

Part 1

Longevity risk and stochastic models

Wenyu Bai

Quantitative Analyst, Redington Partners LLP

Rodrigo Leon-Morales

Investment Consultant, Redington Partners LLP

Muqiu Liu

Quantitative Analyst, Redington Partners LLP

Yi Wang

Quantitative Analyst, Redington Partners LLP

Abstract

A single 'best' model to predict mortality and understand longevity risk exists only in theory. However, our research and tests conducted on four stochastic mortality models reveals the strengths and weaknesses of each model's predictive capabilities while simultaneously analyzing how parameters can be set to appropriately smooth and massage the input data. However, just being able to model longevity risk is not enough. This paper examines how stochastic modeling is better than deterministic modeling, how longevity risk can be incorporated into an overall liability driven investment (LDI) strategy, and how understanding longevity risk can improve the management of defined benefit pension schemes in the U.K.

Longevity risk and stochastic models

In a defined-benefit (DB) pension scheme benefit levels are determined in advance of retirement and guaranteed irrespective of how underlying funds are invested [Barr and Diamond (2006), Barr (2006)]. Benefits are usually calculated by accounting for the number of years an individual was employed plus the salary level of that employee during part or all of the time employed [Barr and Diamond (2006)]. However, due to the decline in equity markets between 2001 and 2003, late recognition of significant increases in life expectancy, and the increasingly stringent pension regulations and accounting standards, the accumulated pension liabilities for DB schemes in the U.K. now represent a significant financial burden for many corporations.

In this environment a large risk facing DB scheme providers is longevity risk, which is “the risk that members of some reference population might live longer on average than anticipated” [Blake et al. (2006)]. Longevity is a real and substantial risk facing pension funds today and being able to understand and quantify it is becoming increasingly important to trustees and sponsors.

Deterministic versus stochastic models

Traditionally, pension scheme actuaries used deterministic mortality tables, which only provide one estimate of mortality, in the process of valuing pension liabilities. However, to accurately quantify longevity risk, just using the traditional actuarial model is not sufficient. In an attempt to create more accurate projections, actuaries carry out sensitivity tests based on the deterministic mortality assumption, by reasonably shifting the future mortality improvement rate up or down, or by adding one year of life expectancy to all ages to observe the magnitude of change. However, deterministic mortality analysis fails to provide a clear picture of the full distribution of the liabilities associated with and impacted by longevity risk. To better understand and predict longevity risk, research and development efforts are now focused on stochastic mortality models.

Longevity risk

There are two types of longevity risk: systematic longevity risk and longevity basis risk. Systematic longevity risk refers to the part of the risk that is non-diversifiable and longevity basis risk is defined as the difference in mortality rates of the reference population against the actual mortality rates of the population analyzed. From a modeling perspective, the longevity basis risk is implied by the difference between the calibration data and the mortality rates specific to the scheme. To calculate longevity risk accurately it is preferable to use scheme specific data to reduce longevity basis risk and, in its absence, sector mortality data should be taken as a proxy. Examples of variables that contribute to basis risk are geographic location and income. When hedging against a population-wide index, longevity basis risk is greater for small, undiversified schemes. Next it is important to understand age, period, and cohort effect.

Age effect

Age effect describes the relationship between age and mortality rates. The probability of dying increases as people become older and these mortality trends exhibit an exponential behavior. An analysis of England and Wales data¹ reveals that the mortality rate for a 38-year-old in 2003 was 0.001308, 68% higher than a 20-year-old in the same year. However, an 85-year-old male has a mortality rate of 0.120463, and is seven times more likely to die than a male who is twenty years younger.

Period effect

The period effect captures the impact that time-specific events have on the number of deaths experienced in a population, including effects such as general health status of the population, availability of health services, and critical weather conditions [Olivieri (2007)]. The time period can span a single year or multiple years. Overall, the period effect reveals a decreasing trend in realized mortality rates². However, during a specific period, there can be a sudden jump in mortality rates, like the increases observed as a result of World War II.

¹ Data was provided by the Office of National Statistics.

² Based on England and Wales data from 1920 to 2003, from HMD.

Longevity risk and stochastic models

Cohort effect

The cohort effect factor captures the influence of year of birth on mortality improvement rates. The GAD has found that “higher than average rate of improvement is a special feature of generations born between 1925 and 1945 (which centered on the generation born in 1931). It is not yet understood precisely why the members of the generation born about 1931 have been enjoying so much lower death rates throughout adult life than the preceding generation”[GAD Report (1995)]. This has been the subject of recent research, but more research is needed to determine how long the cohort effect will last and to what extent it will affect the whole population.

Models analyzed

Traditional actuarial views have traditionally favored the use of deterministic models, using ‘best estimates.’ However, this methodology depends on the modeler’s opinions and cannot quantify the uncertainty around future mortality rates. Stochastic models are more suitable to fully capture the uncertainty lying around mortality improvements and provide a whole distribution of possible outcomes that helps in the decision making process. This section compares four of the leading stochastic mortality models.

Introduction to four stochastic mortality models

The models studied were developed using different underlying assumptions and parameter constraints, but all attempt to stochastically quantify the impact of longevity risk. We built and tested four models from two model families: the Lee-Carter family and CBD family.

The Lee-Carter model is the simplest model we studied and was developed by Professors Ronald Lee and Lawrence Carter. This model has become the “leading statistical model of mortality forecasting in the demographic literature” in the U.S. [Deaton and Paxson (2004)]. Lee and Carter originally calibrated their model to use U.S. mortality data from 1933-1987. Girosi and King (2007) note that the model is “now being applied to all-cause and cause-specific mortality data from many countries and time periods, and all well beyond

the application for which it was designed.” Nevertheless, the Lee-Carter model has thus far failed to fit Australian and U.K. mortality data.

Because of the inability of the original Lee-Carter model to fit these countries’ data, Steven Haberman and Arthur Renshaw designed the Lee-Carter Extension model which laid the groundwork for the development of a wider class of generalized, parametric, non-linear models. The Lee-Carter Extension model allows for the modeling and extrapolation of age-specific cohort effects, something the original Lee-Carter model is unable to model.

The Cairns, Blake, and Dowd (CBD) model was developed by three Professors in the U.K.: Professors Andrew Cairns, of Heriot-Watt University, David Blake, of Cass Business School, and Kevin Dowd, of Nottingham University Business School. The CBD model was developed for and tested using data from males living in England and Wales and has yet to be tested with data from other countries. However, the model has already been taken up widely by actuaries in Germany and is currently being investigated by the U.K.’s Continuous Mortality Investigation Bureau [Pensions Institute (2007)].

The CBD Extension 2 and Extension 3 models are still under review by the designers, who are determining how the models should be calibrated and used for forecasting purposes. In January 2008 they will release a discussion paper detailing findings of their current research.

Figure 1 shows the formulae for the four models tested and the two models that are currently under development (CBD

Model	Formula
Lee-Carter	$\text{Log } m(t,x) = \beta_x^{(1)} + \beta_x^{(2)} k_t^{(2)}$
Lee-Carter Extension	$\text{Log } m(t,x) = \beta_x^{(1)} + \beta_x^{(2)} k_t^{(2)} + \beta^{(3)} \gamma_{t-x}^{(3)}$
CBD	$\text{Logit } q(t,x) = k_t^{(1)} + k_t^{(2)} (x - \bar{x})$
CBD Extension 1	$\text{Logit } q(t,x) = k_t^{(1)} + k_t^{(2)} (x - \bar{x}) + \gamma_{t-x}^{(3)}$
CBD Extension 2	$\text{Logit } q(t,x) = k_t^{(1)} + k_t^{(2)} (x - \bar{x}) + k_{t-x}^{(3)} [(x - \bar{x})^2 - \sigma_x^{(2)}] + \gamma_{t-x}^{(4)}$
CBD Extension 3	$\text{Logit } q(t,x) = k_t^{(1)} + k_t^{(2)} (x - \bar{x}) + \gamma_{t-x}^{(3)} (x_c - x)$

Figure 1 - Formula for the stochastic mortality models

Longevity risk and stochastic models

Extension 2 and 3). Note that for all models the $\beta_x^{(i)}$ functions reflect age effects, the $k_t^{(i)}$ functions reflect period effects, and the $\gamma_c^{(i)}$ functions reflect cohort effects, with $c = t-x$.

Research methodology

The four models tested were calibrated using two samples in order to analyze the sensitivity of each to sample size and to determine its ability to make accurate projections. In the first test the four models were calibrated with population mortality data from the Human Mortality Database (HMD) for ages 20 to 100 using the period 1920 to 1960. Mortality rates were volatile during this period and the data also includes the impact of World War II. We carried out stochastic projections from 1961 to 2003 for the ages of 20 to 100. The projections were then compared against the actual mortality experienced during the period to test the projection accuracy of the models.

In the second test the four models were calibrated for the data using ages 20-100 for the period from 1963 to 1983. Unlike mortality rates in the first test, the mortality rates during this 20-year period were relatively smooth and show improvement over time. Stochastic projections from 1984 to 2003 were carried out for ages 20 to 100 and the actual mortality experience during the period was compared to the test results to evaluate the accuracy of the model's forecasting capability.

Both test results were then compared and the life expectancy and probability of dying was calculated for ages: 25, 35, 45, 55, 65, 75, and 85. The model results were then compared to actual data to determine the accuracy of each model evaluated. After calibrating the models against the 1920-1960 and 1963-1983 periods, 1000 simulations for ages 20-100 were performed for each year between 1961-2003 and 1984-2003 respectively.

Evaluation criteria

For the purpose of evaluating the test results, we calculated the mean squared error (MSE), the sign test, the test statistic, and the Bayesian information criterion (BIC). The MSE is

used to determine the goodness of fit to the original data and calculates the squared difference between realized and projected values. It is one of the key metrics used to rank a models' forecasting accuracy. It is defined as: $MSE(x, T) = [\sum_{i=1-T}^T (\hat{y}_{x,i} - y_{x,i})^2] \div T$, where T is the total number of projected years, and $\hat{y}_{x,i}$ and $y_{x,i}$ are respectively the projected and actual mortality for age x at time i . For the purpose of this research, the MSE was calculated for all ages using following equation: $MSE(T) = \sum_{x=20-100} (x, T)$. The sign test was used to determine whether the projections were consistently under- or overestimated when compared to the actual values.

The outliers test is used to study the model's capability to replicate tail outcomes. If the tails follow the same distribution as the one proposed by each model, then it is reasonable to assume to not expect more outliers than those predicted by the confidence level. If the number of outliers is substantially more than expected, this means the model is underestimating the 'fatness' of the tails of the true underlying distribution of the data.

The BIC test is a statistical criterion for model selection and to determine the optimal number of free parameters. What we look for is parsimony, this means to find a good balance between the sources of uncertainty and the number of parameters. The data used to test the four models was provided by the Human Mortality Database³ (HMD), and included the mortality data for the population of England and Wales for the period 1920 to 2003.

Model testing results

The overall fitting results for the Lee-Carter, the Lee-Carter Extension, CBD and CBD Extension 1 models are shown in Figure 2. With regards to BIC, the Lee-Carter Extension model exhibits the best parsimony for both the 40- and 20-year data sample cases.

In the 40-year test, the MSE reveals that the Lee-Carter model provides the best projection result in each age group and time horizon, with the Lee-Carter Extension model exhib-

3 Data in the HMD is supplied by the Office of National Statistics (ONS).

Longevity risk and stochastic models

	Based on data sample from 1920 to 1960			Based on data sample from 1963 to 1983		
	BIC	MSE	sign test	BIC	MSE	sign test
Lee-Carter	-85517.16	0.00021626	3.38	-12236.11	0.00002981	6.52
Lee-Carter Extension	-40296.28	0.00144592	8.51	-10939.88	0.00028208	19.28
CBD	-388466.86	0.00262512	3.42	-23922.11	0.00002739	14.77
CBD Extension 1	-147934.84	0.03506693	-5.63	-18294.03	0.00002316	21.56

Figure 2 - Overall fitting statistics

iting a fairly reasonable result, and both CBD models providing poor predictions results, especially the CBD Extension 1 model. This is because from the parameterization process, the parameter representing cohort effect in CBD Extension 1 model is significantly affected by the WWII data, which results in poor forecasting.

In the 20-year test, the MSE reveals that the CBD Extension 1 model is the best overall model in projection accuracy. For almost all ages and time horizons, the Lee-Carter model exhibits better forecasting accuracy in mortality rate than the CBD Extension 1 model. However, our analysis concludes that the Lee-Carter Model is poor in predicting the mortality rates in older age groups, with CBD Extension 1 model providing the most accurate predictions for these groups.

It is also important to use the sign test to evaluate the normality of the results. The test statistics reveals that with both sets of data used nearly all models predict a biased mortality rate. This means that the mortality rates are overestimated

compared with the observations and that life expectancies are underestimated by the models in both samples, with the exception being the CBD Extension 1 model in the 40-year test. As the projection result is biased, it is necessary to do a more detailed analysis to understand the significance of these deviations and where are they come from.

For each model we projected 3484 future mortality rates over the 40-year projection period and 1620 future mortality rates over the 20-year projection period. The 90% confidence intervals for each of these mortality rates were also estimated with each model. When the realized mortality rates were compared with the projected confidence intervals, 348 outliers were expected in the 40-year test and 162 in the second because these outliers should be randomly distributed in the projected tables. Figure 3 shows that for both projection periods, the numbers of outliers is greater than expected. The exception was the CBD Extension 1 model, which produced less outliers than expected, indicating that this model overestimates the volatility of the mortality rates.

	1961-2003 projection			1981-2003 projection		
	Based on data sample from 1920 to 1960			Based on data sample from 1963 to 1983		
	Total no. of values projected	Total no. of outliers	Outliers age range	Total no. of values projected	Total no. of outliers	Outliers age range
Lee-Carter	3483	711	70-80 (few) and 80-100 (most)	1620	418	25-35 and 50-70
Lee-Carter Extension	3483	1109	78-100	1620	905	50-70 and 80-100
CBD	3483	31	20-25	1620	635	before 65
CBD Extension 1	3483	2188	all but 60-75	1620	545	before 55

Figure 3 - Analysis of outliers by year

Longevity risk and stochastic models

In the 40-year test for Lee-Carter model there were 711 outliers with almost half falling in the 80-100 year age group and very few in the 70-80 year age group. The data suggests that the Lee-Carter model systematically overestimated the mortality rates and is not correctly estimating the mortality improvements for older ages. The Lee-Carter model exhibited similar results in the 20-year test projections, with most of the outliers concentrated in the 20-35 and 50-70 age groups.

The analysis of the Lee-Carter Extension model produced 1109 outliers in total for the 40-year test, with most outliers concentrated in the 78-100 age group. Here, like the Lee-Carter model, the Lee-Carter Extension is overestimating mortality rates for this age group. Similar results were observed in the 20-year test, with the outliers concentrated in the 20-35, 50-70, and 80-100 age groups.

By design, the CBD models are more sensitive to variations in the sample data, particularly the three-factor CBD Extension Model. For example, in the 40-year sample where data from World War II is included, the estimation and projection of model parameters are disrupted and the CBD models cannot give a reasonable projection. While the CBD two-factor model only produces 31 outliers in the first test⁴, the CBD Extension 1 model produces 2,188 outliers in total and only produces reasonable projections on the variation range for the age groups 60-75. However, the models' performance improved when more stable mortality data from the 20-year test was used. Most of the outliers for both the CBD and CBD Extension 1 models are concentrated in the 20-50 age groups, with only a limited number of outliers represented in older ages.

The overall conclusions that can be drawn from the data tests are that the Lee-Carter models provide more accurate projections of future mortality rates for younger ages but systematically overestimate the mortality rates and fail to project a reasonable confidence interval for older ages. In contrast, CBD models show accurate projections of future

mortality rates and project a confidence interval which covers most realized mortality rates for old ages.

In the universe of pension scheme risk management, if the sample data can be chosen properly, the CBD model is the preferred choice because it can better quantify the underlying volatility of future mortality improvements for older age groups.

Benefits of stochastic mortality modeling

The traditional actuarial valuations will only provide an expected cash flow value (mean) and the impact on liabilities of a one year increase in life expectancy. While these two metrics are very useful, they are unable to quantify extreme case scenarios or provide a confidence level for the actual present value of the liabilities. To resolve this problem, stochastic models introduce Value-at-longevity-risk (VaLR).

Figure 4 shows the various metrics of an example pension fund that the CBD stochastic longevity model produces, which are very useful for a trustee or sponsor to determine a scheme's target return over liabilities and to define the optimal investment portfolio.

Value-at-longevity-risk (VaLR₉₅) is defined as the maximum increase of pension liabilities in one year's time solely due to mortality improvements with a 95% probability. In other words, there is 1 in 20 probability that pension liabilities will exceed 8,476 (95% confidence level) by the end of year. The annual value at longevity risk for this sample shown in Figure 4 is equivalent to 4.877% of total pension liability.

Key statistics	CBD model (£Mil)
Mean	8,073
S.d.	250
95%	8,476
5%	7,679
VaLR	394
As % of total liability	4.877%

Figure 4 - Key pension metrics based on CBD model

⁴ A result of broader projected confidence intervals, caused by the large volatility of mortality experiences within the sample.

Longevity risk and stochastic models

Conclusion

The risk management of pension schemes is now a top priority for trustees and corporate sponsors. In the current environment, those responsible for managing the assets and liabilities of pension funds need to implement the right analytic models and hedging strategies to cope with the uncertainty lying around future financial and economic conditions. Solutions to mitigate interest and inflation risk can be found in the capital markets. However, demographic conditions and longevity risk remain largely underestimated. A first step towards understanding longevity risk is setting up the right models to analyze, forecast, and price longevity risk.

This research studied the forecasting capabilities of four stochastic models. The models were calibrated against total population mortality data from England and Wales from 1920-1960 and projected from 1961 to 2003 and then calibrated for the same target population but from 1963-1983 and projected from 1983 to 2003. Note that the period from 1920-1960 includes those born just before and during World War II. The projections were then compared against actual mortality experienced during the period to test the projection accuracy of the models. This group was important because it benefited from a post-war welfare state, as shown by a strong cohort effect as described by the GAD.

The test results show that the Lee-Carter model and Lee-Carter Extension model do not accurately project the range of future mortality rates for individuals over the age of 65. Conversely, the CBD models have the capability of projecting a reasonable confidence interval which covers most of the mortality observations in the age group of individuals 65 and older. Thus these findings suggest that the Cairns, Blake, and Dowd model (CBD) is more relevant in risk management and for generating results that can be used for decision making.

Stochastic models are better suited for risk monitoring and decision making. A stochastic model produces a more complete picture of the impact of mortality improvements on a pension scheme's cash flows and can help better calculate

a scheme's funding position or help refine a hedging strategy. Additionally, stochastic models can be used to conduct extreme case scenario analyses and to fully integrate longevity risk within the overall risk framework of the scheme. While further research is needed to fully understand how longevity risk can be included in an overall liability driven investment strategy, it is clear that stochastic modeling is the best way to capture and quantify longevity risk. Our current research focuses on how to actually integrate longevity modeling into an overall asset liability management framework.

References

- Barr, N., 2006a, "Pensions: overview of the issues," *Oxford Review of Economic Policy*. 22:1, 1-14
- Barr, N., and P. Diamond, 2006, "The economics of pensions," *Oxford Review of Economic Policy*. 22:1, 15-39
- Blake, D., A. J. G. Cairns, K. Dowd, and R. MacMinn, 2006, "Longevity bonds: financial engineering, valuation and hedging," *Journal of Risk and Insurance*, 73, 647-672
- Cummins, J. D., 2004, "Securitization of life insurance assets and liabilities," preprint, The Wharton School, University of Pennsylvania
- Deaton, A., and C. Paxson, 2004 "Mortality, income, and income inequality over time in the Britain and the United States," National Bureau of Economic Research Technical Report 8534
- Girosi, F., and G. King, 2007, "Understanding the Lee-Carter mortality forecasting method," working paper, Harvard University
- Government Actuary's Department, 1995, National population projections 1992-based, H.M.S.O., London
- Human Mortality Database <http://www.mortality.org/> [cited 5 December 2007]
- Olivieri, A., 2007, "The longevity risk in pension products," Cass Business School lecture notes, June
- Pensions Institute, 2007, "Latest research reveals man could live up to 12 years longer by 2050 than currently predicted," Pensions Institute Press Release, Cass Business School, 26 November
- Pensions Regulator, 2007: <http://www.thepensionsregulator.gov.uk/pdf/DavidNorgroveUKPensionsInvest> [cited 30 November]
- Renshaw A. E., and S. Haberman, 2006, "A cohort-based Extension to the Lee-Carter model for mortality reduction factors," *Insurance: Mathematics and Economics*, 38:3, 556-570