha_x$'s and $\beta_x$'s remain constant over time and the values of $\kappa_t$ are modeled by random walk with drift:

$$\kappa_t = \kappa_{t-1} + \phi + e_t \tag{4}$$

where $e_t \sim N(0, \sigma_{LC}{}^2)$, and $\phi$ is known as the drift parameter. The parameters can be estimated via singular value decomposition (SVD), weighted least square, the maximum likelihood estimation, or approximation method, if there are missing data.

The CBD model, proposed by Cairns et al. (2006), was designed to model mortality rates for higher ages. Assuming that there are two time trends, $k_{C1,t}$ and $k_{C2,t}$, the model, for given age range [$x_1$, $x_1$+$n$-1], and time period [$t_1$, $t_1$+$K$–1], is given by:

$$\text{logit}(q_{x,t}) = \ln\left(\frac{q_{x,t}}{1-q_{x,t}}\right) = k_{C1,t} + k_{C2,t}(x-\bar{x}) + \varepsilon_{x,t}, \tag{5}$$

where

$$\bar{x} = \frac{\sum_{x=x_1}^{x_1+n-1} x}{n} \tag{6}$$

and $q_{x,t}$ is the mortality rate of an individual aged $x$ in year $t$.

The errors are assumed to be independent and identically distributed and the two time trends are modeled by a bivariate random walk with drift:

$$\begin{cases} k_{C1,t} = k_{C1,t-1} + \phi_{C1} + e_{C1,t} \\ k_{C2,t} = k_{C2,t-1} + \phi_{C2} + e_{C2,t} \end{cases}. \tag{7}$$

The Gompertz law probably is a well-known relational model and is often applied to acquire mortality rates for the elderly. The Gompertz law of mortality assumes that the force of mortality at age $x$ takes the form of $\mu_x = BC^x$, where $B > 0$ and $C > 1$. Under this assumption, the probability that an individual at age $x$ would survive to age $x$+1, denoted by $p_x$, is

$$p_x = \exp\left[-BC^x(C-1)/\ln C\right]. \tag{8}$$

Let $l_x$ be the number of individuals alive at age $x$, then the amount that will survive one year later would be

$$l_{x+1} = l_x \cdot \exp\left[-BC^x(C-1)/\ln C\right]. \tag{9}$$

From eq. (7), we also have

$$\ln p_{x+1} / \ln p_x = C \,. \tag{10}$$

Therefore, if the mortality pattern of the observed population follows the Gompertz low, the ratios of $\ln p_{x+1}$ over $\ln p_x$ shall remain fairly constant.

The CK model (Coale and Kisker, 1990) assumes that the central death rate can be modeled as

$$m_x = m_{x-1} \cdot \exp\left[ k_{K,85} + (x-85) \cdot s \right], \text{ for } x \geq 85 \text{, where} \tag{11}$$

$$s = -\frac{\ln\left(\frac{m_{84}}{m_{110}}\right) + 26 k_{K,85}}{325} \quad \text{and} \quad k_{K,85} = \ln\left(\frac{m_{85}}{m_{84}}\right). \tag{12}$$

Given that $m_{110}$ is set at 1.0/0.8 for men/women, the mortality rates above age 85 can be obtained recursively and allow us carry out the extrapolation of the age-extended mortality table. A further revised method used weighed least square on the following quadratic form related to the central death rate:

$$\min_{\alpha_K, s_K} \sum_x w_x \left[ \ln m_x - \alpha_K - k_{K,85}(x-84) + \frac{(x-84)(x-85)}{2} s_K \right]^2 \tag{13}$$

The logistic model was proposed by Perks (1932), who found empirically that the values of force of mortality in a life table which he was examining could be fitted by a logistic function. In other words, the population growth rate declines with population numbers and reaches a limit. In the paper, we assume that the logarithm of $p_x$, the probability that an individual at age $x$ would survive to age $x$+1, is modeled as the following:

$$\eta_x = -\ln p_x = \frac{\exp\left[ \beta_{L,0} + \beta_{L,1} \cdot (x+0.5) \right]}{1 + \exp\left[ \beta_{L,0} + \beta_{L,1} \cdot (x+0.5) \right]}. \tag{14}$$

## 3. The Proposed Synthesis Model

Before introducing the process of combining two mortality models, we notice that mortality models may not use the same target variables, and they can be death rate, central death rate, or force of mortality. The Gompertz model is based on the assumption of the force of mortality, whereas the LC model is based on the central death rate and the CBD model is on death rates. Thus, we need to modify all models to have the same target variable. By definition, the central death rate of a person aged-$x$ at year $t$, denoted by $m_{x,t}$, is calculated as follows:

$$m_{x,t} = \frac{d_{x,t}}{L_{x,t}} = \frac{d_{x,t}}{\int_0^1 l_{x+y,t+y-1} dy}, \tag{15}$$

where $d_{x,t} = l_{x,t} - l_{x+1,t+1}$ is the number of death between age $x$ and age $x$+1 at year t, and $L_{x,t}$ is the exposure, or number of person-years lived between age $x$ and $x$+1 at year t. We first assume that the force of mortality follows the Gompertz assumption, $\mu_{x,t} = B_t C_t^x$, at year t.[2] Assuming that the force of mortality is uniformly distributed within year $t$, we further approximate the exposure (or stationary population) $\mathrm{L}_{x,t}$ of age $x$ at time $t$, under the uniform distribution of death, via

$$L_{x,t} \cong \frac{l_{x,t} + l_{x+1,t}}{2} = \frac{l_{x,t}(1+p_{x,t})}{2}. \tag{16}$$

Then we have

$$m_{x,t} = \frac{d_{x,t}}{L_{x,t}} = \frac{2\left(1-p_{x,t}\right)}{\left(1+p_{x,t}\right)}. \tag{17}$$

Given that $p_{x,t} = \exp\left[-B_t C_t^x (C_t - 1)/\ln C_t\right]$ from the Gompertz model, we can derive the following estimation for the central death rate as follows:

---

[2] Notice that the Gompertz parameters now have subscripts $t$. Instead of setting a fixed Gompertz parameter for the whole period, we estimate the parameter on a year-to-year basis.

$$\ln\left(-\ln\left(\frac{1-\frac{m_{x,t}}{2}}{1+\frac{m_{x,t}}{2}}\right)\right)=\ln\left(B_t(C_t-1)/\ln C_t\right)+x\ln C_t\,. \tag{18}$$

We will combine both the advantages of relational and stochastic mortality models to build the proposed model. The process of constructing synthesis mortality model is similar to that of constructing life tables in Taiwan. The relational models have a better fit of the elderly's mortality rates may not have good estimation for all ages generally. This is the reason for using the Gompertz law for the mortality rates at only ages 65 and over in Taiwan. Thus, we will use the stochastic models for younger ages, apply relational models for the older ages, and design a method to connect these two models in between. In other words, we use two different mortality models before and after a chosen age. The chosen age, namely the transition age (Figure 3), is set so that the morality curve will be separated into two parts: the stochastic model before the transition age and the synthesis model after that.

(Please insert Figure 3 about here)

(Please insert Figure 4 about here)

Figure 4 briefly describes the process of proposed synthesis model. The estimation of stochastic model is performed first, followed by that of relational model for mortality rates of ages over the transition age.

The model combining the LC model and the Gompertz law is called the "LC-G" model. Similarly, the one combining the LC model and the CK/logistic model is referred as the LC-CK/LC-Log model, while that of combining the CBD model and the Gompertz law, the CK model and the logistic model are referred as the CBD-G, the CBD-CK and the CBD-Log model, respectively.

## 4. Empirical Study and Applications

### 4.1 Empirical data and determining the transaction ages

We use empirical data to evaluate the proposed synthesis model described in the previous section. The mortality data used are from four countries: U.S., U.K., Japan, and Taiwan (Source: Human Mortality Database; HMD), with age range 65-99 and time period 1970-2009. We use backcasting (or cross-validation) procedure as mentioned by Dowd et al. (2010), Tsai and Yang (2015) and Tsai and Lin (2017). Since projections results are highly sensitive to calibration periods and accuracy does not consistently increase with increase length of fitting period (Berkum et al. 2016; Terblannche, 2016), we consider the backcasting procedure with 10-year training period and 5-year testing period. We assume that the training period is [$t_0$, $t_0$+9], where $t_0$ is 1970, 1975, 1980, 1985, 1990, and 1995, with 5 years overlap between two consecutive periods and the testing period is [$t_0$+10, $t_0$+14]. The data from training periods are used to construct mortality models and the fitted model is applied to the data from the testing periods.

We first look at the results of the transition age. As defined, the certain transition age point is set so that the morality curve will be separated into two parts: the stochastic model before the transition age and the combined model after that. To find the transition age for each year period [$t_0$:$t_0$+9], we assume the transition age being from 65 to 98 and repeatedly run the synthesis process, as described in Figure 4. The resulted transition age is then set as the one that minimizes the root mean square error.

(Please insert Table 2 about here)

Table 2 summarizes the transition ages of synthesis models for countries at different data periods.[3] Generally speaking, most transition ages lie at the age range of 65-80, except for the cases of Taiwan and Japan. It seems that the trend of elderly's mortality rates is different to that of younger ages and introducing relational models at the older ages may improve the model fit.

[3] We did not apply the transition age process on the Lee-Log model because estimating parameters for the logistic model requires regression on all ages.

The results of CBD model and Gompertz model require further attention, especially for the combination of CBD model and Gompertz model. For this treatment combination, the transition age is 65 in the case of U.S. data and is 98 in the case of Taiwan data (and some of Japan data). This indicates that the Gompertz model has better fits than the CBD model in the case of American elderly, but the CBD model fits better than the Gompertz model in the case of Taiwanese elderly (i.e., we don't need to consider the Gompertz model). Since the CBD and Gompertz models are designed to model the elderly's mortality rates, our results imply that the fitting performance of elderly mortality models is likely to be data dependent, which is mentioned in the previous studies (e.g., Wang et al., 2016). The combination of CBD and the Coale-Kisker model also reveals similar pattern regarding to the determination of the transition age.

Note that, in addition to applying mortality models directly, we can also check whether the empirical data support using these models. As mentioned in the previous section, the ratios of $\ln p_{x+1}$ over $\ln p_x$ shall remain fairly constant if the mortality pattern of the observed population follows the Gompertz low. The results show that we cannot reject the assumption that the mortality rates of elderly population follow the Gompertz law. Yue (2002) proposed using the bootstrap simulation to construct the confidence interval of parameter C and we can use it to verify if the elderly mortality rates follow the Gompertz assumption. The bootstrap simulation (Appendix A) suggests that the Gompertz model is a feasible assumption for the data from four countries, since the intersections of confidence intervals are not empty. We can consider similar tests for other mortality models as part of exploratory data analysis before constructing the synthesis model. We will use the empirical data to evaluate the proposed synthesis model in the next section.

### 4.2 Fitting and forecasting results

We can evaluate the performance of mortality models based on the training data (estimation) or the testing data (prediction). We consider the case of training data first. More parameters in the model usually would have smaller estimation error, but there is risk for using too many parameters (i.e., overfitting). Thus, we use the Akaike information criterion (AIC) to evaluate the model fit for the training data. AIC was proposed by Akaike (1974) and has been widely used for model selection. AIC is defined as:

$$AIC = -2\log(ML) + 2K , \tag{19}$$

where $ML$ is the maximum likelihood under the given model and $K$ is the number of fitted parameters. If the errors of mortality models are assumed to be normally distributed with constant variance, then AIC can be also computed from the least squares regression (Burnham and Anderson, 2002), i.e., redefining the AIC as:

$$AIC = N\log(\frac{RSS}{N}) + 2K , \tag{20}$$

where $RSS$ is the residual sum of squares,

$$RSS = \sum_{j=1}^{T}\sum_{i=0}^{n-1}(\hat{q}_{x+i,j} - q_{x+i,j})^2 \tag{21}$$

and N is the number of observations and K is the total number of estimated regression parameters. Smaller AIC values indicate better performance in estimation. Given that we use the weighted least squares to estimate parameters of our combined models, we applied equation (20) for the calculation of the AIC values of all models, as shown in Figure 5. In general, the synthesis models tend to have smaller AIC values, comparing to those without combining the relational models, although additional parameters are added. For example, the LC+CK model has smaller AIC values than the LC model. However, no models have the smallest AIC values in all cases. Detailed results are in Table 4.

(Please insert Figure 5 about here)

(Please insert Table 4 about here)

Of course, we can use the estimation errors to evaluate the mortality models, such as in terms of root mean square error (RMSE) and mean absolute percentage error (MAPE), which are defined as:

$$RMSE = \frac{1}{T}\sum_{j=1}^{T}\left[\frac{1}{n}\sum_{i=0}^{n-1}(\hat{q}_{x+i,j} - q_{x+i,j})^2\right]^{1/2} \tag{22}$$

and

$$MAPE = \frac{1}{T}\sum_{j=1}^{T}\left[\frac{1}{n}\sum_{i=0}^{n-1}\left|\hat{q}_{x+i,j} - q_{x+i,j}\right|\right], \tag{23}$$

where $T$ is the time period, $n$ is the age range, $\hat{q}_{x+i,j}$ is the observation value and $q_{x+i,j}$ is the estimated value.

We divide the results into two age groups: [65, 84] and [85, 99], for younger and older elderly mortality. Detailed results are in Appendix B-1 and B-2. In terms of estimation errors, the synthesis models have smaller errors for the age group of [85, 99] but the LC and CBD models have smaller errors for the group [65, 84]. The estimation errors generally are larger for the group [85, 99], almost ten times as large as those for the group [65, 84].

It seems that the fitting performance of models is data-dependent and we cannot find a single model that dominates the others. Still, we look for the mortality model(s) which have fitted well in most countries, and we apply the voting method to determine. The voting is to count the scores for each data set, and the model has the best performance gets 3 points and the second/third best has 2/1 points, respectively. Out-performance scores of overall fitting are based on RMSE and MAPE (Table 5). Given that the total out-performance score is 48 points for each comparison, with 4 treatment combinations, models with scores over 32 points can be considered as a performance better than average. The CBD+Gompertz and CBD+CK models are the best among all models.

(Please insert Table 5 about here)

(Please insert Table 6 about here)

We can follow the same procedure to evaluate the forecasting performance. We apply the models, acquired from the training data, to the testing data. Again, we use the RMSE and

MAPE as the criteria for comparison, and the calculation results are shown in Appendix B-3 and B-4. Unlike the results of fitting errors, the synthesis models have smaller forecasting errors, but no mortality models can outperform other models. Thus, we also use the voting method for overall comparison (Table 6). Again, the CBD+Gompertz and CBD+CK models are the best among with respect to forecasting errors. It should be noted that, among all other models, the LC and CBD models have fine performance in fitting for ages 65-84, whereas the CBD+Gompertz and CBD+CK models performs well in fitting and forecasting for the oldest-old (ages 85-99).

One of our motivations for introducing synthesis model is extrapolation beyond highest recorded ages, and thus we will illustrate the extension of mortality rates for age 100 and over. Because not all countries have data beyond age 100, we select U.S. and Japan to demonstrate our approach. Also, due to the data availability, only two data periods are considered: [1990, 1999] and [1995: 2004] for training periods, [2000, 2004] and [2005: 2009] for testing periods. We estimate parameters of mortality models from ages 65 to 99, and apply the acquired models to ages 100 to 105. Note that the LC model is not considered here since it cannot be used to extrapolate.

Figure 6 displays the forecasting results of our synthesis models, compared with the CBD model, whereas Figure 8 shows the average mortality curves of ages 85-105. We should use only the RMSE errors to demonstrate the forecasting results. In general, the CBD+Gompertz and LC+Gompertz models have the smallest forecasting errors in almost all cases, and the average errors are smaller than 30% which is not bad for the ages 85-105. Also, as shown in Figure 6, the combined model CBD+Gompertz always has smaller RMSE than the CBD model for all 8 cases, which is equivalent to p-value = (0.5)8 ≈ 0.004 (i.e., run test). This indicates the CBD+Gompertz model is better than the CBD model for the group of ages 85+. It seems that combining the stochastic and relational models is a feasible approach. Also, it seems that the CBD model may overestimate mortality improvements for ages over 95, resulting in lower mortality curve shown in Figure 7.

(Please insert Figure 6 about here)

(Please insert Figure 7 about here)

**4.3 Pricing Annuity Products**

Other than the empirical study, we also consider applications by applying our synthesis model to annuity products. To simplify the discussion, we consider the single premium of a 10-year life annuity product that pays $1 at the end of each year. The interest rate, *i*, is 2% and the insured ages are 70, 80, and 90. We use the data from 1980~1999 for parameter estimation and the premiums are calculated for insured period 2000~2009 based on the acquired mortality models. We also calculate the actual values of the annuities based on the real mortality data of the insured period. Detailed results of premium calculations and estimated errors (reported as % of actual values) are listed in Appendix C.

Figure 8 shows the annuity premiums derived from all models, compared with the actual values based on the true mortality rates. Not surprisingly, the premiums of mortality models with lower forecasting errors (the proposed approaches) in the previous section are closer to the actual values. This means that the synthesis models can be used in practice. However, all mortality models underestimate the premiums for insured age 70 and overestimate the amount for insured age 90. This indicates that all models tend to over-estimate the mortality rates at younger ages and under-estimate those at older ages.

(Please insert Figure 8 about here)

## 5. Conclusion and Discussions

Post-retirement life has received a lot of attention in recent years. The need for forecasting the life expectancy for the elderly is essential to deal with the longevity risk and mortality models are a popular choice. However, there is no consensus which model is better in modelling the elderly's mortality rates. Also, due to the data availability, it is often necessary to interpolate mortality rates beyond the highest recorded ages. In this paper, we propose a synthesis model, combining relational and stochastic mortality models, in order to achieve a better fit of mortality rates for the elderly and mortality extrapolation.

The idea behind the synthesis is similar to the construction of life tables in Taiwan, where two different approaches are used: one for the elderly and the other for younger ages. Basically, we think that stochastic models are preferred for younger ages and relational models can be used for the elderly. In this study, we consider LC and CBD models for younger ages and Gompertz, CK, and logistic models for the older ages. We use empirical data from the Human Mortality Database (U.S., U.K., Japan, and Taiwan; years 1970-2009) to evaluate the proposed approach. The data are separated into training and testing periods, and the RMSE and MAPE are used as error criteria. The synthesis mortality models generally have smaller fitting errors (and AIC values) in the training periods and smaller forecasting errors in the testing periods.

In addition, we found that the LC and CBD models perform well for mortality rates of age 65 to 84 and the synthesis models are a good choice for older ages. The synthesis models also have stability performance for mortality extrapolation, up to age 105. It seems that the synthesis model performs especially well for the elderly and the oldest-old (i.e., ages 85 and over) and it can be used to annuity products and insurance products related to older ages. However, similar to the results of previous studies, we found that no mortality models can dominate other models at all cases and the fitting results are highly data-dependent. We only include four countries in this study, and we should consider more data in the future.

There are quite a lot of choices for the stochastic and relational models. In addition to use the historical data and the backcasting procedure to choose the feasible combination of models, we can also use the idea of exploratory data analysis to screen possible choices of models. We can adapt the idea of Yue (2002) and develop tools for checking whether the mortality data satisfy certain model assumption. For example, bootstrapping simulation can be used to construct possible range of parameter C for the Gompertz model. This can help to narrow down the possible choices of mortality models.

There is a small difference between the synthesis model and life table construction in Taiwan. We use one single connection (single transition age) between two models, while Taiwan's life tables use a transition period (15 to 20 ages). The advantage of using a transition period is that the mortality rates are smoother around the connect points and this won't create jumps at consecutive ages. This is important in calculating insurance premiums. Also, since no mortality models can dominate, we should impose possible variations (or variance) to the mortality predictions. One possible approach is to use the block bootstrap simulation and derive 1,000 or 10,000 sample paths of predictions. The variance can be calculated via these sample paths. Another possibility is to adapt Bayesian approach via MCMC (Markov Chain Monte Carlo) but it would require more computations.

**References:**

AHCAN, A., MEDVED, D., OLIVIERI, A., and PITACCO, E. 2014. Forecasting Mortality for Small Populations by Mixing Mortality Data. *Insurance: Mathematics and Economics* 54, 12-27.

AKAIKE, H. 1973. Information Theory and an Extension of the Maximum Likelihood Principle. In: Petrov, B.N., Csaki, F. (eds.) Proceedings of the Second International Symposium on Information Theory, pp. 267-281. Akademiai Kiado, Budapest.

BOHK-EWALD, C. and RAU, R. 2017. Probabilistic Mortality Forecasting with Varying Age Specific Survival Improvements. *Genus* 73(1), 1-37.

BŐRGER, M. and SCHUPP, J. 2018. Modeling Trend Processes in Parametric Mortality Models. *Insurance: Mathematics and Economics* 78, 369-380.

BROUHNS, N., DENUIT, M., and VERMUNT, J. K. 2002. A Poisson Log-Bilinear Regression Approach to the Construction of Projected Lifetables. *Insurance: Mathematics and Economics* 31, 373-393.

BROWN, R.L. 1993. Introduction to the Mathematics of Demography, Second Edition, ACTEX Publications.

BURNHAM, K.P. and ANDERSON, D.R. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Second Edition, Springer.

CADENA, M. and DENUIT, M. 2016. Semi-parametric Accelerated Hazard Relational Models with Applications to Mortality Projections. *Insurance: Mathematics and Economics* 68, 1-16

CAIRNS, A. J. G., 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, A. J. G., BLAKE, D., DOWD, K., COUGHLAN, G. D., EPSTEIN, D., ONG, A., and BALEVICH, I. 2009. A Quantitative Comparison of Stochastic Mortality Models

Using Data from England and Wales and the United States. *North American Actuarial Journal* 13, 1-35.

CARNES, B.A., OLSHANSKY, S.J., and GRAHN, D. 2003. Discussed Biological Evidence for Limits to the Duration of Life. *Biogeography* 4(1), 31-45.

CHEN, H. and COX, S. H. 2009. Modeling Mortality with Jumps: Applications to Mortality Securitization. *Journal of Risk and Insurance* 76, 727-751.

COALE, A. and KISKER, E. E. 1990. Defects in Data on Old-Age Mortality in the United States: New Procedures for Calculating Mortality Schedules and Life Tables at the Highest Ages. Asian and Pacific Population Forum 4(1), 1-31.

COX, S. H., LIN, Y., and PEDERSEN, H. 2010. Mortality Risk Modeling: Applications to Insurance Securitization. *Insurance: Mathematics and Economics* 46, 242-253.

DOWD, K., CAIRNS, A. J. G., BLAKE, D., COUGHLAN, G.D., EPSTEIN, D., and KHALAF-ALLAH, M. 2010. Backtesting Stochastic Mortality Models: An Ex Post Evaluation of Multiperiod-Ahead Density Forecasts. *North American Actuarial Journal* 14, 281-298.

HABERMAN, S. and RENSHAW, A. 2009. On Age-Period-Cohort Parametric Mortality Rate Projections. *Insurance: Mathematics and Economics* 45, 255-270.

KANNISTO, V. 1994. Development of Oldest-Old Mortality, 1950-1990: Evidence from 28 Developed Countries. Odense Monographs on Population Aging, vol. 1. Odense University Press, Odense, Denmark.

KANNISTO, V., LAURITSEN, J., THATCHER, A. R. and VAUPEL, J. W. 1994. Reductions in Mortality at Advanced Ages: Several Decades of Evidence from 27 Countries. *Population and Development Review* 20(4), 793-810.

LEE, R.D., and CARTER, L. R. 1992. Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association* 87(419), 659-671.

LEE, W. 2003. A Partial SMR Approach to Smoothing Age-Specific Rates. *Annuals of*

*Epidemiology*, 13(2), 89-99.

LI, J. S.-H., CHAN, W. and CHEUNG, S. 2011. Structural Changes in the Lee-Carter Mortality Indexes: Detection and Implications. *North American Actuarial Journal* 15, 13–31.

MITCHELL, D., BROCKETT, P., MENDOZA-ARRIAGA, R. & MUTHURAMAN, K. 2013. Modeling and Forecasting Mortality Rates. *Insurance: Mathematics and Economics* 52, 275-285.

RENSHAW, A. E. and HABERMAN, S. 2006. A Cohort-Based Extension to the Lee-Carter Model for Mortality Reduction Factors. *Insurance: Mathematics and Economics* 38, 556-570.

PRESTON, S. H., and STOKES, A. 2012. Sources of Population Aging in More and Less Developed Countries. *Population and Development Review*, 38(2), 221–236.

TERBLANCHE, W. 2016. Retrospective Testing of Mortality Forecasting Methods for the Projection of Very Elderly Populations in Australia. *Journal of Forecasting*, 35, 703-717.

THATCHER, A.R., KANNISTO, V., and VAUPEL, J. W. 1999. The Force of Mortality at Ages 80 to 120. Odense Monographs on Population Aging, vol. 5. Odense University Press, Odense, Denmark.

THATCHER, A.R., KANNISTO, V., and ANDREEV, K. 2002. The Survivor Ratio Method for Estimating Numbers at High Ages. *Demographic Research*, 6(1), 2-15

TSAI, C. C. L. and YANG, S. 2015. A Linear Regression Approach to Modeling Mortality Rates of Different Forms. *North American Actuarial Journal* 19, 1-23.

TSAI, C. C. L. and LIN, T. 2017. Incorporating the Bühlmann Credibility into Mortality Models to Improve Forecasting Performances. *Scandinavian Actuarial Journal* 2017/5, 419-440.

WANG, H., YUE, C. J., and CHEN, Y. 2016. A Study of Elderly Motyality Models.

*Journal of Population Studies* 52, 1-42.

WANG, H., YUE, C. J., and CHONG, C. 2018. Mortality Models and Longevity Risk for Small Populations. *Insurance: Mathematics and Economics*, 78, 351-359

YUE, C. J. 2002. Oldest-Old Mortality Rates and the Gompertz Law: A Theoretical and Empirical Study Based on Four Countries. *Journal of Population Studies* 24, 33-57.

YUE, C. J. 2012. Mortality Compression and Longevity Risk. *North American Actuarial Journal* 16(4), 434-448.

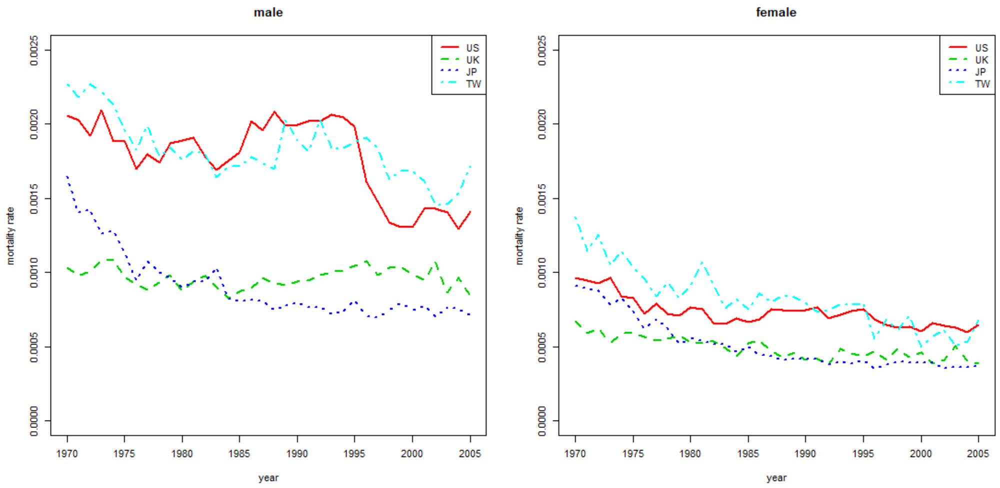
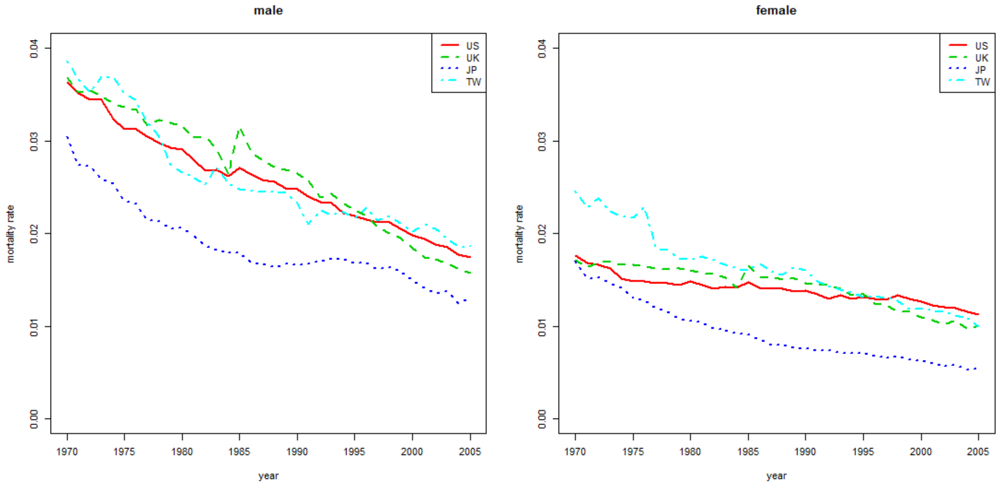
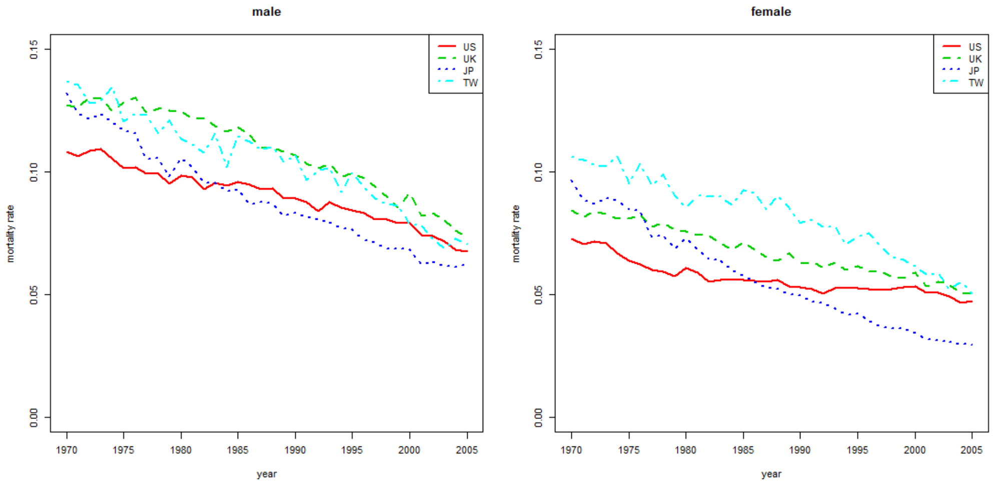



Figure 1. Trend of mortality rates of age 30 for various countries (1970-2005)




(a) Mortality rates of age 65




(b) Mortality rates of age 80

Figure 2. Trends of mortality rates of ages 65 and 80 for various countries (1970-2005)

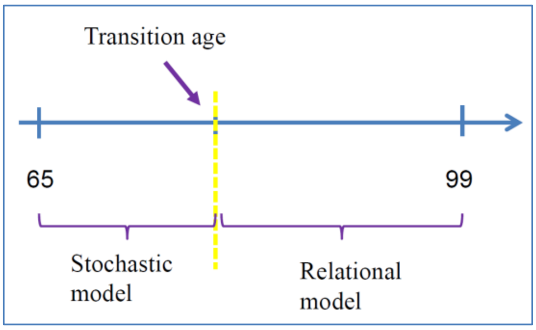


Figure 3. The transition age after which we apply the synthesis process

Step 1: Estimate the stochatic parameters using origianal mortality data of the training period

Step 2: Forecast future mortality rates of testing period using stochastic parameters

Step 3: Assuming that the transition age being from 65 to 98, estimate parameters of relational model based on mortality paths of training period generated by the stochastic model repeatedly. The transition age is set as the certain age that minimized RMSE.

Step 4: For each year of the testing period, estimate parameters of relational model after the transition age.

Figure 4. The synthesis process

**註解 [c1]:** 如果我沒弄錯意思，應該是分別使用隨機及關係模型（都是估計、不是預測！），再依照線性加權確定 transition age。現在這個寫法有點難懂，建議修改這個圖形，或甚至以口頭敘述，不需要這個圖形。
步驟 1、分別套入各一個隨機、關係模型，計算死亡率的估計值及誤差。
步驟 2、以線性加權計算步驟 1 的合成模型，在不同 transition age 的死亡率估計值及誤差。
步驟 3、比較隨機、關係、合成模型的估計誤差，如果合成模型誤差最小紀錄最佳模型的 transition age。

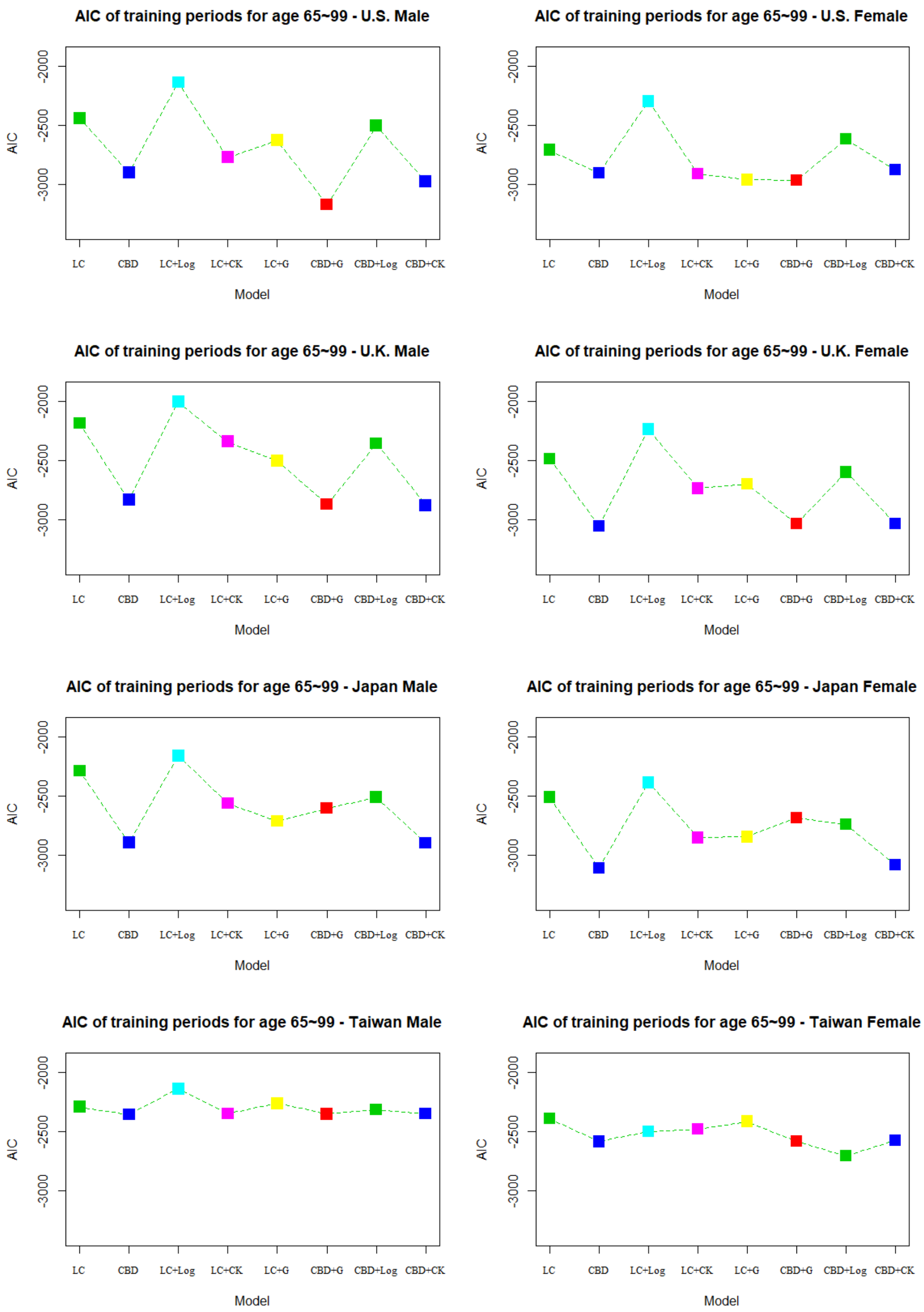



Figure 5. The AIC values of each countries using 1970-2009 data

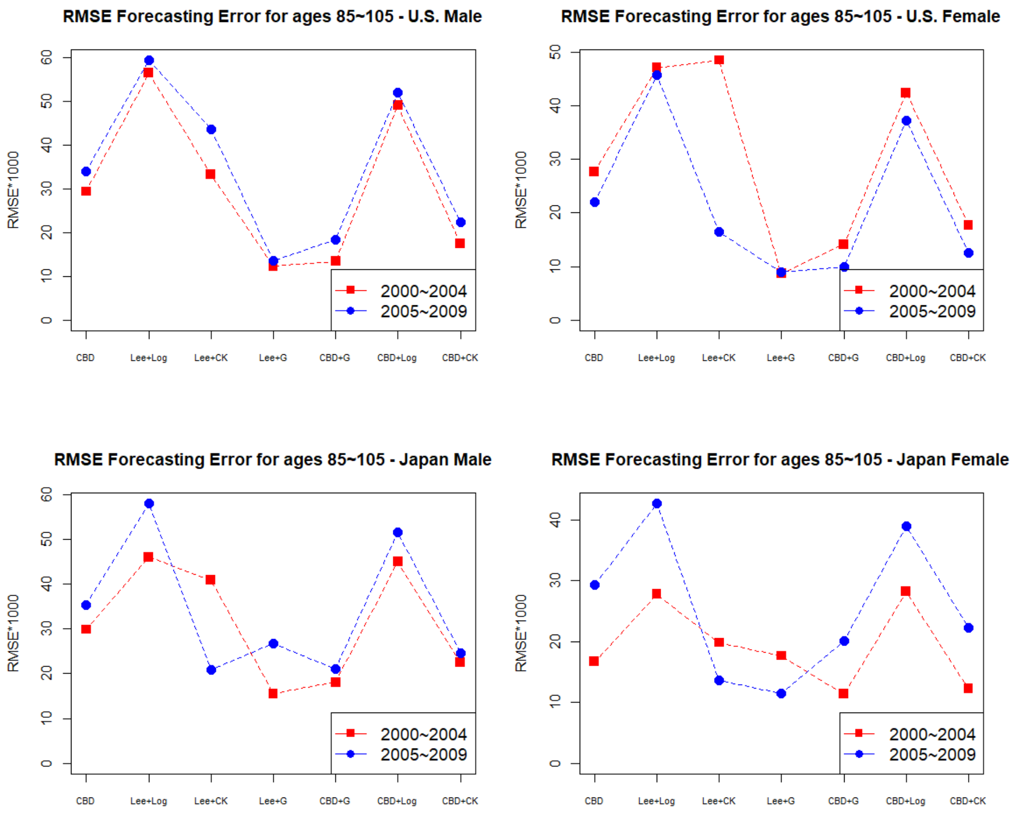



Figure 6. RMSE forecasting error for ages 85~105

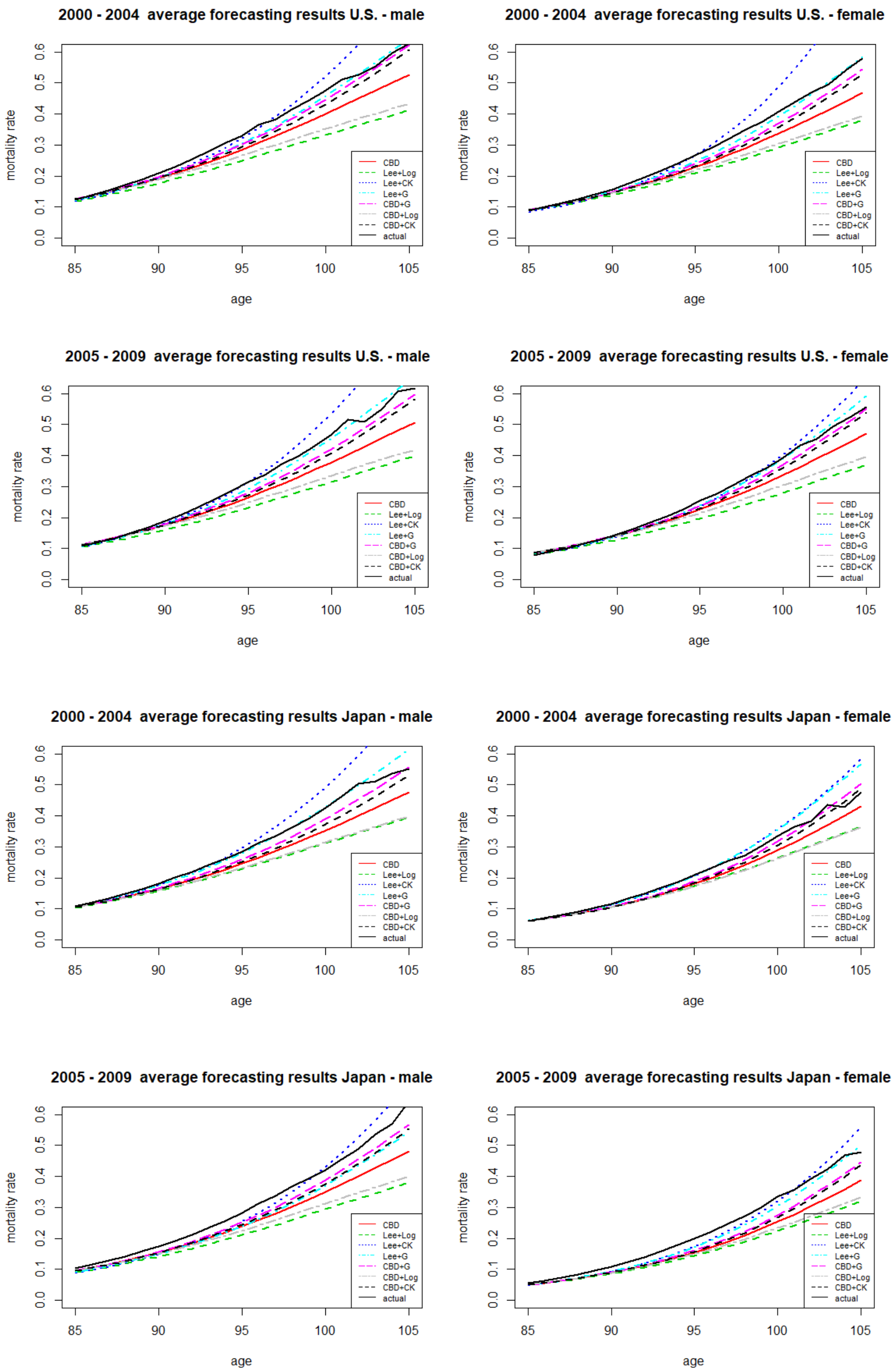



Figure 7. 5-year average mortality curve of ages 85-105 (CBD and synthesis models)

Note: The black solid lines are the observed mortality rates.

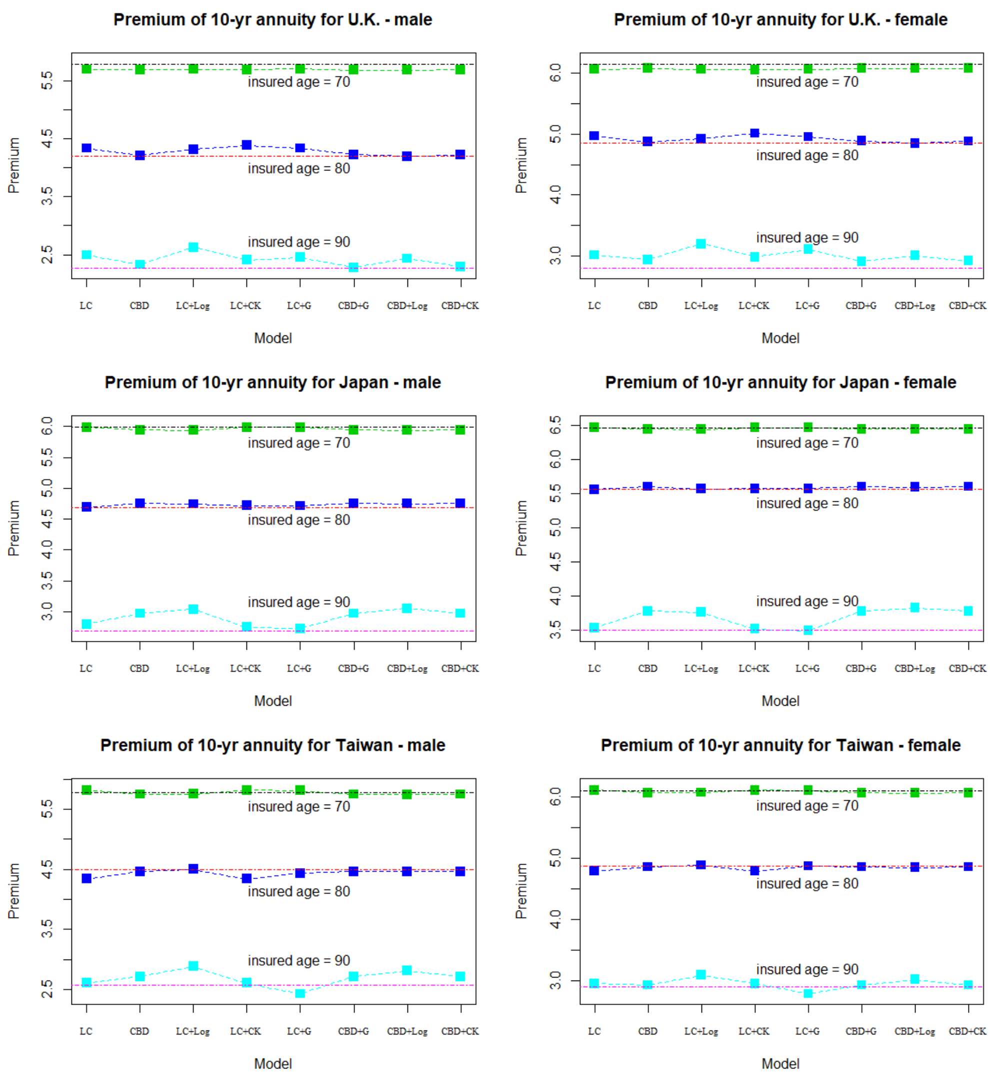



Figure 8. Single premium of 10-year pure annuity for insured ages 70, 80, and 90

Table 1. List of mortality models used in this study

| Model | Target variable | Assumption |
|---|---|---|
| LC | Central death rate | $\ln m_{x,t} = \alpha_x + \beta_x \kappa_t + \varepsilon_{x,t}$ |
| CBD | Mortality rate | $\text{logit}(q_{x,t}) = \ln\left(\frac{q_{x,t}}{1-q_{x,t}}\right) = k_{C1,t} + k_{C2,t}(x-\bar{x}) + \varepsilon_{x,t}$, |
| Gompertz | Force of mortality | $\mu_x = BC^x$ |
| CK | Central death rate | $\min_{\alpha_K, s_K} \sum_x w_x \left[\ln m_x - \alpha_K - k_{K,85}(x-84) + \frac{(x-84)(x-85)}{2} s_K\right]^2$ |
| Logistic | Mortality rate | $\eta_x = -\ln p_x = \frac{\exp\left[\beta_{L,0} + \beta_{L,1} \cdot (x+0.5)\right]}{1+\exp\left[\beta_{L,0} + \beta_{L,1} \cdot (x+0.5)\right]}$ |

Table 2. Transition ages for the synthesis models

| | LC + CK | | | | LC + Gompertz | | | |
|---|---|---|---|---|---|---|---|---|
| Data period | U.S. | U.K. | Japan | Taiwan | U.S. | U.K. | Japan | Taiwan |
| 1970-1979 | 77 | 77 | 67 | 85 | 78 | 65 | 75 | 70 |
| 1975-1984 | 77 | 77 | 73 | 96 | 77 | 65 | 75 | 71 |
| 1980-1989 | 77 | 69 | 76 | 96 | 77 | 68 | 76 | 75 |
| 1985-1994 | 75 | 68 | 74 | 71 | 77 | 75 | 77 | 74 |
| 1990-1999 | 75 | 68 | 75 | 77 | 77 | 65 | 75 | 66 |
| 1995-2004 | 66 | 66 | 75 | 74 | 66 | 66 | 75 | 74 |
| | CBD + CK | | | | CBD + Gompertz | | | |
| Data period | U.S. | U.K. | Japan | Taiwan | U.S. | U.K. | Japan | Taiwan |
| 1970-1979 | 66 | 72 | 98 | 96 | 65 | 86 | 98 | 98 |
| 1975-1984 | 66 | 75 | 95 | 97 | 65 | 81 | 97 | 98 |
| 1980-1989 | 66 | 75 | 95 | 96 | 65 | 77 | 98 | 98 |
| 1985-1994 | 66 | 66 | 98 | 95 | 65 | 67 | 98 | 98 |
| 1990-1999 | 66 | 66 | 69 | 98 | 65 | 69 | 68 | 98 |
| 1995-2004 | 66 | 66 | 66 | 98 | 65 | 69 | 65 | 98 |

Table 4. AIC results for all models

| Country | Model | Male | Female |
|---|---|---|---|
| U.S. | LC | -2442.39 | -2710.06 |
| | CBD | -2900.39 | -2901.45 |
| | LC + Logistic | -2137.96 | -2298.65 |
| | LC + CK | -2770.56 | -2909.07 |
| | LC + Gompertz | -2627.22 | -2960.20 |
| | CBD + Gompertz | -3170.57 | -2962.67 |
| | CBD + Logistic | -2505.06 | -2616.81 |
| | CBD + CK | -2976.2 | -2874.85 |
| U.K. | LC | -2188.61 | -2485.58 |
| | CBD | -2832.27 | -3054.74 |
| | LC + Logistic | -2006.08 | -2238.35 |
| | LC + CK | -2340.22 | -2735.48 |
| | LC + Gompertz | -2502.70 | -2699.24 |
| | CBD + Gompertz | -2871.12 | -3032.61 |
| | CBD + Logistic | -2358.40 | -2598.59 |
| | CBD + CK | -2877.96 | -3033.13 |
| Japan | LC | -2287.47 | -2511.17 |
| | CBD | -2894.01 | -3106.61 |
| | LC + Logistic | -2162.06 | -2387.90 |
| | LC + CK | -2563.81 | -2850.71 |
| | LC + Gompertz | -2711.83 | -2844.52 |
| | CBD + Gompertz | -2602.32 | -2683.34 |
| | CBD + Logistic | -2511.96 | -2740.09 |
| | CBD + CK | -2897.29 | -3078.81 |
| Taiwan | LC | -2292.74 | -2391.17 |
| | CBD | -2357.67 | -2584.49 |
| | LC + Logistic | -2141.92 | -2502.43 |
| | LC + CK | -2351.51 | -2479.17 |
| | LC + Gompertz | -2265.03 | -2415.86 |
| | CBD + Gompertz | -2353.67 | -2580.49 |
| | CBD + Logistic | -2318.17 | -2703.30 |
| | CBD + CK | -2348.34 | -2575.16 |

Note: The numbers with shadow are the best among all models.

Table 5. Fitting Out-performance Scores of all models

| Model | RMSE | | MAPE | | Total |
|---|---|---|---|---|---|
| | 65~84 | 85~99 | 65~84 | 85~99 | |
| LC | 9 | 1 | 15 | 2 | 27 |
| CBD | 9 | 5 | 4 | 7 | 25 |
| LC + Logistic | 1 | 2 | 2 | 2 | 7 |
| LC + CK | 0 | 8 | 4 | 5 | 17 |
| LC + Gompertz | 4 | 7 | 9 | 5 | 25 |
| CBD + Gompertz | 13 | 8 | 8 | 11 | 40 |
| CBD + Logistic | 2 | 3 | 0 | 3 | 8 |
| CBD + CK | 10 | 14 | 6 | 13 | 43 |
| Total score | 48 | 48 | 48 | 48 | 192 |

Note: The numbers with shadow are the best among all models.

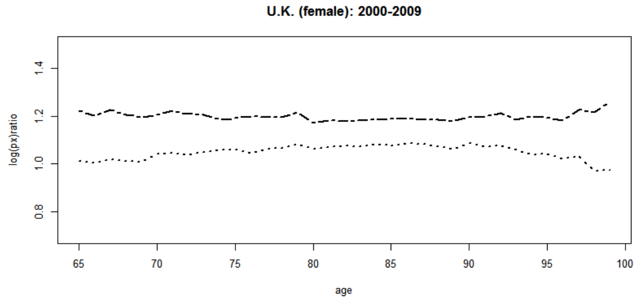
Table 6. Forecasting Out-performance Scores of all models

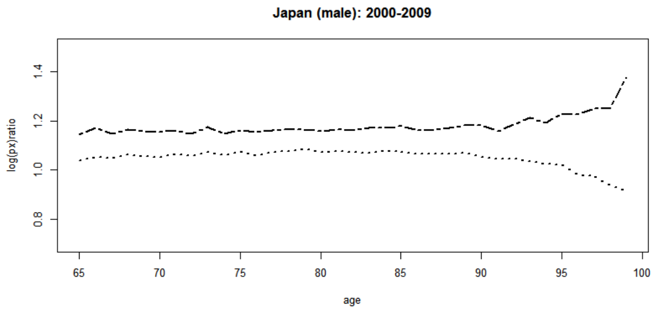
| Model | RMSE | | MAPE | | Total |
|---|---|---|---|---|---|
| | 65~84 | 85~99 | 65~84 | 85~99 | |
| LC | 4 | 0 | 5 | 2 | 11 |
| CBD | 9 | 4 | 9 | 6 | 28 |
| LC + Logistic | 7 | 4 | 8 | 4 | 23 |
| LC + CK | 0 | 2 | 1 | 1 | 4 |
| LC + Gompertz | 2 | 10 | 5 | 11 | 28 |
| CBD + Gompertz | 9 | 10 | 12 | 10 | 41 |
| CBD + Logistic | 9 | 6 | 2 | 2 | 19 |
| CBD + CK | 8 | 12 | 6 | 12 | 38 |
| Total Score | 48 | 48 | 48 | 48 | 192 |

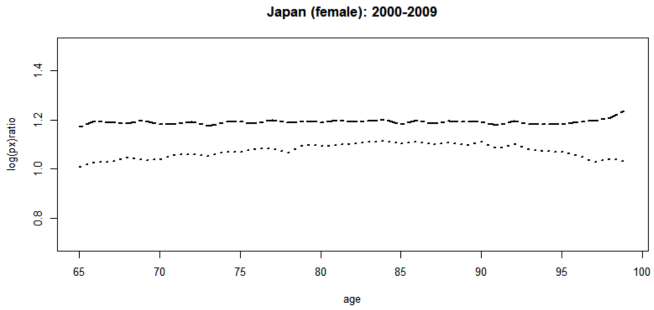
Note: The numbers with shadow are the best among all models.

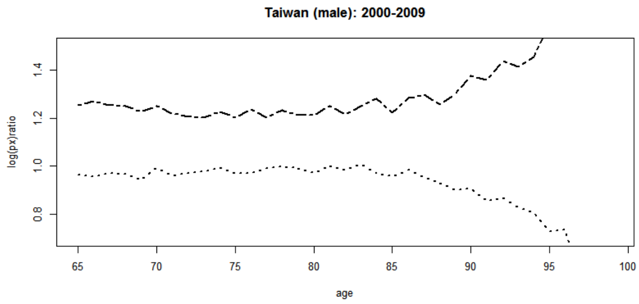
Appendix A: Bootstrap Confidence Intervals of U.S., U.K., Japan, and Taiwan




Japan (male): 2000-2009
log(px)ratio
0.8
1.0
1.2
1.4
65
70
75
80
85
90
95
100
age


Japan (female): 2000-2009
log(px)ratio
0.8
1.0
1.2
1.4
65
70
75
80
85
90
95
100
age


Taiwan (male): 2000-2009
log(px)ratio
0.8
1.0
1.2
1.4
65
70
75
80
85
90
95
100
age


Taiwan (female): 2000-2009
log(px)ratio
0.8
1.0
1.2
1.4
65
70
75
80
85
90
95
100
age

## Appendix B: Empirical Results of Mortality Models

Table B-1. Fitting Errors for ages 65~84

| Country | Model | Male | | Female | |
|---|---|---|---|---|---|
| | | RMSE($\times 10^{-3}$) | MAPE(%) | RMSE($\times 10^{-3}$) | MAPE(%) |
| U.S. | LC | 3.3560 | 3.1277 | 1.4594 | 2.1621 |
| | CBD | 2.1185 | 2.5550 | 1.8546 | 4.0348 |
| | LC + Logistic | 3.5496 | 3.2161 | 1.7389 | 3.4426 |
| | LC + CK | 4.6253 | 3.4927 | 2.4752 | 2.7939 |
| | LC + Gompertz | 3.9362 | 3.2945 | 1.7601 | 2.2613 |
| | CBD + Gompertz | 1.9626 | 2.3263 | 1.6819 | 3.6132 |
| | CBD + Logistic | 2.3927 | 2.9576 | 2.1067 | 4.5994 |
| | CBD + CK | 2.0941 | 2.4997 | 1.7936 | 3.8961 |
| U.K. | LC | 4.9721 | 3.8591 | 2.1960 | 2.6960 |
| | CBD | 2.4688 | 2.3131 | 1.8309 | 3.1763 |
| | LC + Logistic | 4.6858 | 3.7752 | 2.2085 | 2.7896 |
| | LC + CK | 5.8437 | 4.3659 | 2.8245 | 3.0330 |
| | LC + Gompertz | 5.2510 | 4.3488 | 2.6865 | 2.9121 |
| | CBD + Gompertz | 2.4403 | 2.3125 | 1.7862 | 3.0816 |
| | CBD + Logistic | 2.6401 | 2.5159 | 2.0874 | 3.7577 |
| | CBD + CK | 2.4272 | 2.2817 | 1.8207 | 3.1184 |
| Japan | LC | 3.5715 | 3.0268 | 1.7504 | 2.1575 |
| | CBD | 2.1805 | 2.7498 | 1.6379 | 3.7064 |
| | LC + Logistic | 4.1610 | 3.5400 | 2.4842 | 3.7494 |
| | LC + CK | 4.3507 | 3.3342 | 2.5940 | 2.6477 |
| | LC + Gompertz | 4.3719 | 3.2891 | 2.5311 | 2.5368 |
| | CBD + Gompertz | 2.1934 | 2.6369 | 1.6582 | 3.4820 |
| | CBD + Logistic | 2.2652 | 3.0344 | 1.6669 | 4.0903 |
| | CBD + CK | 2.2044 | 2.7541 | 1.6462 | 3.7030 |
| Taiwan | LC | 5.8133 | 4.5638 | 3.9634 | 4.1601 |
| | CBD | 6.8117 | 5.7656 | 4.4357 | 5.1083 |
| | LC + Logistic | 7.8068 | 6.1683 | 4.9867 | 5.3945 |
| | LC + CK | 6.9567 | 5.1280 | 4.6875 | 4.6720 |
| | LC + Gompertz | 6.4367 | 4.7221 | 4.1599 | 4.2192 |
| | CBD + Gompertz | 6.8117 | 5.7656 | 4.4357 | 5.1083 |
| | CBD + Logistic | 6.6875 | 5.7677 | 4.4041 | 5.1468 |
| | CBD + CK | 6.8117 | 5.7656 | 4.4357 | 5.1083 |

Note: The numbers with shadow are the best among all models.

Table B-2. Fitting Errors for ages 85~99

| Country | Model | Male | | Female | |
|---|---|---|---|---|---|
| | | RMSE($\times 10^{-3}$) | MAPE(%) | RMSE($\times 10^{-3}$) | MAPE(%) |
| U.S. | LC | 37.4976 | 11.3087 | 25.5738 | 9.0608 |
| | CBD | 24.1130 | 6.6857 | 24.2132 | 8.4894 |
| | LC + Logistic | 55.4611 | 16.6398 | 44.2728 | 16.0002 |
| | LC + CK | 19.7779 | 7.7307 | 16.8179 | 7.3884 |
| | LC + Gompertz | 24.6415 | 8.7118 | 15.4925 | 6.9827 |
| | CBD + Gompertz | 13.3335 | 3.8946 | 17.7827 | 6.4567 |
| | CBD + Logistic | 39.9940 | 10.4867 | 34.0051 | 11.1086 |
| | CBD + CK | 17.7228 | 5.3219 | 20.6657 | 7.6032 |
| U.K. | LC | 53.8685 | 13.1828 | 35.3625 | 10.6348 |
| | CBD | 24.8964 | 4.7524 | 18.2539 | 5.5562 |
| | LC + Logistic | 66.3180 | 15.4857 | 47.6979 | 14.6451 |
| | LC + CK | 35.8483 | 9.6428 | 20.8625 | 7.8345 |
| | LC + Gompertz | 27.8738 | 7.1356 | 21.7907 | 8.1852 |
| | CBD + Gompertz | 20.6717 | 4.2787 | 16.2918 | 4.5904 |
| | CBD + Logistic | 47.8889 | 9.1208 | 33.8142 | 9.1015 |
| | CBD + CK | 19.8658 | 3.8724 | 15.4734 | 4.9204 |
| Japan | LC | 47.3160 | 12.2084 | 35.0886 | 10.2843 |
| | CBD | 23.1915 | 5.7631 | 17.1989 | 5.6509 |
| | LC+Logistic | 52.6172 | 14.1613 | 38.1043 | 13.0589 |
| | LC+CK | 28.1403 | 8.3231 | 18.0771 | 6.6864 |
| | LC+Gompertz | 22.8317 | 7.0423 | 18.7452 | 6.3077 |
| | CBD+Gompertz | 35.8247 | 8.1371 | 32.9483 | 8.0402 |
| | CBD + Logistic | 38.1077 | 8.6331 | 26.763 | 8.2957 |
| | CBD + CK | 21.6431 | 5.3908 | 16.8237 | 5.4409 |
| Taiwan | LC | 44.5132 | 12.4697 | 38.6652 | 10.8262 |
| | CBD | 57.7218 | 20.0147 | 37.3526 | 10.9785 |
| | LC + Logistic | 51.8190 | 16.1094 | 30.8100 | 8.8296 |
| | LC + CK | 38.1068 | 11.5223 | 31.9804 | 9.5293 |
| | LC + Gompertz | 42.0679 | 12.6552 | 32.4046 | 9.8854 |
| | CBD + Gompertz | 57.7218 | 20.0147 | 37.3526 | 10.9785 |
| | CBD + Logistic | 50.7823 | 16.6938 | 28.6391 | 8.1733 |
| | CBD + CK | 57.7215 | 20.0147 | 37.3525 | 10.9785 |

Note: The numbers with shadow are the best among all models.

Table B-3. Forecasting Errors for ages 65~84

| Country | Model | Male | | Female | |
|---|---|---|---|---|---|
| | | RMSE($\times 10^{-3}$) | MAPE(%) | RMSE($\times 10^{-3}$) | MAPE(%) |
| U.S. | LC | 3.3740 | 3.9212 | 2.0057 | 3.8829 |
| | CBD | 2.8898 | 3.9629 | 2.6157 | 5.6883 |
| | LC + Logistic | 3.7241 | 4.1513 | 2.4206 | 5.1069 |
| | LC + CK | 4.6562 | 4.4027 | 3.0765 | 4.7087 |
| | LC + Gompertz | 3.8357 | 4.0913 | 2.1757 | 3.9097 |
| | CBD + Gompertz | 2.7170 | 3.6410 | 2.4714 | 5.2919 |
| | CBD + Logistic | 3.1676 | 4.4455 | 2.8240 | 6.2341 |
| | CBD + CK | 2.8342 | 3.8672 | 2.5631 | 5.5667 |
| U.K. | LC | 4.0196 | 4.2840 | 2.2153 | 4.4132 |
| | CBD | 3.0639 | 3.8845 | 2.2554 | 4.5885 |
| | LC + Logistic | 3.8397 | 3.9780 | 2.2655 | 4.2714 |
| | LC + CK | 4.9212 | 5.0706 | 2.4642 | 4.4244 |
| | LC + Gompertz | 4.2132 | 4.3535 | 2.3816 | 4.1642 |
| | CBD + Gompertz | 2.8290 | 3.7237 | 2.0652 | 4.4468 |
| | CBD + Logistic | 3.5034 | 4.3308 | 2.6572 | 5.2325 |
| | CBD + CK | 2.9109 | 3.8051 | 2.1679 | 4.5415 |
| Japan | LC | 4.0920 | 6.1659 | 1.8233 | 4.8791 |
| | CBD | 2.6074 | 4.5133 | 1.5196 | 4.4374 |
| | LC + Logistic | 4.6493 | 6.1616 | 2.3510 | 5.4622 |
| | LC + CK | 4.7505 | 6.4669 | 2.3772 | 5.1434 |
| | LC + Gompertz | 4.7055 | 6.3933 | 2.3313 | 5.1169 |
| | CBD + Gompertz | 2.6725 | 4.5106 | 1.5464 | 4.2312 |
| | CBD + Logistic | 2.5333 | 4.5737 | 1.5036 | 4.7498 |
| | CBD + CK | 2.6360 | 4.5406 | 1.5280 | 4.4394 |
| Taiwan | LC | 7.3002 | 8.2912 | 5.3172 | 7.7119 |
| | CBD | 5.9859 | 6.1255 | 4.2835 | 5.9536 |
| | LC + Logistic | 5.8909 | 6.0399 | 4.0245 | 5.2832 |
| | LC + CK | 7.4214 | 8.1515 | 5.2624 | 7.4556 |
| | LC + Gompertz | 7.1015 | 8.1854 | 5.1306 | 7.6294 |
| | CBD + Gompertz | 5.9859 | 6.1255 | 4.2835 | 5.9536 |
| | CBD + Logistic | 5.9436 | 6.2005 | 4.3752 | 6.1213 |
| | CBD + CK | 5.9859 | 6.1255 | 4.2835 | 5.9536 |

Note: The numbers with shadow are the best among all models.

Table B-4. Forecasting Errors for ages 85~99

| Country | Model | Male | | Female | |
|---|---|---|---|---|---|
| | | RMSE($\times 10^{-3}$) | MAPE(%) | RMSE($\times 10^{-3}$) | MAPE(%) |
| U.S. | LC | 39.4826 | 11.5883 | 27.6544 | 9.9862 |
| | CBD | 29.7181 | 8.0761 | 30.3576 | 10.7362 |
| | LC + Logistic | 62.4814 | 18.1527 | 50.6321 | 17.8374 |
| | LC + CK | 22.2683 | 8.1030 | 21.6274 | 9.1569 |
| | LC + Gompertz | 26.1881 | 9.0437 | 18.5894 | 8.2137 |
| | CBD + Gompertz | 16.6971 | 5.0162 | 21.5568 | 8.4009 |
| | CBD + Logistic | 46.0261 | 11.7647 | 39.9645 | 13.1683 |
| | CBD + CK | 22.2178 | 6.5447 | 25.6593 | 9.6790 |
| U.K. | LC | 51.9006 | 12.8568 | 34.7497 | 10.6434 |
| | CBD | 22.6302 | 5.0615 | 19.2496 | 6.0992 |
| | LC + Logistic | 61.9280 | 15.0403 | 48.4727 | 14.9654 |
| | LC + CK | 52.0008 | 11.9610 | 23.1984 | 8.6991 |
| | LC + Gompertz | 27.7059 | 7.5214 | 26.6892 | 9.6011 |
| | CBD + Gompertz | 20.2961 | 4.8595 | 16.0250 | 5.1120 |
| | CBD + Logistic | 42.6221 | 8.7252 | 33.1332 | 9.3357 |
| | CBD + CK | 18.3238 | 4.3022 | 16.9487 | 5.5190 |
| Japan | LC | 45.2495 | 12.9741 | 31.2740 | 10.4128 |
| | CBD | 24.9037 | 7.4609 | 18.5358 | 7.7907 |
| | LC + Logistic | 51.7791 | 15.7953 | 36.8234 | 14.5451 |
| | LC + CK | 30.0134 | 9.6770 | 23.1293 | 8.9763 |
| | LC + Gompertz | 22.3085 | 8.0699 | 18.1496 | 7.1200 |
| | CBD + Gompertz | 30.9035 | 8.0922 | 26.5387 | 8.5839 |
| | CBD + Logistic | 38.4739 | 10.5045 | 28.0938 | 10.5815 |
| | CBD + CK | 22.7956 | 7.0297 | 17.3811 | 7.5203 |
| Taiwan | LC | 47.9116 | 12.6169 | 41.1403 | 11.4193 |
| | CBD | 45.2508 | 14.9925 | 35.6224 | 12.2041 |
| | LC + Logistic | 40.1348 | 12.6390 | 27.9482 | 9.2390 |
| | LC + CK | 44.5323 | 12.7162 | 38.2318 | 11.2188 |
| | LC + Gompertz | 40.9803 | 11.2288 | 31.7637 | 9.9403 |
| | CBD + Gompertz | 45.2508 | 14.9925 | 35.6224 | 12.2041 |
| | CBD + Logistic | 39.9630 | 12.8570 | 27.8872 | 9.9199 |
| | CBD + CK | 45.2506 | 14.9925 | 35.6231 | 12.2042 |

Note: The numbers with shadow are the best among all models.

## Appendix C.

Table C-1. Single premium of 10-year pure annuity paying $1 each year - U.S.

| Model | Male | | | Female | | |
|---|---|---|---|---|---|---|
| | 70 | 80 | 90 | 70 | 80 | 90 |
| Actual Cost of Annuity | 5.8109 | 4.4646 | 2.3788 | 6.1381 | 5.0381 | 2.9376 |
| LC | 5.7842 | 4.4748 | 2.6695 | 6.1014 | 5.0846 | 3.2451 |
| | -0.46% | 0.23% | 12.22% | -0.60% | 0.92% | 10.47% |
| CBD | 5.7420 | 4.4193 | 2.6552 | 6.0971 | 5.0165 | 3.2645 |
| | -1.19% | -1.01% | 11.62% | -0.67% | -0.43% | 11.13% |
| LC + Log | 5.7526 | 4.5020 | 2.8967 | 6.0812 | 5.0576 | 3.4755 |
| | -1.00% | 0.84% | 21.77% | -0.93% | 0.39% | 18.31% |
| LC + CK | 5.7816 | 4.5371 | 2.5336 | 6.0998 | 5.1172 | 3.2039 |
| | -0.50% | 1.63% | 6.51% | -0.62% | 1.57% | 9.06% |
| LC + G | 5.7825 | 4.5059 | 2.6051 | 6.1013 | 5.0917 | 3.2441 |
| | -0.49% | 0.93% | 9.51% | -0.60% | 1.07% | 10.43% |
| CBD + G | 5.7462 | 4.4259 | 2.5790 | 6.0992 | 5.0265 | 3.2149 |
| | -1.11% | -0.86% | 8.42% | -0.63% | -0.23% | 9.44% |
| CBD + Log | 5.7369 | 4.4083 | 2.7410 | 6.0946 | 5.0031 | 3.3172 |
| | -1.27% | -1.26% | 15.23% | -0.71% | -0.69% | 12.92% |
| CBD + CK | 5.7418 | 4.4279 | 2.6255 | 6.0964 | 5.0244 | 3.2481 |
| | -1.19% | -0.82% | 10.37% | -0.68% | -0.27% | 10.57% |

Table C-2. Single premium of 10-year pure annuity paying $1 each year - U.K.

| Model | Male | | | Female | | |
|---|---|---|---|---|---|---|
| | 70 | 80 | 90 | 70 | 80 | 90 |
| Actual Cost of Annuity | 5.7830 | 4.2021 | 2.2652 | 6.1471 | 4.8601 | 2.7991 |
| LC | 5.6945 | 4.3331 | 2.4933 | 6.0703 | 4.9700 | 3.0157 |
| | -1.53% | 3.12% | 10.07% | -1.25% | 2.26% | 7.74% |
| CBD | 5.6866 | 4.2074 | 2.3276 | 6.0856 | 4.8734 | 2.9406 |
| | -1.67% | 0.13% | 2.76% | -1.00% | 0.27% | 5.05% |
| LC + Log | 5.6955 | 4.3147 | 2.6242 | 6.0648 | 4.9259 | 3.1992 |
| | -1.51% | 2.68% | 15.85% | -1.34% | 1.35% | 14.29% |
| LC + CK | 5.6867 | 4.3849 | 2.4093 | 6.0615 | 5.0144 | 2.9784 |
| | -1.67% | 4.35% | 6.36% | -1.39% | 3.17% | 6.41% |
| LC + G | 5.7057 | 4.3335 | 2.4500 | 6.0696 | 4.9565 | 3.1023 |
| | -1.34% | 3.13% | 8.16% | -1.26% | 1.98% | 10.83% |
| CBD + G | 5.6842 | 4.2217 | 2.2783 | 6.0832 | 4.8867 | 2.9111 |
| | -1.71% | 0.47% | 0.58% | -1.04% | 0.55% | 4.00% |
| CBD + Log | 5.6800 | 4.1927 | 2.4322 | 6.0826 | 4.8542 | 3.0042 |
| | -1.78% | -0.22% | 7.37% | -1.05% | -0.12% | 7.33% |
| CBD + CK | 5.6865 | 4.2188 | 2.2901 | 6.0847 | 4.8844 | 2.9200 |
| | -1.67% | 0.40% | 1.10% | -1.02% | 0.50% | 4.32% |

Table C-3. Single premium of 10-year pure annuity paying $1 each year – Japan

| Model | Male | | | Female | | |
|---|---|---|---|---|---|---|
| | 70 | 80 | 90 | 70 | 80 | 90 |
| Actual Cost of Annuity | 5.9874 | 4.6890 | 2.6888 | 6.4584 | 5.5623 | 3.4986 |
| LC | 5.9883 | 4.6916 | 2.7974 | 6.4627 | 5.5616 | 3.5392 |
| | 0.02% | 0.05% | 4.04% | 0.07% | -0.01% | 1.16% |
| CBD | 5.9441 | 4.7565 | 2.9710 | 6.4428 | 5.6021 | 3.7805 |
| | -0.72% | 1.44% | 10.49% | -0.24% | 0.72% | 8.06% |
| LC + Log | 5.9421 | 4.7436 | 3.0383 | 6.4376 | 5.5649 | 3.7626 |
| | -0.76% | 1.16% | 13.00% | -0.32% | 0.05% | 7.55% |
| LC + CK | 5.9854 | 4.7224 | 2.7465 | 6.4619 | 5.5728 | 3.5220 |
| | -0.03% | 0.71% | 2.15% | 0.05% | 0.19% | 0.67% |
| LC + G | 5.9872 | 4.7147 | 2.7255 | 6.4624 | 5.5722 | 3.4956 |
| | 0.00% | 0.55% | 1.36% | 0.06% | 0.18% | -0.09% |
| CBD + G | 5.9441 | 4.7565 | 2.9711 | 6.4428 | 5.6021 | 3.7807 |
| | -0.72% | 1.44% | 10.50% | -0.24% | 0.72% | 8.06% |
| CBD + Log | 5.9404 | 4.7483 | 3.0495 | 6.4417 | 5.5937 | 3.8292 |
| | -0.78% | 1.26% | 13.41% | -0.26% | 0.56% | 9.45% |
| CBD + CK | 5.9441 | 4.7565 | 2.9710 | 6.4428 | 5.6021 | 3.7805 |
| | -0.72% | 1.44% | 10.49% | -0.24% | 0.72% | 8.06% |

Table C-4. Single premium of 10-year pure annuity paying $1 each year – Taiwan

| Model | Male | | | Female | | |
|---|---|---|---|---|---|---|
| | 70 | 80 | 90 | 70 | 80 | 90 |
| Actual Cost of Annuity | 5.7768 | 4.4971 | 2.5655 | 5.7768 | 4.4971 | 2.5655 |
| LC | 5.8175 | 4.3359 | 2.6022 | 5.8175 | 4.3359 | 2.6022 |
| | 0.70% | -3.59% | 1.43% | 0.70% | -3.59% | 1.43% |
| CBD | 5.7461 | 4.4618 | 2.7092 | 5.7461 | 4.4618 | 2.7092 |
| | -0.53% | -0.78% | 5.60% | -0.53% | -0.78% | 5.60% |
| LC + Log | 5.7500 | 4.4972 | 2.8797 | 5.7500 | 4.4972 | 2.8797 |
| | -0.46% | 0.00% | 12.25% | -0.46% | 0.00% | 12.25% |
| LC + CK | 5.8175 | 4.3359 | 2.6022 | 5.8175 | 4.3359 | 2.6022 |
| | 0.70% | -3.59% | 1.43% | 0.70% | -3.59% | 1.43% |
| LC + G | 5.8115 | 4.4286 | 2.4247 | 5.8115 | 4.4286 | 2.4247 |
| | 0.60% | -1.52% | -5.49% | 0.60% | -1.52% | -5.49% |
| CBD + G | 5.7461 | 4.4618 | 2.7092 | 5.7461 | 4.4618 | 2.7092 |
| | -0.53% | -0.78% | 5.60% | -0.53% | -0.78% | 5.60% |
| CBD + Log | 5.7422 | 4.4588 | 2.8066 | 5.7422 | 4.4588 | 2.8066 |
| | -0.60% | -0.85% | 9.40% | -0.60% | -0.85% | 9.40% |
| CBD + CK | 5.7461 | 4.4618 | 2.7092 | 5.7461 | 4.4618 | 2.7092 |
| | -0.53% | -0.78% | 5.60% | -0.53% | -0.78% | 5.60% |