KEY Q-DURATION:
A FRAMEWORK FOR HEDGING LONGEVITY RISK

BY

JOHNNY SIU-HANG LI AND ANCHENG LUO

ABSTRACT

When hedging longevity risk with standardized contracts, the hedger needs to calibrate the hedge carefully so that it can effectively reduce the risk. In this article, we present a calibration method that is based on matching mortality rate sensitivities. Specifically, we introduce a measure called key q-duration, which allows us to estimate the price sensitivity of a life-contingent liability to each portion of the underlying mortality curve. Given this measure, one can easily construct a longevity hedge with a small number of q-forward contracts. We further propose an extension for hedging the longevity risk associated with multiple birth cohorts, and another extension for accommodating population basis risk.

KEYWORDS

Cairns-Blake-Dowd model; q-forwards; Securitization.

1. INTRODUCTION

1.1. Background

Pension plans and insurers selling life annuities are subject to longevity risk, the risk that individuals are living longer than expected. Although the risk may not pose as severe a short-term threat as large falls in asset values, it may possibly undermine the long-term sustainability of a portfolio, and thus requires careful management. In recent years, longevity risk has become a high profile risk, partly because of the current low yield environment, and partly because of changes in regulatory regimes. For instance, Solvency II, which is scheduled to come into effect in 2013, requires insurers operating in the European Union to hold longevity risk solvency capital that is based on either a prescribed stress test or an approved internal risk model.

Longevity risk is systematic, so there is a limit to how much longevity risk an entity can take, given its capital base and risk objectives. Pension plans and

Astin Bulletin 42(2), 413-452. doi: 10.2143/AST.42.2.2182804

annuity writers may transfer their longevity risk exposure to capital markets. For example, a pension plan can take a long position of a contract which pays an amount that increases with its realized survival rate to offset the unexpected increase in its liability. Some investors including hedge funds are interested in acquiring an exposure to longevity risk for earning a risk premium, because the risk has no obvious correlation with typical market risk factors such as stock prices, interest rates and foreign exchange rates. We refer readers to Blake et al. (2006) for a comprehensive discussion of how insurers and pension plans can manage their exposure to longevity risk through securitization.

Longevity securities can be divided into two categories: indemnity (bespoke) and standardized. Indemnity contracts are based on the actual mortality experience of the hedger's own portfolio. An example is the survivor swap agreed between Babcock International and Credit Suisse in 2009. Under the terms of the contract, Babcock's pension plan will swap pre-agreed monthly payments to Credit Suisse in return for monthly payments dependent on the longevity of the plan's own members. Indemnity contracts fully mitigate the hedger's longevity risk exposure, but they have limited liquidity. The counterparties may find it difficult to unwind the deal after it had been done.

Standardized contracts, by contrast, are based on the mortality calculated by reference to a national population index. An example is the 25-year longevity bond jointly announced by BNP Paribas and the European Investment Bank in 2004. This bond makes coupon payments that are proportional to the realized survival rates of English and Welsh males who were aged 65 in 2002. A more recent range of standardized contracts include q-forwards and S-forwards, which are documented in Coughlan (2009) and the website of the Life and Longevity Markets Association (LLMA)[1]. Standardized contracts have lower initial and ongoing data requirements. Also, because they are more transparent to investors, they provide quicker execution and, in theory, greater liquidity. However, in relying on standardized contracts, the hedger needs to calibrate the longevity hedge carefully. Specifically, given a certain range of standardized instruments available in the market, the hedger needs to determine the positions of these instruments, so that the portfolio of hedging instruments can result in the desired level of effectiveness.

Although there has been significant research on the pricing of longevity securities (see, e.g., Cairns et al., 2006; Denuit et al., 2007; Dowd et al., 2006; Li et al., 2011; Wills and Sherris, 2010; Zhou and Li, 2010), the formulation of longevity hedging strategies has not been extensively studied. This paper is devoted to latter problem, contributing a practical strategy for hedging longevity risk.

1.2. Previous work on longevity hedging strategies

In a continuous-time setting, Dahl et al. (2008) derived risk-minimizing strategies in markets where survivor swaps are available. Barbarin (2008) performed

[1] www.llma.org

https://doi.org/10.2143/AST.42.2.2182804 Published online by Cambridge University Press

a similar derivation by considering longevity bonds instead. The results in both studies are built upon some assumed stochastic processes for the evolution of mortality, interest rates and net payments to individuals in the portfolio. For example, in the work of Dahl et al. (2008), it was assumed that the underlying mortality process is driven by a Cox-Ingersoll-Ross (CIR) model and that the net payments to all individuals in the portfolio are identical. These assumptions may be too stringent to fit the actual situation that a hedger is facing.

The strategy considered by Cairns et al. (2011b) is more practical. According to this strategy, given a hedging instrument that pays a random amount $H(T)$ at its maturity date T, the hedger should acquire

$$h^* = -\rho_{LH} \frac{\sigma(L(T))}{\sigma(H(T))}$$

units of the hedging instrument, where $L(T)$ is the random liability value at time T, ρ_{LH} is the coefficient of correlation between $H(T)$ and $L(T)$, and $\sigma(L(T))$ and $\sigma(H(T))$ are the standard deviations of $L(T)$ and $H(T)$, respectively. This strategy minimizes the variance of the hedged position at time T. A major limitation of this strategy is that it is not applicable if multiple hedging instruments are considered.

Another practical strategy is the one proposed by Cairns et al. (2008). This strategy is based on a first order approximation to the (spot) survival function. Given the approximation, one can calculate the notional amounts of q-forwards needed to hedge a liability that depends on the survival function. This strategy requires q-forwards that are associated with all ages involved in the liability. However, it is unlikely that such a wide range of contracts are going to be available in the market, even when the market becomes more mature.

The strategy proposed by Cairns (2011) may be considered as an improved version of the previous one. In more detail, survival functions, after a probit transformation, are approximated by Taylor expansions around the stochastic factors in the assumed mortality projection model. The required notional amounts of q-forwards are functions of the parameters in the Taylor expansions. This method permits efficient dynamic hedging, because it can be applied without recalibration to any portfolio of life contingent cashflows and under any given interest rate term structure model. Nevertheless, this method depends heavily on the assumed mortality projection model, and requires the assumed model to possess the Markov property, which means that the method may not be applicable when, for example, Model M8 (Cairns et al., 2009; Dowd et al., 2010), which contains a second order autoregressive process for its cohort effect term, is used. Furthermore, this strategy does not work with liabilities having non-linear payoff structures, such as guaranteed annuity options.

If the hedger is interested in static rather than dynamic hedging, then it can be more convenient to construct a longevity hedge by considering the liability's q-duration, which measures the liability's price sensitivity to changes in the underlying mortality curve. The q-duration measure was first proposed as a concept by Coughlan et al. (2007) and then described in more detail by Coughlan (2009). A longevity hedge can be formed easily by equating the q-durations of the hedge portfolio and the liability being hedged. A shortcoming of this approach, as pointed out by Plat (2010), is that the estimation of q-durations is difficult, because, in any given year, percentage changes in death probabilities for different ages are not identical. The estimation problem has not been analyzed sufficiently in previous research.

1.3. Our contributions

In this paper, we work along the lines of mortality duration measures. The proposed longevity hedging strategy is based on a measure that we coined as key q-duration. Each key q-duration measures the price sensitivity of a life-contingent liability to a specific *portion* of the underlying mortality curve. We investigate the estimation of key q-durations in great depth. The estimation method we propose takes account of the property that changes in mortality rates at different ages are not identical but are correlated. This property, which is sometimes referred to as age dependence, has been found to be important in the securitization of longevity risk (Wills and Sherris, 2008, 2010).

The mortality duration measure we propose is largely analogous to the key rate duration measure, introduced by Ho (1992), which measures a liability's price sensitivity to a specific segment of an interest rate yield curve. Because the method of key rate duration has been used extensively in the industry for many years, practitioners should find the idea presented in this paper easy to accept and implement. Furthermore, as we are going to show, q-forwards are a functional equivalent of zero-coupon bonds, the instruments on which Ho's (1992) interest rate hedging framework is based.

We call the proposed measure key q-duration, because, as we demonstrate empirically, the evolution of a mortality curve over time is driven primarily by a few latent factors, each of which represents a broad age group. By recognizing this property, one can easily create a longevity hedge with a small number of q-forward contracts. Because only a few contracts are required, the method we propose can reduce the cost of hedging, and more importantly, help the longevity market concentrate liquidity on a restricted number of instruments.

We emphasize that the validity of the hedge created by key q-durations does not depend on a specific stochastic mortality model[2]. The calibration work involves no simulation, which means that it is quick and manageable even for a complicated real-life pension plan. The ease of implementation requires little sacrifice of hedge effectiveness. Our illustrations indicate that a longevity hedge created by key q-durations is almost equally effective as the

2 In the absence of basis risk, the calculation of key q-durations does not require a mortality model. However, when basis risk is present, a model may be needed to estimate an adjustment factor for the difference between the mortality of the two populations involved. Further details are provided in Section 6.

https://doi.org/10.2143/AST.42.2.2182804 Published online by Cambridge University Press

corresponding hedge that is optimized by a computationally intensive method. Besides being easy to implement, our proposed hedging strategy is applicable to, in principle, all types of liabilities, including those with non-linear payoff structures. The shortcoming, relative to the strategy proposed by Cairns (2011), is that more computational effort is required if the hedge is dynamically adjusted.

There are several problems that make longevity risk more difficult to hedge than interest rate risk. Because a typical pension plan contains members who were born in different years, when we construct a longevity hedge, a group of mortality curves have to be considered simultaneously. This means that, in practice, what we need to hedge is the uncertainty arising from the evolution of a two-dimensional surface, which is composed of a collection of mortality curves. To overcome this problem, we generalize key q-durations to a two-dimensional set-up, permitting us to hedge longevity risk associated with multiple birth cohorts using only a small number of q-forward contracts.

Another challenge is population basis risk, which arises from the differences between the mortality experience of the hedger's portfolio and the national population to which the standardized instrument is linked. The issue of population basis risk has recently attracted considerable attention. Cairns et al. (2011a), Dowd et al. (2011), Jarner and Kryger (2011) and Zhou et al. (2011) contributed various multi-population stochastic mortality models for measuring population basis risk. Coughlan et al. (2010) and Stevens et al. (2011) estimated the impact of population basis risk using different measures of hedge effectiveness. However, there has not been much discussion on how a longevity hedge should be built when population basis risk exists. In this paper, we fill this gap by describing how the key q-duration strategy can be adjusted to accommodate population basis risk.

The remainder of this paper is organized as follows. Section 2 defines key q-durations and explains how they can be estimated. Section 3 details the formulation of a hedging strategy using key q-durations. Section 4 uses a synthetic pension plan to illustrate the proposed hedging strategy. Section 5 extends key q-durations to a two-dimensional set-up, which can be applied to portfolios involving multiple birth cohorts. Section 6 further generalizes the strategy to accommodate population basis risk. Finally, Section 7 concludes the paper with some suggestions for further research. The data (historical death counts and exposures from 1961 to 2007) used in our illustrations are obtained from the Human Mortality Database (2011).

2. Key q-duration

2.1. Motivations

The concept of key q-duration is motivated by a few empirical properties of mortality rates.

The first property is that over a given period of time, mortality curves are subject to non-parallel shifts, which do not take any single predetermined shape.

https://doi.org/10.2143/AST.42.2.2182804 Published online by Cambridge University Press

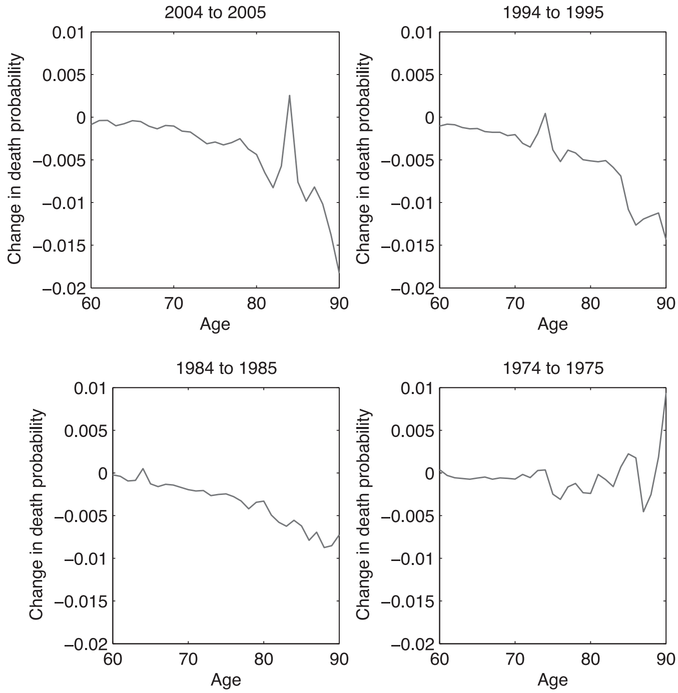

FIGURE 1: The (absolute) changes in the death probabilities, $q(x,t)$, for English and Welsh males aged 60 to 90 over the periods of 1974-1975, 1984-1985, 1994-1995 and 2004-2005.

This property is illustrated in Figure 1, in which we illustrate the shifts of the curve of death probabilities against age for English and Welsh males over different periods of time[3]. This property implies that when we measure a liability's price sensitivity to changes in the underlying mortality curve, we need to use a vector of numbers rather than just a single number, which by itself is not able to take account of different shapes of shift.

The second property is that the evolution of mortality over time can be explained adequately by a small number of latent factors. To see this property, we divide the age range, 60 to 90, into j consecutive age groups, X_i, $i = 1, \ldots, j$, of (approximately) equal size. For example, when $j = 2$, X_1 corresponds to the age group 60-75, while X_2 corresponds to the age group 76-90. Then, for each $j = 1, 2, \ldots$, we fit the following model:

$$\ln(q(x,t)) = \alpha_x + \sum_{i=1}^{j} \kappa_t^{(i)} I(x \in X_i) + \varepsilon(x,t),$$

[3] The death probabilities are calculated from the central death rates obtained from the Human Mortality Database, under the assumption that deaths are uniformly distributed over each year of age. The central death rates have been graduated by cubic splines (see Coughlan et al., 2007). The graduation removes sampling fluctuations, which may conceal the underlying statistical properties.

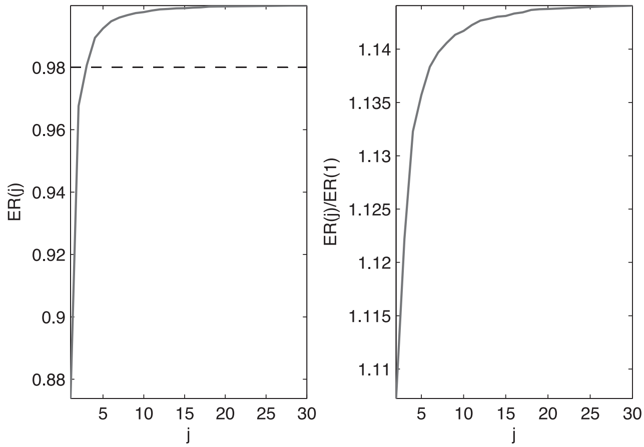

FIGURE 2: Explanation ratios, $ER(j)$, for models with different numbers of stochastic factors. The left panel shows $ER(j)$ for $j = 1, 2, \ldots$, while the right panel shows $ER(j) / ER(1)$ (i.e., $ER(j)$ relative to $ER(1)$), for $j = 2, 3 \ldots$.

where $q(x,t)$ denotes the probability that an individual aged x at the beginning of year t dies during year t, given that the individual has survived to age x; α_x is the average of $\ln(q(x,t))$ over the sample period (1961 to 2007); $\kappa_t^{(i)}$ is a time-varying stochastic factor for the ith age group; I is an indicator function; and $\varepsilon(x,t)$ is the error term[4].

Following Li and Lee (2005), we measure the goodness-of-fit for each fitted model with the following explanation ratio:

$$ER(j) = 1 - \frac{\sum_{x,t}\left(\ln(q(x,t)) - \hat{\alpha}_x - \sum_{i=1}^{j} \hat{\kappa}_t^{(i)} I(x \in X_i)\right)^2}{\sum_{x,t}\left(\ln(q(x,t)) - \hat{\alpha}_x\right)^2},$$

where the hat sign indicates the quantity underneath is an estimate. The explanation ratio can be interpreted as the proportion of the variation in $q(x,t)$ over time explained by the j stochastic factors. A higher value of $ER(j)$ thus indicates a better fit, and $ER(j) = 1$ indicates a perfect fit. Figure 2 depicts the estimated values of $ER(j)$ for $j = 1, 2, \ldots$. When $j \geq 5$, $ER(j)$ is greater than 98% and the benefit from increasing j is very marginal[5]. This property suggests that although we require a vector of numbers for measuring a liability's price sentivity to the underlying mortality curve, the dimension of this vector can be small.

[4] We fit the model by the method of least squares. Readers should keep in mind that this simple model is used for illustrating the empirical properties of historical mortality data only. The model does not fully capture the dependency across ages and time, and therefore should not be used for reserving purposes.

[5] In Appendix A, we perform a deeper analysis of what these latent factors correspond to.

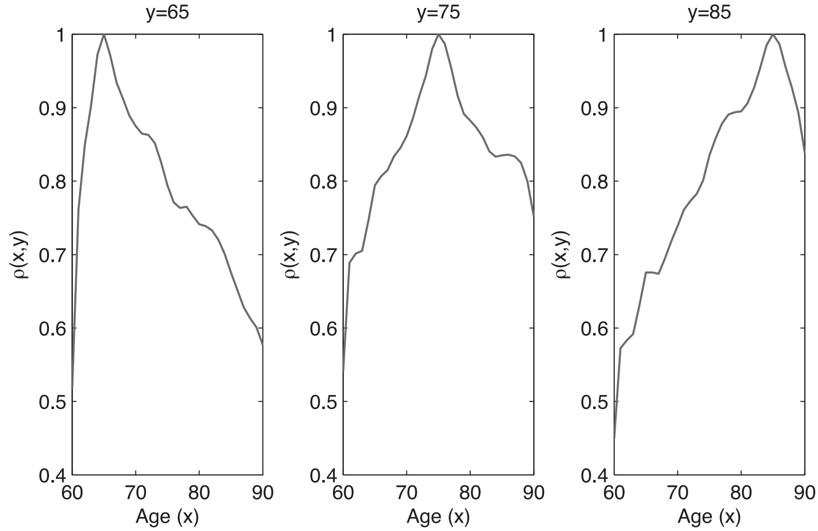

FIGURE 3: Values of $\rho(x, y)$, the sample correlation between the mortality reduction factors at ages x and y, for $y = 65, 75, 85$ and for $60 \leq x \leq 90$.

The third property is the dependence over the age dimension. To demonstrate this property, we consider the following quantity:

$$RF(x,t) = 1 - \frac{q(x,t+1)}{q(x,t)},$$

which can be interpreted as the percentage reduction in the death probability at age x from year t to year $t+1$. Let $\overline{RF}(x)$ be the mean of $RF(x,t)$ over the sample period. We calculate the sample correlation between the mortality reduction factors at ages x and y using the following formula:

$$\rho(x,y) = \frac{\sum_t (RF(x,t) - \overline{RF}(x))(RF(y,t) - \overline{RF}(y))}{\sqrt{\sum_t (RF(x,t) - \overline{RF}(x))^2}\sqrt{\sum_t (RF(y,t) - \overline{RF}(y))^2}}.$$

In Figure 3 we display the calculated values of $\rho(x,y)$[6]. For example, the right panel shows the values of $\rho(x, 85)$ for different values of x. The analysis indicates that mortality reductions at neighboring ages are significantly correlated with one another, but that the strength of the dependence diminishes as the age gap becomes wider. This property should also be taken into consideration when we estimate mortality durations.

[6] The death probabilities involved are calculated using smoothed central death rates from the Human Mortality Database and the assumption that deaths are uniformly distributed over each year of age.

2.2. Defining key q-durations

Properties 1 and 2 described in Section 2.1 motivate us to measure a portfolio's price sensitivity to changes in the underlying mortality curve by a vector of numbers, each of which measures the portfolio's price sensitivity to a shift at a certain key point on the mortality curve.

We let n be the number of key points, and let ages $x_1, x_2, \ldots, x_n$ (from smallest to largest) be the key points on a mortality curve $\mathbf{q}$, the vector of age-specific cohort death probabilities. The death probabilities at the key ages $x_1, x_2, \ldots, x_n$ are called the *key mortality rates*.

A portfolio's price sensitivity to a shift in a key mortality rate is referred to as a *key q-duration*. Suppose that there is a change of $\delta(j)$ in jth key mortality rate. Let $\mathbf{q}$ and $\tilde{\mathbf{q}}(\delta(j))$ be the mortality curves before and after the shift in the jth key mortality rate. We can express the jth key q-duration of a portfolio as

$$KQD(P(\mathbf{q}), j) = \lim_{\delta(j) \to 0} \frac{P(\tilde{\mathbf{q}}(\delta(j))) - P(\mathbf{q})}{\delta(j)},$$

where $P(\mathbf{q})$ denotes the value of the portfolio on the basis of the mortality curve $\mathbf{q}$. The number $KQD(P(\mathbf{q}), j)$ measures the portfolio's price sensitivity to the jth key rate, whereas the vector $\{KQD(P(\mathbf{q}), j);\ j = 1, 2, \ldots, n\}$ as a whole measures the portfolio's price sensitivity to the entire mortality curve[7].

We have the following comments on the choice of the key ages:

1. As we are going to demonstrate in Section 3.2, using q-forwards that are linked to the key mortality rates makes the calibration exercise simple, because in this way the appropriate notional amounts of the q-forwards in the hedge portfolio can be determined independently. Therefore, the choice of key mortality rates (or key ages) depends on what q-forwards are available in the market. In the early stages of the market's development, it is expected that transactions are restricted to a limited number of q-forwards, which are most likely to be linked to representative ages such as 65, 70, 75, etc.
2. In Appendix A, a multivariate factor analysis on the historical values of $R(x, t) = 1 - q(x, t+1) / q(x, t)$, for $x = 60, \ldots, 90$ and for $t = 1961, \ldots, 2006$, is performed. In the analysis, we identify five latent factors that represent the age-specific mortality reductions. The factors identified roughly correspond to five age groups, which are centered at ages 64, 70, 75, 80 and 86. The conclusion from this analysis supports the use of key ages that are spread uniformly over the mortality curve.

---

[7] Plat (2010) employs a duration-convexity approach to approximate the impact of a change in mortality rates on the value of a liability. The use of this approximation is to speed up the calculation of the liability's value-at-risk. Plat's approach is different from our key q-duration approach in that only one number (instead of a vector of numbers) is used for measuring duration, and that the age dependence property is not explicitly modeled.

3. In Section 4.4, we determine the optimal key ages for the illustrative longevity hedge we consider. Among all possible combinations of five key ages, the combination "62, 67, 73, 79, 85" leads to the maximum hedge effectiveness. The optimal key ages found suggest that it makes sense to pick key ages that represent different segments of the mortality curve.
4. The precise combination of key ages that leads to the best hedge effectiveness depends on the structure of the pension/annuity liability as well as the interest rate at which the liability cash flows are discounted. In a high interest rate regime, cash flows in the distant future are not so much important, and therefore the optimal key ages could be lower. The opposite is true in a low interest rate regime. Being dependent on the liability structure and interest rate is a limitation of key q-durations. This problem may be avoided in the method proposed by Cairns (2011), which approximates the underlying survival function rather than the pension/annuity liability. However, this alternative method may potentially require more computational effort.

2.3. Calculating key q-durations

Property 3 (age dependence) described in Section 2.1 means that the change $\delta(j)$ in the jth key mortality rate is associated with changes in other parts of $\mathbf{q}$. There is a need to approximate this association in estimating $KQD(P(\mathbf{q}), j)$.

To capture age dependence, we assume that the change $\delta(j)$ in the jth key mortality rate is accompanied with changes in mortality rates at ages that are close enough to x_j. Let $s(x, j, \delta(j))$ be the shift at age x associated with a change $\delta(j)$ in the jth key mortality rate. The function s, which may be viewed as an analog to a kernel in density estimation, can take different forms. We suggest the following specification, which uses a linear interpolation to approximate the diminishing dependence between two mortality rates as the age gap becomes wider.

For $2 < j < n - 1$,

$$s(x, j, \delta(j)) = \begin{cases} 0 & x \leq x_{j-1} \\ \frac{\delta(j)(x - x_{j-1})}{x_j - x_{j-1}} & x_{j-1} < x \leq x_j \\ \frac{\delta(j)(x_{j+1} - x)}{x_{j+1} - x_j} & x_j < x < x_{j+1} \\ 0 & x \geq x_{j+1} \end{cases} \tag{1}$$

Note that the impact of $\delta(j)$ becomes zero when the next or the previous key age is reached. The specification of s is slightly different for $j = 1$ and $j = n$. For $j = 1$, we change s to $\delta(1)$ when $x \leq x_1$, and for $j = n$, we change s to $\delta(n)$ when $x > x_n$. A similar specification is also used by Ho (1992) to model shifts in an interest rate yield curve.

The shift in the whole mortality curve is approximated by the sum of $s(x, j, \delta(j))$, $j = 1, 2, \ldots, n$. This means that for any age x between x_j and x_{j+1},

https://doi.org/10.2143/AST.42.2.2182804 Published online by Cambridge University Press

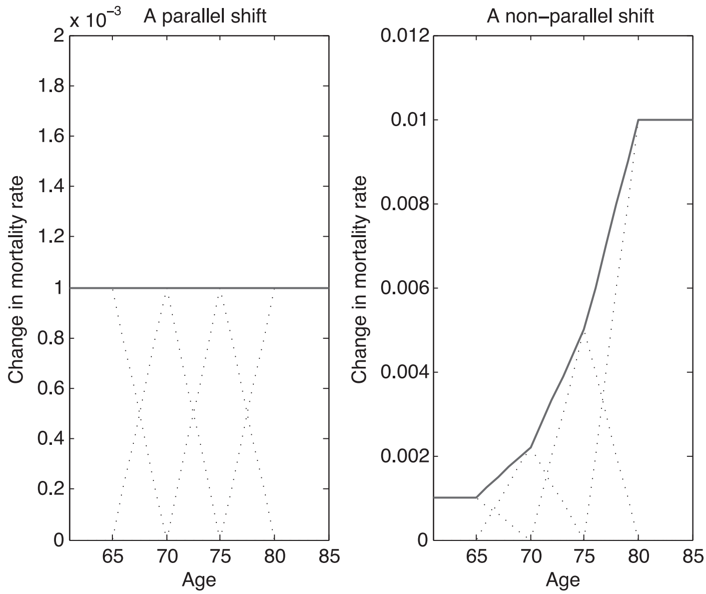

FIGURE 4: Approximation of the shift in a mortality curve as a sum of $s(x, j, \delta(j))$, $j = 1, 2, 3, 4$. The values shown are arbitrary.

the shift of the mortality rate at age x is simply the weighted average of $\delta(j)$ and $\delta(j+1)$, with the weights being $\frac{x_{j+1}-x}{x_{j+1}-x_j}$ and $\frac{x-x_j}{x_{j+1}-x_j}$, respectively. In Figure 4 we display the overall shift resulting from arbitrary shifts in four key mortality rates, which are located at ages $x_1 = 65$, $x_2 = 70$, $x_3 = 75$, and $x_4 = 80$. The diagrams illustrate how parallel and non-parallel shifts can be approximated by the above specification.

In most cases, it is difficult to analytically calculate a key q-duration. For practical purposes, we may estimate $KQD(P(\mathbf{q}), j)$ by using the following algorithm:

1. take $\mathbf{q}$ as the best estimate of the underlying mortality curve;
2. assuming $\delta(j)$ is 10 basis points, calculate $\tilde{\mathbf{q}}(\delta(j))$;
3. set $KQD(P(\mathbf{q}), j)$ to $\frac{P(\tilde{\mathbf{q}}(\delta(j))) - P(\mathbf{q})}{\delta(j)}$.

3. BUILDING A LONGEVITY HEDGE

3.1. Hedging instruments

Our hedging framework is based on q-forward contracts. A q-forward contract is a zero-coupon swap that exchanges on the maturity date a fixed amount, determined at time 0, for a random floating amount that is proportional to an age-specific death probability (the reference rate) for a certain population (the reference population) in some future time (the reference year).

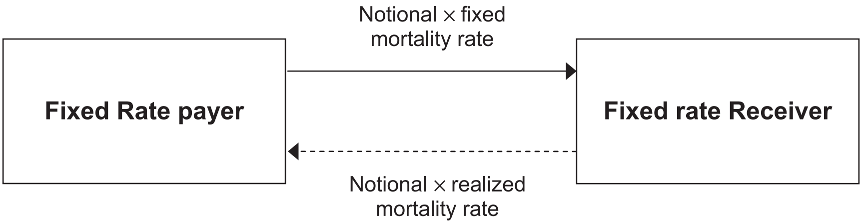

FIGURE 5: Settlement of a q-forward at maturity.

For a q-forward contract with a reference age x^* and a reference year t^*, the reference rate to which the floating payment is linked is $q(x^*, t^*)$, which is a random unknown at time 0. The fixed payment is proportional to the corresponding forward mortality rate[8], which we denote by $q^f(x^*, t^*)$, for the reference population. This rate is a fixed constant, and is determined in such a way that no payment exchanges hands at time 0.

At maturity, a net payment will be made by one counterparty or the other. The settlement that takes place at maturity is illustrated diagrammatically in Figure 5. Mathematically, per \$1 notional, the fixed rate receiver will receive an amount of

$$q^f(x^*, t^*) - q(x^*, t^*)$$

from the fixed rate payer[9].

In practice, the maturity date T^* (the time at which the settlement takes place) may be slightly later than the reference year t^*, because of the time lag in the availability of the mortality index data. From the perspective of a fixed rate receiver, the (random) present value (per \$1 notional) of the mortality forward can be expressed as

$$(1 + r)^{-(T^* - t_0)} (q^f(x^*, t^*) - q(x^*, t^*)),$$

where t_0 is the current date, and r is the interest rate at which cash flows are discounted.

It is easy to see that an entity wishing to hedge longevity risk could enter into a portfolio of q-forwards in which it receives fixed mortality rates and pays realized mortality rates. In this way, at maturity, the q-forwards will pay out to the hedger an amount that increases as future mortality rates fall to offset the unexpected increase in the hedger's liability. Therefore, if the weight

[8] The term 'forward mortality rate' is used in the sample term sheet available on LLMA's website and in JP Morgan's LifeMetrics documentation (Coughlan et al., 2007). More precisely, it should be referred to as the forward death probability instead.

[9] We follow the specification described in the sample q-forward term sheet shown on LLMA's website.

on each q-forward is calibrated properly, the resulting longevity hedge can stabilize the hedger's liability with respect to changes in future mortality rates.

The fixed payer can be an investor wishing to take longevity risk for a risk premium. To attract investors (fixed rate payers), the forward mortality rate must be smaller than the corresponding expected mortality rate, that is,

$$q^f(x^*, t^*) < \mathbb{E}(q(x^*, t^*)),$$

where $\mathbb{E}$ denotes the expectation under the real-world probability measure, so that on average (i.e., if mortality is realized as expected), the investor will be paid. The difference between the expected and forward rates, therefore, indicates the expected risk premium to the investor. In Figure 6 we illustrate the relationship between forward and expected mortality rates at a certain age for different times to maturity. A widening divergence is expected, because investors demand a higher risk premium from a longer term contract, which involves a prediction further into the future.

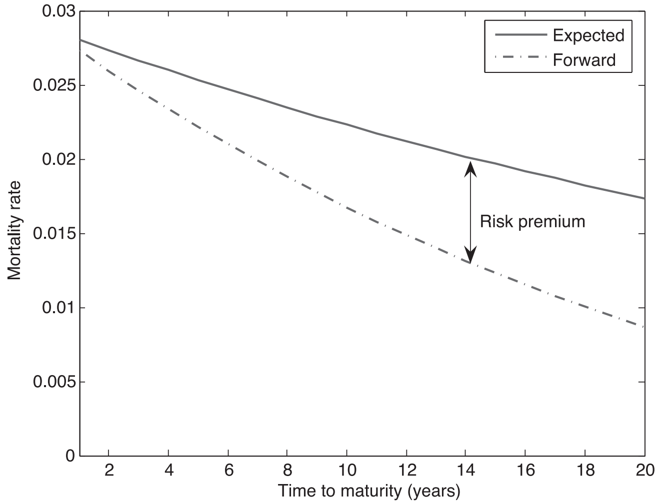

FIGURE 6: Illustrative curves of expected mortality rates and forward mortality rates. The gap between the two curves represents the expected risk premiums for different times to maturity.

3.2. Calibrating a longevity hedge

Consider a liability that depends on the future mortality of a single cohort of individuals, who were all born in year c. On the current date t_0, the (random) present value of this liability is $L(\mathbf{q})$, where $\mathbf{q}$ denotes the vector of (random) future death probabilities for this cohort of individuals.

On the current date t_0, the hedger sets up a longevity hedge for this liability using a combination of q-forward contracts. Suppose that there are n q-forwards that are linked to the cohort of individuals in question. In forming a hedging

https://doi.org/10.2143/AST.42.2.2182804 Published online by Cambridge University Press

strategy, we choose key mortality rates that correspond to the mortality rates to which the n available q-forwards are linked. This is to say that the key mortality rates are

$$q(x_j, t_j), \quad , 1, 2, \ldots, n,$$

where $t_j - x_j = c$, $x_1 < x_2 < \ldots, x_n$ and $t_1 < t_2 < \ldots, t_n$, and that the jth q-forward has a floating leg that depends on the random death probability $q(x_j, t_j)$ and a fixed leg that is depends on the corresponding fixed forward mortality rate $q^f(x_j, t_j)$. On the current date t_0, the hedge portfolio has a value of

$$H(\mathbf{q}) = \sum_{j=1}^{n} w(j) F_j(\mathbf{q}),$$

where $F_j(\mathbf{q})$ is the random value of the jth q-forward (per \$1 notional) on the current date t_0, and $w(j)$ is the notional amount of the jth q-forward in the hedge portfolio. Note that the hedger participates in the q-forward contracts as a fixed rate receiver.

To make the hedge effective, the liability and the hedge portfolio must have similar price sensitivities to the underlying mortality curve $\mathbf{q}$. This can be achieved by setting

$$KQD(L(\mathbf{q}), j) = KQD(H(\mathbf{q}), j), \tag{2}$$

for $j = 1, 2, \ldots, n$.

In general, there is no analytical solution to the key q-durations for the liability. They can be estimated by using the algorithm presented in Section 2.3.

By contrast, the key q-durations for a q-forward can be computed analytically. As we explained in Section 3.1, the (random) value (per \$1 notional) of the jth q-forward on the current date t_0 can be written as

$$F_j(\mathbf{q}) = (1+r)^{-(T_j - t_0)}(q^f(x_j, t_j) - q(x_j, t_j)),$$

where r is the interest rate at which cash flows are discounted, and T_j is the maturity date of q-forward contract[10]. The jth key rate $q(x_j, t_j)$ is only random quantity involved in $F_j(\mathbf{q})$, and $F_j(\mathbf{q})$ is a linear function of $q(x_j, t_j)$. It is therefore obvious that

$$KQD(F_j(\mathbf{q}), j) = -(1+r)^{-(T_j - t_0)}. \tag{3}$$

Furthermore, according to the way we specify the key rate shifts (equation (1)), the impact of a shift in a key mortality rate reduces to zero when the next or the previous key age is reached. It follows that the shift in any key mortality

[10] As we mentioned in Section 3.1, the maturity date T_j may be slightly later than the reference year t_j, because of the time lag in the availability of the mortality index data.

rate other than $q(x_j, t_j)$ has no impact on $F_j(\mathbf{q})$. Therefore, $KQD(F_j(\mathbf{q}), i)$ must be zero for any $i \neq j$.

The property that $KQD(F_j(\mathbf{q}), i) = 0$ for all $i \neq j$ allows us to determine the appropriate notional amounts of the q-forward contracts independently, without the need of solving a system of equations. Specifically, to satisfy equation (2), we need a notional amount of

$$w(j) = \frac{KQD(L(\mathbf{q}), j)}{KQD(F_j(\mathbf{q}), j)}$$

for the jth q-forward contract.

The property that $KQD(F_j(\mathbf{q}), i) = 0$ for all $i \neq j$ also makes mortality forwards an analog to zero-coupon bonds in Ho's (1992) framework for hedging interest rate risk. In more detail, Ho (1992) demonstrated that the price of a zero-coupon bond is sensitive only to a shift in the corresponding key interest rate but not other key rates on the yield curve. It follows that one can easily construct an interest rate hedge with zero-coupon bonds, as the required face amounts of these bonds can be estimated independently.

4. An Illustration

4.1. The hedge

We now illustrate the hedging framework with a simple synthetic pension plan, which pays a pensioner \$1 at the beginning of each year until the pensioner dies or reaches age 91, whichever is earlier. In the illustration, the following assumptions are made.

1. The current date t_0 is the beginning of year 2008. We use this assumption because we use historical mortality data up to and including year 2007.
2. The pensioner is exactly 60 years old on the current date.
3. The pensioner's mortality experience is exactly the same as that of English and Welsh males with the same year of birth. This assumption will be relaxed later when population basis risk and sampling risk are studied.
4. The best estimate of the underlying mortality curve, that is, the vector of

$$\mathbb{E}(q(60+k, 2008+k)), \quad k = 0, 1, \ldots,$$

is based on the central projection made by the Cairns-Blake-Dowd model (Cairns et al., 2006), fitted to the data from the population of English and Welsh males. The specification of the Cairns-Blake-Dowd model is detailed in Appendix B[11].

[11] The best estimate is obtained by switching off the random components in the period effect terms of the Cairns-Blake-Dowd model.

TABLE 1

KEY q-DURATIONS OF THE PENSION LIABILITY AND THE q-FORWARDS IN HEDGE PORTFOLIO. THE RATIO OF $KQD(F_j(\mathbf{q}), j)$ TO $KQD(L(\mathbf{q}), j)$ IS THE REQUIRED NOTIONAL AMOUNT $w(j)$ FOR THE jTH q-FORWARD.

	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$
x_j	65	70	75	80	85
$KQD(F_j(\mathbf{q}), j)$	– 0.8375	– 0.7224	– 0.6232	– 0.5375	– 0.4637
$KQD(L(\mathbf{q}), j)$	– 97.9501	– 37.3005	– 23.1447	– 12.5785	– 6.2496
$w(j)$	116.96	51.63	37.14	23.40	13.48

5. The interest rate at which cash flows are discounted is $r = 3\%$.
6. q-forwards linked to ages 65, 70, 75, 80 and 85 for the same birth cohort are available. In our baseline calculations, we make use of all five available q-forwards, and set the key ages to these five ages.
7. For each mortality forward, the maturity date (the date when the payment is settled) is one year after the reference year. That is, we assume $T_j = t_j + 1$ for $j = 1, 2, 3, 4, 5$.
8. The forward mortality rates are the same as the corresponding best estimate mortality rates. That is,

$$q^f(x_j, t_j) = \mathbb{E}(q(x_j, t_j)), \quad j = 1, 2, 3, 4, 5.$$

   This assumption, which implies zero risk premium, would affect the cost but not the performance of the longevity hedge.

The key q-durations for the q-forwards are calculated analytically using equation (3), while those for the pension liability are calculated numerically using the algorithm described in Section 2.3. The calculated key q-durations are displayed in Table 1. Also shown in the table are the required notional amounts of the q-forwards in the hedge portfolio.

4.2. Hedge effectiveness

We evaluate the effectiveness of the longevity hedge by comparing the variability in the present value of all unexpected cash flows with and without the hedge in place. If the pension plan is unhedged, then on the current date t_0, the (random) present value of the unexpected cash flows from the plan can be expressed as

$$X = L(\mathbf{q}) - L(\mathbb{E}(\mathbf{q})),$$

By contrast, if there is a longevity hedge, then on the current date t_0, the (random) present value of the unexpected cash flows from the plan and the hedge can be written as

$$X^* = L(\mathbf{q}) - L(\mathbb{E}(\mathbf{q})) - H(\mathbf{q}) + H(\mathbb{E}(\mathbf{q})).$$

The longevity hedge is effective if X^* is significantly less variable than X. As such, we can measure hedge effectiveness in terms of the amount of risk reduction, R, which is defined by

$$R = 1 - \frac{\sigma^2(X^*)}{\sigma^2(X)},$$

where $\sigma^2(X)$ and $\sigma^2(X^*)$ are the variances of X and X^*, respectively. A higher value of R means a better hedge effectiveness.

To evaluate the robustness of the hedging strategy relative to the simulation model used, we consider three different stochastic mortality models:

- The original Cairns-Blake-Dowd (CBD) model (Cairns et al., 2006)
- A generalized Cairns-Blake-Dowd (G-CBD) model with a quadratic age effect term and a cohort effect term (Cairns et al., 2009)
- The Lee-Carter (LC) model (Lee and Carter, 1992)

These models are fitted to historical mortality data (years 1961 to 2007, ages 60 to 90) from English and Welsh male population. Complete specifications of these models are provided in Appendix B. For each model, we simulate 5,000 realizations of $\mathbf{q}$. Parameter uncertainty is incorporated into the simulated $\mathbf{q}$ by the parametric bootstrap, which we detail in Appendix C. Then, for each realization of $\mathbf{q}$, we calculate the values of X and X^*. This creates empirical distributions of X and X^*, from which the amount of risk reduction R (based on that particular mortality model) can be calculated. Note that these models are used for evaluating hedge effectiveness only, and are not involved in the calibration of the longevity hedge.

The resulting hedge effectiveness is summarized in Table 2. Using a portfolio of five q-forwards, one achieve a risk reduction of about 97%. The amounts of risk reduction calculated from different simulation models are broadly similar, indicating that the substantial hedge effectiveness is not because of the imposed model, but it would be able to achieve in reality.

TABLE 2

THE AMOUNTS OF RISK REDUCTION (R) RESULTING FROM HEDGE PORTFOLIOS WITH DIFFERENT NUMBER OF q-FORWARDS.

Simulation model	5 q-forwards	4 q-forwards	3 q-forwards
CBD	97.3%	94.9%	83.0%
G-CBD	97.5%	95.5%	83.7%
LC	97.6%	96.0%	84.3%

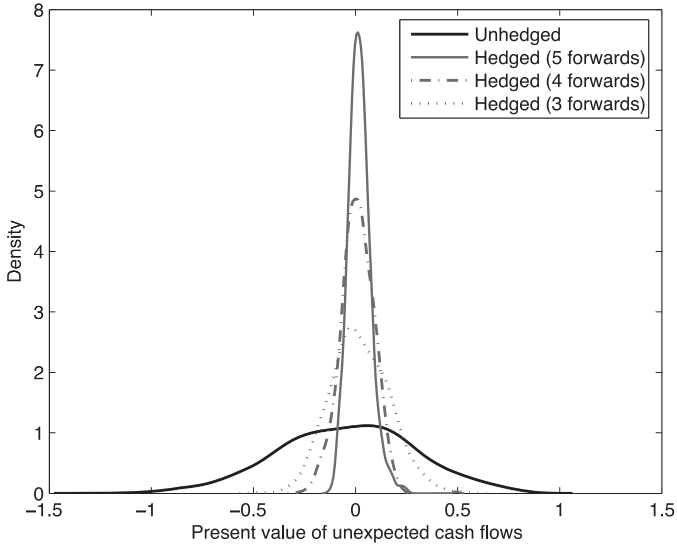

FIGURE 7: The simulated distributions of X, and the simulated distributions of X^* based on hedge portfolios with 3, 4 and 5 mortality forwards. The simulation model used in the CBD model with parameter uncertainty.

To examine how the amount of risk reduction is related to the number of q-forwards used, we repeat the analysis by using four mortality forwards (linked to ages 65, 70, 75 and 80) and three mortality forwards (linked to ages 65, 70, 75). The hedge effectiveness is reduced to about 95% when four q-forwards are used and to about 83% when three q-forwards are used. The relationship between hedge effectiveness and the number of q-forwards used can be visualized from Figure 7, in which we show the simulated distributions of X^* that are based on hedge portfolios with different numbers of q-forwards.

4.3. Proximity to optimality

The longevity hedge based on key q-durations is effective, but may not be the most effective, as key q-durations are only approximate measurements of the sensitivity to $\mathbf{q}$. Here we study if a more effective hedge can be formed with the same q-forwards, and if it can be formed, how much more effective it would be.

Suppose that our sole objective is to maximize the amount of risk reduction, R, produced by the hedge. To achieve this objective, we can gradually adjust the weight on each of the five q-forwards until the maximum value of R is attained. The value of R in each iteration can be estimated by simulating from a stochastic mortality model. In our illustration, the CBD model fitted to English and Welsh male data is used.

We found that, on the basis of the five q-forwards we consider, the maximum achievable amount of risk reduction is 98.5%, which is 1.2 percentage points higher that achieved by using key q-duration. The result indicates that key q-durations can yield a risk reduction that is not too far from the best one can achieve.

Note that the optimization procedure, which involves simulations and a maximization of a multivariable function, demands a lot of computational resources. It works for the synthetic pension plan we consider, but does not seem to be feasible for more sophisticated real-life pension plans. Key q-durations, by contrast, are much easier to calculate and are not dependent on a specific simulation model.

4.4. The optimal combination of key mortality rates

In forming a hedging strategy, we choose key mortality rates that correspond to the rates to which the q-forwards available in the market are linked. In the early stages of the market's development, it is expected that transactions are restricted to a limited number of contracts in which liquidity can be concentrated. Therefore, the choice of key mortality rates is exogenous, depending on the instruments that are available in the market. In our baseline calculations, we assumed that q-forwards linked to ages 65, 70, 75, 80 and 85 for the cohort in question are available, and chose the key mortality rates accordingly.

Now suppose that q-forwards linked to all rates on the underlying mortality curve **q** are available. In this case, which five key mortality rates (q-forwards) should we choose? We answer this question by comparing the hedge effectiveness provided by all possible combinations of key mortality rates. Since the synthetic pension plan involves 31 mortality rates (from ages 60 to 90), there are altogether $\frac{31!}{5! \times 26!} = 169\,911$ possible combinations. The simulations here are based on the CBD model with parameter uncertainty. Except the assumption regarding the availability of q-forwards, all assumptions made in Section 4.1 remain unchanged.

We found that the maximum amount of risk reduction is attained when the key ages are 62, 67, 73, 79 and 85. These ages are roughly evenly spaced, and are close to the key mortality rates used in the baseline calculation. This combination of key mortality rates yields a hedge effectiveness of 98.3%, which is one percentage point higher than that achieved in the baseline calculation. Our findings suggest that, to concentrate liquidity and to attract demand from pension plans, it makes sense for investment banks to issue q-forwards that are linked to rates representing different segments on a mortality curve.

4.5. The effect of sampling risk

In previous illustrations, we assume that there is no sampling risk (or small sample risk), that is, the risk that the realized mortality experience is different from the true mortality rate. Sampling risk is diversifiable, so it does not matter much if the pension plan is sufficiently large. However, for smaller plans, sampling risk may be significant and may affect the hedge's effectiveness.

Let us consider again the synthetic pension plan described in Section 4.1 and suppose that all assumptions (except the assumption that there is no sampling risk) continue to hold. We now assume that the plan involves a single cohort of individuals, who were all born in year 1948 (aged 60 on the current day). We let $l(60)$ be the initial number of pensioners, and $l(x)$ be the number of pensioners who will survive to age x, where $x = 61, 62, \ldots$. We incorporate sampling risk by treating the cohort of pensioners as a random survivorship group. This means that $l(x)$ for $x > 60$ is still a random variable even if the underlying mortality curve $\mathbf{q}$ is completely known.

We model sampling risk with the following binomial death process:

$$l(x+1) \sim \text{Binomial}(l(x), 1 - q(x, 1948 + x)), \quad x = 60, 61, \ldots, 90,$$

which is then incorporated into the procedure for simulating unexpected cash flows as follows:

1. simulate a mortality curve, $\mathbf{q}$ using the fitted CBD model with parameter uncertainty;
2. for each simulated $\mathbf{q}$, simulate the number of survivors $l(x)$, $x = 61, \ldots, 91$;
3. calculate cash flows on the basis of the simulated $l(x)$;
4. repeat steps 1-3 to derive empirical distributions of the unexpected cash flows.

The q-forwards used are the same as those in Section 4.1. Note that there is no change to the key q-durations and hence the notional amounts.

In Figure 8 we depict the simulated distributions of X and X^* for different values of $l(60)$. When $l(60) = 10{,}000$, the hedge appears to be equally effective as that when sampling risk is not taken into account. However, for smaller values of $l(60)$, sampling risk has a significant effect on the reduction in volatility. To quantify the effect of sampling, we calculate the amount of risk reduction, R, for different values of $l(60)$ (see Table 3). The amount of risk reduction is reduced to only 65% when the initial number of pensioners is 500. This analysis

TABLE 3

THE AMOUNTS OF RISK REDUCTION (R) FOR DIFFERENT INITIAL SIZES ($l(60)$) OF THE SYNTHETIC PENSION PLAN. THE SIMULATIONS ARE BASED ON THE CBD MODEL WITH PARAMETER UNCERTAINTY.

$l(60)$	R
$+\infty$	97.3%
10,000	95.2%
3,000	90.1%
1,000	78.3%
500	65.1%

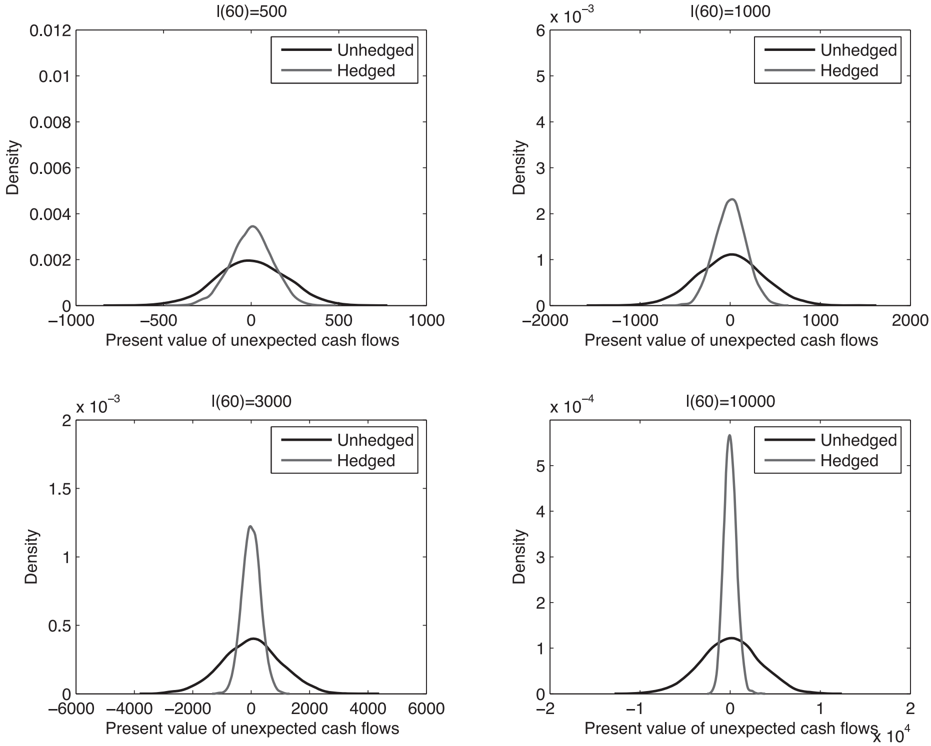

FIGURE 8: Simulated distributions of X and X^* for different initial sizes ($l(60)$) of the synthetic pension plan. The simulations are based on the CBD model with parameter uncertainty.

suggests that for a small pension plan with say fewer than 500 members, an indemnity hedge may be a better alternative to one that is based on standardized instruments. The impact of population size and sampling risk is also considered Cairns et al. (2011b). They look at a different problem of value-hedging a simple liability in 10 years but the conclusions are similar.

4.6. Hedging longevity risk associated with more advanced ages

One may wonder how the longevity hedge would perform if mortality rates at ages beyond 91 are involved. This problem deserves a separate analysis, because investment banks may not be willing to issue q-forwards for more advanced ages, at which the uncertainty in mortality improvement is high. As a matter of fact, the LifeMetrics index, the index to which q-forwards issued by JP Morgan are linked, is available only up to age 89.

Now, let us extend the synthetic pension plan to age 101. That is, we assume that the plan pays the pensioner \$1 at the beginning of each year until the pensioner dies or reaches age 101, whichever is earlier. The q-forward contracts in the hedge portfolio are still linked to ages 65, 70, 75, 80 and 85, as contracts

https://doi.org/10.2143/AST.42.2.2182804 Published online by Cambridge University Press

linked to more advanced ages are unlikely to be available in the market. As before, we calibrate the hedge by using key q-durations, with key ages the same as the ages to which the five q-forwards are linked. All assumptions stated in Section 4.1 remain.

To examine the effectiveness of this hedge, we use a CBD model that is fitted to English and Welsh males mortality data from year 1961 to 2007 and from age 60 to 101. As in previous analyses, parameter uncertainty is incorporated into the simulations by the parametric bootstrap. The resulting amount of risk reduction (R) is 95.1%, which is 2.2 percentage points lower than that for the plan that does not involve ages beyond 90. The reduction in hedge effectiveness is potentially because the risk associated with the higher ages is under-hedged.

It would be interesting to examine the hedge effectiveness for a plan that includes even more advanced ages. However, the lack of reliable mortality data for this age range prohibits us from performing the investigation. In the Human Mortality Database, "data" beyond age 100 are often extrapolated values, based on the extinct cohorts method. Data at extreme ages are not available from other sources such as the LifeMetrics database[12].

5. Extension 1: Hedging Multiple Cohorts

In Section 2, we presented key q-durations in a one-dimensional setting. This simple setting is easy to understand, and is adequate for hedging longevity risk associated with one cohort. However, in practice, pension plans involve multiple birth cohorts, and therefore when we construct a longevity hedge, we may need to consider a group of mortality curves simultaneously. In this section, we generalize key q-durations to a two-dimensional setting, which can then be used to hedge the uncertainty arising from the evolution of a two-dimensional mortality surface that is composed of a collection of mortality curves.

We illustrate the two-dimensional generalization using a synthetic pension plan with members ranging from age 60 to 80 on the current date (the beginning of year 2008). The bar chart in Figure 9 summarizes the initial demographic structure of the plan. We assume again that the plan pays each pensioner \$1 at the beginning of each year until the pensioner dies or reaches age 91, whichever is earlier. For simplicity, we assume further that the plan is closed, that is, there are no new entrants to the plan after the current date.

The plan involves 21 birth cohorts, with the youngest born in 1948 and the oldest in 1928. In total, it is subject to the uncertainty associated with 441 future death rates. Assuming the maximum separation between two key mortality rates in a mortality curve is five ages, our strategy in a one-dimensional setting would require 76 q-forwards. It is unlikely that the longevity market can provide such a wide variety of q-forwards. Even if the required contracts

[12] www.lifemetrics.com

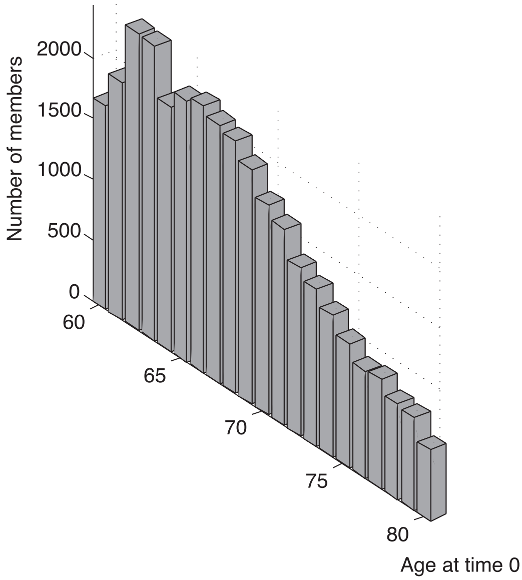

FIGURE 9: The initial demographic structure of the illustrative multi-cohort pension plan.

are available, the complexity of the hedge would make it costly and difficult to manage in the future.

To mitigate this problem, we attempt to consider the potential dependence along the year of birth dimension. Let us consider the group of individuals who were born in year c. We let

$$RF^*(x, c) = 1 - \frac{q(x, c+x+1)}{q(x, c+x)}$$

be their mortality reduction factor at age x[13], and let $\overline{RF}^*(c)$ be the mean of $RF^*(x,c)$ over the sample age range for that cohort. To examine the dependence across cohorts, we calculate the sample correlation between the reduction factors for years of birth c and d with the following formula:

$$\rho^*(c,d) = \frac{\sum_x (RF^*(x,c) - \overline{RF}^*(c))(RF^*(x,d) - \overline{RF}^*(d))}{\sqrt{\sum_x (RF^*(x,c) - \overline{RF}^*(c))^2}\sqrt{\sum_x (RF^*(x,d) - \overline{RF}^*(d))^2}},$$

where the summations are taken over the common sample age range for both cohorts.

[13] We can express $RF^*(x,c)$ as $RF(x,c+x)$, the mortality reduction factor we defined previously in Section 2.1.

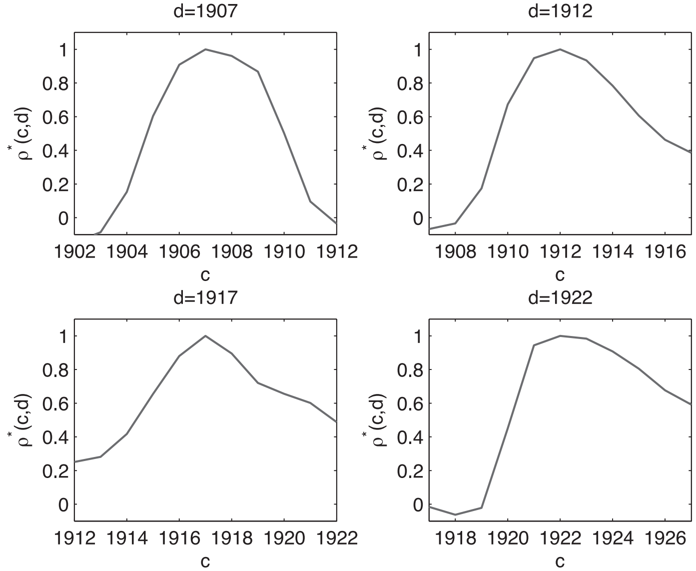

FIGURE 10: Values of $\rho^*(c,d)$, the sample correlation between the mortality reduction factors for years of birth c and d, for $d = 1907, 1912, 1917, 1922$ and for $c = d-5, d-4, \ldots, d+5$.

The calculated values of $\rho^*(c,d)$ for $d = 1907, 1912, 1917, 1922$ and for $c = d-5, d-4, \ldots, d+5$ are displayed in Figure 10[14]. The patterns of $\rho^*(c,d)$ indicate that mortality improvement rates of neighboring cohorts are significantly correlated with one another, but the strength of the correlation tapers off as the birth cohorts are wider apart. The analysis of $\rho^*(c,d)$ motivates us to generalize the way in which a key rate shift is modeled. Specifically, in what follows, we assume that a shift in a key mortality rate would affect mortality rates for not only the same but also the neighboring cohorts.

In the two-dimensional setting, a mortality surface (i.e., the collection of mortality curves) is represented by the key mortality rates for m key cohorts, who were born in years $c_1, \ldots, c_m$ (from smallest to largest), respectively. The mortality curve for key cohort k, $k = 1, \ldots, m$, contains n_k key mortality rates, which are located at ages $x_{1,k}, x_{2,k}, \ldots, x_{n_k,k}$ (from smallest to largest). We identify the jth key mortality rate for the kth key cohort as the (j,k)th key mortality rate.

To hedge the synthetic pension plan, we consider $k = 4$ key cohorts, with $c_1 = 1933$, $c_2 = 1937$, $c_3 = 1941$ and $c_4 = 1945$. Each key cohort contains key mortality rates that are no more than five ages apart. The strategy we use can be visualized from the Lexis diagram in Figure 11. The locations of the 14 key mortality rates used are shown in Table 4. We assume that q-forwards linked to the key mortality rates are available in the market. The q-forward associated

[14] The death probabilities involved in $\rho^*(c,d)$ are computed using smoothed central death rates from the Human Mortality Database and the assumption that deaths are uniformly distributed over each year of age.

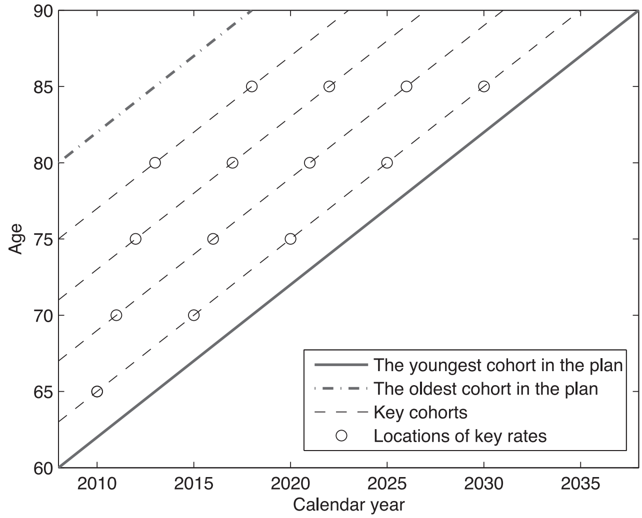

FIGURE 11: The key cohorts used in the hedging strategy for the multi-cohort synthetic pension plan.

with the (j,k)th key mortality rate has a reference age $x_{j,k}$ and a reference year $t_{j,k} = c_k + x_{j,k}$.

Our goal is to model $s(x, c, (j,k), \delta(j,k))$, the impact of the shift $\delta(j,k)$ in the (j,k)th key mortality rate on the mortality rate for age x and for year of birth c. For $j = 2, \ldots, n_k - 1$ and for $k = 2, \ldots, m - 1$, we set

$$s(x, c, (j,k), \delta(j,k)) = \delta(j,k)\, \alpha(x,j,k)\, \beta(c,k),$$

where

$$\alpha(x,j,k) = \begin{cases} 0 & x \leq x_{j-1,k} \\ \frac{x - x_{j-1,k}}{x_{j,k} - x_{j-1,k}} & x_{j-1,k} < x \leq x_{j,k} \\ \frac{x_{j+1,k} - x}{x_{j+1,k} - x_{j,k}} & x_{j,k} < x < x_{j+1,k} \\ 0 & x \geq x_{j+1,k} \end{cases}$$

TABLE 4

LOCATIONS OF THE KEY MORTALITY RATES USED IN HEDGING THE MULTI-COHORT SYNTHETIC PENSION PLAN.

k	n_k	Locations of key mortality rates
1	2	Ages 80, 85
2	3	Ages 75, 80, 85
3	4	Ages 70, 75, 80, 85
4	5	Ages 65, 70, 75, 80, 85

and

$$\beta(c,k) = \begin{cases} 0 & c \le c_{k-1} \\ \frac{c - c_{k-1}}{c_k - c_{k-1}} & c_{k-1} < c \le c_k \\ \frac{c_{k+1} - c}{c_{k+1} - c_k} & c_k < c < c_{k+1} \\ 0 & c \ge c_{k+1} \end{cases}$$

Similar to the one-dimensional setting, the effect of $\delta(j,k)$ diminishes as x is farther way from $x_{j,k}$, and becomes zero when $x_{j-1,k}$ or $x_{j+1,k}$ is reached. The dependence along the year of birth dimension is also modeled in a similar manner. In particular, the impact of $\delta(j,k)$ reduces linearly with the distance between c and c_k, and is reduced to zero when c_{k-1} or c_{k+1} is reached. As an example, in Figure 12 we demonstrate how we model the impact of an arbitrary increase of $\delta(2,2) = 0.1$ in the $(2,2)$th key mortality rate on other mortality rates that are involved in the synthetic multi-cohort pension plan.

The specification of α is slightly different for the first and last key mortality rates in a key cohort. For $j = 1$, we set α to 1 when $x \le x_{1,k}$, and for $j = n_k$, we set α to 1 when $x > x_{n_k,k}$. Also, for $k = 1$, we set β to 1 when $c \le c_1$, and for $k = m$, we set β to 1 when $c > c_m$. It is easy to see that the way we specify s permits a mortality surface to shift in both parallel and non-parallel fashions[15].

Given $s(x, c, (j,k), \delta(j,k))$, it is straightforward to calculate the key q-duration associated with the (j,k)th key mortality rate. Let $\mathbf{Q}$ and $\tilde{\mathbf{Q}}(\delta(j,k))$ be original mortality surface and the mortality surface affected by $\delta(j,k)$, respectively. Then the key q-duration with respect to the (j,k)th key mortality rate is given by

$$KQD(P(\mathbf{Q}),(j,k)) = \lim_{\delta(j,k)\to 0} \frac{P(\tilde{\mathbf{Q}}(\delta(j,k))) - P(\mathbf{Q})}{\delta(j,k)},$$

where $P(\mathbf{Q})$ is the value of the portfolio on the basis of the mortality surface $\mathbf{Q}$. The set of key q-durations $\{KQD(P(\mathbf{Q}), (j,k)); j = 1, \ldots, n_k; k = 1, \ldots, m\}$ as a whole measures the portfolio's price sensitivity to the entire mortality surface. We can adapt the algorithm in Section 2.3 accordingly to estimate key q-durations in a two-dimensional setting.

Let $F_{j,k}(\mathbf{Q})$ be the (random) present value (per \$1 notional) of the q-forward linked to the (j,k)th key mortality rate. Similar to the one-dimensional setting, we have

$$KQD(F_{j,k}(\mathbf{Q}),(j,k)) = -(1+r)^{-(T_{j,k}-t_0)},$$

[15] Setting $\delta(j,k)$ for all j and k to the same value would imply a parallel shift.

https://doi.org/10.2143/AST.42.2.2182804 Published online by Cambridge University Press

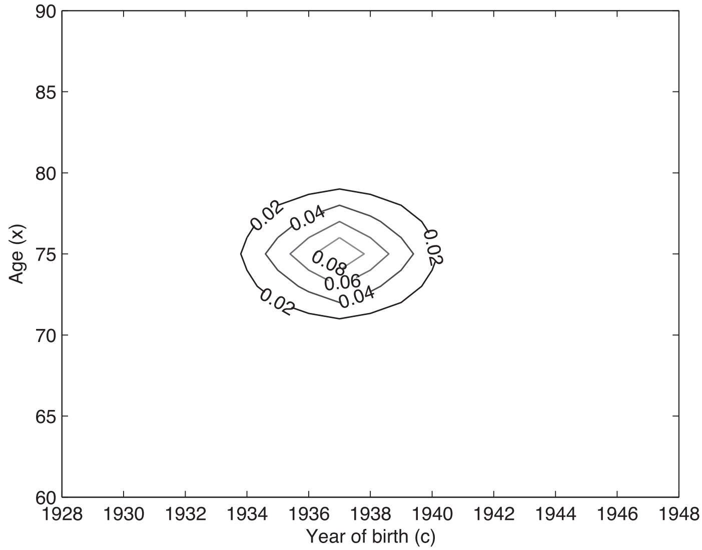

FIGURE 12: The impact of an arbitrary increase of $\delta(2,2) = 0.1$ in the (2, 2)th key mortality rate on other mortality rates that are involved in the synthetic multi-cohort pension plan. The impact is reduced to zero when the next key cohort ($c_3 = 1941$), the previous key cohort ($c_1 = 1933$), the next key age ($x_{3,2} = 80$), or the previous key age ($x_{1,2} = 70$) is reached.

where r is the interest rate at which cash flows are discounted, t_0 is the current date, $T_{j,k}$ is the maturity date of the q-forward[16]. According to the way we specify a key rate shift in a two-dimensional setting, the impact of a shift $\delta(h,i)$ in the (h,i)th key mortality rate becomes zero when the next key age $(x_{h+1,i})$, the previous key age $(x_{h-1,i})$, the next key cohort (c_{i+1}) or the previous key cohort (c_{i-1}) is reached. This implies that $KQD(F_{j,k}(\mathbf{Q}), (h,i)) = 0$ if $h \neq j$ or $i \neq k$. Because of this property, we can determine the notional amounts of the q-forwards independently. The appropriate notional amount of the q-forward linked to the (j,k)th key mortality rate is

$$w(j,k) = \frac{KQD(L(\mathbf{Q}),(j,k))}{KQD(F_{j,k}(\mathbf{Q}),(j,k))}, \tag{4}$$

where $L(\mathbf{Q})$ is the value of the liability being hedged on the current date t_0.

In addition to the aforementioned assumptions, we assume in this illustration that Assumptions 1, 3, 4, 5, 7 and 8 in Section 4.1 still hold. We calculate the required notional amount of each q-forward in the hedge portfolio using equation (4). The details of the q-forwards used are summarized in Table 5.

[16] As we mentioned in Section 3.1, the maturity date $T_{j,k}$ may be slightly later than the reference year $t_{j,k}$, because of the time lag in the availability of the mortality index data.

TABLE 5

THE REFERENCE AGES, REFERENCE YEARS AND NOTIONAL AMOUNTS OF THE Q-FORWARDS IN THE LONGEVITY HEDGE FOR THE SYNTHETIC MULTI-COHORT PENSION PLAN.

j	k	Ref. age ($x_{j,k}$)	Ref. year ($t_{j,k}$)	Notional amt. ($w(j,k)$)
1	1	80	2013	168,950
2	1	85	2018	84,210
1	2	75	2013	236,160
2	2	80	2017	141,660
3	2	85	2022	79,625
1	3	70	2011	331,490
2	3	75	2016	299,540
3	3	80	2021	185,130
4	3	85	2026	102,340
1	4	65	2010	887,660
2	4	70	2015	766,730
3	4	75	2020	543,530
4	4	80	2025	329,080
5	4	85	2030	177,560

As before, we evaluate the effectiveness of the hedge by examining the random present values X and X^*. Figure 13 shows the simulated distributions of X and X^*, which are based on 5,000 realizations of $\mathbf{Q}$ simulated from the CBD model (with parameter uncertainty), fitted to English and Welsh males mortality

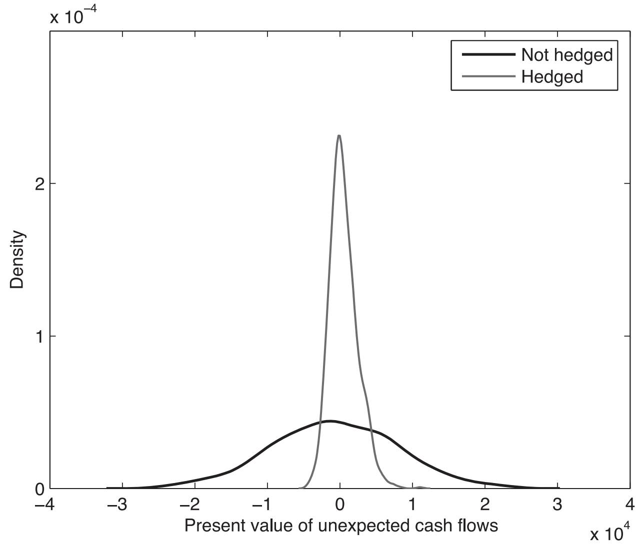

FIGURE 13: Simulated distributions of X and X^* for the multi-cohort synthetic pension plan. The simulations are based on the CBD model with parameter uncertainty.

https://doi.org/10.2143/AST.42.2.2182804 Published online by Cambridge University Press

data from age 60 to 90 and from year 1961 to 2007. The amount of risk reduction, R, provided by the portfolio of 14 mortality forwards is 95.0%. The results indicate that the two-dimensional extension can help us create a highly effective hedge for multi-cohort pension plans with a relatively small number of hedging instruments.

6. Extension 2: Accommodating Population Basis Risk

When a hedger relies on standardized q-forwards to hedge its longevity risk exposure, it is inevitably subject to the risk associated with the difference in the mortality experience between the hedger's population and the population to which the q-forwards are linked. This risk, which is often called population basis risk, is taken into account in the extension that we now present.

For simplicity, the extension below is for a single cohort of individuals, but it can easily be modified for multiple cohorts. Let $\mathbf{q}_1$ and $\mathbf{q}_2$ be the mortality curve for the hedger's population (population 1) and the q-forwards' reference population (population 2), respectively. In general, $\mathbf{q}_1$ and $\mathbf{q}_2$ are different, because of, for example, differing profiles of socioeconomic group, lifestyle and geography.

The existence of population basis risk makes a difference to the calculation of the appropriate notional amounts of the q-forwards. With population basis risk, the appropriate notional amount of the q-forward linked to the jth key mortality rate is given by

$$w(j) = \frac{KQD(V(\mathbf{q}_1), j)}{KQD(F_j(\mathbf{q}_1), j)} = \frac{KQD(V(\mathbf{q}_1), j)}{KQD(F_j(\mathbf{q}_2), j)} \frac{\partial q(x_j, t_j, 1)}{\partial q(x_j, t_j, 2)}, \tag{5}$$

where $q(x_j, t_j, i)$, $i = 1, 2$, is the jth key mortality rates for population i. The calculation of the adjustment factor $\frac{\partial q(x_j, t_j, 1)}{\partial q(x_j, t_j, 2)}$ is not straightforward, and may require a two-population mortality model.

Assume that mortality rates for the two populations follow the augmented common factor model (Li and Lee, 2005), which is defined by

$$\begin{aligned} \ln(m(x,t,i)) &= a(x,i) + B(x)\,K(t) + b(x,i)\,k(t,i) + \varepsilon(x,t,i),\ i = 1,2, \\ K(t) &= \mu + K(t-1) + \xi(t), \\ k(t,i) &= \phi_0(i) + \phi_1(i)\,k(t-1,i) + \zeta(t,i),\ i = 1,2. \end{aligned}$$

where $m(x,t,i)$ is the central death rate (at age x and year t) for population i; $\varepsilon(x,t,i)$ is the error term; μ, $\phi_0(i)$, and $\phi_1(i)$ are constants that do not depend on x and t; and $\xi(t)$ and $\zeta(t,i)$ are innovation terms, i.i.d. normal with zero mean. The model assumes mortality rates for both populations are driven by a common stochastic factor, $K(t)$, which follows a random walk with drift, and a population-specific stochastic factor, $k(t,i)$, which follows a first order auto-regressive model. Parameters $B(x)$ and $b(x,i)$ are the sensitivities to $K(t)$ and $k(t,i)$, respectively.

In a parallel study, Li and Hardy (2011) proved that if the augmented common factor model is assumed, the adjustment factor in equation (5) can be calculated as follows:

$$\frac{\partial q(x,t,1)}{\partial q(x,t,2)} = \frac{q(x,t,1)\,(1+0.5m(x,t,2))^2\,A(x,t,1)}{q(x,t,2)\,(1+0.5m(x,t,1))^2\,A(x,t,2)}, \tag{6}$$

where $A(x,t,i) = B(x)\mu + b(x,i)\,\phi_1(i)^{t-t_0}(\phi_0(i) + (\phi_1(i) - 1)\,k(t_0,i))$ for $i = 1,2$, and t_0 denotes the year when the hedge is set up.

Let us revisit the synthetic pension plan described in Section 4.1. We keep all assumptions stated in Section 4.1 except the assumption that there is no population basis risk. We now allow the hedger's population to be different from the population to which the hedging instruments are linked. Specifically, we assume that the q-forwards in the hedge portfolio are linked to English and Welsh male population, while the hedger's population is either Canadian males, French males or Scottish males.

For each case, we calculate the appropriate notional amounts of the five q-forwards with equations (5) and (6). We then simulate the distributions of the unexpected cash flows, hedged and unhedged, on the basis of the augmented common factor model that is fitted to the corresponding pair of populations. Parameter uncertainty is incorporated into the simulations by the parametric bootstrap, which is detailed in Appendix C. The simulated distributions are displayed in Figure 14. The results indicate that a properly weighted portfolio of q-forwards can reduce a significant amount of risk, even if population basis risk exists.

Also shown in Figure 14 are the corresponding distributions when population basis risk is absent, that is, when the hedge is created with q-forwards that are linked to the hedger's own population. We observe that population basis risk somewhat lowers the effectiveness of a longevity hedge. The impact of population basis risk can also be seen from Table 6, which shows the amounts of risk reduction R for all cases we consider. On average, population basis risk reduces

TABLE 6

THE AMOUNTS OF RISK REDUCTION R WHEN POPULATION BASIS RISK IS PRESENT AND ABSENT. THE VALUES SHOWN ARE BASED ON SIMULATIONS FROM THE COMMON FACTOR MODEL WITH PARAMETER UNCERTAINTY.

Hedger's population	With basis risk	Without basis risk
Scottish males	82.9%	92.7%
French males	90.6%	95.5%
Canadian males	80.3%	95.2%

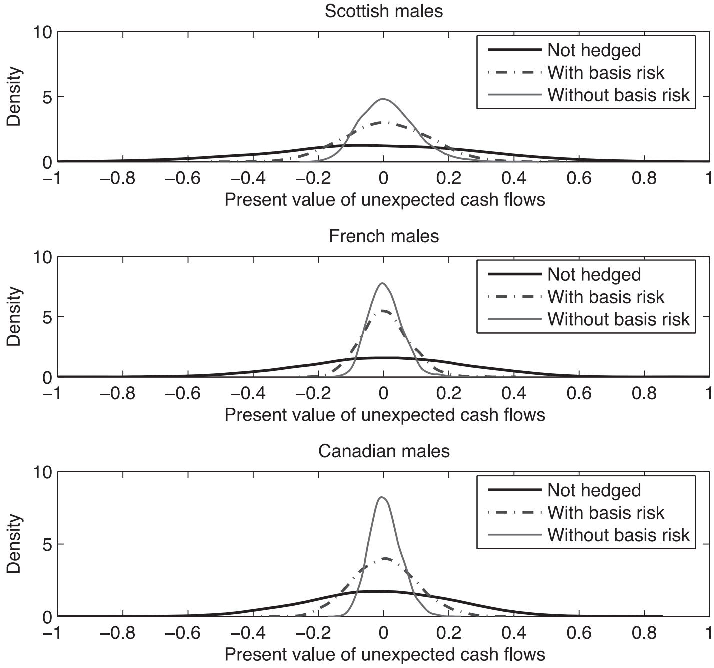

FIGURE 14: Simulated distributions of X and X^*, with and without basis risk. The q-forwards are linked to English and Welsh male population, while the hedger's population is either Scottish males, French males or Canadian males. The simulations are based on the common factor model with parameter uncertainty.

the hedge effectiveness by 10 percentage points. The impact of population basis risk is the highest when the hedger's population is Canadian males, which is possibly the least related to the population on which the q-forwards are based.

7. Concluding Remarks

In this paper, we introduced a measure called key q-duration, which enables us to estimate the price sensitivity of a life-contingent liability to the underlying mortality curve. Given the key q-durations of a portfolio, one can easily hedge the longevity risk associated with the portfolio with only a small number of q-forward contracts. The method we propose is easy to implement, yet a hedge calibrated with our method is almost equally effective as one that is calibrated with a computationally intensive optimization.

In practice, most pension plans involve a spectrum of birth cohorts. To improve applicability, we extended key q-durations to a two-dimensional setting. We illustrated the two-dimensional extension with a synthetic pension plan that involves 21 birth cohorts. The hedge we created could reduce the synthetic plan's longevity risk by 95%, while keeping the number of hedging instruments to a manageable level. A method requiring a modest number of hedging instruments is desirable, because when the market is still in its infancy, transactions are likely to be restricted to a limited number of instruments in which liquidity can be concentrated.

In the absence of any pricing information, it is assumed in our calculations that the forward mortality rates are the same as (not lower than) the corresponding best estimate mortality rates, which equivalently means that the hedge is costless. However, when a hedger is given the relevant forward mortality rates, it can easily incorporate them into the simulations and calculate the cost of the hedge. The hedger can then compare it with the costs associated with other options, including bespoke longevity swaps and buy-ins, that is, the seeking of a financial institution to insure (lock-in) its liabilities. A financially optimal decision can then be made.

When deriving our hedging strategy, it is assumed that the hedger intends to eliminate as much longevity risk as possible. Some entities, however, may only want to transfer a portion of the risk to capital markets. For instance, a life insurer may only want to transfer to capital markets the residual longevity risk that cannot be naturally hedged with its life insurance books. In future research, it would be interesting to adapt the framework of key q-durations so that other hedging objectives, such as hedging 50% of the total risk, can be used.

The calculation of key q-durations requires a central estimate of future mortality and an assumed rate of interest. After the inception of the hedge, both quantities may change as new information is unfolded. This means that key q-durations and hence the optimal hedging strategy may also vary over time. With varying key q-durations, we may be able to achieve a better hedge effectiveness by dynamically adjusting the hedge portfolio. The benefit from dynamic hedging was not studied in this paper, but certainly deserves an investigation when sufficient information about liquidity and transaction costs becomes available.

In the absence of basis risk, the calculation of key q-durations does not require a specific stochastic mortality model. However, when basis risk is present, we need a two-population mortality model to estimate the adjustment term in equation (5). Another avenue for future research to investigate how the adjustment term may change if a different two-population mortality model, for example, the models proposed by Cairns et al. (2011a), is used. It is also warranted to validate the resulting hedging strategy by non-parametric means such as the block bootstrap.

Acknowledgments

This work is supported by a Discovery Grant from the Natural Science and Engineering Research Council of Canada. The authors would like to thank Andrew Cairns and Guy Coughlan for their comments on an earlier draft of this paper.

References

Barbarin, J. (2008) Heath-Jarrow-Morton Modelling of Longevity Bonds and the Risk Minimization of Life Insurance Portfolios. *Insurance: Mathematics and Economics*, **43**, 41-55.

Blake, D., Cairns, A.J.G. and Dowd, K. (2006) Living with Mortality: Longevity Bonds and Other Mortality-Linked Securities. *British Actuarial Journal*, **12**, 153-197.

Brouhns, N., Denuit, M., and Keilegom, I.V. (2005) Bootstrapping the Poisson Log-bilinear Model for Mortality Forecasting. *Scandinavian Actuarial Journal*, **31**, 212-224.

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., and Dowd, K. (2008) Modelling and Management of Mortality Risk: a Review. *Scandinavian Actuarial Journal*, **108**, 79-113.

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.

Cairns, A.J.G. (2011) Modelling and Management of Longevity Risk: Approximations to Survivor Functions and Dynamic Hedging. *Insurance: Mathematics and Economics*, **49**, 438-453.

Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D. and Khalaf-Allah, M. (2011a) Bayesian Stochastic Mortality Modelling for Two Populations. *ASTIN Bulletin*, **41**, 29-55.

Cairns, A.J.G., Dowd, K., Blake, D. and Coughlan, G.D. (2011b) *Longevity Hedge Effectiveness: A Decomposition*. Working paper, Heriot-Watt University.

Coughlan, G. (2009) *Longevity Risk Transfer: Indices and Capital Market Solutions*. In Barrieu, P.M. and Albertini, L. (eds.) The Handbook of Insurance Linked Securities. London: Wiley.

Coughlan, G., Epstein, D., Ong, A., Sinha, A., Hevia-Portocarrero, J., Gingrich, E., Khalaf-Allah, M. and Joseph, P. (2007) LifeMetrics: A Toolkit for Measuring and Managing Longevity and Mortality Risks. Available at http://www.jpmorgan.com/pages/jpmorgan/investbk/solutions/lifemetrics/library.

Coughlan, G.D., Khalaf-Allah, M., Ye, Y., Kumar, S., Cairns, A.J.G., Blake, D. and Dowd, K. (2010) Longevity Hedging: A Framework for Longevity Basis Risk Analysis and Hedge Effectiveness. *North American Actuarial Journal*, **15**, 150-176.

Dahl, M., Melchior, M. and Moller, T. (2008) On Systematic Mortality Risk and Risk-Minimization with Survivor Swaps. *Scandinavian Actuarial Journal*, **2**, 114-146.

Denuit, M., Devolder, P., and Goderniaux, A.-C. (2007) Securitization of Longevity Risk: Pricing Survivor Bonds with Wang Transform in the Lee-Carter Framework. *Journal of Risk and Insurance*, **74**, 87-113.

Dowd, K., Blake, D., Cairns, A.J.G. and Dawson, P. (2006) Survivor swaps. *Journal of Risk and Insurance*, **73**, 1-17.

Dowd, K., Cairns, A.J.G., Blake, D., Coughlan, G.D., Epstein, D. and Khalaf-Allah, M. (2011) A Gravity Model of Mortality Rates for Two Related Populations. *North American Actuarial Journal*, **15**, 334-356.

Dowd, K., Cairns, A.J.G., Blake, D., Coughlan, G.D., Epstein, D. and Khalaf-Allah, M. (2010) Backtesting Stochastic Mortality Models: An Ex-Post Evaluation of Multi-Period-Ahead Density Forecasts. *North American Actuarial Journal*, **14**, 281-298.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. and Tatham, R.L. (2006) *Multivariate Data Analysis*, Sixth Edition. Upper Saddle River, RJ: Prentice Hall.

Ho, T.S.Y. (1992) Key Rate Durations: Measures of Interest Rate Risks. *Journal of Fixed Income*, **2**, 29-44.

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 April 2011).

Jarner, S.F. and Kryger, E.M. (2011) Modelling Adult Mortality in Small Populations: The SAINT Model. *ASTIN Bulletin*, **41**, 377-418.

Johnson, R. and Wichern, D. *Applied Multivariate Statistical Analysis*, 6th Edition. Upper Saddle River, NJ: Prentice Hall.

LEE, R. and CARTER, L. (1992) Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association*, **87**, 659-671.
LI, N. and LEE, R. (2005) Coherent Mortality Forecasts for a Group of Population: An Extension of the Lee-Carter Method. *Demography*, **42**, 575-594.
LI, J.S.-H. and HARDY, M.R. (2011) Measuring Basis Risk in Longevity Hedges. *North American Actuarial Journal*, **15**, 177-200.
LI, J.S.-H., NG, A.C.Y. and CHAN, W.S. (2011) On the Calibration of Mortality Forward Curves. *Journal of Futures Markets*, **31**, 941-970.
PLAT, R. (2010) One-Year Value-at-Risk for Longevity and Mortality. *Insurance: Mathematics and Economics*, **49**, 462-470.
SAUNDERS, D.R. (1962) Trans-varimax: Some Properties of the Ratiomax and Equamax Criteria for Blind Orthogonal Rotation. Paper presented at the meeting of the American Psychological Association, St. Louis.
STEVENS, R., DE WAEGENAERE, A. and MELENBERG, B. (2011) Longevity Risk and Natural Hedge Potential in Portfolios of Life Insurance Products: the Effect of Investment Risk. CentER Discussion Paper No. 2011-036.
WILLS, S. and SHERRIS, M. (2008) Integrating Financial and Demographic Longevity Risk Models: An Australian Model for Financial Applications. UNSW Australian School of Business Research Paper No. 2008ACTL05.
WILLS, S. and SHERRIS, M. (2010) Securitization, Structuring and Pricing of Longevity Risk. *Insurance: Mathematics and Economics*, **46**, 173-185.
ZHOU, R. and LI, J.S.-H. (2010) A Cautionary Note on Pricing Longevity Index Swaps. *Scandinavian Actuarial Journal*, first published on: 04 September 2010 (iFirst).
ZHOU, R., LI, J.S.-H. and TAN, K.S. (2011) Pricing Standardized Mortality Securitizations: A Two-Population Model with Transitory Jump Effects. Paper presented in the Seventh International Longevity Risk and Capital Markets Solutions Conference, Frankfurt, Germany.

APPENDIX A

A Factor Analysis of Historical Annual Mortality Reductions

The objective of this appendix is to identify a limited number of latent factors that represent the age-specific mortality reductions. We achieve this goal by using a factor analysis, a common technique in multivariate statistics. For brevity, we will restrict our discussion to points necessary for describing the application of a factor analysis to the data in consideration. We refer readers to Hair et al. (2006) and Johnson and Wichern (2007) for fuller details about the subject of factor analysis.

We perform the factor analysis with mortality data from English and Welsh males population for years 1961 to 2007 and for ages 60 to 90. We first graduate the crude central death rates to remove sampling fluctuations. Assuming uniform distribution of deaths over each year of age, we calculate death probabilities, which are then used to compute the realized annual mortality reduction,

$$RF(x,t) = 1 - \frac{q(x, t+1)}{q(x,t)},$$

for $x = 60, 61, \ldots, 90$ and $t = 1961, 1962, \ldots, 2006$. We can regard the values of $RF(x,t)$ for $t = 1961, 1962, \ldots, 2006$ as realizations of the random variable $RF(x)$, the random annual mortality reduction at age x.

For notational convenience, we let $r_1 = RF(60)$, $r_2 = RF(61)$, ..., $r_{31} = RF(90)$, and let $\mathbf{r} = (r_1, \ldots, r_{31})'$ be the vector of random annual mortality reductions. The factor analysis allows us to identify k latent factors, where $k < 31$, that represent the 31 elements in $\mathbf{r}$. The analysis is based on the factor model, which can be expressed as

$$\mathbf{r} = \mathbf{p} + \mathbf{Lf} + \varepsilon, \tag{7}$$

where $\mathbf{p} = (p_1, \ldots, p_{31})'$ is a vector of constants; $\mathbf{f}$ is a $k \times 1$ random vector, with elements $f_1, \ldots, f_k$, which are called the common factors; $\mathbf{L}$ is a $31 \times k$ matrix of unknown constants, called factor loadings; and the elements, $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{31}$, of the 31×1 random vector ε are called the specific factors. The following assumptions are made on the random quantities in the model:

1. The expectations of ε_i, $i = 1, \ldots, 31$, and f_j, $j = 1, \ldots, k$, are all zero.
2. The random vectors $\mathbf{f}$ and ε are uncorrelated.
3. The variance-covariance matrix for the random vector $\mathbf{f}$ is an identity matrix of order k.
4. $\text{Var}(\varepsilon_i) = \psi_i$, $i = 1, \ldots, 31$; $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$ for $i \neq j$.

The factor model implies that the random mortality reduction at each age can be viewed as a linear combination of all common factors and one specific factor. Specifically, we have

$$r_i = c_i + l_{i,1} f_1 + \ldots + l_{i,k} f_k + \varepsilon_i, \quad i = 1, \ldots, 31,$$

where $l_{i,j}$, the (i, j)th element of $\mathbf{L}$, is the factor loading of r_i on the jth common factor f_j. The following two properties are important to our analysis:

- $\text{corr}(r_i, f_j) = \dfrac{l_{i,j}}{\sqrt{\text{Var}(r_i)}}$

  This property arises from Assumption 3. It follows that for a given i (i.e., age), the reduction in mortality is the most related to the common factor with the largest factor loading.

- $\text{Var}(r_i) = l_{i,1}^2 + l_{i,2}^2 + \ldots + l_{i,k}^2 + \psi_i$

  This property arises from Assumptions 2 to 4. It follows that for a given i (i.e., age), the common factor with the largest factor loading offers the largest explanation to the variation in the annual mortality reduction for that age.

Note that the factor loadings are not unique. In particular, if $\mathbf{L}$ satisfies equation (7), then $\mathbf{L}^* = \mathbf{L\Gamma}$, where $\mathbf{\Gamma}$ is an orthogonal matrix, will also satisfy equation (7). When the first estimate of the factor loadings are not readily interpretable, it is customary to transform these loadings by post multiplication using an orthogonal matrix so that a meaningful interpretation is possible. This process is referred to as factor rotation.

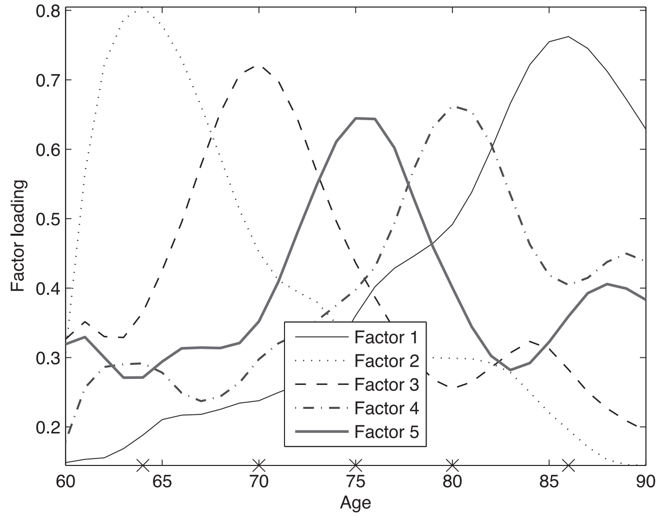

FIGURE 15: Estimated factor loadings for the five latent factors at different ages. For each latent factor, the age at which the factor loading is maximum is marked by a cross on the horizontal axis.

We consider a factor model with $k = 5$ latent factors. The factor model is estimated by the method of maximum likelihood. The first maximum likelihood estimates are not readily interpretable, so a factor rotation is required. We use the equamax rotation proposed by Saunders (1962)[17].

In Figure 15 we display the estimated (rotated) factor loadings for the five latent factors at different ages. We have the following observations and conclusions:

1. From age 60 to 67, the factor loadings for Factor 2 are the highest among all latent factors. By using the two statistical properties discussed above, we can conclude that mortality reductions at these ages are the most associated with Factor 2, and that Factor 2 can be interpreted as a latent factor representing this group of ages.
2. Similarly, Factor 1 represents ages 83 to 90, Factor 3 represents ages 68 to 73, Factor 4 represents ages 77 to 82, and Factor 5 represents ages 74 to 76. The five latent factors represent age groups that divide the mortality curve into portions of approximately the same length.
3. If we were to choose five key ages (key mortality rates), it would be sensible to choose ages that correspond to the latent factors or equivalently the five consecutive age groups.

[17] The equamax rotation determines the orthogonal matrix $\mathbf{\Gamma}$ such that

$$\frac{1}{p}\sum_{j=1}^{k}\left(\sum_{i=1}^{31} l_{i,j}^{*4} - \frac{k}{2p}\left(\sum_{i=1}^{31} l_{i,j}^{*2}\right)^2\right),$$

where $l_{i,j}^{*}$ is the (i,j)th element in $\mathbf{L}^* = \mathbf{L\Gamma}$, is maximized.

https://doi.org/10.2143/AST.42.2.2182804 Published online by Cambridge University Press

4. The ages at which the factor loadings are highest are marked by crosses on the horizontal axis. The factor loading for Factor 5, for example, is the maximum at age 75, which means that the annual mortality reduction at age 75 is the most correlated with this latent factor. Hence, considering only the statistical properties of the mortality data, it makes sense to choose these ages at the key ages. The key ages (65, 70, 75, 80 and 85) used in our illustrative hedge are very close to these ages.

Appendix B

Stochastic Mortality Models Used for Evaluating Hedge Effectiveness

The following three stochastic mortality models are used in Section 4.2 for evaluating hedge effectiveness.

- The original Cairns-Blake-Dowd (CBD) model

  The CBD model can be expressed as

$$\ln\left(\frac{q(x,t)}{1-q(x,t)}\right) = \kappa_t^{(1)} + \kappa_t^{(2)}(x-\bar{x}), \tag{8}$$

  where $\bar{x}$ is the average age over the age range we consider, and $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ are period effect indexes. In particular, we may regard $\kappa_t^{(1)}$ as the overall mortality level at time t and $\kappa_t^{(2)}$ as the slope of the mortality curve (in logit scale) at time t. This model has no identifiability problems, and therefore parameter constraints are not required.

  The period effect indexes $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ are modeled by a bivariate random walk with drift, that is,

$$\kappa_t = \kappa_{t-1} + \mu + CZ(t), \tag{9}$$

  where $\kappa_t = (\kappa_t^{(1)}, \kappa_t^{(2)})'$, $\mu = (\mu_1, \mu_2)'$ is a constant 2×1 vector, C is a constant 2×2 upper triangular matrix, and $\{Z(t)\}$ is a sequence of i.i.d. 2-dimensional standard normal random vectors.

- A generalized Cairns-Blake-Dowd (G-CBD) model with a quadratic age effect term and a cohort effect term

  The G-CBD model can be expressed as

$$\ln\left(\frac{q(x,t)}{1-q(x,t)}\right) = \kappa_t^{(1)} + \kappa_t^{(2)}(x-\bar{x}) + \kappa_t^{(3)}\left((x-\bar{x})^2 - \hat{\sigma}_x^2\right) + \gamma_c^{(4)}, \tag{10}$$

  where $\kappa_t^{(1)}$, $\kappa_t^{(2)}$, and $\kappa_t^{(3)}$ are period effect indexes, $c = t - x$ denotes the year of birth, $\gamma_c^{(4)}$ is a cohort effect index, and $\hat{\sigma}_x^2$ is the mean of $(x - \bar{x})^2$ over the age range we consider.

https://doi.org/10.2143/AST.42.2.2182804 Published online by Cambridge University Press

This G-CBD model differs from the original CBD model in two ways. First, it contains a cohort risk factor $\gamma_c^{(4)}$ that is explicitly linked to the year of birth c. Second, it includes a quadratic term $\kappa_t^{(3)}((x-\bar{x})^2-\hat{\sigma}_x^2)$ to capture the potential curvature in the relationship between $\ln\left(\frac{q(x,t)}{1-q(x,t)}\right)$ and x.

This model has an identifiability problem. To stipulate parameter uniqueness, we use the constraints $\sum_{x,t}\gamma_{t-x}^{(4)}=0$, $\sum_{x,t}(t-x)\gamma_{t-x}^{(4)}=0$, and $\sum_{x,t}(t-x)^2\gamma_{t-x}^{(4)}=0$. The summations are taken over the entire sample age range and sample period.

The period effect indexes are modeled by a trivariate random walk with drift:

$$\kappa_t = \kappa_{t-1} + \mu + CZ(t), \tag{11}$$

where $\kappa_t = (\kappa_t^{(1)}, \kappa_t^{(2)}, \kappa_t^{(3)})'$, $\mu = (\mu_1, \mu_2, \mu_3)'$ is a constant 3×1 vector, C is a constant 3×3 upper triangular matrix, and $\{Z(t)\}$ is a sequence of i.i.d. 3-dimensional standard normal random vectors. The cohort effect index is modeled by a second order autoregressive process:

$$\gamma_c^{(4)} = \phi_0 + \phi_1\gamma_{c-1}^{(4)} + \phi_2\gamma_{c-2}^{(4)} + a_c,$$

where $\{a_c\}$ is a sequence of i.i.d. standard normal random variables, and ϕ_0, ϕ_1 and ϕ_2 are constants.

- The Lee-Carter (LC) model

The LC model can be expressed as

$$\ln(m(x,t)) = \alpha_x + \beta_x\kappa_t, \tag{12}$$

where $m(x,t)$ is the central death rate at age x in year t, α_x indicates the average level of mortality at age x, κ_t is a period effect index, β_x indicates the the sensitivity of $\ln(m(x,t))$ to changes in κ_t at age x. Note that the LC model is based on central death rates. To compute death probabilities, we can use the relation

$$q(x,t) = \frac{m(x,t)}{1+0.5m(x,t)}, \tag{13}$$

which results from the assumption that deaths are uniformly distributed over each year of age.

This model has an identifiability problem. We use the constraints $\sum_t\kappa_t=0$ and $\sum_x\beta_x=1$ to stipulate parameter uniqueness. The summations are taken over the entire sample period and sample age range, respectively.

The period effect index κ_t is modeled by a random walk with drift:

$$\kappa_t = \kappa_{t-1} + \mu + a_t, \tag{14}$$

https://doi.org/10.2143/AST.42.2.2182804 Published online by Cambridge University Press

where μ is a constant and $\{a_t\}$ is a sequence of i.i.d. standard normal random variables.

All mortality models in this paper are fitted by the method of maximum likelihood. Let us define $D(x,t)$ by the number of deaths at age x and in year t, and $E(x,t)$ by the corresponding exposures to the risk of death. In constructing the likelihood function, we treat $D(x,t)$ as independent Poisson responses, that is,

$$D(x,t) \sim \text{Poisson}(\hat{D}(x,t)),$$

where $\hat{D}(x,t) = E(x,t)m(x,t)$ is the expected number of deaths at age x and in year t. This gives the following log-likelihood, which is applicable to all models considered:

$$l = \sum_{x,t}(D(x,t)\ln(\hat{D}(x,t)) - \hat{D}(x,t) - \ln(D(x,t)!)), \tag{15}$$

where $D(x,t)!$ stands for $D(x,t)$ factorial. The summation is taken over all x in the sample age range and all t in the sample period. For models that are based on $m(x,t)$, the likelihood function can be obtained by substituting the model equation directly into equation (15). For models that are based on $q(x,t)$, the likelihood function can be obtained by substituting the model equation into equation (13) and then into equation (15).

The maximization of the likelihood function can be accomplished by an iterative Newton-Raphson method, in which parameters are updated one at a time. The updating of a typical parameter θ proceeds according to

$$u(\theta) = \theta - \frac{\partial l/\partial\theta}{\partial^2 l/\partial\theta^2},$$

where $u(\theta)$ is the updated value of θ in the iteration. The parameter constraints (if any) are applied at the end of each iteration.

Appendix C

Inclusion of Parameter Uncertainty

To incorporate parameter uncertainty into the simulations of future mortality, we use the parametric bootstrap (Brouhns et al., 2005), in which distributions of model parameters are obtained by repeated estimations from pseudo samples. In this appendix, we describe how the parametric bootstrap is implemented.

The description below is based on the Lee-Carter model, but it can be adapted easily to other stochastic mortality models by modifying a few steps accordingly.

1. For each x in the sample age range and each t in the sample period, simulate a realization of $D(x,t)$ from Poisson $(\hat{D}(x,t))$. This gives a pseudo sample of death counts.
2. On the basis of the pseudo sample, re-estimate parameters α_x, β_x and κ_t in equation (12) by maximizing the log-likelihood, which is given in equation (15).
3. Reestimate the parameters in the random walk (equation (14)) on the basis of the reestimated κ_t's.
4. Simulate future values of κ_t's using the reestimated random walk.
5. Calculate future values of $m(x,t)$, using the reestimated α_x's and β_x's (Step 2) and the simulated future values of κ_t (Step 4).
6. Perform the steps above 5,000 times to obtain 5,000 simulated mortality scenarios.

The algorithm above incorporates both stochastic uncertainty and parameter uncertainty. In particular, stochastic uncertainty is taken into account in Step 4, while parameter uncertainty is taken into account in Steps 2 and 3.

JOHNNY SIU-HANG LI (corresponding author)
Department of Statistics and Actuarial Science
University of Waterloo
Waterloo, Ontario, Canada, N2L3G1
Email: shli@uwaterloo.ca

ANCHENG LUO
Department of Statistics and Actuarial Science
University of Waterloo
Waterloo, Ontario, Canada, N2L3G1
Email: a2luo@uwaterloo.ca