risks

MDPI

Article

Disentangling Trend Risk and Basis Risk with Functional Time Series

Yanxin Liu [1,*] and Johnny Siu-Hang Li [2]

[1] Department of Finance, University of Nebraska-Lincoln, Lincoln, NE 68588, USA
[2] Department of Finance, The Chinese University of Hong Kong, Hong Kong SAR, China; johnny.li@cuhk.edu.hk
* Correspondence: yanxin.liu@unl.edu

Abstract: In recent multi-population stochastic mortality models, one critical scientific issue is the vague distinction between trend risk and population basis risk. In particular, the cross- and auto-correlations between the innovations of the latent factors representing the common trend and the population-specific trends are often assumed to be non-existent, although they are possibly statistically significant. While it is theoretically possible to capture such correlations by treating the latent factors as a vector time series, the resulting model would contain a large number of parameters, which may in turn lead to robustness problems. In this paper, we address these issues by the use of the product–ratio model. Contrary to the prevalent assumption of non-existent correlations, the latent factors under the product–ratio model are approximately uncorrelated. This permits us to disentangle trend risk and population basis risk, thereby sparing us from the need to use a heavily parameterized vector time-series process. Compared to the augmented common factor model, our approach demonstrates improved robustness in terms of correlation structures and hedging performance, offering a new perspective on treating cross- and auto-correlations between latent factors in mortality modeling.

Keywords: population basis risk; mortality model; functional time series; index-based hedge

Citation: Liu, Yanxin, and Johnny Siu-Hang Li. 2023. Disentangling Trend Risk and Basis Risk with Functional Time Series. *Risks* 11: 208. https://doi.org/10.3390/risks11120208

Academic Editor: Montserrat Guillén

Received: 30 September 2023
Revised: 14 November 2023
Accepted: 25 November 2023
Published: 28 November 2023

1. Introduction

Recent advancements in public health and medical systems have significantly improved human mortality rates. For instance, the life expectancy at birth in the U.S. has increased dramatically from 60.91 in 1935 to 78.70 in 2010, a 30% increase over 75 years[1]. This pattern is also observed in other developed countries such as Canada. However, this remarkable progress in longevity is sometimes viewed as a social or economic burden. This is because the increased life expectancy substantially extends the duration of pension payments, posing a financial risk to insurance companies and pension plan providers. In actuarial studies, this is referred to as mortality risk (or longevity risk), which is the risk that insurance companies face financial strain due to policyholders living longer than expected.

Mortality risk is generally categorized into two types: trend risk and idiosyncratic risk. Idiosyncratic risk refers to the uncertainty associated with the realized death probability. This risk can be mitigated effectively by increasing the population base, thanks to the law of large numbers. Trend risk, however, is associated with the uncertainty in the increasing trends of life expectancy. It is a systematic risk that affects everyone and cannot be reduced simply by expanding the population base.

In longevity research, various stochastic mortality models have been developed to address mortality trend risk. The Lee–Carter model (Lee and Carter 1992), for example, decomposes log-central death rates into an age-period structure. The period effects, which reflect the dynamics of mortality over time, are modeled as a time series (typically random walk) to forecast future mortality, capturing trend risk. The CBD model (Cairns et al. 2006)

focuses on modeling death probability, with its decomposition including two period effects modeled as a bi-variate random walk, thus addressing mortality trend risk. Some researchers have extended these models to include cohort effects, assuming that mortality trend risk also exists along the year-of-birth direction. Notable models in this category include the Renshaw–Haberman model (Renshaw and Haberman 2006), models M6/M7/M8 (Cairns et al. 2009), and the Plat model (Plat 2009). More recently, the Heat-Wave model Li and Liu (2020) was proposed, differentiating short-term and long-term mortality improvements in its assessment of trend risk.

To manage emerging mortality risks, many pension providers use insurance-based solutions like buy-ins/buy-outs and longevity swaps. However, as some studies suggest (e.g., Graziani (2014) and Michaelson and Mulholland (2014)), the insurance industry can only absorb a limited portion of longevity risk, highlighting the need to find additional participants, such as capital market investors. In standardized index-based mortality derivatives, the payouts are linked to standardized mortality indexes. These financial products appeal to a broader range of participants because of their cost efficiency, high liquidity, and the way they mitigate concerns about information asymmetry. However, the use of standardized index-based mortality derivatives introduces population basis risk, stemming from the mismatch between the population tied to the liability and the one linked to the hedging instrument. Traditional mortality models, which typically focus on single populations, do not account for this risk.

To address population basis risk, various multi-population extensions to traditional mortality models have been proposed. For example, based on the Lee–Carter modeling framework, Li and Lee (2005) proposed the augmented common factor model (ACF model), which assumes the existence of a common factor governing the mortality co-movement of all the underlying populations. Any deviations from the common trend are then absorbed by the population-specific factors. The ACF structure is then applied by Lyu et al. (2021) to model cause-of-death mortality data. Russolillo et al. (2011) considered a three-way Lee-Carter model which incorporates the population effects in addition to the age and period effects in the original Lee-Carter model. This modeling approach is later generalized by Dong et al. (2020) via the use of Tucker decomposition. Based on the CBD framework, Cairns et al. (2011) proposed a mortality model designed for two populations of similar size. Dowd et al. (2011), on the other hand, developed a gravity model for populations with imbalance size. Notably, under the gravity model, the population with a bigger size is the main driving force while the population with a smaller size is orbiting the other population, similar to being attracted by gravity. Zhou et al. (2014, 2019) considered a multi-population model where the underlying period effects from different populations are assumed to follow the vector auto-regressive (VAR) model and the vector error correction (VECM) model. Furthermore, Chen et al. (2015) used a factor copula approach to capture mortality dependence among different populations. Tsai and Zhang (2019) combined Bühlmann credibility with a multi-population mortality model and proposed a new modeling method called the multi-dimensional Bühlmann credibility approach.

In most existing multi-population models, the distinction between trend risk and population basis risk is vague. In particular, the cross- and auto-correlations between the innovations of the latent factors representing the common trend and the population specific trends are often assumed to be non-existent, although they are possibly statistically significant. While it is theoretically possible to capture such correlations by treating the latent factors as a vector time series, the resulting model would contain a large number of parameters, which may in turn lead to robustness problems.

In this paper, we address the issue by the use of the product–ratio model (PR model) developed by Hyndman et al. (2013). The PR model includes a product model that uses the geometric mean of sub-populations to capture the overall mortality trend, and a ratio model that specifically models deviations between sub-populations. Our paper contributes to the literature in the following aspects. First, we specifically target the cross- and/or auto-correlation of latent factors, which is assumed to be non-existent and not studied

extensively in the previous literature. Second, we utilize the property of the PR model that the latent factors are approximately uncorrelated and provides a new perspective on the disentangling of trend risk and population basis risk. Third, via the numerical comparison between the ACF model and the PR model, we found that the approach we consider yields term structures of correlations and hedging performances that are more robust with respect to how cross- and auto-correlations between the latent factors are treated.

The remainder of the paper is organized as follows. Section 2 describes the multi-population mortality models considered in this paper. Specifically, we consider the augmented common factor model and the product–ratio model. Section 3 provides the estimation results of the models and conducts the numerical analysis on various time series structures for the assumed models. Section 4 studies the hedge performance of the assumed models. And finally, Section 5 concludes the paper.

2. Mortality Models for Multiple Populations

2.1. Augmented Common Factor Model

We first present the augmented common factor model (the ACF model) developed by Li and Lee (2005), which can be used to model mortality rates for multiple populations. The ACF model extends the Lee–Carter model (Lee and Carter 1992) to include a common factor that describes the mortality trend shared by all populations, and a population-specific factor which accommodate any population-specific deviation from the common trend. The ACF model with two populations is defined as follow:

$$\log(m_{x,t}^{(i)}) = a_x^{(i)} + B_x K_t + b_x^{(i)} k_t^{(i)} + e_{x,t}^{(i)} \quad \text{for } i = 1, 2, \tag{1}$$

where $m_{x,t}^{(i)}$ represents the central death rate for individuals at age x in year t from population i. Parameters $a_x^{(i)}$, B_x, and $b_x^{(i)}$, are the age-specific parameters, with $a_x^{(i)}$ representing the level of (log-)mortality at age x from population i, B_x and $b_x^{(i)}$ capturing the sensitivity of mortality rate to changes in the common trend and population-specific deviation. Parameter K_t represents the co-movement of mortality shared by all populations and would generally be modeled by a random walk. Parameters $k_t^{(1)}$ and $k_t^{(2)}$ are population-specific deviations from the common trend, which would be modeled by some mean-reverting processes.

We assume the error terms $e_{x,t}^{(i)}$ in Equation (1) are normally distributed, that is, $e_{x,t}^{(i)} \overset{iid}{\sim} N(0, \sigma_i^2)$. The parameter estimation of the ACF model is then achieved through maximum likelihood estimation (MLE). We assume $\sum_x B_x = \sum_x b_x^{(i)} = 1$ and $\sum_t K_t = \sum_t k_t^{(i)} = 0$ to ensure the uniqueness of parameter estimation (Hunt and Blake 2020).

After obtaining the MLE estimates, the sequences of K_t, $k_t^{(1)}$, and $k_t^{(2)}$ would be fitted to some time-series for the purpose of mortality forecast. Based on the same MLE estimates, we consider three time-series structures, denoted by the ACF0 model, ACF1 model, and ACF2 model, respectively. The definition of these three variants are summarized below:

- ACF0 Model:

  In the ACF0 model, we assume the three mortality indices K_t, $k_t^{(1)}$, and $k_t^{(2)}$ are independent. Under this assumption, K_t, $k_t^{(1)}$, and $k_t^{(2)}$ can be modeled by the following expression:

$$\begin{cases} \Delta K_t &= \mu^c + \eta_t^c \\ k_t^{(1)} &= \mu^{(1)} + \phi^{(1)} k_{t-1}^{(1)} + \eta_t^{(1)} \\ k_t^{(2)} &= \mu^{(2)} + \phi^{(2)} k_{t-1}^{(2)} + \eta_t^{(2)} \end{cases},$$

  where

$$\begin{pmatrix} \eta_t^c \\ \eta_t^{(1)} \\ \eta_t^{(2)} \end{pmatrix} \sim MVN(\vec{0}, Q), \quad Q = \begin{pmatrix} Q_{11} & 0 & 0 \\ 0 & Q_{22} & 0 \\ 0 & 0 & Q_{33} \end{pmatrix}.$$

There are eight parameters in this model, which include $\mu^c, \mu^{(1)}, \mu^{(2)}, \phi^{(1)}, \phi^{(2)}, Q_{11}, Q_{22}$, and Q_{33}.

- ACF1 Model:
  The ACF1 model extends the ACF0 model to incorporate the correlations between different sequences. As a result, the Q matrix is no longer diagonal and we have

$$\begin{pmatrix} \eta_t^c \\ \eta_t^{(1)} \\ \eta_t^{(2)} \end{pmatrix} \sim MVN(\vec{0}, Q), \quad Q = \begin{pmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{12} & Q_{22} & Q_{23} \\ Q_{13} & Q_{23} & Q_{33} \end{pmatrix}.$$

  Comparing to the ACF0 model, the number of unknown parameters increase from eight to eleven due to the addition of off-diagonal covariance parameters Q_{12}, Q_{13}, and Q_{23}.
- ACF2 Model:
  In the ACF2 model, we consider a more complicated autoregressive structure, the vector autoregressive (VAR) structure, for K_t, $k_t^{(1)}$, and $k_t^{(2)}$ sequences. We consider the simplest lag-1 vector autoregression model, the VAR(1) model, which can be expressed as follow:

$$\begin{cases} \Delta K_t &= \mu^c + \phi_{11}\Delta K_{t-1} + \phi_{12}k_{t-1}^{(1)} + \phi_{13}k_{t-1}^{(2)} + \eta_t^c \\ k_t^{(1)} &= \mu^{(1)} + \phi_{21}\Delta K_{t-1} + \phi_{22}k_{t-1}^{(1)} + \phi_{23}k_{t-1}^{(2)} + \eta_t^{(1)} \\ k_t^{(2)} &= \mu^{(2)} + \phi_{31}\Delta K_{t-1} + \phi_{32}k_{t-1}^{(1)} + \phi_{33}k_{t-1}^{(2)} + \eta_t^{(2)} \end{cases},$$

  where

$$\begin{pmatrix} \eta_t^c \\ \eta_t^{(1)} \\ \eta_t^{(2)} \end{pmatrix} \sim MVN(\vec{0}, Q), \quad Q = \begin{pmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{12} & Q_{22} & Q_{23} \\ Q_{13} & Q_{23} & Q_{33} \end{pmatrix}.$$

  In this model, a fully vectorized autoregressive structure is used. Even in the simplest case, the VAR(1) structure, the number of unknown parameters increase to eighteen. For a more complicated autoregressive VAR(p) ($p > 1$) structure, the number of unknown parameters would be $(18 + 9 \times (p-1))$.

2.2. Product–Ratio Model

In the traditional modeling approach, the number of unknown parameters increases dramatically when we additionally consider cross- and auto-correlations between the innovations of the common trend and the population specific trends. To overcome this over-parameterization problem, we consider the product–ratio model (the PR model) by Hyndman et al. (2013), under which the innovations are uncorrelated by construction. The PR model includes two parts, the product model and the ratio model, which are defined as follow:

- The product model is defined by

$$\log(g_{x,t}) = \mu_x^p + \sum_{i=1}^{m} B_{i,x}K_{i,t} + e_{x,t}, \tag{2}$$

  where $g_{x,t} = \sqrt{m_{x,t}^{(1)} m_{x,t}^{(2)}}$ is the square root of the product of $m_{x,t}^{(1)}$ and $m_{x,t}^{(2)}$, μ_x^p is the average level at age x for the product model, m is the number of components being considered, $e_{x,t}$ is the error.
  In the product model, the quantity $\log(g_{x,t})$ captures the average of (log-) mortality rates under different populations. The product of $B_{i,x}$ and $K_{i,t}$ is then used to reflect the evolution of (average) mortality at age x.

- The ratio model is defined by

$$\log(r_{x,t}) = \mu_x^r + \sum_{i=1}^{n} b_{i,x} k_{i,t} + w_{x,t} \tag{3}$$

where $r_{x,t} = \sqrt{m_{x,t}^{(1)} / m_{x,t}^{(2)}}$ is the square root of the ratio of $m_{x,t}^{(1)}$ and $m_{x,t}^{(2)}$, μ_x^r is the average level at age x for the ratio model, n is the number of components being considered, $w_{x,t}$ is the error.
In the ratio model, the quantity $\log(r_{x,t})$ captures the deviation between the two populations. The product of $b_{i,x}$ and $k_{i,t}$ is then used to reflect the evolution of such deviation at age x.

The parameter estimation of the PR model is achieved through the demography package (Hyndman 2023) in R. We specify $m = 1$ and $n = 2$ such that the PR model is comparable to the ACF model defined in the previous section. The notation of $B_{1,x}$ and $K_{1,t}$ in the product model can be simplified to B_x and K_t, respectively.

The PR model does not directly model the log-central death rate. Instead, using Equations (2) and (3), the expression of the (log-) central death rates under the PR model can be computed via:

$$\log(m_{x,t}^{(1)}) = (\mu_x^p + \mu_x^r) + B_x K_t + b_{1,x} k_{1,t} + b_{2,x} k_{2,t} + e_{x,t} + w_{x,t},$$

$$\log(m_{x,t}^{(2)}) = (\mu_x^p - \mu_x^r) + B_x K_t - b_{1,x} k_{1,t} - b_{2,x} k_{2,t} + e_{x,t} - w_{x,t}.$$

Similar to the ACF model, the sequences of K_t, $k_{1,t}$, and $k_{2,t}$ would be fitted to some time-series for the purpose of forecasting future mortality rates. We consider three variants of the PR model that are the same as those described in Section 2.1.

3. Numerical Analysis

3.1. Data

For illustration purpose, we consider U.S. males and Canadian males with age range between 20 and 100, and a sample period from 1950 to 2019[2]. We obtain the mortality data for these two populations from the Human Mortality Database (HMD). The observations from the most recent years (i.e., year 2020 and year 2021) are excluded from this study to avoid the impact of COVID-19, so that we can be more focused on trend risk and basis risk.

3.2. Parameter Estimation

We fit the ACF model and the PR model to the U.S. and Canadian male populations. The parameter estimates for the ACF model and the PR model are shown in Figure 1 and Figure 2, respectively. The following observations can be made:

- The ACF Model
  The ACF model uses the common factor to capture the general mortality trend shared by all populations, and the population-specific factor to capture any deviation from the common trend. The estimated values of K_t are therefore representing the co-movement of the mortality dynamic shared by the two countries, and B_x representing the common age sensitivity to such co-movement.
  For the U.S. and Canada, although they are not identical, these two countries do share a great amount of similarities in economy, culture, living style, and healthcare system. As a result, the estimated levels $a_x^{(i)}$ as well as the population-specific factors $b_x^{(i)}$ and $k_t^{(i)}$ from the two countries are quite similar to each other, causing a vague distinction between trend risk and basis risk.
- The PR Model
  The PR model, on the other hand, decomposes trend risk and population basis risk by construction, meaning that these two risks are separated in the first place when

the product model and ratio model are defined. As a result, the product model would be solely focusing on trend risk while the ratio model is solely focusing on basis risk, respectively, leading to a clear distinction between the two.

- In the product model, the estimated μ_p represents the mean level of the general trend. It carries very similar shape to both of $a_x^{(1)}$ and $a_x^{(2)}$ in the ACF model. The estimated K_t carries similar downward pattern as that in the ACF model, representing general mortality improvement overtime. For the estimated B_x, it presents similar peak (around age 70) and trough (around age 30) to those shown in the ACF model, with different magnitudes.
- In the ratio model, the estimated μ_r describes the difference in the mean level between the two populations. It can be easily observed in the pattern of μ_r that the biggest difference between the two populations lies in the working age males (from 25 to 65). The gap closes out quickly toward older ages.
  Besides the difference in the mean level, the ratio model further decomposes the difference between the U.S. and Canada into two components, $b_x^{(1)}k_t^{(1)}$ and $b_x^{(2)}k_t^{(2)}$. The upward pattern of $k_t^{(1)}$ suggests that the differences between the two countries are getting bigger over time, and the pattern of $b_x^{(1)}$ shows such a gap is more significant in young adults around age 20. The remaining residuals not explained through the first component are then captured by the second component $b_x^{(2)}k_t^{(2)}$.

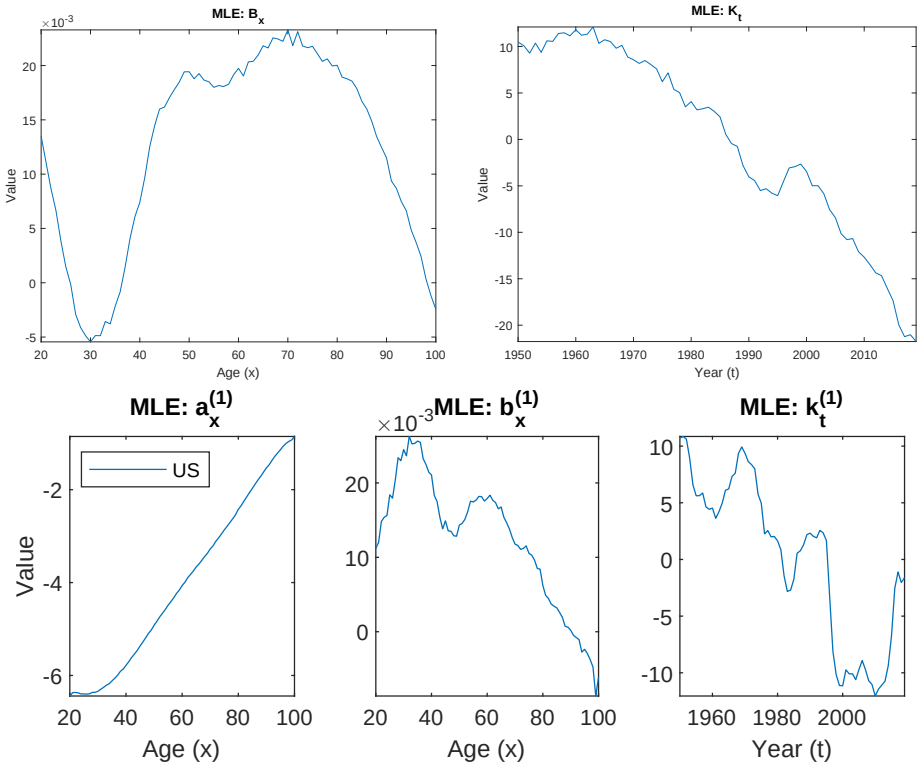

Figure 1. *Cont.*

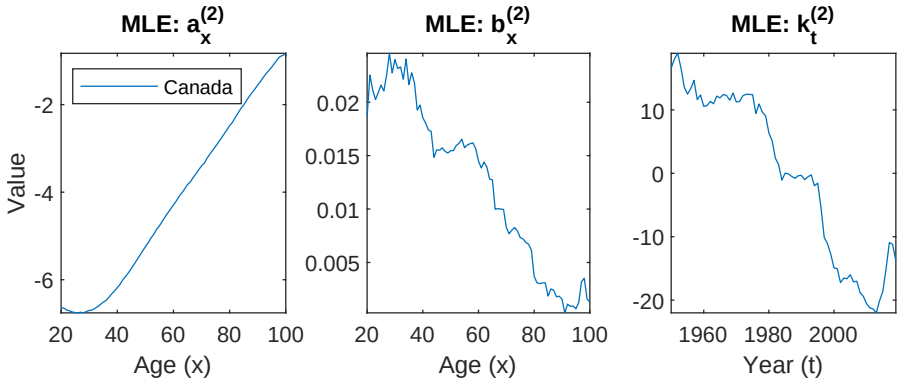

Figure 1. Parameter estimates of the Augmented Common Factor model fitted to the U.S. and Canadian males with ages from 20 to 100 and years from 1950 to 2019.

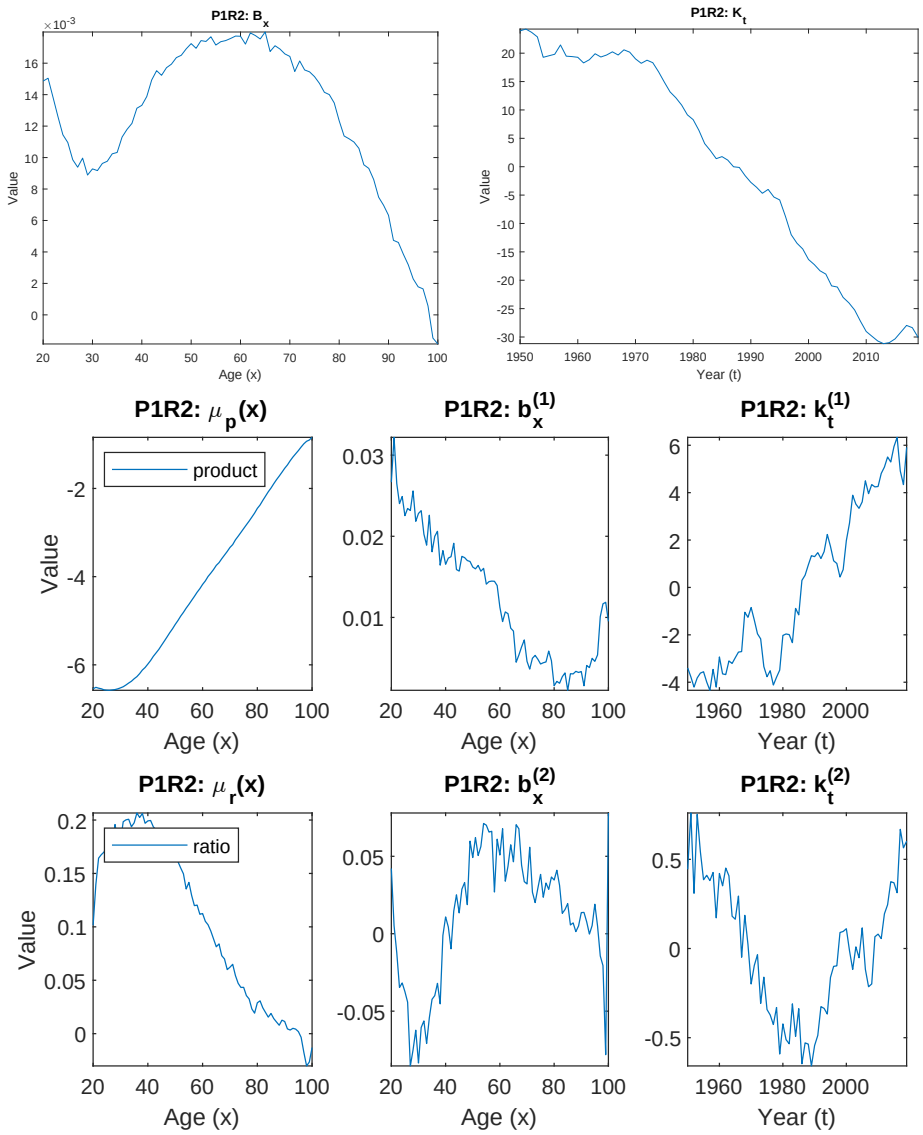

Figure 2. Parameter estimates of the product–ratio model fitted to the U.S. and Canadian males with ages from 20 to 100 and years from 1950 to 2019.

3.3. The Time-Series Process

Once we obtain the estimates of K_t, $k_t^{(1)}$, and $k_t^{(2)}$, we will fit these sequences to the VAR time series models as described in Section 2. In this subsection, we first explore the sample cross-correlation matrix (SCCM) of the time series sequences. We then present the model selection result for the three time series structures described in Section 2. Finally, we analyze the term structure of the death probability under different time series structures.

3.3.1. Sample Cross-Correlation Matrix

The sample cross-correlation matrix (SCCM) is typically used to measure the strength of linear dependence between time series. Considering the multivariate sequence $(\Delta K_t, k_t^{(1)}, k_t^{(2)})$ obtained in ACF model and PR model, we present the values of SCCM in Table 1. For the reader's convenience, we also include the simplified SCCM in this table.

Table 1. Sample cross-correlation matrix of the estimated sequence of $(\Delta K_t, k_t^{(1)}, k_t^{(2)})$ under the ACF model and the PR model. The symbol "." means insignificance (at $\alpha = 1\%$ significance level), "+" means positive significance.

ACF Model						**P/R Model**					
$\boldsymbol{\Delta K_t}$	$\boldsymbol{k_t^{(1)}}$	$\boldsymbol{k_t^{(2)}}$	ΔK_t	$k_t^{(1)}$	$k_t^{(2)}$	$\boldsymbol{\Delta K_t}$	$\boldsymbol{k_t^{(1)}}$	$\boldsymbol{k_t^{(2)}}$	$\boldsymbol{\Delta K_t}$	$\boldsymbol{k_t^{(1)}}$	$\boldsymbol{k_t^{(2)}}$
Lag 0											
1.0000	0.1797	0.2752	+	.	.	1.0000	–0.0405	0.3563	+	.	.
0.1797	1.0000	0.9426	.	+	+	–0.0405	1.0000	0.0180	.	+	.
0.2752	0.9426	1.0000	.	+	+	0.3563	0.0180	1.0000	.	.	+
Lag 1											
0.0668	0.2995	0.3395	.	.	.	0.2952	–0.0118	0.2574	.	.	.
0.1238	0.9612	0.8934	.	+	+	–0.0062	0.9484	–0.0333	.	+	.
0.2453	0.9377	0.9688	.	+	+	0.3023	0.0684	0.8464	.	.	+
Lag 2											
0.3110	0.3811	0.3538	.	.	.	0.2163	–0.0193	0.3509	.	.	.
0.0895	0.8990	0.8371	.	+	+	–0.0487	0.9092	–0.0683	.	+	.
0.2414	0.9178	0.9348	.	+	+	0.2803	0.1131	0.8269	.	.	+
Lag 3											
0.1718	0.3854	0.3461	.	.	.	0.0640	–0.0482	0.3242	.	.	.
0.0817	0.8311	0.7815	.	+	+	–0.0924	0.8635	–0.1339	.	+	.
0.2139	0.8959	0.9005	.	+	+	0.2535	0.1753	0.7405	.	.	+
Lag 4											
0.0070	0.4049	0.3551	.	+	.	0.1223	–0.0878	0.3519	.	.	.
0.0786	0.7685	0.7310	.	+	+	–0.1899	0.8075	–0.1704	.	+	.
0.1759	0.8736	0.8643	.	+	+	0.1557	0.2055	0.6686	.	.	+
Lag 5											
0.1093	0.3565	0.2957	.	.	.	0.0949	–0.0788	0.3239	.	.	.
0.0840	0.7078	0.6851	.	+	+	–0.2548	0.7553	–0.2155	.	+	.
0.1583	0.8397	0.8243	.	+	+	0.1600	0.2280	0.6172	.	.	+

By comparing the results between the ACF model and the PR model, the following observations can be made. First, for the ACF model, some off-diagonal components are statistically significant. This finding highlights the existence of cross-correlations between different latent states in the ACF model.

On the other hand, for the PR model, all the off-diagonal components from lag 0 to lag 5 are insignificant. These insignificant values serve as the advantageous outcome of the PR model, since it is more capable of removing the cross-correlation among different

latent states. With this advantage, a time series model with simpler structure could be used under the PR model.

The differing cross-correlations observed in the two models primarily stem from the different response variables they use. In the ACF model, the response variables are set to be the log mortality rates from sub-populations, which are highly correlated. On the other hand, the PR model operates differently: its product model uses the square root of the geometric mean to capture the joint movement of multiple populations, while its ratio model employs the square root of the ratio between two populations to capture their divergence. Hyndman et al. (2013) demonstrated that these response variables in the PR model are largely uncorrelated, particularly when the variances among the sub-populations are relatively similar.

A final observation lies in the significant SCCM in the diagonal components of $k_t^{(1)}, k_t^{(2)}$ from lag 1 to 5. It suggests that $k_t^{(1)}$, $k_t^{(2)}$ are serially correlated with their previous values, and, as such an auto-regressive structure is needed for each of $k_t^{(1)}$ and $k_t^{(2)}$.

3.3.2. AICs and Likelihood Ratio Test

In this section, we compare the fitting performance of the time series models within each of the ACF modeling framework and the PR modeling framework. Since the two sets of model variants, (ACF0, ACF1, and ACF2) and (PR0, PR1, and PR2), are both nested models, we use the Akaike information criterion (AIC) to compare the fitting of different model variants. The value of AIC is computed as follow:

$$\text{AIC} = -2(\mathcal{L}) + 2(\mathcal{N}),$$

where $\mathcal{L}$ is the maximized log-likelihood and $\mathcal{N}$ is the number of unknown parameters in the model. The term $2(\mathcal{N})$ can be viewed as a penalty term penalizing for the complexity of the model. Therefore, AIC can be viewed as a trade-off between the fitting performance and the model complexity. According to its definition, a model with a smaller value of AIC is preferred. In the following, we use restricted model to refer to a model with simpler model structure, and unrestricted model to a model with more complex structure within the nested models.

Table 2 shows the values of AIC for different time series models within the ACF framework (left panel) and within the PR framework (right panel). Let us first consider the results within the ACF framework. By comparing the pairs of "ACF0 vs. ACF1", "ACF0 vs. ACF2", "ACF1 vs. ACF2", we can find that a more complicated model is always resulting in a greater value of $\mathcal{L}$, the maximized log-likelihood. This result is expected since these are all nested models where the restricted model is always a special case of an unrestricted model. More importantly, if we compare the AIC values, the unrestricted models are always preferred since they have smaller AIC values. It suggests that, under the ACF model framework, the restricted model does not provide adequate fit to the data, partly due to the significant cross-correlation among different latent states as presented in previous section.

Next we focus on the results for the PR framework. The same observation can be made regarding the value of $\mathcal{L}$ when we compare the three pairs of "PR0 vs. PR1", "PR0 vs. PR2", "PR1 vs. PR2". However, regarding the value of AIC, a conclusion opposite to that for the ACF framework will be drawn, since the restricted models are always having smaller AIC values. It matches the result represented in previous section where the cross-correlation coefficients are insignificant for the PR framework. It can be viewed as the main advantage of the PR model under which by construction separates the co-movement from the deviation of the two populations, and therefore disentangling trend risk and basis risk. Finally, it should be noted that the results between the two modeling framework are not directly comparable, since they are based on different estimated values of $(K_t, k_t^{(1)}, k_t^{(2)})$.

Table 2. The log-likelihood and AIC for different model variants.

	ACF0	**ACF1**	**ACF2**	**PR0**	**PR1**	**PR2**
$\mathcal{N}$	8	11	18	8	11	18
$\mathcal{L}$	–338.5748	–297.5167	–285.9478	–144.2575	–141.3853	–135.7845
AIC	693.1496	617.0334	607.8957	304.5149	304.7706	307.5689

Besides the comparison of the AIC values, we have also considered likelihood ratio (LR) tests to compare the goodness-of-fit of these nested models. The null hypothesis and alternative hypothesis of a LR test is summarized as follow:

$$\begin{aligned} H_0: &\quad \theta = \theta_0, \quad \text{representing the restricted model} \\ H_1: &\quad \theta = \theta_1, \quad \text{representing the unrestricted model} \end{aligned}$$

The result of the LR test is summarized in Table 3. If p-value is less than 5%, it means the null is rejected and the unrestricted model is preferred, and vice versa. The same conclusion can be drawn; that is, among the ACF framework, the unrestricted model is preferred, due to the fact that the restricted model fails to adequately capture the cross-correlation among different latent states. On the other hand, among the PR model, the restricted model is preferred since the PR model by construction has separated common-factor from the population-specific deviation.

Table 3. Likelihood ratio test result for the three variants of the ACF model and the PR model.

$H_0: \theta = \theta_0$ **Restricted Model**	$H_1: \theta = \theta_1$ **Unrestricted Model**	p**-Value**
ACF0	ACF1	0.0000
ACF0	ACF2	0.0000
ACF1	ACF2	0.0016
PR0	PR1	0.1247
PR0	PR2	0.0756
PR1	PR2	0.1301

3.3.3. Term Structure of Correlation

Finally, we present the term structure of correlation between $q^{(1)}_{x,t}$ and $q^{(2)}_{x,t}$, the death probabilities calculated for the two populations. Following Cairns et al. (2019), the term structure of correlation is defined to be $Corr(q^{(1)}_{x,t}, q^{(2)}_{x,t})$, for $t = 1, \ldots, T$. For illustration, we select $x = 70$, and provide the term structure of correlation between $q^{(1)}_{70,t}$ and $q^{(2)}_{70,t}$ over the period of 2020 to 2069 (that is, a 50-year forecast window) in Figure 3. We assume constant force of mortality, so $q_{x,t}$ and $m_{x,t}$ is linked through the following equation:

$$m_{x,t} = -\log(1 - q_{x,t}).$$

The key findings are summarized into four aspects.

- ACF0 vs. ACF1
  Let us first focus on ACF0 model and ACF1 model. For these models, the only difference is the specification of Q matrix. To be more specific, In ACF0, the off-diagonal elements of Q matrix are zeros, while in ACF1, those elements are all non-zeros. The term structures of correlation under ACF0 and ACF1 have similar shape. However, a 10% increase is observed when the model does incorporate the off-diagonal covariance elements.
- ACF2 vs. ACF1 or ACF0
  The mortality indices in the ACF2 model follows a VAR(1) model where additional parameters are used to capture the cross-correlation among different period effects.

Comparing to ACF1 and ACF0, a dramatic change of shape (the black dotted line) is observed for the ACF2, reflecting those additional interactions.

- PR0 vs. PR1 vs. PR2
  Within the PR framework, the resulting term structures of correlation from the three models carry two features. The first feature is related to the magnitudes of the lines, with the values being greater for more complex models. The second feature is related to the shape of the lines, where all three lines yield very similar patterns. This finding is consistent with the observations we made in Sections 3.3.1 and 3.3.2. It also justifies the argument made by Hyndman et al. (2013) that the latent factors are uncorrelated in the PR model.
- PR vs. ACF
  Since the two models have different model specification, the shape of the curve may not be the same. For the PR model, the three variants yield similar term structure of correlations, while for the ACF model, the ACF0 variant and ACF1 variant have omitted the cross-correlation between different latent states, and thereby not adequately fitting the data and leading to the worst term structure of correlations. When the full VAR structure is used in the ACF2 variant, its term structure of correlation would be similar to those in the PR framework.

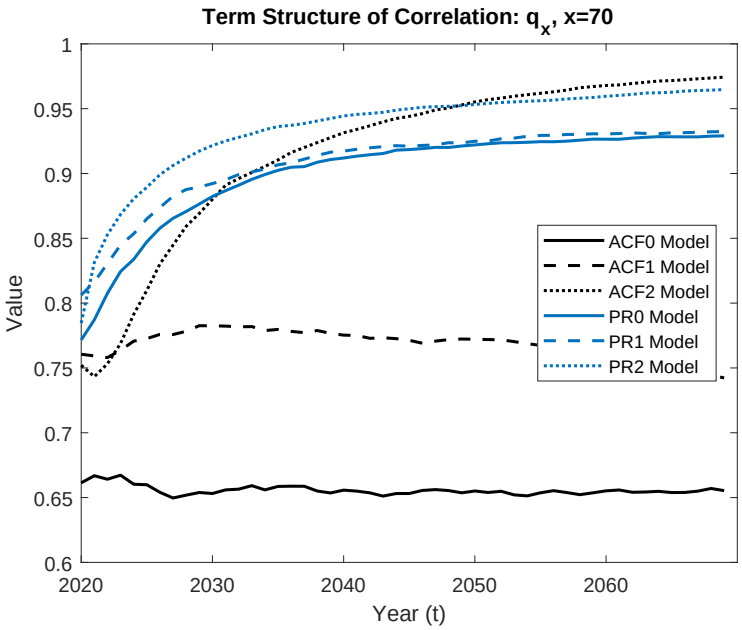

Figure 3. The term structure of correlation between $q^{(1)}_{x,t}$ and $q^{(2)}_{x,t}$, $x = 65$.

4. Hedge Performance

In this section, we apply the ACF models and the PR models to longevity hedges and study their hedge performances.

4.1. Basic Set Up

We first describe the set up of the hedge. Suppose it is now time t_0 (the end of year t_0). We assume the hedger has a longevity liability that is a life annuity that pays USD 1at the end of each year, provided that the annuitant is alive. Let $L(x,t)$ denotes the value of a life annuity sold to individuals aged x at time t. We consider a deferred life annuity that sold to age x_0 in T_d years. The liability being hedged is $L(x_0, t_0 + T_d)$, which can be expressed as

$$L(x_0, t_0 + T_d) = \sum_{s=1}^{w-x_0} e^{-r \cdot s} \prod_{k=1}^{s} p^{(P_L)}_{x_0+k-1, t_0+T_d+k},$$

where w is the maximum age, r is the interest rate and P_L is the underlying population of the liability.

The hedging instrument is a q-forward, which is a zero-coupon swap with the fix leg attached to the forward mortality rate determined at the beginning of the contract, and the floating leg being the realized mortality rate at maturity. We use $H(x,t)$ to represent the value of a q-forward which is linked to reference age x and maturity time t. For simplicity, we assume only one q-forward is used in the hedge. The q-forward being used is denoted by $H(x_h, t_0+T_h)$, which has a reference age x_h and maturity T_h years from now. The value of $H(x_h, t_0+T_h)$ can be expressed as

$$H(x_h, t_0+T_h) = e^{-r\cdot T_h}\left(\hat{q}^{(P_H)}_{x_h,t_0+T_h} - q^{(P_H)}_{x_h,t_0+T_h}\right)$$

We assume the hedger's goal is to minimize volatility of their longevity exposure over the long run. The hedging target is set to be

$$\min_N(\mathrm{Var}(V_L - N\times V_H)),$$

where N is the notional amount or the hedge ratio of the portfolio, V_L and V_H are the values of the liability and the q-forward at time (t_0+T_d). We obtain both V_L and V_H through simulation with detailed procedure provided in Appendix B. The hedge ratio N is obtained by matching the sensitivity of V_L and V_H with respect to $K_{t_0+T_d}$, i.e., $N = \Delta^c_L/\Delta^c_H$. The procedure for calculating Δ^c_L and Δ^c_H is provided in Appendix A. The effectiveness of the hedge is measured by HE, which is computed via the following equation

$$HE = 1 - \frac{\mathrm{Var}(V_L - N\times V_H)}{\mathrm{Var}(V_L)}.$$

4.2. Illustration of Hedge Performance

To illustrate the hedge result, we consider the following set up for the liability and the q-forward.

- The liability is a life annuity sold to individuals aged 70 ($x_0 = 70$) 20 years from now ($T_d = 20$). The liability is linked to population 1, the U.S. males. The maximum age is set to be $w = 100$.
- The q-forward has a reference age of 65 ($x_h = 70$) and will matures 20 years from now ($T_h = 20$). The q-forward is linked to population 2, the Canadian males.
- The interest rate is assumed to be 1% per annum.

To assess the impact of different VAR structures, we allow the simulation model (that is, the model used to generate the future mortality sample paths) to be different from the model used for hedge calibration. The values of hedge effectiveness and the associated notional amounts of the hedging instrument are provided in Tables 4 and 5, respectively. In these two tables, each column corresponds to a particular simulation model and each row corresponds to a particular calibration model. For example, for the column of ACF0 and the row of ACF1, we use ACF0 to simulate V_L and V_H. We use ACF1 to derive the hedge ratio N. The values in the cell are the corresponding HE and N.

For the ACF framework, the hedge effectiveness is highly dependent on the model used for simulating the underlying mortality experience. The resulting HE ranges from 47.1669% to 77.8786%, with a maximum difference being as large as 30%. This reveals the model risk embedded in the ACF framework. When the time series model is not specified properly, the outcome of the hedge would also suffer. However, in practice, the hedger is always subject to model uncertainty, since the true underlying mortality model is unknown.

The PR framework, on the other hand, yields more robust result in hedging. Considering different combinations of simulation model and calibration model, very little change is observed in the resulting HE, where its value ranges from 68.1193% and 73.0943% with maximum difference of 4.9750% which is way smaller than that in the ACF framework.

Table 4. The value of hedging effectiveness (HE) in the longevity hedge under the ACF framework (left panel) and the PR framework (right panel), using different combinations of simulation model and hedge calibration model.

		Simulation		
Calibration	Model	ACF0	ACF1	ACF2
	ACF0	52.6756%	61.7846%	77.8786%
	ACF1	52.6756%	61.7846%	77.8786%
	ACF2	47.1669%	53.5626%	70.2547%
		Max diff. = 30.7117%		

		Simulation		
Calibration	Model	PR0	PR1	PR2
	PR0	68.7645%	68.2410%	73.0943%
	PR1	68.7645%	68.2410%	73.0943%
	PR2	68.6478%	68.1193%	72.9632%
		Max diff. = 4.9750%		

Table 5. The notional amounts of q-forward in the longevity hedge under the ACF framework (left panel) and the PR framework (right panel), using different combinations of simulation model and hedge calibration model.

		Simulation		
Calibration	Model	ACF0	ACF1	ACF2
	ACF0	182.9765	182.9765	182.9765
	ACF1	182.9765	182.9765	182.9765
	ACF2	141.4138	141.4138	141.4138

		Simulation		
Calibration	Model	PR0	PR1	PR2
	PR0	146.9599	146.9599	146.9599
	PR1	146.9599	146.9599	146.9599
	PR2	146.5070	146.5070	146.5070

5. Conclusions

In this paper, we apply the PR model developed by Hyndman et al. (2013) to the U.S. male population and Canadian male population, with a goal of studying its performance in disentangling mortality trend risk and population basis risk. Under the PR model, the trend risk and basis risk are separated by construction when the product model and the ratio model are specified. Based on the U.S. and Canada's data, we found that the cross-correlations of different latent states under the PR model are insignificant and the resulting time series structure is more parsimonious. This finding is in great contrast to the result obtained under the ACF framework, where cross-correlations are presented and the associated time series structure is heavily parameterized. We then apply both models to a longevity hedge. The hedge performance as well as the notional amounts of the hedging instruments under the PR framework are more robust.

In this paper, we only consider two populations. For future research, we may extend the PR model to more than two populations. The extension is not trivial. Based on the current structure, the ratio model which focus on capturing deviations among different populations need to be extended to accommodate any pair-wise combinations. If we have a large population size, say 10 populations, the equations required in the ratio model would be increase significantly to $C_{10}^{9} = 45$, which is way more than the number of equations required in the traditional ACF model. As a final item of note, when applying the PR model, we should always consider populations similar mortality experience. As pointed out by Hyndman et al. (2013), the geometric mean and the ratio of the two populations are roughly uncorrelated when they share a similar mortality experience, which advocates the use of the PR model in this case. Extra caution should be taken when dealing with populations that have significantly different mortality experiences.

Author Contributions: Conceptualization, J.S.-H.L.; Methodology, Y.L. and J.S.-H.L.; Validation, Y.L.; Formal analysis, Y.L. and J.S.-H.L.; Investigation, Y.L.; Writing—original draft, Y.L. and J.S.-H.L.; Writing—review & editing, J.S.-H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: https://www.mortality.org accessed on 1 September 2023.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Partial Derivatives of Liability and Hedging Instrument

Appendix A.1. The ACF Model:

Liability:

$$\begin{aligned} V_L &= e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-rs}{}_sP_{x_0}^{(P_L)} \\ &= e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-rs}\prod_{u=1}^{s} p_{x_0+u-1,t_0+T_d+u}^{(P_L)} \\ &= e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-rs}\prod_{u=1}^{s} e^{-m_{x_0+u-1,t_0+T_d+u}^{(P_L)}} \\ &= e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-rs}e^{-\sum_{u=1}^{s} m_{x_0+u-1,t_0+T_d+u}^{(P_L)}} \\ &= e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-rs}\exp(-\sum_{u=1}^{s}\exp(a_{x_0+u-1}+B_{x_0+u-1}\cdot K_{t_0+T_d+u}+b_{x_0+u-1}^{(P_L)}\cdot k_{t_0+T_d+u}^{(P_L)})) \end{aligned}$$

At time t_0+T_d, we project the life table and we have

$$V_L = e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-rs}\exp(-\sum_{u=1}^{s}\exp(a_{x_0+u-1}+B_{x_0+u-1}\cdot K_{t_0+T_d}+b_{x_0+u-1}^{(P_L)}\cdot k_{t_0+T_d}^{(P_L)})).$$

The partial derivatives w.r.t. $K_{t_0+T_d}$ and $k_{t_0+T_d}^{(P_L)}$ can then be computed as follows:

$$\Delta_L^c = \frac{\partial V_L}{\partial K_{t_0+T_d}} = e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-r\cdot s}\frac{\partial A}{\partial K_{t_0+T_d}}$$

$$\Delta_L^{(P_L)} = \frac{\partial V_L}{\partial k_{t_0+T_d}^{(P_L)}} = e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-r\cdot s}\frac{\partial A}{\partial k_{t_0+T_d}^{(P_L)}}$$

where

$$A = \exp(-\sum_{u=1}^{s}\exp(a_{x_0+u-1}+B_{x_0+u-1}(K_{t_0+T_d})+b_{x_0+u-1}^{(P_L)}\cdot(k_{t_0+T_d}^{(P_L)})))$$

$$\frac{\partial A}{\partial K_{t_0+T_d}} = A[-\sum_{u=1}^{s} B_{x_0+u-1}\exp(a_{x_0+u-1}+B_{x_0+u-1}(K_{t_0+T_d})+b_{x_0+u-1}^{(P_L)}\cdot(k_{t_0+T_d}^{(P_L)}))]$$

and

$$\frac{\partial A}{\partial k_{t_0+T_d}^{(P_L)}} = A[-\sum_{u=1}^{s} b_{x_0+u-1}^{(P_L)}\exp(a_{x_0+u-1}+B_{x_0+u-1}(K_{t_0+T_d})+b_{x_0+u-1}^{(P_L)}\cdot(k_{t_0+T_d}^{(P_L)}))]$$

Q-forward:

$$\begin{aligned} V_H &= e^{-rT_h}(\hat{q}_{x_h,t_0+T_h}-q_{x_h,t_0+T_h}) \\ &= -e^{-rT_h}[1-e^{-m_{x_h,t_0+T_h}^{(P_H)}}-1+e^{-\hat{m}_{x_h,t_0+T_h}^{(P_H)}}] \\ &= -e^{-rT_h}[e^{-\hat{m}_{x_h,t_0+T_h}^{(P_H)}}-e^{-\exp(a_{x_h}^{(P_H)}+B_{x_h}K_{t_0+T_h}+b_{x_h}^{(P_H)}k_{t_0+T_h}^{(P_H)})}]. \end{aligned}$$

The partial derivatives of H_j w.r.t. $K_{t_0+T_h}$ and $k_{t_0+T_h}^{(P_H)}$ can be computed as follows:

$$\Delta_H^c = \frac{\partial H}{\partial K_{t_0+T_h}} = -e^{rT_h}\cdot e^{-\exp(a_{x_h}^{(P_H)}+B_{x_h}K_{t_0+T_h}+b_{x_h}^{(P_H)}k_{t_0+T_h}^{(P_H)})}\cdot B_{x_h}\exp(a_{x_h}^{(P_H)}+B_{x_h}K_{t_0+T_h}+b_{x_h}^{(P_H)}k_{t_0+T_h}^{(P_H)})$$

$$\Delta_H^{(P_H)} = \frac{\partial H}{\partial k_{t_0+T_h}^{(P_H)}} = -e^{rT_h}\cdot e^{-\exp(a_{x_h}^{(P_H)}+B_{x_h}K_{t_0+T_h}+b_{x_h}^{(P_H)}k_{t_0+T_h}^{(P_H)})}\cdot B_{x_h}\exp(a_{x_h}^{(P_H)}+b_{x_h}^{(P_H)}K_{t_0+T_h}+b_{x_h}^{(P_H)}k_{t_0+T_h}^{(P_H)})$$

Appendix A.2. The PR Model:

Liability:

$$\begin{aligned} V_L &= e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-rs}{}_sP_{x_0}^{(P_L)} \\ &= e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-rs}\prod_{u=1}^{s} p_{x_0+u-1,t_0+T_d+u}^{(P_L)} \\ &= e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-rs}\prod_{u=1}^{s} e^{-m_{x_0+u-1,t_0+T_d+u}^{(P_L)}} \\ &= e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-rs}e^{-\sum_{u=1}^{s} m_{x_0+u-1,t_0+T_d+u}^{(P_L)}} \end{aligned}$$

Since that in the PR model, we have

$$\log(m_{x,t}^{(1)}) = (\mu_x^p + \mu_x^r) + B_xK_t + b_{1,x}k_{1,t} + b_{2,x}k_{2,t} + e_{x,t} + w_{x,t},$$

$$\log(m_{x,t}^{(2)}) = (\mu_x^p - \mu_x^r) + B_xK_t - b_{1,x}k_{1,t} - b_{2,x}k_{2,t} + e_{x,t} - w_{x,t}.$$

Therefore, if P_L is selected to be population 1, the value of V_L can be computed as

$$\begin{aligned} V_L &= e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-rs}\exp(-\sum_{u=1}^{s}\exp((\mu_{x_0+u-1}^p + \mu_{x_0+u-1}^r)+ \\ &\quad B_{x_0+u-1}\cdot K_{t_0+T_d} + b_{1,x_0+u-1}\cdot k_{1,t_0+T_d} + b_{2,x_0+u-1}\cdot k_{2,t_0+T_d})). \end{aligned}$$

Similarly, if P_L is selected to be population 2, we have

$$\begin{aligned} V_L &= e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-rs}\exp(-\sum_{u=1}^{s}\exp((\mu_{x_0+u-1}^p - \mu_{x_0+u-1}^r)+ \\ &\quad B_{x_0+u-1}\cdot K_{t_0+T_d} - b_{1,x_0+u-1}\cdot k_{1,t_0+T_d} - b_{2,x_0+u-1}\cdot k_{2,t_0+T_d})). \end{aligned}$$

In the following derivation, we assume P_L is population 1 and P_H is population 2. It can be easily modified to adapt different situations.

At time $t_0 + T_d$, we project the life table. This is denoted by

$$\tilde{m}_{x,t_0+T_d}^{(P_L)} = \exp\Big((\mu_x^p + \mu_x^r) + B_x\cdot K_{t_0+T_d} + b_{1,x}\cdot k_{1,t_0+T_d} + b_{2,x}\cdot k_{2,t_0+T_d}\Big).$$

We have

$$V_L = e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-rs}\exp\Big(-\sum_{u=1}^{s}\tilde{m}_{x_0+u-1,t_0+T_d}^{(P_L)}\Big).$$

The partial derivatives with respect to $K_{t_0+T_d}$, k_{1,t_0+T_d} and k_{2,t_0+T_d} can then be computed as follows:

$$\Delta_L^c = \frac{\partial V_L}{\partial K_{t_0+T_d}} = e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-r\cdot s}\frac{\partial A}{\partial K_{t_0+T_d}}$$

$$\Delta_{1,L} = \frac{\partial V_L}{\partial k_{1,t_0+T_d}} = e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-r\cdot s}\frac{\partial A}{\partial k_{1,t_0+T_d}}$$

$$\Delta_{2,L} = \frac{\partial V_L}{\partial k_{2,t_0+T_d}} = e^{-r\cdot T_d}\sum_{s=1}^{w-x_0} e^{-r\cdot s}\frac{\partial A}{\partial k_{2,t_0+T_d}}$$

where

$$\begin{aligned} \frac{\partial A}{\partial K_{t_0+T_d}} &= A\Big(-\sum_{u=1}^{s} B_{x_0+u-1}\cdot\tilde{m}_{x_0+u-1,t_0+T_d}^{(P_L)}\Big) \\ \frac{\partial A}{\partial k_{1,t_0+T_d}} &= A\Big(-\sum_{u=1}^{s} b_{1,x_0+u-1}\cdot\tilde{m}_{x_0+u-1,t_0+T_d}^{(P_L)}\Big) \\ \frac{\partial A}{\partial k_{2,t_0+T_d}} &= A\Big(-\sum_{u=1}^{s} b_{2,x_0+u-1}\cdot\tilde{m}_{x_0+u-1,t_0+T_d}^{(P_L)}\Big) \end{aligned}$$

and

$$A = \exp\Big(-\sum_{u=1}^{s}\tilde{m}_{x_0+u-1,t_0+T_d}^{(P_L)}\Big)$$

Q-forward:

$$\begin{aligned} V_H &= e^{-rT_h}(\hat{q}_{x_h,t_0+T_h} - q_{x_h,t_0+T_h}) \\ &= -e^{-rT_h}\left(1 - e^{-m^{(P_H)}_{x_h,t_0+T_h}} - 1 + e^{-\hat{m}^{(P_H)}_{x_h,t_0+T_h}}\right) \\ &= -e^{-rT_h}\left(e^{-\hat{m}^{(P_H)}_{x_h,t_0+T_h}} - e^{-\tilde{m}^{(P_H)}_{x_h,t_0+T_h}}\right). \end{aligned}$$

As previously mentioned, we assume P_H to be population 2. Therefore, $\tilde{m}^{(P_H)}_{x,t_0+T_h}$ can be computed as

$$\tilde{m}^{(P_H)}_{x,t_0+T_h} = \exp\Big((\mu^p_x - \mu^r_x) + B_x \cdot K_{t_0+T_h} - b_{1,x} \cdot k_{1,t_0+T_h} - b_{2,x} \cdot k_{2,t_0+T_h}\Big).$$

The partial derivatives of H_j with respect to $K_{t_0+T_h}$ and $k^{(P_H)}_{t_0+T_h}$ can be computed as follows:

$$\begin{aligned} \Delta^c_H &= \tfrac{\partial H}{\partial K_{t_0+T_h}} = (-1)e^{rT_h} \cdot B_{x_h} \cdot \tilde{m}^{(P_H)}_{x_h,t_0+T_h} \cdot e^{-\tilde{m}^{(P_H)}_{x_h,t_0+T_h}} \\ \Delta_{1,H} &= \tfrac{\partial H}{\partial k_{1,t_0+T_h}} = e^{rT_h} \cdot b_{1,x_h} \cdot \tilde{m}^{(P_H)}_{x_h,t_0+T_h} \cdot e^{-\tilde{m}^{(P_H)}_{x_h,t_0+T_h}} \\ \Delta_{2,H} &= \tfrac{\partial H}{\partial k_{2,t_0+T_h}} = e^{rT_h} \cdot b_{2,x_h} \cdot \tilde{m}^{(P_H)}_{x_h,t_0+T_h} \cdot e^{-\tilde{m}^{(P_H)}_{x_h,t_0+T_h}} \end{aligned}$$

Appendix B. Simulation Procedure

The simulation procedure of the value of the liability being hedged is summarized as follows:

- Simulate M sample paths of K_{t_0+T}, k_{1,t_0+T} and k_{2,t_0+T}. We denote these simulated paths as ω_j, for $j = 1, 2, \ldots, M$. Accordingly, for each simulated path ω_j, the values of the corresponding state variables are $K_{t_0+T}(\omega_j), k_{1,t_0+T}(\omega_j)$ and $k_{2,t_0+T}(\omega_j)$.
- For each simulated path ω_j, $j = 1, 2, \ldots, M$, we project the life table at time $t_0 + T$, based on the simulated $K_{t_0+T}(\omega_j), k_{1,t_0+T}(\omega_j)$ and $k_{2,t_0+T}(\omega_j)$. The methodology of how we project the life table is summarized as follows:
  1. Based on $K_{t_0+T}(\omega_j), k_{1,t_0+T}(\omega_j)$ and $k_{2,t_0+T}(\omega_j)$, we first compute the central death rate of an individual aged x, $m_{x,t_0+T}(\omega_j)$:
     - ACF Model

$$\begin{aligned} \log(m^{(1)}_{x,t_0+T}(\omega_j))) &= a^{(1)}_x + B_x K_{t_0+T}(\omega_j)) + b^{(1)}_x k^{(1)}_{t_0+T}(\omega_j)) \\ \log(m^{(2)}_{x,t_0+T}(\omega_j))) &= a^{(2)}_x + B_x K_{t_0+T}(\omega_j)) + b^{(2)}_x k^{(2)}_{t_0+T}(\omega_j)) \end{aligned}$$

     - PR Model

$$\begin{aligned} \log(m^{(1)}_{x,t_0+T}(\omega_j)) &= (\mu^p_x + \mu^r_x) + B_x K_{t_0+T}(\omega_j) + b_{1,x} k_{1,t_0+T}(\omega_j) + b_{2,x} k_{2,t_0+T}(\omega_j), \\ \log(m^{(2)}_{x,t_0+T}(\omega_j)) &= (\mu^p_x - \mu^r_x) + B_x K_{t_0+T}(\omega_j) - b_{1,x} k_{1,t_0+T}(\omega_j) - b_{2,x} k_{2,t_0+T}(\omega_j). \end{aligned}$$

  2. Then we compute the death probability $q_{x,t_0+T}(\omega_j)$, based on the value of $m_{x,t_0+T}(\omega_j)$. Throughout this study, we assume constant force of mortality. Therefore, $q_{x,t_0+T}(\omega_j)$ can be computed through the following equation:

$$q_{x,t_0+T}(\omega_j) = 1 - \exp(-m_{x,t_0+T}(\omega_j)),$$

     and accordingly

$$p_{x,t_0+T}(\omega_j) = 1 - q_{x,t_0+T}(\omega_j).$$

3. We use ${}_sP_{x,t_0+T}(\omega_j)$ to denote the probability that an individual aged x at time t_0+T survives to age $x+s$, for a particular path ω_j. The survival probability is computed as

$$ {}_sP_{x,t_0+T}(\omega_j) = \prod_{k=1}^{s} p_{x+k-1,t_0+T}(\omega_j). $$

- For each projected life table, we calculate the value of the life annuity $V_L(\omega_j)$, which is

$$ V_L(\omega_j) = \sum_{s=1}^{w-x_0} e^{-r\cdot s} {}_sP^{(P_L)}_{x_0,t_0+T_d}(\omega_j), $$

Similarly, the value of the q-forward V_H is

$$ V_H(\omega_j) = e^{-r\cdot T_h}\left(\hat{q}^{(P_H)}_{x_h,t_0+T_h} - q^{(P_H)}_{x_h,t_0+T_h}(\omega_j)\right), $$

where

$$ q^{(P_H)}_{x_h,t_0+T_h}(\omega_j) = 1 - {}_1P^{(P_H)}_{x_h,t_0+T_h}(\omega_j), $$

and $\hat{q}^{(P_H)}_{x_h,t_0+T_h}$ is the pre-defined forward mortality rate at time t_0.

Notes

[1] Data source: Human Mortality Database (HMD).

[2] For an expanded data set, the same MLE method can be used.

References

Cairns, Andrew J. G., David Blake, and Kevin Dowd. 2006. A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *Journal of Risk and Insurance* 73: 687–718. [CrossRef]

Cairns, Andrew J. G., David Blake, Kevin Dowd, Guy D. Coughlan, and Marwa Khalaf-Allah. 2011. Bayesian stochastic mortality modelling for two populations. *ASTIN Bulletin: The Journal of the IAA* 41: 29–59.

Cairns, Andrew J. G., David Blake, Kevin Dowd, Guy D. Coughlan, David Epstein, Alen Ong, and Igor Balevich. 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. [CrossRef]

Cairns, Andrew J. G., Malene Kallestrup Lamb, Carsten Rosenskjold, David P. Blake, and Kevin Dowd. 2019. *Modelling Socio-Economic Differences in the Mortality of Danish Males Using a New Affluence Index*. London: Pensions Institute, Cass Business School, City University London.

Chen, Hua, Richard MacMinn, and Tao Sun. 2015. Multi-population mortality models: A factor copula approach. *Insurance: Mathematics and Economics* 63: 135–46. [CrossRef]

Dong, Yumo, Fei Huang, Honglin Yu, and Steven Haberman. 2020. Multi-population mortality forecasting using tensor decomposition. *Scandinavian Actuarial Journal* 8: 754–75. [CrossRef]

Dowd, Kevin, Andrew J. G. Cairns, David Blake, Guy D. Coughlan, and Marwa Khalaf-Allah. 2011. A gravity model of mortality rates for two related populations. *North American Actuarial Journal* 15: 334–56. [CrossRef]

Graziani, George. 2014. Longevity risk—A fine balance. *Journal of Retirement* 1: 35–37.

Hunt, Andrew, and David Blake. 2020. Identifiability in age/period mortality models. *Annals of Actuarial Science* 14: 461–99. [CrossRef]

Hyndman, Rob. 2023. *Demography: Forecasting Mortality, Fertility, Migration and Population Data*. R Package Version 2.0. Available online: https://CRAN.R-project.org/package=demography (accessed on 30 September 2023).

Hyndman, Rob J., Heather Booth, and Farah Yasmeen. 2013. Coherent mortality forecasting: The product-ratio method with functional time series models. *Demography* 50: 261–83. [CrossRef] [PubMed]

Lee, Ronald D., and Lawrence R. Carter. 1992. Modeling and forecasting u. s. mortality. *Journal of the American Statistical Association* 87: 659–71.

Li, Johnny Siu-Hang, and Yanxin Liu. 2020. The heat wave model for constructing two-dimensional mortality improvement scales with measures of uncertainty. *Insurance: Mathematics and Economics* 93: 1–26. [CrossRef]

Li, Nan, and Ronald Lee. 2005. Coherent mortality forecasts for a group of populations: An extension of the lee-carter method. *Demography* 42: 575–94. [CrossRef] [PubMed]

Lyu, Pintao, Anja De Waegenaere, and Bertrand Melenberg. 2021. A multipopulation approach to forecasting all-cause mortality using cause-of-death mortality data. *North American Actuarial Journal* 25: S421–56. [CrossRef]

Michaelson, Avery, and Jeff Mulholland. 2014. Strategy for increasing the global capacity for longevity risk transfer: Developing transactions that attract capital markets investors. *Journal of Alternative Investment* 17: 18–27. [CrossRef]

Plat, Richard. 2009. On stochastic mortality modeling. *Insurance: Mathematics and Economics* 45: 393–404. [CrossRef]
Renshaw, Arthur E., and Steven Haberman. 2006. A cohort-based extension to the lee-carter model for mortality reduction factors. *Insurance: Mathematics and Economics* 38: 556–70. [CrossRef]
Russolillo, Maria, Giuseppe Giordano, and Steven Haberman. 2011. Extending the lee–carter model: A three-way decomposition. *Scandinavian Actuarial Journal* 2011: 96–117. [CrossRef]
Tsai, Cary Chi-Liang, and Ying Zhang. 2019. A multi-dimensional bühlmann credibility approach to modeling multi-population mortality rates. *Scandinavian Actuarial Journal* 5: 406–31. [CrossRef]
Zhou, Rui, Guangyu Xing, and Min Ji. 2019. Changes of relation in multi-population mortality dependence: An application of threshold vecm. *Risks* 7: 14. [CrossRef]
Zhou, Rui, Yujiao Wang, Kai Kaufhold, Johnny Siu-Hang Li, and Ken Seng Tan. 2014. Modeling period effects in multi-population mortality models: Applications to solvency ii. *North American Actuarial Journal* 18: 150–67. [CrossRef]