Recreating Sustainable Retirement

Resilience, Solvency, and Tail Risk

EDITED BY

Olivia S. Mitchell,
Raimond Maurer, and
P. Brett Hammond

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

First Edition published in 2014

Impression: 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2014940448

ISBN 978–0–19–871924–3

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

Contents

Part III. Regulatory and Political Risk

Part IV. Implications for Plan Sponsors

Chapter 5

Modeling and Management of Longevity Risk

Andrew Cairns

This chapter considers recent developments in the modeling 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 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 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.

We start with a review of developments in the modeling of longevity risk. We consider how three distinctively different approaches to modeling have 'interbred' 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, have not received much attention from the modeling community but are beginning to cause practitioners some anxiety.

We then move on to discuss the question of robustness. There are many outputs from a modeling 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 modeling will be fruitless.

Modeling Challenges

Recent years have seen the development of new stochastic models for future improvements in mortality rates. One element of this chapter 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 et al. 2006*b*), and it is normally straightforward to fit these models to the usual datasets and to get a better fit. However, a question remains as to whether 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 modeling and consequent measurement of longevity risk, we must also think about the management of that risk. The transfer of longevity risk from pension plans to reinsurers, insurers, and the capital markets (for example, hedge funds specializing 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 U.K., but with large and notable transactions in the Netherlands (e.g. Aegon 2013) and the U.S. more recently. Again we consider which transactions are easy and which ones are difficult. For an actuarial consultancy, it is easy to recommend a customized 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 customized 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 consequence, therefore, their use gives rise to basis risk. But for many pension plans (the hedgers), some residual risk might be acceptable if the hedge is relatively cheap compared to the customized longevity swap. But many consultants will completely 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?); and
- Perceived reputational risk from the consultant's perspective if he/she recommends an index-linked solution that subsequently requires topping up (a customized 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 customized 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 and 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 below.

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.

Model Development: A Genealogy

We next review briefly some of the key developments in modeling over the past 20+ years, before discussing in a later section where efforts might be focused in the future on the development of new models (especially in a multifactor setting). We choose to refer here to the modeling '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 5.1.

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 $1-q(t, x) \approx \exp[-m(t, x)]$. Stochastic mortality modeling in demography and actuarial work can mainly be traced back to the model of Lee and Carter (1992) (model M1 in Table 5.1). 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 has been proposed as alternatives to the

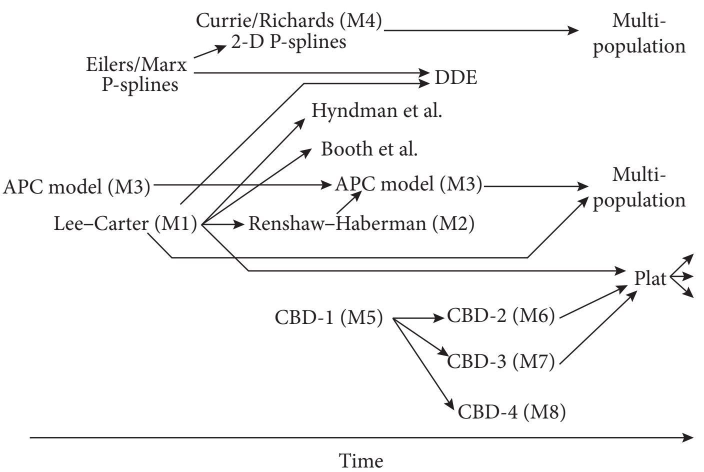

Figure 5.1. Timeline for the development of stochastic mortality models.

Note: Arrows indicate the influence that individual models have had on the development of later generations.

Source: Cairns et al. (2008).

TABLE 5.1 Formulae for the mortality models

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	$\log m(t,x) = \sum_{i,j} \theta_{ij} \mathrm{B}_{ij}^{ay}(x, t)$
M5	$\operatorname{logit} q(t,x) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x})$
M6	$\operatorname{logit} q(t,x) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \gamma_{t-x}^{(3)}$
M7	$\operatorname{logit} q(t,x) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \tilde{x}) + \kappa_t^{(3)}((x - \tilde{x})^2 - \hat{\sigma}_x^2) + \gamma_{t-x}^{(4)}$
M8	$\operatorname{logit} q(t,x) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \gamma_{t-x}^{(3)}(x_c - x)$

Notes: The functions $\beta_x^{(i)}$, $\kappa_t^{(i)}$, and $\gamma_{t-x}^{(i)}$ are age, period, and cohort effects, respectively. The $\mathrm{B}_{ij}^{ay}(x,t)$ 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 $(x - \bar{x})^2$. See text.
Source: Cairns et al. (2009).

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 Booth et al. 2002 and Hyndman and Ullah 2007), by adding additional age-period effects. Other models had distinctly different roots. Currie et al. (2004), building on Eilers and Marx (1996), proposed the use of two-dimensional P-splines (M4 in Table 5.1). Cairns et al. (2006*b*) (CBD) proposed a two-factor model with parametric age effects in contrast to the fully non-parametric Lee–Carter model (M5 in Table 5.1).

Analysis of underlying mortality data in the early 2000s (Willets 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 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, 2011*a*) and Dowd et al. (2011*a, b*), who used a wide range of criteria to compare different models, as well as providing a framework for developing and analyzing other new models in the future. Of the models considered in these comparative studies, several fit historical data well but M2 and M8 were found to have significant (and apparently insurmountable) problems with robustness (see also Continuous Mortality Investigation, CMI 2007), leading to a recommendation that these models not be used in practical work except with extreme caution, and then only 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 a 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. (2006*b*, 2009). Plat's work has since been developed further by Börger et al. (2011).

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. These include the work of Li and Lee (2005), Cairns et al. (2011*b*), Li and Hardy (2011), Jarner and Kryger (2011), and Dowd et al. (2011*a*). 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.

In addition to the models discussed above, a variety of other approaches has been proposed. Cairns et al. (2006*a*) reviewed how arbitrage-free frameworks for modeling interest-rate risk and credit risk can be adapted to form different frameworks for modeling mortality risk. The models covered in Figure 5.1 and Table 5.1 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 (Olivier and Jeffery 2004; Miltersen and Persson 2005; Smith 2005; Bauer 2006; Bauer and Russ 2006; Cairns 2007). In a similar spirit, Cairns et al. (2008) describe in more detail the Survivor Credit Offered Rate (SCOR) market model. Compared to the extended family models illustrated in Figure 5.1, 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 2013) or income (Kallestrup-Lamb et al. 2013), are also under consideration, but such approaches are constantly hindered by the lack of good quality data on relevant covariates.

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 then found that, using the Bayes Information Criterion, the more complex models (e.g. M7) fitted the historical data much better. Additionally, an analysis of standardized residuals reveals that 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 5.2, 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 5.2, right).

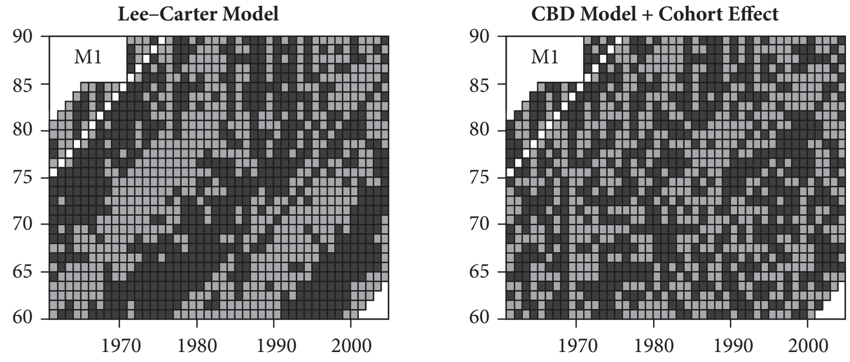

Figure 5.2. Standardized errors for actual versus expected deaths.

Notes: $\epsilon(t,x) = (D(t,x) - \hat{m}(t,x)E(t,x)) / \sqrt{\hat{m}(t,x)E(t,x)}$ where the $E(t,x)$ are the exposures, the $D(t,x)$ 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.

Source: Author's computations from England and Wales data on males aged 60 to 89 from 1961 to 2005.

But this raises a potential problem. Models such as Lee–Carter and the basic CBD approach (Cairns et al. 2006*b*) are known to be simple and robust, but then violate the underlying assumptions when they are fitted to the data (specifically that deaths are conditionally independent and have a Poisson distribution). The more 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 will be less robust. Backing this up, Dowd et al. (2010*a, b*) compared six models and found that complex models that fit historical data much better 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.

Data Reliability

Model fitting generally makes the assumption that the exposures data, $E(t, x)$, are accurate. However, for many national datasets and, potentially, smaller specialized sub-populations, it is acknowledged that exposures are estimates and sometimes quite poor estimates of true values. This issue was mentioned in passing in the discussion of U.S. mortality data in Cairns et al. (2009). More recently, the Office for National Statistics in the U.K. (ONS 2012) made significant revisions to estimated exposures from 2001 to 2011 for higher ages in the U.K. (including England and Wales). The U.K. carries out population censuses every ten years (the most recent 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, standardized errors were bigger in magnitude than they ought to be under the conditional Poisson model. 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.

Applications of Models

The models themselves have a number of applications. As a starting point, the 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*a*), and Dowd et al. (2010*c*).

A larger body of papers has 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 a 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. 2011*b*).

More recent work has focused on the use of models to develop and assess hedging strategies (see Dahl et al. 2008; Coughlan et al. 2011; Dowd et al. 2011*c*; and Li and Luo 2012; Cairns 2013; Cairns et al. 2014). 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 discussed later in this chapter.

Robustness

A key theme in this chapter 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 might be taken. The assessment of robustness takes many forms.

Model Fit

Models M1 to M8 in Table 5.1 consist of combinations of age, period, and cohort effects. We wish to know how robust 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); and the method of calibration. Additionally, it is important to ask whether estimated age, period, and cohort effects are robust relative to uncertainties in the estimated exposures. Where results are found to be sensitive to these choices, it could be that the sensitivity is just a manifestation of identifiability constraints (as discussed, for example, by Cairns et al. 2009) or 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). A Bayesian or frequentist approach might be taken, smoothing might be imposed, and the objective being optimized might differ (for example, maximum likelihood or a more simple form of linear regression).

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; and the choice of stochastic model for simulating future period and cohort effects? Moreover, analysts must explore the robustness of forecasts relative to the more general treatment of model and parameter risk and uncertainty in exposures data.

Business Decisions

Related to the forecasts of future mortality rates, one must ask how robust, relative to the factors discussed above, 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); and risk management decisions (such as the choice of hedging instrument and the number of units of that instrument) are.

Future Developments

The preceding sections have revealed a tension between the need for robustness on the one hand, and the temptation to add complexity to models to better explain smaller and smaller details of single population data on the other. We focus next on the development of models to meet the needs of industry and for a better understanding of the objectives that longevity risk hedgers seek to optimize.

Modeling

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

First, 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 methods, 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 customized and index-linked derivative 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; or 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. the central trend and uncertainty around that) in order to price the deal accurately. This needs a multipopulation model.

Last, life insurers themselves might seek to transfer longevity risk to third parties, and so the same issues as for pension 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. (2013) 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 difference 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.[1] With this type of approach, multipopulation modeling will 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 two cohort effects. With three populations, the number of correlation parameters starts to become unmanageable. Instead, we can seek to minimize the number of non-zero correlations: for example, correlations between the principal period effects, $\kappa 1(t)$ (affecting the level of mortality) might be found to be significant, while correlations between the slope and curvature period effects between the two populations might be negligible. Alternatively, we might seek to establish a correlation between some linear combinations of the period effects with zero correlation otherwise. A second modeling challenge concerns the treatment of exposures. As remarked earlier, modelers have, in the past, always treated exposures as accurate point estimates or (as in Cairns et al. 2009) treated 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. We also must consider how, in individual populations, exposures might from time to time be revised up or down. These revisions could potentially result in significant changes in base mortality tables and also in central trajectories.

Risk Appetite

Derisking Glide Paths

We will now discuss how a pension plan might choose between the various hedging options.[2] Anecdotal evidence based on recent deals and professional magazines (e.g. Khiroya and Penderis 2012) points to one situation as being typical for what consultants recommend to U.K. 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 characterized by a number of features. For pensions in payment, the plan should seek to hedge the liabilities in a way that minimizes 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 buy-in, or a customized longevity swap. For active members where pensions are linked to future salary increases, the plan continues to invest in a mixed portfolio of risky assets (e.g. 60 percent equities, 40 percent 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 customized 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.

Against this background we 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? A candidate for this lies 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 5.3). In some sense, this utility is consistent with the strategies 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 5.3. If the funding level is below 100 percent, then the plan should adopt a more risky investment strategy until it can get back up to 100 percent funding, at which point it should derisk as a one-off, irreversible transaction. But, this logic only follows if the plan has no unhedgeable liabilities such as salary risk for active members. In that case, it is less clear that 100 percent removal of risk for one sub-population is actually optimal.

Now consider the setting where there is a mixture of member classes (e.g. actives, deferred pensioners, and pensioners). An open question is the following: is there a

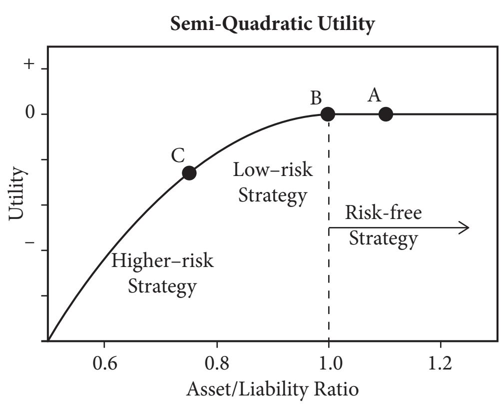

Figure 5.3. Semi-quadratic utility function for a pension plan.

Source: Cairns et al. (2008).

realistic formulation of risk appetite (e.g. a utility function) under which it is locally optimal to totally derisk in relation to one population, and to maintain a substantially risky strategy for a different sub-population? Intuition suggests that it will be difficult, if not impossible, to find such a formulation. Total derisking of one population suggests that the plan sponsor and trustees are totally intolerant of risk. In this case, for active members, the plan should also hedge all hedgeable risks (e.g. price inflation risk and longevity risk), leaving only the residual non-hedgeable risks (e.g. the difference between salary inflation and price inflation). Countering this criticism, one might argue that typical hedging strategies are illiquid and cannot be reversed easily without incurring substantial cost. But, taking this into account, it might still be preferable to gradually derisk the actives' liabilities in a planned series of 'irreversible' hedging transactions.

Size Matters

We will now consider other reasons why a pension plan might consider alternatives to bulk buyouts and customized longevity swaps.[3]

In Figure 5.4 we present a stylized view of the relative costs of four options for a pension that contains only pensions in payment. The four options (relative to inaction) are as follows.

Individual Buyout

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).

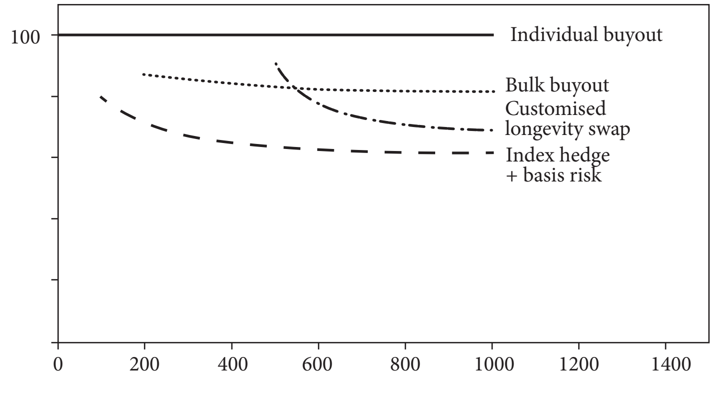

Figure 5.4. Potential prices per unit for different longevity hedging instruments as a function of the size of a transaction.

Note: Quantities and relationships are illustrative only and have no scientific basis.

Source: Cairns et al. (2008).

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 Customized Longevity Swap

This is part of a buy-in strategy involving additional hedges, for example, against inflation risk in pensions in payment. Customized 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.

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.

This list of options is not exhaustive: for further longevity risk-management options, see Blake et al. (2006), Coughlan et al. (2007), Cairns et al. (2008), and <www.llma.org>.

Figure 5.4 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 5.5 presents a stylized view of this in a way that is consistent with Figure 5.4, and plots the difference in expected utility of a given strategy relative to the individual buyout strategy. Figure 5.5 assumes a strictly concave and strictly increasing utility implying that the plan always has some appetite for risk, rather than (as in Figure 5.3) zero appetite for risk above some threshold.

We include no hedging as one option. The normalized 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 customized longevity swaps achieve essentially the same endpoint as individual buyouts using different vehicles, and so the differences between

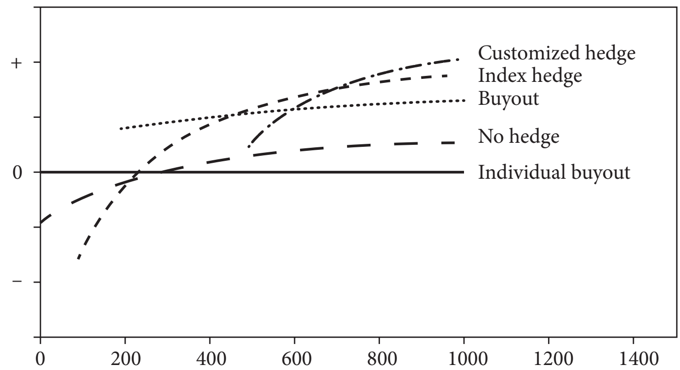

Figure 5.5. Potential expected utilities for different longevity hedging strategies relative to the individual annuitization strategy (normalized to have zero utility).

Note: Quantities and relationships are illustrative only and have no scientific basis.

Source: Cairns et al. (2008).

the three simply reflect 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 5.4. 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 normalized utility.

We have constructed our stylized plot so that the optimal hedge will depend on the size of the pension plan. For small transactions up to 'liability sizes' of 200 (Figures 5.4 and 5.5), individual buyouts are optimal even though index-linked hedges are available over some of that range. Bulk buyouts take over between liability sizes in the range 200 to 440, index hedges between 440 and 680, and finally, customized 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 customized 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 comes into play: price per unit of risk as a function of scale; sampling risk; basis risk; and risk aversion.

Conclusion

In this chapter we 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 customized 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 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.

Acknowledgments

Part of this research was conducted while the author was in receipt of a research grant from Netspar, the Netherlands.

Notes

1. As an example, let log $\mathrm{m} \cdot (t,x) = \beta_1(x) + \kappa_2(t) + \kappa_3(t)(x = \bar{x})$ be an adaptation of the CBD (Cairns et al. 2006*b*) 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 $m(t/x) = \dot{m}(t,x) + R(t,x)$. Lastly, deaths follow the usual Poisson model $D(t,x) \sim Poisson(m(t/x)E(t,x))$.
2. The use of the expression 'glide path' is an interesting one. Relative to a 'flight' path it suggests no further contributions from the sponsor (which is what they would like), but also only limited controls relative to powered flight.
3. An example of this is the Pall Pension Plan longevity hedge for active members transacted with J.P. Morgan LifeMetrics early in 2011.

References

Aegon (2013). 'Aegon Grows Earnings and Sales in Q4; Proposes Increased Final Dividend.' The Hague, Netherlands: Aegon, February 15. <http://www.aegon.com/Documents/aegon-com/Sitewide/Quarterly-results/2012-Q4/EN/2012-Q4-Full-version-English.pdf>.

Bauer, D. (2006) 'An Arbitrage-Free Family of Longevity Bonds,' University of Ulm Working Paper. Ulm, Germany: University of Ulm. <http://www.mortalityrisk.org>.

Bauer, D., and J. Russ (2006). 'Pricing Longevity Bonds Using Implied Survival Probabilities,' University of Ulm Working Paper. Ulm, Germany: University of Ulm. <http://www.mortalityrisk.org>.

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

Börger, M. D. Fleischer, and N. Kuksin (2011). 'Modeling Mortality Trend Under Modern Solvency Regimes,' *ASTIN Bulletin*, 41: 1–38.

Booth, H., J. Maindonald, and L. Smith (2002). 'Applying Lee–Carter Under Conditions of Variable Mortality Decline,' *Population Studies*, 56: 325–336.

Brouhns, N., M. Denuit, and J. K. Vermunt (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 Generalization of the Olivier–Smith Model for Stochastic Mortality,' in *Proceedings of the 1st IAA Life Colloquium*. Stockholm, 2007.

Cairns, A. J. G. (2011). 'Modeling 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,' *Journal of Risk and Insurance*, 80(3): 621–648.

Cairns, A. J. G., D. Blake, and K. Dowd (2006*a*). 'Pricing Death: Frameworks for the Valuation and Securitization of Mortality Risk,' *ASTIN Bulletin*, 36: 79–120.

Cairns, A. J. G., D. Blake, and K. Dowd (2006*b*). '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., D. Blake, and K. Dowd (2008). 'Modelling and Management of Mortality Risk: A Review,' *Scandinavian Actuarial Journal*, 2008: 79–113.

Cairns, A. J. G., D. Blake, K. Dowd, G. D. Coughlan, D. Epstein, A. Ong, and I. Balevich (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., D. Blake, K. Dowd, G. D. Coughlan, D. Epstein, and M. Khalaf-Allah, (2011*a*). 'Mortality Density Forecasts: An Analysis of Six Stochastic Mortality Models,' *Insurance: Mathematics and Economics*, 48: 355–367.

Cairns, A. J. G., D. Blake, K. Dowd, G. D. Coughlan, and M. Khalaf-Allah (2011*b*). 'Bayesian Stochastic Mortality Modelling for Two Populations,' *ASTIN Bulletin*, 41: 29–59.

Cairns, A. J. G., K. Dowd, D. Blake, and G. D. Coughlan (2014). 'Longevity Hedge Effective ness: A Decomposition,' *Quantitative Finance*, 14(2): 217–235.

Cairns, A. J. G., G. Mavros, G. Streftaris, and T. Kleinow (2013). 'New Directions in the Modeling of Longevity Risk,' Presented at the International AFIRERM/Life/ASTIN Colloquium, Mexico City, October 2012.

Continuous Mortality Investigation (CMI) (2007). 'Stochastic Projection Methodologies: Lee-Carter Model Features, Example Results and Implications,' CMI Working Paper No. 25. London, U.K.: Institute and Faculty of Actuaries.

Coughlan, G., D. Epstein, A. Sinha, and P. Honig (2007). '*q*-Forwards: Derivatives for Transferring Longevity and Mortality Risk' [website]. <http://www.lifemetrics.com>.

Coughlan, G. D., M. Khalaf-Allah, Y Ye, S. Kumar, A. J. G. Cairns, D. Blake, and K. Dowd (2011). 'Longevity Hedging 101: A Framework for Longevity Basis Risk Analysis and Hedge Effectiveness,' *North American Actuarial Journal*, 15: 150–176.

Currie, I. D., M. Durban, and P. H. C. Eilers (2004). 'Smoothing and Forecasting Mortality Rates,' *Statistical Modeling*, 4: 279–298.

Currie, I. D. (2011). 'Modeling and Forecasting the Mortality of the Very Old,' *ASTIN Bulletin*, 41: 419–427.

Dahl, M., M. Melchior, and T. Møller (2008). 'On Systematic Mortality Risk and Risk Minimization with Survivor Swaps,' *Scandinavian Actuarial Journal*, 2–3: 114–146.

Delwarde, A., M. Denuit, and P. Eilers (2007). 'Smoothing the Lee–Carter and Poisson Log-Bilinear Models for Mortality Forecasting: A Penalized Log-Likelihood Approach,' *Statistical Modeling*, 7: 29–48.

Denuit, M., S. Haberman, and A. E. Renshaw (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., A. J. G. Cairns, D. Blake, G. D. Coughlan, D. Epstein, and M. Khalaf-Allah (2010*a*). 'Evaluating the Goodness of Fit of Stochastic Mortality Models,' *Insurance: Mathematics and Economics*, 47: 255–265.

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

Dowd, K., D. Blake, and A. J. G. Cairns (2010*c*). 'Facing up to Uncertain Life Expectancy: The Longevity Fan Charts,' *Demography*, 47: 67–78.

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

Dowd, K., D. Blake, and A. J. G. Cairns (2011*b*). '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., D. Blake, A. J. G. Cairns, G. D. Coughlan (2011*c*). 'Hedging Pension Risks with the Age-Period-Cohort Two-Population Gravity Model,' Presented at the Seventh International Longevity Risk and Capital Markets Solutions Conference, Frankfurt, September.

Eilers, P. H. C., and B. D. Marx (1996). 'Flexible Smoothing with B-Splines and Penalties,' *Statistical Science*, 11: 89–121.

Hyndman, R. J., and S. M. Ullah (2007). 'Robust Forecasting of Mortality and Fertility Rates: A Functional Data Approach,' *Computational Statistics and Data Analysis*, 51: 4942–4956.

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

Kallestrup-Lamb, M., A. B. Kock, and J. T. Kristensen (2013). 'Lassoing the Determinants of Retirement,' CREATES Research Paper 2013-21. Aarhus, Denmark: Aarhus University Department of Economics and Business.

Khiroya, N., and M. Penderis (2012). 'Using LDI as Part of a Glide Path De-Risking Program,' in Professional Pensions, ed., *The Guide to Risk Reduction*. London, U.K.: Professional Pensions, pp. 8–9.

Kleinow, T., and A. J. G. Cairns (2013). 'Mortality and Smoking Prevalence: An Empirical Investigation in Ten Developed Countries,' *British Actuarial Journal*, 18(2): 452–466.

Lee, R. D., and L. R. Carter (1992). 'Modeling and Forecasting U.S. Mortality,' *Journal of the American Statistical Association*, 87: 659–675.

Li, J. S.-H., M. R. Hardy, and K. S. Tan (2009). 'Uncertainty in Model Forecasting: An Extension to the Classic Lee-Carter Approach,' *ASTIN Bulletin*, 39: 137–164.

Li, J. S.-H., and A. Luo (2012). 'Key *q*-Duration: A Framework for Hedging Longevity Risk,' *ASTIN Bulletin*, 42: 413–452.

Li, J. S.-H., and M. R. Hardy (2011). 'Measuring Basis Risk in Longevity Hedges,' *North American Actuarial Journal*, 15: 177–200.

Li, N., and R. Lee (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 S.-A. Persson (2005). 'Is Mortality Dead? Stochastic Forward Force of Mortality Determined by No Arbitrage,' University of Bergen Working Paper. Bergen, Norway: University of Bergen.

Olivier P., and T. Jeffery (2004) 'Stochastic Mortality Models,' Presentation to the Society of Actuaries of Ireland.

Olivieri, A., and E. Pitacco (2009). 'Stochastic Mortality: The Impact on Target Capital,' *ASTIN Bulletin*, 39: 541–563.

Office for National Statistics (ONS) (2012). *Population Estimates for England and Wales, Mid-2002 to Mid-2010 Revised (National)*. Newport, South Wales: Office for National Statistics.

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. London, U.K.: Pensions Institute.

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: 556–570.

Richards, S. J., and I. D. Currie (2009). 'Longevity Risk and Annuity Pricing with the Lee-Carter Model,' *British Actuarial Journal*, 15: 317–343.

Richards, S. J., J. G. Kirkby, and I. D. Currie (2006). 'The Importance of Year of Birth in Two-Dimensional Mortality Data,' *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.

Willets, R. C. (2004). 'The Cohort Effect: Insights and Explanations,' *British Actuarial Journal*, 10: 833–877.

Zhou, R., and J. S.-H. Li (2013). 'A Cautionary Note on Pricing Longevity Index Swaps,' *Scandinavian Actuarial Journal*, 2013: 1–23.

Zhou, R., J. S.-H. Li, and K. S. Tan (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.