



## City Research Online

### City, University of London Institutional Repository

---

**Citation:** Mayhew, L. & Smith, D. (2009). Whither human survival and longevity or the shape of things to come (Actuarial Research Paper No. 189). London, UK: Faculty of Actuarial Science & Insurance, City University London.

This is the unspecified version of the paper.

This version of the publication may differ from the final published version.

---

**Permanent repository link:** <https://openaccess.city.ac.uk/id/eprint/2320/>

**Link to published version:**

**Copyright:** City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

**Reuse:** Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

---

---

---

City Research Online:

<http://openaccess.city.ac.uk/>

[publications@city.ac.uk](mailto:publications@city.ac.uk)

---



**Cass Business School**  
CITY UNIVERSITY LONDON

**Faculty of Actuarial  
Science and Insurance**

**Actuarial Research Paper  
No. 189**

**Whither Human Survival and Longevity  
*or*  
The Shape of Things to Come**

**Leslie Mayhew and David Smith**

**June 2009**

**ISBN 978-1-905752-21-8**

**Cass Business School  
106 Bunhill Row  
London EC1Y 8TZ  
Tel +44 (0)20 7040 8470  
[www.cass.city.ac.uk](http://www.cass.city.ac.uk)**

"Any opinions expressed in this paper are my/our own and not necessarily those of my/our employer or anyone else I/we have discussed them with. You must not copy this paper or quote it without my/our permission".

# **Whither Human Survival and Longevity** *or* **The Shape of Things to Come**

## **Contents**

1. Introduction
2. Modelling approach and basic ideas
3. The empirical model
4. Developing projections
5. Future projections
6. Conclusions

References

Annexes

## **Author email addresses**

Lesmayhew@googlemail.com

D.A.Smith@city.ac.uk

# Whither Human Survival and Longevity *or* The Shape of Things to Come

Leslie Mayhew and David Smith  
Faculty of Actuarial Science and Insurance  
Cass Business School

## Abstract

With the continuing increases in life expectancy, populations are ageing rapidly. Governments are concerned for the future of pensions and health care for which population forecasts are an important component for planning purposes. In this paper we focus on human survival rather than mortality rates which are the more usual starting point when estimating future populations. Using a simple model we link basic measures of life expectancy to the shape of the human survival function and consider its various forms. We then use the simple model as the basis for investigating actual survival in England and Wales from 1841 onwards and investigate the concept of a 'maximum age'. We show how the model can be used in a predictive sense and demonstrate in two tests that show our model would have given more accurate results than comparable government forecasts using the same base information. We then go on to show that, based on trends in life expectancy, official population forecasts could undershoot the population at age 50+ by 0.6m, with consequent financial implications for pensions, health and social care.

## 1. Introduction

In June 2009 Henry Allingham of the UK celebrated his 113<sup>th</sup> birthday, one of a tiny number of people that have made it to such an old age. On present trends the next 20 years or so will see a six fold increase in the number of centenarians in this country to around 60,000 and so Henry Allingham will potentially be joined by many others, not only in the UK but also around the world (Manton et al, 1991; Coale, 1996; Vaupel, 1998; Wilmoth and Robine, 2003; Robine and Saito, 2003).

However, amidst the celebrations of his achievement it is still the case that around 14% of English males will die before their 65<sup>th</sup> birthday and 31% before their 75<sup>th</sup> birthday. The highly publicised increases in life expectancy have thus to an extent masked the fact that life span is still hugely variable even in prosperous countries.

Whilst to a degree life span inequalities are inevitable, society continues to attach a high priority to reducing the number of premature deaths from whatever causes or influences whether biomedical or societal in origin. As a result, policy makers are pulled in two directions. At one extreme they wrestle with the problem of how society will be able to cope with the burgeoning numbers of elderly in terms of health care costs and pensions, and at the other, of promoting a healthy living and risk-avoiding culture so that people are able to live longer active and more productive lives.

We have reached the present position in which the majority of UK citizens live until they are at least 80 years old as a result of many factors and influences spread over 100 years. These include massive reductions in childhood mortality through better nutrition and immunisation programmes that has led to mortality under 30 years of age virtually disappearing, better welfare safety nets leading to poverty reduction, improved health care, central heating, more recently life style changes from reductions in smoking, and finally better treatment and management of chronic diseases in old age particularly heart disease.

Nevertheless, inequality in life span is arguably the most fundamental inequality that exists among human populations (Rogers, 1995; Carey, 2003; Edwards and Tuljapurkar, 2005; Wachter, 2003) and so progress in reducing it is of significant interest in public policy terms. For example the UK government target is to reduce health inequalities by 10% by 2010 as measured by infant mortality and life expectancy at birth between the worst performing areas of the UK and the rest. To what extent is the achievement of this target supported by trends in  $S(x)$ , the proportion of the population surviving to age  $x$ , through time and by when is it likely to be accomplished?

Such changes in the way we live spread over many years, prompts a number of questions. For example, whether inequalities are reducing or increasing through time, or whether there are upper limits to life span or if patterns differ between countries, in particular those at different stages of economic development or with different cultures. An interesting question for example is whether countries experiencing rapid development and economic growth exhibit similar patterns in human survival to already developed countries, and if there are differences whether they are attributable to differences in governance or approaches to welfare.

The literature on the subject of human survival is considerable, going back in some cases centuries (Olshansky and Carnes, 1997). Among the many papers on the subject, Fries (1980) made the important observation that, although more people were living longer for all these reasons, the shape of the human survival curve  $S(x)$  was becoming more rectangular in shape, suggesting that there were possible biological limits to life span.

This has sparked considerable debate in the literature ever since in which the arguments broadly resolve into two camps. One agrees with Fries which is that from a rational standpoint there must be biological limits to life expectancy and it is only a matter of 'when' and not 'if' they limit is reached. The other camp points to the fact that there has been an unbroken linear rise in life expectancy of about three months a year for at least 150 years and that there are no signs of this abating (Oeppen and Vaupel, 2006). In the leader board of life expectancies in different countries, the position has changed many times over the years with, for example, New Zealand in the first half of last century leading the way, then Scandinavia, briefly Switzerland, and now Japan. Supporters of this viewpoint hence argue that, taking the world as a whole, it is premature to talk of upper limits for so long as this trend persists.

Following in Fries's footsteps there have been numerous efforts to demonstrate progress towards rectangularisation hypothesis applying an array of measures to  $S(x)$  including the interquartile range, standard deviations and Gini coefficients (Wilmoth

and Horiuchi. 1999; Kanisto, 2000; Lynch and Brown, 2001). Their ultimate focus is on the ‘compression of human mortality’ and around an upper limit to life span rather than on the driving forces. Edwards and Tuljapurkar (2005), using the Kulber-Leibler divergence to compare changes over time, found evidence of differences in  $S(x)$  according to socio-economic inequality, associated factors such as educational status and ethnicity as well as for example gender.

For an actuary, pension provider or health economist, maximum life span is arguably mainly of academic interest as compared with  $S(x)$ . The reason is that very few people will reach the maximum life span while changes in the numbers of people reaching different ages will determine the future benefits to be paid out in the case of life assurance and pension provision, and will inform the demand for social security and health care from the public purse. Such information is customarily supported by official published population projections. A problem is that there has been a tendency to under-estimate improvements in mortality especially at older ages, resulting in population forecasts that are too low (Bengtsson, and Keilman, 2003).

In this paper we will argue that there is very little sense of overall strategic direction or indication of speed of change in human survival, so that we do not know if this year’s projections will alter significantly from one year to the next or how this will affect future planning assumptions. An example of this would be the answer to the question of how far pension age should be raised, and if life expectancy increases faster than we had planned for in our calculations, how many more people will survive to pensionable age than we had anticipated? A simple look-up table or chart may be more practicable in this regard than repeated model runs of different population scenarios (Blake and Mayhew, 2006).

Taking these considerations into account, measures of rectangularisation do not directly lead to the answer of how many people will survive to age  $x$  between now and 2020 and nor how many deaths would be avoided as a result each year in the intervening period. What would be helpful are methods to predict the shape of  $S(x)$  going forward from which can be derived many useful measures including how many lives will survive to any given age. To do this we need to understand patterns of survival in more detail.

In this paper we develop a model that builds on regularities in the survival curve to shed light on some of these important questions, whether life expectancy is measured at birth or at any other age. The paper has five objectives; they are to:

- model and predict changes in the shape of the survival curve through time so that for example knowing where life expectancy will be in the future will enable accurate estimates of the percentages surviving to different ages
- use the model to investigate whether there is evidence of a human ‘maximum’ age and, if so, how many years will elapse until that point is reached
- investigate whether there are signs that life-span inequalities are reducing or increasing with time and at what rate and in which age range



- estimate how many extra lives will be saved each year and at which ages as survival rates improve
- test whether the approach produces more accurate estimates of survivorship than currently used methods that spring from mortality models in the first instance.

We use data for England and Wales concentrating on the period 1841 to 2003 to illustrate our findings. First, however, we outline the modelling approach and develop the mathematical framework we need to develop the basic ideas used in the analysis. Later sections apply the model to historical data to interpret the past and as a basis for building projections in which technical details on how projections are constructed are included. Since survival is the converse of death, the method is theoretically complementary to customary methods based on projecting mortality rates although differences arise due to technical differences of approach.

Limitations of space restrict the number of empirical examples that can be provided in this paper, although we have tested the method on data from several countries. In presenting our results, we consider the case of predicting survivorship from 1981 to the present using then GAD forecasts and our forecasts based on the model but from a 1981 perspective and then with a similar procedure for 1991. We then compare the performance of both models with the realised survivorship rates and compare what happens when the model is rolled forward to 2020. A concluding section reflects on the principal findings and overall approach and discusses follow-on work.

The origins of the techniques described are derived from research in two other unrelated fields both of which use queuing theory. These are worthy of mention because they illustrate a commonality of mathematical processes under different guises that suggests, in turn, a range of other possible applications. The first of these is described in Mayhew (1987) which is concerned with estimating the time elapsed by social security claims in the system and was used for setting social security targets, and the second is based on time to discharge times in accident and emergency departments (Mayhew and Smith, 2007). Both in fact have their basis in types of survival curve, but the distributions are known by different names in the queuing literature (usually the distribution of ‘time spent in the system’).

The fact that in both cases there are statistically robust empirical regularities between average time spent and the time taken to process a given percentage of social security claims or patients in an A&E department encouraged us to believe that similar regularities are observable in other fields, in this case human survival. Such an example is given in Mayhew (2001) which investigates the issue of long term care provision in Japan. Trends in life expectancy are used to project the Japanese survival curve in future years and hence the numbers of people likely to need long term care going forward. The key difference is that sojourns in these environments are ideally as short as possible but in life most want to remain in the system as long as possible ~ a life in slow lane!

Among the results obtained in this paper the following are of particular significance:

- There has been a fundamental change in the pattern of survivorship from 1946 with the result that previously indicated trends of a convergence in survival ages has stalled. At the oldest ages the survival ages are becoming more divergent with no sign of slowing down so that some people will live ever longer. These and other findings could mean that the Government target of reducing health inequalities which relies on a convergence in life expectancy in different areas of the country may not be achievable.
- The underlying model that generates these patterns and trends, described in the first part of the paper, can be adapted to produce future life tables that appear to be more accurate than those officially used as the basis for population projections. This finding is based on a comparison of our model with GAD projections made in 1981 and 1991 and comparing the results with what actually occurred.
- In projecting forward to 2020 our model shows greater survivorship than official GAD forecasts suggesting that there will be even higher numbers of older people than is currently being planned for. For example using 2001 as our base year we estimate that the number of males aged 50+ in 2020 will be approximately 0.7m higher than GAD's own principal 2001 based projection for 2020. If correct this finding has implications for pensions, life insurance and services for older people such as health and social care that rely on population projections for planning and distribution purposes.

## 2. Modelling approach and basic ideas

Mathematically the survival curve denoted by  $S(x)$  denotes the probability of surviving to age  $x$  whereas  $e(x)$  defines life expectancy at age  $x$ . With no loss of generality we initially develop our ideas into the form of a simple, stylised model in order to gain insights into the processes linking survival and life expectancy. This allows us to derive a basic relationship between the survival processes and simple dynamics that allow us to estimate lives saved, and mortality at different ages.

To calibrate the model in practice we use a form of the conventional Gompertz-Makeham survival curve in order to investigate the actual relationship between lives saved at different ages over time. However, we do not place restrictions on the values that the parameters are able to take. By exploiting the regularities in the data and trends over time we are also able to make a qualified guess at whether populations are converging towards a maximum life span, how many lives are saved each year of life expectancy, and the speed at which the survival curve is converging or otherwise.

Imagine a stationary population in which there is a constant number of births and deaths, no migration and which is subjected to the same mortality regime each year. Consider Figures 1(a) – (c), which show the mortality curves ABC for three such hypothetical populations *at a given point in time*. The vertical axis shows the number of survivors  $l_x$  and the horizontal axis age  $x$ . We define the point  $x_1$  in each case as the onset of mortality, the age at which death begins, which can range from zero upwards. For simplicity we assume there are no deaths before this age. We call the point where BC cuts the age axis as  $x_2$  or the maximum age to which anyone lives.

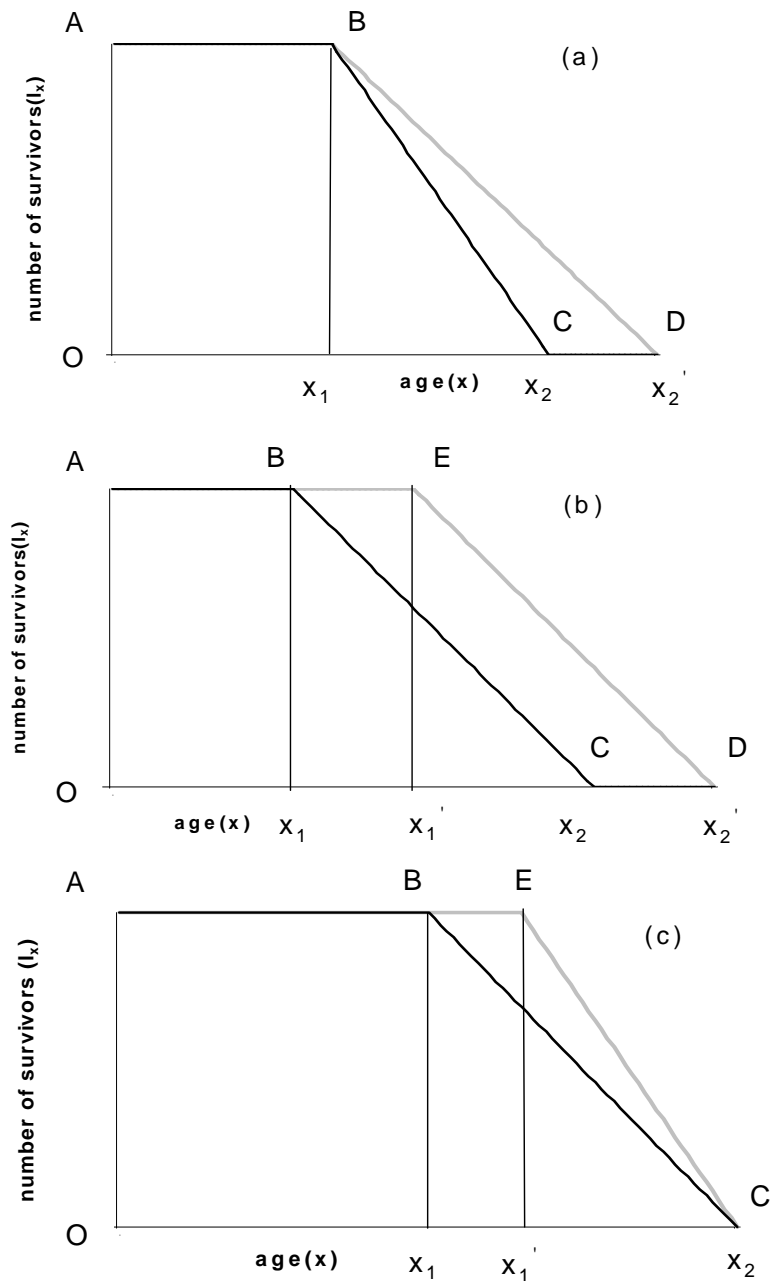


Figure 1: Three hypothetical mortality models (a), (b) and (c)

Both  $x_1$  and  $x_2$  are somewhat fuzzy quantities in the real world. In developed countries we could assume the onset of mortality ( $x_1$ ) occurs from around 50 years onwards. Alternatively this may be defined by reference to particular percentile of deaths, for example 10%, to remove the effects of accidents and rare diseases that affect few of the population. Similarly, rather than look at the final death in a population it may be more appropriate to assume that  $x_2$  is the age of death of the 90<sup>th</sup> percentile. However, our purpose is to use them as conceptually useful devices to anchor and compare distributions and mortality processes rather than to determine them empirically.

Now imagine the age distribution of the population at another point in time. In model (a), we see that  $x_1$  is unchanged whilst  $x_2$  the oldest age has advanced to  $x_2'$  (point D). In other words the onset of mortality is unchanged but now some people live to older ages. The consequence of this is a decline in the mortality gradient BD compared with BC. In model (b) we see that both  $x_1$  and  $x_2$  have advanced by the same amount, such that the mortality gradient is the same before and after. In model (c)  $x_2$  has remained constant but  $x_1$  has advanced to  $x_1'$  (point E) with the effect that the mortality gradient is steeper. For reasons that will become apparent shortly we call (a) the dispersed mortality model, (b) the parallel mortality model and (c) the compressed mortality model

What do the models tell us about the evolution of populations following one or other of these evolutionary paths? Qualitatively speaking, model (a) might be thought of benefiting older people more than younger people, model (b) all age groups equally and model (c) younger generations before older generations.

### Some relationships and properties of the simple model

Table 1 is a list of basic parameters comparing each of the three models. Assume the number of survivors at a given age is  $l_x$ . Line 1 in the parameter list (table 1) shows the population for each case as a function of  $x_1$  and  $x_2$ . This is simply the area under the population curve. Line 2 shows life expectancy  $e_{x_1}$  at  $x_1$  as a function of  $x_1$  and  $x_2$ . Line 3 shows the mortality gradient or number of deaths at each age where death can occur. As shown, this value must increase if the difference between the age when people start dying and the age at which all lives have ceased decreases. This is to ensure a stable population as an increase in the number of deaths at each age is needed so that total deaths are the same.

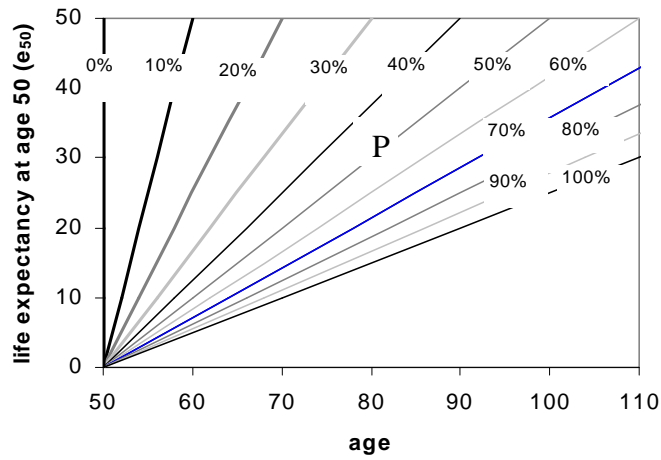
Line 4 is a measure of the degree to which the mortality curve is rectangular in shape, following Wilmoth and Horiuchi (1999). The nearer the ratio is to one the closer the mortality curve is to a rectangle. It is calculated by comparing the area under the mortality curve from birth with the encompassing rectangle whose base ranges from the age at which mortality age to  $x_2$ .

Line 5 shows the theoretical relationship between cumulative mortality  $p(x)$ , life expectancy at  $x_1$  and a specified age,  $x$ , that is greater than  $x_1$ . These equations are most easily understood in the context of Figure 2a-c. Figure 2a depicts the dispersion case in which  $x_1$  is assumed fixed while  $x_2$  increases. On the vertical axis is life expectancy at 50 years, which is assumed for illustrative purposes to equate with  $x_1$ . Each line represents the cumulative mortality of 0%, 10%, 20% 30% etc. and can be used to determine the age at which the specified percentage of the population has died by. The 0% line corresponds to the vertical axis and so is invariant with life expectancy; however, note that all lines converge at the origin. Take for example point P equating to a life expectancy of a further 30 years at age 50 - the graph shows that 50% of the population would die by age 80. Further analysis shows that each one-

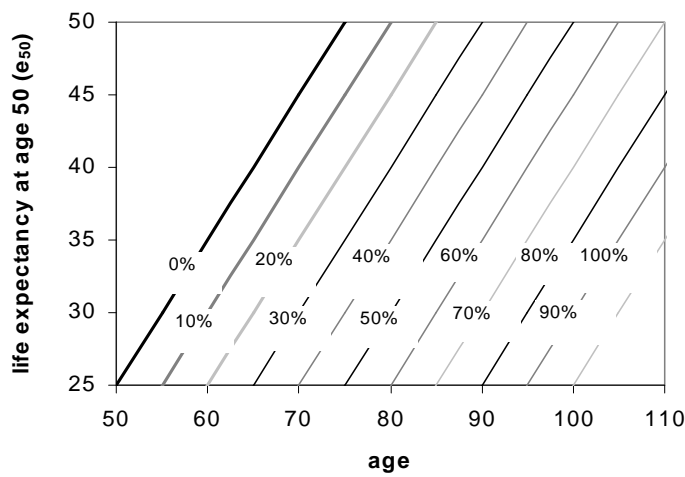
year advance in life expectancy at  $x_1$  translates into 2 extra years in maximum age as can be easily verified.

|    | <i>Parameter</i>   | <i>Mortality dispersion (A)</i>                         | <i>Parallel Mortality (B)</i>                              | <i>Mortality compression (C)</i>                        |
|----|--|---|--|---|
| 1  | Population size  | $\frac{l_x}{2}(x_1 + x_2')$                             | $\frac{l_x}{2}(x_1' + x_2')$                               | $\frac{l_x}{2}(x_1' + x_2')$                            |
| 2  | Life expectancy at age $x_1$ ( $e_{x_1}$ )                               | $\frac{1}{2}(x_2' - x_1)$                               | $\frac{1}{2}(x_1' + x_2') - x_1$                           | $\frac{1}{2}(x_1' + x_2) - x_1$                         |
| 3  | Number of deaths at each year of age where death is possible             | $\frac{l_{x_1}}{x_2' - x_1}$                            | $\frac{l_{x_1}}{x_2' - x_1'}$                              | $\frac{l_{x_1}}{x_2 - x_1'}$                            |
| 4  | Coefficient of rectangularisation  | $\frac{x_2' + x_1}{2x_2'}$                              | $\frac{x_2' + x_1'}{2x_2'}$                                | $\frac{x_2 + x_1'}{2x_2'}$                              |
| 5  | Proportion of population deceased as a function of $e_{x_1}$ and age $x$ | $p(x) = \frac{x - x_1}{2e_{x_1}}$                       | $p(x) = \frac{x - x_1'}{2(e_{x_1} - (x_1' - x_1))}$        | $p(x) = \frac{x - x_1'}{2(e_{x_1} - (x_1' - x_1))}$     |
| 6  | Survival function as a function of $e_{x_1}$                             | $1 - \frac{x - x_1}{2e_{x_1}}$                          | $1 - \frac{x - x_1'}{2(e_{x_1} - (x_1' - x_1))}$           | $1 - \frac{x - x_1'}{2(e_{x_1} - (x_1' - x_1))}$        |
| 7  | Force of Mortality at age $x$  | $\frac{1}{2e_{x_1} - x + x_1}$                          | $\frac{1}{2(e_{x_1} - (x_1' - x_1)) - x + x_1'}$           | $\frac{1}{2(e_{x_1} - (x_1' - x_1)) - x + x_1'}$        |
| 8  | Inter percentile range, $\Delta x$                                       | $2e_{x_1} \Delta p(x)$                                  | $2(e_{x_1} + x_1 - x_1') \Delta p(x)$                      | $2(x_2 - x_1 - e_{x_1}) \Delta p(x)$                    |
| 9  | Deaths avoided or person years saved                                     | $\frac{l_{x_1}}{2}(x_2' - x_2)$                         | $l_{x_1}(x_2' - x_2)$                                      | $\frac{l_{x_1}}{2}(x_1' - x_1)$                         |
| 10 | Adjustment cycle (years)   | $x_2' - x_1$  | $x_2' - x_1$   | $x_2 - x_1$   |
| 11 | Mean age of Population, $\bar{x}$  | $\frac{x_1 + x_2'}{3} - \frac{x_1 x_2'}{3(x_1 + x_2')}$ | $\frac{x_1' + x_2'}{3} - \frac{x_1' x_2'}{3(x_1' + x_2')}$ | $\frac{x_1' + x_2}{3} - \frac{x_1' x_2}{3(x_1' + x_2)}$ |
| 12 | Median age   | $\frac{1}{4}(x_1 + x_2)$                                | $\frac{1}{4}(x_1' + x_2')$                                 | $\frac{1}{4}(x_1' + x_2)$                               |

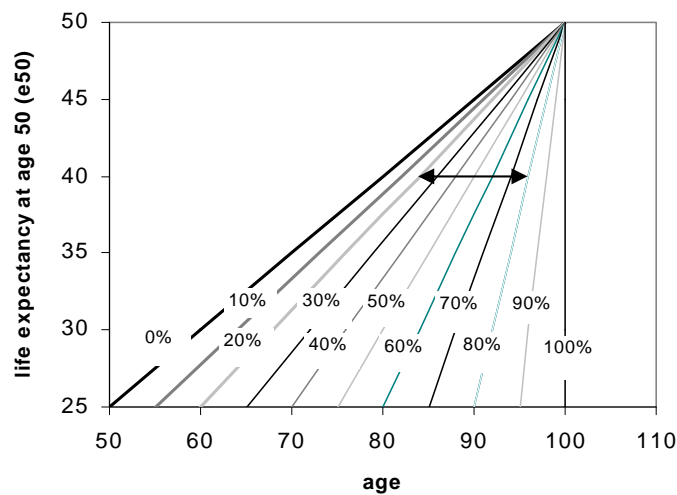
Table 1: Some parameters of interest



(a)



(b)



(c)

Figure 2: Cumulative mortality for each case: (a), (b) and (c)

In Figure 2b we have the same axes as before but cumulative mortality is now assumed to be parallel; that is to say each advance in life expectancy has exactly the same proportional effect on the chances of survival at all ages above the given age

and so all age groups appear to benefit equally. To aid the illustration  $x_1$  is again initially set to 50 and  $x_2$  to 100, where both parameters are assumed to advance in lock step. Minimum life expectancy in this case is 25 years rising to an assumed maximum of 50 years, which is why the vertical axis starts at 25 and not 0 as in case (a). Now only the 0% mortality curve passes through the origin, whereas previously all the percentiles did so. As for this model, the increase in values for  $x_1$  and  $x_2$  must be the same value, each one-year increase in life expectancy advances both  $x_1$  and  $x_2$  by one year, as can be easily verified.

In 2c cumulative mortality from 50 converges to a point as  $x_1$  advances while  $x_2$  is held constant. As with case (b), minimum life expectancy at 50 is 25 years assuming  $x_1$  equals 50 and  $x_2$  equals 100. Similarly, only the 0% mortality curve passes through the origin, but note that the lines meet when the onset of mortality  $x_1$  equals the maximum age  $x_2$ . Life expectancy at this point is given by  $x_2 - x_1$  so for example if  $x_2$  equals 100 and  $x_1$  equals 50 the maximum life expectancy is  $100 - 50 = 50$ . In other words all people live to 100! Since  $x_2$  is fixed each one-year advance in life expectancy equates to a 1-year delay in the onset of mortality,  $x_1$ , up to a maximum of  $x_2$ , and so the rate at which the lines converge depends on advances in life expectancy.

It is noteworthy that the median age (denoted by the 50<sup>th</sup> percentile) is identical in all three models with a cumulative mortality gradient of one. This can be seen by putting  $p(x) = 0.5$  in each equation in line 5 and simplifying to give  $e_{x_1} = x - x_1$  in each case. Put another way, it means that a one-year increase in life expectancy always equates to a rise in median age of one year regardless of which model one uses.

Line 6 shows the survival function as a function of  $e_{x_1}$ . This is simply 1 less the proportion of lives that have died as stated in line 5.

Line 7 shows the force of mortality calculated using

$$\mu(x) = \frac{-\frac{d}{dx}S(x)}{S(x)}$$

By definition we can also reverse the process to get the survival function from the force of mortality.

So for the Mortality Dispersion Model:

$$\begin{aligned}
S(x) &= \exp \left\{ - \int_{x_1}^x \frac{1}{2e_{x_1} - (t - x_1)} dt \right\} \\
&= \exp \left\{ \ln \left[ 2e_{x_1} - (t - x_1) \right]_{x_1}^x \right\} \\
&= \exp \left\{ \ln \left[ 2e_{x_1} - (x - x_1) \right] - \ln \left[ 2e_{x_1} - (x_1 - x_1) \right] \right\} \\
&= \exp \left\{ \ln \left[ \frac{2e_{x_1} - (x - x_1)}{2e_{x_1}} \right] \right\} \\
&= 1 - \frac{(x - x_1)}{2e_{x_1}}
\end{aligned}$$

Wilmoth and Horiuchi (1999) noted the degree of convergence (or divergence) in mortality curves provided one possible measure of rectangularisation, and suggested the age difference between the inter-quartile range for this purpose. In line 8, we generalise this concept and call it the inter-percentile range. Consider model (a) and assume the inter-percentile range is 0.8 i.e. (90%-10%) and that life expectancy at  $x_1$  is 20 years. The Inter-percentile age range is  $2 \times 20 \times 0.8$  or 32 years but if life expectancy is 30 years the range is 48 years or 50% larger.

If we now consider model (b) the IPR is, not surprisingly, constant for any given value of life expectancy. For example, if  $2(e_{x_1} + x_1 - x'_1) = 10$ , i.e. the expectation of life at age  $x'_1$  is 10 years, and the IPR is 0.8 then the age range would be 8 years. In the convergent case, it is obvious from Figure 2c the IPR reduces as life expectancy increases. Plugging similar numbers into model (c) as before and assuming a maximum age  $x_2$  equal to 100 then when life expectancy is 40 we obtain an IPR of 16 (calculated as  $2 \times 0.8 \times (100 - 50 - 40)$ ) and similarly an IPR of 48 years when life expectancy is 20 years. The line arrow in Figure 2c denotes this case.

On some more general points we note that the derivative or slope of the cumulative mortality curves  $\frac{d}{dx} e_{x_1}$  is independent of both  $x_1$  and  $x_2$ . This means that changes in these parameters will not affect the general shape of the graphs, but they will cause a scale shift on the horizontal and vertical axes. It is further evident that the slopes in models (a) and (c) are functions of  $p$  whereas in model (b) the slope is a constant, as noted above. It is normally assumed that mortality rates decline in old age. In our models age specific mortality rates are given by  $\frac{1}{x_2 - x}$  which means that in model (a) mortality rates decline through time, in model (b) they are constant but shift to the right, while in model (c) they increase but become more compacted into a smaller age range.



Of the three models, it is readily apparent that (c) comes closest to the model originally proposed by Fries (1980). This because it is the only case among the three in which the coefficient of rectangularization (line 4) converges to a value of one, which occurs when  $x_1$  equals  $x_2$ . The other two models approach a value of one but only as  $x'_2$  becomes unrealistically large. Since rectangularization has been thoroughly studied by Wilmoth and Horiuchi (1999), we do not pursue it further here but simply note the range of outcomes.

### Population Size

The table shows there are different equations for the three distinct models. However this does not need to be the case, as all models have an  $x'_1$  and an  $x'_2$ , it is just that the value is sometimes the same as the original. If we are happy to allow the situation where  $x'_1 = x_1$  or  $x'_2 = x_2$  then we can have one equation for all three models: e.g.

$$\text{Population size} = l_x x'_1 + l_x \frac{(x'_2 - x'_1)}{2} = l_x \frac{(x'_2 + x'_1)}{2}$$

Similarly, by using 'new' values for all models we are able to get life expectancy to always equal

$$\frac{1}{2}(x'_2 - x'_1) + (x'_1 - x_1) = \frac{1}{2}(x'_2 + x'_1) - x_1$$

As the population is assumed to be stationery the total number of deaths per year must equal  $l_x$ . Therefore the number of deaths at a particular age, where this age must be greater than  $x'_1$ , can always be expressed as:

$$\frac{l_{x_1}}{x'_2 - x'_1}$$

In terms of the degree of rectangularisation, we have

$$\text{Area of rectangle} = l_x x'_2$$

$$\text{Area under mortality curve} = l_x x'_1 + \frac{1}{2} l_x (x'_2 - x'_1) = \frac{1}{2} l_x (x'_2 + x'_1)$$

$$\text{Co-efficient} = \frac{l_x (x'_2 + x'_1)}{2 l_x x'_2} = \frac{x'_2 + x'_1}{2 x'_2}$$

The proportion of population deceased as a function of  $e_{x_1}$  and age  $x$  is then,

$$\frac{x - x'_1}{x'_2 - x'_1} = \frac{x - x'_1}{2[e_{x_1} - (x'_1 - x_1)]} \quad \text{where } x \geq x'_1$$

And the Inter percentile range (IPR) is

$$\Delta p(x'_2 - x'_1) = \Delta p 2[e_{x_1} - (x'_1 - x_1)]$$

By general reasoning the increase in the stationary population must be the lives saved over the period.

$$l_x \left[ \frac{(x'_2 - x_2)}{2} + \frac{(x'_1 - x_1)}{2} \right]$$

Table 1 can therefore be simplified to:

|    | <i>Parameter</i>   | <i>Equation</i>  |
|----|--|--|
| 1  | Population size  | $\frac{l_x}{2}(x'_1 + x'_2)$   |
| 2  | Life expectancy at age $x_1$ ( $e_{x_1}$ )                               | $\frac{1}{2}(x'_2 + x'_1) - x_1$                                     |
| 3  | Number of deaths at each year of age, $b$                                | $\frac{l_{x_1}}{x'_2 - x'_1}$  |
| 4  | Coefficient of rectangularisation  | $\frac{x'_2 + x'_1}{2x'_2}$  |
| 5  | Proportion of population deceased as a function of $e_{x_1}$ and age $x$ | $p(x) = \frac{x - x'_1}{2[e_{x_1} - (x'_1 - x_1)]}$                  |
| 6  | Survival function as a function of $e_{x_1}$                             | $1 - \frac{x - x'_1}{2[e_{x_1} - (x'_1 - x_1)]}$                     |
| 7  | Force of Mortality at age $x$  | $\frac{1}{2(e_{x_1} - (x'_1 - x_1)) - x + x'_1}$                     |
| 8  | Inter percentile range   | $\Delta x = \Delta p 2[e_{x_1} - (x'_1 - x_1)]$                      |
| 9  | Deaths avoided or person years saved                                     | $l_x \left[ \frac{(x'_2 - x_2)}{2} + \frac{(x'_1 - x_1)}{2} \right]$ |
| 10 | Adjustment cycle (years)   | $x'_2 - x_1$   |
| 11 | Mean age of Population, $\bar{a}$  | $\frac{x'_1 + x'_2}{3} - \frac{x'_1 x'_2}{3(x'_1 + x'_2)}$           |
| 12 | Median age   | $\frac{1}{4}(x'_1 + x'_2)$   |

Table 2: Some parameters of interest for the combined model

## Mortality Rates

We need to understand how populations adjust to changes in mortality prospects for subsequent cohorts through time in order to assess the impact of improved survival on the lives saved at each age. Again we adopt a simplified approach, by considering each of the models separately and then bring them together in a rather clumsy general equation.

### Model A

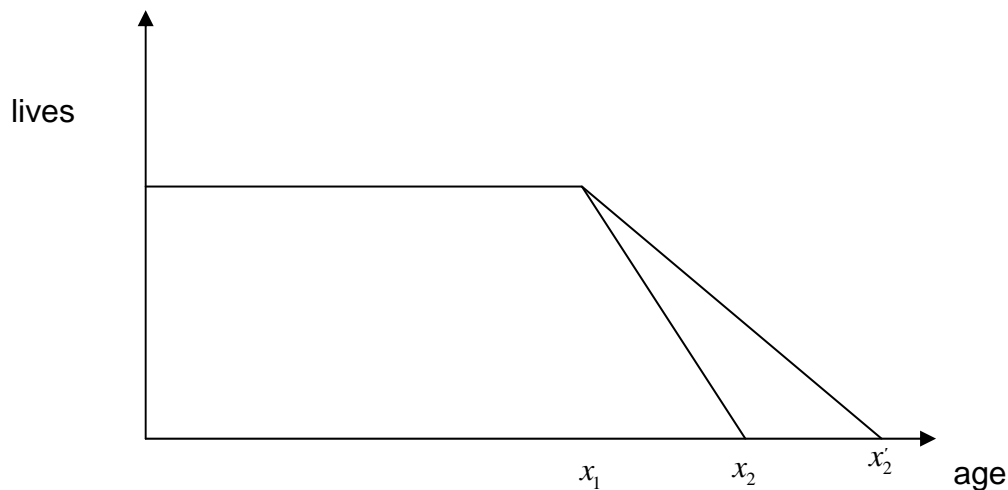


Figure 3: Change in survival under model A

If we consider a particular year where the population suddenly become 'healthier' i.e. they are the first people to have the potential to reach age  $x_2'$ , then we need time to elapse until they reach the age  $x_1$ .

For  $t \leq x_1$  i.e. before these new 'healthier' lives reach  $x_1$

$$\text{Death rate} = (x_2 - x_1) \frac{l_x}{(x_2 - x_1)} = l_x$$

This is of course intuitive as we start with a stable population and if  $l_x$  people are born then the same number must die.

For  $x_1 < t \leq x_2$  i.e. when the new 'healthier' lives reach  $x_1$  and die at the new lower rate but we still have some older lives dying at the original rate

$$\text{Death rate} = (x_2 - t) \frac{l_x}{(x_2 - x_1)} + (t - x_1) \frac{l_x}{(x_2' - x_1)}$$

The first term involves the older lives who die at the original rate. As  $t$  increases this number reduces (as the population with original mortality rates dies out). The second

term involves the newer lives dying at the newer, lower rate. As  $t$  increases the value of this term increases as a larger number of people from this population reach ages where they can die. The overall death rate decreases as  $t$  increases over this range of values. The population is therefore growing over this period and growth reaches a maximum at time  $t = x_2$ .

For  $x_2 < t \leq x'_2$  i.e. all the older lives have died off but no new lives have yet reached the new maximum age  $x'_2$  i.e. the population is still increasing.

$$\text{Death rate} = (t - x_1) \frac{l_x}{(x'_2 - x_1)}$$

As there are only new lives left we only have the new death rate but we haven't reached the stable population yet. As  $t$  increases the number of people in the total population and the population exposed to mortality increases and the number of deaths also increases.

For  $x'_2 < t$  i.e. the new maximum age has been reached and the population is again stable

$$\text{Death rate} = (x'_2 - x_1) \frac{l_x}{(x'_2 - x_1)} = l_x$$

Deaths under model A

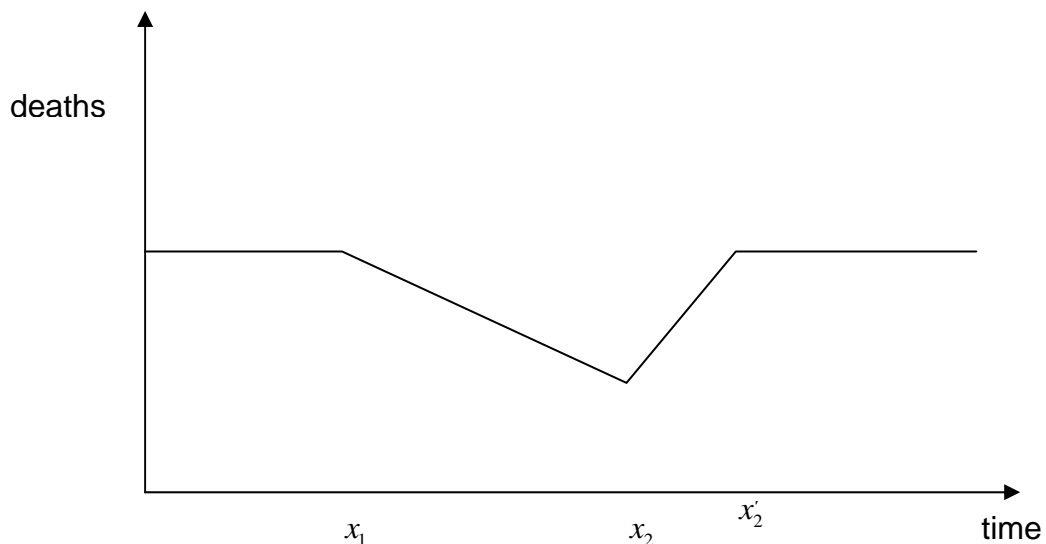


Figure 4: Deaths under model A

Note that the gradients will not usually be the same (they will only be the same if  $x'_2 = 2x_2 - x_1$ ).

Model B

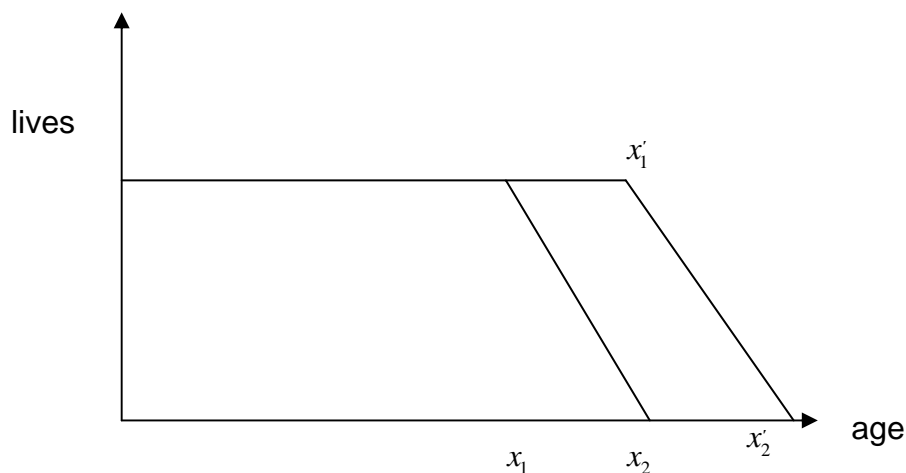


Figure 5: Change in survival under model B

This time we have the difficulty of whether  $x'_1$  is greater or less than  $x_2$  i.e. how large the shift is.

**Case (a)**  $x'_1 < x_2$

For  $t \leq x_1$  i.e. before these new 'healthier' lives reach  $x_1$  and don't die.

$$\text{Death rate} = (x_2 - x_1) \frac{l_x}{(x_2 - x_1)} = l_x$$

Which is the same as previously.

For  $x_1 < t \leq x'_1$  i.e. the new 'healthier' lives reach  $x_1$  and don't die but we still have some older lives dying.

$$\text{Death rate} = (x_2 - x_1) \frac{l_x}{(x_2 - x_1)} - (t - x_1) \frac{l_x}{(x_2 - x_1)} = (x_2 - t) \frac{l_x}{(x_2 - x_1)}$$

The first term is the same as above while the second term represents the lives that no longer die. As  $t$  increases this number increases which means that the total death rate decreases as  $t$  increases.

For  $x'_1 < t \leq x_2$  i.e. the new 'healthier' lives reach  $x'_1$  and start to die at the new rate (which is actually the same as the old rate) and we still have some older lives dying.

However, the rate is changing as more of the new lives start getting exposed to ages they can die at.

$$\text{Death rate} = (x_2 - t) \frac{l_x}{(x_2 - x_1)} + (t - x_1) \frac{l_x}{(x'_2 - x'_1)} = \frac{l_x (x_2 - x'_1)}{(x'_2 - x'_1)}$$

The first term is the same as above (i.e. decreases as  $t$  increases as the old lives die out) while the second term are the new lives that are starting to die. As  $t$  increases this number increases. Note that for this period deaths are constant as the number of new lives dying is the same as the reduction in old lives dying (as the mortality rate for a given age where mortality occurs is not changed as the slope is identical).

For  $x_2 < t \leq x'_2$  i.e. all the older lives have died off and the number of new lives dying is increasing as they start to reach the new maximum age.

$$\text{Death rate} = (t - x'_1) \frac{l_x}{(x'_2 - x'_1)}$$

As there are only new lives left we only have the new death rate but we haven't reached the stable population yet. As  $t$  increases the number of people in the population and the number of deaths both increase.

For  $x'_2 < t$  i.e. the new maximum age has been reached and the population is again stable

$$\text{Death rate} = (x'_2 - x'_1) \frac{l_x}{(x'_2 - x'_1)} = l_x$$

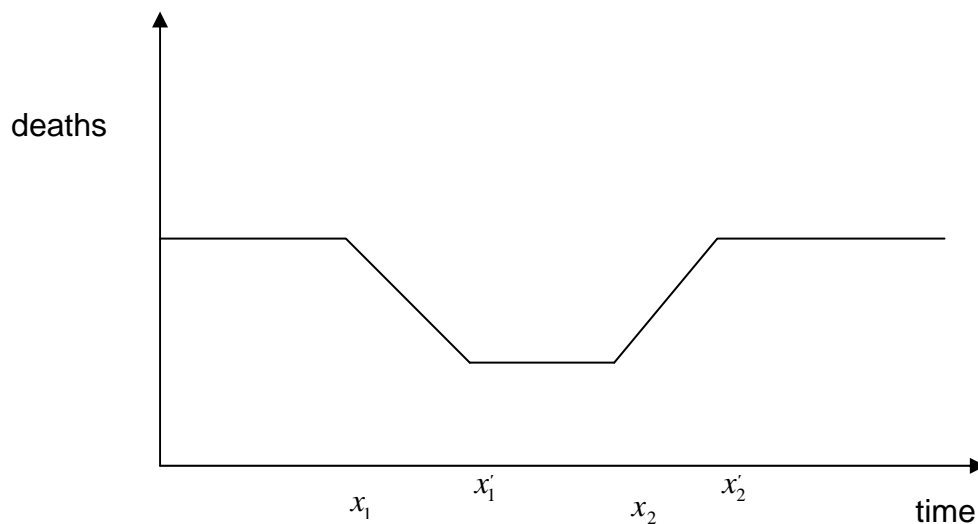


Figure 6: Deaths under model B

Note that the gradients are the same.

**Case (b)**  $x'_1 < x_2$

For  $t \leq x_1$  i.e. before the new 'healthier' lives reach  $x_1$

$$\text{Death rate} = (x_2 - x_1) \frac{l_x}{(x_2 - x_1)} = l_x$$

Which is the usual death rate as seen in case a.

For  $x_1 < t \leq x_2$  i.e. the new 'healthier' lives reach  $x_1$  and don't die but we still have some older lives dying.

$$\text{Death rate} = (x_2 - x_1) \frac{l_x}{(x_2 - x_1)} - (t - x_1) \frac{l_x}{(x_2 - x_1)} = (x_2 - t) \frac{l_x}{(x_2 - x_1)}$$

This is the same death rate as in case (a) though the times when this process occurs differ.

For  $x_2 < t \leq x'_1$  i.e. the new lives haven't reached  $x'_1$  so aren't dying but all the old lives have died off => no one dies during this period!

$$\text{Death rate} = 0$$

For  $x'_1 < t \leq x'_2$  i.e. all the older lives have died off and the number of new lives dying is increasing as they pass the age where deaths start to occur,  $x'_1$ , and start to reach the new maximum age  $x'_2$ .

$$\text{Death rate} = (t - x'_1) \frac{l_x}{(x'_2 - x'_1)}$$

Again, we have seen this death rate in case (a) but with the time that this applies being different.

For  $x'_2 < t$  i.e. the new maximum age has been reached and the population is again stable

$$\text{Death rate} = (x'_2 - x'_1) \frac{l_x}{(x'_2 - x'_1)} = l_x$$

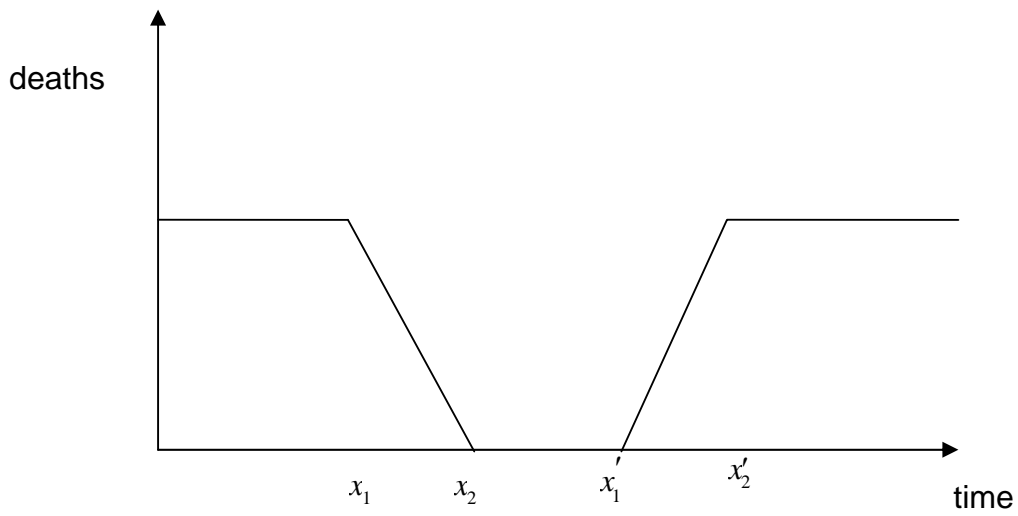


Figure 7: Deaths under model B: alternative case ~ note that the gradients will be the same

Model C

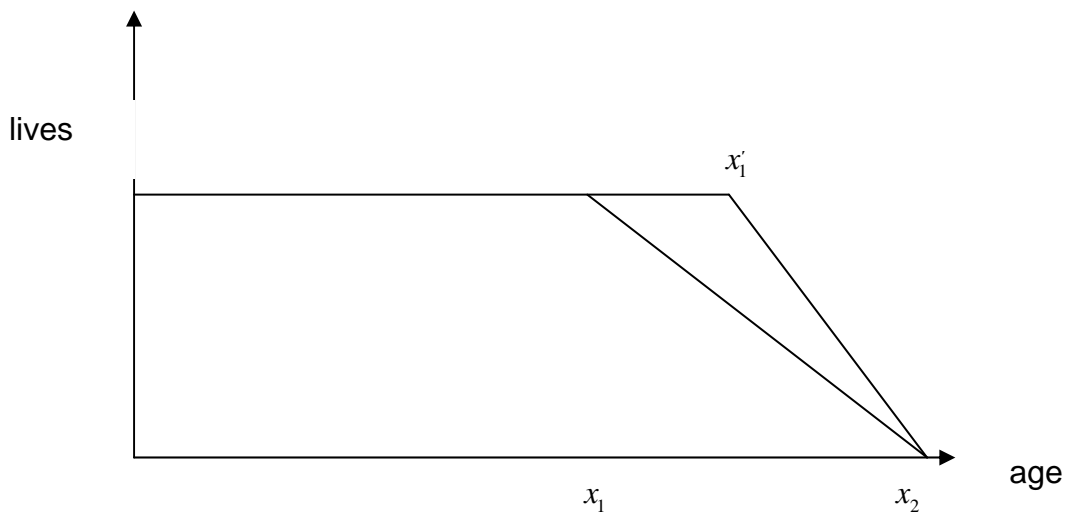


Figure 8: Effect of an increase in survival with no change in maximum age

For  $t \leq x_1$  i.e. before these new 'healthier' lives reach  $x_1$

$$\text{Death rate} = (x_2 - x_1) \frac{l_x}{(x_2 - x_1)} = l_x$$

Which is our usual result.



$x_1 < t \leq x'_1$  i.e. the new 'healthier' lives reach  $x_1$  and don't die but we still have some older lives dying

$$\text{Death rate} = (x_2 - x_1) \frac{l_x}{(x_2 - x_1)} - (t - x_1) \frac{l_x}{(x_2 - x_1)} = (x_2 - t) \frac{l_x}{(x_2 - x_1)}$$

This is the same as Model B.

For  $x'_1 < t \leq x_2$  i.e. the new 'healthier' lives reach  $x'_1$  and start to die at the new rate and we still have some older lives dying at the old rate.

$$\text{Death rate} = (x_2 - t) \frac{l_x}{(x_2 - x_1)} + (t - x'_1) \frac{l_x}{(x_2 - x'_1)}$$

This is similar to Model 2 but the rates that the new and old population experience are different

For  $x_2 < t$  i.e. the old lives die out and the population becomes stable once again

$$\text{Death rate} = (x'_2 - x'_1) \frac{l_x}{(x'_2 - x'_1)} = l_x$$

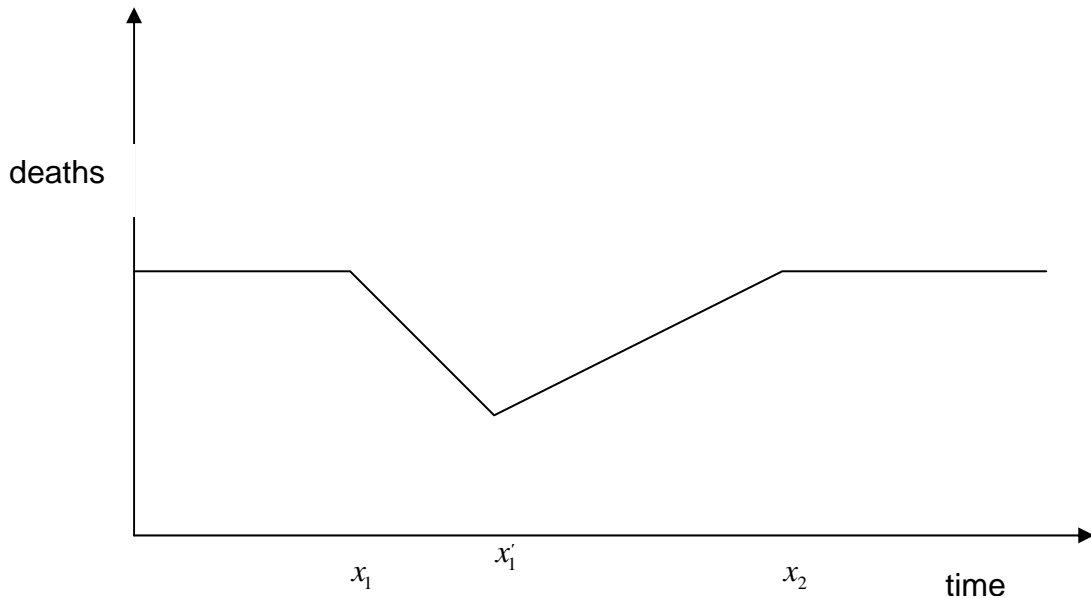


Figure 9: Transition in annual number of deaths over adjustment period

### Generalised Model

We can see above that there are sections where two of the models experience similar rates to each other. In fact, what drives the three models are similar ideas i.e. how many 'old' and 'new' lives are dying at a given point in time. Allowing  $x'_1 = x_1$  or

$x'_2 = x_2$  then we can generalise the model initially into two parts; old lives and new lives.

### Old lives

$$\text{death rate} = \begin{cases} (x_2 - x_1) \frac{l_x}{(x_2 - x_1)} = l_x & t \leq x_1 \\ (x_2 - t) \frac{l_x}{(x_2 - x_1)} & x_1 < t \leq x_2 \\ 0 \frac{l_x}{(x_2 - x_1)} = 0 & x_2 < t \end{cases}$$

### New lives

$$\text{death rate} = \begin{cases} 0 \frac{l_x}{(x'_2 - x'_1)} = 0 & t \leq x'_1 \\ (t - x'_1) \frac{l_x}{(x'_2 - x'_1)} & x'_1 < t \leq x'_2 \\ (x'_2 - x'_1) \frac{l_x}{(x'_2 - x'_1)} = l_x & x'_2 < t \end{cases}$$

Of course, bringing the two equations together in this form brings problems with trying to set the boundaries. However, by using 'Max' and 'Min' functions we can combine both sets of equations into one equation.

$$\text{death rate} = (x_2 - \text{Max}[\text{Min}(t, x_1), x_2]) \frac{l_x}{(x_2 - x_1)} + (\text{Min}[\text{Max}(t, x'_1), x'_2] - x'_1) \frac{l_x}{(x'_2 - x'_1)}$$

## 3. The empirical model

The theoretical model above allows many convenient results to be obtained. We now turn to the issues of whether such geometric shapes occur in reality by using the predictions and insights generated by the theory on real data. Our particular interests are ascertaining whether such patterns hold in reality and if they alter through time; if populations are converging to a maximum age with the passage of time; as life expectancy improves how many lives are saved year on year in different age bands, both in the past and future; and finally, using the model generated to forecast age-specific survival.

For the UK life data exists for many years and we are easily able to obtain the expected life and the percentiles at the ages where death occurs. We focus here on life tables for England and Wales for which there is a long established annual record

as far back as 1841. By studying these it is hoped that we can demonstrate that one of the three types of models detailed above can be used to model the results and forecasts we seek.

In the next section we illustrate aspects of continuous or broken evolution, convergence, divergence or parallelism from past data. In particular, results will be presented that show to what extent life is on track to converge to some upper limit and over what time period. This will be supported by a range of illustrative outputs including long-term projections of the annual number of avoided deaths at different ages based on trends in life expectancy and other potentially useful insights.

### Data Sources

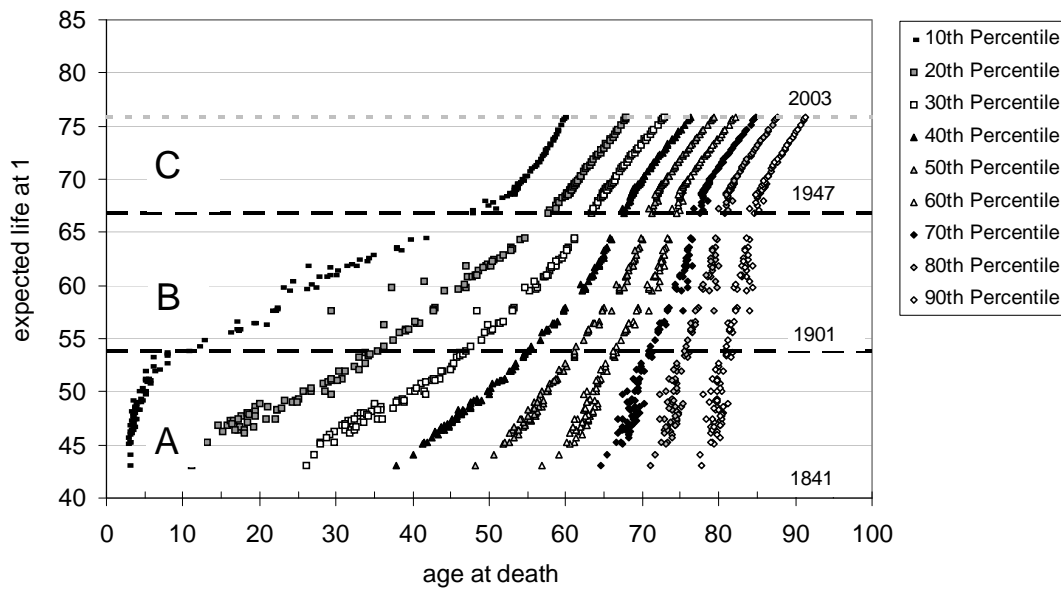
Life tables for England and Wales from 1841 to 2003 are contained in the Human Mortality Database (HMD, 2007), which has been available since 2000 and is the product of a joint collaboration between the Department of Demography at the University of California at Berkeley and the Max Plank Institute for Demographic Research. The HMD was created to provide detailed mortality and population data to researchers, and others interested in the history of human longevity. Its main goal is to document the longevity revolution of the modern era and to facilitate research into its causes and consequences (for more information, see [www.mortality.org](http://www.mortality.org).)

The database contains original calculations of death rates and life tables for 33 national populations, as well as the raw data used in constructing those tables. A complete statement of the methodology used is contained in the methods protocol (Wilmoth et al, 2007). We illustrate examples of each predicted form of the survival curve using different starting ages and time windows. Our start ages are 1 year, 50 years and 80 years with a span of data running from 1841 to 2003, which are split into different time periods. Different time windows or starting ages produce distinctive patterns that are capable of particular interpretation and analysis going forward and backward in time.

### Characterisation of changes based on life expectancy at age 1

In the first example, based on a starting age of 1, we split the series into three distinct eras, each spanning 4 to 5 decades, from 1841 onward. We then use the cut-off points partitioning these eras to similarly compare the patterns obtained at age 50 and 80 respectively. In the accompanying figures, life expectancy is plotted against the percentiles of people surviving to a given age. These percentiles are shown as the 10<sup>th</sup>, 20<sup>th</sup>, 30<sup>th</sup> etc to the 90<sup>th</sup> percentile. The choice of percentile is arbitrary and in the model any combination of values can be used. Since the starting population is 100,000 lives a horizontal line connecting any two adjacent percentiles for a given life expectancy represents 10,000 deaths between the ages indicated on the horizontal axis. As percentile lines move closer together, it is therefore indicative of increased mortality rates within the age ranges indicated by a compaction of mortality into a smaller age range. The percentiles indicate the underlying model; if the percentiles are divergent than model (a), parallel then model (b), and if convergent then model (c).

(a) England and Wales (males)



(b) England and Wales (females)

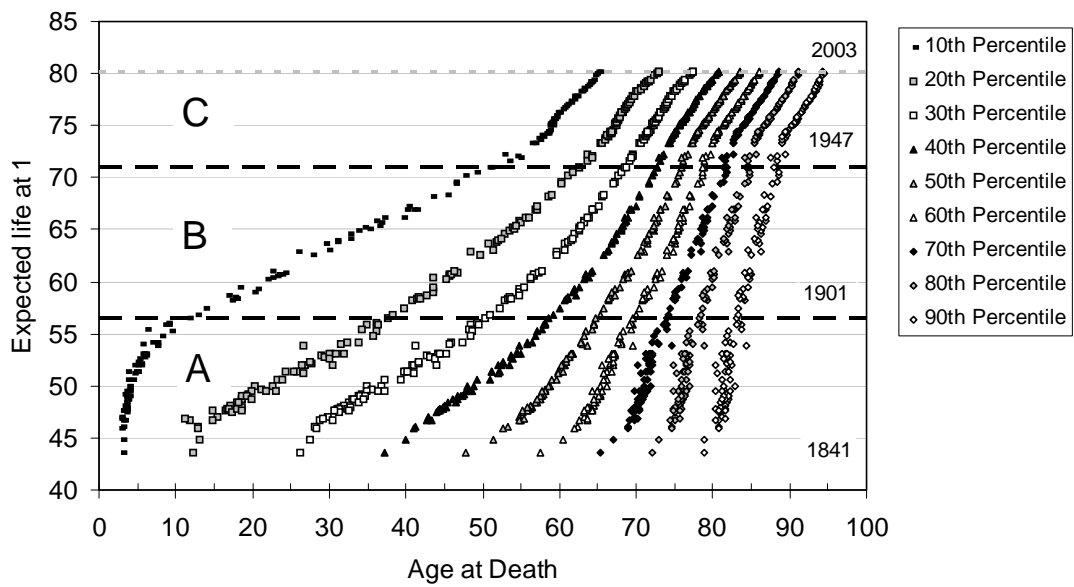


Figure 10: Mortality percentiles as a function of life expectancy at 1, males (A) and females (B) from England and Wales

From the case with a starting age of 1 year in Figures 10 (a) and (b), it appears we can identify three eras as follows:

A – Starting in 1841 and ending in 1900 an era of rising life expectancy at age 1 from around 47 to 54 years (1.2 years per decade) for males and from 48 to 57 years for females (1.5 years per decade). During this era there was persistent high infant and childhood mortality, but reducing health inequalities at older ages as indicated by the convergence in the 20<sup>th</sup> to 90<sup>th</sup> mortality percentiles.

B - Starting in 1901 and ending in 1946, an era of rising life expectancy from around 54 to 67 years (2.8 years per decade) for males and from 57 years to 71 years (3.1 years per decade) for females. This more rapid improvement in life expectancy is mainly due to falling infant and childhood mortality and the continuing reduction in health inequalities as indicated by convergence in the 10<sup>th</sup> to 90<sup>th</sup> mortality percentiles.

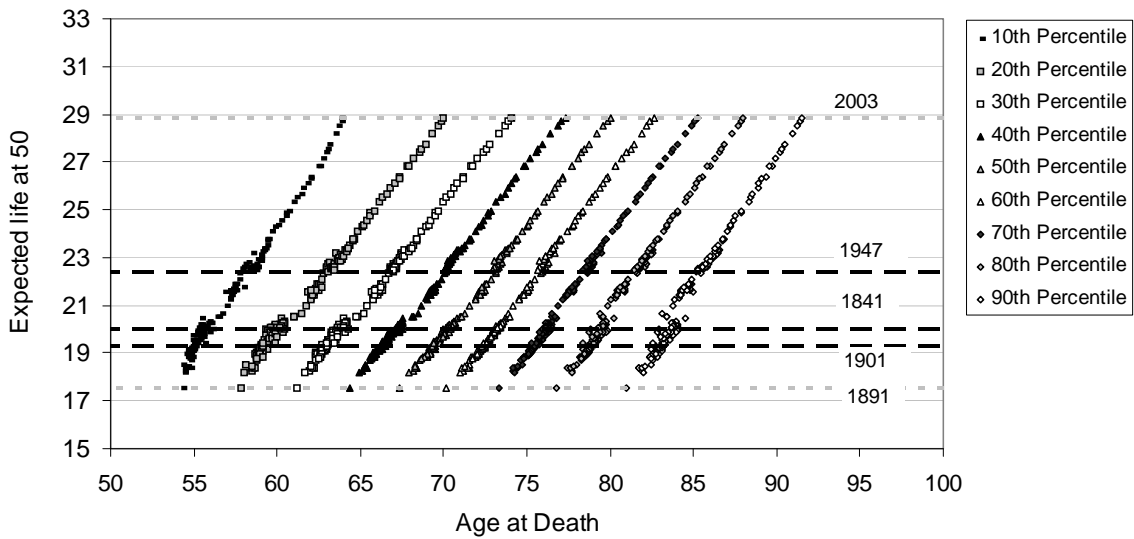
C - Starting in 1947 to the present, an era of continuing rises in life expectancy from 67 to 76 years (1.6 years per decade) for males and from 71 years to 80 years (1.6 years per decade) for females. This slower improvement in life expectancy compared to B is due to the fact that childhood mortality had already been virtually eliminated at the start of this period so little improvement was achieved from reducing this further. Also the fall in convergence in other percentiles meant improvements in life expectancy was improved by a more parallel behaviour. The improvement in the lower percentiles was thus lower than before while the improvement at the higher percentile was greatest in this era.

We are able to split the results into 3 distinct eras because of more or less year on year improvements in life expectancy at this starting age. An exception is male years spanning the First World War, which interrupt the long run trend and have therefore been removed from Figure 10 (a). rather than starting at age 1, we could have used a start age of 0, but in this case we found that there was pattern loss due to high mortality at birth which tends to mask the underlying trend in the 10<sup>th</sup> percentile series in particular.

#### *Starting age of 50*

For later starting ages, continuous improvements are more restricted in time. Figures 11 (a) and (b) show the patterns for males and females with a starting age of 50 years. There are no net improvements in life expectancy before 1901 and only limited improvements up to 1947. From 1947 to the present there has been a significant acceleration with more unbroken gains in life expectancy of around 6 to 7 years for males and females as compared with half that between 1901 and 1946. A key difference with the previous case is that the pattern now appears to be parallel rather than convergent and is unchanged over the whole historical period.

(a) England and Wales (males)



(b) England and Wales (females)

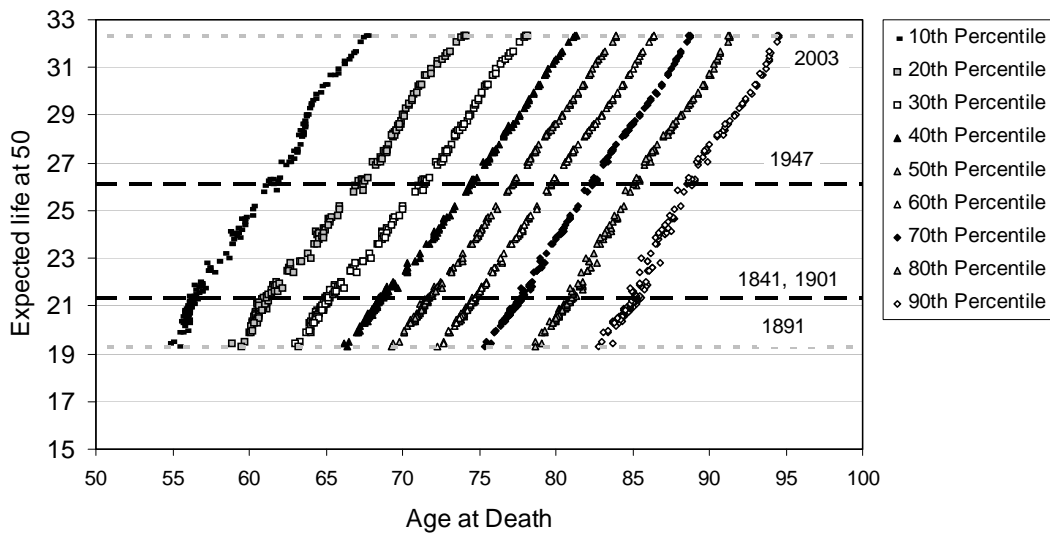
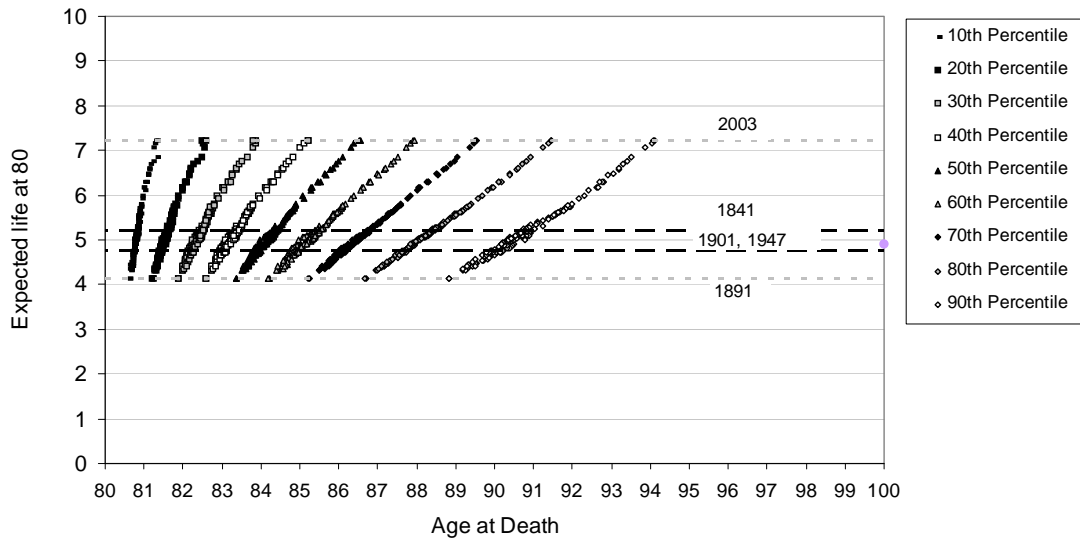


Figure 11: Mortality percentiles as a function of life expectancy at 50, (a) males and (b) females England and Wales

*Starting age of 80*

Figures 12 (a) and (b), consider the pattern based on a starting age of 80 years. In this case it is evident that mortality percentiles acquire a divergent pattern indicating an increasing spread in the probability of death at older age coupled with rises in life expectancy. As in the previous case, up to 1901 the results show that improvements in life expectancy were uneven and that there were no net gains over the period with the pattern holding more or less constant whether life expectancy was increasing or decreasing. Post-1947 however, there has been more or less continuous improvement with life expectancy increases for males and females of between 2.3 and 2.6 years over the period.

(a) England and Wales (males)



(b) England and Wales (females)

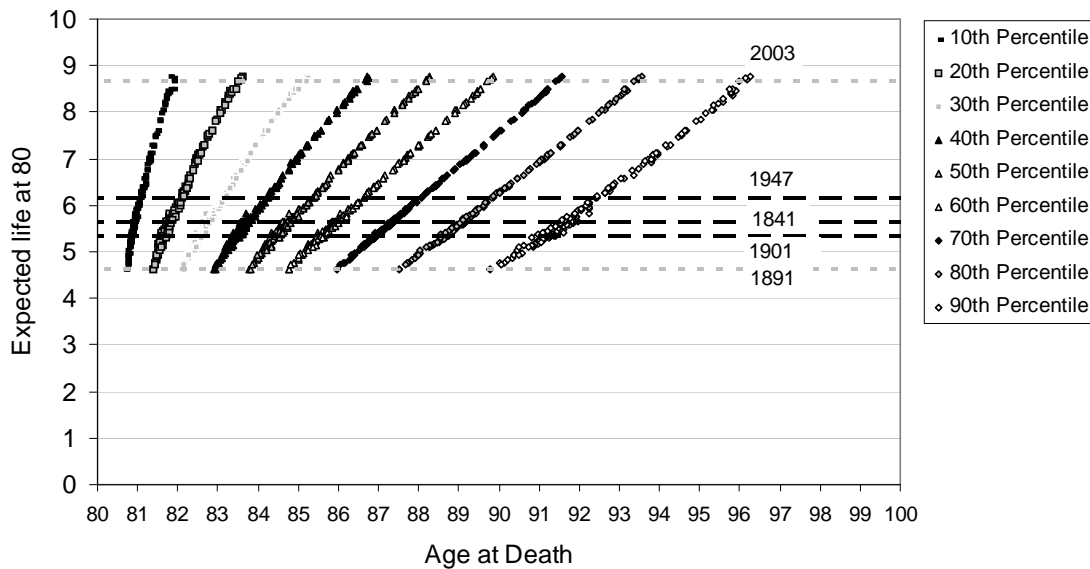


Figure 12: Mortality percentiles as a function of life expectancy at 80, (a) males and (b) females England and Wales

Testing for convergence and maximum age

If these trends are to be used for forecasting future population survival an important question is how such patterns evolve, and whether a maximum age is indicated for example in convergent cases. However, as we have seen, continuation of such patterns is not necessarily guaranteed and if used to predict the future it is dependent on both the pace of progress in life expectancy and any changes in mortality patterns and the extent to which these can be anticipated.

There would appear to be two main ways to project the data trends forward. The first is to project the percentiles forward using a basic relationship with the expected future lifetime as seen in the graphs above. A time frame can then be obtained by looking at how quickly expected future lifetime is increasing with respect to calendar year. The second method is to project the percentiles by comparing them directly to the calendar year.

To some extent this is a theoretical exercise as convergence fizzled out after 1946 but there is value in ascertaining what the maximum age would have been projected to be pre-1947, and how long it would have been expected to take to reach that point. The results can then be benchmarked against what has actually occurred

Consider the convergent pattern for females already indicated in section B of earlier Figure 10 covering the period 1901 to 1947. For the purposes of our analysis survival data for female lives for the year 1918 has been removed, as the influenza pandemic creates distortions in this year. It can be seen that the early deciles are converging towards the later ones though whether the later percentiles are converging or are parallel is harder to ascertain by sight.

To see whether convergence is occurring we fitted a linear regression to each decile. With these regressions we are able to determine which deciles are converging and also predict the expected life at age 1 required and the age of death where convergence would occur. The results are given below in Table 3 and presented in Figure 13.

| Converging percentiles                | Age of death where convergence occurs | Required expected future lifetime at age 1 |
|---------------------------------------|---------------------------------------|--|
| 10 <sup>th</sup> and 20 <sup>th</sup> | 83.91                                 | 83.16                                      |
| 20 <sup>th</sup> and 30 <sup>th</sup> | 84.06                                 | 83.24                                      |
| 30 <sup>th</sup> and 40 <sup>th</sup> | 86.14                                 | 84.96                                      |
| 40 <sup>th</sup> and 50 <sup>th</sup> | 88.47                                 | 87.46                                      |
| 50 <sup>th</sup> and 60 <sup>th</sup> | 92.28                                 | 92.58                                      |
| 60 <sup>th</sup> and 70 <sup>th</sup> | 98.82                                 | 103.19                                     |
| 70 <sup>th</sup> and 80 <sup>th</sup> | 102.89                                | 110.86                                     |
| 80 <sup>th</sup> and 90 <sup>th</sup> | 106.05                                | 117.8                                      |

*Table 3: Details of when deciles converge assuming linear regression*

The above table can be easily interpreted when looking at one individual line. For example the top line states that the 10<sup>th</sup> and 20<sup>th</sup> percentiles will converge when the expected future lifetime at age 1 is 83.16 years and the age of death for these lives is 83.91.

Also, if we look at the required expected future lifetime at age 1 the column indicates that the earlier deciles are converging first and that the last adjacent deciles to converge are the 80<sup>th</sup> and 90<sup>th</sup> percentiles which will only converge when expected future lifetime at age 1 reaches 117.8 years.



However, there seems to be a problem when we look at the two columns together and it is easiest to illustrate with the last line. According to this line the 80<sup>th</sup> and 90<sup>th</sup> percentiles will converge when the expected future lifetime at age 1 is 117.8 but the age at death for these lives will be 106.05.

The problem is that the expected age at death is less than the expected future lifetime but these are the higher percentiles so the age of death should be higher than the expected future lifetime. This is seen in Figure 13 which is based on female data in the era from 1901 to 1947 and includes the fitted regression lines projected to 1988.

As can be seen the inconsistency with expected future lifetime and age at death is that the graph inverts i.e. the 10<sup>th</sup> percentile would cross the 90<sup>th</sup> percentile before the 80<sup>th</sup> and 90<sup>th</sup> percentiles are anywhere near close. In fact the 80<sup>th</sup> and 90<sup>th</sup> percentiles cross when the expected future lifetime is 117.8 by which time the 10<sup>th</sup> percentile lives will be dying at age 178!

There are a number of ways that we can get around this problem and most focus on looking at how the percentiles will behave once they have met. The first method that will lead to the quickest convergence is to consider that as the 10<sup>th</sup> percentile crosses the other percentiles it dominates and the improvement in mortality continues at the same pace that the 10<sup>th</sup> percentile has shown. In this case convergence will occur where the 10<sup>th</sup> and 90<sup>th</sup> percentiles converge. This takes place when the lives are expected to die at age 94.45.

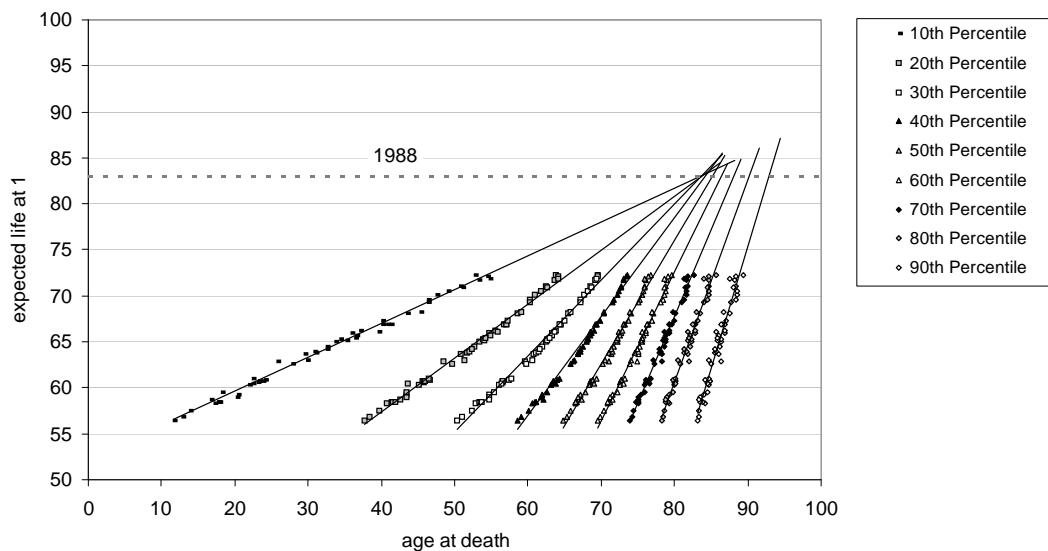


Figure 13: Convergent case based on females from England and Wales, 1901 to 1946 with fitted regression lines

However, there is very little justification in the above assumption. As causes of death among the prematurely dying are a consequence of social, environmental or lifestyle changes it is hard to mount a case that improvements will continue at such a pace. Therefore when the 10<sup>th</sup> percentile meets the 20<sup>th</sup> percentile it is far more likely that the 20<sup>th</sup> percentile's rate of mortality improvement carries on into the future. If this is the case then total convergence will only occur when the last percentiles meet, in this

case the 80<sup>th</sup> and 90<sup>th</sup> percentiles which will meet when the lives are expected to die at age 106.05.

*Projecting time to convergence*

While obtaining the ages of possible convergence is useful we also want to know if this will be in 10 years time or 1,000 years time. As we are basing the values of the percentiles on expected future lifetime we can analyse how fast this value is increasing and from this determine our convergence values.

For the years 1901-1951 a simple linear regression gives the equation

$$\text{Expected future lifetime} = \text{calendar year} * 0.3109 - 517.26$$

Under our first scenario we had the 10<sup>th</sup> and 90<sup>th</sup> percentiles converging when lives were expected to die at age 94.45. This equates to a nominal expected future lifetime of 87.038 (this is nominal as if we have convergence expected future lifetime would be 94.45-1 = 93.45). The value of 87.038 is obtained, assuming our regression is correct, in the calendar year 2001.65. In other words we should now have had convergence of mortality for females in the UK!

Under the second scenario we had the assumption that the slower improving percentile would dominate when percentiles meet. We can therefore reproduce table 3 from above and now calculate the year that these percentiles merge.

| Converging percentiles                | Age of death where convergence occurs | Required expected future lifetime at age 1 | Calendar Year this occurs |
|---------------------------------------|---------------------------------------|--|---------------------------|
| 10 <sup>th</sup> and 20 <sup>th</sup> | 83.91                                 | 83.16                                      | 1988                      |
| 20 <sup>th</sup> and 30 <sup>th</sup> | 84.06                                 | 83.24                                      | 1989                      |
| 30 <sup>th</sup> and 40 <sup>th</sup> | 86.14                                 | 84.96                                      | 1994                      |
| 40 <sup>th</sup> and 50 <sup>th</sup> | 88.47                                 | 87.46                                      | 2003                      |
| 50 <sup>th</sup> and 60 <sup>th</sup> | 92.28                                 | 92.58                                      | 2019                      |
| 60 <sup>th</sup> and 70 <sup>th</sup> | 98.82                                 | 103.19                                     | 2055                      |
| 70 <sup>th</sup> and 80 <sup>th</sup> | 102.89                                | 110.86                                     | 2080                      |
| 80 <sup>th</sup> and 90 <sup>th</sup> | 106.05                                | 117.8                                      | 2103                      |

*Table 4: Table showing age of death at convergence for different mortality percentiles and the calendar year when this would be expected to occur*

However, we have seen that expected future lifetime has become a nominal value as the way that we assume the percentiles' behaviour changes when meeting another percentile will affect the real expected future lifetime. We can therefore look at when the percentiles converge by projecting the percentiles instead. Table 5 predicts the year of convergence by projecting either the expected future lifetime (as above), or by looking at the age of death where convergence occurs and projecting the lower percentile or the higher percentile to this age.

| Converging percentiles                | Calendar Year this occurs according to expected future lifetime | Calendar year obtained projecting lower percentile | Calendar year obtained projecting upper percentile |
|---------------------------------------|---|--|--|
| 10 <sup>th</sup> and 20 <sup>th</sup> | 1988  | 1989   | 1988   |
| 20 <sup>th</sup> and 30 <sup>th</sup> | 1989  | 1989   | 1989   |
| 30 <sup>th</sup> and 40 <sup>th</sup> | 1994  | 1994   | 1995   |
| 40 <sup>th</sup> and 50 <sup>th</sup> | 2003  | 2003   | 2003   |
| 50 <sup>th</sup> and 60 <sup>th</sup> | 2019  | 2020   | 2021   |
| 60 <sup>th</sup> and 70 <sup>th</sup> | 2055  | 2056   | 2058   |
| 70 <sup>th</sup> and 80 <sup>th</sup> | 2080  | 2084   | 2087   |
| 80 <sup>th</sup> and 90 <sup>th</sup> | 2103  | 2111   | 2119   |

Table 5: Table showing the calendar year of convergence for different percentiles based on projecting the percentiles forward

### Testing for parallelism

After 1951 the pattern changes from convergence to an apparently more parallel regime, at least visually, in which there is no convergence and therefore no upper age limit indicated. However, it is hard to find an occurrence of perfect parallel lines which last for any length of time. To illustrate this, below is a plot for male lives from 1952-2003 aged 50.

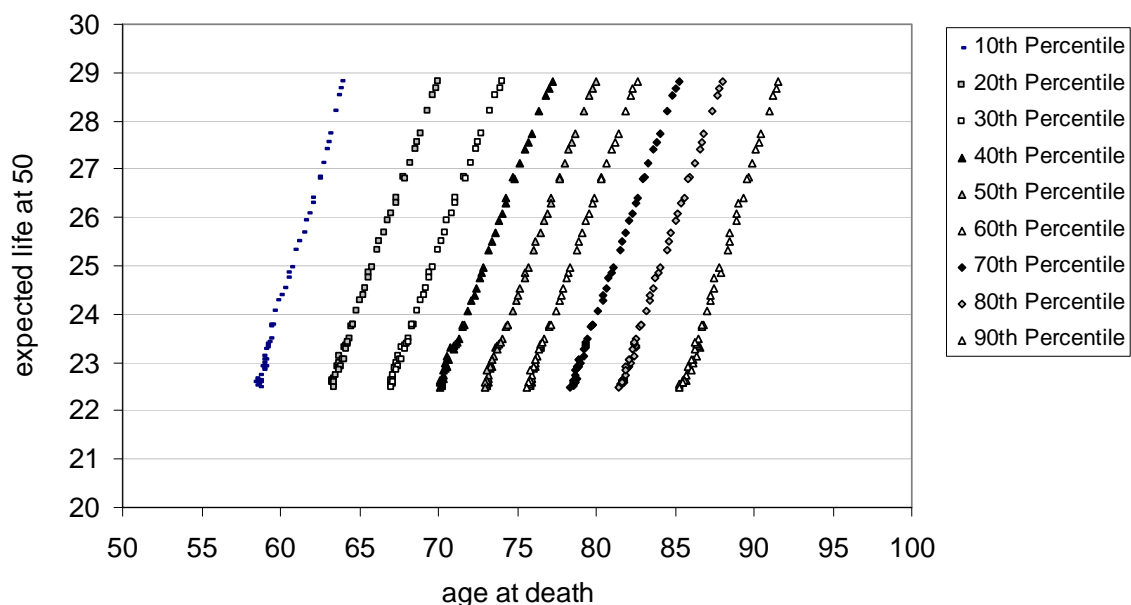


Figure 14: Mortality percentiles for males from England Wales at age 50

Looking at the plot we can see that the lines are parallel with the exception of the first two deciles which appear to be diverging. Fitting regressions to these percentiles and testing for convergence gives the results in Table 6.

| Converging percentiles                | Age of death where convergence occurs | Required expected future lifetime at age 50 |
|---------------------------------------|---------------------------------------|---|
| 10 <sup>th</sup> and 20 <sup>th</sup> | N/A                                   | N/A   |
| 20 <sup>th</sup> and 30 <sup>th</sup> | N/A                                   | N/A   |
| 30 <sup>th</sup> and 40 <sup>th</sup> | N/A                                   | N/A   |
| 40 <sup>th</sup> and 50 <sup>th</sup> | 338.78                                | 262.49                                      |
| 50 <sup>th</sup> and 60 <sup>th</sup> | 287.95                                | 216.61                                      |
| 60 <sup>th</sup> and 70 <sup>th</sup> | 187.63                                | 124.89                                      |
| 70 <sup>th</sup> and 80 <sup>th</sup> | 139.26                                | 79.53                                       |
| 80 <sup>th</sup> and 90 <sup>th</sup> | 158.3                                 | 98.35                                       |

Table 6: Table showing age of death at convergence and expected future life time at age 50

This table is very different to the previous table. For the 10<sup>th</sup>, 20<sup>th</sup> and 30<sup>th</sup> percentiles we have divergence so we do not have values for convergence. Even the other percentiles which do converge do not do so for a number of years. For example, the first deciles to converge are predicted to be the 70<sup>th</sup> and 80<sup>th</sup> percentiles but this will only occur when the expected future lifetime of a 50 year old male is 79.53 years and these particular lives will live until they are 139.26 years old!

#### Testing for divergence

The divergence case appears to be restricted to the oldest starting ages. A convenient way to consider the pace of inter-decile divergence is either as a function of life expectancy, as in this illustration (Figure 15) at age 80, or a function of time.

Figure 15 shows that the inter-decile range (the gap in years between the 10<sup>th</sup> and 90<sup>th</sup> percentile) increases in direct proportion to life expectancy. For each one year increase in life expectancy, the female inter-decile range increases by 1.2 years (male by 1.5 years).

Note that there is a slight hint that the line may be non-linear, that the pace of divergence is slowing and that that divergence cannot proceed indefinitely in the current phase of human evolution.

The data points range from 1841 to 2003, but increases have not been continuous through time. As Figure 16 shows, until 1950 the inter-decile range for females fluctuated between 10 and 12 years, with minima occurring between 1880 and 1890. In this case the best fit curve is a second-degree polynomial ( $R^2 = 0.91$ )

Since 1950 however the inter-decile range it has increased to almost a 15-year spread and may reach 16 years by 2020 based on current trends. In practical terms this means that for any female reaching age 80 there is a widening spread of age of death and that once life expectancy at 80 attains 11 years there will be a 10% chance of living to 100.

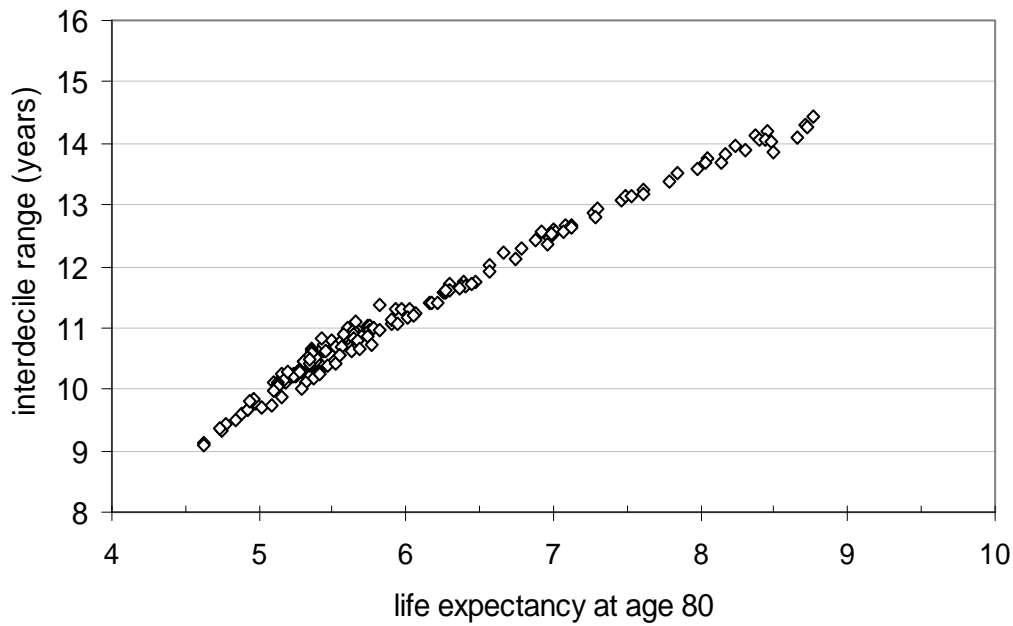


Figure 15: Inter-decile range as a function of life expectancy at 80, females

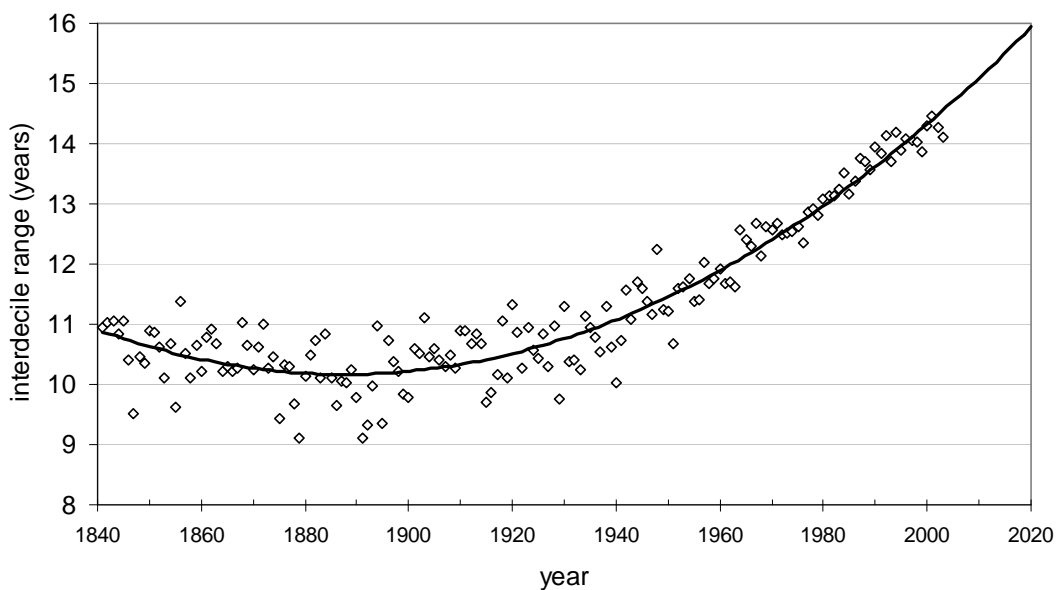


Figure 16: Trend in the inter-decile range for females from 1841

#### 4. Development of projections

In order to develop projections of these trends and produce future population survival estimates, we can split the tasks into five stages as follows:

1. Establish relationship between the percentiles and the expectation of life at a given age
2. Establish nature of linear relationship between calendar year and expectation of life

3. Project forward expectation of life using the relationship found in stage 2
4. Derive survival percentiles using the projected expectation of life and the relationship derived in stage 1
5. Compare projections with projections from an independent source and with actual realised values.

The Gompertz-Makeham formula

A problem with using the full life table for this process is the volume of data required. A more suitable solution can thus be to determine a function that fits the data so that only a few parameters are required. The function chosen to fit the data is a form of the Gompertz-Makeham Model. The Gompertz-Makeham Model provides a function for the force of mortality and is defined as:

$$\mu_x = A + Bc^x \text{ or } \mu_x = A + Be^{\gamma x} \text{ where } \gamma = \ln(c)$$

Traditionally the curve is fit over all ages (though not very young ages) and it is assumed that the constant A mainly deals with non-age related deaths such as accidents. As a result this constant is constrained by the fact that it cannot be negative as this would imply people coming back to life due to accidents. However, for this fit we have allowed the constant to be negative if this gives a better fit overall for the ages we are interested in (50-85).

The survival function for a life is simply defined as:

$$S(x) = \exp\left\{-\int_0^x \mu_s ds\right\}$$

However, we are interested in the survival function for a life already aged 50 so:

$$S(x) = \exp\left\{-\int_{50}^x \mu_s ds\right\} \quad x \geq 50$$

For a population assumed to have 100,000 people alive at aged 50 then multiplying the equation above by 100,000 will give the number of people alive at age  $x$ .

$$L_x = 100,000 \exp\left\{-\int_{50}^x \mu_s ds\right\} \quad x \geq 50$$

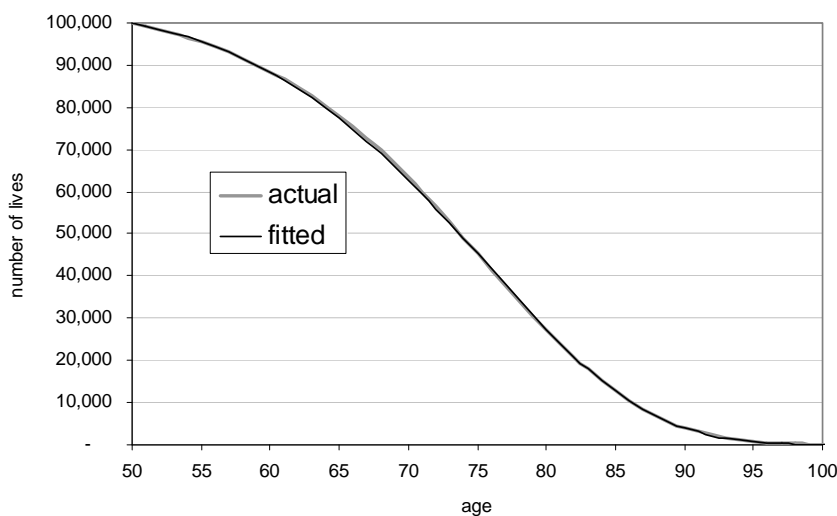
So using the Gompertz-Makeham formula and an age  $t$  where lives are expected to start dying from, then the survival function  $S(x)$  is:

$$\begin{aligned}
S(x) &= \exp \left\{ - \int_t^x \mu(s) ds \right\} \\
&= \exp \left\{ - \int_t^x A + Be^{\alpha s} ds \right\} \\
&= \exp \left\{ - \left[ As + \frac{Be^{\gamma s}}{\alpha} \right]_t^x \right\} \\
&= \exp \left\{ - \left[ Ax - At + \frac{Be^{\gamma x}}{\gamma} - \frac{Be^{\gamma t}}{\gamma} \right] \right\} \\
&= \exp \left\{ A(t-x) + \frac{B(e^{\gamma t} - e^{\gamma x})}{\gamma} \right\}
\end{aligned}$$

### *Fitting the Curve*

The curve was fitted using iteration. We focused on the ages 50-85 which we split into six groups each with six ages i.e. 50-55, 56- 61, etc. The parameter A has most affect on the first two groups, B has most affect on the middle two groups and finally C has most affect on the last two groups. The iterative method worked by finding the group with the largest difference between expected and observed and changed the relevant parameter. This process was repeated until all deviations were of an acceptable amount.

By allowing the parameters to be flexible, the fits for all the years were more than adequate. Figure 17 below shows the fit for the population curve for the year 1973 compared to the actual lives recorded (1973 is used as a representative year approximately midway through the period under investigation).



*Figure 17: Chart showing fit between the actual and predicted survival based on data from 1973*

It can be seen that the fit captures the general shape of the curve very well. While the actual curve is higher for long periods, the gap between the curves is never large as the mortality rates are very similar. The parameters that were fitted for 1953-2003 in steps of 10 years are given in Table 7.

| Parameter | 1953     | 1963     | 1973     | 1983     | 1993     | 2003    |
|-----------|----------|----------|----------|----------|----------|---------|
| A         | -0.00318 | -0.00639 | -0.00542 | -0.00331 | -0.00317 | 0.00062 |
| B         | 0.00016  | 0.00024  | 0.00021  | 0.00011  | 0.00008  | 0.00001 |
| C         | 1.08803  | 1.0832   | 1.08346  | 1.09091  | 1.0927   | 1.11208 |
| $\gamma$  | 0.08437  | 0.07992  | 0.08016  | 0.08701  | 0.08865  | 0.10623 |

Table 7: Parameter values for the Gompertz-Makeham formula

#### Technical observation

There are trends in the fitted parameters though as each fit was done independently there is no smoothing of parameters to highlight the trends. The data from 1953 shows the randomness that mortality data always demonstrate and had noticeably lower mortality than the surrounding years. Ignoring 1953 we can see that parameter B has fallen while parameter C has increased. This is expected as the fall in parameter B means that less lives are dying at younger ages and parameter C means that more lives are dying at older ages (which is expected as although mortality rates are improving there are more lives reaching these later ages). However, we would expect parameter A to also fall as well as this affects younger lives but this has actually increased (i.e. become less negative). As discussed earlier, the parameter A is not normally allowed to go negative as it is seen as the lives that die through non-age related deaths. However, to get a better fit we allowed this to happen and so parameter B meant too many younger lives die and the negative parameter A compensated for this excessive loss. As B has decreased to capture the shape of the curve later on, parameter A hasn't had to adjust the lives as much and so has become larger.

#### Calculating the lives saved from 1953 to 2003

The number of lives in the population is simply the area under the curve. The easiest way to determine the population at any point should therefore be to integrate the survival function. Unfortunately, the survival function we have used is not an integrateable function. We therefore need to use a trapezoid approximation to get our population. The population aged  $x$ ,  $l_x$  is given by:

$$l_x = l \frac{1}{2} (S(x) + S(x+1))$$

where  $l$  is the population size for the first age being considered. Of course, we can make this approximation more accurate by dividing the year into smaller intervals but the difference that this makes is negligible compared to the random error that occurs when observing mortality rates and additionally when fitting the Gompertz-Makeham formula.

Table 8 below gives the population size for each of the six selected years while Table 9 gives the difference in population when compared to 1953 i.e. the lives saved



assuming 100,000 lives were aged 50 and the populations at any point in time are stable. Based on Table 8 we see that compared with 1953, 2.9m would be alive in 2003 as compared with 2.3m in 1953; this represents an accelerating trend as is seen in Table 9 which compares lives saved group by age. It can be seen that in terms of numbers the greatest change is in the 70-79 age group with large changes also seen in the 60-69 and 80-89 age groups. In percentage terms the changes increase as the age group gets older.

| Age range | 1953      | 1963      | 1973      | 1983      | 1993      | 2003      |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 50-59     | 945,493   | 948,745   | 951,703   | 959,817   | 968,880   | 974,978   |
| 60-69     | 752,275   | 752,548   | 768,155   | 802,122   | 837,397   | 880,654   |
| 70-79     | 430,495   | 422,652   | 451,946   | 504,878   | 566,651   | 667,014   |
| 80-89     | 121,532   | 116,636   | 137,988   | 169,537   | 219,995   | 313,694   |
| 90+       | 11,949    | 11,596    | 16,349    | 21,668    | 35,750    | 61,291    |
| total     | 2,261,744 | 2,252,177 | 2,326,141 | 2,458,022 | 2,628,673 | 2,897,631 |

Table 8: Population size grouped by age

| Age range | 1953 | 1963   | 1973   | 1983    | 1993    | 2003    |
|-----------|------|--------|--------|---------|---------|---------|
| 50-59     | n/a  | 3,253  | 6,211  | 14,324  | 23,387  | 29,485  |
| 60-69     | n/a  | 273    | 15,881 | 49,847  | 85,123  | 128,379 |
| 70-79     | n/a  | -7,842 | 21,451 | 74,383  | 136,156 | 236,519 |
| 80-89     | n/a  | -4,897 | 16,455 | 48,004  | 98,463  | 192,162 |
| 90+       | n/a  | -353   | 4,399  | 9,719   | 23,801  | 49,342  |
| total     | n/a  | -9,566 | 64,397 | 196,277 | 366,930 | 635,887 |

Table 9: Lives saved compared to 1953 grouped by age

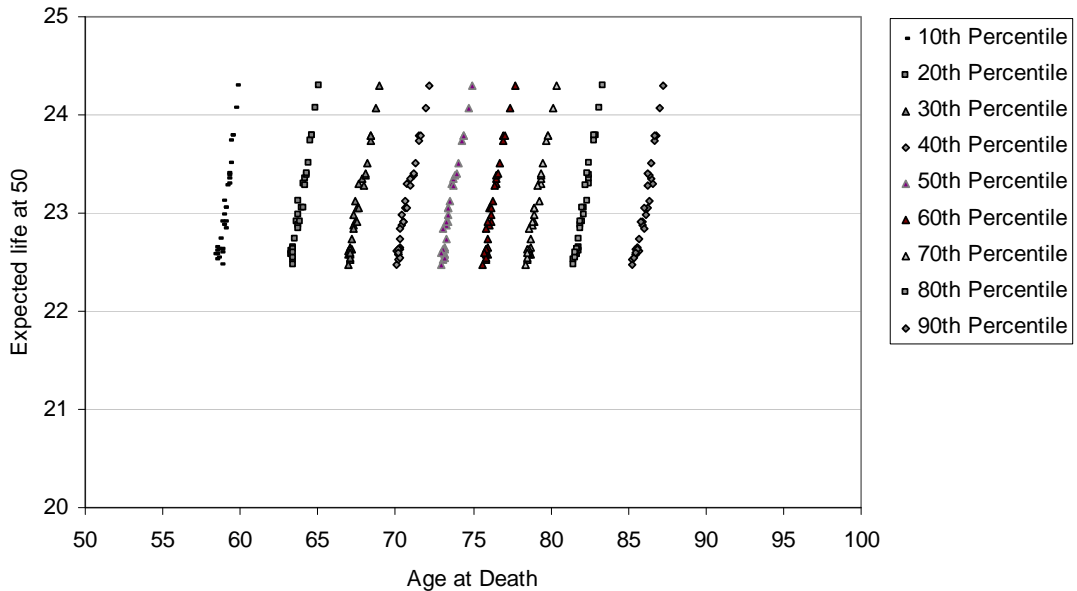
### ***Justification of Projections going Forward***

It can be seen above that the Gompertz-Makeham formula can fit the data well and allow calculations on the number of lives ‘saved’. We have previously shown that the life tables show systemic patterns in terms of their evolution over time. The question arises as to whether we can combine these findings into a tool that allows us to construct life tables in the future based on projected trends. The validity of the approach will shortly be thoroughly tested to see if it is capable of more accurate forecasts than those officially published.

Just because all of the stages listed earlier can be carried out does not justify this as a suitable projection method. Our method of validation involved backward testing by comparing our results with forward projections from base years in 1981 and 1991 produced at the time by the Government Actuary’s Department and with what actually occurred. In doing so we wanted to establish whether our approach predicted actual change more accurately than Government Actuary Department’s (GAD)’s own forecasts at the time. By using these projections we have sufficient time from the date of the prediction to the last easily available life table (2003) to assess the quality of the projection.

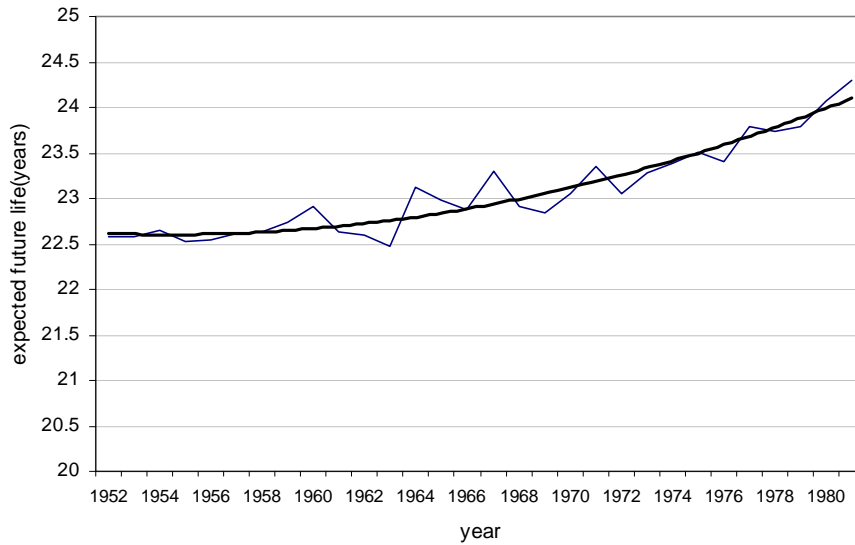
We used data from 1952 to 1981 to generate our own 1981 projection and data from 1952 to 1991 to generate our 1991 projection. It can be argued that the GAD projections would not have been able to use all of this data because of the time taken

to calculate their projections but as our method is very quick we should have been able to use all of the data available. Using 1952-1981 we get the following relationships between the percentiles and the expectation of future life at age 50 shown in Figure 18 for English and Welsh males. Of course, this data makes up a subset of the data in Figure 14 above and all the illustrated percentiles can certainly justify a straight line fit.

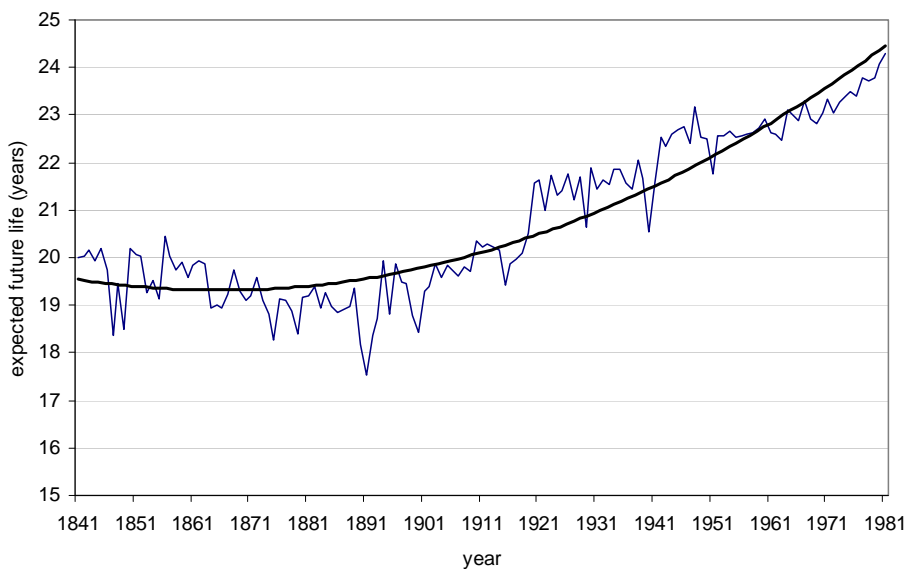


*Figure 18: Mortality percentiles for England and Wales males at age 50 for data from 1952-1981*

For life expectancy and calendar year a polynomial of order 2 was used to capture the shape as shown in Figure 19. It can be argued that this is not the correct equation to use as it implies that expected life pre-1952 is actually higher the further back we go. However, this phenomenon is often seen and a similar shape is observed if we use data from 1841-1981 as is shown in Figure 20. As we are only needing to project until 2020 (the length of the GAD projection) it was decided that the fit using 1952-1981 data would be sufficient to demonstrate whether our method produced comparable or different results to GAD.



*Figure 19: Expected life at age 50 for England and Wales males for 1952-1981 with regression*



*Figure 20: Expected life at age 50 for England and Wales males for 1841-1981 with regression*

### ***Suitability of Projections***

It should be noted that the GAD projection was obtained by constructing life tables using the projected mortality rates.

To show how suitable the projections are we can now compare the projections from our model, the GAD projection and actual survival for the years 1990 and 2000. The key graphs showing the life tables for 1990 and 2000 are given in Figures 21 and 22. These compare GAD forecasts with our own model and with actual survival using subsequently published life tables.

Looking at these two snapshots there are some clear similarities. For both years and at all ages the number of lives is highest for the realised values, followed by our model, followed by the GAD projection. It can also be seen that for the older ages our projection is closer to GAD than the realised value while for the younger ages our projection is closer to realised than it is to the GAD projection.

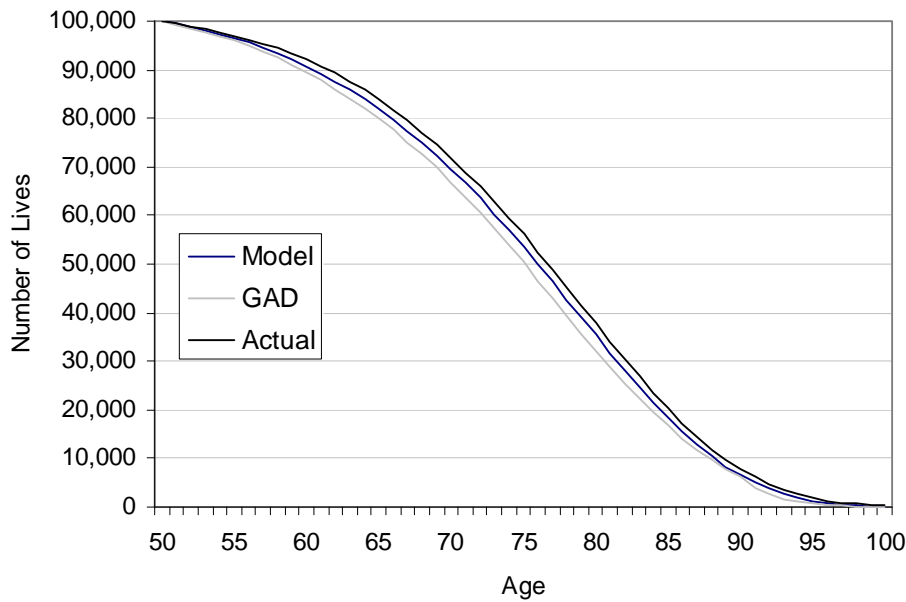


Figure 21: Comparison of projected life tables for 1990 for England and Wales males

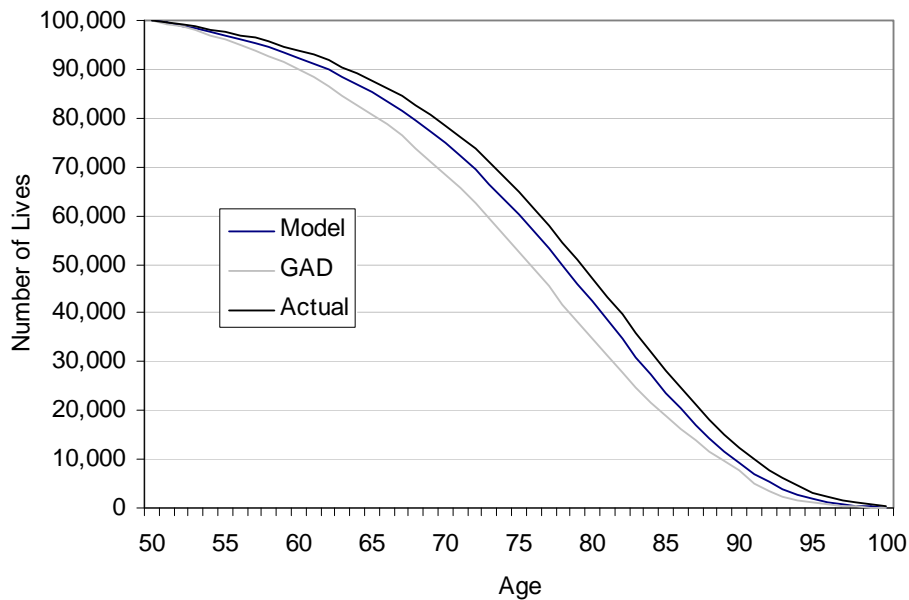
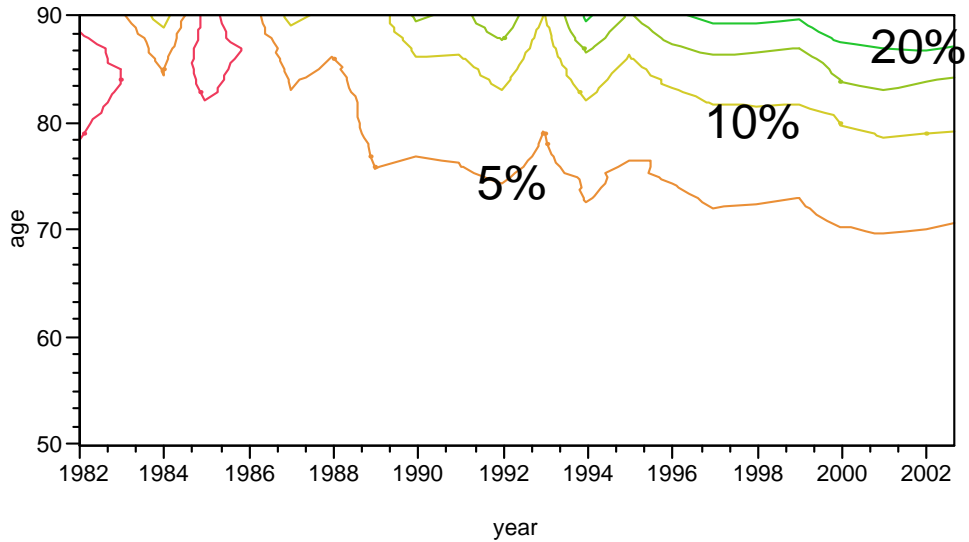


Figure 22: Comparison of projected life tables for 2000 for England and Wales males

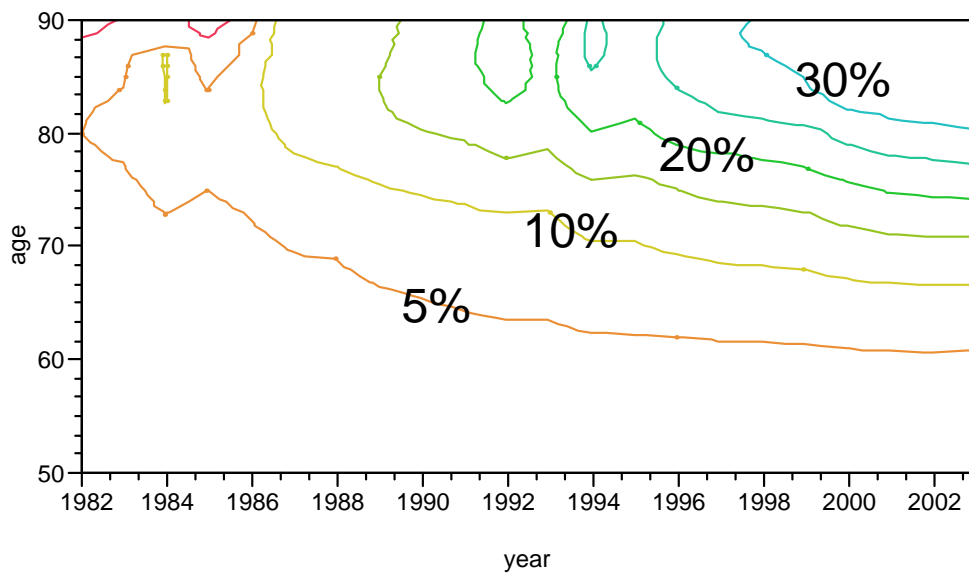
#### Continuous comparisons

Rather than looking at just a couple of snapshot graphs it is possible to look at differences in continuous time by the use of contour plots. These plots work in a similar way to contour lines on geographical maps where the size of the 'hill' in this

case is the size of the difference at a particular time and particular age. Each contour measures a constant percentage difference either between our model and actual data or between GAD forecasts and actual data. Figure 23 (A) shows a plot for our model and Figure 23 (B) based on GAD projections.



(A) Percentage difference between actual and model as a function of age and calendar year



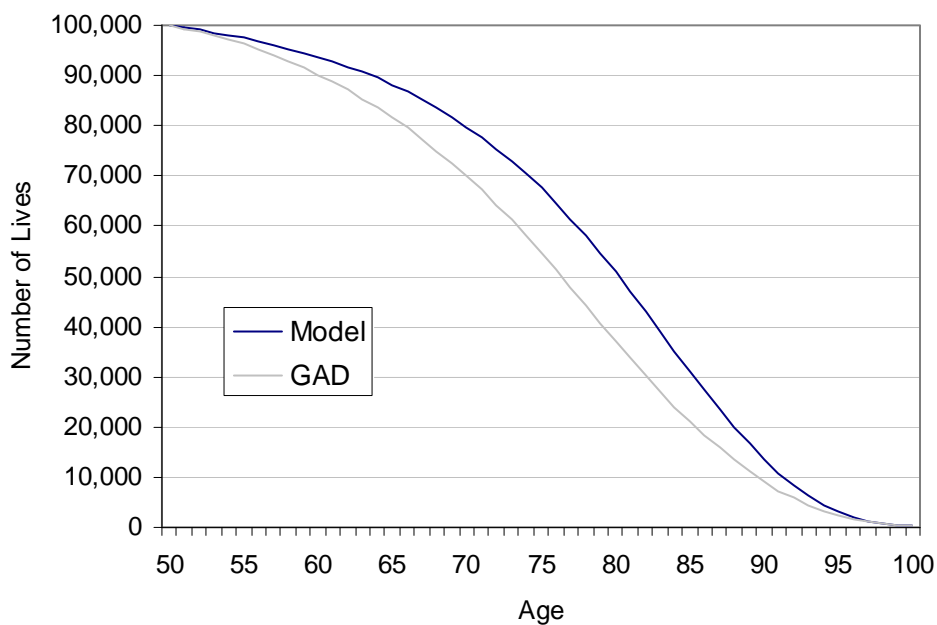
(B) Percentage difference between Actual and GAD as a function of age and calendar year

Figure 23: Contour plot showing differences between actual and model and actual and GAD

The contour maps illustrate that the two snapshots are correct in the impression they give; namely that at all ages and for all years our model was able to get a more accurate projection to the mortality observed than the GAD projections. As the model currently fits a life-table to each year individually there is less smoothness in the deviations compared to the GAD model. This is represented by the contour lines being more irregular.

## 5. Future Projections

Although we cannot compare with reality we can still see the differences between the GAD projection and our model for the years up to 2020. If there are significant differences this could have implications for pensions, health care and other public services that rely for their planning on accurate population forecasts.



*Figure 24: Comparison of projected life tables for 2010 for England and Wales males*

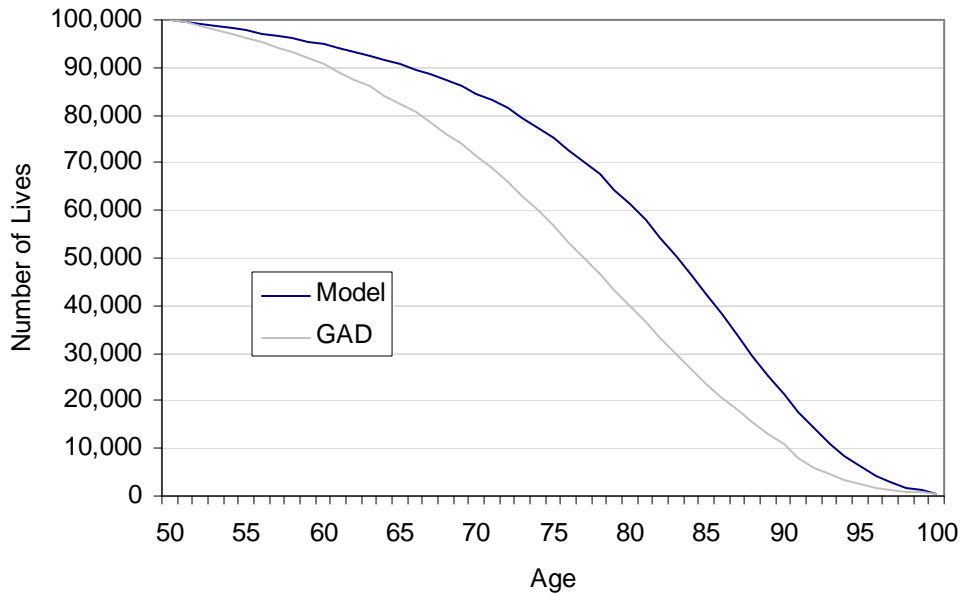


Figure 25: Comparison of projected life tables for 2020 for England and Wales males

We can see that the trends already apparent in the first two graphs continue with the difference in projected lives increasing at all ages as lives get projected further into the future. This is further illustrated by using a contour map of the difference between our model and GAD for all years from 1981-2020 shown in Figure 26.

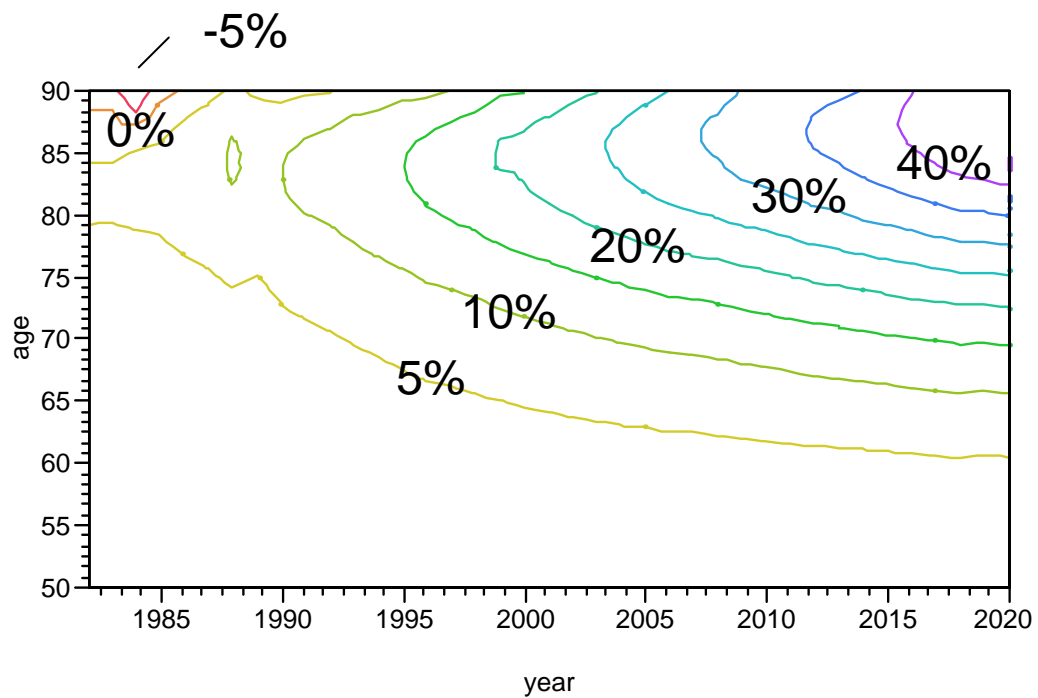


Figure 26: Contour plot showing differences between model and GAD

As can be seen the deviations as a percentage between the two models increase with both year of projection and age that we are studying.

### 1991 Based Projections

Using the same five stage process as highlighted above, a projection can be based in 1991 that uses the data from 1952-1991. For this data we have the following relationship between the percentiles and the future expectation of life as shown in Figure 27.

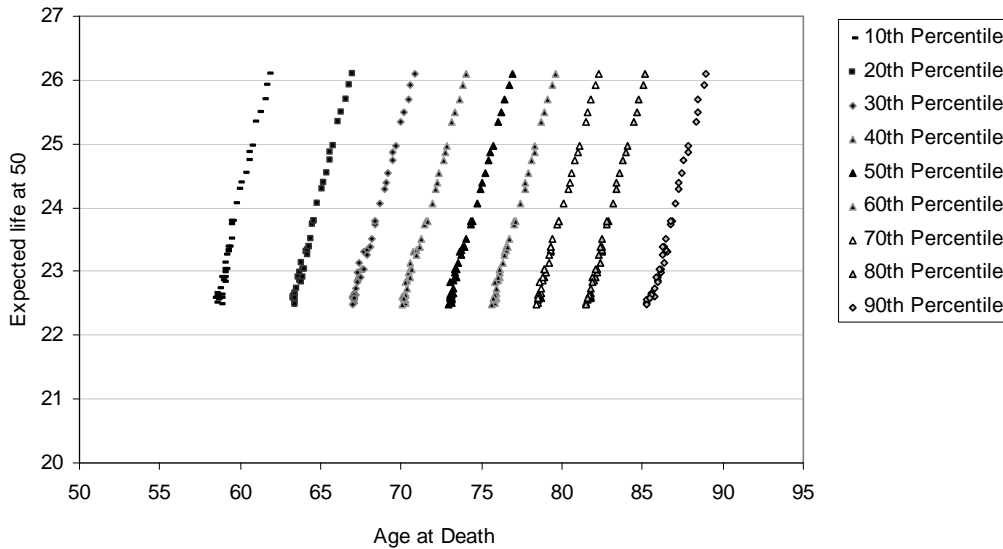


Figure 27: Mortality percentiles for England and Wales males at age 50 for data from 1952-1981

As expected, there is no reason to change our assumption that there is a basic linear relationship between the percentiles and the expectation of life over this data range.

We also have the following relationship between expectation of life and calendar year shown in Figure 28. Once again, fitting a polynomial of order two appears to capture the shape of this relationship although once again we see a point of inflection around the year 1956 which implies that people were expected to live longer in the past.



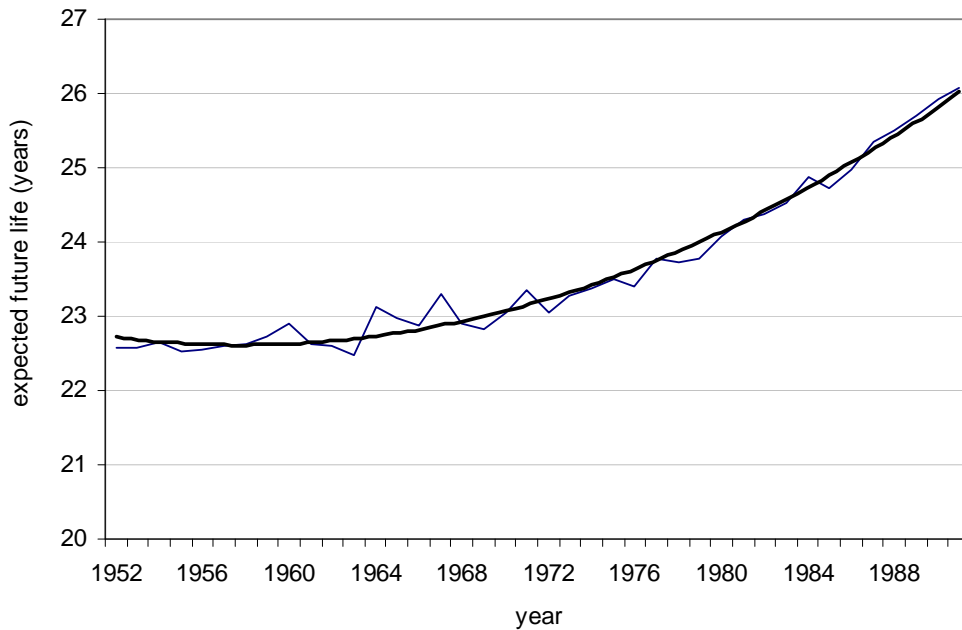


Figure 28: Expected life at age 50 for England and Wales males for 1952-1991 with regression

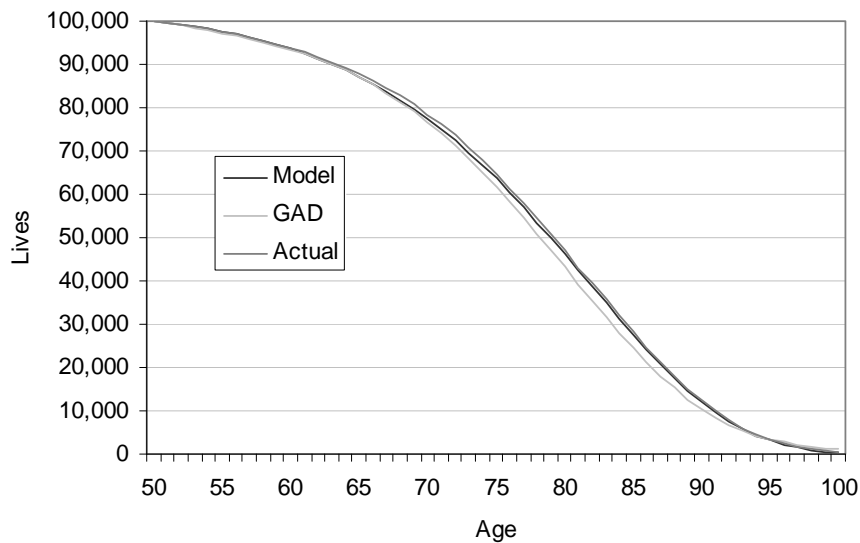


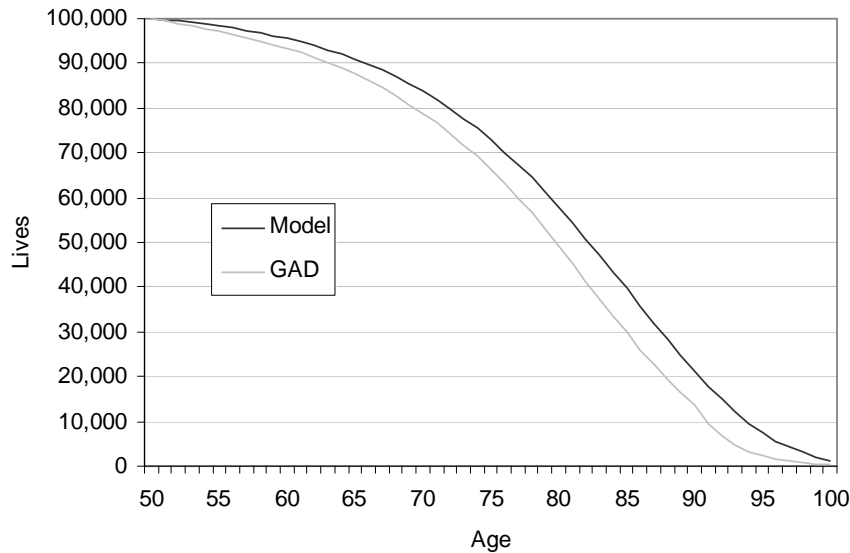
Figure 29: Comparison of projected life tables for 2000 for England and Wales males

From Figure 29 we can see that once again we have the same pattern of our model predicting less lives at each age than was realised but with GAD underestimating the number of lives by a greater extent. Similarly we see again that at the younger ages our model is closer to GAD and for later ages we are closer to the realised values.

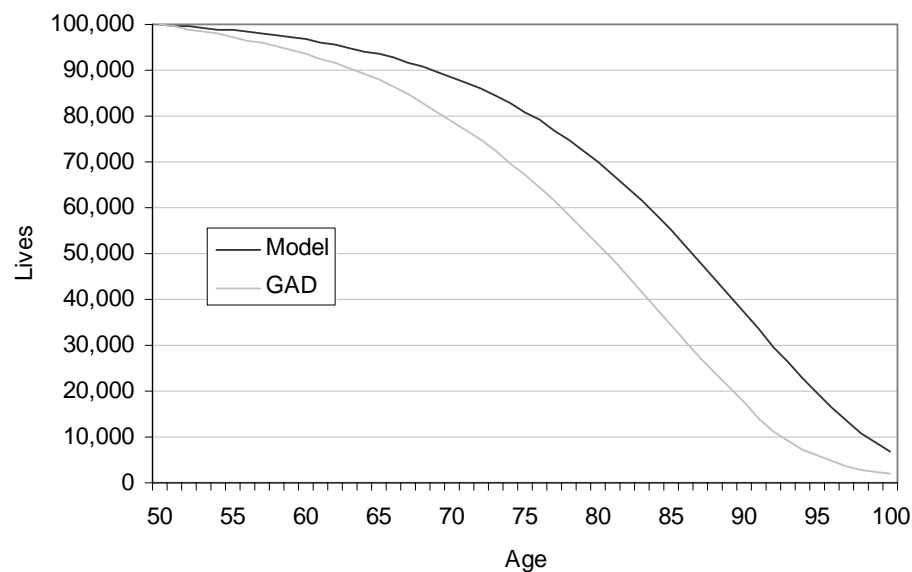
When we observe the above graph and compare it to the graph for the 1990 projection using data up to 1981, we can see that although both projections are for nine years the model is far closer to the realised values in the 2000 projection using 1991 data. This

is because the assumed increase in life expectancy is greater for the later projection as we have more recent data that captures this increasing rate.

Once again we can compare GAD projections and our model for the future even though we cannot of course compare with realised values. This is shown in Figures 30 and 31.



*Figure 30: Comparison of projected life tables for 2010 for England and Wales males*

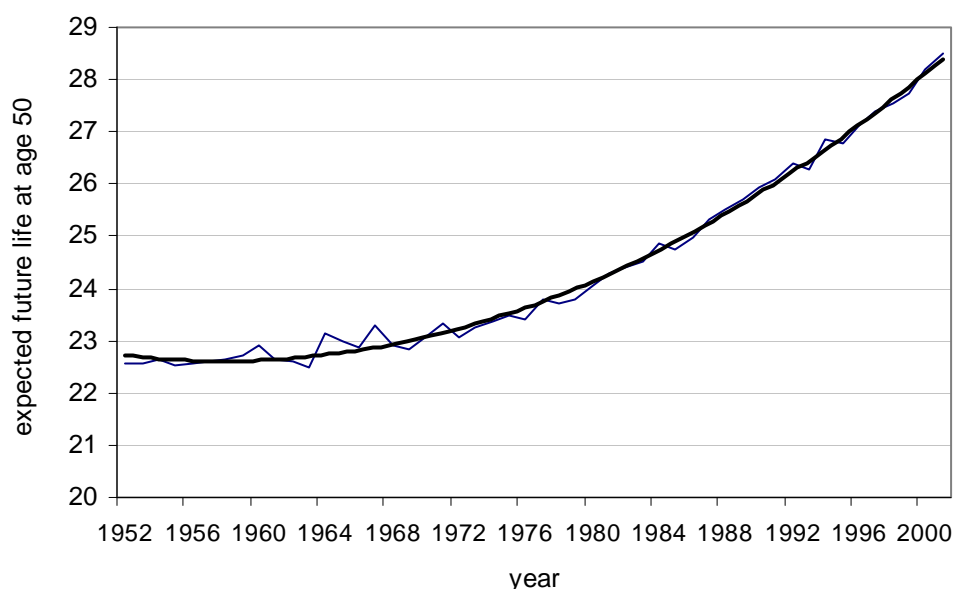


*Figure 31: Comparison of projected life tables for 2020 for England and Wales males*

Again we see the difference between the projections increasing as we move further into the future. This is due to GAD believing that the increase in life expectancy will slow down in the future whereas we are predicting that the trend will continue for the foreseeable future. There is less difference in the shapes of the curve for 2020 compared to the 1981 projection.

### *2001 based projections*

The 1981 and 1991 based projections allow us to compare our model and GAD's model to realised outcomes. However, the last major GAD projection was 2001 and so it is useful to see how our model and GAD's model for this time will project future survival rates and population sizes. The usual stages were carried out to fit our model and it is useful to note how well the second order polynomial fits the expectation of life data for 1952-2001 as shown in Figure 32.



*Figure 32: Male life expectancy at age 50 from 1952 to 2001*

The percentiles are similar to previous fits so are not shown here and the life tables show similar characteristics in that our model predicts more survivorship at each age with the difference increasing with time and age. Hence it is unnecessary these life tables are not shown here.

### Projections of Population

A more useful demonstration of the results is to compare the number of lives predicted by our model and GAD in the age bands 50-59, 60-69, 70-79, 80-89 for the year 2020 based on the 2001 projection. This was carried out using the projection model detailed on the GAD website with the same assumptions for immigration that this model used. It was also assumed that birth rates and mortality for ages up to and including 49 were the same as GAD leading to both models having the same number of 50 year olds in each calendar year. For ages 50 and above the mortality rates used were those from the model fitted above. The results for male lives are given in table 10 below.

| age   | GAD 2020   | Model      | Diff    | Diff % |
|-------|------------|------------|---------|--------|
| 50-59 | 3,788,205  | 3,809,512  | 21,306  | 0.56%  |
| 60-69 | 3,014,841  | 3,111,925  | 97,084  | 3.22%  |
| 70-79 | 2,324,314  | 2,504,966  | 180,653 | 7.77%  |
| 80-89 | 978,574    | 1,164,099  | 185,525 | 18.96% |
| total | 10,105,934 | 10,590,502 | 484,568 | 4.79%  |

*Table 10: Comparison of male population projected from 2001, model versus GAD*

As expected our model predicts higher numbers of lives in each age band compared to GAD with the percentage increasing as age increases. For the age groups 50-59 and 60-69 the percentage change is small as the difference in assumed mortality rates is small for each age in these age groups. For the 70-79 age group the difference is a lot larger though it should be noted that the percentage differences for this age group would have been a lot larger if we had used projections from 1981 or 1991 as the projected mortality rates of our model and GAD have closed with each projection as we saw earlier.

The largest percentage change by far is the 80-89 year age group. This is because the difference in mortality rates between the models is most prominent in this age range. One concern with this model is the function we have used to predict our future life expectancy. Although we have shown that the model has predicted more accurately than the GAD model for 1981 and 1991 it can be argued that using our model to predict expected future life expectancy in say 2050 is going to give a very large number if our expected trend continues. However, we are not stating that our model should be used for such large projections; merely that over shorter periods of time life expectancy trends continue rather than sharply changing other for events such as war and influenza.

However, to demonstrate that the projection is certainly feasible to 2020 we also fitted an ARIMA model to the data. For the data 1952-2001 the ARIMA model had trouble fitting a suitable model. However for the data 1963-2001 an ARIMA(4,1,0) model fitted very well (see Annex B). Table 11 gives the projected life expectancy from our model, the ARIMA model and also 95% confidence limits from the ARIMA model. We note that apart from 2002 the predicted expected future life expectancy from our model is within the 95% confidence interval of the ARIMA model.

Therefore, although the expectation of life is high it is certainly within a realistic scenario. It should also be noted that our predictions in 1981 and 1991 would also have been considered to be high and yet have turned out to have underestimated the improvements in life expectancy. Note that for females the situation is different and the rate of acceleration is lower and so the potential for error appears to be much lower (see Figure A3, Annex A). This confirmed by the fact that our projections are much closer to official projections over the period.

| Year | Model | ARIMA Forecast | ARIMA Lower 95% Confidence | ARIMA Upper 95% Confidence |
|------|-------|----------------|----------------------------|----------------------------|
| 2002 | 28.95 | 28.58          | 28.25                      | 28.92                      |
| 2003 | 29.23 | 28.99          | 28.62                      | 29.36                      |
| 2004 | 29.52 | 29.31          | 28.87                      | 29.75                      |
| 2005 | 29.82 | 29.42          | 28.81                      | 30.02                      |
| 2006 | 30.12 | 29.70          | 28.97                      | 30.43                      |
| 2007 | 30.43 | 30.05          | 29.25                      | 30.86                      |
| 2008 | 30.74 | 30.17          | 29.21                      | 31.14                      |
| 2009 | 31.06 | 30.39          | 29.28                      | 31.50                      |
| 2010 | 31.39 | 30.72          | 29.51                      | 31.93                      |
| 2011 | 31.73 | 30.89          | 29.54                      | 32.23                      |
| 2012 | 32.07 | 31.05          | 29.55                      | 32.55                      |
| 2013 | 32.41 | 31.35          | 29.74                      | 32.96                      |
| 2014 | 32.76 | 31.55          | 29.81                      | 33.29                      |
| 2015 | 33.12 | 31.69          | 29.81                      | 33.58                      |
| 2016 | 33.49 | 31.96          | 29.95                      | 33.96                      |
| 2017 | 33.86 | 32.18          | 30.05                      | 34.31                      |
| 2018 | 34.24 | 32.32          | 30.05                      | 34.58                      |
| 2019 | 34.62 | 32.54          | 30.15                      | 34.93                      |
| 2020 | 35.01 | 32.78          | 30.27                      | 35.28                      |

*Table 11: ARIMA forecasts compared with model: males 2002 to 2020*

## 6. Conclusions

In this paper we have offered a new way of charting the development of populations through an analysis of trends in human survival based on ordinary life tables. We sought to explain the various different shapes of survival curves starting with a simple model. This conjectured three basic patterns of survival: (a) ‘divergent’, (b) ‘convergent’ or (c) ‘parallel’. Qualitatively speaking (a) might be thought of benefiting older people more than younger people, (b) all age groups equally and (c) younger generations before older generation. Variant (a) predicts that as life expectancy increases inequalities in age of death would increase over time; only variant (c) predicts a maximum age to which everybody would survive and is thus consistent with the process of increasing rectangularisation of the survival curve conjectured by other researchers; in the parallel variant (b) inequalities in age at death would persist but that increases in life expectancy would be shared equally across the population.

The simple model provided many useful technical results as well as hypotheses which were used as a basis for explaining real life tables over a long period. We used data from England and Wales spanning 160 years of evolution for our analysis. The results show a more complex and nuanced picture than portrayed in the three simple variants depending on era and on starting age. We observed that these trends could alter so that from time to time this led to different survival developments paths. Three eras were identified based on survival patterns from age 1. The first was pre-1901 in which there was persistently high childhood mortality coupled with convergence and hence reduced inequalities between the 20<sup>th</sup> and 90<sup>th</sup> survival percentiles. The second ending in 1946 was characterised by more rapid improvements in life expectancy following improvements in infant and childhood mortality. In this period our results showed that trends appeared to predict convergence towards a maximum age that would be

reached in this century, although it was difficult to pin down both the value of the maximum or the year of convergence with accuracy.

The process suggested sequential convergence with the 10<sup>th</sup> and 20<sup>th</sup> percentiles converging first and then the 20<sup>th</sup> with the 30<sup>th</sup> and so on. The reasoning for this is as follows. As causes of death among the prematurely dying are a consequence of social, environmental or lifestyle changes it is hard to mount a case that improvements will continue at such a pace going forward. Therefore when the 10<sup>th</sup> percentile meets the 20<sup>th</sup> percentile it is far more likely that the 20<sup>th</sup> percentile's rate of mortality improvement carries on into the future. If this is so then total convergence will only occur when the last percentiles meet, in this case the 80<sup>th</sup> and 90<sup>th</sup> percentiles which will meet when the lives are expected to die ( i.e. in this case at age 106.05).

It was clear that the convergence had further to go at higher percentiles, but after 1946 the pattern changed from the convergent to the parallel variant. We are now in an era of continuing rises in life expectancy from 67 to 76 years (1.6 years per decade) for males and from 71 years to 80 years (1.6 years per decade) for females. The slower improvement in life expectancy compared to the previous era is due to the fact that childhood mortality had all but been eliminated at the start of this period so little improvement was achieved from reducing this further. The reasons for the changes post 1946 are attributable to factors such as developments in medical science, improvements in welfare, improved health care, and fewer people working in hazardous occupations such as coal mining, and also a decline in smoking. But the improvements have been selective and higher socio-economic groups have benefited more than others. At the oldest ages the probability of survival is indicating a growing variance with no sign of slowing down so that some people will live ever longer.

When the starting age for investigating survival patterns is varied to begin later in life in this case at age 50, the pattern obtained is an accurate representation of the parallel variant. This applies from when records began, although improvements in life expectancy were not continuous through time as the analysis showed. In contrast the pattern from age 80 onwards was a clear example of the divergent variant. It means that the gains in life expectancy at age 80 are being shared ever less equally between survivors to that age. In other words life expectancy is stretched more for some than for others, with obvious implications for areas such as long term care of the elderly. In summary, although life expectancy is increasing, there is no sign of convergence at other ages and there is slower growth in improvement in the first decile of mortality. These findings could mean that the Government target of reducing health inequalities which relies on a convergence in life expectancy may not be achievable.

In using these findings to inform current population projections, it needs to be confirmed whether such changes of trend are predictable going forward or whether we are due for another change of direction (e.g. life expectancy improvement level off or there is a new variant). We have already noted that there has been a lot of concern among actuaries and demographers that governments around the world repeatedly underestimate how long people will live by under estimating improvements in mortality rates at older ages in future years. We therefore tested whether our model would have produced more accurate estimates of survival than the UK Government's own forecasts using information that would have been available at the time. In tests

over two different periods we found that our model gave considerably more accurate results but still underestimated observed survival at the end of both test periods.

We then projected expected population numbers using our model based on data available in 2001 and compared it to the GAD projections from that year. As predicted the percentage of difference between our model and GAD's increased as the age increased with our model predicting more lives. The difference between our model and GAD's comes from two possible sources. The first is that GAD expects a slower improvement in life expectancy than we do. We use a second order polynomial so life expectancy is increasing at a quicker rate than in their approach. The effect of this may be appreciated from Figure 32 which showed the pattern of male life expectancy at aged 50 from 1952 to 2001 including a fitted regression line.

The second potential source of difference is that is that GAD, through caution, expects improvements at the older ages to slow down as life expectancy potentially hits some 'barrier' such as a biological limit to life. We, on the other hand, are assuming that improvements will keep on continuing at a constant rate. However as our tests indicated we can show that GAD's error percentage compared with actual outcome have been wrong before. Since population forecasts are updated annually there is plenty of time for GAD to correct an underestimate of a population forecast that is for 15 years hence. However, for the pensions industry which is used to looking much further ahead and pricing products accordingly more accurate information could be important.

We carried out similar modelling for female lives and the relevant figures are included at Annex A. The main item to note is how much closer our prediction of population size is with GAD's figures. The main reason for this is that the expected increase in female life expectancy has not been increasing as rapidly for females as males in recent years and so the prediction in life expectancy for our model and GAD's is a lot closer. The main discrepancy again is in the 80-89 age group caused again by GAD taking the view that life increases at the older ages will be less rapid than those we assume.

As for the assumption about continuing life improvement our results are based around the long term trend from (1841-2001) so there is arguably no need to assume that life expectancy will have to slow down in the near term. Since countries such as Japan have extra years of expected life compared to the UK then there is a clear argument that these extra years in the UK can be obtained before any natural limits are found. When we used more recent short term cuts of historical data to forecast life expectancy we found if anything that by 2020 it could be even higher than we predict but overshooting is a possibility.

We therefore used ARIMA techniques to ensure that projections did not breach confidence limits and that if they did the year in which this occurs could define the effective time horizon. This showed that our model is likely to be less reliable for males than for females farther into the future for the reason that the male life expectancy is changing quicker. Nevertheless, the results show that the male model will remain within the 95% confidence intervals until 2020. It is consequently suggested that results obtain are likely to be a more realistic forecast than current official projections for the 50+ age group.

The implications of our results are considerable, even in the next decade. By 2020 life expectancy for males at age 50 will be just below that of females who would be expected to live for 35 years. This compares with a male life expectancy of 22.5 years in 1960, a difference of 12.5 years. In addition, a male that reaches age 50 in 2020 would have a 4.5% chance of reaching 100 as compared with a female who would have an 8.8% chance (see Table C1, at Annex C). This compares with 0.54% for males and 1.7% for females in 2001 and 0.014% and 0.098% in 1951. A male reaching age 80 in 2020 would have 6.4% chance of reaching 100, as compared with a female who would have a 12.3% chance. These findings thus underline the speed at which the ageing population will make its presence felt within the next decades. Finally we have already noted that the population age 50+ will be 0.65m higher than expected based on current forecasts.

## References

- Bengtsson, and Keilman, T. N. (2003) Perspectives on Mortality Forecasting Swedish National Social Insurance Board, Social Insurance Studies, Published by: Riksförsäkringsverket (RFV).
- Blake, D. and Mayhew, L. (2006) On the sustainability of the UK state pension system in the light of population ageing and declining fertility. *The Economic Journal*, 116, F286–F305.
- Carey, J.R. (2003) Life Span: A Conceptual Overview, *Population and Development Review*, Vol. 29, Supplement: Life Span: Evolutionary, Ecological, and Demographic Perspectives, pp. 1-18.
- Coale, A. J. (1996) Age Patterns and Time Sequence of Mortality in National Populations with the Highest Expectation of Life at Birth, *Population and Development Review*, Vol. 22, No. 1, pp. 127-135
- Edwards, R.D. and S. Tuljapurkar (2005) Inequality in Life Spans and a New Perspective on Mortality Convergence across Industrialized Countries, *Population and Development Review*, Vol. 31, No. 4 (Dec., 2005), pp. 645-674
- Fries, J.F. 1980 Aging, natural death, and the compression of morbidity. *New England Journal of Medicine*, N Engl J Med. 303(3):130-5.
- Human Mortality Database*. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at [www.mortality.org](http://www.mortality.org) or [www.humanmortality.de](http://www.humanmortality.de).
- Kannisto V, J Lauritsen, A R Thatcher and J Vaupel (???) Reductions in Mortality at advanced ages: Several decades of evidence from 27 countries.
- Kannisto, V. (2000) Measuring the compression of mortality, *Demographic Research*, Vol 3, 6, [www.demographic-research.org/volumes/vol3/6](http://www.demographic-research.org/volumes/vol3/6).



Lynch S., M. and J. S. Brown (2001) Reconsidering Mortality Compression and Deceleration: An Alternative Model of Mortality Rates, *Demography*, Vol. 38, No. 1, pp. 79-95.

Manton, K. G., E Stallard and H. D. Tolley (1991) Limits to Human Life Expectancy: Evidence, Prospects, and Implications. *Population and Development Review*, Vol. 17, No. 4, pp. 603-637.

Manton K.G. and E. Stallard (1996) Strategies to maximize health and functioning and increase life expectancy. In *Resources and population – Natural, institutional, and demographic dimensions of development* eds. Colombo B, Demeny P and Perutz M, chapter 17, 200-223, Clarendon Press, Oxford.

Mayhew (1987) Resources inputs and performance outputs in social security. *Journal of the Operational Research Society*, Vol. 38,10, pp 913-928

Mayhew (2001) The Japanese Longevity Revolution and the Implications for Health Care Finance and Long-term care. Interim report, IR-01-010, IIASA, Laxenburg, Austria.

Mayhew, L. and Smith D.A. (2007) Using queuing theory to analyse the Government's 4-h completion time target in Accident and Emergency departments, *Health Care Management Science*, DOI 10.1007/s10729-007-9033-8

Myers, G. C. and K. G. Manton (1984) Compression of mortality: Myth or reality? *The Gerontologist* 24: 24: 346-353.

Oeppen, J. and Vaupel, J.W (2006) The Linear Rise in Life Expectancy: The Linear Rise in the Number of Our Days, in *Perspectives on Mortality Forecasting*, Ed. Tommy Bengtsson, Swedish Social Insurance Agency, Social Insurance Studies No. 3.

Olshansky, S. J and B. A. Carnes (1977) Ever Since Gompertz, *Demography*, Vol. 34, No. 1, *The Demography of Aging* (Feb., 1997), pp. 1-15.

Robine, J-M., and Y. Saito (2003) Survival beyond Age 100: The Case of Japan., *Population and Development Review*, Vol. 29, Supplement: Life Span: Evolutionary, Ecological, and Demographic Perspectives, pp. 208-228

Rogers, R.G. (1995) Socio-demographic Characteristics of Long-Lived and Healthy Individuals, *Population and Development Review*, Vol. 21, No. 1, pp. 33-58.

Vaupel, J.W. (1998) Demographic Analysis of Aging and Longevity, *The American Economic Review*, Vol. 88, No. 2, *Papers and Proceedings of the Hundred and Tenth Annual Meeting of the American Economic Association*, pp. 242-247.

Vaupel, J (2001) Demographic Insights into Longevity, *Population: An English Selection*, Vol. 13, No. 1, *Biodemographic Perspectives on Human Longevity*, pp. 245-259, Institut National d'Études Démographiques.

Wachter, K.W. (2003) Hazard Curves and Life Span Prospects, Population and Development Review, Vol. 29, Supplement: Life Span: Evolutionary, Ecological, and Demographic Perspectives, pp. 270-291

Wilmoth J.R. Horiuchi S (1999) Rectangularization revisited: variability of age at death within human populations. Demography. 36 (4) , 475-495.

Wilmoth, J.R and J-M Robine (2003) The World Trend in Maximum Life Span, Population and Development Review, Vol. 29, Supplement: Life Span: Evolutionary, Ecological, and Demographic Perspectives, pp. 239-257

Wilmoth, J.R., K. Andreev, D. Jdanov, D.A. Gleijeses assisted by C. Boe, M. Bubenheim, D. Philipov, V. Shkolnikov, P. Vachon (2007 revision) Methods Protocol for the Human Mortality Database, [www.mortality.org/Public/Docs/MethodsProtocol.pdf](http://www.mortality.org/Public/Docs/MethodsProtocol.pdf).

### Annex A: Females

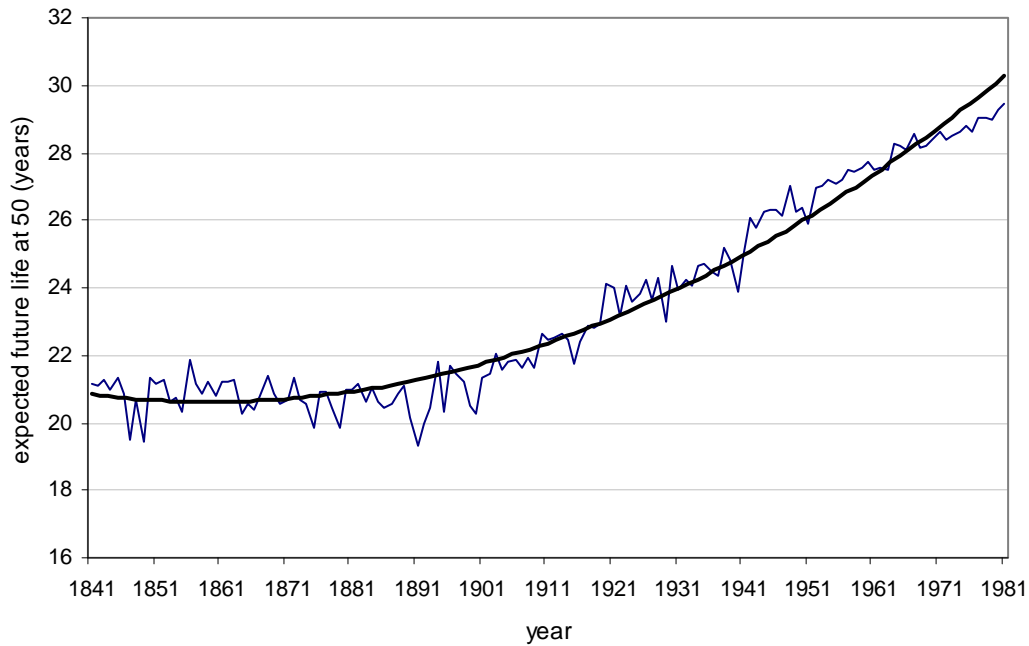


Figure A1: Female life expectancy at age 50, 1841-1981

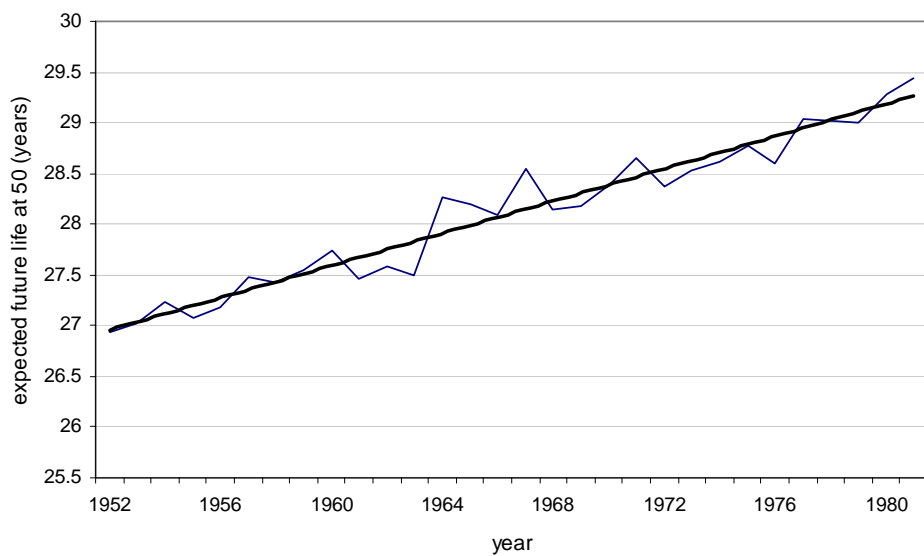


Figure A2: Female life expectancy at 50, 1952-1981

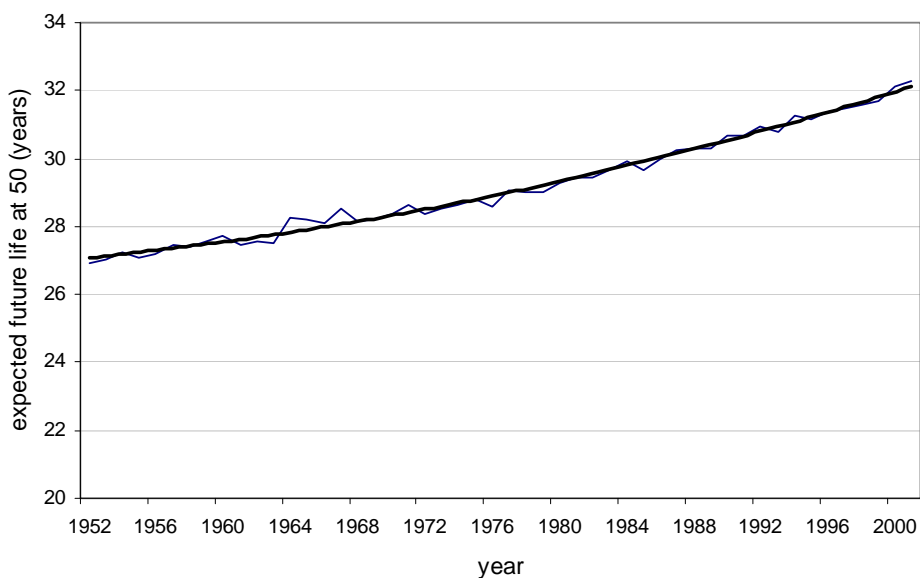


Figure A3: Female life expectancy at 50, 1952-2001

| age   | GAD 2020   | Model      | Diff    | Diff % |
|-------|------------|------------|---------|--------|
| 50-59 | 3,962,913  | 3,963,685  | 771     | 0.02%  |
| 60-69 | 3,203,880  | 3,202,956  | -924    | -0.03% |
| 70-79 | 2,632,919  | 2,645,805  | 12,886  | 0.49%  |
| 80-89 | 1,344,369  | 1,442,493  | 98,124  | 7.30%  |
| total | 11,144,081 | 11,254,939 | 110,858 | 0.99%  |

Table A1: Comparison of female population projected from 2001, model versus GAD

## Annex B: ARIMA

Fitting a second order polynomial to the expected life as noted in the main body of the paper causes problems in two ways. Firstly, when looking back we have the problem that life expectancy would start to increase the further back we go which of course does not happen.

The second problem is that when forecasting forward we find that the projected life expectancy will quickly exceed a value that can be considered to be in the range of likely values. Now the paper with this current projection method is looking to project until 2020 which as we are using data up to 2001 can be considered to be a 19 year projection. However, it could be argued that even by then the projection of expected life may be too large. To try to address this problem we decided to fit an ARIMA model to the data so that we could derive confidence intervals for the expected life at age 50 up to 2020.

The purpose of this paper is to keep the methodology simple and so the ARIMA fitting is not part of the model; rather it is a sensibility check on the values derived. Therefore, there will be little discussion on the fitting of the model.

To ensure that we achieved the best fit for our time series we used the ARIMA model with no preconceptions of the parameter values. Therefore we fitted the normal form of the nonseasonal ARIMA model i.e. using the parameters ARIMA(p,d,q) model, where:

- p** is the number of autoregressive terms,
- d** is the number of non-seasonal differences, and
- q** is the number of lagged forecast errors in the prediction equation.

The main problem we had with our main data set (1841-2001) is that the volatility of life expectancy due to wars and illness meant that we were not able to fit a suitable ARIMA model because of these random ‘jumps’ in the expected life.

We then used our standard reduced data set of (1952-2001). While this allowed us to fit an ARIMA model the fit was not particularly good. We concluded that this was due to the volatility at the start of the period. We therefore concentrated our attention on the years 1963-2001. Using this reduced data set we were able to fit a satisfactory ARIMA model of the form

$$\text{ARIMA}(4, 1, 0)$$

The value of the coefficient (and their significance) for this model (as derived by Minitab) are shown in Table A2:

| Type     | Coefficient | SE of Coefficient | T     | P     |
|----------|-------------|-------------------|-------|-------|
| AR 1     | -0.5341     | 0.1471            | -3.63 | 0.001 |
| AR 2     | -0.0529     | 0.1572            | -0.34 | 0.739 |
| AR 3     | 0.6133      | 0.1558            | 3.94  | 0.000 |
| AR 4     | 0.6234      | 0.1490            | 4.18  | 0.000 |
| Constant | 0.06034     | 0.03075           | 1.96  | 0.058 |

*Table B1: ARIMA results*

Using this model we were thus able to calculate the confidence intervals as presented in the main text.

### **Annex C: Probability of reaching 100 years at ages 50 and 80**

(a) Males

| Current Age | Base year |        |        |        |        |
|-------------|-----------|--------|--------|--------|--------|
|             | 1951      | 1981   | 2001   | 2010   | 2020   |
| 50          | 0.014%    | 0.129% | 0.535% | 1.152% | 4.483% |
| 80          | 0.063%    | 0.411% | 1.103% | 1.967% | 6.370% |

(b) Females

| Current Age | Base year |        |        |        |         |
|-------------|-----------|--------|--------|--------|---------|
|             | 1951      | 1981   | 2001   | 2010   | 2020    |
| 50          | 0.098%    | 0.738% | 1.741% | 4.876% | 8.801%  |
| 80          | 0.264%    | 1.381% | 2.730% | 7.297% | 12.298% |

*Table C1: Probabilities of reaching 100 years by person becoming given age in base year: (a) males; (b) females*

**Faculty of Actuarial Science and Insurance**

**Cass Business School**

---

Copyright 2009 © Faculty of Actuarial Science and Insurance,  
Cass Business School  
106 Bunhill Row,  
London EC1Y 8TZ.

ISBN 978-1-905752-21-8



## FACULTY OF ACTUARIAL SCIENCE AND INSURANCE

### Actuarial Research Papers since 2001

| <b>Report Number</b> | <b>Date</b>    | <b>Publication Title</b>   | <b>Author</b>                                       |
|----------------------|----------------|--|---|
| 135.                 | February 2001. | On the Forecasting of Mortality Reduction Factors.<br>ISBN 1 901615 56 1   | Steven Haberman<br>Arthur E. Renshaw                |
| 136.                 | February 2001. | Multiple State Models, Simulation and Insurer<br>Insolvency. ISBN 1 901615 57 X  | Steve Haberman<br>Zoltan Butt<br>Ben Rickayzen      |
| 137.                 | September 2001 | A Cash-Flow Approach to Pension Funding.<br>ISBN 1 901615 58 8   | M. Zaki Khorasane                                   |
| 138.                 | November 2001  | Addendum to "Analytic and Bootstrap Estimates of<br>Prediction Errors in Claims Reserving". ISBN 1 901615 59 6   | Peter D. England                                    |
| 139.                 | November 2001  | A Bayesian Generalised Linear Model for the Bornhuetter-<br>Ferguson Method of Claims Reserving. ISBN 1 901615 62 6                                    | Richard J. Verrall                                  |
| 140.                 | January 2002   | Lee-Carter Mortality Forecasting, a Parallel GLM<br>Approach, England and Wales Mortality Projections.<br>ISBN 1 901615 63 4                           | Arthur E. Renshaw<br>Steven Haberman.               |
| 141.                 | January 2002   | Valuation of Guaranteed Annuity Conversion Options.<br>ISBN 1 901615 64 2  | Laura Ballotta<br>Steven Haberman                   |
| 142.                 | April 2002     | Application of Frailty-Based Mortality Models to Insurance<br>Data. ISBN 1 901615 65 0   | Zoltan Butt<br>Steven Haberman                      |
| 143.                 | Available 2003 | Optimal Premium Pricing in Motor Insurance: A Discrete<br>Approximation.   | Russell J. Gerrard<br>Celia Glass                   |
| 144.                 | December 2002  | The Neighbourhood Health Economy. A Systematic<br>Approach to the Examination of Health and Social Risks<br>at Neighbourhood Level. ISBN 1 901615 66 9 | Les Mayhew  |
| 145.                 | January 2003   | The Fair Valuation Problem of Guaranteed Annuity<br>Options : The Stochastic Mortality Environment Case.<br>ISBN 1 901615 67 7                         | Laura Ballotta<br>Steven Haberman                   |
| 146.                 | February 2003  | Modelling and Valuation of Guarantees in With-Profit and<br>Unitised With-Profit Life Insurance Contracts.<br>ISBN 1 901615 68 5                       | Steven Haberman<br>Laura Ballotta<br>Nan Want       |
| 147.                 | March 2003.    | Optimal Retention Levels, Given the Joint Survival of<br>Cedent and Reinsurer. ISBN 1 901615 69 3  | Z. G. Ignatov Z.G.,<br>V. Kaishev<br>R.S. Krachunov |
| 148.                 | March 2003.    | Efficient Asset Valuation Methods for Pension Plans.<br>ISBN 1 901615707   | M. Iqbal Owadally                                   |
| 149.                 | March 2003     | Pension Funding and the Actuarial Assumption<br>Concerning Investment Returns. ISBN 1 901615 71 5  | M. Iqbal Owadally                                   |

| <b>Report Number</b> | <b>Date</b>           | <b>Publication Title</b>  | <b>Author</b>   |
|----------------------|-----------------------|---|---|
| 150.                 | Available August 2004 | Finite time Ruin Probabilities for Continuous Claims Severities   | D. Dimitrova<br>Z. Ignatov<br>V. Kaishev                            |
| 151.                 | August 2004           | Application of Stochastic Methods in the Valuation of Social Security Pension Schemes. ISBN 1 901615 72 3   | Subramaniam Iyer  |
| 152.                 | October 2003.         | Guarantees in with-profit and Unitized with profit Life Insurance Contracts; Fair Valuation Problem in Presence of the Default Option <sup>1</sup> . ISBN 1-901615-73-1   | Laura Ballotta<br>Steven Haberman<br>Nan Wang                       |
| 153.                 | December 2003         | Lee-Carter Mortality Forecasting Incorporating Bivariate Time Series. ISBN 1-901615-75-8  | Arthur E. Renshaw<br>Steven Haberman                                |
| 154.                 | March 2004.           | Operational Risk with Bayesian Networks Modelling. ISBN 1-901615-76-6   | Robert G. Cowell<br>Yuen Y, Khuen<br>Richard J. Verrall             |
| 155.                 | March 2004.           | The Income Drawdown Option: Quadratic Loss. ISBN 1 901615 7 4   | Russell Gerrard<br>Steven Haberman<br>Bjorn Hojgarrd<br>Elena Vigna |
| 156.                 | April 2004            | An International Comparison of Long-Term Care Arrangements. An Investigation into the Equity, Efficiency and sustainability of the Long-Term Care Systems in Germany, Japan, Sweden, the United Kingdom and the United States. ISBN 1 901615 78 2 | Martin Karlsson<br>Les Mayhew<br>Robert Plumb<br>Ben D. Rickayzen   |
| 157.                 | June 2004             | Alternative Framework for the Fair Valuation of Participating Life Insurance Contracts. ISBN 1 901615-79-0  | Laura Ballotta  |
| 158.                 | July 2004.            | An Asset Allocation Strategy for a Risk Reserve considering both Risk and Profit. ISBN 1 901615-80-4  | Nan Wang  |
| 159.                 | December 2004         | Upper and Lower Bounds of Present Value Distributions of Life Insurance Contracts with Disability Related Benefits. ISBN 1 901615-83-9  | Jaap Spreeuw  |
| 160.                 | January 2005          | Mortality Reduction Factors Incorporating Cohort Effects. ISBN 1 90161584 7   | Arthur E. Renshaw<br>Steven Haberman                                |
| 161.                 | February 2005         | The Management of De-Cumulation Risks in a Defined Contribution Environment. ISBN 1 901615 85 5.  | Russell J. Gerrard<br>Steven Haberman<br>Elena Vigna                |
| 162.                 | May 2005              | The IASB Insurance Project for Life Insurance Contracts: Impact on Reserving Methods and Solvency Requirements. ISBN 1-901615 86 3.   | Laura Ballotta Giorgia<br>Esposito Steven<br>Haberman               |
| 163.                 | September 2005        | Asymptotic and Numerical Analysis of the Optimal Investment Strategy for an Insurer. ISBN 1-901615-88-X   | Paul Emms<br>Steven Haberman  |
| 164.                 | October 2005.         | Modelling the Joint Distribution of Competing Risks Survival Times using Copula Functions. ISBN 1-901615-89-8   | Vladimir Kaishev<br>Dimitrina S, Dimitrova<br>Steven Haberman       |
| 165.                 | November 2005.        | Excess of Loss Reinsurance Under Joint Survival Optimality. ISBN1-901615-90-1   | Vladimir K. Kaishev<br>Dimitrina S. Dimitrova                       |



| <b>Report Number</b> | <b>Date</b>    | <b>Publication Title</b>  | <b>Author</b>   |
|----------------------|----------------|---|---|
| 166.                 | November 2005. | Lee-Carter Goes Risk-Neutral. An Application to the Italian Annuity Market. ISBN 1-901615-91-X  | Enrico Biffis<br>Michel Denuit                            |
| 167.                 | November 2005  | Lee-Carter Mortality Forecasting: Application to the Italian Population. ISBN 1-901615-93-6   | Steven Haberman<br>Maria Russolillo                       |
| 168.                 | February 2006  | The Probationary Period as a Screening Device: Competitive Markets. ISBN 1-901615-95-2  | Jaap Spreeuw<br>Martin Karlsson                           |
| 169.                 | February 2006  | Types of Dependence and Time-dependent Association between Two Lifetimes in Single Parameter Copula Models. ISBN 1-901615-96-0                              | Jaap Spreeuw  |
| 170.                 | April 2006     | Modelling Stochastic Bivariate Mortality ISBN 1-901615-97-9   | Elisa Luciano<br>Jaap Spreeuw<br>Elena Vigna.             |
| 171.                 | February 2006  | Optimal Strategies for Pricing General Insurance. ISBN 1901615-98-7   | Paul Emms<br>Steve Haberman<br>Irene Savoulli             |
| 172.                 | February 2006  | Dynamic Pricing of General Insurance in a Competitive Market. ISBN1-901615-99-5   | Paul Emms   |
| 173.                 | February 2006  | Pricing General Insurance with Constraints. ISBN 1-905752-00-8  | Paul Emms   |
| 174.                 | May 2006       | Investigating the Market Potential for Customised Long Term Care Insurance Products. ISBN 1-905752-01-6   | Martin Karlsson<br>Les Mayhew<br>Ben Rickayzen            |
| 175.                 | December 2006  | Pricing and Capital Requirements for With Profit Contracts: Modelling Considerations. ISBN 1-905752-04-0  | Laura Ballotta  |
| 176.                 | December 2006  | Modelling the Fair Value of Annuities Contracts: The Impact of Interest Rate Risk and Mortality Risk. ISBN 1-905752-05-9                                    | Laura Ballotta<br>Giorgia Esposito<br>Steven Haberman     |
| 177.                 | December 2006  | Using Queuing Theory to Analyse Completion Times in Accident and Emergency Departments in the Light of the Government 4-hour Target. ISBN 978-1-905752-06-5 | Les Mayhew<br>David Smith                                 |
| 178.                 | April 2007     | In Sickness and in Health? Dynamics of Health and Cohabitation in the United Kingdom. ISBN 978-1-905752-07-2  | Martin Karlsson<br>Les Mayhew<br>Ben Rickayzen            |
| 179.                 | April 2007     | GeD Spline Estimation of Multivariate Archimedean Copulas. ISBN 978-1-905752-08-9   | Dimitrina Dimitrova<br>Vladimir Kaishev<br>Spiridon Penev |
| 180.                 | May 2007       | An Analysis of Disability-linked Annuities. ISBN 978-1-905752-09-6  | Ben Rickayzen   |
| 181.                 | May 2007       | On Simulation-based Approaches to Risk Measurement in Mortality with Specific Reference to Poisson lee-Carter Modelling. ISBN 978-1-905752-10-2             | Arthur Renshaw<br>Steven Haberman                         |

| <b>Report Number</b> | <b>Date</b>  | <b>Publication Title</b>   | <b>Author</b>                                 |
|----------------------|--------------|--|---|
| 182.                 | July 2007    | High Dimensional Modelling and Simulation with Asymmetric Normal Mixtures. ISBN 978-1-905752-11-9  | Andreas Tsanakas<br>Andrew Smith              |
| 183.                 | August 2007  | Intertemporal Dynamic Asset Allocation for Defined Contribution Pension Schemes. ISBN 978-1-905752-12-6  | David Blake<br>Douglas Wright<br>Yumeng Zhang |
| 184.                 | October 2007 | To split or not to split: Capital allocation with convex risk measures. ISBN 978-1-905752-13-3   | Andreas Tsanakas                              |
| 185.                 | April 2008   | On Some Mixture Distribution and Their Extreme Behaviour. ISBN 978-1-905752-14-0   | Vladimir Kaishev<br>Jae Hoon Jho              |
| 186.                 | October 2008 | Optimal Funding and Investment Strategies in Defined Contribution Pension Plans under Epstein-Zin Utility. ISBN 978-1-905752-15-7                          | David Blake<br>Douglas Wright<br>Yumeng Zhang |
| 187.                 | May 2008     | Mortality Risk and the Valuation of Annuities with Guaranteed Minimum Death Benefit Options: Application to the Italian Population. ISBN 978-1-905752-16-4 | Steven Haberman<br>Gabriella Piscopo          |
| 188.                 | January 2009 | The Market Potential for Privately Financed Long Term Care Products in the UK. ISBN 978-1-905752-19-5  | Leslie Mayhew                                 |
| 189.                 | June 2009    | Whither Human Survival and Longevity or the Shape of things to Come. ISBN 978-1-905752-21-8  | Leslie Mayhew<br>David Smith                  |

### **Statistical Research Papers**

|     |                 |  |                             |
|-----|-----------------|--|-----------------------------|
| 1.  | December 1995.  | Some Results on the Derivatives of Matrix Functions. ISBN 1 874 770 83 2   | P. Sebastiani               |
| 2.  | March 1996      | Coherent Criteria for Optimal Experimental Design. ISBN 1 874 770 86 7   | A.P. Dawid<br>P. Sebastiani |
| 3.  | March 1996      | Maximum Entropy Sampling and Optimal Bayesian Experimental Design. ISBN 1 874 770 87 5                                   | P. Sebastiani<br>H.P. Wynn  |
| 4.  | May 1996        | A Note on D-optimal Designs for a Logistic Regression Model. ISBN 1 874 770 92 1   | P. Sebastiani<br>R. Settini |
| 5.  | August 1996     | First-order Optimal Designs for Non Linear Models. ISBN 1 874 770 95 6   | P. Sebastiani<br>R. Settini |
| 6.  | September 1996  | A Business Process Approach to Maintenance: Measurement, Decision and Control. ISBN 1 874 770 96 4                       | Martin J. Newby             |
| 7.  | September 1996. | Moments and Generating Functions for the Absorption Distribution and its Negative Binomial Analogue. ISBN 1 874 770 97 2 | Martin J. Newby             |
| 8.  | November 1996.  | Mixture Reduction via Predictive Scores. ISBN 1 874 770 98 0   | Robert G. Cowell.           |
| 9.  | March 1997.     | Robust Parameter Learning in Bayesian Networks with Missing Data. ISBN 1 901615 00 6                                     | P. Sebastiani<br>M. Ramoni  |
| 10. | March 1997.     | Guidelines for Corrective Replacement Based on Low Stochastic Structure Assumptions. ISBN 1 901615 01 4.                 | M.J. Newby<br>F.P.A. Coolen |
| 11. | March 1997      | Approximations for the Absorption Distribution and its Negative Binomial Analogue. ISBN 1 901615 02 2                    | Martin J. Newby             |

| <b>Report Number</b> | <b>Date</b>    | <b>Publication Title</b>  | <b>Author</b>   |
|----------------------|----------------|---|---|
| 12.                  | June 1997      | The Use of Exogenous Knowledge to Learn Bayesian Networks from Incomplete Databases. ISBN 1 901615 10 3                                       | M. Ramoni<br>P. Sebastiani  |
| 13.                  | June 1997      | Learning Bayesian Networks from Incomplete Databases. ISBN 1 901615 11 1  | M. Ramoni<br>P. Sebastiani  |
| 14.                  | June 1997      | Risk Based Optimal Designs. ISBN 1 901615 13 8  | P. Sebastiani   |
| 15.                  | June 1997.     | Sampling without Replacement in Junction Trees. ISBN 1 901615 14 6  | H.P. Wynn<br>Robert G. Cowell   |
| 16.                  | July 1997      | Optimal Overhaul Intervals with Imperfect Inspection and Repair. ISBN 1 901615 15 4   | Richard A. Dagg<br>Martin J. Newby  |
| 17.                  | October 1997   | Bayesian Experimental Design and Shannon Information. ISBN 1 901615 17 0  | P. Sebastiani.<br>H.P. Wynn   |
| 18.                  | November 1997. | A Characterisation of Phase Type Distributions. ISBN 1 901615 18 9  | Linda C. Wolstenholme   |
| 19.                  | December 1997  | A Comparison of Models for Probability of Detection (POD) Curves. ISBN 1 901615 21 9  | Wolstenholme L.C  |
| 20.                  | February 1999. | Parameter Learning from Incomplete Data Using Maximum Entropy I: Principles. ISBN 1 901615 37 5   | Robert G. Cowell  |
| 21.                  | November 1999  | Parameter Learning from Incomplete Data Using Maximum Entropy II: Application to Bayesian Networks. ISBN 1 901615 40 5                        | Robert G. Cowell  |
| 22.                  | March 2001     | FINEX : Forensic Identification by Network Expert Systems. ISBN 1 901615 60X  | Robert G. Cowell  |
| 23.                  | March 2001.    | Wren Learning Bayesian Networks from Data, using Conditional Independence Tests is Equivalent to a Scoring Metric ISBN 1 901615 61 8          | Robert G Cowell   |
| 24.                  | August 2004    | Automatic, Computer Aided Geometric Design of Free-Knot, Regression Splines. ISBN 1-901615-81-2   | Vladimir K Kaishev,<br>Dimitrina S. Dimitrova,<br>Steven Haberman<br>Richard J. Verrall |
| 25.                  | December 2004  | Identification and Separation of DNA Mixtures Using Peak Area Information. ISBN 1-901615-82-0   | R.G. Cowell<br>S.L. Lauritzen<br>J Mortera,   |
| 26.                  | November 2005. | The Quest for a Donor : Probability Based Methods Offer Help. ISBN 1-90161592-8   | P.F. Mostad<br>T. Egeland.,<br>R.G. Cowell<br>V. Bosnes<br>Ø. Braaten                   |
| 27.                  | February 2006  | Identification and Separation of DNA Mixtures Using Peak Area Information. (Updated Version of Research Report Number 25). ISBN 1-901615-94-4 | R.G. Cowell<br>S.L. Lauritzen<br>J Mortera,   |

| <b>Report Number</b> | <b>Date</b>   | <b>Publication Title</b>   | <b>Author</b>  |
|----------------------|---------------|--|--|
| 28.                  | October 2006  | Geometrically Designed, Variable Knot Regression Splines : Asymptotics and Inference. ISBN 1-905752-02-4                 | Vladimir K Kaishev<br>Dimitrina S.Dimitrova<br>Steven Haberman<br>Richard J. Verrall |
| 29.                  | October 2006  | Geometrically Designed, Variable Knot Regression Splines : Variation Diminishing Optimality of Knots. ISBN 1-905752-03-2 | Vladimir K Kaishev<br>Dimitrina S.Dimitrova<br>Steven Haberman<br>Richard J. Verrall |
| 30.                  | November 2008 | Scheduling Reentrant Jobs on Parallel Machines with a Remote Server. ISBN 978-1-905752-18-8                              | Konstantin Chakhlevitch<br>Celia Glass   |

***Papers can be downloaded from***

***<http://www.cass.city.ac.uk/arc/actuarialreports.html>***