

The CBD Mortality Indexes: Modeling and Applications

Wai-Sum Chan, Johnny Siu-Hang Li* and Jackie Li

Presented at the Living to 100 Symposium

Orlando, Fla.

January 8–10, 2014

Copyright 2014 by the Society of Actuaries.

All rights reserved by the Society of Actuaries. Permission is granted to make brief excerpts for a published review. Permission is also granted to make limited numbers of copies of items in this monograph for personal, internal, classroom or other instructional use, on condition that the foregoing copyright notice is used so as to give reasonable notice of the Society's copyright. This consent for free limited copying without prior consent of the Society does not extend to making copies for general distribution, for advertising or promotional purposes, for inclusion in new collective works or for resale.

* Corresponding author. Address: Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1. Email: shli@uwaterloo.ca

The CBD Mortality Indexes: Modeling and Applications

Wai-Sum Chan, Johnny Siu-Hang Li* and Jackie Li

Abstract

Most extrapolative stochastic mortality models are constructed in a similar manner. Specifically, when they are fitted to historical data, one or more series of time-varying parameters are identified. By extrapolating these parameters to the future, we can obtain a forecast of death probabilities and consequently cash flows arising from life contingent liabilities. In this paper, we first argue that, among various time-varying model parameters, those encompassed in the Cairns-Blake-Dowd (CBD) model (also known as Model M5) are most suitably used as indexes to indicate levels of longevity risk at different time points. We then investigate how these indexes can be jointly modeled with a more general class of multivariate time-series models, instead of a simple random walk that takes no account of cross-correlations. Finally, we study the joint prediction region for the mortality indexes. Such a region, as we demonstrate, can serve as a graphical longevity risk metric, allowing practitioners to compare the longevity risk exposures of different portfolios readily.

Keywords: Cairns-Blake-Dowd model; VARIMA models; Prediction regions.

1 Introduction

As a functional equivalent of stock indexes, mortality indexes can be constructed to summarize levels of mortality at different time points. Mortality indexes have a wide range of applications. Their trends, for instance, indicate the speed at which the mortality in a population has been changing. The degree of longevity risk, that is, the risk associated with under- or over-estimating mortality improvements, can hence be observed from the deviations between the expected and actual trajectories of a mortality index. More importantly, with appropriate mortality indexes, one can construct standardized mortality-linked securities, such as longevity bonds and swaps, which can be traded between investors and pension plans who wish to hedge their longevity risk exposures. Relative to custom-made hedging instruments, standardized securities are easier for investors to analyze and are more conducive to the development of liquidity.

Several investment banks have launched mortality indexes over the past few years. In 2006, Credit Suisse launched a longevity index that is based on the life expectancy at birth of the US

*Corresponding author. Address: Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1. Email: shli@uwaterloo.ca

general population. In 2007, Goldman Sachs introduced the QxX Index, whose value is based on the number of survivors in the underlying reference pool, which consisted of 46,290 lives initially.¹ Also in 2007, JP Morgan started the LifeMetrics Index, which provides mortality rates (both crude and graduated) and period life expectancy levels across various ages, by gender, for the U.S., England and Wales, the Netherlands and Germany. In 2008, Deutsche Börse began to offer standardized Xpect Cohort Indexes for England and Wales, the Netherlands and Germany. Each Xpect Cohort Index tracks the number of survivors in a certain birth cohort on a monthly basis.

The aforementioned mortality indexes are all non-parametric. Being model-free may be seen as an advantage, but without the aid of a model, an index can convey only a limited amount of information. For instance, an index that is based on the life expectancy at birth summarizes longevity improvements at a highly aggregate level, indicating nothing about how the underlying mortality curve has actually evolved over time. Given that shifts in a mortality curve are generally not uniform and take no fixed shape (see, e.g., Li and Luo, 2012), a pension plan may find it difficult to build a longevity hedge using such an index. To represent the evolution of a mortality curve in a non-parametric mean, there is a need to use a rather large number of indexes. For instance, this may be achieved by using indexes that are linked to mortality rates at various ages. Nevertheless, the use of a large number of mortality indexes makes tracking difficult. It will also be difficult for the market to concentrate liquidity if there exists a large number of mortality indexes.

A model-based construction method may improve the information content of mortality indexes. In recent years, a number of extrapolative stochastic mortality models have been developed. These models are constructed in a similar manner. Specifically, when they are fitted to historical data, one or more series of time-varying parameters are identified. By extrapolating these parameters to the future, a mortality forecast can be obtained. These parameters, which contain rich information about the evolution of mortality in a population over time, can potentially be used as mortality indexes. The first objective of this paper is to investigate the possibility of constructing mortality indexes using the time-varying parameters in a prevalent stochastic mortality model. This idea is somewhat parallel to the construction of the Chicago Board Options Exchange (CBOE) implied volatility indexes, which are based on the volatility parameter in the Black-Scholes option pricing model.

Not all stochastic mortality models are suitable for creating indexes. We believe that model-based mortality indexes should satisfy three criteria. First, despite being small in dimension, the vector of indexes should represent the varying age-pattern of mortality improvement, rather than merely the overall level of mortality. Second, the model on which the indexes are based should possess what we refer to as the new-data-invariant property. This means that when an additional year of mortality data becomes available and the model is updated accordingly, indexes in previous years will not be affected. This criterion is crucially important, because it would be impossible to

¹Goldman Sachs ceased to run its QxX Index in December 2009.

track an index if its historical values are revised from time to time. Third, the mortality indexes should be readily interpretable, so that they can be communicated easily to hedgers, investors and even the general public. Among the six prevalent stochastic mortality models discussed in Dowd et al. (2010), we find that the original Cairns-Blake-Dowd (CBD) model (Cairns et al., 2006) seems the most suitable for constructing mortality indexes. We therefore propose to use the two time-varying parameters, usually denoted by $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$, in the CBD model jointly as mortality indexes.

The second objective of this paper is to model the joint dynamics of the CBD mortality indexes over time. This problem, to our knowledge, has not been extensively studied. Except Sweetings (2011), who employed a piece-wise linear function, researchers often use a simple bivariate random walk with drift to model $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$. A shortcoming of a bivariate random walk is that it does not capture any cross-correlations, which include, for example, the association between the change in one CBD mortality index during the current year and the change in the other CBD mortality index in a previous year. In this paper, we show empirically that cross-correlations between changes in the two CBD mortality indexes can be significant. To incorporate cross-correlations, we propose modeling the CBD mortality indexes with a general class of Vector Autoregressive Integrated Moving-Average (VARIMA) models. We also explain how an optimal VARIMA model for the CBD mortality indexes can be identified using a multivariate generalization of the Box and Jenkins (1976) approach.

When forecasting the CBD mortality indexes, it is important to consider both indexes jointly. This is because, as we explain later in this paper, the association between the two CBD mortality indexes has a significant impact on the longevity risk exposure of a portfolio. It is therefore inadequate to represent forecast uncertainty using a marginal prediction interval for each CBD mortality index. The third objective of this paper is to develop a joint prediction region for measuring the uncertainty surrounding both CBD mortality indexes simultaneously. We demonstrate that this prediction region, when plotted on a Cartesian plane, can serve as a graphical risk metric, allowing practitioners to compare the longevity risk exposures of different portfolios readily. The proposed graphical risk metric is a complement to a mortality/longevity fan chart (Blake et al., 2008; Dowd et al., 2010). The former measures the entire longevity risk profile at a single time point, while the latter measures the uncertainty associated with a particular death/survival probability over time.

The rest of this paper is organized as follows. Section 2 defines the CBD mortality indexes and evaluates the properties of the indexes. Section 3 presents the VARIMA model for modeling the dynamics of the CBD mortality indexes and details the procedure for identifying an optimal VARIMA model. Section 4 explains how a joint prediction region for the CBD mortality indexes can be constructed, and demonstrates how the prediction region can be used as a graphical longevity risk metric in practice. Section 5 concludes the paper.

2 Developing the CBD Mortality Indexes

2.1 The New-Data-Invariant Property

The first objective of this paper is to construct mortality indexes by using the time-varying parameters in a stochastic mortality model. The candidate models we consider are the six prevalent stochastic mortality models that were analyzed in Dowd et al. (2010). In describing these models, the following conventions are used:

- $m_{x,t}$ is the central death rate for age x in year t ;
- $q_{x,t}$ is the probability that an individual aged exactly x at exact time t will die between t and $t + 1$;
- $\beta_x^{(i)}$, $i = 1, 2, 3$, are age-specific parameters;
- $\kappa_t^{(i)}$, $i = 1, 2, 3$, are time-varying parameters;
- $\gamma_c^{(i)}$, $i = 3, 4$, where $c = t - x$ denotes year of birth, are cohort-related parameters;
- n_a is the number of ages covered in the sample age range;
- \bar{x} is the mean age over the sample age range;
- $\hat{\sigma}_x^2$ is the mean of $(x - \bar{x})^2$ over the sample age range.

Table 1 summarizes the specifications of the six candidate models. Some of these models require parameter constraints to ensure identifiability. The number of parameter constraints required for each model is shown in parentheses in Table 1. We refer readers to Cairns et al. (2009) and Dowd et al. (2010) for further discussions on the identifiability problem.

We require the model on which the mortality indexes are based to be completely robust with respect to an extension of the sample period. That is, when an additional year of mortality data becomes available and the stochastic model is updated accordingly, indexes in previous years will not be affected. This property, which we call the new-data-invariant property, is crucially important, because it would be impossible to track an index if its historical values are revised from time to time.

To examine the new-data-invariant property, we fit the six candidate models to data for English and Welsh males. The data were obtained from the Human Mortality Database (2012). We use an age range of 40-90, because some of the models are not able to capture the accident hump at younger ages and the data beyond age 90 may not be reliable.² Three different sample

²According to the Human Mortality Database, raw population counts by single year of age for English and Welsh males are available up to age 89 only. The population counts beyond age 89 are only estimates that are derived using the extinct cohort method.

Model M1: The Lee-Carter Model	
$\ln(m_{x,t}) = \beta_x^{(1)} + \beta_x^{(2)} \kappa_t^{(2)}$	(2 constraints)
Model M2: The Renshaw-Haberman Model	
$\ln(m_{x,t}) = \beta_x^{(1)} + \beta_x^{(2)} \kappa_t^{(2)} + \beta_x^{(3)} \gamma_{t-x}^{(3)}$	(4 constraints)
Model M3: The Age-Period-Cohort Model	
$\ln(m_{x,t}) = \beta_x^{(1)} + n_a^{-1} \kappa_t^{(2)} + n_a^{-1} \gamma_{t-x}^{(3)}$	(3 constraints)
Model M5: The Original Cairns-Blake-Dowd Model	
$\ln\left(\frac{q_{x,t}}{1-q_{x,t}}\right) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x})$	(No constraint)
Model M6: The Cairns-Blake-Dowd Model with a Cohort Effect Term	
$\ln\left(\frac{q_{x,t}}{1-q_{x,t}}\right) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \gamma_{t-x}^{(3)}$	(2 constraints)
Model M7: The Cairns-Blake-Dowd Model with Cohort Effect and Quadratic Terms	
$\ln\left(\frac{q_{x,t}}{1-q_{x,t}}\right) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \kappa_t^{(3)}((x - \bar{x})^2 - \hat{\sigma}_x^2) + \gamma_{t-x}^{(4)}$	(3 constraints)

Table 1: Specifications of the six stochastic mortality models under consideration. The number of required parameter constraints for each model is shown in parentheses.

periods, namely 1950-1989, 1950-1999 and 1950-2009, are considered. If the new-data-invariant property holds, then the three sets of parameter estimates during the overlapping years must be identical. The estimation results are shown graphically in Figure 1. We observe that among all six candidate models, the original CBD model (Model M5) is the only model that possesses the data-invariant property.

Technically speaking, there are two reasons that the (original) CBD model possesses the data-invariant property.

The first reason is related to the model's log-likelihood. In general, assuming Poisson death counts, the log-likelihood for the six candidate models can be expressed as

$$l = \sum_{x=x_0}^{x_1} \sum_{t=t_0}^{t_1} D_{x,t} \ln(E_{x,t} m_{x,t}) - E_{x,t} m_{x,t} - \ln(D_{x,t}!) = \sum_{t=t_0}^{t_1} \lambda(t),$$

where $[x_0, x_1]$ is the sample age range, $[t_0, t_1]$ is the sample period, $D_{x,t}$ is observed number of deaths at age x and in year t , and $E_{x,t}$ is the corresponding number of persons at risk. The value of $m_{x,t}$ depends on the model parameters via the specifications shown in Table 1. For models that are built for $q_{x,t}$, an assumption is needed to connect $q_{x,t}$ and $m_{x,t}$. We use the assumption that the force of mortality is constant between integer ages, which implies $m_{x,t} = -\ln(1 - q_{x,t})$. The

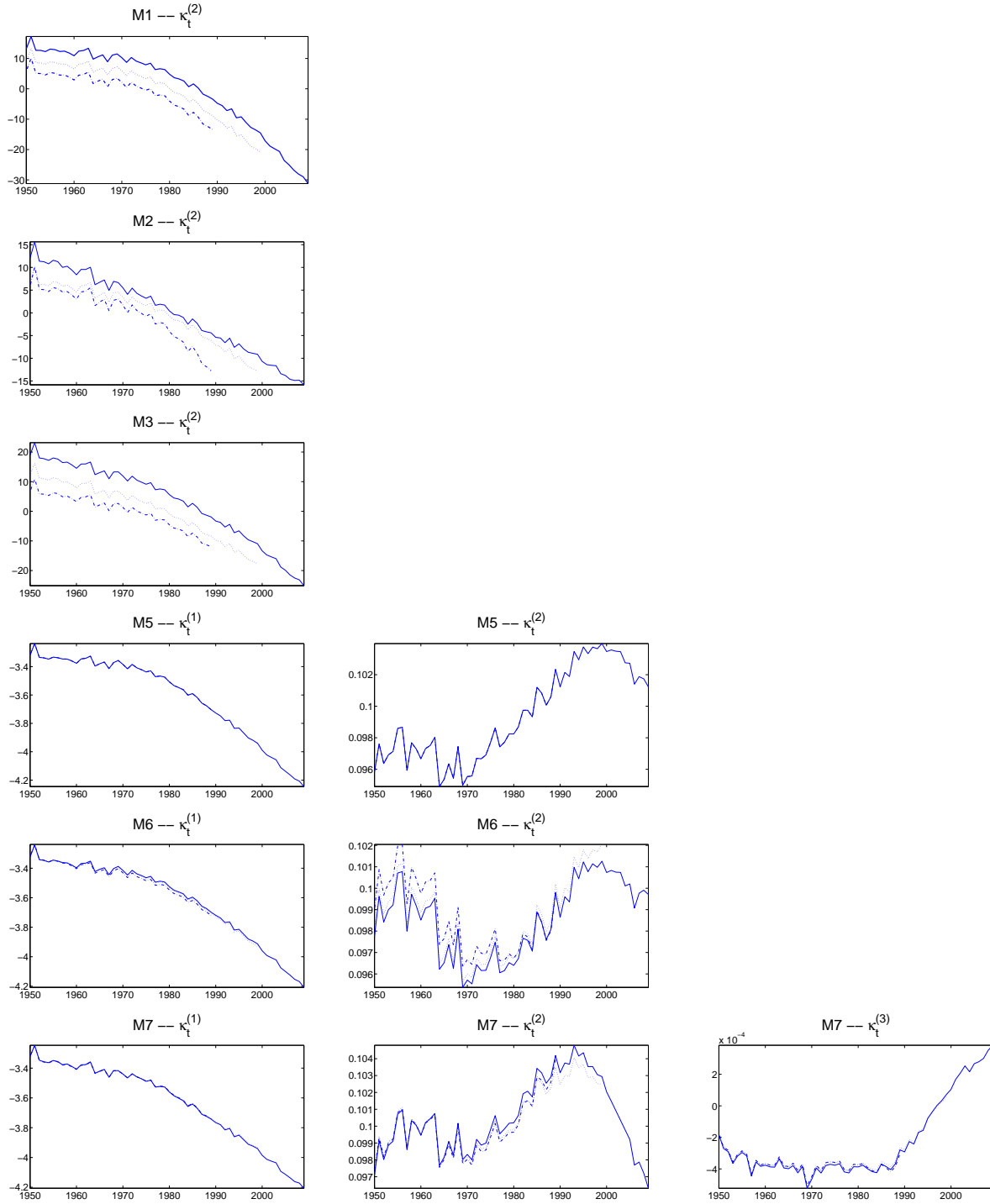


Figure 1: Estimates of the time-varying parameters in the six stochastic mortality models. Estimates based on sample periods 1950-1989, 1950-1999 and 1950-2009 are displayed in solid lines, dotted lines and dot-dashed lines, respectively.

function $\lambda(t)$ represents the contribution to the log-likelihood from data for year t . The form of $\lambda(t)$ depends on which candidate model is used.

For the (original) CBD model, we have

$$\lambda(t) = \sum_{x=x_0}^{x_1} D_{x,t} \ln(E_{x,t} \ln(1 + e^{\kappa_t^{(1)} + (x-\bar{x})\kappa_t^{(2)}})) - E_{x,t} \ln(1 + e^{\kappa_t^{(1)} + (x-\bar{x})\kappa_t^{(2)}}) - \ln(D_{x,t}!).$$

We observe that for $t \neq s$, there are no model parameters that are encompassed in both $\lambda(t)$ and $\lambda(s)$. Hence, extending the sample period, that is, adding extra terms $\lambda(t)$, $t = t_1 + 1, t_1 + 2, \dots$, to the log-likelihood, will not affect the maximum likelihood estimates of $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ for $t = t_0, \dots, t_1$.

However, the log-likelihoods for the other five candidate models do not exhibit such a feature. As an illustration, let us consider the Lee-Carter model. For this model, we have

$$\lambda(t) = \sum_{x=x_0}^{x_1} D_{x,t} (\beta_x^{(1)} + \beta_x^{(2)} \kappa_t^{(2)} + \ln(E_{x,t})) - E_{x,t} e^{\beta_x^{(1)} + \beta_x^{(2)} \kappa_t^{(2)}} - \ln(D_{x,t}!).$$

We observe that for $t \neq s$, $\lambda(t)$ and $\lambda(s)$ contain the same age-specific parameters. It follows that if the sample period is extended, then the values of $\beta_x^{(1)}$ and $\beta_x^{(2)}$ for $x = x_0, \dots, x_1$ and $\kappa_t^{(2)}$ for $t = t_0, \dots, t_1$ that maximize the original log-likelihood, $\sum_{t=t_0}^{t_1} \lambda(t)$, may no longer maximize the revised log-likelihood, $\sum_{t=t_0}^{t_1} \lambda(t) + \sum_{t>t_1} \lambda(t)$. Hence, the estimates of $\kappa_t^{(2)}$ before year $t_1 + 1$ will change as data for year $t_1 + 1$ and onwards are added to the estimation process, implying that the new-data-invariant property does not hold.

The second reason is related to the parameter constraints that are required to fix a unique parameterization. To illustrate, let us consider the Lee-Carter model again. The usual parameter constraints that are used in fitting the Lee-Carter model are $\sum_{x=x_0}^{x_1} \beta_x^{(2)} = 1$ and $\sum_{t=t_0}^{t_1} \kappa_t^{(2)} = 0$. Suppose that one more year of data is available and is included in the estimation process. The second parameter constraint requires the revised time-varying parameters to satisfy $\sum_{t=t_0}^{t_1+1} \kappa_t^{(2)} = 0$, which implies that unless $\kappa_{t_1+1}^{(2)}$ is zero, the values of $\kappa_t^{(2)}$ in the original sample period $[t_0, t_1]$ must change. Among all six candidate models, the (original) CBD model is the only one that requires no parameter constraint.

We acknowledge that the original CBD model may not give the best goodness-of-fit among the six candidate models. However, the top priority in our problem is the tractability of the resulting mortality indexes. We need a model that is completely robust with respect to an extension of the sample period, instead of one that is all-rounded. The use of the CBD model to produce mortality indexes may be compared with the use of the Black-Scholes model to produce implied volatility indexes. We know that both models may not be completely accurate, but they do produce indexes that are tractable by investors and other users.

2.2 Intuitions behind the CBD Mortality Indexes

Recall that the CBD model can be expressed as

$$\ln\left(\frac{q_{x,t}}{1-q_{x,t}}\right) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}),$$

which implies that for a given year t , the value of $q_{x,t}$ after a logit transformation³ is linearly related with age x . Given this simple structure, the CBD mortality indexes, that is, parameters $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ in equation above, can be interpreted easily.

The first CBD mortality index, $\kappa_t^{(1)}$, represents the level of the mortality curve (the curve of $q_{x,t}$ in year t) after a logit transformation. A reduction in $\kappa_t^{(1)}$, that is, a parallel downward shift of the logit-transformed mortality curve, represents an overall mortality improvement. The impact of a reduction in $\kappa_t^{(1)}$ on the logit-transformed mortality curve is illustrated on the left panel of Figure 2.

The second CBD mortality index, $\kappa_t^{(2)}$, represents the slope of the logit-transformed mortality curve. An increase in $\kappa_t^{(2)}$, that is, an increase in the steepness of the logit-transformed mortality curve, means that mortality (in logit scale) at younger ages (below the mean age \bar{x}) improves more rapidly than at older ages (above the mean age \bar{x}). The impact of an increase in $\kappa_t^{(2)}$ on the logit-transformed mortality curve is illustrated in the middle panel of Figure 2. Note that when the first index is fixed, a change in the second index has no impact on the mortality at the mean age $\bar{x} = 65$.

The two CBD mortality indexes can be used to represent a logit-transformed mortality curve with any slope and level. The right panel of Figure 2 shows the impact on the logit-transformed mortality curve when the two indexes are changed simultaneously. In Figure 3 we depict the mortality curves (in the original scale) implied by four different hypothetical pairs of the CBD mortality indexes. We observe that the two CBD mortality indexes, when used jointly, are able to capture different patterns of shifts in the underlying mortality curve. To achieve this ability with mortality indexes that are built in a non-parametric manner, a much larger number of indexes may be required.

Although each CBD mortality index has its own meaning, it is important to consider them jointly, because the association between them has a significant impact on the longevity risk exposure of a portfolio. To illustrate, let us consider the following two examples.

First, we consider closed pension plans, which are commonplace in the UK. Their financial obligations are positively related to future mortality improvements at old ages. More specifically, payouts from closed pension plans are larger when, of course, the overall mortality improvement is higher than expected, which happens when future values of $\kappa_t^{(1)}$ are lower than expected. Moreover, for a fixed overall mortality improvement, the problem to pension plans that are closed to new entrants would be worse when the mortality improvement is more concentrated at older

³The logit transformation of a real number w is given by $\ln(w/(1-w))$.

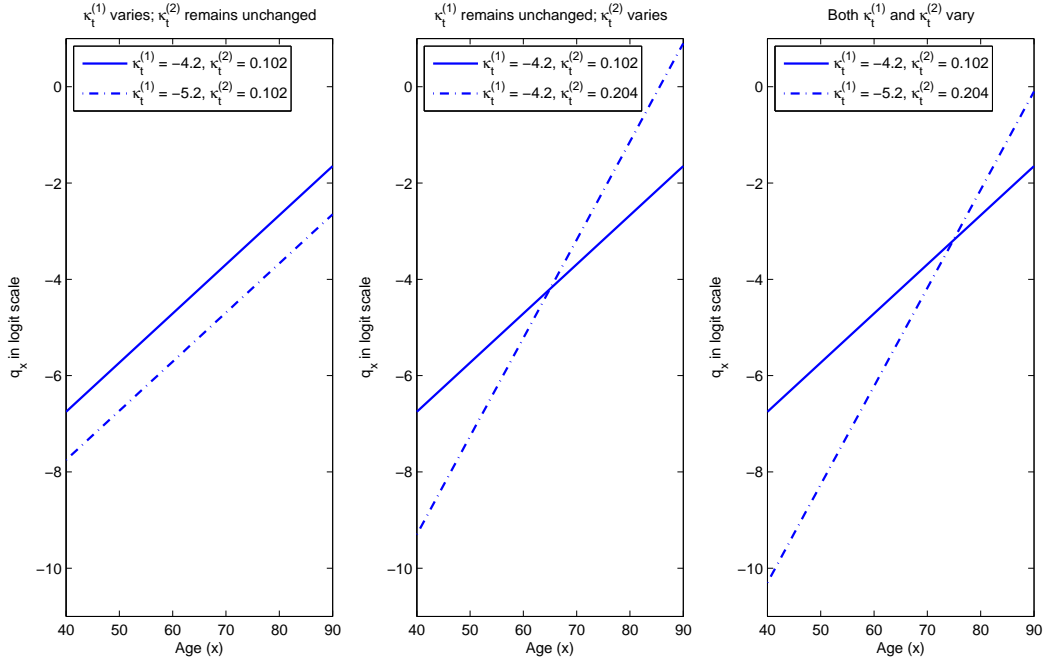


Figure 2: Changes in the logit-transformed mortality curve when either one or both of two CBD mortality indexes change.

ages, which happens when future values of $\kappa_t^{(2)}$ are lower than expected. Therefore, these plans will be subject to a double hit if both CBD mortality indexes turn out to be low.

Second, we consider insurance companies that concentrate on term-life insurances, which are typically sold to young people.⁴ Their financial obligations are negatively related to future mortality improvements at young ages. In more detail, payouts from these insurance companies are larger when the overall mortality improvement is smaller than expected, which occurs when future values of $\kappa_t^{(1)}$ are larger than expected. Also, for a fixed overall mortality improvement, their situation would be worse when the mortality improvement is more concentrated at older ages, which happens when future values of $\kappa_t^{(2)}$ are smaller than expected. Overall, for these insurance companies, the worst scenario occurs when $\kappa_t^{(1)}$ turns out to be large but $\kappa_t^{(2)}$ turns out to be small.

The association between the two CBD mortality indexes will be addressed in Section 3 when we model the dynamics of the two indexes over time, and also in Section 5 when we introduce a longevity risk metric that is based on the two indexes.

⁴For instance, 98% of the customers of Lifenet Insurance (a Japanese insurance company that focuses on selling term-life insurances) in FY3/11 were below age 60. Source: www.lifenet-seimei.co.jp.

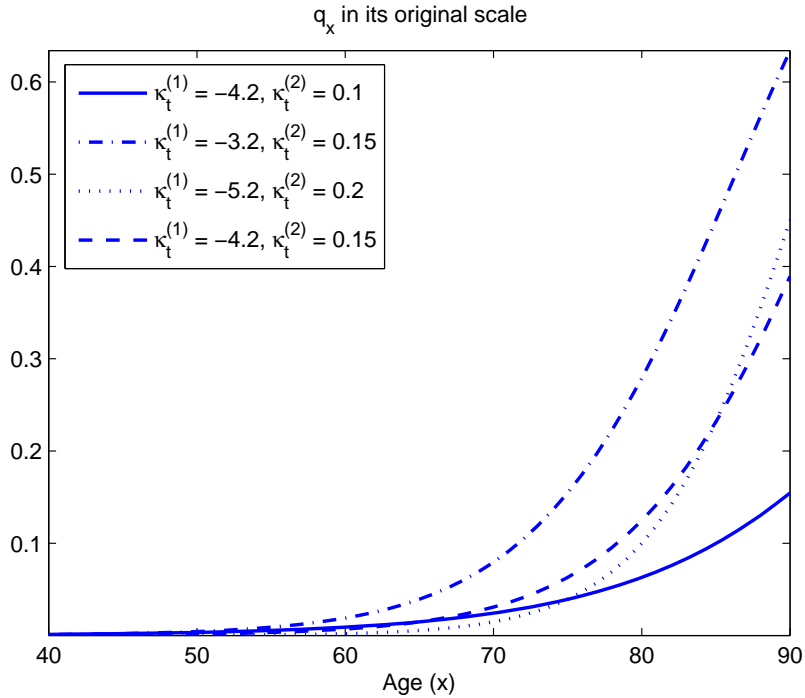


Figure 3: Curves of q_x (in its original scale) for different pairs of CBD mortality indexes.

2.3 Other Properties of the CBD Mortality Indexes

We have just demonstrated that the CBD mortality indexes satisfy the following three criteria for model-based mortality indexes:

1. The indexes can represent the varying age-pattern of mortality improvement, rather than the overall level of mortality only.
2. The model on which the indexes are based possesses the new-data-invariant property.
3. The indexes are readily interpretable.

On top of these three criteria, Sweeting (2010) suggests several other desirable criteria that mortality indexes, in general, should fulfill. In what follows, we explain these additional criteria, and discuss whether or not the CBD mortality indexes can satisfy them.

- Unambiguous

The population on which the mortality indexes are based must be defined in detail. This criterion is not difficult to fulfill if the CBD mortality indexes are based on major national populations, whose compositions are usually well documented.

- Transparent

The method used to calculate index values must be clear. While there exist multiple methods for estimating the CBD model, the index provider can use one method, for example, maximum likelihood, consistently over time. Methods for estimating the CBD model are detailed in many journal articles (e.g., Cairns et al., 2006, 2009), and a computer program for fitting the model is also available in the public domain (www.cbdmodel.com). Given the relevant data, one can easily replicate the calculations made by the index provider.

- Objective

The method used to calculate index values should have as little subjective input as possible. The CBD mortality indexes meet this criterion, because given a data sample and an age range $[x_0, x_1]$, the maximum likelihood estimation of the CBD model requires no subjective input.

- Measurable

The mortality experience of the reference population should be capable of being measured. The criterion is automatically satisfied if the CBD mortality indexes are based on major national populations.

- Timely and regular

Mortality data for the reference population should be available timely, so that index values can be made available in accordance with a pre-arranged timetable. This criterion is more related to the data provider than how the mortality indexes are constructed. The fulfillment of this criterion requires the data provider's cooperation.

- Appropriate

The indexes should reflect the compositions of the populations requiring hedging. If the CBD mortality indexes are based on national populations, then this criterion may not be met, because the mortality experience between the pension plan that requires hedging and the population of the country in which the plan is located could be quite different. This problem may be ameliorated by basing the CBD mortality indexes on data from various pension plans, in a way that is similar to Deutsche Börse's Xpect Club Vita Indexes.

- Popular

The indexes should be few enough, so that securities based on them will be liquid. The CBD mortality indexes meet this criterion very well, because for each reference population, there are only two CBD mortality indexes, $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$.

- Relevant, highly correlated and reflective of current hedging needs

The mortality indexes should allow hedgers to create effective longevity hedges. In Section 5, we propose to create forward contracts that are linked to the CBD mortality indexes.

Using an adapted version of the existing mortality duration and convexity measures, we may be able to calibrate an effective longevity hedge that is composed of such forward contracts.

- Stable and specified in advance

The indexes should be stable and defined in advance as far as possible. Unless the availability of data changes in the future, this criterion is met by the CBD mortality indexes. Note that this criterion is related to the new-data-invariant property, without which the index values will need to be revised from time to time.

3 Vector Time-Series Modelling

We now turn to the modeling of the CBD mortality indexes. In previous studies of the CBD model, including the original work of Cairns et al. (2006), the CBD time-varying parameters are often modeled by a bivariate random walk with drift. This simple process, however, does not fully capture the association between the two CBD mortality indexes. In more detail, it implies that after first differencing, the two indexes do not exhibit any serial- and cross-correlations, but such correlations, as we are about to demonstrate, can be significant.

This problem can be overcome by using a Vector Autoregressive Integrated Moving-Average (VARIMA) process instead of a bivariate random walk. In what follows, we present a VARIMA process that is applicable to the CBD mortality indexes. We also explain how an optimal VARIMA model for the CBD mortality indexes can be identified using a multivariate generalization of the Box and Jenkins (1976) approach. For brevity, we shall restrict the discussion to points necessary for describing the applications in this paper. Fuller details can be found in Tiao and Box (1981) and Wei (2006).

To illustrate the method, we use the CBD mortality indexes that are estimated from data for English and Welsh males over the age range [40, 90] and the sample period [1950, 2009]. The values of these indexes over the sample period are displayed graphically in Figure 1 (row 4).

3.1 Model Specification

Let us begin with the definition of a Vector Autoregressive Moving-Average (VARMA) process. The VARMA(p, q) process for a general bivariate vector time-series $\mathbf{Z}_t = (Z_t^{(1)}, Z_t^{(2)})'$ is defined as follows:

$$\mathbf{Z}_t = \mathbf{C}_0 + \sum_{i=1}^p \Phi_i \mathbf{Z}_{t-i} + \sum_{j=1}^q \Theta_j \epsilon_{t-j} + \epsilon_t, \quad (1)$$

where \mathbf{C}_0 is a (2×1) intercept vector, and Φ_i 's and Θ_j 's are (2×2) autoregressive and moving-average coefficient matrices, respectively. The residual (2×1) vectors ϵ_t are independently and identically distributed bivariate normal vectors with mean zero and variance-covariance matrix

Σ . The duplet (p, q) , where $p, q \geq 0$, is called the order of the VARMA process. If p (or q) equals zero, then the first (or second) summand in the equation above is omitted accordingly.

A VARMA process can be characterised by its cross-correlation matrices, denoted by $\mathbf{\Gamma}(l)$, for $l = 0, 1, \dots$. The (i, j) th element in $\mathbf{\Gamma}(l)$ is the correlation coefficient between $Z_t^{(i)}$ and $Z_{t-l}^{(j)}$, $i, j = 1, 2$. When $p = 0$, that is, \mathbf{Z}_t is a vector MA(q) process, the cross-correlation matrices for $l > q$ are all zero.

The order of a VARMA process is also related to its partial autoregression matrices, denoted by $\mathbf{P}(l)$, for $l = 0, 1, \dots$. The partial autoregression matrix $\mathbf{P}(l)$ at lag l is defined as the last coefficient matrix of the following VAR with order l :

$$\mathbf{Z}_{t+l} = \mathbf{\Phi}_{l,1} \mathbf{Z}_{t+l-1} + \mathbf{\Phi}_{l,2} \mathbf{Z}_{t+l-2} + \dots + \mathbf{\Phi}_{l,l} \mathbf{Z}_t + \boldsymbol{\epsilon}_{l,t+l}. \quad (2)$$

When $q = 0$, that is, \mathbf{Z}_t is a vector AR(p) process, the partial autoregression matrices for $l > p$ are all zero.

The VARMA process is suitable only if \mathbf{Z}_t is weakly stationary, which means that the mean vector and covariance matrix of \mathbf{Z}_t are both constant over time. If \mathbf{Z}_t is not a stationary time-series, then we may transform it to a stationary one by differencing. This means that when \mathbf{Z}_t is non-stationary, we model $\Delta \mathbf{Z}_t = \mathbf{Z}_t - \mathbf{Z}_{t-1}$ instead. If $\Delta \mathbf{Z}_t$ is still non-stationary, then we move on to $\Delta^2 \mathbf{Z}_t = \Delta \mathbf{Z}_t - \Delta \mathbf{Z}_{t-1}$, and so on. Let d be the number of differencing required to achieve stationarity. We call the resulting process,

$$\Delta^d \mathbf{Z}_t = \mathbf{C}_0 + \sum_{i=1}^p \mathbf{\Phi}_i \Delta^d \mathbf{Z}_{t-i} + \sum_{j=1}^q \boldsymbol{\Theta}_j \boldsymbol{\epsilon}_{t-j} + \boldsymbol{\epsilon}_t,$$

a Vector Autoregressive Integrated Moving-Average VARIMA(p, d, q) process. The triplet (p, d, q) is called the order of the VARIMA process.

3.2 Model Identification

The multivariate generalization of the Box and Jenkins approach for model identification relies heavily on the ‘cut-off’ properties of $\mathbf{\Gamma}(l)$ and $\mathbf{P}(l)$. In practice, matrices $\mathbf{\Gamma}(l)$ and $\mathbf{P}(l)$ are not known, but we can estimate them using the sample cross-correlation matrix (SCCM) and the sample partial autoregression matrix (SPAM), respectively.

Given a time-series data sample $\{\mathbf{Z}_t = (Z_t^{(1)}, Z_t^{(2)})', t = 1, 2, \dots, n\}$, the (i, j) th element of the lag- l SCCM, $\hat{\mathbf{\Gamma}}(l)$, can be calculated with the following formula:

$$\hat{\gamma}_{ij}(l) = \frac{\sum_{t=l+1}^n (Z_t^{(i)} - \bar{Z}^{(i)})(Z_{t-l}^{(j)} - \bar{Z}^{(j)})}{\sqrt{\sum_{t=1}^n (Z_t^{(i)} - \bar{Z}^{(i)})^2 \sum_{t=1}^n (Z_t^{(j)} - \bar{Z}^{(j)})^2}}, \quad (3)$$

where $i, j = 1, 2$, n is the sample size, and $\bar{Z}^{(1)}$ and $\bar{Z}^{(2)}$ are the sample averages of the corresponding component series. If the time-series follows the VARMA process specified by equation

(1), then the standard error of each element of the SCCM is approximately $1/\sqrt{n}$. We can use this property to test the significance of $\mathbf{\Gamma}(l)$. The SCCM helps us identify q in the VARIMA order. It also assists us with determining if \mathbf{Z}_t is stationary or not, because if \mathbf{Z}_t is non-stationary, then the SCCM will show no decay to zero.

The lag- l sample partial autoregression matrix, $\hat{\mathbf{P}}(l)$, and the standard errors of its elements can be obtained by fitting equation (2) by least squares. Tiao and Box (1981) recommend using the likelihood ratio statistic to test the null hypothesis $\mathbf{P}(l) = 0$ against the alternative $\mathbf{P}(l) \neq 0$. To conduct such a test, we compute

$$\mathbf{\Xi}(l) = |\hat{\mathbf{\Sigma}}(l)| / |\hat{\mathbf{\Sigma}}(l-1)|,$$

where $\hat{\mathbf{\Sigma}}(l)$ is the matrix of residual sum of squares and cross products after fitting a vector AR(l) to the data. Using Bartlett's (1938) approximation, the likelihood statistic

$$M(l) = -(n - 3l - \frac{3}{2}) \ln[\mathbf{\Xi}(l)]$$

is, under the null hypothesis, asymptotically distributed as χ^2 with 4 degrees of freedom. The SPAM helps us identify p in the VARIMA order.

In practice the SCCM and SPAM are crowded with numbers, making it difficult to recognize the cut-off patterns. To alleviate the problem, Tiao and Box (1981) suggest summarizing these matrix elements using indicator symbols +, - and ·, where + denotes a value greater than twice the estimated standard error, - denotes a value less than twice the estimated standard error, and · denotes an insignificant value.

We now illustrate how the underlying process for the CBD mortality indexes can be identified. From now on, we set \mathbf{Z}_t to $(\kappa_t^{(1)}, \kappa_t^{(2)})'$, the vector of the two CBD mortality indexes. The first step is to calculate the SCCM and SPAM for \mathbf{Z}_t using the data sample $\{\mathbf{Z}_t = (\kappa_t^{(1)}, \kappa_t^{(2)})', t = 1950, \dots, 2009\}$. The calculated SCCM and SPAM in terms of the (+, -, ·) symbols are displayed in Table 2. Since the SCCM shows no decay to zero, we conclude that \mathbf{Z}_t is non-stationary, and apply first order differencing to it.

By replacing \mathbf{Z}_t in equations (2) and (3) with $\Delta\mathbf{Z}_t$, we calculate the SCCM and SPAM for $\Delta\mathbf{Z}_t$. The calculated values of the SCCM and SPAM for $\Delta\mathbf{Z}_t$ are shown in Table 3. The SCCM now cuts off at lag 1, which means no further differencing is necessary. Also, on the basis of the M -statistic, which has a critical value of $\chi_{4,0.95}^2 = 9.45$ at 5% level of significance, the SPAM cuts off at lag 5. The cut-off patterns of the SCCM and SPAM suggest us to model the CBD mortality indexes with either VARIMA(0,1,1) or VARIMA(5,1,0). However, the estimated VARIMA(0,1,1) process does not pass the diagnostic checks that we are going to detail in Section 3.4. For this reason, a VARIMA(5,1,0) specification is chosen.

The results displayed in Table 3 also point out the inadequacy of modeling the CBD mortality indexes with a bivariate random walk. In particular, a bivariate random walk is not able to pick up any of the many significant serial- and cross-correlations (displayed as '+' and '-' signs) that exist in the CBD mortality indexes after first differencing.

	lag (l)							
	1	2	3	4	5	6	7	8
	(a) Sample cross-correlation matrices (SCCM)							
	$\begin{pmatrix} + & - \\ - & + \end{pmatrix}$	$\begin{pmatrix} + & - \\ - & + \end{pmatrix}$	$\begin{pmatrix} + & - \\ - & + \end{pmatrix}$	$\begin{pmatrix} + & - \\ - & + \end{pmatrix}$	$\begin{pmatrix} + & - \\ - & + \end{pmatrix}$	$\begin{pmatrix} + & - \\ - & + \end{pmatrix}$	$\begin{pmatrix} + & - \\ - & + \end{pmatrix}$	$\begin{pmatrix} + & - \\ - & + \end{pmatrix}$
	(b) Sample partial autoregression matrices (SPAM)							
	$\begin{pmatrix} + & - \\ \cdot & + \end{pmatrix}$	$\begin{pmatrix} + & \cdot \\ + & \cdot \end{pmatrix}$	$\begin{pmatrix} + & \cdot \\ + & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ + & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ + & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & + \\ \cdot & \cdot \end{pmatrix}$
$M(l)$	319.70	15.47	9.87	5.46	6.24	9.13	6.68	1.58

Table 2: The SCCM and SPAM for the vector of the CBD mortality indexes.

	lag (l)							
	1	2	3	4	5	6	7	8
	(a) Sample cross-correlation matrices (SCCM)							
	$\begin{pmatrix} - & - \\ - & - \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} + & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$
	(b) Sample partial autoregression matrices (SPAM)							
	$\begin{pmatrix} \cdot & \cdot \\ \cdot & - \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} + & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} + & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} + & - \\ + & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & - \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$
$M(l)$	12.42	1.88	8.08	7.31	14.88	7.79	4.98	7.81

Table 3: The SCCM and SPAM for the vector of the CBD mortality indexes after first order differencing has been applied.

3.3 Model Estimation

Having chosen the tentative VARIMA order, asymptotically efficient estimates of the model parameters can be obtained by using maximum likelihood. We estimate parameters in the specified VARIMA(5,1,0) model using the conditional likelihood method (see Wilson, 1973). The parameter estimates and their approximate standard errors are displayed in part (a) of Table 4. The standard errors can be used to test for the significance of the parameters.

Further gains in the efficiency of the parameter estimates may be achieved by eliminating parameters that are found to be statistically insignificant. Hence, we impose zero restrictions on the parameters that are found to be insignificant, and re-estimate the remaining parameters by the exact likelihood method (see Reinsel, 1997, Chapter 5). The final parameter estimates are displayed in part (b) of Table 4. Zero restrictions are not imposed on the intercept vector \mathbf{C}_0 , because otherwise the model will imply that the expected values of the CBD mortality indexes in future will remain at their 2009 levels forever. Such a projection result is counter-intuitive.

3.4 Diagnostic Checking

To prevent model mis-specification, a detailed diagnostic analysis of the residuals is needed. This includes an examination of the residuals' SCCM and SPAM, which can be calculated by replacing \mathbf{Z}_t in equations (2) and (3) with the residuals of the estimated model at time t . Table 5 shows the calculated SCCM and SPAM for the residuals from the final estimated model which we presented in part (b) of Table 4. The matrices indicate that the final VARIMA(5,1,0) model provides an adequate fit, as the residuals are free from significant serial- and cross-correlations.

At this stage, the M -statistic provides an additional criterion for checking residual serial correlations. Its values for the first eight lags are all smaller than $\chi_{4,0.95}^2 = 9.45$, the critical value at 5% level of significance. We can therefore conclude that the first eight partial autoregression matrices for the residuals are insignificant.

On top of this model, we have also estimated other VARIMA models that have more compact specifications, including a VARIMA(0,1,1) model that is suggested by the SCCM in Table 3. We found that VARIMA(5,1,0) is the most parsimonious specification that adequately captures the serial- and cross-correlations contained in the data.

4 A Graphical Risk Metric

4.1 K1 and K2 Risks

Assuming the CBD model holds, future death probabilities for all ages depend exclusively on the future values of the two CBD mortality indexes. This dependence motivates us to split longevity risk into two constituents, which we call "K1 risk" and "K2 risk". The former, which arises from

(a) Full Model

C_0	Φ_1	Φ_2	Φ_3
$\begin{pmatrix} -.006 \\ (.005) \\ .0002 \\ (.0016) \end{pmatrix}$	$\begin{pmatrix} -.282 & -1.426 \\ (.149) & (3.400) \\ -.017 & -.363 \\ (.006) & (.144) \end{pmatrix}$	$\begin{pmatrix} -.099 & 1.457 \\ (.141) & (3.582) \\ -.011 & -.054 \\ (.006) & (.152) \end{pmatrix}$	$\begin{pmatrix} .406 & -2.188 \\ (.135) & (3.512) \\ .000 & .101 \\ (.006) & (.149) \end{pmatrix}$
	Φ_4	Φ_5	Σ
	$\begin{pmatrix} .311 & -3.333 \\ (.138) & (3.458) \\ .005 & .238 \\ (.006) & (.147) \end{pmatrix}$	$\begin{pmatrix} .272 & -7.827 \\ (.129) & (3.181) \\ .016 & .009 \\ (.005) & (.135) \end{pmatrix}$	$\begin{pmatrix} .000350 & .000008 \\ .000008 & .000001 \end{pmatrix}$

(b) Final Model

C_0	Φ_1	Φ_2	Φ_3
$\begin{pmatrix} -.008 \\ (.004) \\ .0001 \\ (.0001) \end{pmatrix}$	$\begin{pmatrix} -.277 & 0 \\ (.122) & \\ -.014 & -.326 \\ (.006) & (.121) \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ \\ -.011 & 0 \\ (.005) & \end{pmatrix}$	$\begin{pmatrix} .365 & 0 \\ (.100) & \\ 0 & 0 \end{pmatrix}$
	Φ_4	Φ_5	Σ
	$\begin{pmatrix} .201 & 0 \\ (.102) & \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} .231 & -6.159 \\ (.114) & (2.667) \\ .011 & 0 \\ (.004) & \end{pmatrix}$	$\begin{pmatrix} .000368 & .000008 \\ .000008 & .000001 \end{pmatrix}$

Table 4: Estimation of a VARIMA (5,1,0) process for the vector of the CBD mortality indexes. Standard errors of the parameter estimates are given in parentheses.

	lag (l)							
	1	2	3	4	5	6	7	8
(a) Sample cross-correlation matrices (SCCM)	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$
(b) Sample partial autoregression matrices (SPAM)	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & + \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$	$\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$
$M(l)$	4.11	4.00	2.80	7.01	2.97	2.08	4.00	3.16

Table 5: Diagnostic checking of the residuals of the estimated VARIMA(5,1,0) model for the vector of the CBD mortality indexes.

the uncertainty in $\kappa_t^{(1)}$, is the risk associated with overall mortality improvements, whereas the latter, which arises from the uncertainty in $\kappa_t^{(2)}$, is the risk associated with the differentials in mortality improvements across different age groups.

We can better understand K1 and K2 risks by studying the Cartesian coordinate plane shown in Figure 4. In the diagram, the horizontal and vertical axes respectively represent the realized values of $\kappa_T^{(1)}$ and $\kappa_T^{(2)}$ in a future year T . K1 and K2 risks can be seen as the random variations along the horizontal and vertical dimensions, respectively. The two dotted lines indicate $\hat{\kappa}_T^{(1)}$ and $\hat{\kappa}_T^{(2)}$, the best estimates of $\kappa_T^{(1)}$ and $\kappa_T^{(2)}$, which can be calculated readily by switching off the random terms in the VARIMA process presented in the previous section.

The two dotted lines divide the Cartesian coordinate plane into four quadrants, which facilitate us to compare and contrast the K1 and K2 risks that different financial institutions are subject to. For instance, the lower left-hand quadrant represents the worst outcome to pension plans, in particular those that are closed to new entrants. This is because, as we argued in Section 2.2, their payouts are the highest when future $\kappa_t^{(1)}$'s and $\kappa_t^{(2)}$'s are lower than expected. On the other hand, the lower right-hand quadrant represents the worst outcome to life insurers that focus on selling term life insurances to younger people, because, as we explained in Section 2.2, their payouts are the highest when future $\kappa_t^{(1)}$'s are higher than expected but $\kappa_t^{(2)}$'s are lower than expected.

The diagram in Figure 4 also offers us some new insights into natural hedging. Let us suppose that an insurance company naturally hedges the longevity risk associated with its book of life annuities by acquiring a book of life insurances that were sold to young people. Such an action, according to the diagram, may offset some of the company's exposure to K1 risk, but it may at the same time bring more K2 risk to the company. The potential increase in K2 risk exposure

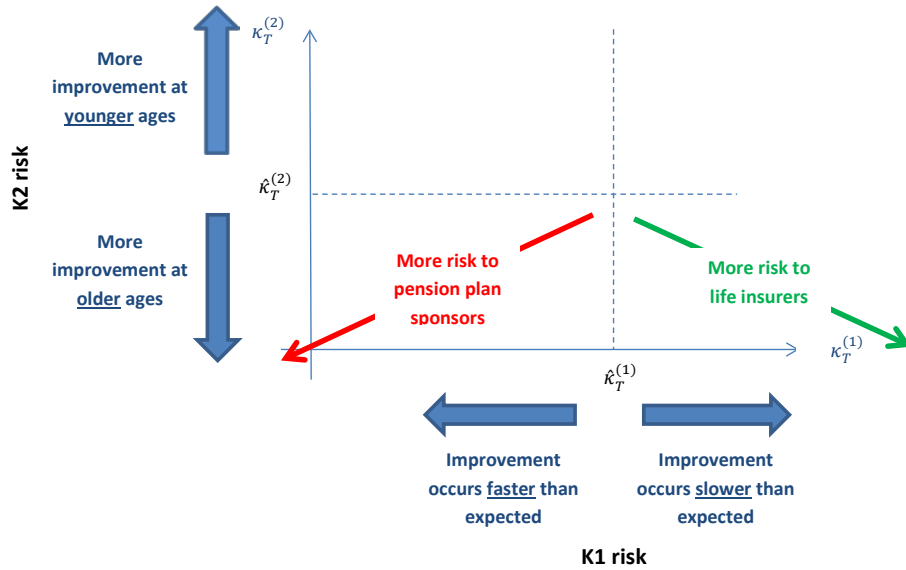


Figure 4: A graphical exposition of K1 and K2 risks.

offers an alternative explanation as to why natural hedging may not work perfectly in practice, even if annuitants and insured lives share exactly the same characteristics such as social economic status and medical underwriting.

The concepts discussed above will be used heavily in the next subsection, where we introduce a graphical risk metric for measuring longevity risk.

4.2 The Concept of Joint Prediction Regions

When quantifying the uncertainty associated with a time-series variable, a marginal prediction interval is often used. We call $[l_t^{(i)}, h_t^{(i)}]$ a marginal prediction interval for the time- t value of the i th CBD mortality index with coverage probability $0 < 1 - \alpha \leq 1$ if

$$\Pr(l_t^{(i)} \leq \kappa_t^{(i)} \leq h_t^{(i)}) = 1 - \alpha, \quad i = 1, 2, \quad t > t_1. \quad (4)$$

The interval $[l_t^{(i)}, h_t^{(i)}]$ should cover $100(1 - \alpha)\%$ of the possible realizations of $\kappa_t^{(i)}$. The values of $l_t^{(i)}$ and $h_t^{(i)}$ can be found readily by using either simulations or analytical means.

A major drawback of the interval defined by equation (4) is that it treats the two CBD mortality indexes in isolation. However, as we have emphasized throughout this paper, the association between the two indexes has important financial implications. We should therefore consider the two indexes jointly in not only modeling but also conveying information about the uncertainty surrounding them.

To better express the uncertainty surrounding the CBD mortality indexes, we can use a joint prediction region instead. We call \mathbf{J}_t a joint prediction region for the time- t values of the CBD

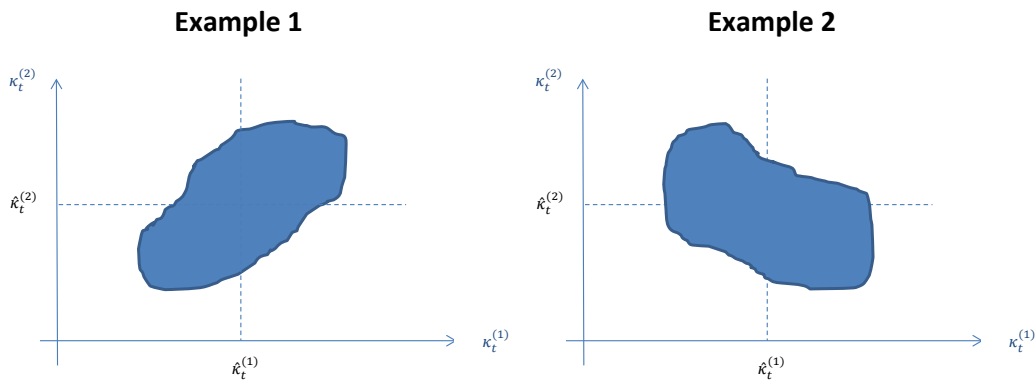


Figure 5: Two hypothetical joint prediction regions for the CBD mortality indexes.

mortality indexes with coverage probability $0 < 1 - \alpha \leq 1$ if

$$\Pr((\kappa_t^{(1)}, \kappa_t^{(2)}) \in \mathbf{J}_t) = 1 - \alpha, \quad t > t_1. \quad (5)$$

The region \mathbf{J}_t should contain $100(1 - \alpha)\%$ of the possible combinations of the two CBD mortality indexes at time t .

When plotted on a Cartesian coordinate plane, the region \mathbf{J}_t can serve as a graphical metric of longevity risk. First, the area of \mathbf{J}_t conveys information about the aggregate level of longevity risk at time t . A larger area means, of course, a higher aggregate level of longevity risk. Second, the shape of the region conveys information about the longevity risk profile of a particular financial institution. To illustrate, let us consider the two hypothetical joint prediction regions displayed in Figure 5.

The areas of these two joint prediction regions are similar, representing similar levels of aggregate longevity risk. However, to different financial institutions, these two joint prediction regions are highly different. Using the arguments presented in previous sections, Example 1 represents more risk for a closed pension plan, while Example 2 represents more risk to an insurance company that focuses on selling term life insurances to younger people. A joint prediction region for the CBD mortality indexes therefore allows us to understand the riskiness of a particular portfolio or financial institution.

4.3 Constructing Joint Prediction Regions

For a given confidence level, a marginal prediction interval is not unique. For example, one can construct a $100(1 - \alpha)\%$ marginal prediction interval for $\kappa_t^{(i)}$ by setting $l_t^{(i)}$ to $-\infty$ and $h_t^{(i)}$ to the $100(1 - \alpha)$ th percentile of $\kappa_t^{(i)}$, or by setting $l_t^{(i)}$ and $h_t^{(i)}$ to the $100\frac{\alpha}{2}$ th and $100(1 - \frac{\alpha}{2})$ th percentiles of $\kappa_t^{(i)}$, respectively.

Likewise, for a fixed value of α , a joint prediction region is not unique. There exist different methods for constructing joint prediction regions (see, e.g., Lütkepohl, 1991; Chan et al., 1999).

We propose to use a numerical method, which is not difficult to implement and is in line with the context of our application. The proposed method is summarized as follows:

- Calculate $\hat{\kappa}_t^{(1)}$ and $\hat{\kappa}_t^{(2)}$, the best estimates of the two CBD mortality indexes, by switching off the random components in the VARIMA model.
- From the VARMA model, simulate, say N , pairs of $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$.
- Calculate $s_t^{(1)}$ and $s_t^{(2)}$, the standard deviations of the simulated $\kappa_t^{(1)}$'s and $\kappa_t^{(2)}$'s.
- For each simulated pair of $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$, calculate its weighted distance to the best estimate with the following formula:

$$\sqrt{\left(\frac{\kappa_t^{(1)} - \hat{\kappa}_t^{(1)}}{s_t^{(1)}}\right)^2 + \left(\frac{\kappa_t^{(2)} - \hat{\kappa}_t^{(2)}}{s_t^{(2)}}\right)^2}.$$

- Sort the N simulated pairs of $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ by their distances to the best estimate. Choose the $\lceil(1 - \alpha)N\rceil$ pairs with the shortest distances.
- Draw a convex hull to enclose the $\lceil(1 - \alpha)N\rceil$ chosen pairs of simulated $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$. Geometrically speaking, the convex hull is the smallest convex set that contains the selected $\lceil(1 - \alpha)N\rceil$ points. The use of a convex hull prevents the joint prediction region from overstating the underlying uncertainty.
- The convex hull drawn is a $100(1 - \alpha)\%$ joint prediction region for the two CBD mortality indexes. By construction, it contains a randomly selected pair of indexes in the simulated sample with probability (no less than) $1 - \alpha$. Note that the coverage probability might be slightly greater than $1 - \alpha$ due to rounding.

4.4 Illustrations

On the basis of the VARIMA(5,1,0) model estimated in Section 3, we construct 99.5% joint prediction regions for the CBD mortality indexes for English and Welsh males in years 2022, 2032 and 2042. The results, which are based on $N = 5000$ simulations, are presented in Figure 6. In each diagram, the dots represent the simulated index values, the dotted lines represent the best estimates of the indexes, and the polygon represents the estimated joint prediction region.

The following observations can be made regarding the patterns of the joint prediction regions:

1. The area of the joint prediction region becomes larger over time. This agrees with our expectation that forecast uncertainty increases as we predict farther into the future.
2. As we predict farther into the future, the vertical line shifts leftwards, indicating that the overall level of mortality is reducing. On the other hand, the horizontal line shifts upwards

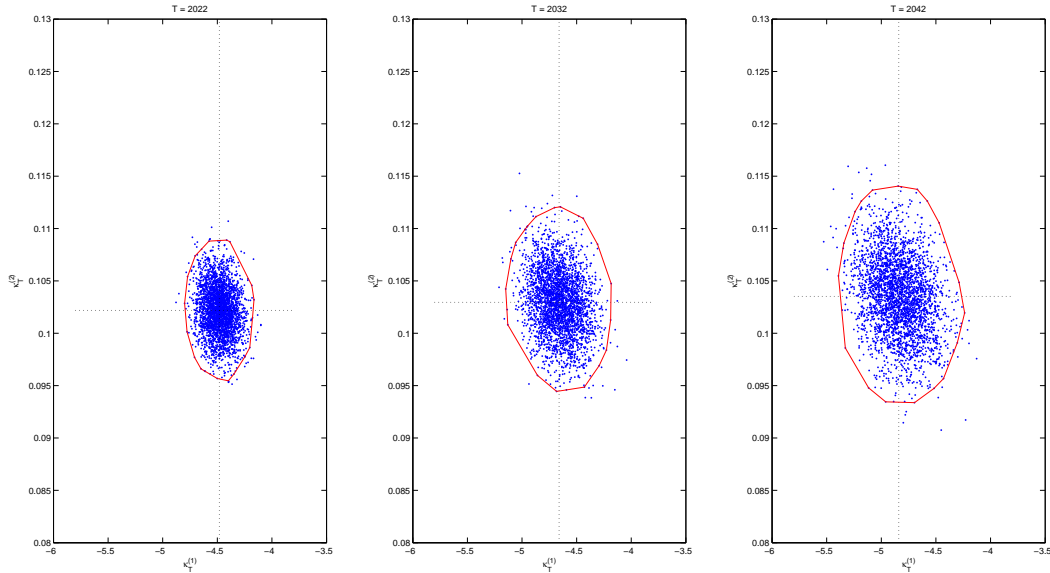


Figure 6: Joint prediction regions for the CBD mortality indexes for English and Welsh males in years 2022, 2032 and 2042. The simulated values are produced from the best fitting VARIMA(5,1,0).

slowly, indicating that that mortality at younger ages (below the mean age $\bar{x} = 65$) is improving (slightly) faster than at older ages. Overall, the centroid of the joint prediction region moves to the upper-left-hand corner over time.

3. The tilts of the joint prediction regions are minimal. For a given year, the portions of the joint prediction region falling into the two lower quadrants are similar in size. We conclude that for this particular population, pension providers and life insurers are subject to similar levels of risk.

To further illustrate the concept, we also produce joint prediction regions for Canadian males. The data for this population were also obtained from the Human Mortality Database (2012). We use the same age range (40-90) as before, but a shorter sample period (1950-2007). The sample period here ends in 2007, because mortality data for this population are available only up to that year. By fitting a CBD model to the data, we calculate the values of the CBD mortality indexes for Canadian males over the sample period. We then model the CBD mortality indexes for this population using a VARIMA(3,1,0) model, which is identified and estimated using the methods described in Section 3.

By using $N = 5000$ simulations from the VARIMA(3,1,0) model, we obtain 99.5% joint prediction regions for the CBD mortality indexes for Canadian males in years 2022, 2032 and 2042 (see Figure 7). The axes in this figure are the same as those in Figure 6, allowing us to compare

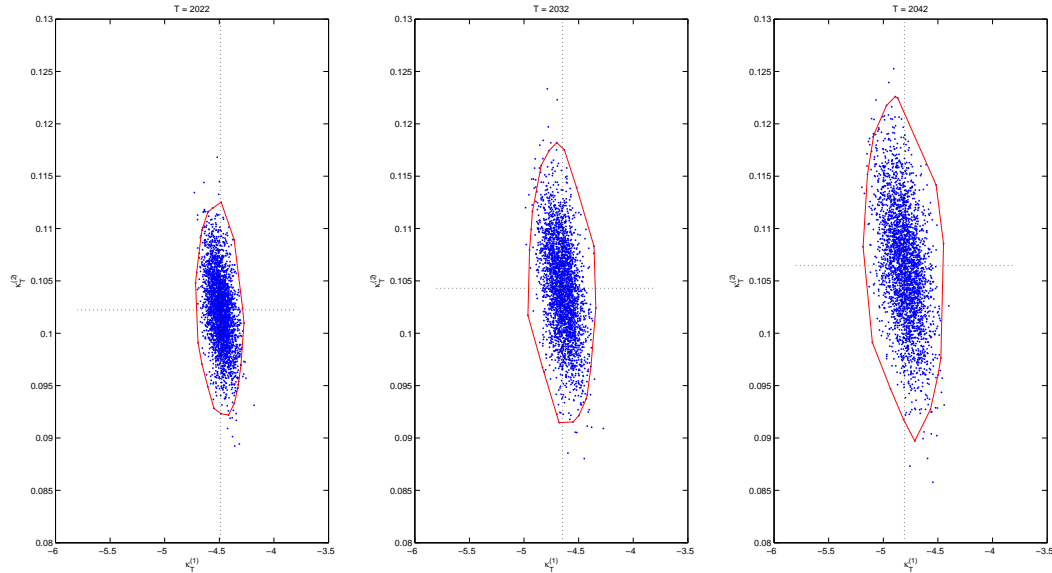


Figure 7: Joint prediction regions for the CBD mortality indexes for Canadian males in years 2022, 2032 and 2042. The simulated values are produced from the best fitting VARIMA(3,1,0).

the joint prediction regions of the two populations. The following interesting points can be made from the comparison:

1. The joint prediction regions for Canadian males are narrower and taller than the corresponding ones for English and Welsh males. This means that Canadian males are subject to less K1 risk but more K2 risk than English and Welsh males. In laymen's terms, what this means is that for Canadian males, there is less uncertainty associated with overall mortality improvements, but more uncertainty associated with how mortality improvements are different among different age groups.
2. The joint prediction regions for Canadian males appears to be more tilted. In particular, we observe that for each year, the joint prediction region and the simulated index values roughly form a diagonal, running from the upper-left-hand to the lower-right-hand corner.
3. Let us focus on the lower part of each joint prediction region. The lower-left-hand quadrant is much smaller than the lower right quadrant. A pattern like this means that in Canada, life insurers selling term life insurance are subject to relatively more risk than pension plan providers.

5 Discussion and Conclusion

In this paper, we studied the possibility of constructing mortality indexes with a model-based approach. We proposed to use parameters $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ in the original CBD model as mortality indexes, because they satisfy three properties, namely (1) able to represent a varying age-pattern of mortality improvement, (2) new-data-invariant, and (3) interpretable, which we believe to be crucial to model-based mortality indexes. On top of that, the CBD mortality indexes also meet other criteria, for example, objective and transparent, that mortality indexes in general should fulfill.

We argued, with some examples, that the association between the two CBD mortality indexes has a significant impact on the longevity risk exposure of a portfolio. To address this issue, we proposed to model the dynamics of the two mortality indexes jointly with a VARIMA process. Relative to a random walk, which is often employed to model $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$, a VARIMA process can better capture the cross-correlations between the two mortality indexes. We also proposed to express the uncertainty surrounding the two CBD mortality indexes simultaneously with a joint prediction region. Such a region, as we demonstrated, can be used as a graphical longevity risk metric, allowing practitioners to compare the longevity risk exposures of different portfolios readily.

Ignoring the potential cross- and serial correlations in the historical index data may result in a mis-estimation of K1 and K2 risks. In Figure 8 we show the joint prediction regions that are created from a random walk with drift estimated to the historical CBD mortality indexes for English and Welsh males. To evaluate the importance of modeling cross- and serial correlations, let us compare the diagrams in Figure 8 with those in Figure 6, which are created on the basis of an optimal VARIMA(5,1,0) model. The two processes produce very similar central estimates of future $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ values, but they predict very different associations between the two indexes. As previously mentioned, the association between the indexes has a significant impact on a portfolio's longevity risk exposure, and therefore using a random walk with drift may lead us to an incorrect conclusion about the riskiness of a portfolio.

A limitation of our modeling approach is that it does not incorporate interrupting phenomena such as trend changes and jumps, which are sometimes observed in the CBD mortality indexes, especially those in the early 20th century. One possible avenue of future research is to generalize the basic VARIMA model to include features such as structural breakpoints (Li et al., 2011) and outliers (Li and Chan, 2005, 2007). It would also be interesting to see how the resulting joint prediction regions for the CBD mortality indexes will change when these additional features are incorporated.

Another caveat is that in generating the joint prediction regions in Section 4.4, we did not incorporate parameter risk, which arises because the true parameter values can never be known. In future research, it is warranted to include parameter uncertainty in the estimation and simulation

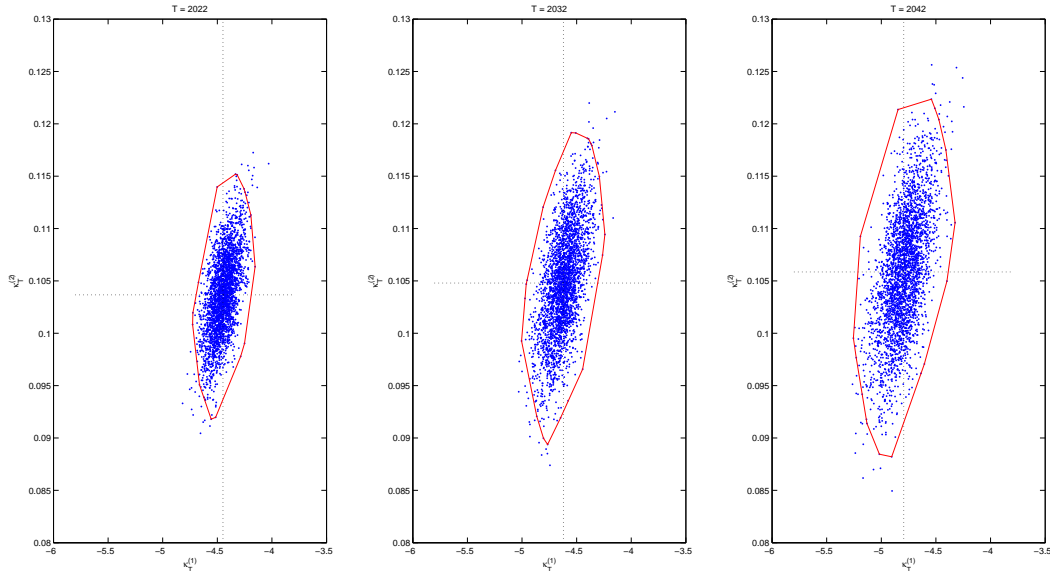


Figure 8: Joint prediction regions for the CBD mortality indexes for English and Welsh males in years 2022, 2032 and 2042. The simulated values are produced by a random walk with drift.

procedures, possibly by using Bayesian approaches, in which parameters are treated as random variables instead of fixed constants. By conducting the required simulations in a Bayesian set-up, we can produce joint prediction regions that take parameter risk into account.

It is possible to write derivative securities that are linked to the two CBD mortality indexes. Inspired by the previously developed q-forwards and S-forwards,⁵ we propose, as a concept, a standardized longevity security called K-forward, which can be regarded as a zero-coupon swap that exchanges on the maturity date a fixed amount for a random amount that is proportional to a CBD mortality index at some future time. Mathematically, to the fixed receiver, the payoff from a K-forward that is linked to the i th CBD mortality index is

$$X \times (\tilde{\kappa}_{t^*}^{(i)} - \kappa_{t^*}^{(i)}), \quad i = 1, 2,$$

where X is the notional amount, t^* is the maturity date, and $\tilde{\kappa}_{t^*}^{(i)}$ is the forward value of the CBD mortality index.⁶ The forward value is determined at time 0, and is chosen in such a way that no payment exchanges hands at the inception of trade. A pension plan wishing to hedge its longevity risk exposure may write, as a fixed receiver, K-forwards that are linked to the two CBD mortality indexes. In this way, payouts from the K-forwards can offset the worst outcome, which happens when both indexes turn out to be lower than expected.

⁵We refer readers to Coughlan (2009) and the website of the Life and Longevity Markets Association (LLMA) for details regarding q-forwards and S-forwards.

⁶For simplicity, we ignore the potential lag in the availability of the index data.

In a parallel study, Tan et al. (2013) propose a mortality duration measure that permits us to calibrate a longevity hedge that is formed by a collection of K-forwards. They found that a longevity hedge created by K-forwards can reduce the variance of a pension plan's present value by more than 90%. The trading of K-forwards thus has a strong potential to bring economic benefits to hedgers. Moreover, to achieve the same level of hedge effectiveness, a K-forward hedge requires fewer contracts than a q-forward hedge does. It follows that the demand per K-forward contract is likely to be higher than that per q-forward contract. Consequently, it may be easier for the market to concentrate liquidity if K-forwards are traded instead of other standardized contracts.

Tan et al. (2013) also investigate the potential benefit from increasing the number of mortality indexes. They found that adding a time-varying parameter that measures the curvature in the age effect to the original CBD model sometimes improves goodness-of-fit, depending on the age range and the population being modeled. Including K-forward contracts linked to this additional CBD mortality index may also enhance hedge effectiveness by up to 2.7 percentage points. To visualize the potential associations among all three indexes, they generalize the joint prediction region in Section to a three-dimensional version, in which the volume of the region represents the aggregate level of risk that an insurer or pension plan sponsor is facing. The pairwise association between any two of the three indexes can be visualized by examining the corresponding cross-section of the three-dimensional joint prediction region.

Acknowledgments

The authors gratefully acknowledge financial support from SCOR on the longevity project, from which this paper is extracted. This work is also supported by a Discovery Grant from the Natural Science and Engineering Research Council of Canada, and a research grant from the Research Grants Council of the Hong Kong Special Administrative Region (General Research Fund Project No. CUHK 440812H).

References

- Bartlett, M.S. (1938). Further aspects of the theory of multiple regression. *Proceedings of the Cambridge Philosophical Society*, **34**, 33-40.
- Blake, D., Cairns, A.J.G. and Dowd, K. (2008). Longevity risk and the Grim Reaper's toxic tail: the survivor fan charts. *Insurance: Mathematics and Economics*, **42**, 1062-1066.
- Box, G.E.P. and Jenkins, G.M. (1976). *Time series analysis: forecasting and control*. Second Edition. Holden-Day Press, San Francisco.
- Chan, W.S., Cheung, S.H. and Wu, K.H. (1999). On exact joint forecast regions for vector autoregressive models. *Journal of Applied Statistics*, **26**, 35-44.

- 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.
- Coughlan, G. (2009). Longevity risk transfer: Indices and capital market solutions. In Barriau, P.M. and Albertini, L. (eds.) *The Handbook of Insurance Linked Securities*. London: Wiley.
- Dowd, K., D. Blake, and A.J.G. Cairns (2010). Facing up to uncertain life expectancy: the longevity fan charts. *Demography*, **47**, 67-78.
- Dowd, K., Cairns, A.J.G., Blake, D., Coughlan, G.D., Epstein, D., and Khalaf-Allah, M. (2010). Back-testing stochastic mortality models: An ex-post evaluation of multi-period-ahead density forecasts. *North American Actuarial Journal*, **14**, 281-298.
- Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute of Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded on 1 August 2012).
- Li, S.H. and Chan, W.S. (2005). Outlier analysis and mortality forecasting: The United Kingdom and Scandinavian countries. *Scandinavian Actuarial Journal*, **3**, 187-211.
- Li, S.H. and Chan, W.S. (2007). The Lee-Carter model for forecasting mortality revisited. *North American Actuarial Journal*, **11**, 68-89.
- Li, J.S.-H., Chan, W.S. and Cheung, S.H. (2011). Structural changes in the Lee-Carter mortality indexes: detection and implications. *North American Actuarial Journal*, **15**, 13-31.
- Li, J.S.-H. and Luo, A. (2012). Key q-duration: A framework for hedging longevity risk. *ASTIN Bulletin*, in press.
- Lütkepohl, H. (1991). Introduction to multiple time series analysis. Springer-Verlag, Berlin.
- Reinsel, G.C. (1997). *Elements of multivariate time series analysis*. Second Edition. Springer Verlag.
- Sweeting, P.J. (2010). Longevity indices and pension fund risk. Pension Institute Discussion Paper PI-1004
- Sweeting, P.J. (2011). A trend-change extension of the Cairns-Blake-Dowd model. *Annals of Actuarial Science*, **5**, 143-162.
- Tan, C.I., Li, J, Li, J.S.-H., and Balasooriya, U. (2013). Parametric Mortality Indexes: From Index Construction to Hedging Strategies. Working paper.
- Tiao, G.C. and Box, G.E.P. (1981). Modelling multiple time series with applications. *Journal of the American Statistical Association*, **76**, 802-816.
- Wei, W.W.S. (2006). *Time Series Analysis: Univariate and Multivariate Method*. Second Edition. Addison Wesley/Pearson.

Wilson, G.T. (1973). The estimation of parameters in multivariate time series models. *Journal of the Royal Statistical Society*, **B35**, 76-85.