DISCUSSION PAPER PI-0822

Stochastic Mortality made Easy

Paul Sweeting

June 2008

ISSN 1367-580X

The Pensions Institute
Cass Business School
City University
106 Bunhill Row London
EC1Y 8TZ
UNITED KINGDOM

http://www.pensions-institute.org/

Stochastic Mortality made Easy

Paul Sweeting

Abstract

The purpose of this paper is to give actuaries an easy-to-use approach to modelling stochastic mortality. Whilst the approach described can be used with tailor-made projections, it can also be applied to published base tables and improvement factors. The methodology is not particularly new or ground-breaking; however, it is hopefully accessible and will allow actuaries to use a stochastic approach to mortality projections more easily.

1. Introduction

The importance of being able to model future mortality rates accurately has become increasingly important. This is partly because past projections have been shown to be lacking, but also because falling real interest rates have made the values of long-term pension and insurance policies much more sensitive to projected longevity. However, as well as being able to model expected rates accurately, it is important to model the uncertainty around these expected rates. Otherwise, we are left using best estimates with margins for prudence, an approach that is out of step with the stochastic approaches used elsewhere in actuarial work and more broadly in the world of finance and investment.

A number of mortality models have been developed in recent years, and these have been able increasingly to explain changes in mortality rates. However, they frequently exhibit one or more of the following characteristics:

- the method used to parameterise the model makes it difficult to generate sample paths (stochastic projections are not straightforward);
- parameterisation itself is not straightforward (the models are difficult to fit); and
- the writing style required for publication in refereed journals can make it hard to translate the models into practical solutions (the models are difficult to understand).

What I try to do in this paper is therefore to give a practical approach to modelling mortality stochastically. This means that the approach outlined is not wholly original – the important work has been covered by other authors. Nor will this approach be as theoretically robust as some of those produced by other authors. However, I believe that it is more important at this stage to produce an approach that can – and will – be used.

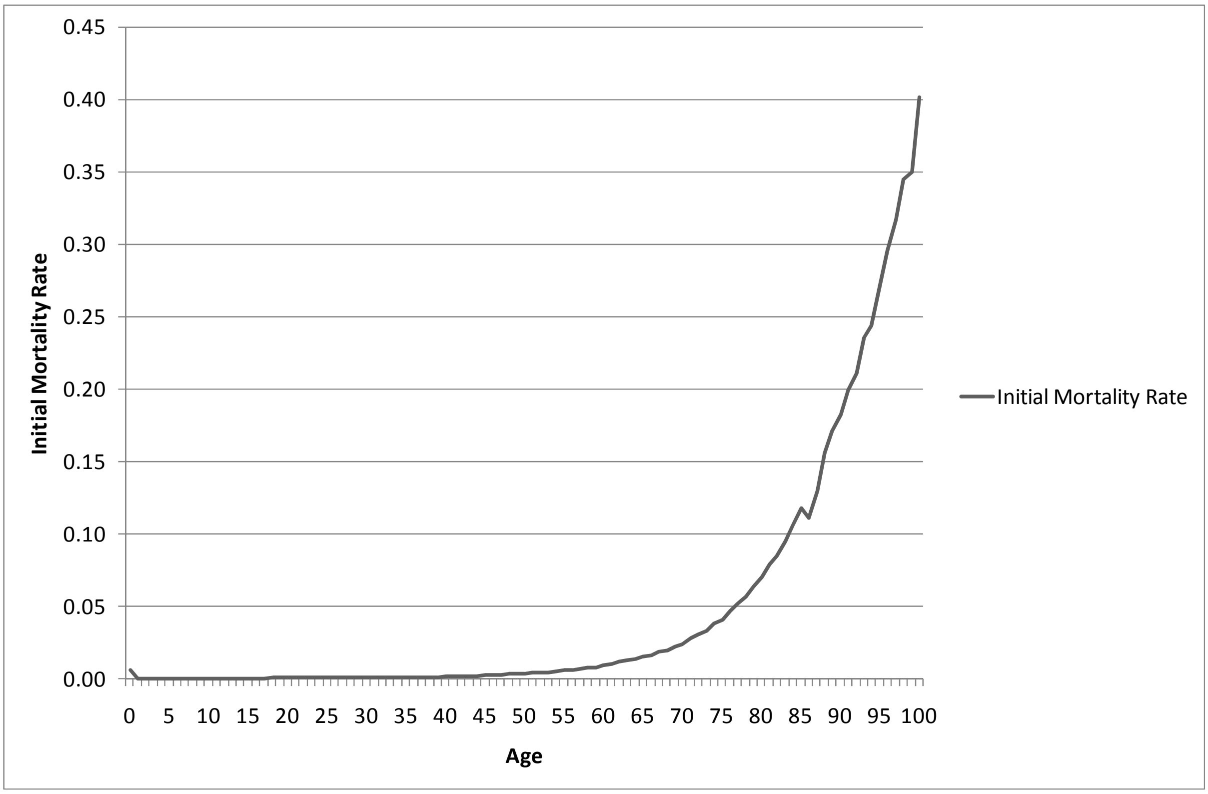
Cairns *et al* (2008) categorise mortality risk as unsystematic and systematic. The former occurs when the number of deaths is random even when the true mortality rate is known. The larger the population is, the smaller is the unsystematic mortality risk. Systematic mortality risk, on the other hand, is the component of mortality risk that cannot be diversified, in other words, uncertainty over the true future mortality rate. It is the latter that I am concerned with modelling here, although the former should be of interest to many pension schemes.

The structure of this paper is as follows. First, I give an overview of the structure of mortality rates, looking at variation by time and age. I then develop a simple model to describe the structure of mortality rates and their deviations from expected values. Finally, I show how these results can be used to generate sample paths, or series of stochastic mortality rates and, ultimately, stochastic annuity rates.

2. The Structure of Mortality Rates

The basic structure of a mortality rate curve will be familiar to most actuaries, and an example is shown below in Figure 2.1. This demonstrates the features of being generally lower at younger ages and then increasing exponentially (or in line with some similar function) for higher ages.

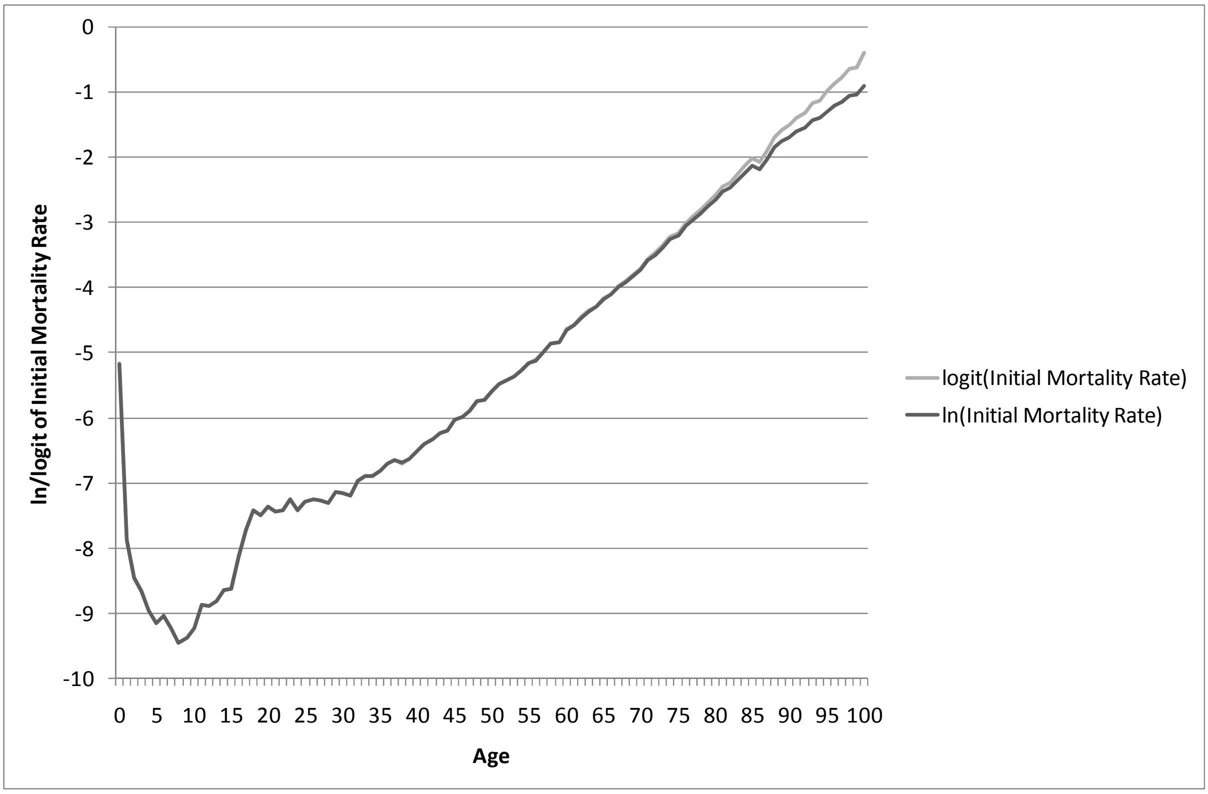
Figure 2.1 – Initial Mortality Rates for Male Lives aged 0 to 100, England and Wales, 2005



Source: Human Mortality Database – University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany)

Whilst familiar, this functional form is not particularly tractable – some sort of transformation is needed to give a function that can be analysed more easily. Cairns *et al* (2007) find that two functions that give more useful function forms for q_x, the initial mortality rate. These are $\ln(q_x)$, the natural logarithm of the mortality rate; and the logit of the initial mortality rate, $\ln(q_x/(1-q_x))$. These are shown in Figure 2.2. It appears – at least for English and Welsh males in 2005 over a wide range of ages – that the logit function does a slightly better job of transforming the initial mortality rate into a linear function, but for most of the range there is very little difference. The natural logarithm provides an easier function to work with for later analysis; however, if transforming initial mortality rates using natural logarithms, adding random variation and then converting back into initial rates, there is the chance that the resulting rate will be greater than unity. Whilst results can be artificially constrained to be less than this, is unsatisfactory. I therefore use the logit function, as the inverse of this function – $\exp(z)/(1+\exp(z))$ – is constrained to be between zero and unity.

Figure 2.2 – Natural Logarithm and Logit of Initial Mortality Rates for Male Lives aged 0 to 100, England and Wales, 2005

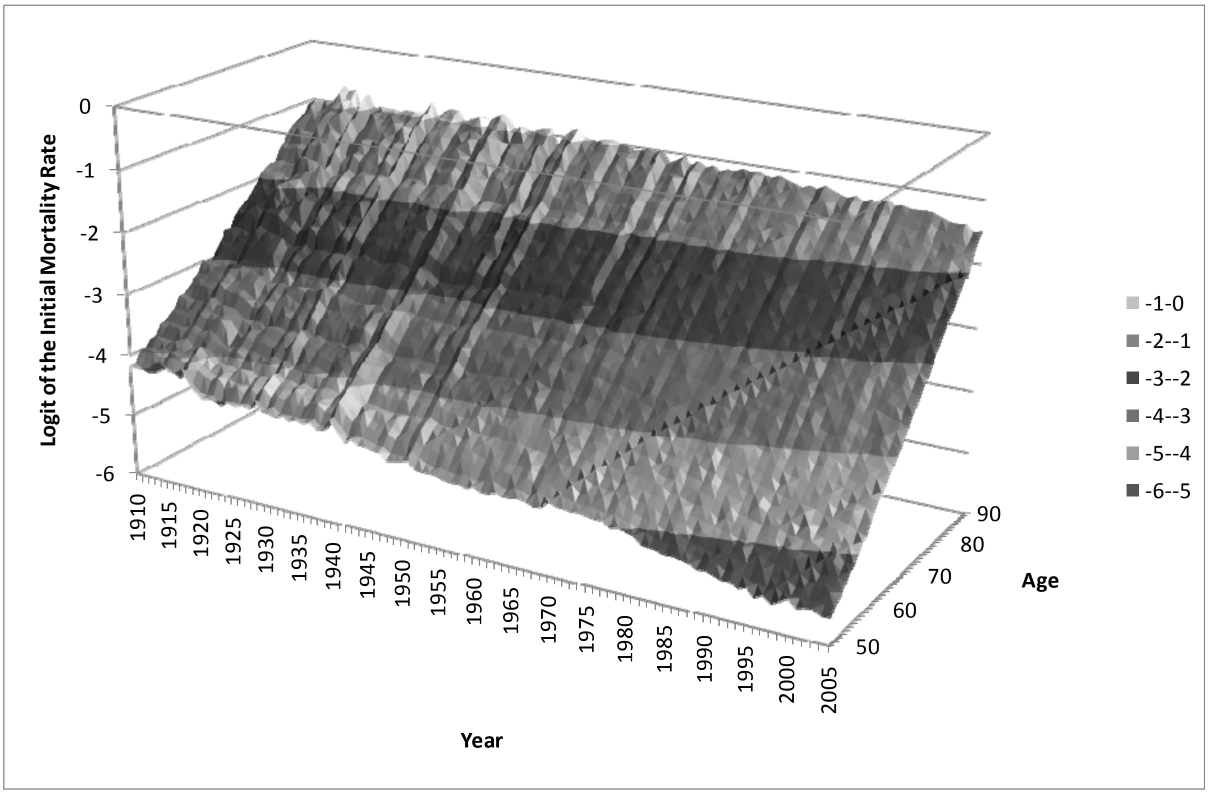


Source: Human Mortality Database – University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany); author's calculations

The range for which this chart is linear is of interest, as this is the range of ages that can easily be modelled. However, this chart considers rates only over a single year, and to project rates into the future some analysis of historical rates. The range of ages over which the logit of mortality rates is linear is wider in 2005 than it was in earlier years. This means that the "useful" range of ages that can be used for modelling is around age 50 to age 90. Other work frequently uses age 60 to age 90, for example Cairns *et al* 2007; however, if the results here are to be applied to, say, the RMV00 table then starting the age range at age 50 is helpful. The progress of historical rates in this age range is shown in the surface chart in Figure 2.3.

The range of years shown in this chart is 1910 to 2005, despite the fact that data is available going back to 1841. The reason that the data range is limited is that the ultimate data sources change through the whole period from 1841 onwards, but prior to 1910 the mortality rates appear artificially smoothed. I have therefore used the smaller range show.

Figure 2.3 – Logit of Initial Mortality Rates for Male Lives aged 50 to 90, England and Wales, 1910 to 2005



Source: Human Mortality Database – University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany); author's calculations

It is clear from the above chart that the year-to-year changes are the most important factors in describing mortality rates. This can be confirmed by noting that the most obvious patterns are ridges and peaks changing across the years. However, there are other more subtle patterns that it is helpful to model. Most important is the cohort effect. This has been described by many authors, but a good overview is given by Willets (2004). The cohort effect refers to the fact that people born in certain years have experienced greater mortality improvements than those born either before or after the period in question. For the UK, this refers to people born between 1925 and 1945. The cohort effect is described by diagonal lines on the surface in Figure 2.3.

3. An Easy Approach

The first stage in this approach is to specify the model I use. This is:

(3.1) $$\text{logit}(q_{x,y}) = \alpha_y + x\beta_y + \gamma_c + \varepsilon_{x,y}$$

Where:

- $q_{x,y}$ is the initial rate of mortality for a life aged x in year y;
- α_y is the constant in year y;
- β_y is the slope in year y; and
- γ_c is the constant for cohort c, where $c = y - x$
- $\varepsilon_{x,y}$ is an error term

This is similar to model M6 described in Cairns *et al* (2007), including the fact that their model uses the logit of the initial mortality rate. Model M6 is also described as a generalisation of the Cairns-Blake-Dowd model – from Cairns et al (2006) – allowing for the cohort effect.

It is worth noting that much of the analysis in the literature uses central rather than initial mortality rates. This is because the central rates can be taken to be as approximately equal to the force of mortality. Being an instantaneous measure, this lends itself more easily for use in some models. However, given that most practitioners will actually use initial rates of mortality – being interested in the proportion of an initial population that will survive – I carry out all calculations here in terms of initial rates.

I fit the data to this model using least squares regression. This is a basic approach which does not make use of much of the information in the data. In particular, it ignores the actual number of deaths and the exposed to risk at each age and in each year. Other methods, such as the Poisson model used by Brouhns *et al* (2002) and others since, do make use of this information – but the computational requirements are much greater than for the approach I use.

In order to specify the equation robustly, I do not include cohorts –or lives – for which there are five or fewer observations. This means the earliest cohort is for lives born in 1825 (so men who were 85 in 1910; 86 in 1911; 87 in 1912; and so on up to 90 in 1915) and the latest is for lives born in 1950 (so for men aged 50 in 2000; 51 in 2001; 52 in 2002; and so on up to age 55 in 2005). This means that in (1) above, c ranges from 1825 to 1950. Subject to this constraint and remembering that $c = y - x$, x ranges from 50 to 90 and y ranges from 1910 to 2005.

So that the equation can actually be specified, some of the parameters need to be excluded. I omit α_{2005}, β_{2005}, γ_{1949} and γ_{1950}.

Carrying out this regression gives the building blocks for the next stage of the process. There are two approaches that can be used here. The first involves projecting mortality rates and stochastic variation around them; the second involves deriving only the stochastic variation and applying this to an existing projection. I describe only the second approach, for two main reasons. First, the simplistic nature of this approach is unlikely to give central projections of mortality that are as well-informed as some of the official projections; secondly, given that a range of accepted projection bases exists, it is probably more helpful to provide a methodology for the application of stochastic variation around these projections.

Although a cohort factor has been modelled, its main purpose is to reduce the degree of unexplained variation in the rest of the model; only α_y and β_y, the level and slope factors with respect to age for each year must be analysed. The analysis takes the following stages:

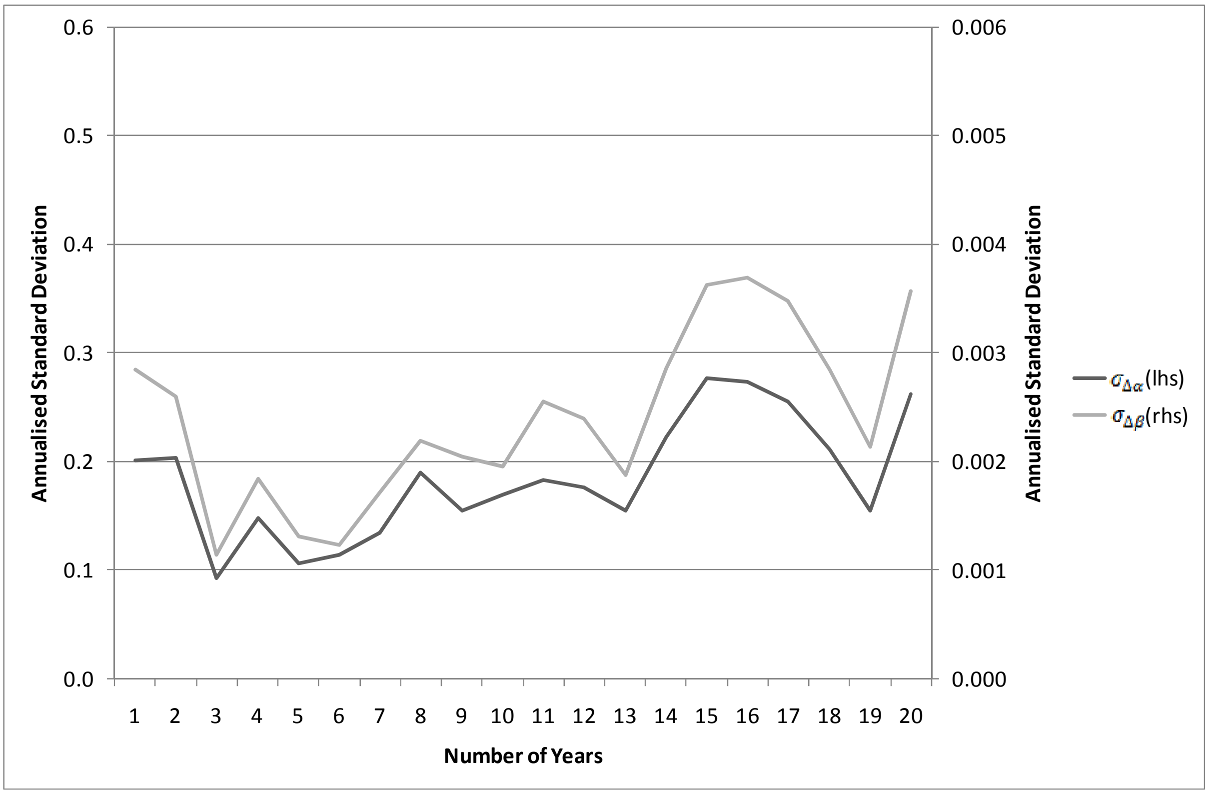
- for α_y, linear regression is applied to the coefficients α_0 to α_n, where $y = 0$ for the first year of observation;
- the results are then used to estimate $\hat{\alpha}_{2n}$;
- the difference between $\hat{\alpha}_{2n}$ and α_{2n} is then calculated as $\Delta\alpha_{2n} = \hat{\alpha}_{2n} - \alpha_{2n}$;
- the process is then repeated using the coefficients α_{2n} to α_{3n}, and so on until all coefficients have been analysed;
- at this point, the same process is carried out with β_0 to β_n, to give $\hat{\beta}_n$ and then $\Delta\beta_n$; and

- the standard deviations of $\Delta\alpha$ and $\Delta\beta$ are then calculated and annualised as $\sigma_{\Delta\alpha}$ and $\sigma_{\Delta\beta}$ respectively, and the correlation between them is calculated as $\rho_{\Delta\alpha,\Delta\beta}$.

It is worth explaining in words what $\sigma_{\Delta\alpha}$, $\sigma_{\Delta\beta}$ and $\rho_{\Delta\alpha,\Delta\beta}$ represent. The first two terms describe the extent to which estimates of the shape of future mortality curves (given as a linear function of the logit of the initial rate) differ from the actual rates in terms of the level and the rate of change with age; the third term describes how the difference in one of these estimates is correlated with the difference in the second.

Clearly the choice of n is key here. A larger value of n is more consistent with a longer projection period, but gives fewer observations. Both standard deviations show broadly increasing values with time, as shown in Figure 3.1 below. However, the number of observations on which this conclusion is based becomes increasingly small. Having said this, using the 20-year values of a value of 0.262 for $\sigma_{\Delta\alpha}$ and 0.00358 for $\sigma_{\Delta\beta}$ seems reasonable, even conservative.

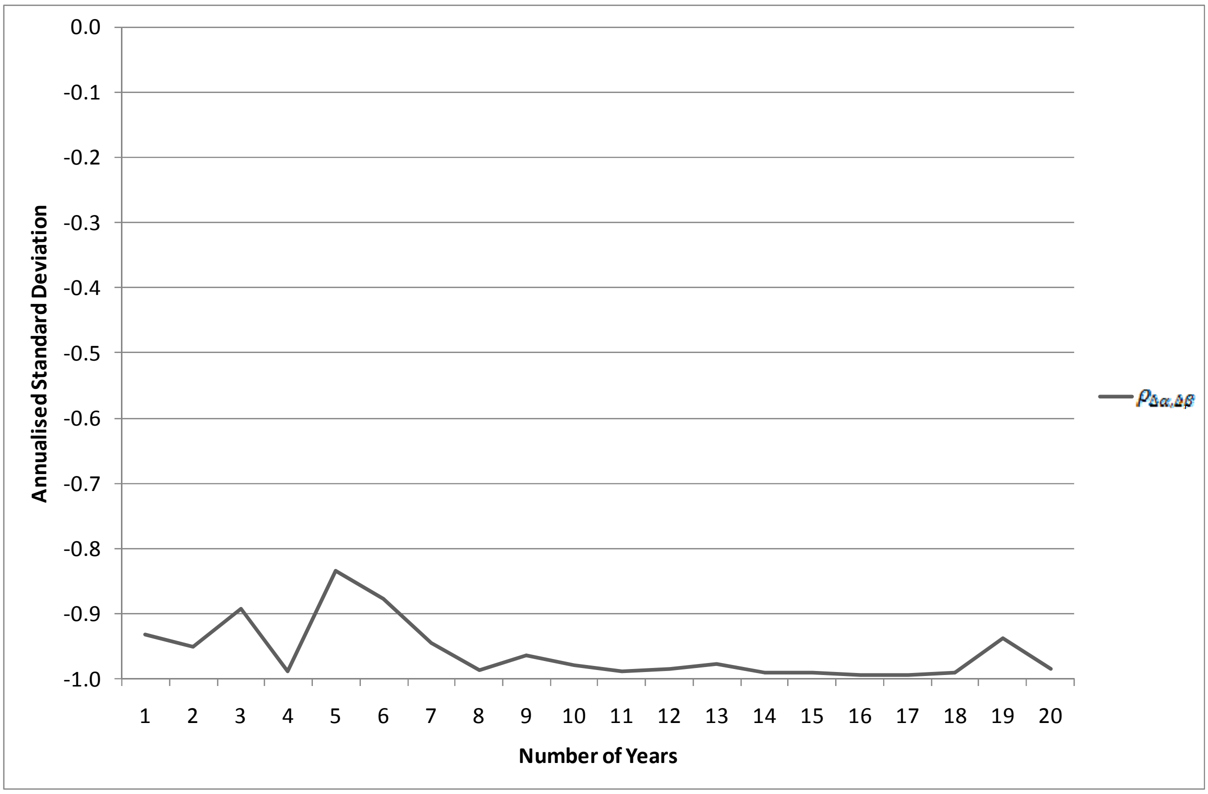
Figure 3.1 – Annualised Standard Deviations for Model Parameters



Source: author's calculations

Moving on to the correlation between these $\Delta\alpha$ and $\Delta\beta$, it becomes clear that something strange is happening, as the correlation is highly negative, frequently being less than -0.9, as shown in figure 3.2. This is because both of these variables have exhibited changes in trend which a purely analytical process – as used here – fails to pick up, giving this strong negative correlation.

Figure 3.2 – Correlation between Model Parameters



Source: author's calculations

This feature can be used to make stochastic simulation more straightforward. If the correlation is taken to be -1, then the two-factor model reduces to a one-factor formulation.

4. Applying These Results

What I now have is a basic one-factor mortality model. In order to apply this to an existing table, the following stages are followed:

- for the scenario s, generate a normal random variable $V_{f,s}$ with zero mean and unit standard deviation for each future year f;
- multiply $V_{f,s}$ by 0.262 to get $\sigma_{\Delta\alpha}$ and by 0.00358 to get $\sigma_{\Delta\beta}$ for each year;
- calculated the stochastic adjustment $A_{x,f,s}$ for future year f, scenario s and each age x as:

(4.1) $$A_{x,f,s} = \sum_{\varphi=0}^{f}\left(0.262V_{\varphi,s} - 0.00358V_{\varphi,s}x\right)$$

- construct a "central projection" of $q_{x,f}$ for each future year f and each age x using published base tables and projection factors for each age and future year of interest;
- calculate the logit of each of these values, $lq_{x,f}$:

(4.2) $$lq_{x,f} = \ln\left(q_{x,f}/\left(1-q_{x,f}\right)\right)$$

- apply the stochastic adjustment for future year f, scenario s and each age x to give the logit of the simulated mortality rate, $lq_{x,f,s}$:

(4.3) $lq_{x,f,s} = lq_{x,f} + A_{x,f,s}$

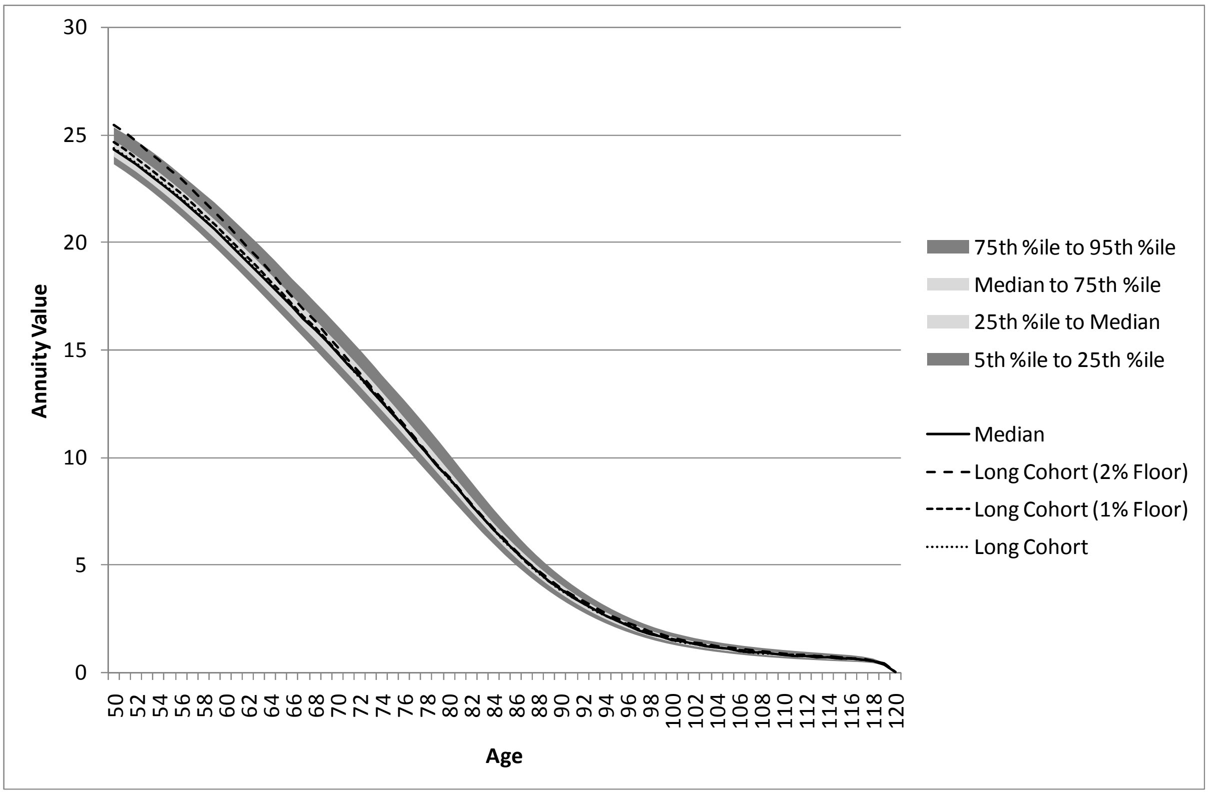
- convert this back to an initial rate to give the simulated initial rate of mortality for future year f, scenario s and age x:

(4.4) $q_{x,f,s} = \exp(lq_{x,f,s})/(1 + \exp(lq_{x,f,s}))$

These mortality rates can then be used to calculate annuity factors using commutation functions or other approaches. An Excel spreadsheet is provided to show how all of this can be done in practice.

To see how much difference this stochastic mortality approach makes, I carry out some calculations of annuity values. The first results are given in Figure 4.1. These show selected percentiles of annuity rates calculated from stochastic initial mortality rates based on RMV00 and the long cohort projection basis (with year zero being 2007), compared with the annuity rates calculated using a deterministic with the same base table and long cohort with 0%, 1% and 2% floors. A real interest rate of 2% is used.

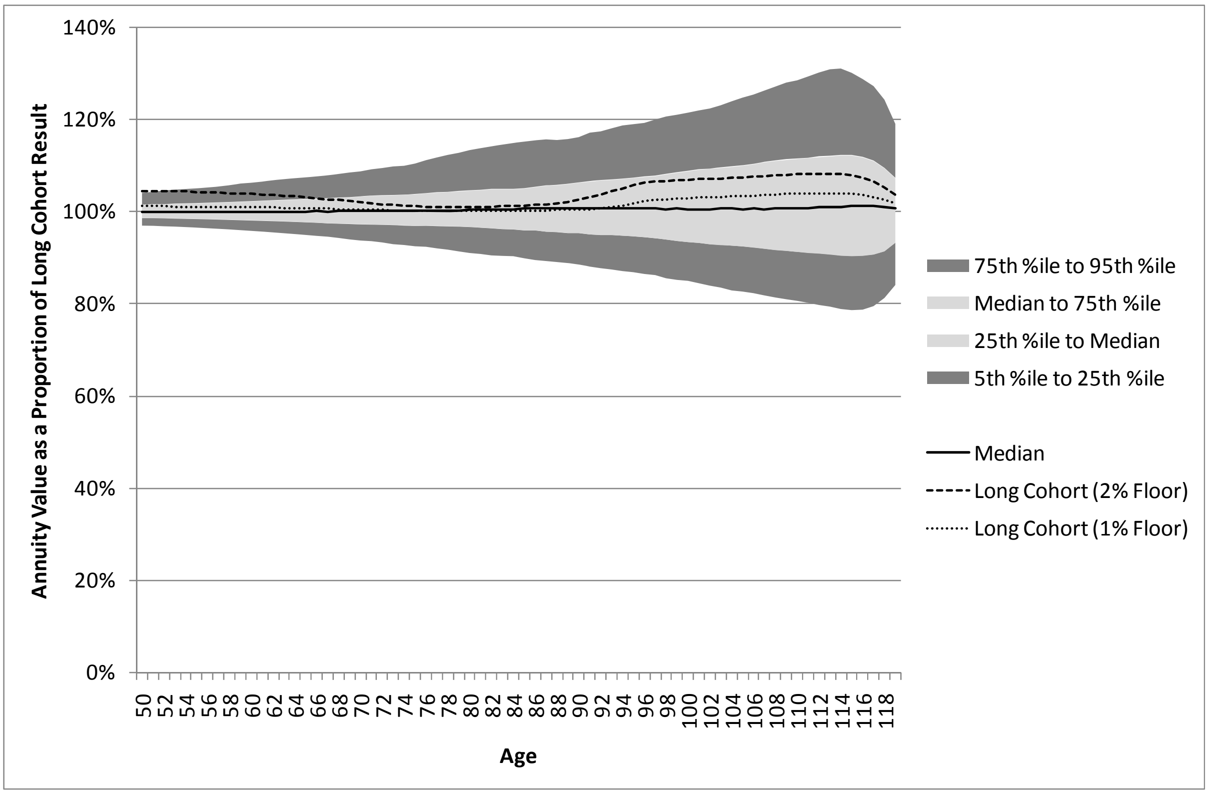
Figure 4.1 – Comparison of Annuity Rates



Source: CMI, author's calculations

Reassuringly, the median from the stochastic calculations sits very close to the result from the deterministic long cohort calculation, but it is interesting to see that even at relatively young ages the deterministic projections fall below the 95th percentile of the stochastic calculations, and even below the 75th percentile for most of the range. This is easier to see if the results are presented as a proportion of the deterministic long cohort result, as they are in Figure 4.2, below.

Figure 4.1 – Comparison of Annuity Rates



Source: CMI, author's calculations

Another way of thinking about the difference is to consider the difference in values for an annuity book made up of a stable population of pensioners. If the age range of the stable population is 50 years of age and higher, then relative to a portfolio valued using the long cohort approach the 75th percentile using a stochastic approach is 2.6% higher, whilst the 95th percentile is 7.3% higher. This compares with an increase of 0.7% if a 1% floor is added to the long cohort basis, or an increase of 3.0% if the floor is at the 2% level.

This confirms what other studies have shown – that allowing for uncertainty in mortality projections can have a big impact on the level of reserves that might be required.

5. Limitations of this Approach

Whilst this method does (hopefully) give a useable approach for modelling mortality stochastically, it is important to recognise the limitations of this approach. In summary, these are:

- theoretically, the central (rather than the initial) rate of mortality should be modelled, and the result converted into an initial rate;
- the logit (or the natural logarithm) of mortality is not necessarily a linear function of age for ages outside the range modelled here – in particular, mortality at extreme old ages is often lower than predicted by a linear model, and the "accident hump" at younger ages is not picked up using a simple approach;

- mortality improvements are independent from year to year – whilst there are independent random fluctuations in mortality each year, the bigger risk is that the trend in mortality improvements will change;
- the method used does not allow for the number volume of available data at different ages

Hopefully, easy-to-use models will appear that address these limitations, but hopefully the approach presented here will be useful in the meantime.

References

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., Blake, D., and Dowd, K. (2006). “A Two-Factor Model for Stochastic Mortality with Parameter Uncertainty: Theory and Calibration", *Journal of Risk and Insurance*, 73, 687-718.

Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong A. and Balevich I., (2007). “A quantitative comparison of stochastic mortality models using data from England & Wales and the United States”, *Pensions Institute Discussion Paper PI-0701.*

Cairns, A. J. G., Blake, D. and Dowd, K., (2008). “Modelling and Management of Mortality Risk: A Review”, *Pensions Institute Discussion Paper PI-0814.*

Continuous Mortality Investigation, (2004). *Report No 21*.

Continuous Mortality Investigation, (2007). *CMI Library of Mortality Projections*.

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

University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). (2008) “Human Mortality Database” *available at www.mortality.org or www.humanmortality.de (data downloaded on 10 June 2008).*