Annals of Actuarial Science, Vol. 13, part 1, pp. 166–197. © Institute and Faculty of Actuaries 2018
doi:10.1017/S1748499518000179
First published online 4 July 2018

Assessing basis risk in index-based longevity swap transactions

Jackie Li*
Actuarial Studies and Business Analytics, Macquarie University, Sydney, NSW 2109, Australia

Johnny Siu-Hang Li
Statistics and Actuarial Science, University of Waterloo, 200 University Avenue West Waterloo, ON, Canada N2L 3G1

Chong It Tan
Actuarial Studies and Business Analytics, Macquarie University, Sydney, NSW 2109, Australia

Leonie Tickle
Actuarial Studies and Business Analytics, Macquarie University, Sydney, NSW 2109, Australia

Abstract

In this paper, we carry out an investigation on modelling basis risk and measuring risk reduction in a longevity hedge constructed by index-based longevity swaps. We derive the fitting procedures of the M7-M5 and common age effect+Cohorts models and define the level of longevity risk reduction. Based on a wide range of hedging scenarios of pension plans, we find that the risk reduction levels are often around 50%–80% for a large plan, while the risk reduction estimates are usually smaller than 50% for a small plan. Moreover, index-based hedging looks more effective under a more precise hedging scheme. We also perform a detailed sensitivity analysis on the hedging results. The most important modelling features are the behaviour of simulated future variability, portfolio size, speed of reaching coherence, data size and characteristics, simulation method, and mortality structural changes.

Keywords

Longevity basis risk; Two-population mortality projection model; Index-based longevity hedging; Longevity swap; Hedge effectiveness

1. Introduction

It is a global phenomenon that life expectancy continues to rise. This trend has been persistent and driven by significant progresses in hygiene, nutrition, lifestyle, medical knowledge, and health care. Although it is definitely no mean feat for humans, it presents a significant challenge to pension plan sponsors and annuity providers. There is the so-called longevity risk that pension plans or annuity portfolios pay more than expected because of unanticipated improvements in mortality. Roughly speaking, this risk has two components, systematic longevity risk and non-systematic longevity risk, in which only the latter can be mitigated by increasing the plan or portfolio size.

*Correspondence to: Jackie Li, Actuarial Studies and Business Analytics, Macquarie University, Sydney, NSW 2109, Australia. Tel: +612 98508576. E-mail: jackie.li@mq.edu.au

There are generally three broad approaches for financial institutions to tackle longevity risk (Cairns *et al*., 2008). The first is the use of insurance and reinsurance, in which the unwanted risk is passed on to an insurer or reinsurer after paying a premium. The second is natural hedging (Li & Haberman, 2015), which is basically a diversification strategy and exploits the opposite changes in the values of annuities and life insurances. The third approach, capital market solutions, has attracted much attention in recent years. These solutions include insurance securitisation, mortality- or longevity-linked securities, and derivatives. Insurance securitisation involves securitising a class of business as a complex package and then selling the highly structured securities to capital market investors. Moreover, some currently popular bespoke de-risking methods like buy-ins, buy-outs, and longevity swaps are also customised transactions for hedging specific portfolios. On the other hand, standardised mortality- or longevity-linked securities and derivatives have their cash flows linked to an index or *reference* population, but not the (*book*) population underlying the portfolio being hedged. There would then be a potential mismatch between the hedging instrument and the portfolio, due to certain demographic differences (e.g. age profile, sex, socioeconomic status). There are also two other important issues. First, the experience of a small portfolio tends to be more volatile and is more likely to deviate further from that of the reference population. Second, the payoff structures of the hedging tool and the portfolio would usually be quite different. These three sources of discrepancies give rise to *longevity basis risk*, the assessment of which is under much research in the latest actuarial literature.

While most longevity transactions thus far have been customised, index-based solutions and standardised products could potentially draw more interest from financial institutions both inside and outside the insurance sector. They have substantial potential to allow effective risk management at lower costs and provide significant capital savings. As noted above, however, the potential mismatch between the hedging instrument and the pension or annuity portfolio to be hedged leads to longevity basis risk, which is composed of *demographic* basis risk (demographic or socioeconomic differences), *sampling* basis risk (random outcomes of individual lives), and *structural* basis risk (differences in payoff structures). For small- to medium-sized pension plans or annuity portfolios, the mortality experience would be more volatile, which means that sampling basis risk and structural basis risk would have more impact and hence need to be properly allowed for. Haberman *et al.* (2014) and Villegas *et al.* (2017) proposed a decision tree framework as a practical guide for choosing a two-population mortality projection model for assessing demographic basis risk. The major options are namely the M7-M5 model, the CAE+Cohorts (common age effect + cohort effect) model, and the characterisation approach. In this paper, we put the earlier work in Haberman *et al.* (2014) into practice and measure longevity basis risk under various practical conditions. In particular, we take all the three basis risk components into account and investigate a variety of hedging scenarios using UK population and industry data sets.

The structure of this paper is as follows. Section 2 analyses the historical mortality patterns and improvements as reflected in three UK data sets. Section 3 provides the details of the M7-M5 and CAE +Cohorts models we have used. Section 4 explains the simulation technique and the calibration method adopted. Section 5 examines a range of hedging scenarios of pension portfolios using index-based longevity swaps, and calculates the corresponding levels of longevity risk reduction. Section 6 performs an extensive sensitivity analysis on the hedging results by making a series of changes to the original model settings and assumptions, and also the time series processes. Lastly, Section 7 concludes the paper.

2. Data Features

The UK mortality data used in this paper are collected from the Continuous Mortality Investigation (CMI; Self-Administered Pension Scheme Mortality Investigation), Office for National Statistics

(ONS; by deprivation subgroups), and Human Mortality Database (HMD 2017; England and Wales). The first two data sets are taken as the underlying experience of the pension or annuity portfolio being hedged. The last data set is assumed to represent the experience of the reference population of the hedging instrument. The retirement age range of 60–89 is covered in the analysis, while the data of ages 90+ are scarce and so are excluded.

Figure 1 plots the logit mortality rates of different CMI groups of male pensioners under normal retirement at ages 60–69, 70–79, and 80–89. The solid line corresponds to the pension range of £1–£8,500 p.a., the dashed line refers to £8,500+ p.a., and the dotted lines depict their potential underlying linear trends over time. It can be seen that the higher pension groups have lower mortality and more volatile experience in general, but the differences between the two pension groups tend to reduce over age. Moreover, the differences in mortality levels between industries appear quite random and are rather small, especially at older ages. Table 1 states the average mortality rate and the average annual rate of improvement in mortality rate for each group during the period. The higher pension groups generally have lower mortality levels, and the differences in average mortality rates between industries look quite small, though the pensioners in financials seem to have slightly lower mortality levels than the others. Moreover, the improvement rates range largely from around 1%–5% p.a. and have a tendency to reduce over age. The lower pension groups, which have higher mortality rates, seem to experience greater improvements in mortality for many of the cases. The differences in improvement rates between industries are mostly randomly scattered, except that the pensioners in local authority appear to enjoy greater mortality improvements than the rest.

Figure 2 shows the logit mortality rates of England IMD (index of multiple deprivation) quintile groups of males for the same age groups as above. The five different lines with progressively lighter shades represent the most deprived areas to the least consecutively, in which the IMD is an overall relative measure of deprivation, based on a number of socioeconomic factors such as income, employment, and education. It is very clear that mortality increases with the deprivation level for all age groups, but the differences between the quintile groups become smaller at older ages. The declining temporal trends are all fairly steady, as the data exposures are very large. Table 2 lists the average mortality rate and the average annual rate of improvement in mortality rate for each group during the period. Less deprived areas have lower mortality levels and also higher improvement rates than more deprived areas. The improvement rates range from 1.3% to 3.6% p.a. and are highest at ages 70–79.

Figure 3 displays the logit mortality rates of the English and Welsh male population from 2000 to 2014. The decreasing mortality trends over time are stable for all age groups with little variability, due to the very large population exposures. Table 3 gives the average mortality rate and the average annual rate of mortality improvement for each age group. The improvement rates range from 2.3% to 3.5% p.a. They are lowest at ages 80–89, and like the IMD quintile groups, those lives aged 70–79 experience the highest improvement rates.

Between the CMI pensioners (Table 1) and IMD quintile groups (Table 2), the mortality levels of the lower pension groups are comparable to those of the third most deprived areas, whereas the higher pension groups have slightly lower mortality than the least deprived areas. One interesting observation is that while the lower pension (less wealthy) groups seem to have greater mortality improvements in many cases, the less deprived (wealthier) areas clearly have higher improvement rates. This conflicting difference may be explained by two reasons. First, the former uses only the pension amount in dividing the groups, but the latter adopts the IMD based on a combination of

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

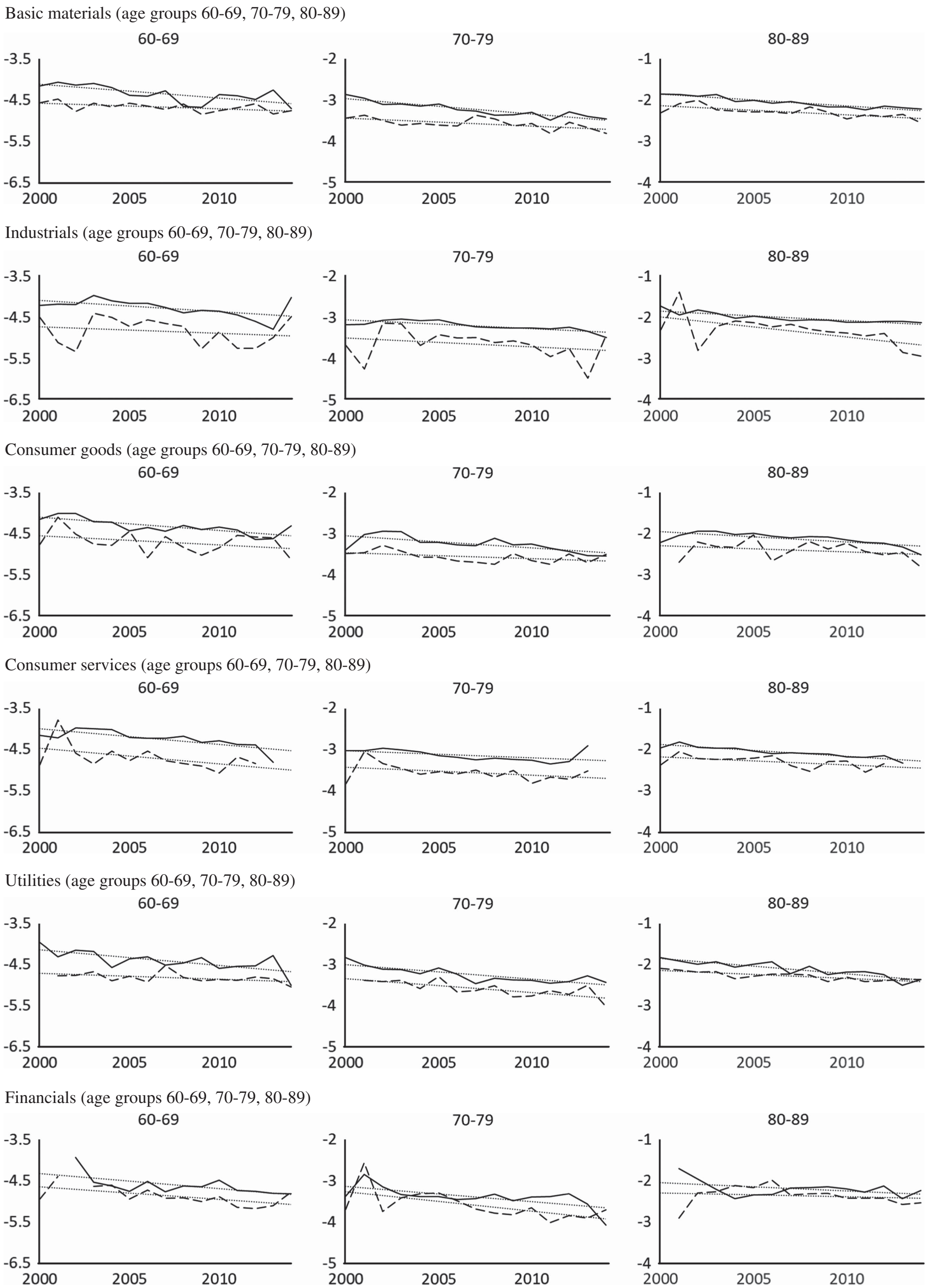

Figure 1. Logit mortality rates of Continuous Mortality Investigation male pensioners from 2000 to 2014. The solid line corresponds to the pension range of £1–£8,500 p.a., the dashed line refers to £8,500+ p.a. (for technology, the split is by £4,500 p.a. instead), and the dotted lines depict their potential underlying linear trends over time.

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

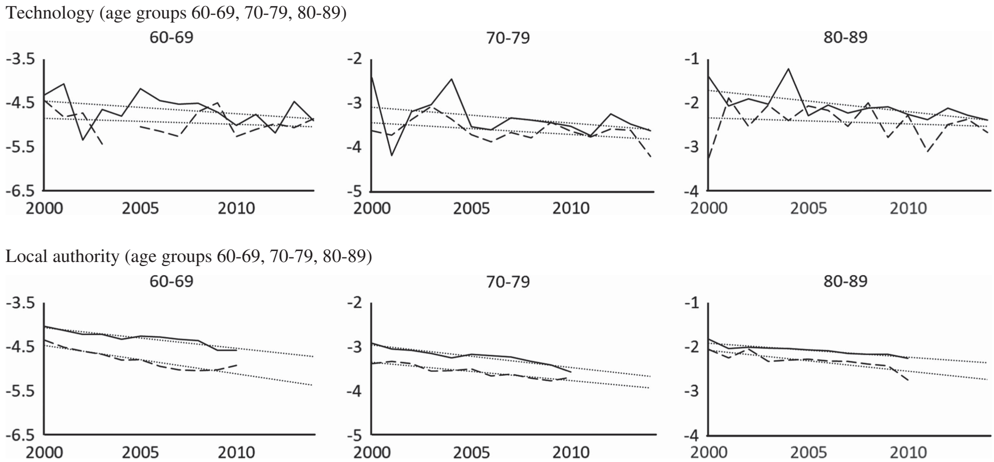

Figure 1. Continued.

seven domains, so the two sets of results may correspond to individuals with different demographic features. Second, the data exposures of each CMI industry are much smaller than those of the IMD quintile groups, and the experience is unavoidably a lot more volatile. Between the CMI pensioners and English and Welsh population (Table 3), the lower pension groups have mortality levels and improvement rates roughly comparable to those of the English and Welsh population.

3. Model Descriptions

In this paper, we apply the M7-M5 model and the CAE+Cohorts model in Haberman *et al.* (2014) to the three data sets. The M7-M5 model is a two-population extension of the Cairns–Blake–Dowd model (Cairns *et al.*, 2006, 2009) and has two major components:

$$\text{logit}\, q_{x,t}^{R} = \kappa_{t,1}^{R} + (x-\bar{x})\kappa_{t,2}^{R} + \left((x-\bar{x})^{2}-\sigma^{2}\right)\kappa_{t,3}^{R} + \gamma_{t-x}^{R} = \eta_{x,t}^{R} \qquad \text{(reference component)}$$

$$\text{logit}\, q_{x,t}^{B} - \text{logit}\, q_{x,t}^{R} = \kappa_{t,1}^{B} + (x-\bar{x})\kappa_{t,2}^{B} = \eta_{x,t}^{B} \qquad \text{(book component)}$$

In the reference component, $q_{x,t}^{R}$ is the mortality rate of the reference population at age x in year t, $\kappa_{t,1}^{R}$, $\kappa_{t,2}^{R}$, and $\kappa_{t,3}^{R}$ refer to the level, slope, and curvature, respectively, of the mortality curve across age in year t, and γ_{t-x}^{R} describes the cohort effect of those lives born in year t–x. In the book component, the difference in the logit mortality rate between the book and reference populations in year t, $\text{logit}\, q_{x,t}^{B} - \text{logit}\, q_{x,t}^{R}$, is captured by another two parameters $\kappa_{t,1}^{B}$ and $\kappa_{t,2}^{B}$. The two terms $\bar{x} = \frac{1}{\text{no. of ages}}\sum_{x} x$ and $\sigma^{2} = \frac{1}{\text{no. of ages}}\sum_{x}(x-\bar{x})^{2}$, and $\eta_{x,t}^{R}$ and $\eta_{x,t}^{B}$ denote the overall reference and book components. The mortality rates can then be expressed as $q_{x,t}^{R} = \frac{\exp\left(\eta_{x,t}^{R}\right)}{1+\exp\left(\eta_{x,t}^{R}\right)} = \frac{1}{\exp\left(-\eta_{x,t}^{R}\right)+1}$ and $q_{x,t}^{B} = \frac{\exp\left(\eta_{x,t}^{B}+\eta_{x,t}^{R}\right)}{1+\exp\left(\eta_{x,t}^{B}+\eta_{x,t}^{R}\right)} = \frac{1}{\exp\left(-\eta_{x,t}^{B}-\eta_{x,t}^{R}\right)+1}$. Moreover, there are three identifiability constraints $\sum_{c}\gamma_{c}^{R} = 0$, $\sum_{c} c\gamma_{c}^{R} = 0$, and $\sum_{c} c^{2}\gamma_{c}^{R} = 0$, which ensure that the cohort parameters have a unique set of solutions and fluctuate around zero over successive cohorts.

While Haberman *et al.* (2014) used the binomial distribution assumption, we model the random numbers of deaths at age x in year t of the reference and book populations with the Poisson

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

Table 1. Average mortality levels and improvements of Continuous Mortality Investigation male pensioners from 2000 to 2014.

	Lower pension group		Higher pension group	
Industry\ages	Average level	Improvement (p.a. (%))	Average level	Improvement (p.a. (%))
Basic materials				
60–69	0.0129	3.3	0.0093	1.3
70–79	0.0385	3.5	0.0275	1.8
80–89	0.1136	2.5	0.0921	2.0
Industrials				
60–69	0.0136	2.7	0.0080	1.6
70–79	0.0380	2.1	0.0259	2.1
80–89	0.1155	2.1	0.0905	4.3
Consumer goods				
60–69	0.0133	3.2	0.0093	2.2
70–79	0.0378	2.8	0.0278	1.4
80–89	0.1076	2.2	0.0843	1.3
Consumer services				
60–69	0.0145	3.8	0.0096	3.8
70–79	0.0422	1.6	0.0285	1.8
80–89	0.1139	2.5	0.0929	1.7
Utilities				
60–69	0.0124	3.7	0.0081	1.4
70–79	0.0380	3.3	0.0273	3.3
80–89	0.1096	3.3	0.0936	1.7
Financials				
60–69	0.0102	3.6	0.0078	3.1
70–79	0.0332	3.6	0.0283	4.6
80–89	0.1010	1.7	0.0871	0.8
Technology				
60–69	0.0100	2.8	0.0074	1.3
70–79	0.0374	3.4	0.0267	2.5
80–89	0.1185	4.1	0.0854	1.2
Local authority				
60–69	0.0135	4.6	0.0085	6.4
70–79	0.0391	4.8	0.0280	4.0
80–89	0.1117	2.7	0.0906	4.3

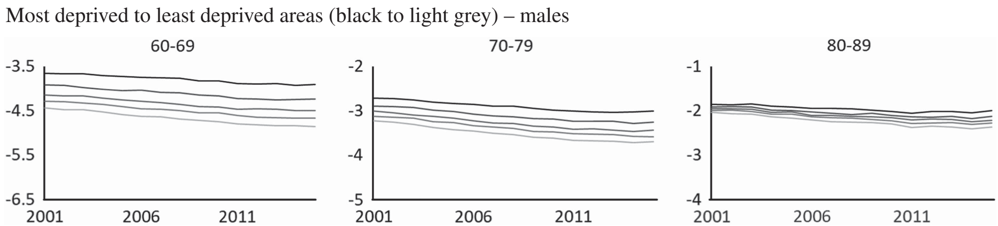

Figure 2. Logit mortality rates of England index of multiple deprivation quintile groups from 2001 to 2015 for age groups 60–69, 70–79, and 80–89.

distribution instead, i.e., $D^R_{x,t} \sim \text{Poisson}\left(e^R_{x,t} m^R_{x,t}\right)$ and $D^B_{x,t} \sim \text{Poisson}\left(e^B_{x,t} m^B_{x,t}\right)$, which are in agreement with most of the current literature. The terms $e^R_{x,t}$ and $e^B_{x,t}$ are the corresponding central exposed to risk measures. We further assume that the force of mortality is constant within each

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

Table 2. Average mortality levels and improvements of England index of multiple deprivation quintile groups from 2001 to 2015.

	Males	
Group\ages	Average level	Improvement (p.a. (%))
Most deprived areas		
60–69	0.0221	2.1
70–79	0.0522	2.4
80–89	0.1237	1.3
Second most deprived areas		
60–69	0.0162	2.6
70–79	0.0428	2.9
80–89	0.1139	1.7
Third most deprived areas		
60–69	0.0129	2.8
70–79	0.0367	3.3
80–89	0.1082	1.9
Fourth most deprived areas		
60–69	0.0111	3.0
70–79	0.0331	3.4
80–89	0.1032	2.2
Least deprived areas		
60–69	0.0093	3.2
70–79	0.0291	3.6
80–89	0.0963	2.4

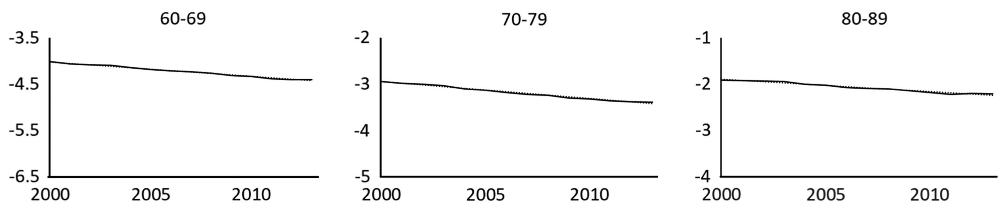

Figure 3. Logit mortality rates of English and Welsh male population from 2000 to 2014 for age groups 60–69, 70–79, and 80–89.

Table 3. Average mortality levels and improvements of English and Welsh male population from 2000 to 2014.

Ages	Average level	Improvement (p.a. (%))
60–69	0.0145	3.1
70–79	0.0398	3.5
80–89	0.1119	2.3

age-time cell and so it is equal to the central death rate. Then the central death rates of the reference and book populations can be expressed as $m_{x,t}^{R} = -\ln\left(1 - q_{x,t}^{R}\right) = \ln\left(1 + \exp\left(\eta_{x,t}^{R}\right)\right)$ and $m_{x,t}^{B} = -\ln\left(1 - q_{x,t}^{B}\right) = \ln\left(1 + \exp\left(\eta_{x,t}^{B} + \eta_{x,t}^{R}\right)\right)$. Under the Poisson assumption, the log likelihood

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

functions of the two populations are:

$$l^R = \sum_{x,t}\left(d^R_{x,t}\ \ln e^R_{x,t} + d^R_{x,t}\ \ln m^R_{x,t} - e^R_{x,t} m^R_{x,t} - \ln\left(d^R_{x,t}\,!\right)\right)$$
$$l^B = \sum_{x,t}\left(d^B_{x,t}\ln e^B_{x,t} + d^B_{x,t}\ln m^B_{x,t} - e^B_{x,t} m^B_{x,t} - \ln\left(d^B_{x,t}\,!\right)\right)$$

in which $d^R_{x,t}$ and $d^B_{x,t}$ are the observed numbers of deaths. The iterative updating schemes for estimating the model parameters are given in the Appendix.

For projecting and simulating future mortality rates, we follow Haberman *et al.* (2014) and model the time series of $\kappa^R_{t,1}$, $\kappa^R_{t,2}$, and $\kappa^R_{t,3}$ with a multivariate random walk with drift (MRWD):

$$\begin{pmatrix}\kappa^R_{t,1}\\ \kappa^R_{t,2}\\ \kappa^R_{t,3}\end{pmatrix} = \begin{pmatrix}d_1\\ d_2\\ d_3\end{pmatrix} + \begin{pmatrix}\kappa^R_{t-1,1}\\ \kappa^R_{t-1,2}\\ \kappa^R_{t-1,3}\end{pmatrix} + \begin{pmatrix}\epsilon_{t,1}\\ \epsilon_{t,2}\\ \epsilon_{t,3}\end{pmatrix}$$

The parameters d_1, d_2, and d_3 are the drift terms (slopes of the linear trends), and $\varepsilon_{t,1}$, $\varepsilon_{t,2}$, and $\varepsilon_{t,3}$ are the multivariate normal error terms with mean zero and covariance matrix Σ. The future variability of $\kappa^R_{t,1}$, $\kappa^R_{t,2}$, and $\kappa^R_{t,3}$ increase over time. Moreover, we model the time series of γ^R_c with an auto-regressive integrated moving average process (ARIMA(1,1,0)):

$$\gamma^R_c - \gamma^R_{c-1} = \phi_0 + \phi_1\left(\gamma^R_{c-1} - \gamma^R_{c-2}\right) + \omega_c$$

The slope of the long-term linear trend is equal to $\phi_0/(1-\phi_1)$ if $|\phi_1| < 1$. The size of autocorrelations depends on the value of ϕ_1. The term ω_c is the normal error term with mean zero and variance σ^2_ω and is independent of $\left(\epsilon_{t,1}, \epsilon_{t,2}, \epsilon_{t,3}\right)$. If $|\phi_1| < 1$, the future variability of γ^R_c is finite and increases over time. Furthermore, we model the time series of $\kappa^B_{t,1}$ and $\kappa^B_{t,2}$ with a vector autoregressive process of order one (VAR(1)):

$$\begin{pmatrix}\kappa^B_{t,1}\\ \kappa^B_{t,2}\end{pmatrix} = \begin{pmatrix}\varphi_{1,0}\\ \varphi_{2,0}\end{pmatrix} + \begin{bmatrix}\varphi_{1,1} & \varphi_{1,2}\\ \varphi_{2,1} & \varphi_{2,2}\end{bmatrix}\begin{pmatrix}\kappa^B_{t-1,1}\\ \kappa^B_{t-1,2}\end{pmatrix} + \begin{pmatrix}\xi_{t,1}\\ \xi_{t,2}\end{pmatrix}$$

The size of autocorrelations depends on the values of $\varphi_{1,1}$, $\varphi_{1,2}$, $\varphi_{2,1}$, and $\varphi_{2,2}$. The two terms $\xi_{t,1}$ and $\xi_{t,2}$ are the bivariate normal error terms with mean zero and covariance matrix Ψ. If it is required that the projected book-to-reference ratio of mortality rates at each age should converge approximately in the long term[1], all the eigenvalues of the matrix $\begin{bmatrix}\varphi_{1,1} & \varphi_{1,2}\\ \varphi_{2,1} & \varphi_{2,2}\end{bmatrix}$ must be smaller than one in magnitude, and so both $\kappa^B_{t,1}$ and $\kappa^B_{t,2}$ tend to a constant over time. Under such conditions, the future variability of $\kappa^B_{t,1}$ and $\kappa^B_{t,2}$ are bounded across time. Note that $\left(\epsilon_{t,1}, \epsilon_{t,2}, \epsilon_{t,3}, \omega_c\right)$ and $\left(\xi_{t,1}, \xi_{t,2}\right)$ are independent in each year t. More technical details of time series modelling can be found in Tsay (2002). In fact, using other time series processes could produce very different effects. A sensitivity analysis on time series modelling is provided in section 6.

The CAE+Cohorts model is a two-population extension of the Lee & Carter (1992) model and also has two main components:

$$\text{logit}\, q^R_{x,t} = \alpha^R_x + \beta^R_x \kappa^R_t + \gamma^R_{t-x} = \eta^R_{x,t} \qquad \text{(reference component)}$$

$$\text{logit}\, q^B_{x,t} - \text{logit}\, q^R_{x,t} = \alpha^B_x + \beta^R_x \kappa^B_t = \eta^B_{x,t} \qquad \text{(book component)}$$

[1] As noted in Cairns *et al.* (2011), it would be reasonable to expect that at each age x, the projected (central estimate) ratio of $q^B_{x,t}/q^R_{x,t}$ will not diverge as $t \to \infty$.

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

In the reference component, α_x^R depicts the mortality schedule over age x, κ_t^R is the mortality index reflecting the overall mortality improvement over time t, with β_x^R as the age-specific sensitivity measure, and γ_{t-x}^R incorporates the cohort effect of the lives who were born in year t–x. In the book component, the difference in the logit mortality rate between the two populations is described by another Lee–Carter structure with the parameters α_x^B, β_x^R (the same β_x^R as that in the reference component), and κ_t^B. There are five identifiability constraints $\sum_x \beta_x^R = 1$, $\sum_t \kappa_t^R = 0$, $\sum_c \gamma_c^R = 0$, $\sum_c (c-\bar{c})\gamma_c^R = 0$, and $\sum_t \kappa_t^B = 0$, in which $\bar{c} = \frac{1}{\text{no. of cohorts in reference}} \sum_c c$. These constraints make sure that all the model parameters have unique solutions. The mortality rates and log likelihood functions can be treated similarly as under the M7-M5 model.

We then model the time series of κ_t^R as a random walk with drift (RWD), i.e., $\kappa_t^R = d + \kappa_{t-1}^R + \epsilon_t$. The parameter d is the drift term (slope of the linear trend) and ε_t is the normal error term with mean zero and variance σ_ϵ^2. The future variability of κ_t^R increases over time. Moreover, we model the time series of γ_c^R as an ARIMA(1,1,0), like in the M7-M5 model, and assume that ε_t and ω_c are independent. Finally, we model the time series of κ_t^B as an autoregressive process of order one, AR(1), i.e., $\kappa_t^B = \varphi_0 + \varphi_1 \kappa_{t-1}^B + \xi_t$. The long-term mean of κ_t^B is equal to $\varphi_0/(1-\varphi_1)$ if $|\varphi_1| < 1$. The extent of autocorrelations depends on the size of φ_1. The term ξ_t is the normal error term with mean zero and variance σ_ξ^2. If it is required that the projected book-to-reference ratio of mortality rates at each age should converge approximately in the long run, $|\varphi_1|$ must be smaller than one and so the projected κ_t^B converges to $\varphi_0/(1-\varphi_1)$ over time. (The closer φ_1 is to zero, the faster the process converges.) Under this condition, the future variability of κ_t^B is bounded across time. The error terms (ϵ_t, ω_c) and ξ_t are treated as independent.

The BIC (Bayesian information criterion) values of fitting the two models to the different data sets and the time series parameter estimates are provided in the Appendix.

4. Future Simulations

From a modelling or regulatory perspective (e.g. Comité Européen des Assurances, 2007), evaluating uncertainty of future outcomes needs to cover all of process error, parameter error, and model error (or process risk, parameter uncertainty/risk, and model uncertainty/risk). Here we adapt the residuals bootstrapping method (Koissi *et al.*, 2006, Li, 2014) to incorporate both process error (variability in the time series) and parameter error (uncertainty in estimating parameters) in simulating future mortality rates:

(a) The residuals from fitting the M7-M5 model or CAE+Cohorts model to the actual data are resampled, with replacement, for each age-time cell within all x and t. The standardised deviance residuals of the two populations are computed by using the two formulae $r_{x,t}^R = \frac{1}{\sqrt{\hat{\phi}^R}} \text{sign}\left(d_{x,t}^R - e_{x,t}^R \hat{m}_{x,t}^R\right) \sqrt{2\left(d_{x,t}^R \ln\left(d_{x,t}^R / \left(e_{x,t}^R \hat{m}_{x,t}^R\right)\right) - d_{x,t}^R + e_{x,t}^R \hat{m}_{x,t}^R\right)}$ and $r_{x,t}^B = \frac{1}{\sqrt{\hat{\phi}^B}} \text{sign}\left(d_{x,t}^B - e_{x,t}^B \hat{m}_{x,t}^B\right) \sqrt{2\left(d_{x,t}^B \ln\left(d_{x,t}^B / \left(e_{x,t}^B \hat{m}_{x,t}^B\right)\right) - d_{x,t}^B + e_{x,t}^B \hat{m}_{x,t}^B\right)}$. The dispersion parameters are estimated by $\hat{\phi}^R = \frac{1}{n_d^R - n_p^R} \sum_{x,t} 2\left(d_{x,t}^R \ln\left(d_{x,t}^R / \left(e_{x,t}^R \hat{m}_{x,t}^R\right)\right) - d_{x,t}^R + e_{x,t}^R \hat{m}_{x,t}^R\right)$ and $\hat{\phi}^B = \frac{1}{n_d^B - n_p^B} \sum_{x,t} 2\left(d_{x,t}^B \ln\left(d_{x,t}^B / \left(e_{x,t}^B \hat{m}_{x,t}^B\right)\right) - d_{x,t}^B + e_{x,t}^B \hat{m}_{x,t}^B\right)$, where n_d^R and n_p^R are the number of data points and the number of effective parameters of the reference population, and n_d^B and n_p^B are those of the book population.

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

(b) The inverse functions of the residuals formulae in step (a) are used to turn the resampled residuals into a pseudo sample of the numbers of deaths $d_{x,t}^{R(i)}$ and $d_{x,t}^{B(i)}$ for all x and t, in which the superscript (i) represents the ith scenario.

(c) The M7-M5 model or CAE+Cohorts model is fitted to the pseudo data sample from step (b) and correspondingly the model parameters $\left(\kappa_{t,1}^{R(i)}, \kappa_{t,2}^{R(i)} \kappa_{t,3}^{R(i)}, \gamma_c^{R(i)}, \kappa_{t,1}^{B(i)}, \kappa_{t,2}^{B(i)}, \text{or}\, \alpha_x^{R(i)}, \beta_x^{R(i)}, \kappa_t^{R(i)}, \gamma_c^{R(i)},\right.$ $\left.\alpha_x^{B(i)}, \kappa_t^{B(i)}\right)$ are calculated. The step allows for parameter error.

(d) The time series processes are fitted to the temporal model parameters of the pseudo data sample $\left(\kappa_{t,1}^{R(i)}, \kappa_{t,2}^{R(i)}, \kappa_{t,3}^{R(i)}, \gamma_c^{R(i)}, \kappa_{t,1}^{B(i)}, \kappa_{t,2}^{B(i)}, \text{or}\, \kappa_t^{R(i)}, \gamma_c^{R(i)}, \kappa_t^{B(i)}\right)$ from step (c) to simulate their future values. This step allows for process error.

(e) Samples of future mortality rates, $q_{x,t}^{R(i)}$ and $q_{x,t}^{B(i)}$, for all x and future t, are generated from incorporating the calculated parameters from step (c) and the simulated values from step (d) into $\eta_{x,t}^{R(i)}$ and $\eta_{x,t}^{B(i)}$. This set of future mortality rates forms one random future scenario.

(f) Steps (a) to (e) are repeated to produce 5,000 random future scenarios.

(g) For each random scenario, the future number of survivors in the pension portfolio over time is simulated as $l_{x+1,t+1}^{B(i)} \sim \text{Binomial}\left(l_{x,t}^{B(i)}, 1-q_{x,t}^{B(i)}\right)$. The notation $l_{x,t}^{B(i)}$ is the future number of lives aged x at time t in the book population. Different starting values of $l_{x,0}^{B}$ at the valuation date may be used in turn to examine the effect of different initial portfolio sizes. (In principle, the future number of lives in the reference population can also be simulated in the same fashion, but we omit this step for computation convenience, as the size of the reference population here is very large for a binomial distribution assumption.)

Note that while Haberman *et al.* (2014) chose to ignore parameter error of the reference population and did not perform the bootstrapping on the reference population, we still carry out the bootstrapping for both populations in order to formalise the overall procedure. In fact, we have experimented with their approach but realise that it tends to underestimate demographic basis risk. Moreover, the extent of model error (uncertainty in model selection) can be assessed by comparing the results obtained from different models and assumptions. This way requires some form of subjective or qualitative judgement, which may involve certain individual bias. An alternative, being more theoretically sound, is to adopt a Bayesian framework to incorporate all of process error, parameter error, and model error simultaneously. However, this Bayesian approach is technically demanding (e.g. Markov chain Monte Carlo simulation) and much more computationally intensive (Li, 2014), particularly for use in insurance or actuarial practice.

Using the simulated environment, we implement a simple numerical optimisation procedure to determine the optimal positions of the hedging instruments. In this paper, we set an objective to minimise the 99.5% value-at-risk (VaR) (minus the mean) of the present value of the aggregate pension portfolio position. In practice, other risk measures (e.g. standard deviation and 99.5% expected shortfall) or a mix of different objectives (e.g. risk minimisation with a targeted level of profitability) may also be used, depending on the purpose of the analysis. Note that when the number of unknown quantities to be estimated increases (e.g. multiple index-based derivatives are available for hedging), there may be a higher chance of the procedure getting trapped in a local optimum rather than obtaining the true global optimum. Some simple, practical methods to mitigate this problem include using other sensible starting values, applying alternative numerical algorithms, simplifying the model settings where possible, and repeating the optimisation procedure by setting some or all the previous solutions as the new starting values.

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

5. Hedging Results

The effectiveness of an index-based longevity hedge can be described as how much longevity risk is transferred away. The remaining part can then be seen as an outcome of longevity basis risk. To be consistent with Coughlan *et al.* (2011), we define the level of longevity risk reduction for a longevity hedge imposed on a pension plan or annuity portfolio as:

$$\text{longevity risk reduction} = \left(1 - \frac{\text{risk(hedged)}}{\text{risk(unhedged)}}\right) \times 100\,\%$$

The terms risk(unhedged) and risk(hedged) are the portfolio's aggregate longevity risk before and after the hedging. This metric provides the proportion of the portfolio's initial longevity risk that is being hedged away. We apply some risk measures commonly used in the literature (Dowd & Blake, 2006), including the variance, standard deviation, 99.5% VaR, and 99.5% expected shortfall (conditional VaR), to the random present value of the portfolio liability. The 99.5% VaR is of particular interest in practice, as it is embedded in the calculation of the Solvency Capital Requirement (SCR) under Solvency II[2]. Measuring the levels of risk reduction in different index-based longevity hedges could help identify potential opportunities for capital savings.

In this section, we first consider a simple hypothetical case study of a pension plan to illustrate the concept. The current date is taken as the start of the calendar year 2014. Suppose all the pensioners in the plan are now aged exactly 65 and each pension pays $1 per year on survival from ages 66 to 90. The pension plan is already closed and there are no more new members. The pensioners (book population) have the same mortality experience as that reflected in the CMI data or the ONS data (starting from year 2000 or 2001). The pension plan sponsor intends to minimise the longevity risk exposure by constructing a longevity hedge with standardised longevity swaps. (Note: Bespoke longevity swaps are by far the most commonly used hedging instruments in current practice.) Assume that an index-based 25-year longevity swap for the same birth cohort as the pensioners with annual exchange of payments is available in the de-risking market, in which the reference population in the floating leg is the English and Welsh male population. The payments on the fixed leg of the swap are calculated from the central estimates of future mortality rates, based on the HMD data from year 1980. This simplifying assumption implies a zero risk premium and would affect only the price but not the effectiveness of the hedge. Let the interest rate be 1% p.a. flat during the whole period. (Note: UK Gilt 10-year and 30-year yields were 1.05% p.a. and 1.67% p.a. as at 24 April 2017.) Later we will cover a more complex pension plan with multiple cohorts, in which there are two cases of an open pension plan and a closed one.

The present value of future liability of the pension plan is equal to $\sum_{t=1}^{25} l_{65+t,t}^{B(i)} (1+r)^{-t}$, in which r is the interest rate. Moreover, the present value of future cash inflows of the longevity swap as a floating rate receiver is expressed as $\sum_{t=1}^{25} \left({}_tp_{65}^{R(i)} - {}_tp_{65}^{R\,;\,\text{forward}}\right)(1+r)^{-t}$, where the random future survivor index ${}_tp_{65}^{R(i)}$ and the forward survivor index ${}_tp_{65}^{R\,;\,\text{forward}}$ are calculated by adapting the survival probability formula ${}_tp_{65}^{R} = \left(1-q_{65,0}^{R}\right)\left(1-q_{66,1}^{R}\right)\ldots\left(1-q_{65+t-1,t-1}^{R}\right)$. Then the present value of the aggregate pension plan position after taking the longevity hedge is stated as $\sum_{t=1}^{25} l_{65+t,t}^{B(i)}(1+r)^{-t} - w \sum_{t=1}^{25}\left({}_tp_{65}^{R(i)} - {}_tp_{65}^{R\,;\,\text{forward}}\right)(1+r)^{-t}$, and the cash outflow of the net position at

[2] Under Solvency II, the SCR is the amount of capital required to cover all losses over a 1-year time horizon with a probability of at least 99.5%. By contrast, the focus of this paper is on the present value of the entire future liability.

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

each time $t=1, 2, \ldots, 25$ is given by $l_{65+t,t}^{B(i)} - w\left({}_tp_{65}^{R(i)} - {}_tp_{65}^{R;\text{forward}}\right)$. The weight w is the notional amount of the longevity swap required to minimise the 99.5% VaR (minus the mean) of the present value of the aggregate position.

For the IMD quintile groups, the portfolio's initial longevity risk levels before hedging, with regard to the standard deviation, 99.5% VaR (minus the man), and 99.5% expected shortfall (minus the mean), are roughly about 1.5%, 3%, and 4% of the portfolio's expected present value. Table 4 provides the estimated levels of longevity risk reduction under different initial portfolio sizes. The longevity risk reduction estimates using the standard deviation, 99.5% VaR, and 99.5% expected shortfall (conditional VaR) vary from about 60% to 95% for the three largest portfolio sizes. These estimates are quite close between the three risk measures in general. On the other hand, the estimates using the variance, which has a different scale to the others, are usually larger at 90% or over for the bigger portfolios. Moreover, as the initial portfolio size becomes smaller, sampling basis risk in the future simulations grows in significance, and consequently the risk reduction level decreases. But this decline in risk reduction looks a little uneven, and is not too significant until the portfolio size drops to 25,000 and below. For a small portfolio size of 1,000, the risk reduction levels decrease to around 30%–40% only. (Those reductions smaller than 50% are shaded in all the tables.) These results suggest that sampling basis risk is negligible when the portfolio has more than, say, around 20,000 lives, but the effect can be significant when the portfolio size is down to just a few thousand.

Furthermore, for the bigger portfolios, the third most deprived areas have greater risk reductions than the others. It implies that the simulated ${}_tp_{65}^{B(i)}$ and ${}_tp_{65}^{R(i)}$ have higher dependence for the third most deprived areas. The people living in these areas have medium income, employment, education, and so on, and they can potentially be matched more closely by the overall reference population. In contrast, the most deprived areas have the smallest risk reduction estimates for the larger portfolio sizes, ranging from around 60%–70%. This observation suggests that the simulated ${}_tp_{65}^{B(i)}$ of the most deprived groups may deviate more significantly from the simulated ${}_tp_{65}^{R(i)}$ of the general population when compared to the other groups.

Besides the overall present value, we also investigate the individual cash flows and compute their risk reduction levels specifically in each future year. Figure 4 shows the risk reduction effect (in terms of 99.5% VaR) of the cash flows in each year for a portfolio size of 100,000 males under the CAE +Cohorts model. It can be seen that the risk reduction estimates of the individual cash flows are actually very small in the early years, but they increase progressively over time to reach a high level in the later years. This effect appears to arise from the fact that the RWD in the reference component produces unbounded future variability whereas the AR(1) in the book component yields bounded future variability. The variability in the latter would reduce in significance relatively across time, meaning that the simulated differences between the two populations would become less important comparatively, resulting in a lower level of demographic basis risk being modelled for the longer term. This model implication is in line with the usual view that two related populations' mortality improvements may diverge temporarily but would likely move back in line as time goes by. Its significance will further be tested in section 6.

We now study a more realistic hypothetical scenario of a pension plan with multiple cohorts. Suppose there are currently 30,000 pensioners in the plan, and their demographic structure, from ages 60 to 89, is displayed in Figure 5. We consider two cases of an open pension plan and a closed plan. For the open pension plan, 1,400 new members join the plan at age 60 every year, and an index-based longevity

Table 4. Level of longevity risk reduction (in % of initial longevity risk) in a hypothetical scenario of England index of multiple deprivation quintile groups of a single male cohort.

	M7-M5				CAE + Cohorts			
Group\size	Variance	SD	99.5% VaR	99.5% ES	Variance	SD	99.5% VaR	99.5% ES
Most deprived areas								
Infinite	87	64	71	70	90	68	68	69
100,000	87	64	72	70	90	69	68	68
50,000	86	63	70	69	90	68	67	68
25,000	85	62	67	67	87	64	63	63
10,000	83	59	64	62	85	62	59	59
5,000	80	55	59	57	80	56	52	52
2,500	73	48	51	50	73	48	46	48
1,000	60	37	38	40	57	34	33	34
Second most deprived areas								
Infinite	96	81	83	83	98	85	87	86
100,000	96	79	81	80	97	83	83	83
50,000	95	78	79	79	97	81	81	81
25,000	94	76	77	77	95	78	76	77
10,000	92	71	71	71	91	71	67	67
5,000	88	65	62	62	87	64	62	62
2,500	80	55	54	53	77	52	52	52
1,000	65	41	41	41	59	36	33	34
Third most deprived areas								
Infinite	100	94	94	94	100	95	94	94
100,000	99	90	91	91	99	90	90	90
50,000	99	88	88	89	98	87	86	86
25,000	98	84	84	84	97	82	81	81
10,000	95	77	74	75	93	74	71	70
5,000	91	69	69	69	88	65	62	63
2,500	83	59	56	57	78	53	51	50
1,000	66	42	40	40	58	35	33	34
Fourth most deprived areas								
Infinite	98	88	81	76	99	88	84	82
100,000	98	86	81	75	99	88	87	85
50,000	97	82	78	73	98	85	84	82
25,000	96	80	76	71	96	81	79	79
10,000	94	75	74	71	93	73	70	69
5,000	90	68	67	64	87	64	62	61
2,500	82	57	57	55	77	52	50	50
1,000	65	41	42	41	58	35	33	34
Least deprived areas								
Infinite	97	83	75	70	98	87	83	82
100,000	96	80	76	71	98	85	81	79
50,000	94	77	72	65	97	83	80	78
25,000	95	78	73	69	96	79	78	77
10,000	91	69	66	64	91	71	69	67
5,000	88	65	64	63	86	63	60	60
2,500	80	55	54	52	75	50	49	48
1,000	61	37	34	36	56	34	32	33

Note: VaR, value-at-risk; ES, expected shortfall.

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

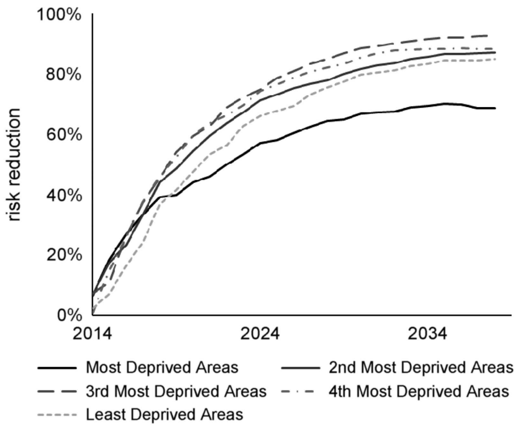

Figure 4. Level of longevity risk reduction (in % of initial longevity risk; in terms of 99.5% value-at-risk) of individual cash flows in a hypothetical scenario of England index of multiple deprivation quintile groups of a single cohort of 100,000 males under CAE + Cohorts model.

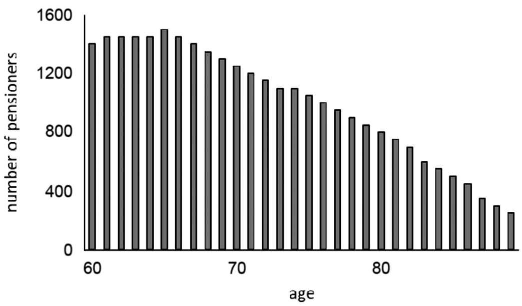

Figure 5. Initial demographic structure of a hypothetical pension plan with multiple cohorts.

swap for the cohort aged 60 now and with a maturity of 30 years is used to construct the hedge. On the other hand, for the closed pension plan, there are no new members, and two index-based longevity swaps for the two cohorts aged 60 and 70 now, with maturities of 30 and 20 years, respectively, are used for the hedge. Each pension pays \$1 per year on survival from ages 61 to 90. The duration of the hedging scheme is 30 years from the current date, and we only consider the portfolio cash flows during this period. All the other previous settings are kept equal.

Note that for the open pension plan, those aged 31–59 now will join the plan later in the future, while the current pensioners are aged 60–89. For demonstration purposes, we use a single longevity swap for the cohort aged 60 now, which is in the middle of the age range under consideration. For the closed pension plan, we use two longevity swaps for the two cohorts aged 60 and 70 now, in which the first one covers the entire duration and the second one refers roughly to the weighted average age of the current pensioners. Table 5 shows that the longevity risk reduction estimates regarding the standard deviation and the extreme risk measures range largely from about 40% to 60% for the open pension plan, and from around 60%–80% for the closed pension plan. These estimates are generally smaller than those in Table 4, because the earlier scenario involves one swap to match one cohort exactly, while the two portfolios here use only one or two swaps to hedge multiple cohorts approximately. Consequently, the extent of demographic basis risk, in terms of age or cohort differences, is greater in these cases.

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

Table 5. Level of longevity risk reduction (in % of initial longevity risk) in a hypothetical scenario of England index of multiple deprivation quintile groups of 30,000 males of multiple cohorts.

	M7-M5				CAE + Cohorts			
Group\plan	Variance	SD	99.5% VaR	99.5% ES	Variance	SD	99.5% VaR	99.5% ES
Most deprived areas								
Open	75	50	53	51	59	36	38	37
Closed	81	56	61	59	81	57	55	55
Second most deprived areas								
Open	82	57	53	52	67	42	42	38
Closed	92	72	73	72	93	74	71	71
Third most deprived areas								
Open	84	60	59	60	69	44	43	41
Closed	96	80	79	79	96	79	78	78
Fourth most deprived areas								
Open	84	60	59	58	70	45	48	46
Closed	95	77	77	73	95	78	78	75
Least deprived areas								
Open	81	57	55	55	71	46	46	44
Closed	92	71	70	66	94	75	75	73

Note: VaR, value-at-risk; ES, expected shortfall.

Moreover, the closed pension plan has larger risk reduction estimates than the open pension plan, as the closed plan uses two swaps to hedge only the current pensioners with no new members.

When the life market is immature, the index-based swap transactions available may be restricted to a few major cohorts due to liquidity issues (e.g. Coughlan *et al.*, 2011), and an "approximate" hedge can be constructed like the demonstration above. If longevity swaps for multiple cohorts and with varying maturities become available in the market, they may be exploited to reduce the extent of demographic basis risk and improve the overall hedge effectiveness. However, the number of pensioners at each age in the example here is only around a thousand or less, and calibrating one unique swap to each single cohort separately would induce much sampling basis risk. An alternative approach is to calibrate multiple swaps altogether as a group to the whole plan, instead of mapping specifically one swap on each cohort. The overall hedge effectiveness could then be improved, although implicitly there would be some offsetting effects between the swaps of different cohorts. We now test eleven index-based longevity swaps for the cohorts aged 35, 40, 45, … , 85 at present on the open pension plan, with a (delayed) maturity of 5, 10, 15, 20, 25 years for those who have not joined the plan as yet and maturities of 30, 25, 20, 15, 10, 5 years for the current pensioners. We also try six index-based longevity swaps for the cohorts aged 60, 65, 70, … , 85 now on the closed pension plan, with maturities 30, 25, 20, … , 5 years, respectively. Table 6 illustrates that this alternative approach of using "aggregate calibration"' with 5-year age buckets increases the risk reduction estimates in many of the cases. The effect is more obvious for the open pension plan, which comprises both existing pensioners and new members but has involved only one swap in the previous setting. Though there would be intricate subsidising effects across the swaps of different cohorts, it appears that the aggregate amount of demographic basis risk is decreased and the overall hedge effectiveness is improved. Comparatively, the effect of including more swaps is mild for the closed pension plan, which suggest that the two swaps in the previous setting with 10-year age buckets would already be adequate to cover the lives of multiple cohorts in that case and more precise hedging would not lead to much improvement.

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

Table 6. Level of longevity risk reduction (in % of initial longevity risk) in a hypothetical scenario of England index of multiple deprivation quintile groups of 30,000 males of multiple cohorts (with swaps of 5-year age buckets).

	M7-M5				CAE + Cohorts			
Group\plan	Variance	SD	99.5% VaR	99.5% ES	Variance	SD	99.5% VaR	99.5% ES
Most deprived areas								
Open	85	62	68	67	93	73	72	72
Closed	84	60	67	66	87	64	63	62
Second most deprived areas								
Open	95	77	78	78	96	81	81	80
Closed	93	73	67	74	93	74	74	73
Third most deprived areas								
Open	97	84	84	84	97	84	84	83
Closed	97	82	83	83	96	81	79	79
Fourth most deprived areas								
Open	96	79	77	70	97	83	83	83
Closed	96	79	77	74	96	79	78	75
Least deprived areas								
Open	95	77	75	69	96	79	80	78
Closed	92	72	71	66	94	76	75	74

Note: VaR, value-at-risk; ES, expected shortfall.

For the CMI pensioners, the initial longevity risk levels before hedging, with regard to the standard deviation and the tail risk measures, are approximately around 1.5%, 3.5%, and 4.5% of the expected present value. Tables 7–9 present the risk reduction results for the CMI pensioners, which have similar patterns to those of the IMD quintile groups. For instance, the longevity risk reduction estimates regarding the standard deviation and the tail risk measures are mostly in the range of around 50%–80% for a portfolio size of 100,000. When the portfolio size is just 1,000, the estimates decrease to about 30%–40% only. Also, more precise hedging often leads to better hedge effectiveness. On the other hand, for the bigger portfolios, the CMI pensioners generally have smaller risk reduction estimates than the IMD quintile groups. Note that the IMD groups are significant sectors of the entire population and so they can be matched more closely by the reference population. But for the smaller portfolios, the differences in the hedging results between the CMI pensioners and the IMD groups are less clear, which again highlight the importance of sampling basis risk. In addition, the two models produce more different results between them for the CMI pensioners, which may be due to the smaller data sizes of the CMI data set. The noisy patterns in a smaller data sample may have been captured somewhat differently under the unique properties of each model.

In the CMI data set, the financials and technology industries have only around 10,000 lives or less per year in their data. The high sampling variability in a small dataset may flow through to the temporal parameters computed, leading to an overestimation of demographic basis risk. The characterisation approach in Haberman *et al.* (2014) may then be adopted, in which the book data (with a small size) are not modelled directly but a combination of some alternative proxy data (with large sizes) is modelled instead. For instance, comparing the average mortality levels and improvement rates between Tables 1 and 2, the lower (higher) pension group in financials may be proxied by the fourth most (least) deprived IMD group. For a pension plan with 100,000 lives all aged 65, assuming half of the lives have a low pension and the other half have a high pension, the risk reduction

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

Table 7. Level of longevity risk reduction (in % of initial longevity risk) in a hypothetical scenario of Continuous Mortality Investigation pensioners of a single male cohort.

Group\size	M7-M5 Variance	M7-M5 SD	M7-M5 99.5% VaR	M7-M5 99.5% ES	CAE + Cohorts Variance	CAE + Cohorts SD	CAE + Cohorts 99.5% VaR	CAE + Cohorts 99.5% ES
Basic materials			(normal retirement; lower pension group)					
100,000	87	65	63	63	95	77	74	73
1,000	60	37	36	38	57	34	37	36
Industrials			(normal retirement; lower pension group)					
100,000	77	52	51	39	81	56	50	49
1,000	54	32	35	27	52	30	25	28
Consumer goods			(normal retirement; lower pension group)					
100,000	81	57	39	29	92	72	66	62
1,000	57	35	31	25	55	33	27	31
Commercial services			(normal retirement; lower pension group)					
100,000	95	78	79	79	97	83	79	79
1,000	66	41	41	41	55	33	30	28
Utilities			(normal retirement; lower pension group)					
100,000	65	41	58	56	93	73	70	69
1,000	50	29	30	33	54	32	27	28
Local authority			(normal retirement; lower pension group)					
100,000	94	75	73	58	98	85	83	83
1,000	59	36	33	25	54	32	30	27

Note: VaR, value-at-risk; ES, expected shortfall.

Table 8. Level of longevity risk reduction (in % of initial longevity risk) in a hypothetical scenario of Continuous Mortality Investigation pensioners of 30,000 males of multiple cohorts.

Group\plan	M7-M5 Variance	M7-M5 SD	M7-M5 99.5% VaR	M7-M5 99.5% ES	CAE + Cohorts Variance	CAE + Cohorts SD	CAE + Cohorts 99.5% VaR	CAE + Cohorts 99.5% ES
Basic materials			(normal retirement; lower pension group)					
Open	76	51	50	49	66	42	43	44
Closed	84	60	60	58	88	66	64	62
Industrials			(normal retirement; lower pension group)					
Open	67	43	46	29	58	35	32	30
Closed	69	44	45	37	71	47	42	39
Consumer goods			(normal retirement; lower pension group)					
Open	70	45	38	29	64	40	35	36
Closed	79	54	45	37	86	62	56	49
Commercial services			(normal retirement; lower pension group)					
Open	84	60	58	57	65	41	40	40
Closed	93	73	74	72	93	74	70	70
Utilities			(normal retirement; lower pension group)					
Open	54	32	47	46	63	39	38	35
Closed	71	47	56	51	89	66	63	60
Local authority			(normal retirement; lower pension group)					
Open	81	56	54	47	66	42	44	44
Closed	91	69	67	55	94	75	73	72

Note: VaR, value-at-risk; ES, expected shortfall.

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

Table 9. Level of longevity risk reduction (in % of initial longevity risk) in a hypothetical scenario of Continuous Mortality Investigation pensioners of 30,000 males of multiple cohorts (with swaps of 5-year age buckets).

	M7-M5				CAE + Cohorts			
Group\plan	Variance	SD	99.5% VaR	99.5% ES	Variance	SD	99.5% VaR	99.5% ES
Basic materials	(normal retirement; lower pension group)							
Open	88	65	65	63	94	75	74	74
Closed	85	62	62	59	90	69	69	67
Industrials	(normal retirement; lower pension group)							
Open	76	51	56	37	83	58	55	55
Closed	72	47	48	40	71	46	41	40
Consumer goods	(normal retirement; lower pension group)							
Open	79	54	46	30	91	70	66	64
Closed	79	54	45	38	85	61	55	49
Commercial services	(normal retirement; lower pension group)							
Open	95	77	79	78	96	80	76	76
Closed	93	74	76	75	94	75	72	71
Utilities	(normal retirement; lower pension group)							
Open	62	38	63	62	92	72	72	71
Closed	71	46	57	50	87	64	61	61
Local authority	(normal retirement; lower pension group)							
Open	91	70	68	54	96	80	80	78
Closed	91	70	68	57	94	76	75	74

Note: VaR, value-at-risk; ES, expected shortfall.

estimates regarding the standard deviation and the tail risk measures are 85%, 86%, and 82%, respectively, using the M7-M5 model, and are 82%, 78%, and 76% using the CAE+Cohorts model, via the characterisation approach. For the former case, the VAR(1) process in the book components of the two characterising groups has four dimensions in total, while for the latter case, it has only two dimensions. We find that it is practically difficult for the VAR(1) process in the first case to possess the convergence property. So we adopt a reduced version of the time series process, which is to fit a bivariate VAR(1) process to each characterising group and assumes that the error terms of the two bivariate processes are correlated. Finally, it should be noted that the mapping in the example above is not precise at all and better proxies can be sought if book data with more details on categorising pensioners or policyholders are available.

6. Sensitivity Analysis

In this section, we carry out a sensitivity analysis on the hedging results by making a series of changes to the initial model settings and assumptions. The tested items include the interest rate, swap weight, older ages, data fitting period, simulation method, other mortality projection models, and additional model features. When a change is made, other things are kept equal unless otherwise specified. This analysis allows us to have a better understanding of the robustness of the hedging results under different conditions. Table 10 shows that in general, as the interest rate increases, the risk reduction estimates for the present value decreases. The later individual cash flows have greater risk reduction and are affected more by a higher interest rate than the earlier cash flows. As such, the aggregate risk reduction decreases as the interest rate rises. But from an interest rate of 1%–5% p.a., the changes in the estimates shown in the table are only 9% or less in magnitude. Furthermore, the hedging results under a variable interest rate

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

Table 10. Levels of longevity risk reduction (in % of initial longevity risk; M7-M5 (left figure) versus CAE + Cohorts (right figure)) under different interest rate assumptions.

100,000 males	CMI pensioners (basic materials; normal retirement; lower pension)			IMD group (most deprived areas)		
Interest rate (p.a. (%))	SD	99.5% VaR	99.5% ES	SD	99.5% VaR	99.5% ES
1	65/77	63/74	63/73	64/69	72/68	70/68
2	63/77	61/73	60/73	63/69	69/67	68/67
3	62/76	59/71	58/71	62/68	69/67	69/68
4	60/75	58/72	57/71	62/68	68/67	67/68
5	58/74	55/71	54/71	61/68	67/66	66/67
CIR	53/52	55/55	53/55	57/57	62/58	61/59

Note: CMI, Continuous Mortality Investigation; IMD, index of multiple deprivation; VaR, value-at-risk; ES, expected shortfall; CIR, Cox–Ingersoll–Ross.

environment simulated from the discretised Cox–Ingersoll–Ross (CIR) model[3] are also given in the table. The model is fitted to the most recent few years of historically low interest rates. The risk reduction estimates are clearly smaller under variable interest rates than under constant interest rates, reflecting the influence of interest rate risk. Although it is difficult to foresee how long the current low interest rates would last, it can be perceived that higher interest rates with more fluctuations would reduce the hedge effectiveness further. Interest rate swaps and government bonds may then be incorporated into the hedging scheme to reduce the impact of interest rate risk (Tsai *et al*., 2011).

So far, the estimation of the optimal positions and hedging results has been based on simulated scenarios from the fitted M7-M5 model or CAE+Cohorts model and numerical optimisation with regard to the 99.5% VaR. If the standard deviation or the 99.5% expected shortfall (conditional VaR) is taken as the risk measure instead in the optimisation process, it can be seen from Table 11 that the resulting differences are minimal. This observation is probably a result of the simulated distribution of the portfolio present value, which looks largely symmetric without much skewness. Moreover, when the swap size is simply set as the initial portfolio size (i.e. a hedge ratio of one), the corresponding differences in risk reduction are also very small, because the original, numerically optimised hedge ratio is actually quite close to one. Furthermore, when the swap weight is computed from a "wrong" model which is different to the one being used to generate the simulations, there is still not much change in the risk reduction estimates.

We also test a wider age range from 60 to 99. The ONS population data are not split by single age for ages 90+ and the corresponding HMD old-age proportions are used as a proxy here. The two models are re-fitted to the new age range. Suppose that there are four different pensions, in which one pays £1 per year on survival from ages 66 to 90 for those pensioners aged 65 at present (as in the original settings), one from ages 66 to 100 for those aged 65 now, one from ages 76 to 100 for those aged 75 now, and one from ages 86 to 100 for those aged 85 presently. Assume that index-based longevity swaps for the same birth cohorts as the pensioners are available with maturities of 25, 35, 25, and

[3] Note that the CIR model is a short interest rate model and is probably too simplified for long-term projection. Its use here is merely to demonstrate the potential effect of interest rate risk on the hedge effectiveness. For more realistic long-term projection, other important features such as economic cycles and future impact of monetary policies should properly be allowed for.

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

Table 11. Levels of longevity risk reduction (in % of initial longevity risk; M7-M5 (left figure) versus CAE + Cohorts (right figure)) using different swap weights.

100,000 males	CMI pensioners (basic materials; normal retirement; lower pension)			IMD group (most deprived areas)		
Swap weight	SD	99.5% VaR	99.5% ES	SD	99.5% VaR	99.5% ES
Optimisation (99.5% VaR)	65/77	63/74	63/73	64/69	72/68	70/68
Optimisation (99.5% ES)	63/77	63/74	63/73	64/68	71/67	70/69
Optimisation (SD)	65/77	63/74	63/72	64/69	71/67	70/68
One-to-one	65/77	62/74	62/74	63/66	70/65	69/66
Wrong model	65/77	62/72	61/71	63/67	70/67	69/68

Note: CMI, Continuous Mortality Investigation; IMD, index of multiple deprivation; VaR, value-at-risk; ES, expected shortfall.

15 years, respectively. The figures in the first two rows of Table 12 suggest that the models are fairly robust to incorporating advanced ages into the modelling process, except for some estimates of the IMD quintile group, which may be caused by the use of approximate population data for ages 90+. Also, it is observed that the risk reduction levels are higher for those pensions with longer durations, as the later individual cash flows have greater risk reduction effect than the earlier cash flows.

Table 13 demonstrates that the differences in risk reduction between two fitting periods for the reference population starting from years 1980 and 1990 are within 5%. Using a slightly shorter data period (starting from 1 or 2 years later) for the book population, the risk reduction levels change by 6% or less. It seems that the two models are quite robust to different data fitting periods, but book data of a longer period are needed to carry out a more thorough testing.

Besides residuals bootstrapping, a simpler parametric method is to simulate the errors of the fitted time series processes directly to generate random future mortality rates. This method, however, allows for only process error but not parameter error, and may understate longevity basis risk. The figures in Table 14 reveal that the parametric method does underestimate demographic basis risk and so overestimates the level of longevity risk reduction by around 8%–20% in magnitude. Proper allowance for parameter error is important, though the parametric method is generally much faster than the residuals bootstrapping process.

Based on the previous simulations, the differences in the risk reduction estimates between the M7-M5 and CAE+Cohorts models become more obvious when the data size is small or when the hedging scheme is not precise enough. Since model error is not negligible under certain conditions, three extra models are considered here for further comparison. The first model is a cohort extension of the first approach in Carter & Lee (1992), which applies the Lee & Carter (1992) model to each population and then co-models the two mortality indices:

$$\text{logit}\, q_{x,t}^{R} = \alpha_x^R + \beta_x^R \kappa_t^R + \gamma_{t-x}^R \qquad \text{(reference population)}$$

$$\text{logit}\, q_{x,t}^{B} = \alpha_x^B + \beta_x^B \kappa_t^B + \gamma_{t-x}^R \qquad \text{(book population)}$$

$$\mathrm{K}_t = \Theta + \mathrm{K}_{t-1} + \Delta_t \qquad \text{(bivariate random walk with drift`)}$$

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

Table 12. Levels of longevity risk reduction (in % of initial longevity risk; M7-M5 (left figure) versus CAE + Cohorts (right figure)) using older ages in modelling.

100,000 males	CMI pensioners (basic materials; normal retirement; lower pension)			IMD group (most deprived areas)		
Pay age range	SD	99.5% VaR	99.5% ES	SD	99.5% VaR	99.5% ES
66–90 (old)	65/77	63/74	63/73	64/69	72/68	70/68
66–90 (new)	64/78	62/76	59/74	52/59	60/62	58/63
66–100	76/86	76/85	74/84	56/61	69/65	69/66
76–100	74/80	73/78	72/76	56/58	65/62	64/63
86–100	65/64	63/63	63/59	55/53	60/55	61/57

Note: CMI, Continuous Mortality Investigation; IMD, index of multiple deprivation; VaR, value-at-risk; ES, expected shortfall.

Table 13. Levels of longevity risk reduction (in % of initial longevity risk; M7-M5 (left figure) versus CAE + Cohorts (right figure)) using different data fitting periods.

100,000 males	CMI pensioners (basic materials; normal retirement; lower pension)			IMD group (most deprived areas)		
Data period	SD	99.5% VaR	99.5% ES	SD	99.5% VaR	99.5% ES
From 1980 (reference)	65/77	63/74	63/73	64/69	72/68	70/68
From 1990 (reference)	63/76	62/70	58/68	62/66	74/65	71/65
From 1 year later (book)	64/77	64/76	65/75	64/70	72/70	73/70
From 2 years later (book)	63/76	63/73	61/72	69/73	77/74	76/74

Note: CMI, Continuous Mortality Investigation; IMD, index of multiple deprivation; VaR, value-at-risk; ES, expected shortfall.

Table 14. Levels of longevity risk reduction (in % of initial longevity risk; M7-M5 (left figure) versus CAE + Cohorts (right figure)) using different simulation methods.

100,000 males	CMI pensioners (basic materials; normal retirement; lower pension)			IMD group (most deprived areas)		
Simulation method	SD	99.5% VaR	99.5% ES	SD	99.5% VaR	99.5% ES
Bootstrapping	65/77	63/74	63/73	64/69	72/68	70/68
Parametric	73/87	73/86	72/84	84/83	84/82	84/81

Note: CMI, Continuous Mortality Investigation; IMD, index of multiple deprivation; VaR, value-at-risk; ES, expected shortfall.

The notation $K_t = \left(\kappa_t^R, \kappa_t^B\right)'$, Θ is the vector drift term, and Δ_t is the bivariate normal error term. The second model is a cohort extension of the third approach in Carter & Lee (1992), which assumes that the two mortality indices follow a co-integrated process (Li & Hardy, 2011):

$$\text{logit}\, q_{x,t}^R = \alpha_x^R + \beta_x^R \kappa_t^R + \gamma_{t-x}^R \qquad \text{(reference population)}$$

$$\text{logit}\, q_{x,t}^B = \alpha_x^B + \beta_x^B \kappa_t^B + \gamma_{t-x}^R \qquad \text{(book population)}$$

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

$$\kappa_t^R = \theta + \kappa_{t-1}^R + \delta_t \quad \text{(RWD)}$$

$$\kappa_t^B = a_0 + a_1\kappa_t^R + \omega_t \quad \text{(co-integrated process)}$$

The parameters θ, a_0, and a_1 define the co-integrated process, and δ_t and ω_t are independent normal error terms. The final one is a cohort extension of the model proposed by Zhou *et al.* (2013), which uses a common age-specific sensitivity measure between the two populations and a weakly stationary AR(1) process for the difference between the two mortality indices:

$$\text{logit}\, q_{x,t}^R = \alpha_x^R + \beta_x^R\kappa_t^R + \gamma_{t-x}^R \quad \text{(reference population)}$$

$$\text{logit}\, q_{x,t}^B = \alpha_x^B + \beta_x^R\kappa_t^B + \gamma_{t-x}^R \quad \text{(book population)}$$

$$\kappa_t^R = \theta + \kappa_{t-1}^R + \delta_t \quad \text{(RWD)}$$

$$\kappa_t^R - \kappa_t^B = b_0 + b_1\left(\kappa_{t-1}^R - \kappa_{t-1}^B\right) + \omega_t \quad \text{(AR(1) process)}$$

The notation θ, b_0, and b_1 are the parameters of the time series processes, and δ_t and ω_t are independent normal error terms.

The first two extra models above are "non-coherent" while the last model is "coherent" (Li & Lee, 2005). When a two-population mortality projection model is coherent, the projected (central estimate) ratio of future mortality rates between the two populations at each age tends to a constant in the long term. The common view in the literature is that two related populations' future mortality trends may deviate in the short term but would move more consistently over the long term. A lack of this long-term coherence in the modelling may result in an overestimation of longevity basis risk. Apart from the coherence of the central estimates, the simulated future variability is also an important factor. Under the first two models, both kappa parameters of the two populations have unbounded future variability. But under the last model, while the kappa parameter of the reference population also has unbounded future variability, the difference between the kappa parameters of the two populations has bounded future variability. The major implication is that the two populations' future mortality movements could diverge more freely under the first two models, especially in the long term, but they would become more consistent over time under the last model.

Table 15 provides the risk reduction estimates computed from the M7-M5 and CAE+Cohorts models and also the three extra models discussed above. The first two extra models, which are non-coherent and have unbounded future variability, produce very small risk reduction estimates. It appears that there is a significant overestimation of demographic basis risk. By contrast, the M7-M5 model, the CAE+Cohorts model, and the last extra model are all coherent and have bounded future variability between the two populations. Their risk reduction estimates are broadly in agreement with one another and reflect a more proper allowance for demographic basis risk[4].

[4] Besides using the various models, we have also followed the Solvency II Standard Formula and tested different pairs of "longevity shocks" to the book and reference populations. Based on the past data, we find that index-based longevity hedging would reduce any portfolio loss significantly when there are considerable unanticipated mortality improvements. Even if the longevity shock is larger for the book population than for the reference population, due to the existence of longevity basis risk, the reduction in the portfolio loss from an index-based hedge would still be sizable, given that the shocks on the two populations are in the same direction. These scenario testings suggest that the very low risk reduction levels produced by the first extra model do not look reasonable.

Table 15. Levels of longevity risk reduction (in % of initial longevity risk) under different two-population mortality projection models.

100,000 males	CMI pensioners (basic materials; normal retirement; lower pension)			IMD group (most deprived areas)		
Model	SD	99.5% VaR	99.5% ES	SD	99.5% VaR	99.5% ES
M7-M5	65	63	63	64	72	70
CAE + Cohorts	77	74	73	69	68	68
Extra model 1	7	8	4	28	22	22
Extra model 2	45	52	51	53	57	55
Extra model 3	77	74	74	67	64	64

Note: CMI, Continuous Mortality Investigation; IMD, index of multiple deprivation; VaR, value-at-risk; ES, expected shortfall.

In spite of the differences in the structures between the various types of two-population mortality projection models, the coherence property and the behaviour of simulated future variability, which depend largely on the assumed time series processes, are the major factors in determining the calculated level of longevity risk reduction.

In the past century, mortality improvement trends have not always been so smooth over time. There could be one-off events such as wars, epidemics, and catastrophes, leading to temporary, sharp changes in mortality levels (i.e. mortality jumps). There could also be long-term effects like medical developments and climate changes, resulting in permanent shifts in mortality improvement rates (i.e. structural changes). Figure 6 displays the mortality index of the English and Welsh male population, in which there are a few "mortality spikes" (red) before 1950 and a shift in the improvement trend (blue) during around 1970. It is difficult to allow for them in the modelling process as these events are either rare or hard to detect precisely. Based on the HMD data of several developed countries, we examine three additional scenarios here using some arbitrary assumptions. The first scenario involves structural changes, the second one involves mortality jumps, and the final one has both. We adjust the last extra model in Table 15 to incorporate these effects, which are implicitly assumed to have the same kind of impact on the two populations:

$$\kappa_t^{R*} = \theta_t + \kappa_{t-1}^{R*} + \delta_t \qquad \text{(RWD modified with variable drift } \theta_t, \text{ i.e. structural changes)}$$

$$\kappa_t^{R} = \kappa_t^{R*} + N_t\, Y_t \qquad \text{(RWD further modified with mortality jumps } N_t\, Y_t\text{)}$$

$$\theta_t = (2\theta, \theta, \theta/2)' \qquad \left(\text{structural changes' transition matrix } \begin{pmatrix} 0.99 & 0.01 & 0 \\ 0.01 & 0.98 & 0.01 \\ 0 & 0.01 & 0.99 \end{pmatrix}\right)$$

$$\Pr(N_t = 0) = 0.99,\ \Pr(N_t = 1) = 0.01 \qquad \text{(frequency of a mortality jump)}$$

$$Y_t \sim \text{Normal}(30, 10^2) \qquad \text{(severity of a mortality jump)}$$

in which the parameter values are very approximately deduced from the frequencies and severities of such historical incidents as the World Wars, Spanish Flu, and past structural changes in mortality improvement. These additional features can be switched off by simply setting θ_t as a constant and N_t as zero.

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

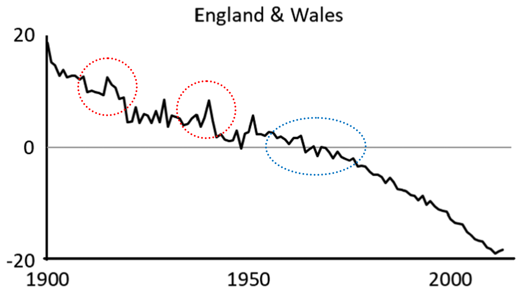

Figure 6. Mortality index of English and Welsh male population under CAE + Cohorts model.

Table 16. Levels of longevity risk reduction (in % of initial longevity risk) using models with additional features deduced from historical incidents.

100,000 males	CMI pensioners (basic materials; normal retirement; lower pension)			IMD group (most deprived areas)		
Additional feature	SD	99.5% VaR	99.5% ES	SD	99.5% VaR	99.5% ES
M7-M5/CAE	65/77	63/74	63/73	64/69	72/68	70/68
Extra model 3	77	74	74	67	64	64
+ Structural	83	85	84	74	80	79
+ Jump	80	72	71	70	64	65
+ Structural and jump	91	91	91	77	78	78
+ Larger structural on both groups	86	90	89	80	87	86
+ Larger structural on book only	79	75	75	72	71	70

Note: CMI, Continuous Mortality Investigation; IMD, index of multiple deprivation; VaR, value-at-risk; ES, expected shortfall.

Table 16 illustrates that the inclusion of structural changes increases the risk reduction estimates with regard to the tail risk measures by more than 10% in magnitude, but the integration of mortality jumps produces very little changes in risk reduction. As structural changes are assumed to have an ongoing impact on both populations in the model, the two populations' mortality levels would move more consistently over the long term, and so demographic basis risk would reduce. In contrast, the temporary effects of mortality jumps do not have much influence on the risk reduction levels. Considering the importance of structural changes, we further consider two more scenarios: (a) $\theta_t = (3\theta, \ \theta, \ \theta/2)'$; and (b) $\kappa_t^R - \kappa_t^{B*} = b_0 + b_1\left(\kappa_{t-1}^R - \kappa_{t-1}^{B*}\right) + \omega_t$, $\eta_t = (0.2\theta, \ 0, \ 0)'$, $\iota_t = \iota_{t-1} + \eta_t$, $\kappa_t^B = \kappa_t^{B*} + \iota_t$, in which η_t and θ_t are independent. The first one has a bigger drop in mortality than the previous cases when there is a structural change, and the second one involves a possibility of a greater structural impact on the book population than on the reference population, where the model becomes non-coherent. In the first case, the risk reduction estimates are larger than previously, because the impact of structural changes on both populations are bigger at the same time. In the second case, the risk reduction levels are lower, as there is a possibility of long-term deviations between the two populations which lead to higher demographic basis risk.

As mentioned earlier, the M7-M5 and CAE+Cohorts models are coherent, under which the projected book-to-reference ratio of future mortality rates at each age converges to a constant in the long term. But in the short term, the two projected trends could deviate from each other, the extent to which

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

Table 17. Levels of longevity risk reduction (in % of initial longevity risk) using different sets of time series processes.

100,000 males	CMI pensioners (basic materials; normal retirement; lower pension)			IMD group (most deprived areas)		
Reference component/book component	SD	99.5% VaR	99.5% ES	SD	99.5% VaR	99.5% ES
(Original)						
M7-M5 (MRWD/VAR(1))	65	63	63	64	72	70
CAE + Cohorts (RWD/AR(1))	77	74	73	69	68	68
(1)						
M7-M5 (RWD/VAR(1))	62	60	60	60	67	69
(2)						
M7-M5 (correlated MRWD and VAR(1))	66	63	61	60	66	64
CAE + Cohorts (correlated RWD and AR(1))	79	76	75	66	65	66
(3)						
M7-M5 (VARIMA(1,1,0)/VAR(1))	56	53	52	52	62	64
CAE + Cohorts (ARIMA(1,1,0)/AR(1))	72	69	68	62	63	63
(4)						
M7-M5 (MRWD/VAR(2))	45	55	47	35	58	35
CAE + Cohorts (RWD/AR(2))	75	73	72	69	72	71
(5)						
M7-M5 (MRWD/BRW)	56	36	37	44	41	37
CAE + Cohorts (RWD/RW)	13	6	7	20	19	15

Note: CMI, Continuous Mortality Investigation; IMD, index of multiple deprivation; VaR, value-at-risk; ES, expected shortfall; MRWD, multivariate random walk with drift; VAR, vector autoregressive process; RWD, random walk with drift; AR, autoregressive; VARIMA, vector autoregressive integrated moving average; ARIMA, autoregressive integrated moving average; BRW, bivariate random walk without drift; RW, random walk without drift.

depends on how quickly the fitted time series process in the book component converges in the projections. For instance, under the CAE+Cohorts model, the projected values of κ_t^B reach the constant $\varphi_0/(1-\varphi_1)$ faster if φ_1 is closer to zero. Moreover, under the two models, the fitted time series processes in the reference component generate unbounded future variability, while those in the book component produce bounded future variability. The former variability would dominate gradually while the latter would reduce in its influence. Consequently, the extent of demographic basis risk would decrease over time in the future simulations.

To further investigate the significance of times series modelling assumptions, we now consider five possible variations. The first one is to use univariate RWD processes in the reference component under the M7-M5 model. The second alternative is to remove the independence assumption between the time series error terms of the two components. The third one is to use integrated autoregressive processes in the reference component. The fourth is to select a higher order of two for the autoregressive processes in the book component. The final alternative is to apply random walk *without* drift processes in the book component.

Table 17 gives the level of longevity risk reduction for each set of time series processes. When univariate RWD processes rather than the MRWD are used in the reference component under the M7-M5 model, the risk reduction levels are slightly lower, though the differences seem to be immaterial. When the time series error terms of the two components are treated as correlated instead of independent, the changes in the risk reduction estimates are mostly negligible. It appears that both of these correlation assumptions do not have a significant impact on the computation of hedge effectiveness.

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

If the MRWD (RWD) in the reference component under the M7-M5 (CAE+Cohorts) model is replaced by the vector autoregressive integrated moving average VARIMA(1,1,0) process (ARIMA (1,1,0)), the risk reduction estimates drop by about 10% (5%) in magnitude. Though the integrated autoregressive processes, like the random walk processes, generate unbounded future variability, the precise levels of their variability across time would be different to those of the random walk processes, leading to different risk reduction levels. Moreover, the projected paths of the integrated autoregressive processes would take some time to converge to the ultimate linear trends and the speed of convergence depends on the values of the autoregressive parameters computed.

When a higher order of two is used for the VAR process in the book component under the M7-M5 model, the risk reduction estimates are much smaller. Though the VAR(2) is still weakly stationary, its speed of convergence is slower than that of the original VAR(1), because of the longer lag and greater autoregressive effects. Thus the projected trends of the two populations would deviate for a longer period before the final convergence, resulting in lower risk reduction levels. On the other hand, the AR(2) in the book component under the CAE+Cohorts model still coverages quickly, so there is not much change in risk reduction.

When the autoregressive processes in the book component are replaced by the random walk *without* drift processes (bivariate BRW or univariate RW), the two mortality projection models are still coherent because of the flat projected trends of the book component. But Table 17 shows that the risk reduction estimates then become very small. The random walk processes produce unbounded future variability, and so the two populations' mortality movements could diverge significantly in the future simulations, especially in the long term. It can be seen again that the assumed behaviour of simulated future variability of the book component is critical in the calculation of longevity risk reduction.

7. Concluding Remarks

In this paper, we have conducted an extensive study on modelling longevity basis risk and measuring longevity risk reduction in index-based swap hedging. We first examine the past mortality rates and improvements based on the CMI, ONS, and HMD data sets. Then we derive in detail the fitting procedures of the M7-M5 and CAE+Cohorts models and the bootstrapping method. After defining the level of longevity risk reduction, we investigate a range of hedging scenarios including open or closed pension plans of a single cohort or multiple cohorts with different portfolio sizes. In particular, we use index-based longevity swaps to construct the longevity hedge. The major finding is that the risk reduction levels are often around 50%–80% for a large portfolio, whereas the risk reduction estimates are usually smaller than 50% for a small portfolio. While the precise level of risk reduction depends on the specific hedging scenario being considered, index-based longevity hedging looks more effective for a larger pension plan and under a more tailored hedging scheme.

Finally, we carry out a thorough sensitivity testing on the hedging results by making a variety of changes to the original model settings and assumptions, and also the time series processes. The most significant modelling assumptions and settings include the behaviour of simulated future variability of the book component, portfolio size, speed of reaching coherence between the two populations, data size and characteristics, simulation method, and mortality structural changes. Relatively, the other conditions tested have limited influence on the calculated hedging results. In building an index-based longevity hedge for a pension plan or annuity portfolio, one should contemplate these factors and perform adequate testing on their potential impact on hedge effectiveness. Further research is needed when more data of longer periods and for different types of pension plans can be collected in the future.

Acknowledgements

Much of this work forms part of the longevity basis risk research project (Phase 2) completed by Macquarie University and sponsored by the Institute and Faculty of Actuaries and the Life and Longevity Markets Association. The authors would like to thank the editor and the referees for their valuable comments and suggestions which enhance the presentation of this paper. The authors thank the CMI and the ONS for their support in providing their data sets. The authors also thank Kam Kuen (Kenny) Mok, Sixian (Alice) Tang, Kenneth Wong, and Jia (Jacie) Liu for their excellent research assistance.

References

Brouhns, N., Denuit, M. & Vermunt, J.K. (2002). A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics*, **31**, 373–393.

Cairns, A.J.G., Blake, D. & Dowd, K. (2006). A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. *Journal of Risk and Insurance*, **73**(4), 687–718.

Cairns, A.J.G., Blake, D. & Dowd, K. (2008). Modelling and management of mortality risk: a review. *Scandinavian Actuarial Journal*, **2008**(2–3), 79–113.

Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D., Epstein, D., Ong, A. & 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), 1–35.

Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D. & Khalaf-Allah, M. (2011). Bayesian stochastic mortality modelling for two populations. *ASTIN Bulletin*, **41**(1), 29–59.

Carter, L.R. & Lee, R.D. (1992). Modeling and forecasting US sex differentials in mortality. *International Journal of Forecasting*, **8**, 393–411.

Comité Européen des Assurances. (2007). Solvency II Glossary. CEA Groupe Consultatif, Brussels.

Coughlan, G.D., Khalaf-Allah, M., Ye, Y., Kumar, S., Cairns, A.J.G., Blake, D. & Dowd, K. (2011). Longevity hedging 101: a framework for longevity basis risk analysis and hedge effectiveness. *North American Actuarial Journal*, **15**(2), 150–176.

Dowd, K. & Blake, D. (2006). After VaR: the theory, estimation, and insurance applications of quantile-based risk measures. *Journal of Risk and Insurance*, **73**(2), 193–229.

Haberman, S., Kaishev, V., Millossovich, P., Villegas, A., Baxter, S., Gaches, A., Gunnlaugsson, S. & Sison, M. (2014). Longevity basis risk: a methodology for assessing basis risk. Cass Business School and Hymans Robertson LLP, Institute and Faculty of Actuaries (IFoA) and Life and Longevity Markets Association (LLMA), London, UK.

Human Mortality Database (HMD) (2017). University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany). Available online at the address www.mortality.org [accessed 6-Apr-2017].

Koissi, M.C., Shapiro, A.F. & Högnäs, G. (2006). Evaluating and extending the Lee-Carter model for mortality forecasting: bootstrap confidence interval. *Insurance: Mathematics and Economics*, **38**, 1–20.

Lee, R.D. & Carter, L.R. (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association*, **87**(419), 659–671.

Li, J. (2014). A quantitative comparison of simulation strategies for mortality projection. *Annals of Actuarial Science*, **8**(2), 281–297.

Li, J. & Haberman, S. (2015). On the effectiveness of natural hedging for insurance companies and pension plans. *Insurance: Mathematics and Economics*, **61**, 286–297.

Li, J.S.H. & Hardy, M.R. (2011). Measuring basis risk in longevity hedges. *North American Actuarial Journal*, **15**(2), 177–200.

Li, N. & Lee, R. (2005). Coherent mortality forecasts for a group of populations: an extension of the Lee-Carter method. *Demography*, **42**(3), 575–594.

Tsai, J.T., Tzeng, L.Y. & Wang, J.L. (2011). Hedging longevity risk when interest rates are uncertain. *North American Actuarial Journal*, **15**(2), 201–211.

Tsay, R.S. (2002). *Analysis of Financial Time Series*. John Wiley & Sons, New York, USA.

Villegas, A.M., Haberman, S., Kaishev, V.K. & Millossovich, P. (2017). A comparative study of two-population models for the assessment of basis risk in longevity hedges. *ASTIN Bulletin*, **47**(3), 631–679.

Zhou, R., Li, J.S.H. & Tan, K.S. (2013). Pricing standardized mortality securitizations: a two-population model with transitory jump effects. *Journal of Risk and Insurance*, **80**(3), 733–774.

Appendix

The parameters in the reference component of the M7-M5 model are first computed using the updating equation $\theta^* = \theta - \frac{\partial l}{\partial \theta} \Big/ \frac{\partial^2 l}{\partial \theta^2}$ (Brouhns *et al*., 2002) and the following iterative updating scheme:

(a) Set the initial parameter values of $\kappa_{t,1}^R$, $\kappa_{t,2}^R$, $\kappa_{t,3}^R$, and γ_c^R to zero and calculate $\hat{\eta}_{x,t}^R$ and $\hat{m}_{x,t}^R$ for all x and t, which are the fitted values of $\eta_{x,t}^R$ and $m_{x,t}^R$.

(b) Update $\kappa_{t,1}^R$ for all t.

(c) Recalculate $\hat{\eta}_{x,t}^R$ and $\hat{m}_{x,t}^R$ for all x and t and then update $\kappa_{t,2}^R$ for all t.

(d) Recalculate $\hat{\eta}_{x,t}^R$ and $\hat{m}_{x,t}^R$ for all x and t and then update $\kappa_{t,3}^R$ for all t.

(e) Recalculate $\hat{\eta}_{x,t}^R$ and $\hat{m}_{x,t}^R$ for all x and t and then update γ_c^R for all c.

(f) Adjust $\kappa_{t,1}^R$, $\kappa_{t,2}^R$, $\kappa_{t,3}^R$, and γ_c^R for all t and c to incorporate the three identifiability constraints.

(g) Recalculate $\hat{\eta}_{x,t}^R$ and $\hat{m}_{x,t}^R$ for all x and t and then calculate l^R.

(h) Repeat steps (b) to (g) until the increase in l^R is less than 10^{-11}.

Table A.1 lists the various equations we have derived for implementing this updating scheme.

Taking the computed parameters $\hat{\kappa}_{t,1}^R$, $\hat{\kappa}_{t,2}^R$, $\hat{\kappa}_{t,3}^R$, and $\hat{\gamma}_c^R$ from above as given (i.e. conditional likelihood), the parameters in the book component of the M7-M5 model are estimated via the updating equation again and the iterative updating scheme below:

(a) Set the initial parameter values of $\kappa_{t,1}^B$ and $\kappa_{t,2}^B$ to zero and calculate the fitted values $\hat{\eta}_{x,t}^B$ and $\hat{m}_{x,t}^B$ for all x and t.

(b) Update $\kappa_{t,1}^B$ for all t.

(c) Recalculate $\hat{\eta}_{x,t}^B$ and $\hat{m}_{x,t}^B$ for all x and t and then update $\kappa_{t,2}^B$ for all t.

(d) Recalculate $\hat{\eta}_{x,t}^B$ and $\hat{m}_{x,t}^B$ for all x and t and then calculate l^B.

(e) Repeat steps (b) to (d) until the increase in l^B is less than 10^{-11}.

Empirically, we find that the tolerance level of 10^{-11} is often sufficient to ensure the parameters converge in the iterative estimation process.

The parameters in the reference component of the CAE+Cohorts model are calculated by the same updating equation and the iterative updating scheme below:

(a) Set the initial parameter values of α_x^R as the average (over time) logit mortality rate of the reference population observed at age x, β_x^R as $\frac{1}{\text{no. of ages}}$, and κ_t^R and γ_c^R as zero. Then calculate the fitted values $\hat{\eta}_{x,t}^R$ and $\hat{m}_{x,t}^R$ for all x and t.

(b) Update α_x^R for all x.
(c) Recalculate $\hat{\eta}_{x,t}^R$ and $\hat{m}_{x,t}^R$ for all x and t and then update κ_t^R for all t.
(d) Recalculate $\hat{\eta}_{x,t}^R$ and $\hat{m}_{x,t}^R$ for all x and t and then update β_x^R for all x.
(e) Recalculate $\hat{\eta}_{x,t}^R$ and $\hat{m}_{x,t}^R$ for all x and t and then update γ_c^R for all c.
(f) Adjust α_x^R, β_x^R, κ_t^R, and γ_c^R for all x, t, and c to incorporate the first four identifiability constraints.
(g) Recalculate $\hat{\eta}_{x,t}^R$ and $\hat{m}_{x,t}^R$ for all x and t and then calculate l^R.
(h) Repeat steps (b) to (g) until the increase in l^R is less than 10^{-11}.

Table A.2 provides the equations we have deduced for running the updating scheme.

Conditioning on the calculated $\hat{\alpha}_x^R$, $\hat{\beta}_x^R$, $\hat{\kappa}_t^R$, and $\hat{\gamma}_c^R$, the parameters in the book component of the CAE+Cohorts model are estimated by the updating equation and the iterative updating scheme as follows:

(a) Set the initial parameter values of α_x^B as the average (over time) logit mortality rate of the book population observed at age x minus $\hat{\alpha}_x^R$, and κ_t^B as zero. Then calculate the fitted values $\hat{\eta}_{x,t}^B$ and $\hat{m}_{x,t}^B$ for all x and t.
(b) Update α_x^B for all x.
(c) Recalculate $\hat{\eta}_{x,t}^B$ and $\hat{m}_{x,t}^B$ for all x and t and then update κ_t^B for all t.
(d) Adjust α_x^B and κ_t^B for all x and t to incorporate the last identifiability constraint.
(e) Recalculate $\hat{\eta}_{x,t}^B$ and $\hat{m}_{x,t}^B$ for all x and t and then calculate l^B.
(f) Repeat steps (b) to (e) until the increase in l^B is less than 10^{-11}.

Table A.3 gives the BIC values of fitting the M7-M5 and CAE+Cohorts models to the various data sets. In each cell, the left figure is the BIC value for the reference component, and the right figure is the BIC value for the book component. For all the data sets (except the book component of the least deprived areas), the M7-M5 model produces the lowest BIC values, but the differences are rather small. We have also examined the residuals by age, year, and cohort, and overall there are no significant patterns in the residuals. Note that the BIC and the residuals examination provide a useful way to assess how well the models are fitted to the data, but the fitted models do not necessarily produce accurate forecasts. Table A.4 lists the parameter estimates of the different time series processes. As expected, all the drift terms are negative, representing overall mortality improvements. Moreover, all the values of φ_1 are smaller than one in magnitude, which means that the fitted AR(1) processes are weakly stationary.

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

Table A.1. Equations derived for iterative updating scheme under M7-M5 model.

Updating equation for reference component	
$\frac{\partial l^R}{\partial \kappa^R_{t,i}} = \sum_x \left(\frac{d^R_{x,t}}{m^R_{x,t}} - e^R_{x,t}\right)\frac{\partial m^R_{x,t}}{\partial \kappa^R_{t,i}}$	$\frac{\partial^2 l^R}{\partial\left(\kappa^R_{t,i}\right)^2} = \sum_x \left(d^R_{x,t}\frac{m^R_{x,t}\frac{\partial^2 m^R_{x,t}}{\partial\left(\kappa^R_{t,i}\right)^2} - \left(\frac{\partial m^R_{x,t}}{\partial \kappa^R_{t,i}}\right)^2}{\left(m^R_{x,t}\right)^2} - e^R_{x,t}\frac{\partial^2 m^R_{x,t}}{\partial\left(\kappa^R_{t,i}\right)^2}\right)$
$\frac{\partial l^R}{\partial \gamma^R_{t-x}} = \sum_{\substack{\text{cohort}\\ t-x}} \left(\frac{d^R_{x,t}}{m^R_{x,t}} - e^R_{x,t}\right)\frac{\partial m^R_{x,t}}{\partial \gamma^R_{t-x}}$	$\frac{\partial^2 l^R}{\partial\left(\gamma^R_{t-x}\right)^2} = \sum_{\substack{\text{cohort}\\ t-x}} \left(d^R_{x,t}\frac{m^R_{x,t}\frac{\partial^2 m^R_{x,t}}{\partial\left(\gamma^R_{t-x}\right)^2} - \left(\frac{\partial m^R_{x,t}}{\partial \gamma^R_{t-x}}\right)^2}{\left(m^R_{x,t}\right)^2} - e^R_{x,t}\frac{\partial^2 m^R_{x,t}}{\partial\left(\gamma^R_{t-x}\right)^2}\right)$
$\frac{\partial m^R_{x,t}}{\partial \kappa^R_{t,1}} = q^R_{x,t}$	$\frac{\partial^2 m^R_{x,t}}{\partial\left(\kappa^R_{t,1}\right)^2} = \exp\left(-\eta^R_{x,t}\right)\left(q^R_{x,t}\right)^2$
$\frac{\partial m^R_{x,t}}{\partial \kappa^R_{t,2}} = (x-\bar{x})q^R_{x,t}$	$\frac{\partial^2 m^R_{x,t}}{\partial\left(\kappa^R_{t,2}\right)^2} = (x-\bar{x})^2\exp\left(-\eta^R_{x,t}\right)\left(q^R_{x,t}\right)^2$
$\frac{\partial m^R_{x,t}}{\partial \kappa^R_{t,3}} = \left((x-\bar{x})^2-\sigma^2\right)q^R_{x,t}$	$\frac{\partial^2 m^R_{x,t}}{\partial\left(\kappa^R_{t,3}\right)^2} = \left((x-\bar{x})^2-\sigma^2\right)^2\exp\left(-\eta^R_{x,t}\right)\left(q^R_{x,t}\right)^2$
$\frac{\partial m^R_{x,t}}{\partial \gamma^R_{t-x}} = q^R_{x,t}$	$\frac{\partial^2 m^R_{x,t}}{\partial\left(\gamma^R_{t-x}\right)^2} = \exp\left(-\eta^R_{x,t}\right)\left(q^R_{x,t}\right)^2$
Updating equation for book component	
$\frac{\partial l^B}{\partial \kappa^B_{t,i}} = \sum_x \left(\frac{d^B_{x,t}}{m^B_{x,t}} - e^B_{x,t}\right)\frac{\partial m^B_{x,t}}{\partial \kappa^B_{t,i}}$	$\frac{\partial^2 l^B}{\partial\left(\kappa^B_{t,i}\right)^2} = \sum_x \left(d^B_{x,t}\frac{m^B_{x,t}\frac{\partial^2 m^B_{x,t}}{\partial\left(\kappa^B_{t,i}\right)^2} - \left(\frac{\partial m^B_{x,t}}{\partial \kappa^B_{t,i}}\right)^2}{\left(m^B_{x,t}\right)^2} - e^B_{x,t}\frac{\partial^2 m^B_{x,t}}{\partial\left(\kappa^B_{t,i}\right)^2}\right)$
$\frac{\partial m^B_{x,t}}{\partial \kappa^B_{t,1}} = q^B_{x,t}$	$\frac{\partial^2 m^B_{x,t}}{\partial\left(\kappa^B_{t,1}\right)^2} = \exp\left(-\eta^B_{x,t}-\eta^R_{x,t}\right)\left(q^B_{x,t}\right)^2$
$\frac{\partial m^B_{x,t}}{\partial \kappa^B_{t,2}} = (x-\bar{x})q^B_{x,t}$	$\frac{\partial^2 m^B_{x,t}}{\partial\left(\kappa^B_{t,2}\right)^2} = (x-\bar{x})^2\exp\left(-\eta^B_{x,t}-\eta^R_{x,t}\right)\left(q^B_{x,t}\right)^2$
Identifiability constraints	
$\begin{pmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\end{pmatrix} = \begin{pmatrix}\sum_c 1 & \sum_c c & \sum_c c^2\\ \sum_c c & \sum_c c^2 & \sum_c c^3\\ \sum_c c^2 & \sum_c c^3 & \sum_c c^4\end{pmatrix}^{-1}\begin{pmatrix}\sum_c \gamma^R_c\\ \sum_c c\gamma^R_c\\ \sum_c c^2\gamma^R_c\end{pmatrix}$	
adjusted $\kappa^R_{t,1} = \kappa^R_{t,1} + \lambda_1 + \lambda_2(t-\bar{x}) + \lambda_3\left((t-\bar{x})^2+\sigma^2\right)$	adjusted $\kappa^R_{t,2} = \kappa^R_{t,2} - \lambda_2 - 2\lambda_3(t-\bar{x})$
adjusted $\kappa^R_{t,3} = \kappa^R_{t,3} + \lambda_3$	adjusted $\gamma^R_{t-x} = \gamma^R_{t-x} - \lambda_1 - \lambda_2(t-x) - \lambda_3(t-x)^2$

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

Table A.2. Equations derived for iterative updating scheme under CAE + Cohorts model.

Updating equation for reference component	
$\frac{\partial l^R}{\partial \alpha_x^R} = \sum_t \left(\frac{d_{x,t}^R}{m_{x,t}^R} - e_{x,t}^R \right) \frac{\partial m_{x,t}^R}{\partial \alpha_x^R}$	$\frac{\partial^2 l^R}{\partial (\alpha_x^R)^2} = \sum_t \left(d_{x,t}^R \frac{m_{x,t}^R \frac{\partial^2 m_{x,t}^R}{\partial (\alpha_x^R)^2} - \left(\frac{\partial m_{x,t}^R}{\partial \alpha_x^R} \right)^2}{(m_{x,t}^R)^2} - e_{x,t}^R \frac{\partial^2 m_{x,t}^R}{\partial (\alpha_x^R)^2} \right)$
$\frac{\partial l^R}{\partial \kappa_t^R} = \sum_x \left(\frac{d_{x,t}^R}{m_{x,t}^R} - e_{x,t}^R \right) \frac{\partial m_{x,t}^R}{\partial \kappa_t^R}$	$\frac{\partial^2 l^R}{\partial (\kappa_t^R)^2} = \sum_x \left(d_{x,t}^R \frac{m_{x,t}^R \frac{\partial^2 m_{x,t}^R}{\partial (\kappa_t^R)^2} - \left(\frac{\partial m_{x,t}^R}{\partial \kappa_t^R} \right)^2}{(m_{x,t}^R)^2} - e_{x,t}^R \frac{\partial^2 m_{x,t}^R}{\partial (\kappa_t^R)^2} \right)$
$\frac{\partial l^R}{\partial \beta_x^R} = \sum_t \left(\frac{d_{x,t}^R}{m_{x,t}^R} - e_{x,t}^R \right) \frac{\partial m_{x,t}^R}{\partial \beta_x^R}$	$\frac{\partial^2 l^R}{\partial (\beta_x^R)^2} = \sum_t \left(d_{x,t}^R \frac{m_{x,t}^R \frac{\partial^2 m_{x,t}^R}{\partial (\beta_x^R)^2} - \left(\frac{\partial m_{x,t}^R}{\partial \beta_x^R} \right)^2}{(m_{x,t}^R)^2} - e_{x,t}^R \frac{\partial^2 m_{x,t}^R}{\partial (\beta_x^R)^2} \right)$
$\frac{\partial l^R}{\partial \gamma_{t-x}^R} = \sum_{\substack{\text{cohort} \\ t-x}} \left(\frac{d_{x,t}^R}{m_{x,t}^R} - e_{x,t}^R \right) \frac{\partial m_{x,t}^R}{\partial \gamma_{t-x}^R}$	$\frac{\partial^2 l^R}{\partial (\gamma_{t-x}^R)^2} = \sum_{\substack{\text{cohort} \\ t-x}} \left(d_{x,t}^R \frac{m_{x,t}^R \frac{\partial^2 m_{x,t}^R}{\partial (\gamma_{t-x}^R)^2} - \left(\frac{\partial m_{x,t}^R}{\partial \gamma_{t-x}^R} \right)^2}{(m_{x,t}^R)^2} - e_{x,t}^R \frac{\partial^2 m_{x,t}^R}{\partial (\gamma_{t-x}^R)^2} \right)$
$\frac{\partial m_{x,t}^R}{\partial \alpha_x^R} = q_{x,t}^R$	$\frac{\partial^2 m_{x,t}^R}{\partial (\alpha_x^R)^2} = \exp\left(-\eta_{x,t}^R\right) \left(q_{x,t}^R\right)^2$
$\frac{\partial m_{x,t}^R}{\partial \kappa_t^R} = \beta_x^R q_{x,t}^R$	$\frac{\partial^2 m_{x,t}^R}{\partial (\kappa_t^R)^2} = (\beta_x^R)^2 \exp\left(-\eta_{x,t}^R\right) \left(q_{x,t}^R\right)^2$
$\frac{\partial m_{x,t}^R}{\partial \beta_x^R} = \kappa_t^R q_{x,t}^R$	$\frac{\partial^2 m_{x,t}^R}{\partial (\beta_x^R)^2} = (\kappa_t^R)^2 \exp\left(-\eta_{x,t}^R\right) \left(q_{x,t}^R\right)^2$
$\frac{\partial m_{x,t}^R}{\partial \gamma_{t-x}^R} = q_{x,t}^R$	$\frac{\partial^2 m_{x,t}^R}{\partial (\gamma_{t-x}^R)^2} = \exp\left(-\eta_{x,t}^R\right) \left(q_{x,t}^R\right)^2$
Updating equation for book component	
$\frac{\partial l^B}{\partial \alpha_x^B} = \sum_t \left(\frac{d_{x,t}^B}{m_{x,t}^B} - e_{x,t}^B \right) \frac{\partial m_{x,t}^B}{\partial \alpha_x^B}$	$\frac{\partial^2 l^B}{\partial (\alpha_x^B)^2} = \sum_t \left(d_{x,t}^B \frac{m_{x,t}^B \frac{\partial^2 m_{x,t}^B}{\partial (\alpha_x^B)^2} - \left(\frac{\partial m_{x,t}^B}{\partial \alpha_x^B} \right)^2}{(m_{x,t}^B)^2} - e_{x,t}^B \frac{\partial^2 m_{x,t}^B}{\partial (\alpha_x^B)^2} \right)$
$\frac{\partial l^B}{\partial \kappa_t^B} = \sum_x \left(\frac{d_{x,t}^B}{m_{x,t}^B} - e_{x,t}^B \right) \frac{\partial m_{x,t}^B}{\partial \kappa_t^B}$	$\frac{\partial^2 l^B}{\partial (\kappa_t^B)^2} = \sum_x \left(d_{x,t}^B \frac{m_{x,t}^B \frac{\partial^2 m_{x,t}^B}{\partial (\kappa_t^B)^2} - \left(\frac{\partial m_{x,t}^B}{\partial \kappa_t^B} \right)^2}{(m_{x,t}^B)^2} - e_{x,t}^B \frac{\partial^2 m_{x,t}^B}{\partial (\kappa_t^B)^2} \right)$
$\frac{\partial m_{x,t}^B}{\partial \alpha_x^B} = q_{x,t}^B$	$\frac{\partial^2 m_{x,t}^B}{\partial (\alpha_x^B)^2} = \exp\left(-\eta_{x,t}^B - \eta_{x,t}^R\right) \left(q_{x,t}^B\right)^2$
$\frac{\partial m_{x,t}^B}{\partial \kappa_t^B} = \beta_x^R q_{x,t}^B$	$\frac{\partial^2 m_{x,t}^B}{\partial (\kappa_t^B)^2} = (\beta_x^R)^2 \exp\left(-\eta_{x,t}^B - \eta_{x,t}^R\right) \left(q_{x,t}^B\right)^2$
Identifiability constraints	
$\tilde{\alpha}_x^R = \alpha_x^R + \beta_x^R \frac{1}{\text{no. of years in ref}} \sum_t \kappa_t^R + \frac{1}{\text{no. of cohorts in ref}} \sum_c \gamma_c^R$	$\tilde{\beta}_x^R = \frac{1}{\sum_x \beta_x^R} \beta_x^R$
$\tilde{\kappa}_t^R = \sum_x \beta_x^R \left(\kappa_t^R - \frac{1}{\text{no. of years in ref}} \sum_t \kappa_t^R \right)$	$\tilde{\gamma}_{t-x}^R = \gamma_{t-x}^R - \frac{1}{\text{no. of cohorts in ref}} \sum_c \gamma_c^R$
$g = \frac{\sum_t (t-\bar{t}) \tilde{\kappa}_t^R}{\sum_t (t-\bar{t})^2}$ $\quad \bar{t} = \frac{1}{\text{no. of years in ref}} \sum_t t$	$h = -\frac{\sum_x 1 \sum_c (c-\bar{c}) \tilde{\gamma}_c^R}{\sum_c (c-\bar{c})^2}$
adjusted $\alpha_x^R = \tilde{\alpha}_x^R + \frac{h}{\sum_x 1} (x-\bar{x})$	adjusted $\beta_x^R = \frac{g}{g-h} \tilde{\beta}_x^R - \frac{h}{\sum_x 1 (g-h)}$
adjusted $\kappa_t^R = \frac{g-h}{g} \tilde{\kappa}_t^R$	adjusted $\gamma_{t-x}^R = \tilde{\gamma}_{t-x}^R + \frac{h}{\sum_x 1} (t-x-\bar{t}+\bar{x})$
adjusted $\alpha_x^B = \alpha_x^B + \beta_x^R \frac{1}{\text{no. of years in book}} \sum_t \kappa_t^B$	adjusted $\kappa_t^B = \kappa_t^B - \frac{1}{\text{no. of years in book}} \sum_t \kappa_t^B$

https://doi.org/10.1017/S1748499518000179 Published online by Cambridge University Press

Table A.3. Bayesian information criterion values of fitting M7-M5 and CAE + Cohorts models to different data sets.

Groups	M7-M5	CAE + Cohorts
Basic materials	13,040/2,848	13,089/2,915
Industrials	13,040/2,929	13,089/3,037
Consumer goods	13,040/2,702	13,089/2,786
Consumer services	13,040/3,042	13,089/3,133
Utilities	13,040/2,600	13,089/2,686
Financials	13,040/2,065	13,089/2,133
Local authority	13,040/2,630	13,089/2,700
Most deprived areas	13,040/4,439	13,089/4,562
Second most deprived areas	13,040/4,248	13,089/4,352
Third most deprived areas	13,040/4,187	13,089/4,217
Fourth most deprived areas	13,040/4,211	13,089/4,254
Least deprived areas	13,040/4,191	13,089/4,167

Table A.4. Estimated parameters of various time series processes.

	M7-M5		CAE + Cohorts	
Groups	$d_1/d_2/d_3$	$\varphi_{1,0}/\varphi_{1,1}/\varphi_{1,2}/\varphi_{2,0}/\varphi_{2,1}/\varphi_{2,2}$	d	φ_0/φ_1
Basic materials	−0.0244/0.000453/ 0.0000239	−0.0138/−0.0658/−3.670/0.00445/ 0.0124/0.308	−0.700	−0.102/0.0990
Industrials	−0.0244/0.000453/ 0.0000239	0.00988/0.424/−3.837/0.00151/ 0.0299/0.834	−0.700	0.436/0.693
Consumer goods	−0.0244/0.000453/ 0.0000239	−0.0175/0.201/−3.525/0.00351/ 0.0180/0.492	−0.700	0.551/0.217
Consumer services	−0.0244/0.000453/ 0.0000239	0.0268/0.470/2.863/0.000645/ 0.0194/0.277	−0.700	0.267/0.386
Utilities	−0.0244/0.000453/ 0.0000239	−0.0748/−0.147/−1.162/0.00520/ −0.0109/−0.0101	−0.700	−0.228/−0.173
Financials	−0.0244/0.000453/ 0.0000239	−0.252/0.150/8.024/0.0152/0.0102/ −0.385	−0.700	0.619/0.327
Local authority	−0.0244/0.000453/ 0.0000239	−0.0450/0.489/−3.033/0.000419/ −0.0626/−0.119	−0.700	−0.267/0.126
Most deprived areas	−0.0244/0.000453/ 0.0000239	−0.0244/0.899/−5.115/−0.0132/ 0.0104/0.288	−0.700	0.209/0.885
Second most deprived areas	−0.0244/0.000453/ 0.0000239	0.0224/0.843/−1.341/−0.0138/ 0.0672/−0.743	−0.700	0.0881/0.726
Third most deprived areas	−0.0244/0.000453/ 0.0000239	−0.0151/0.331/0.849/0.00153/ −0.0735/0.615	−0.700	−0.00947/0.296
Fourth most deprived areas	−0.0244/0.000453/ 0.0000239	−0.0272/0.803/0.146/0.0103/ −0.0316/−0.219	−0.700	−0.0208/0.712
Least deprived areas	−0.0244/0.000453/ 0.0000239	−0.0655/0.714/−0.806/0.00991/ −0.0115/0.201	−0.700	−0.116/0.653