Modelling and Management of Longevity Risk

Andrew J.G. Cairns
Maxwell Institute for Mathematical Sciences, Edinburgh, and
Department of Actuarial Mathematics and Statistics
Heriot-Watt University, Edinburgh, EH14 4AS, United Kingdom.
WWW: http://www.ma.hw.ac.uk/∼andrewc
E-mail: A.J.G.Cairns@hw.ac.uk

First version: March 2013
This version: March 14, 2013

Abstract

In this article we review the state of play in the use of stochastic models for the measurement and management of longevity risk. A focus of the discussion concerns how robust these models are relative to a variety of inputs: something that is particularly important in formulating a risk management strategy. On the modelling front much still needs to be done on robust multipopulation mortality models, and on the risk management front we need to develop a better understanding of what the objectives are of pension plans that need to be optimised. We propose a variety of ways forward on both counts.

Keywords: Longevity Risk, Stochastic mortality models, Robustness, Risk management.

1 Introduction

This paper considers recent developments on the modelling and management of longevity risk: the risk that, in aggregate, people live longer than anticipated. There are a number of aspects to this problem. First, we need to develop good models that will help us to measure and understand the risks that will arise in the future, with longevity risk being one out of a number of risks such as interest-rate risk and other market risks. Pension plan trustees and sponsors then need to consider the results of this exercise in relation to the plan's risk stated risk appetite and risk tolerances. Finally, they need to make active risk management decisions on how best to manage the plan's exposure to longevity risk as part of a bigger package of good risk management.

1.1 Structure of the paper

This paper starts with a review of developments in the modelling of longevity risk. We consider how three distinctively different approaches to modelling have 'inter-bred' in recent years and we discuss some difficulties with the most recent and also more complex models. Alongside this we discuss uncertainties in the underlying population data that, to date, has not received much attention from the modelling community but is beginning to cause practitioners some anxiety.

We then move on to discuss the question of robustness. There are many outputs from a modelling exercise, but, here, our ultimate goal is to ensure that a particular model produces recommendations for risk management actions that are robust, and which the end users can understand and trust. Without this endpoint, the efforts of those researchers who do the modelling will be fruitless.

2 Modelling challenges

Recent years have seen the development of a wide range of new stochastic models for future improvements in mortality rates. One element of this paper is to challenge the usefulness of all of these models. Our hypothesis is that developing new models is relatively easy. That is, additional features can easily be added to existing models such as the Lee-Carter model (Lee and Carter, 1992) or the CBD model (Cairns, Blake and Dowd, 2006b) and it is normally straightforward to fit these models to the usual datasets and to get a better fit. However, the question remains as to whether or not this added complexity actually improves our ability to forecast future developments in mortality. Answering this question is much more difficult, if it can be answered at all.

Alongside the modelling and consequent measurement of longevity risk, we also have to think about the management of that risk. The transfer of longevity risk from pension plans to reinsurers, insurers and to the capital markets (for example, hedge funds specialising in insurance-linked securities) is a relatively new phenomenon as plan sponsors have begun to get a better grip on the risks inherent in the running of these plans. This market has been slowly gaining momentum with most activity in the UK, but with large and notable transactions in the Netherlands (e.g. Aegon 2012) and the USA more recently. Again we consider what transactions are easy and which ones are difficult. For an actuarial consultancy, it is easy to recommend a customised longevity swap. This would be part of a package of over-the-counter transactions that hedge out the interest-rate, inflation and longevity risks that are embedded in a portfolio of pensions in payment. Recommending a longevity swap is "easy" from the consultant's perspective because the end result of zero risk is guaranteed (notwithstanding counterparty risk). All that remains is to negotiate a good price for the swap or, perhaps, to conclude that the price is too high and that the plan should wait until market conditions and the plan funding position improve. But is a customised longevity swap actually the best solution? Alternatives do exist in the form of q-forwards and S-forwards (see www.LLMA.org). These are derivative securities whose payoffs are linked to an index of mortality rather than the pension plan's own mortality. As a consquence, therefore, their use gives rise to basis risk. But for many pension plans (the hedgers), some residual risk might be acceptable of the hedge is relatively cheap compared to the customised longevity swap. But many consultants will complete avoid consideration of such contracts for a variety of reasons:

- assessment of basis risk is difficult and, perhaps, beyond the capabilities of the consultant;
- assessment of the risk appetites of the plan trustees and sponsor is difficult;
- communication of the nature of the underlying derivatives (e.g. q-forwards) is difficult (what does a q-forward have to do with long-term survivorship?);
- perceived reputational risk from the consultant's perspective if he/she recommends an index-linked solution that subsequently requires topping up (a customised swap might be suboptimal but the reputational risk is minimal).

A significant issue concerns establishment of the risk appetites of a plan sponsor and trustees. The use of a customised longevity swap seems to be consistent with zero appetite for risk. But the paradox here is that pension plans seem to be left with two parts: the part of the plan that deals with pensions in payment is completely intolerant of risk; and the pre-retirement liabilities and associated assets. Typically, for the latter portion of the plan, trustees and sponsors are apparently happy to continue with a risky, equity driven, investment strategy. This apparent paradox is discussed further in Section 6.

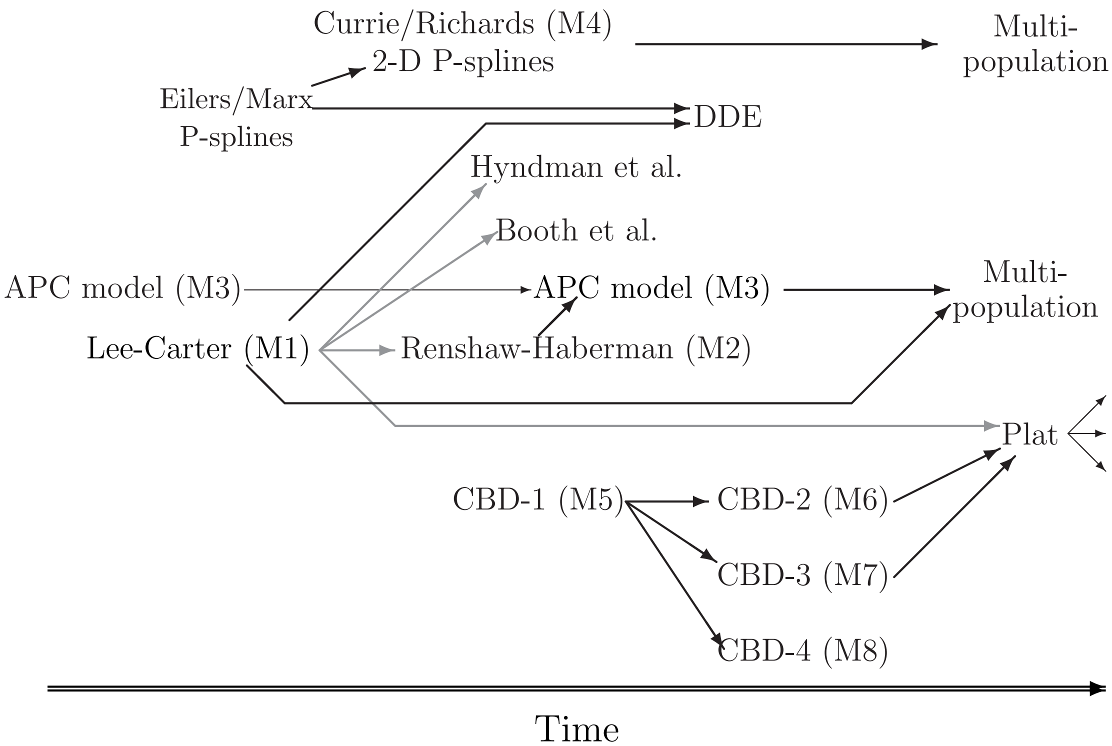

Figure 1: Timeline for the development of stochastic mortality models. Arrows indicate the influence that individual models have had on the development of later generations. (Source: Cairns, 2012.)

With the above discussion in mind, the objective of this paper is to focus minds on the development of a longevity risk management strategy for pension plans and annuity providers that we can have confidence in; that we believe is (close to) optimal; and that we know is robust.

3 Model development: a genealogy

We now review briefly some of the key developments in modelling over the last 20+ years before discussing in a later section where efforts might be focussed in the future on the development of new models (especially in a multifactor setting). We choose to refer here to the modelling 'genealogy' because the majority of new models can be thought of as being modifications (that is, the descendants) of earlier models. This is illustrated in Figure 1.

Models for mortality are typically expressed in terms of the death rate, $m(t,x)$, for age x in year t or the corresponding mortality rate (probability of death), $q(t,x)$. A commonly used approximation that links the two is that $1-q(t,x) \approx \exp[-m(t,x)]$.

Model	formula
M1	$\log m(t,x) = \beta_x^{(1)} + \beta_x^{(2)}\kappa_t^{(2)}$
M2	$\log m(t,x) = \beta_x^{(1)} + \beta_x^{(2)}\kappa_t^{(2)} + \beta_x^{(3)}\gamma_{t-x}^{(3)}$
M3	$\log m(t,x) = \beta_x^{(1)} + \kappa_t^{(2)} + \gamma_{t-x}^{(3)}$
M4	$\log m(t,x) = \sum_{i,j} \theta_{ij} B_{ij}^{ay}(x,t)$
M5	$\text{logit } q(t,x) = \kappa_t^{(1)} + \kappa_t^{(2)}(x-\bar{x})$
M6	$\text{logit } q(t,x) = \kappa_t^{(1)} + \kappa_t^{(2)}(x-\bar{x}) + \gamma_{t-x}^{(3)}$
M7	$\text{logit } q(t,x) = \kappa_t^{(1)} + \kappa_t^{(2)}(x-\bar{x}) + \kappa_t^{(3)}\left((x-\bar{x})^2 - \hat{\sigma}_x^2\right) + \gamma_{t-x}^{(4)}$
M8	$\text{logit } q(t,x) = \kappa_t^{(1)} + \kappa_t^{(2)}(x-\bar{x}) + \gamma_{t-x}^{(3)}(x_c - x)$

Table 1: Formulae for the mortality models: The functions $\beta_x^{(i)}$, $\kappa_t^{(i)}$, and $\gamma_{t-x}^{(i)}$ are age, period and cohort effects respectively. The $B_{ij}^{ay}(x,t)$ are B-spline basis functions and the θ_{ij} are weights attached to each basis function. $\bar{x}$ is the mean age over the range of ages being used in the analysis. $\hat{\sigma}_x^2$ is the mean value of $(x-\bar{x})^2$. (Source: Cairns et al., 2009.)

Stochastic mortality modelling in demography and actuarial work can mainly be traced back to the model of Lee and Carter (1992) (model M1 in Table 1). However, the medical statistics literature does contain the Age-Period-Cohort model (APC) which pre-dates the Lee-Carter model (see, for example, Osmond, 1985). It is only since 2000, that a variety of models have been proposed as alternatives to the Lee-Carter model to address its deficiencies (although, because of its simplicity, the Lee-Carter model does still have its supporters). Some of these new models can be thought of as direct descendants of the Lee-Carter model (such as Hyndman and Ullah, 2007, and Booth et al., 2002), by adding additional age-period effects. Other models had distinctly different roots. Currie, Durban and Eilers (2004) (building on Eilers and Marx, 1996) proposed the use of two-dimensional P-splines (M4 in Table 1). Cairns, Blake and Dowd (2006b) (CBD) proposed a two-factor with parametric age effects in contrast to the fully non-parametric Lee-Carter model (M5 in Table 1).

Basic analysis of underlying mortality data in the early 2000's (Willetts, 2004) revealed patterns in the data related to year of birth that could not be easily explained through the use of age-period models. This gave rise to a number of new models based on the three approaches that built cohort effects into the model: Renshaw and Haberman (2006) building on Lee-Carter (M2); Cairns et al. (2009) building upon CBD(M6, M7 and M8); and Richards et al. (2006) building on Currie et al. (2004) (M4). The growing number of models led to the comprehensive studies of Cairns et al. (2009, 2011a) and Dowd et al. (2011a,b), who used a wide range of criteria to compare different models, as well as providing a framework for developing and analysing other new models in the future. Of the models considered in these comparative studies, several fitted historical data well but M2 and M8 were found to have significant (and apparently insurmountable) problems with robustness (see, also, CMI, 2007) leading to a recommendation that these models are not used in practical work except with extreme caution and only then in expert hands.

Other strands of work have sought to take the best features of the different approaches to create new models. Delwarde et al. (2007) introduced the use of P-splines into the Lee-Carter model. Plat (2009) and Currie (2011) added non-parametric age effect into the CBD family of models (M5, M6 and M7) with the key benefit that these models could be extended to a wider range of ages than was previously recommended by Cairns et al. (2006b, 2009). Plat's work has since been developed further by Börger et al. (2013).

Most recently, new models have begun to emerge that attempt to model mortality in multiple populations, by adapting standard single-population models. So far these have focused on the simpler single-population models, and includes the work of Li and Lee (2005), Cairns et al. (2011b), Li and Hardy (2011), Jarner and Kryger (2011) and Dowd et al. (2011a). Much work remains to be done in this direction, but a better understanding of multipopulation dynamics is central to the development of a vibrant market in longevity transactions.

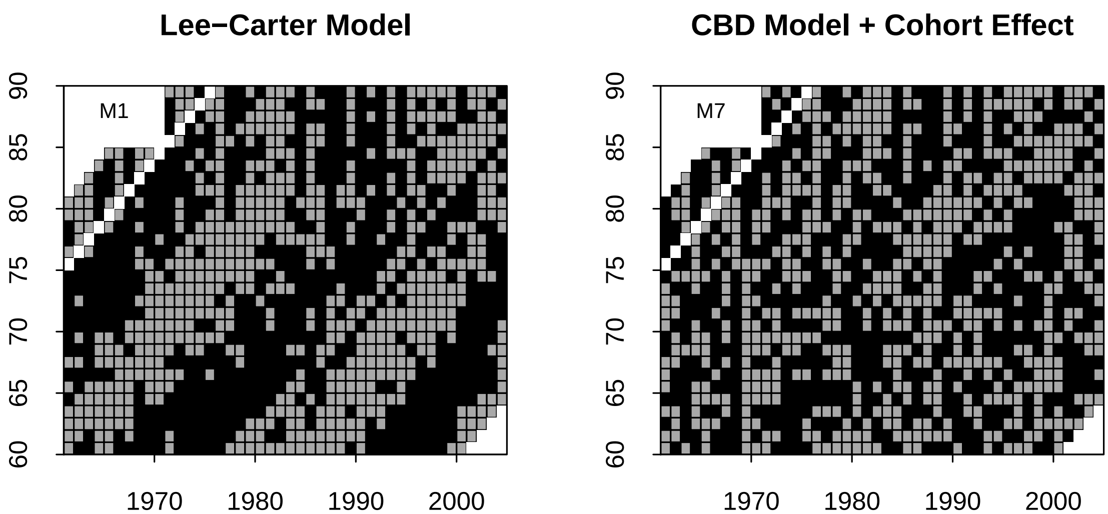

Figure 2: Standardised errors $\epsilon(t,x) = (D(t,x) - \hat{m}(t,x)E(t,x))/\sqrt{\hat{m}(t,x)E(t,x)}$ where the $E(t,x)$ are the exposures, $D(t,x)$ are the actual deaths and the $\hat{m}(t,x)$ are the estimated death rates under the Lee-Carter model (M1; left) and the CBD model, M7, with three period effects and a cohort effect. Black (t,x) cells correspond to $\epsilon(t,x) < 0$, gray cells correspond to $\epsilon(t,x) \geq 0$ and white cells correspond to missing or excluded data. If the model is true then the $\epsilon(t,x)$ should be independent and approximately standard normal. Data: England and Wales males aged 60 to 89 from 1961 to 2005.

Besides the models discussed above, a variety of other approaches have been proposed. Cairns, Blake and Dowd (2006a) reviewed how arbitrage-free frameworks for modelling interest rate risk and credit risk can be adapted to form different frameworks for modelling mortality risk. The models covered in Figure 1 and Table 1 can best be described as 'short-rate' models in the interest-rate context. Of the alternatives proposed, most progress has been made on so-called 'forward rate' models (Miltersen and Persson, 2005, Olivier and Jeffery, 2004, Smith, 2005, Cairns, 2007, Bauer, 2006 and Bauer and Russ, 2006). In a similar spirit, Cairns, Blake and Dowd (2008) describe in more detail the SCOR (Survivor Credit Offered Rate) market model. Compared to the extended family models illustrated in Figure 1, these forward-rate models bring with them greater challenges in terms of complexity and calibration, but they also offer good prospects for efficient market consistent valuation from one time period to the next. Finally, other avenues that concern the use of additional covariates such as smoking prevalence (Kleinow and Cairns, 2011) or income (Kallestrup-Lamb, Blake, Cairns and Dowd, work in progress) are also under consideration, but such approaches are constantly hindered by the lack of good quality data.

So why do we need all of the extra complexity that these models bring? The answer lies with the quality of the fit of the model to historical data. Cairns et al. (2009) compared eight models and found that, using the Bayes Information Criterion, the more complex models (e.g. M7) fitted the historical data much better. Additionally, an analysis of standardised residuals reveals that the simple models such as Lee-Carter and CBD violated key assumptions such as conditional independence of the death count in individual (t, x) cells. Figure 2), for example, shows strong diagonal clusters of gray and black cells (left-hand plot) when, in fact, these should be distributed randomly throughout the plot. This contrasts with the more complex CBD model with a cohort effect (M7) (which includes a cohort effect) where the plot of residuals is much more random (Figure 2, right).

But this raises a potential problem. Models such as Lee-Carter and the basic CBD (Cairns et al., 2006b) are known to be simple and robust, but violate the underlying assumptions when they are fitted to the data (specifically that deaths are conditionally independent and have a Poisson distribution. The model complex models such as the CBD-M7 or Plat (2009) fit much better and satisfy the underlying assumptions. But, as a general rule of thumb, greater complexity brings with it an increased possibility that forecasts are less robust. Backing this up, Dowd et al. (2010,a,b) compared six models and found that complex models that fitted historical data much did not obviously outperform simple models in out-of-sample forecasting (nor did they underperform).

A final problem with more complex models is that the more random processes we have in a single-population model, the more complex it becomes to extend the model to multiple populations.

3.1 Data reliability

Model fitting generally makes the assumption that the exposures data, $E(t, x)$, is accurate. However, for many national datasets and, potentially, smaller specialised sub-populations, it is acknowledged that exposures are estimates and sometimes quite poor estimates of the true values. This issue was mentioned in passing in the discussion of US mortality data in Cairns et al. (2009). More recently, the Office for National Statistics in the UK (ONS, 2012) made significant revisions to estimated exposures from 2001 to 2011 for higher ages in the UK (including England and Wales). The UK carries out population censuses every 10 years (the last being in 2011). Even in the census years population estimates are subject to error, and between censuses the ONS needs to estimate population sizes at each age through estimates of deaths and net migration. In their analysis, Cairns et al. (2009) noted that even for the best fitting models, standardised errors were bigger in magnitude than they ought to be under the conditional Poisson model, and one explanation for this is the fact that exposures are approximations. Indeed, for at least some smaller countries with much better systems in place for estimating population sizes at each age, it seems that the standard mortality models fit better: a fact that might be the result of greater accuracy of the exposures.

3.2 Applications of models

The models themselves have a number of applications.

As a starting point the basic outputs of models need to be communicated to end users in a clear way. Various graphical methods, in particular, have been proposed by Renshaw and Haberman (2006), Cairns et al. (2009, 2011) and Dowd et al. (2010c).

A larger body of papers have sought to consider the pricing of longevity-linked financial contracts. Solvency II and related issues have been discussed by Olivieri and Pitacco (2009) and, with its one-year time horizon, Plat (2010), annuity pricing by Richards and Currie (2009), and pricing in a more general context by Zhou and Li (2013) and Zhou et al. (2011). This includes a requirement to calculate prices or values at future points in time which creates a challenge in its own right: namely that most stochastic mortality models do not give rise to simple analytical formulae for even annuity prices. Some papers, therefore, propose methods for calculating approximate values for key quantities (see, for example, Denuit et al., 2010, Cairns, 2011, Dowd et al, 2011b).

More recent work has focused on the use of models to develop and assess hedging strategies including Dahl et al. (2008), Coughlan et al. (2011), Cairns et al. (2013a), Cairns (2013), Dowd et al. (2011c) and Li and Luo (2012). Much more needs to be done in this direction, in particular, to persuade end users to consider a wider range of risk management options: a topic that is discussed later in this paper.

4 Robustness

A key theme in this paper is the need for robustness in the models, forecasts and decisions that we might take in the measurement and management of longevity risk. If any elements lack robustness, then end users will not have sufficient trust in what is being recommended and, potentially a significantly suboptimal decision will be taken.

The assessment of robustness takes many forms.

4.1 Model fit

Models M1 to M8 in Table 1 consist of combinations of age, period and cohort effects. We wish to know how robust are the estimated age, period and cohort effects are relative to changes in:

- the range of ages used to calibrate the model;
- the range of years (especially adding one new year's data);
- the method of calibration.

Additionally, are estimated age, period and cohort effects robust relative to uncertainties in the estimated exposures (as in subsection 3.1). Where results are found to be sensitive to these choices, is the sensitivity just a manifestation of identifiability constraints (as discussed, for example, by Cairns et al, 2009) or is it a genuine lack of robustness?

The method of calibration relates to the underlying statistical assumptions (for example, the conditional independent Poisson assumption – see, Brouhns et al., 2002; see also Li et al., 2009), is a Bayesian or frequentist approach being taken, how much smoothing has been imposed, and what is the objective being optimised (for example, maximum likelihood or a more simple form of linear regression)?

4.2 Model forecasts

In a similar vein, how robust are stochastic forecasts (both central trajectories and the level of uncertainty around that trend) to changes in

- the range of ages used to calibrate the model;
- the range of years (especially adding one new year's data);
- the method of calibration;
- the choice of stochastic model for simulating future period and cohort effects?

And how robust are forecasts relative to the more general treatment of model and parameter risk and uncertainty in exposures data?

4.3 Business decisions

Finally, and related to the forecasts of future mortality rates, how robust are

- financial variables such as the market consistent value of liabilities, and the prices of, for example, q-forwards;
- risk management metrics (such as hedge effectiveness);
- risk management decisions (such as the choice of hedging instrument and the number of units of that instrument);

relative to the factors listed in the preceeding subsections.

5 Future developments

The preceding sections have revealed a tension between the needs for robustness on the one hand and the temption to add complexity to models to explain better smaller and smaller details of single population data. We will focus here on the development of models that will meet the needs of industry and on development of a better understanding of the objectives that longevity-risk hedgers seek to optimise.

5.1 Modelling

The key challenge on the modelling front is to develop robust multipopulation models. There are several reasons for this.

- Pension plans seek to *measure* accurately trends in mortality rates for their own membership: both central trends and uncertainty around that. In the majority of cases pension plans either have relatively small populations or limited amounts of historical mortality data for their own population, and this makes it difficult to develop a reliable single-population stochastic mortality model. The use of a two-population model means that limited data for the pension plan itself can be augmented by, for example, a much larger national dataset. The use of Bayesian mathods, as in Cairns et al. (2011) means that missing data can be easily dealt with, including earlier years for which pension plan mortality data has been discarded.
- Pension plans seeking to manage their longevity risks need robust multipopulation models that will allow them to compare the various customised and index-linked derivatove solutions. Such models are necessary for both price establishment and comparison as well as the assessment of residual risk (such as basis risk in index-linked hedges).
- Life insurers seeking to measure accurately trends in mortality rates and the uncertainty around them need good multipopulation models because they have exposure, potentially, to many populations"

  - males and females;
  - different contract types (e.g. assurances and annuities);
  - smokers and non-smokers;
  - multinational portfolios.

- Life insurers might bid to take over pension liabilities from pension plans. The underlying risks being transferred need to be measured accurately (i.e. central trend and uncertainty around that) in order to price the deal accurately. This needs a multipopulation model.
- Life insurers, themselves, might seek to transfer longevity risk to third parties, and so the same issues as for paneion plans apply but, perhaps, on a different scale.

As remarked earlier, if a stochastic mortality model has, as its stochastic drivers, additional numbers of processes then this makes extension to two or more populations much more challenging because of the need to consider correlations between all of the driving processes in both populations. Therefore, there is a need to develop a new approach that goes back to basics and focuses on models with fewer period effects in particular. For example, an approach being developed by Cairns et al. (2013b) moves away from the usual assumption that deaths in different (t, x) cells have a conditionally *independent* Poisson distribution. Their approach is to model the different between actual and expected as a mixture of traditional Poisson errors and a new residuals process $R(t, x)$ that allows for correlation between individual cells. As an example, let

$$\log \tilde{m}(t,x) = \beta_1(x) + \kappa_2(t) + \kappa_3(t)(x - \bar{x})$$

be an adaptation of the CBD (Cairns et al., 2006b) and Plat (2009) models. This is used to model the long-term developments in mortality. *Local* mortality adds the residuals process, $R(t, x)$, thus

$$\log m(t,x) = \tilde{m}(t,x) + R(t,x).$$

Lastly, deaths follow the usual Poisson model

$$D(t,x) \sim \text{Poisson}(m(t,x)E(t,x)).$$

With this type of approach, multipopulation modelling should focus only on correlation between the processes driving the long-term $\tilde{m}(t, x)$ processes: that is, assuming that the $R(t, x)$ processes for each population are independent (an assumption that obviously needs verification!).

A different and, perhaps, less radical approach is to start with more complex single population models but reduce the number of correlated processes between populations. An example would be the model M7 in Cairns et al. (2009) which has three period effects and one cohort effect in each population. If we have two populations then a full time series model needs to consider correlations between six period effects (that is, 15 correlation parameters) and between two cohort effects. With three populations the number of correlation parameters starts to become unmanageable. Instead, we can seek to minimise the number of non-zero correlations: for example, is the main correlation between the principal period effects, $\kappa_1(t)$, affecting the level of mortality, while correlations between the slope and curvature period effects between the two populations are negligible. More theoretically, we might seek to establish a correlation between some linear combinations of the period effects with zero correlation otherwise.

The second modelling challenge concerns the treatment of exposures. As remarked earlier, modellers have, in the past, always treated exposures as accurate point estimates or (as in Cairns et al., 2009) by treating specific cohorts as missing data. There is an urgent need to develop a new statistical methodology that considers exposures themselves as being subject to uncertainty. A key question then is to consider whether or not *ex ante* forecasts that assume that exposures are accurate are themselves robust. Secondly, we need to consider how, in individual populations, exposures might, from time to time, be revised up or down. Thes revisions could, potentially, result in significant changes in base mortality tables and also in central trajectories.

6 Risk Appetite

6.1 Derisking glide paths

We will now discuss how a pension plan might choose between the various hedging options. Anecdotal evidence based on recent deals and professional magazines (e.g Khiroya and Penderis, 2012) points to the following situation as being the typical approach that consultants recommend to UK pension plans. Consultants typically refer to a derisking glide path, especially for defined benefit plans that are closed to new members and potentially have no further accrual for existing members. This glide path is characterised by a number of features.

- For pensions in payment, the plan should seek the hedge the liabilities in a way that minimises or even eliminates the risk of deficit for that subset of the pension plan membership. For a fully derisked position this means one of a collection of individual buyouts, a bulk buyout (both of which transfer legal responsibility for payment of the pension to the insurer), a bulk buyin or a customised longevity swap.
- For active members where pensions are linked to future salary increases, the

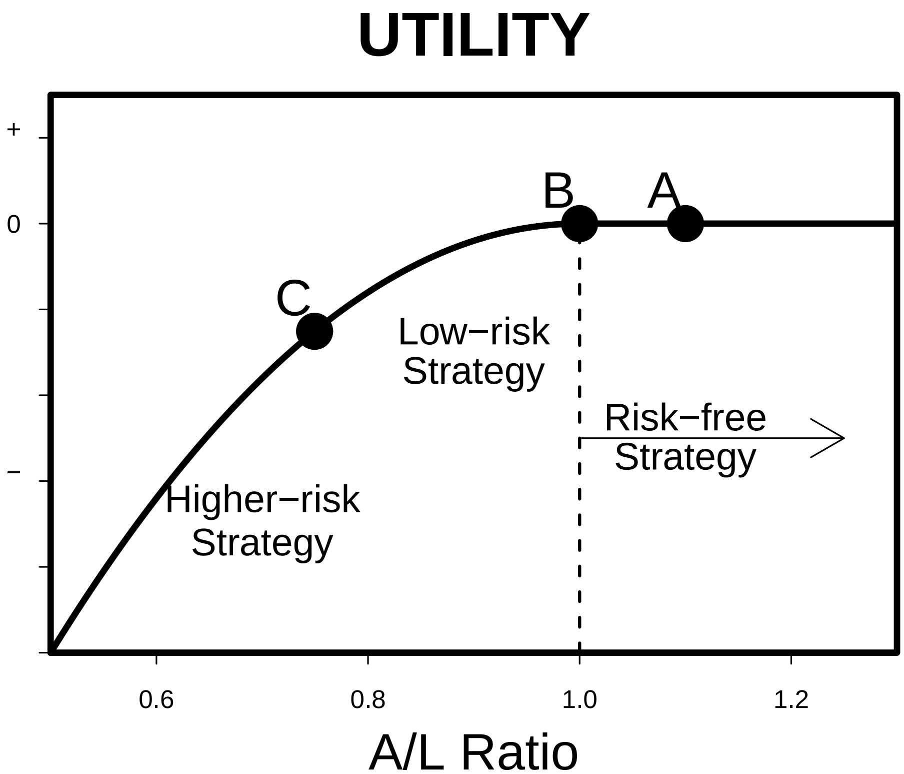

Figure 3: Semi-quadratic utility function for a pension plan. (Source: Cairns, 2012.)

plan continues to invest in a mixed portfolio of risky assets (e.g. 60% equities, 40% bonds).

- Where the plan is in deficit, then derisking activities are deferred until the funding level has improved. In this situation a generally more risky asset strategy is adopted to increase the chances of achieving a fully funded position.
- Intermediate options might be considered for deferred pensioners and also active members if the defined benefit does not include future salary increases. In this case customised buyouts and longevity swaps are potentially very expensive due to significantly elevated levels of longevity risk inherent in such transactions relative to pensions that are already in payment to older plan members.

With this background in mind we have to ask what type of risk appetite or objective do the pension plan trustees have in mind that results in the derisking glide paths described above? We can find a candidate for this in the realm of utility theory. Specifically, we consider a semi-quadratic utility function of the form $u(x) = -(1 - x)^2$ if the funding level, $x < 1$ and $u(x) = 0$ if $x \geq 1$ (see Figure 3).

In some sense, this utility is consistent with the stratgies recommended above to follow a derisking glide path. For pensions in payment, in particular, once the plan is fully funded then derisking means that there is no chance to fall below the bliss point, B, in Figure 3. And if the funding level is below 100% then the plan should adopt a more risky investment strategy until it can get back up to 100% funding, at which point it should derisk as a one-off, irreversible transaction. However, this logic only follows if the plan has no unhedgeable liabilities such as salary risk for active members. In this case it is less clear that 100% removal of risk for one sub-population is actually optimal.

In the same setting with a mixture of member classes, there is an implicit conclusion that it is locally optimal to totally derisk in relation to one population and to maintain a substantially risky strategy for a different sub-population. This might be valid if these were isolated entities but, in reality, they are two parts of one larger entity with a single risk appetite for that entity.

6.2 Size matters

We will now consider other reasons why a pension plan might cosnider alterntives to bulk buyouts and customised longevity swaps.

In Figure 4 we present a stylised view of the relative costs of four options for a pension that contains only pensions in payment.[1]

- Individual buy out. The plan buys individual annuities one by one for its pensioners. In this case the cost does not depend on the size of the plan (that is, the number of members).
- A bulk buyout of the full set of pensioners. This type of transaction enjoys economies of scale, so that the price per unit falls as the size of the plan increases. Additionally, the price will fall with size because sampling risk in the runoff of the liabilities will, relatively, be smaller. There is a minimum size to this type of transaction, below which the receiver (for example, a monoline insurer) would not be interested in taking over the liabilities.
- A customised longevity swap. This is part of a buy in strategy involving additional hedges, for example, against inflation risk in pensions in payment. Customised longevity swaps have a much higher threshold for engagement with the receiver of the longevity risk than bulk buyouts. The price of a longevity swap would also reward scale and reductions in sampling risk and there *might* be a crossover of the price per unit of risk relative to bulk buyout.

[1] For further longevity risk-management options, see Blake et al. (2006), Coughlan et al. (2007), Cairns et al. (2008) and www.llma.org.

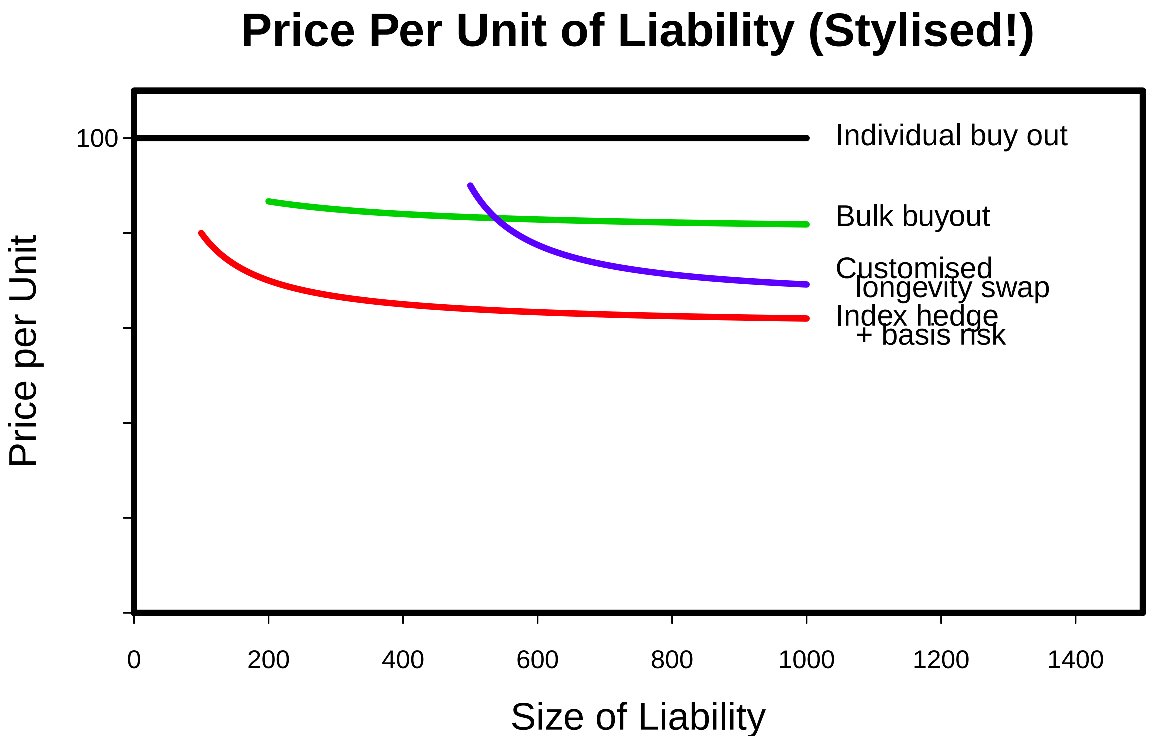

Figure 4: Potential prices per unit for different longevity hedging instruments as a function of the size of a transaction. Quantities and relationships are illustrative only and have no scientific basis. (Source: Cairns, 2012.)

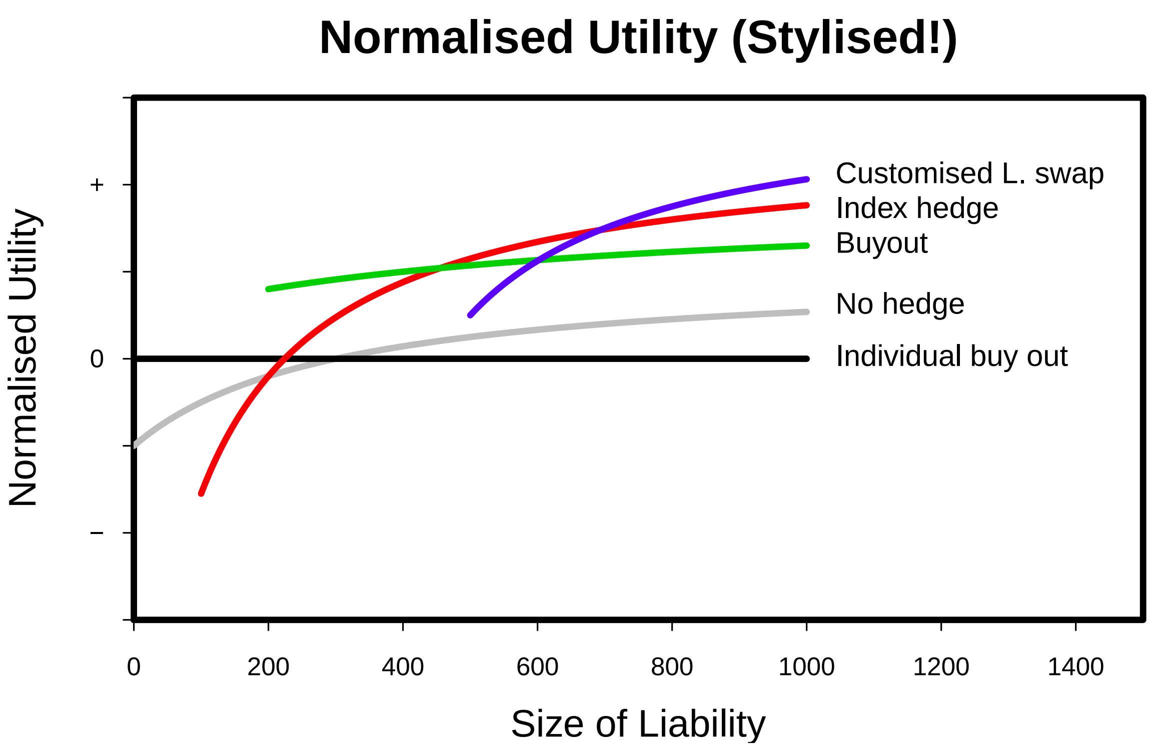

Figure 5: Potential expected utilities for different longevity hedging strategies relative to the individual annuitisation strategy (normalised to have 0 utility). Quantities and relationships are illustrative only and have no scientific basis. (Source: Cairns, 2012.)

- Use of index-linked longevity hedging instruments. This type of transaction has a much lower threshold for engagement (in theory, a single q-forward or S-forward contract). In theory, the price should not reflect the size of the transaction, but, in practice, the expenses related to the contract would push up the price per unit of smaller deals.

Figure 4 and the remarks above point to lower prices for larger plans, but potentially, as we move further up the scale, transactions might become so large that the receivers' appetite for taking on longevity risk diminishes. So the price per unit might actually have to rise in order to balance supply and demand.

Now consider the impact of each of these types of transaction on a pension plan's expected utility. Figure 5 presents a stylised view of this in a way that is consistent with Figure 4, and plots the difference in expected utility of a given strategy relative to the individual buyout strategy. Figure 5 assumes a strictly concave and strictly increasing utility implying that the plan always has some appetite for risk rather than (as in Figure 3) zero appetite for risk above some threshold.

- We include no hedging as one option. The normalised utility increases with scale relative to individual buyout because the plan benefits from lower levels of sampling risk. The two curves cross over because individual buyout includes expenses and a risk premium.
- Bulk buyout and customised longevity swaps achieve essentially the same endpoint as individual buyouts using different vehicles and so the differences between the three simply reflects the different prices per unit of risk and the scale thresholds for bulk buyouts and longevity swaps.
- The curve for an index hedge falls from right to left because of two factors. First, the increasing cost of a smaller transaction size as in Figure 4. Second, the relative level of basis risk that arises with an index-linked transaction rises as the size of the plan gets smaller and this pushes down further the normalised utility.

The way that we have constructed our stylised plot means that the optimal hedge will depend on the size of the pension plan. For small transactions up to 200, individual buyouts are optimal even though index-lonked hedges are available over some of that range. Bulk buyouts take over between 200 and 440, index hedges between 440 and 680, and finally customised longevity swaps are optimal above 680. However, we stress that in practice the bands over which each strategy might be optimal will vary substantially from situation to situation without any guarantee that the order is the same as that presented here or that individual strategies will be optimal at any level of scale (for example, higher levels of risk aversion will push down the utility of the index-hedge relative to the customised transactions).

The point of this example, though, is to show that, particularly if the pension plan has some appetite for risk at all funding levels, then all options should be considered, and that there is no default option that will always come out top. Instead a variety of factors come into play: price per unit of risk as a function of scale; sampling risk; basis risk; and risk aversion.

7 Summary

In this paper we have attempted to contrast aspects of longevity risk measurement and management that are easy versus those that are difficult. Building new and ever more complex models and the recommendation of customised longevity hedges are tasks that are (relatively) easy. In contrast, the development of models that are robust and fit for purpose in a multipopulation setting is much tougher. Robustness, in particular, is a particular criterion that cannot be ignored or glossed over: without a proper analysis of robustness practitioners will not engage with a model or, therefore, use it in the development of risk management strategies.

A rigorous assessment of all of the risk management options including index-linked hedges is also a much tougher call, and this includes a proper prior assessment of the pension plan's risk appetite and risk tolerances.

References

Bauer, D. (2006) “An arbitrage-free family of longevity bonds”, Working paper, University of Ulm. See www.mortalityrisk.org

Bauer, D., and Russ, J. (2006) “Pricing longevity bonds using implied survival probabilities”, Working paper, University of Ulm. See www.mortalityrisk.org

Blake, D., Cairns, A.J.G., and Dowd, K. (2006) Living with mortality: Longevity bonds and other mortality-linked securities. *British Actuarial Journal*, 12: 153-197.

Börger, M., Fleischer, D., and Kuksin. N.2013 Modeling mortality trend under modern solvency regimes. To appear in *ASTIN Bulletin*.

Booth, H., Maindonald, J., and Smith, L. (2002) Applying Lee-Carter under conditions of variable mortality decline. *Population Studies*, 56: 325-336.

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

Cairns, A.J.G. (2007) “A multifactor generalisation of the Olivier-Smith model for stochastic mortality”, In Proceedings of the 1st IAA Life Colloquium, Stockholm, 2007.

Cairns, A.J.G. (2011) Modelling and management of longevity risk: approximations to survival functions and dynamic hedging. *Insurance: Mathematics and Economics*, 49: 438-453.

Cairns, A.J.G. (2013) Robust hedging of longevity risk. To appear in Journal of Risk and Insurance.

Cairns, A.J.G., Blake, D., and Dowd, K. (2006a) "Pricing death: Frameworks for the valuation and securitization of mortality risk", *ASTIN Bulletin,* 36: 79-120.

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

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

Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D., Epstein, D., Ong, A., and Balevich, I. (2009) A quantitative comparison of stochastic mortality models using data from England & Wales and the United States. *North American Actuarial Journal*, 13: 1-35.

Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D., Epstein, D., and Khalaf-Allah, M. (2011a) Mortality density forecasts: an analysis of six stochastic mortality models. *Insurance: Mathematics and Economics*, 48: 355-367.

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

Cairns, A.J.G., Blake, D., Dowd, K., and Coughlan, G.D. (2013a) Longevity hedge effectiveness: A decomposition. To appear in *Quantitative Finance*.

Cairns, A.J.G., Mavros, G., Streftaris, G., and Kleinow, T. (2013b) New directions in the modelling of longevity risk. Presented at the International AFIR-ERM/Life/ASTIN Colloquium, Mexico City, October 2012.

CMI (2007) Stochastic projection methodologies: Lee-Carter model features, example results and implications. *Continuous Mortality Investigation, Working Paper* 25.

Coughlan, G., Epstein, D., Sinha, A., and Honig, P. (2007) q-forwards: Derivatives for transferring longevity and mortality risk. Available at www.lifemetrics.com.

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

Currie, I.D., Durban, M. and Eilers, P.H.C. (2004) "Smoothing and forecasting mortality rates", *Statistical Modelling*, 4: 279-298.

Currie, I.D. (2011) Modelling and forecasting the mortality of the very old. *ASTIN Bulletin*, 41: 419-427.

Dahl, M., Melchior, M., and Møller, T. (2008) On systematic mortality risk and risk minimisation with survivor swaps. *Scandinavian Actuarial Journal*, 2008(2-3): 114-146.

Delwarde, A., Denuit, M., and Eilers, P. (2007) "Smoothing the Lee-Carter and Poisson log-bilinear models for mortality forecasting: A penalised log-likelihood approach", *Statistical Modelling*, 7: 29-48.

Denuit, M., Haberman, S., and Renshaw, A.E. (2010) Comonotonic approximations to quantiles of life annuity conditional expected present values: Extensions to general ARIMA models and comparison with the bootstrap. *ASTIN Bulletin*, 40: 331-349.

Dowd, K., Cairns, A.J.G., Blake, D., Coughlan, G.D., Epstein, D., and Khalaf-Allah, M. (2010a) Evaluating the Goodness of Fit of Stochastic Mortality Models. *Insurance: Mathematics and Economics*, 47: 255-265.

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

Dowd, K., Blake, D., and Cairns, A.J.G. (2010c) Facing up to uncertain life expectancy: The longevity fan charts. *Demography*, 47: 67-78.

Dowd, K., Cairns, A.J.G., Blake, D., Coughlan, G.D., and Khalaf-Allah, M. (2011a) A gravity model of mortality rates for two related populations. *North American Actuarial Journal*, 15: 334-356.

Dowd, K., Blake, D., and Cairns, A.J.G. (2011b) A computationally efficient algorithm for estimating the distribution of future annuity values under interest-rate and longevity risks. *North American Actuarial Journal*, 15: 237-247.

Dowd, K., Blake, D., Cairns, A.J.G., Coughlan, G.D. (2011c). Hedging pension risks with the age-period-cohort two-population gravity model. *Seventh International Longevity Risk and Capital Markets Solutions Conference*, Frankfurt, September 2011.

Eilers, P.H.C., and Marx, B.D. (1996) Flexible smoothing with B-splines and penalties. *Statistical Science*, 11: 89-121.

Hyndman, R.J., and Ullah, S.M. (2007) Robust forecasting of mortality and fertility rates: A functional data approach, *Computational Statistics and Data Analysis*, 51: 4942-4956.

Khiroya, N., and Penderis, M. (2012) Using LDI as part of a glide path de-risking programme. In *The Guide to Risk Reduction, August 2012*, *Professional Pensions*, pp 8-9.

Jarner, S.F., and Kryger, E.M. (2011) Modelling adult mortality in small populations: The SAINT model. *ASTIN Bulletin*, 41: 377-418.

Kleinow, T., and Cairns, A.J.G., (2011) Mortality and smoking prevalence: An empirical investigation in ten developed countries. To appear in *British Actuarial Journal*.

Lee, R.D., and Carter, L.R. (1992) "Modeling and forecasting U.S. mortality", *Journal of the American Statistical Association*, 87: 659-675.

Li, J.S.-H., Hardy, M.R., and Tan, K.S. (2009) Uncertainty in model forecasting: An extension to the classic Lee-Carter approach. *ASTIN Bulletin,* 39: 137-164.

Li, J.S.-H., and Luo, A. (2012) Key q-duration: A framework for hedging longevity risk. *ASTIN Bulletin*, 42: 413-452.

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

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

Miltersen, K.R., and Persson, S.-A. (2005) "Is mortality dead? Stochastic forward force of mortality determined by no arbitrage" Working paper, University of Bergen.

Olivieri, A. and Pitacco, E. (2009) Stochastic mortality: The impact on target capital. *ASTIN Bulletin*, 39: 541-563.

Olivier, P., and Jeffery, T. (2004) "Stochastic mortality models", Presentation to the Society of Actuaries of Ireland: see www.actuaries.ie (Events and Papers).

ONS (2012) Population estimates for England and Wales, Mid-2002 to Mid-2010 Revised (National) *Office for National Statistics*, 13 December 2012.

Osmond, C. (1985) "Using age, period and cohort models to estimate future mortality rates", *International Journal of Epidemiology*, 14: 124-129.

Plat, R. (2009) Stochastic portfolio specific mortality and the quantification of mortality basis risk. *Insurance: Mathematics and Economics*, 45: 123-132.

Plat, R. (2010) One-year value-at-risk for longevity and mortality. Pensions Institute Working Paper PI-1015.

Renshaw, A.E., and Haberman, S. (2006) "A cohort-based extension to the Lee-Carter model for mortality reduction factors", *Insurance: Mathematics and Economics,* 38: 556-570.

Richards, S.J., and Currie, I.D. (2009) Longevity risk and annuity pricing with the Lee-Carter model. *British Actuarial Journal*, 15: 317-343.

Richards, S.J., Kirkby, J.G., and Currie, I.D. (2006) "The importance of year of birth in two-dimensional mortality data", (with discussion) *British Actuarial Journal*, 12: 5-61.

Smith, A.D. (2005) "Stochastic mortality modelling", Workshop on the Interface between Quantitative Finance and Insurance, International Centre for the Mathematical Sciences, Edinburgh. See www.icms.org.uk/archive/meetings/2005/quantfinance/

Willets, R.C. (2004) The cohort effect: Insights and explanations. *British Actuarial Journal*, 10: 833-877.

Zhou, R., and Li, J.S.H. (2010) A cautionary note on pricing longevity index swaps. *Scandiavian Actuarial Journal*, 2013: 1-23.

Zhou, R., Li, J.S.H., and Tan, K.S. (2011) Economic pricing of mortality-linked securities in the presence of population basis risk. *Geneva Papers on Risk and Insurance: Issues and Practice*, 36: 544-566.