UCC

University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

A Mortality Experience Study for Evaluating the Longevity Risk of An Annuity Product

December 1, 2020

Declaration

I, the undersigned hereby declare that the project report entitled "**A Mortality Experience Study for Evaluating the Longevity Risk of An Annuity Product**" written and submitted by me to University College Cork in partial fulfilment of requirements for the award of the degree of Master of Actuarial Science under the guidance of Ms.Linda Daly, is my original work and interpretations drawn therein are based on material collected by myself.

Dedication

I would like to dedicate this report to my friends, family, teachers and everyone who has helped me over the past year. Particularly I would like to acknowledge my parents for their unconditional love and support.

Abstract

This paper aims to assess the longevity risk associated with a specific term annuity product. To do this we compare six different models explaining mortality improvements amongst Irish males and find that models that allow for parameter uncertainty and contain period and cohort effects perform significantly better. We also highlight the advantages of taking a stochastic approach to modelling mortality when estimating longevity risk. Some models show volatility in their parameter estimates but we find that one model, an extension to the Cairns, Blake and Dowd (2006)[1] model, exhibits parameter stability across different time periods in our data set. This model shows us that there have been approximately linear improvements in mortality rates over time and that these improvements have been more significant at lower ages than at higher ages. This model also shows that younger cohorts tend to carry a significantly higher level of longevity risk when estimating future mortality rates.

Contents

List of Tables

List of Figures

1 Introduction

Significant improvements in mortality have been observed in many countries over the last century. Increasing prosperity, better nutrition, improvements in public health, and technological advances have fuelled increases in life expectancy. According to data from The World Bank [17], life expectancy at birth in Ireland has risen from around 54 to 82 years since the early 1900s. Plotted below in Figure 1 we can see how life expectancy has been changing in Ireland over the past 60 years.

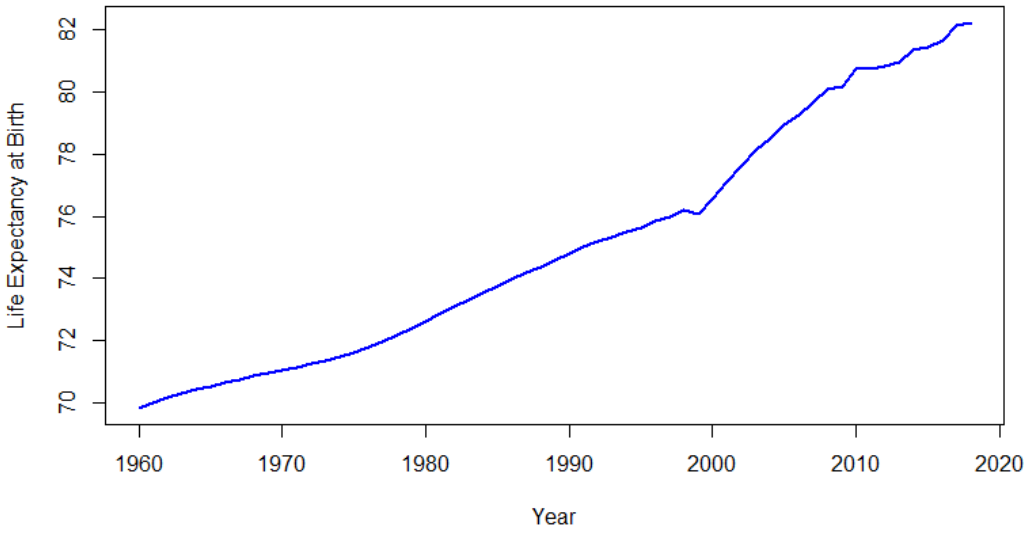


Figure 1: Life Expectancy at birth in Ireland from 1961 to 2017

In the early half of the last century increases in life expectancy were driven mainly due to improvements in sanitation, housing and education, causing a steady decline in infantile and early-life mortality, which was mainly due to infections and complications at birth. However, since the 1960's this trend has shifted and due to positive developments in the prevention and treatment of chronic diseases that affect people at older ages, such as heart disease, relatively higher improvements in later life mortality have been observed. Our study mainly focuses on mortality improvements amongst individuals at mid to later ages and the uncertainty surrounding the future of these mortality rates.

Past mortality projections produced by governmental agencies, actuaries, academics, etc. have consistently underestimated mortality improvements. For example, forecasts based off US projections made in 1928 predicted that biological and practical impediments would cap maximum life expectancy at birth to 65 for both sexes. Non-Maori New Zealand women were already living longer than this in 1921, and by 1996 Japanese women had outlived the revised projection of 85 made in 1990 and 2001.

Underestimating future life expectancy can have serious financial consequences for individuals, governments, insurance companies, employers and any providers of pensions or life annuity products. Most pension funds do not fully account for future improvements in mortality and life expectancy and as a result of this retirees run the risk of outliving their resources and can be forced to reduce their standard of living at older ages. For providers of pension plans and/or life annuities, unexpected improvements in mortality that are not accounted for can have a serious strain on retirement finances. This is partly why here, in Ireland, the state pension retirement age is set to increase to 67 as of next year and then to 68 in 2028.

Unfortunately, even with a considerable effort focused towards improving the accuracy of mortality projections, it is clear from past trends that there will always be an underlying

uncertainty when it comes to future mortality rates and life expectancy. This is why it is necessary to take a stochastic rather than a deterministic approach to modelling mortality as this allows us to attach probabilities to different projections and create a distribution of possible outcomes rather than a single projection which in turn enables us to assess the uncertainty associated with these projections.

The aim of this study is to produce reliable forecasts for mortality rates among Irish males and, most importantly, gauge the level of uncertainty associated with these forecasts. With this, we can estimate the "longevity risk" faced by certain life annuity products. Longevity risk in this context refers to the risk that future life expectancy is greater than anticipated. Estimating this risk is extremely important because if it is understood then it can be taken into account when planning pension funds or providing life annuity products meaning that any potential improvements in life expectancy would have a negligible effect on retirement finances. Although longevity risk is a very real concern for insurers, it is difficult to quantify. Several techniques exist, as outlined by Crawford et al (2008) [7], but many believe that a measure has yet to be developed that is accurate enough to enable insurers to evaluate the capital requirements for managing this risk with much confidence. This study only aims to provide a rough measure that gives a clearer picture of the longevity risk associated with a specific annuity product. To do this we need to get a feel for how mortality rates might look over the next few decades.

A number of models have been developed in previous studies that attempt to shed light on the uncertainty surrounding future life expectancy by analyzing the improvement in general mortality over time and predict how the trend may evolve. A ground-breaking model was introduced by Lee & Carter (1992)[12] that provided a unique new method of forecasting mortality. The distinct new feature of this model was the use of a stochastic process to capture the uncertainty in future mortality improvement.

The prevailing narrative amongst researchers since the introduction of this model, and the availability of new data, is that the underlying process driving mortality improvements has an inherently random element in its nature that is best described by a stochastic process. As mentioned, several models have been developed over the last twenty years, some of which are either direct extensions of the original Lee-Carter model such as those introduced by [2], [16], others are of separate families of models, such as the Cairns-Blake-Dowd (CBD) ([1]) model and it's extensions that still utilize the stochastic approach to modelling mortality but differ in other aspects. Some of these models include what is known as a 'cohort effect' which allows for the effect of an individual's year of birth on their rate of mortality improvement over time. Thorough quantitative comparisons have been performed on these models, including work done by [4], which show that allowing for the cohort effect yields significant improvement in forecasting capabilities.

In this paper we consider a select few of these models, as well as others, and perform our own comparisons to try and determine which model might provide us with the most reliable forecasts for our Irish males data and ultimately use these to estimate the longevity risk associated with a specific life annuity product, as mentioned earlier.

1.1 Notation

Here we briefly discuss any notation we used to ensure that it is clear and consistent throughout the paper.

Calendar year t is defined as running from time t to $t+1$. We define $m_{t,x}$ as the death rate for someone aged x exactly during the calendar year t. Specifically:

$$m_{t,x} = \frac{d_{t,x}}{E^c_{t,x}}$$

Where $d_{t,x}$ is defined as the observed number of deaths at age x exactly in a calendar year t and $E^c_{t,x}$ is the corresponding central exposure at age x exactly in a calendar year t.

A second measure of mortality is the mortality rate $q_{t,x}$. This is the probability that an individual aged exactly x at exact time t will die between times t and $t+1$ and is defined as:

$$m_{t,x} = \frac{d_{t,x}}{E^0_{t,x}}$$

Where $E^0_{t,x}$ is the initial exposure for age x at time t. A commonly used assumption that relates central and initial exposures, sometimes called the "Actuarial assumption", is that

$$E^0_{t,x} \approx E^c_{t,x} + \frac{1}{2}d_{t,x}$$

A third measure is the force of mortality $\mu_{t,x}$. This is the instantaneous death rate at exact time t for an individual aged exactly x at exact time t.

For a given x and t, the values of $m_{t,x}$ and $q_{t,x}$ tend to very similar. With some commonly adopted assumptions, we can formalise a precise relationship between these two measures.

- Assumption 1: The force of mortality $\mu_{t,x}$ remains constant over the calendar year t for each integer age.
- Assumption 2: We assume that our population is stationary, i.e. that the size of the population remains constant across all ages.

These assumptions imply that

(a) $m_{t,x} = \mu_{t,x}$

(b) $q_{t,x} = 1 - exp[-\mu_{t,x}] = 1 - exp[-m_{t,x}]$

These assumptions tend not to hold exactly but the resulting relationship is widely accepted and is generally felt to give a good approximation. Relationship (b) is also useful for the third model being considered in this paper which is formulated in terms of $q_{t,x}$.

We also define the longevity risk of an annuity as:

$$LR = \frac{\text{99.5\% quantile of the distribution of price}}{\text{Mean price}}$$

Where price refers to the expected present value of benefits per policy.

2 The Data

Before fitting any models it is always important to let the observed data speak for itself. We need to assess what trends or characteristics exist in our data to assure that any assumptions underlying a proposed model are consistent with our data.

We will now discuss the general characteristics of the Ireland males data. Since the primary goal of this study is to estimate how much longevity risk the SIPP annuity product is exposed to, we use data at higher ages only (ages 50 to 89) when fitting and comparing our models. This data consists of the observed number of deaths at each integer age and their corresponding central exposures for the years 1961 to 2016.

2.1 Death Rates

Death rates for Irish males between 1961 and 2016 are plotted below in Figure 2. This plot shows death rates across all ages and years in our study and gives us a general impression of how death rates behave over time. Here we can see the rate at which death rates increase with respect to age tends to be higher at older ages and that general death rates have been decreasing over time. We can also see that the surface of this plot is more rough and jagged at older ages indicating that changes in death rates are more erratic at these ages. We need to take a look at our data from different perspectives to verify these observations.

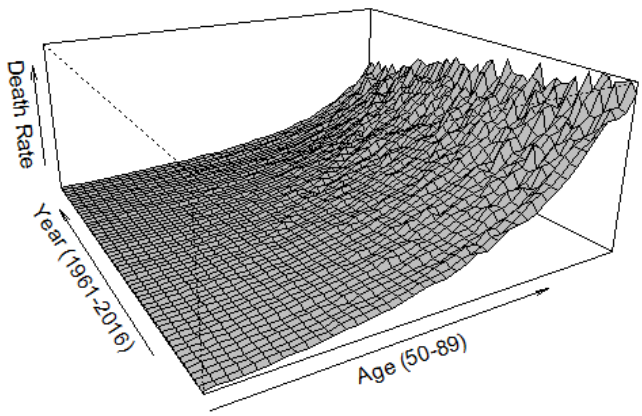
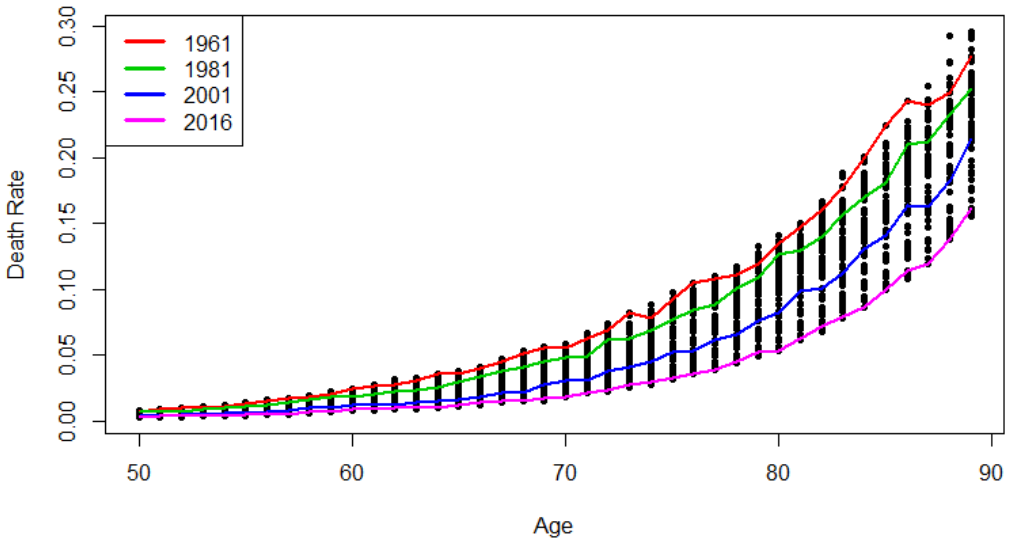
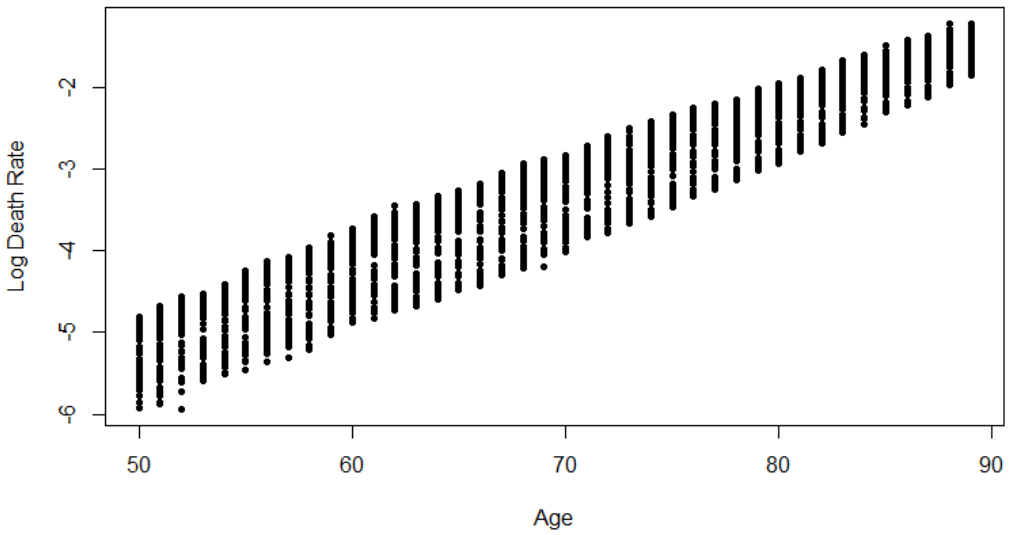


Figure 2: Crude mortality rates for Ireland between 1961 and 2016

First, we analyze the relationship between death rates and age in our data. Figure 3 (a) shows the crude death rates plotted against age and curves representing death rates from the years 1961, 1981, 2001 and 2016. The purpose of adding these curves is to illustrate that the relationship between death rates and age seems to be consistent across time. This plot shows us that death rates seem to increase exponentially with age for the range of ages under study. The linearity of log death rates in Figure 3 (b) supports this.



(a) Crude death rates with age



(b) Log Death Rates with age

Figure 3: Plots displaying the exponential relationship between death rates and age

Next we would like to analyze the relationship between time and general mortality.

Looking at Figure 4 we see that general mortality has clearly been improving over time and that the improvement from year to year is slightly random as discussed earlier. What's also interesting is that the rate of change tended to be more erratic in the past and has gradually smoothed out over time. This could be a characteristic of the trend underlying general mortality improvement or, most likely, it could be due to the due to the lower quality of past data compared to more recent data.

Another important aspect of the relationship between general mortality and time is that mortality improvement is not constant across all age groups. Figure 5 below shows average death rates for each 10 year age group relative to their respective average death rates in the 1960's. Here we see that the greatest proportional improvement was observed for the

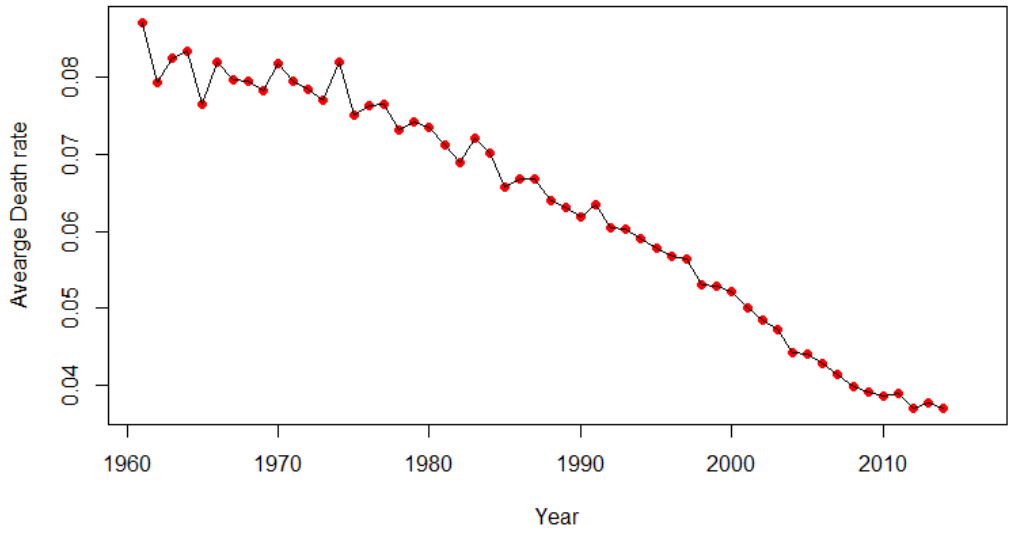


Figure 4: Average mortality rates across ages 50-89 for years 1961-2016

60-69 year old age group whose most recent observed death rate is roughly a third of what it was in the 1960s while the 80-89 year old age group experienced the lowest proportional improvement, their death rate has halved since the 1960s.

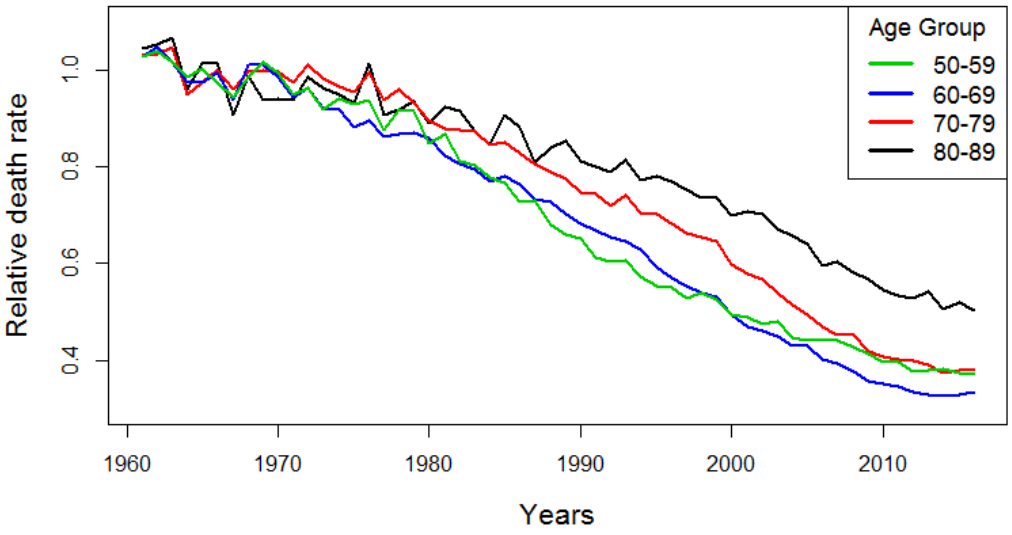


Figure 5: Relative death rates across different age groups

This is a pattern that can be observed in other populations as well and we may look for models that can capture this variability between age groups in their forecasts.

Another feature of our data we wish to analyze is the relationship between $logit(m_{t,x})$ and age. This is relevant to some of the models we will be fitting that make assumptions about the pattern of this relationship.

Looking at Figure 6 we see that the relationship appears to be linear for our data set but it should be noted that curvature has been observed in similar plots for different populations [4].

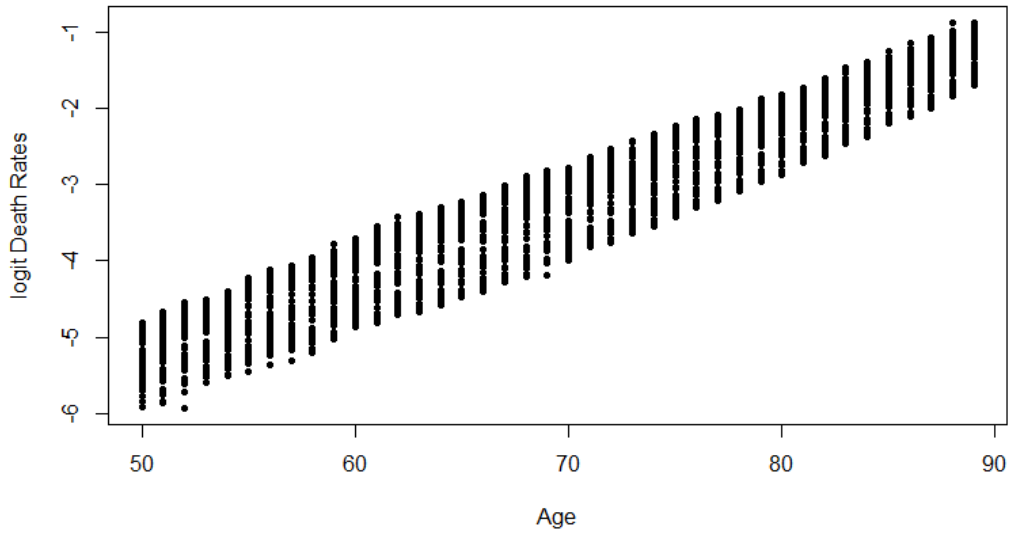


Figure 6: $logit(m_{t,x})$ vs. Age for Irish males data

As mentioned, some models which we will introduce use the *logit* link function to model death rates. By doing this we are applying binomial regression which, in the context of our study, assumes that the number of deaths $D_{t,x}$ follows a binomial distribution with parameter $E^c_{t,x}\hat{m}_{t,x}$:

$$D_{t,x} \sim Bin(E^c_{t,x}\hat{m}_{t,x})$$

This is in contrast to other models we will be analyzing which make use of Poisson regression by applying the *log* link function to model death rates and assume that:

$$D_{t,x} \sim Poi(E^c_{t,x}\hat{m}_{t,x})$$

To verify these assumptions we take a look at our observed deaths $d_{t,x}$ in our data:

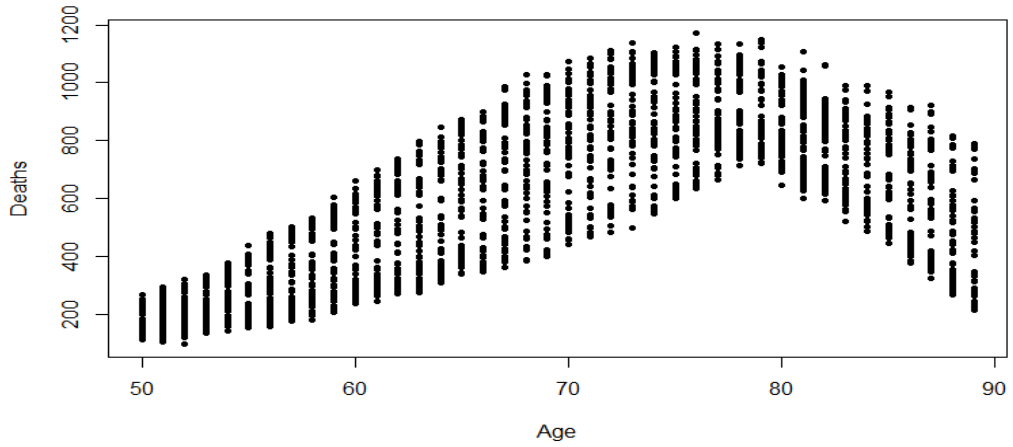


Figure 7: Number of deaths at ages 50-89 for years 1961-2016

Poisson regression is widely applied to death count data as it is useful when:

- describing the occurrence of a rare event in a population
- the number of events can only take on integer values
- events occur independently
- two events cannot occur at exactly the same instant

All of these assumptions are fairly valid for death count data, hence the common usage of Poisson regression, but another assumption made when using Poisson regression is that the mean and the variability of the number of events are equal. This could pose an issue as death count data does tend to be quite volatile especially in older age groups. This is known as "over-dispersion"; when the variability in a data set is much greater than the mean. When using binomial regression over-dispersion is not an issue as an extra free parameter is included that allows for differences between the mean and variability in the number of events.

All this being said, either approach is generally appropriate for death count data and similar results are obtained when using either.

Finally, we attempt to visualise the existence of the cohort effect in our data. Below is a contour plot of mortality improvement at each age for each year in our data.

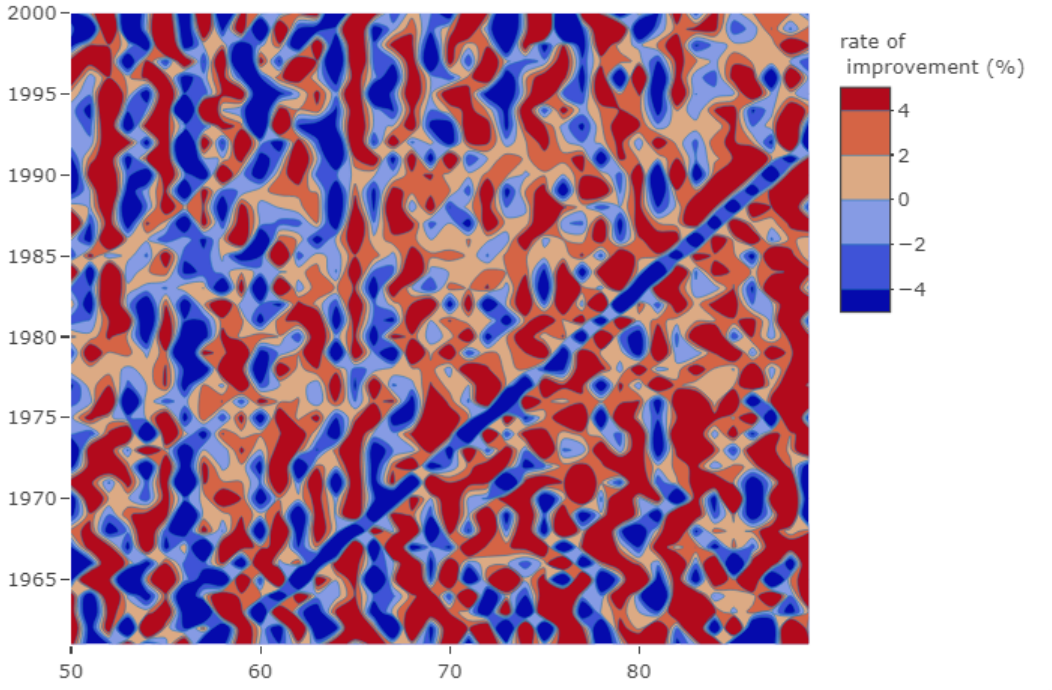


Figure 8: Improvement rates in mortality for Irish males by calendar year and age relative to mortality rates at the same age in the previous year. Red cells imply that mortality is improving and blue cells imply that mortality is deteriorating

Here we can see a clear blue diagonal which corresponds to the 1903 cohort. This is strong evidence for the existence of a cohort effect in our data set. What this figure is showing us is that, despite the improvement of general mortality over time, the 1903 cohort consistently experienced higher mortality rates than previous years.

3 Modelling $m_{t,x}$

In this section, we introduce the models that we propose to fit to our empirical mortality data and explain the assumptions underlying each one and how each model is fitted. We then seek to compare each model based off of chosen metrics that measure the goodness of fit as well as forecast accuracy and plausibility. Below is a table formally outlining each model and how they will be referenced throughout the paper.

Model	Formula
A	$m_x = e^{\alpha+\beta \cdot x} \quad (\alpha, \beta > 0)$
B	$m_x = e^{\alpha+\beta \cdot x} \quad \& \quad D_{t,x} \sim Poi(E^c_{t,x}\hat{m}_{t,x})$
C	$logit(m_{t,x}) = \kappa^{(1)}_t + \kappa^{(2)}_t(x - \bar{x})$
D	$logit(m_{t,x}) = \kappa^{(1)}_t + \kappa^{(2)}_t(x - \bar{x}) + \gamma^{(3)}_{t-x}$
E	$logit(m_{t,x}) = \kappa^{(1)}_t + \kappa^{(2)}_t(x - \bar{x}) + \kappa^{(3)}_t((x - \bar{x})^2 - \hat{\sigma}^2_x) + \gamma^{(4)}_{t-x}$
F	$log(m_{t,x}) = \beta^{(1)}_x + \beta^{(2)}_x\kappa^{(2)}_t + \beta^{(3)}_x\gamma^{(3)}_{t-x}$

3.1 Model A

Firstly we fit the Gompertz mortality model. This law, first introduced by Benjamin Gompertz in 1825 [10], is a very well known and simple model depending on only two parameters, α & β. In the context of modelling mortality, α and β can be interpreted as the general level of mortality in the population and the effect of age on death rates. The Gompertz parametric description of mortality, often referred to as a law, is that mortality rates increase exponentially with age (above some threshold age, typically assumed to be around 40 to 45). Looking back to our analysis in section 2, this assumption certainly agrees with the relationship between age and death rates observed in our data. This simple model has performed well at describing mortality during mid to later life in a variety of different populations and across different time periods and has not only been utilized in the field of actuarial science but also fields such as biology and medicine.

Although the Gompertz model is widely used, there is no clear consensus on the best way to estimate the parameters α and β. For Model A, we take the logarithm of death rates which transforms the right-hand side of our equation into a sum of two components which can be easily estimated using linear regression:

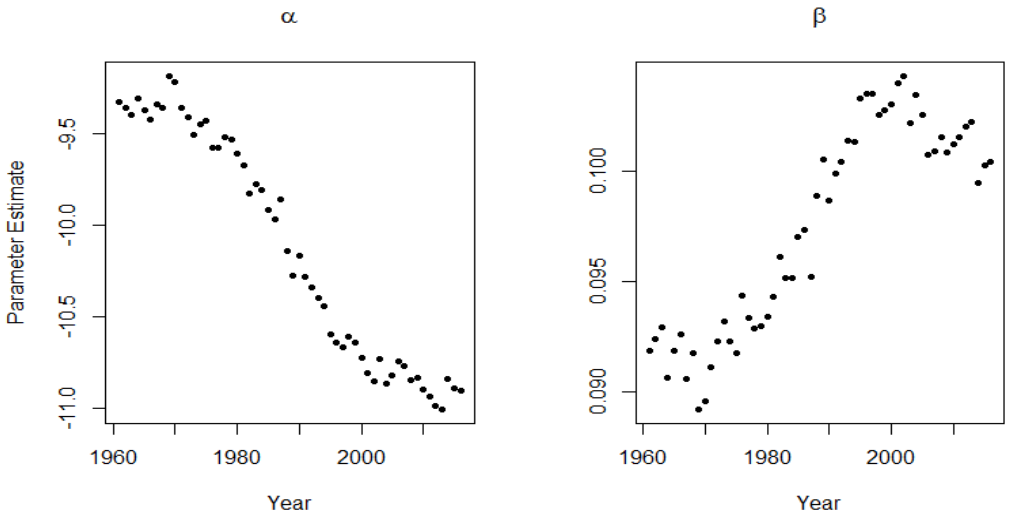
$$log(m_x) = \alpha + \beta x$$

This model, although popular, is known for giving unreliable results at older ages (80+) where the uncertainty surrounding our data increases. This highlights one of the main disadvantages of the deterministic approach to mortality modelling. We mentioned earlier that many researchers believe that a stochastic model is more appropriate than a stochastic one when modelling mortality. The main difference between the two, put simply, is that a deterministic model takes in numbers as inputs (which are assumed to be well known and understood) and gives numbers as outputs and any uncertainty in the output is external to the model, whereas a stochastic model includes a random component that uses a distribution as one of the inputs, and results in a distribution for the output. These distributions reflect the uncertainty in the data which we input to the model and the uncertainty in the output. So, at older ages, where exposure is very low and death rates become increasingly random this deterministic approach tends to break down and fails to capture the real trends in our data.

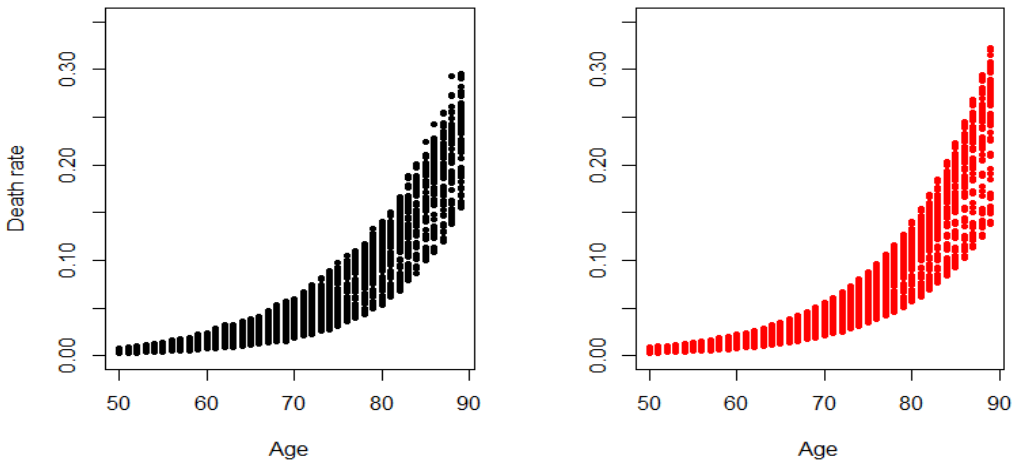
Another issue with model A is that it does not include time as a predictor variable. Without taking account for a period effect this model cannot properly capture the variability in death rates for different age groups over time. Model A also does not take a cohort effect into

consideration which, we will show, does provide significant improvements when modelling death rates.

The fact that Model A does not take a period effect into account raises issues when studying a data set such as ours that contains multiple years worth of data. There are a number of different methods of tackling this issue, some of these include modifying the model by adding time in as a predictor variable but these proved to not be as successful as the method of fitting the model to each year independently (56 separate fits in our case) and collecting all the fitted values from each one. The reason for this is that to add time in as a predictor variable, you must assume some form of parametric relationship between time and mortality but as we've seen the true nature of this relationship carries too much uncertainty for any sort of assumption to be appropriate. By doing this we obtain 56 estimates for each of the parameters, one for each year. Plotted below are the parameter estimates obtained from fitting model A and also the fitted values for m_x from model A in red alongside the observed death rates from our data.



(a) Parameter estimates from model A for each year 1961-2016



(b) Observed (left) and fitted (right) values of m_x from model A across years 1961-2016

Figure 9: Parameter estimates and fitted death rates from model A

Looking at the fitted values in Figure 9 (b) first we see that Model A fits the observed data quite well up until around age 80 where the quality of fit seems to diminish. This was expected since, as discussed earlier, the randomness in the data for these ages is not fully taken into account by the model. Looking at the parameter estimates plotted in Figure 9 (a)

we see that the general mortality level, α, has been decreasing over time and, interestingly, this trend seems to be negatively correlated with the age effect, β. The effect of age on death rates seems to have been increasing over time meaning that, as general mortality decreases, the severity of the relationship between aging and rising mortality rates has been increasing.

3.2 Model B

For this next model, we continue with the assumption that death rates $m_{t,x}$ follow Gompertz' law. We discussed earlier that there is no clear consensus on the best way to fit this model and for model A we used a linear model to model log-death rates and applied the typical method of least squares to estimate the model parameters. For model B we take a slightly different approach and assume that the random variable D_x, the number of deaths at age x, follows a Poisson distribution with parameter $E_x\hat{m}_x$. Including an "offset" term, which appears as the log of exposure, allows us to model log death rates instead of just death counts at each age.

$$log(E(D_x|x)) = \beta_0 + \beta_1 x$$

$$log(\frac{E(D_x|x)}{E_x}) = \beta_0' + \beta_1' x$$

$$log(E(D_x|x)) = log(E_x) + \beta_0' + \beta_1' x$$

or

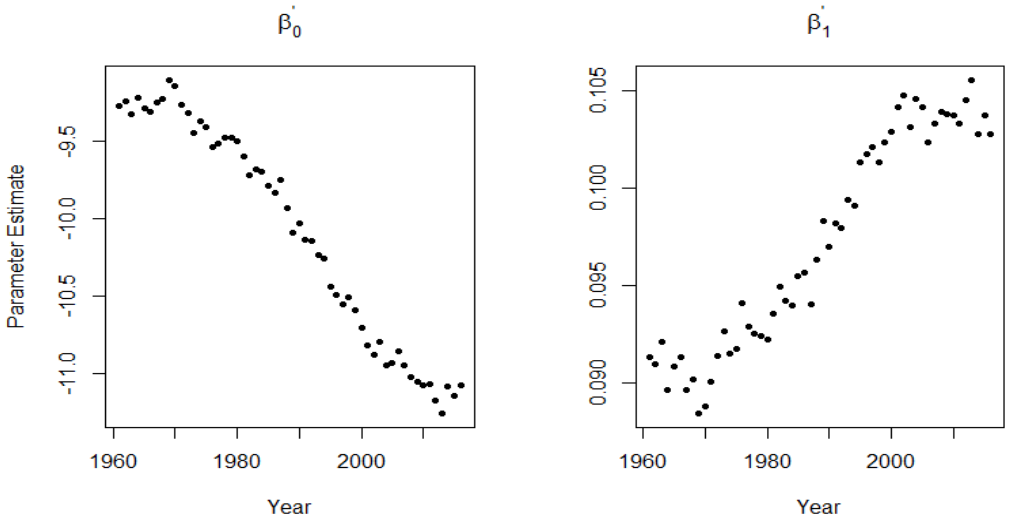
$$\log(\hat{m}_x) = \beta_0' + \beta_1' x$$

As we can see this model is identical to model A but the key difference here is that we have imposed the additional assumption that D_x follows a Poisson distribution and so instead of using least squares the model parameters are estimated by maximizing the log-likelihood that the fitted values, $\hat{D}_x$, follow a Poisson distribution.

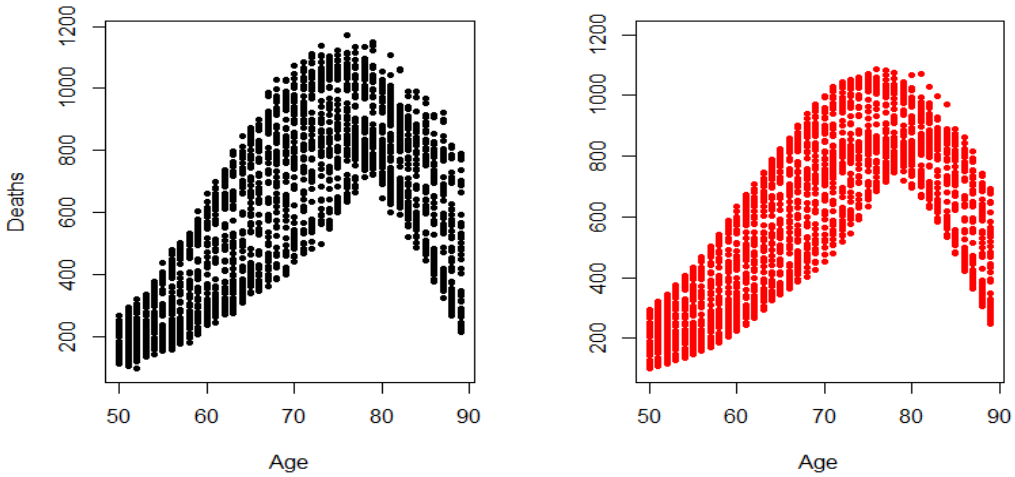
Much like model A, we could have modified this model adding time as a predictor variable but instead, we chose to apply the model to each year individually and collect the fitted values from each model to obtain a full fit for our data set.

Plotted below are the parameter estimates from model B, the fitted values for the number of deaths compared to the observed numbers as well as fitted values for $\hat{m}_x$ alongside the observed data.

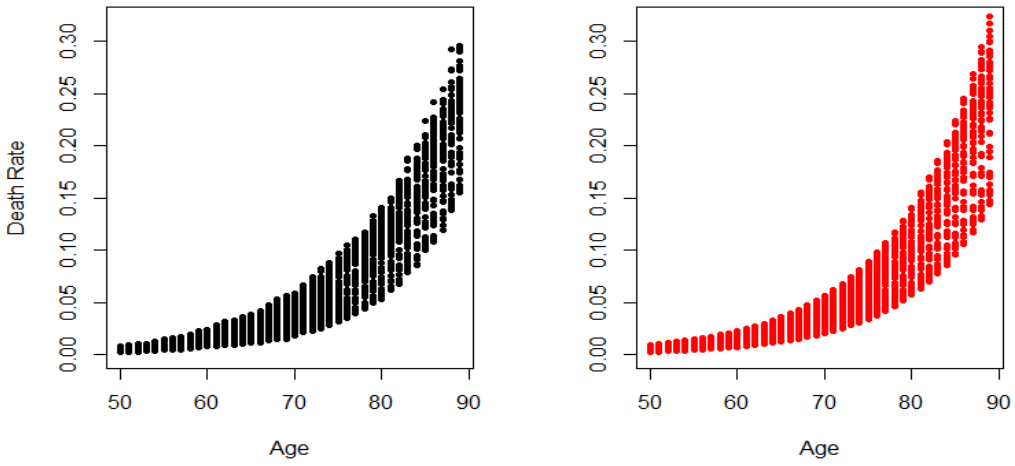
Looking at Figure 10 (b) we see that the variability in the fitted number of deaths for each age group is lower than in the observed data. This highlights the issue with applying Poisson regression for our data, due to over-dispersion in the observed number of deaths at each age, model B fails to capture the true variability in the data by assuming that the mean and variance of $D_{t,x}$ are equal. This causes a problem particularly for older age groups where over-dispersion is most prominent causing the accuracy of the models fitted values to deteriorate for these older age groups, which can be seen in Figure 10 (c). The parameters for this model can be interpreted in the same way as those for model A. β_0' can be thought of as the general level of mortality and β_1' the effect of age on the increase of death rates. We can see in Figure 10 (a) that the parameter estimates for this model are nearly identical to those of model A. These tell us that general mortality has been decreasing over time and that the effect of age seems to be increasing.



(a) Parameter estimates for model B



(b) Observed (left) and fitted (right) number of deaths across years 1961-2016



(c) Observed (left) and fitted (right) values of m_x from model B across years 1961-2016

Figure 10: Fitted values from model B

3.2.1 Stochastic model notation

Next we introduce a series of stochastic models that all take the form $log(m_{t,x}) = \sum_i \beta_x^{(i)} \kappa_t^{(i)} \gamma_{t-x}^{(i)}$ or $logit(m_{t,x}) = \sum_i \beta_x^{(i)} \kappa_t^{(i)} \gamma_{t-x}^{(i)}$ where:

- The $\beta_x^{(i)}$ functions will reflect age related effects
- The $\kappa_t^{(i)}$ functions will reflect period related effects
- The $\gamma_{t-x}^{(i)}$ functions will reflect cohort related effects

3.2.2 The cohort effect

The cohort effect, in the context of this study, refers to the effect of an individual's year of birth on mortality. Including this effect allows for a model to capture the variation in the rates at which mortality has been improving for different age groups in different years. The inspiration for the inclusion of this effect comes from observed phenomenon from previous studies, for example, Willets (2004) [19] found that people born in the U.K. between 1925 and 1945 experienced more rapid improvement in mortality than generations born on either side of this period. Also see Cairns et al (2009)[4] for further evidence supporting the existence of this effect. Since our data set contains an age range of 40 years and 56 years worth of data we have $40 + 56 - 1 = 95$ extra parameters to estimate when adding a cohort effect to a model, or in other words, there are 95 years of birth to consider.

3.3 Model C

Blake et al (2006)[1] fitted the following model to mortality rates $q_{t,x}$:

$$logit(m_{t,x}) = \beta_x^{(1)} \kappa_t^{(1)} + \beta_x^{(2)} \kappa_t^{(2)}$$

Where simple parametric forms were assumed for $\beta_x^{(1)}$ and $\beta_x^{(2)}$:

$$\beta_x^{(1)} = 1$$

$$\beta_x^{(2)} = (x - \bar{x})$$

Where $\bar{x}$ is the mean age in the sample range (in our case $\bar{x} = 69.5$). This gives the following model:

$$logit(q_{t,x}) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x})$$

For the purposes of our study we can use this same model and apply it to death rates giving us our model C:

$$logit(m_{t,x}) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x})$$

The model fit is obtained using a useful `R` package `StMoMo` which is designed for fitting stochastic mortality models. Parameter estimates are obtained by maximizing the model log-likelihood and forecast distributions can be obtained by using simulation techniques which will be discussed in later sections. Villegas et al(2015) [18] provide a much more thorough explanation of how models are fit using this package but, for the purpose of this study, the important thing to note is that this package allows for extra uncertainty when obtaining parameter estimates which hopefully should reflect the uncertainty in our data. The rest of the models in this paper will also be fit using StMoMo.

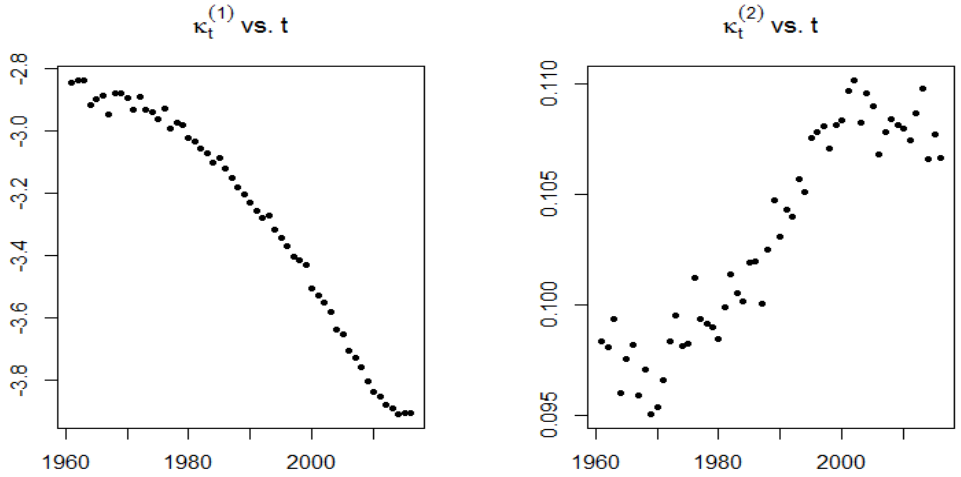
The next thing to note about this model is that we now include period effects $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ as well as an age effect. This allows us to fit to the entire matrix of data rather than a single year at a time which, as we will see, provides a better fit and more reliable forecasts. $\kappa_t^{(1)}$ can be interpreted as the change in general mortality over time and $\kappa_t^{(2)}$ captures the effect of time on the relationship between age and death rates. By assuming a parametric form for

the age effect $\beta_x^{(2)}$ we are assuming an underlying smoothness in the relationship between age and death rates, similar to models A and B, but now the variable age is centered at the mean age for our data. All this changes is the interpretation of the intercept which now refers to a male aged 69.5 rather than 50.

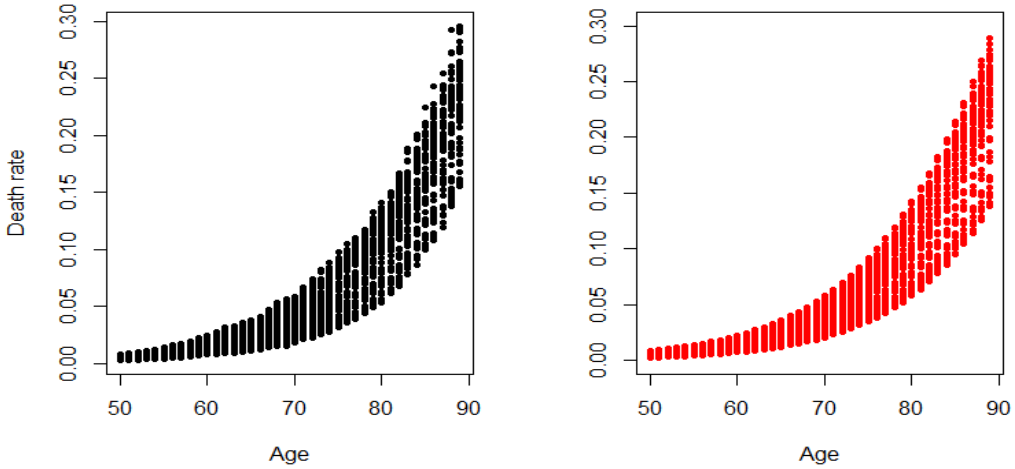
Notice also that we are now using the *logit* link function to model death rates. We assume a linear relationship between age and logit death rates which, as shown in Figure 6, seems an appropriate assumption to make for our data set. The use of the logit link function also implies that this model assumes the number of deaths, $D_{t,x}$, to follow a binomial distribution with parameter $E_{t,x}^c \hat{m}_{t,x}$:

$$D_{t,x} \sim Binomial(E_{t,x}^c \hat{m}_{t,x})$$

The main advantage here compared to using Poisson regression is that there is no restrictions placed on the variance of $D_{t,x}$. This approach allows for a larger variance in $D_{t,x}$. Below we have plotted the parameter estimates and fitted values for model C.



(a) Parameter estimates for Model C



(b) Observed (left) and fitted (right) values of m_x from Model C across years 1961-2016

Figure 11: Parameter estimates and fitted values from Model C

From Figure 11 (a) we see that model C describes mortality much the same as models A and B. The $\kappa_t^{(1)}$ plot shows general mortality improving over time and $\kappa_t^{(2)}$ is increasing but is more scattered than before suggesting that maybe these parameters aren't as strongly correlated as models A and B would suggest. $\kappa_t^{(2)}$ is the slope of our model which has been increasing over time meaning that mortality improvements have been greater in younger age groups. This agrees with our data, we saw in figure 5 that people under the age of 70 have experienced relatively higher rates of mortality than those over the age of 70. Looking at the fitted death rates in Figure 11 (b) we see that the model seems to be quite accurate up to later ages (around 80+) where it begins to stray from our observed data. In this case, mortality rates seem to be underestimated for older ages.

3.4 Model D

Next we introduce a extension of model C that includes a cohort effect:

$$logit(m_{t,x}) = \beta_x^{(1)}\kappa_t^{(1)} + \beta_x^{(2)}\kappa_t^{(2)} + \beta_x^{(3)}\gamma_{t-x}^{(3)}$$

For this model simple parametric forms are assumed for $\beta_x^{(1)}$, $\beta_x^{(2)}$ and $\beta_x^{(3)}$:

$$\beta_x^{(1)} = 1$$

$$\beta_x^{(2)} = (x - \bar{x})$$

$$\beta_x^{(3)} = 1$$

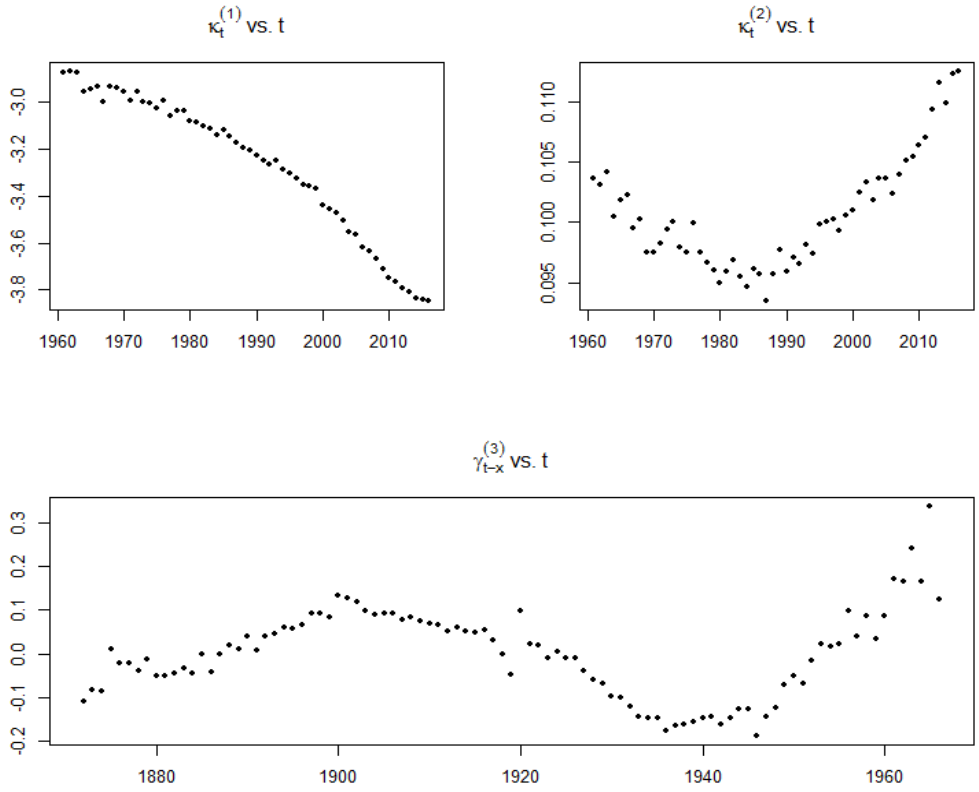
Thus,

$$logit(m_{t,x}) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \gamma_{t-x}^{(3)}$$

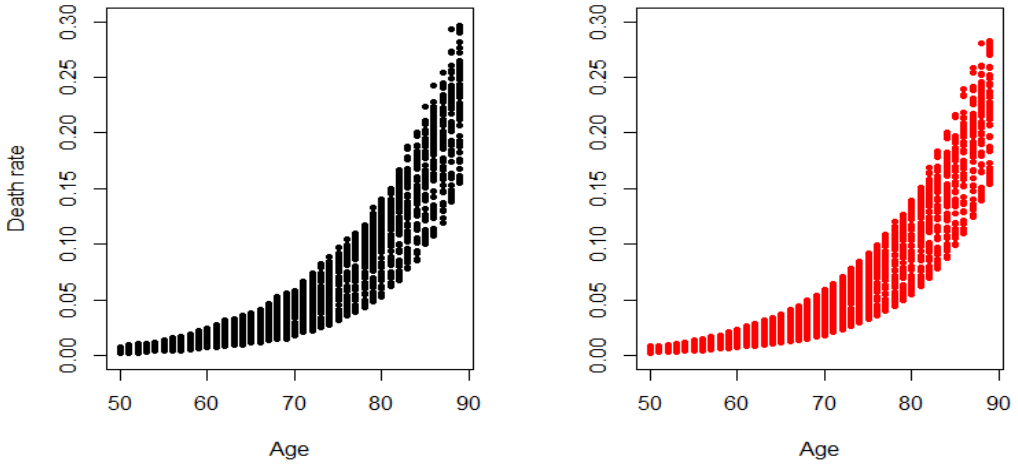
Here we see the first inclusion of the cohort effect. As discussed earlier, this effect describes the effect on an individuals year of birth on their mortality. Due to the nature of this effect it will be centered around 0 where cohorts who experience greater improvements in their mortality rates during their lifetime relative to others will have will have $\gamma_{t-x}^{(3)} < 0$ while cohorts who experience less improvement in their mortality rates during their lifetime relative to others will have will have $\gamma_{t-x}^{(3)} > 0$.

Commenting briefly on the assumptions of the model, here we are again using binomial regression to model deaths, $D_{t,x}$, and assuming that $logit(m_{t,x})$ is linear with respect to age much like Model C which agrees with our data. The significant difference with this model is that we are assuming the existence of a cohort effect. We have already seen evidence behind the existence of the cohort effect in our data in Figure 8.

Below are the parameter estimates and fitted values for model D.



(a) Parameter estimates from Model D



(b) Observed (left) and fitted (right) values of m_x from Model D across years 1961-2016

Figure 12: Parameter estimates and fitted values from Model D

From Figure 12 (a) we see that the general mortality level $\kappa_t^{(1)}$ is decreasing as expected but now we see a different pattern in the period effect linked to age $\kappa_t^{(2)}$, this is partly due to the presence of the cohort effect. What these parameter estimates are now telling us is that younger age groups didn't begin to experience the relatively better rates of improvement until around 1990 where we see the estimates for $\kappa_t^{(2)}$ begin to increase. From the $\gamma_{t-x}^{(3)}$ estimates we see that cohorts born around 1900 experienced worse rates of mortality improvement relative to other cohorts while cohorts born around 1940 experienced better rates of mortality improvement compared to other cohorts in our data set. The first observation here is certainly consistent with our data as we saw in figure 8 that the 1903 cohort seemed to

carry negative rates of improvement consistently through our data.

We can also see from Figure 12 (b) that, visually, this model appears to be giving the most accurate fit to our data of any model so far which is significant since this is the only model so far to include a cohort effect. The improvement in the goodness of fit will be assessed formally in a later section.

3.5 Model E

This model is an extension of Model D adding in a quadratic term into the age effect. The model is initially expressed as:

$$\text{logit}(m_{t,x}) = \beta_x^{(1)}\kappa_t^{(1)} + \beta_x^{(2)}\kappa_t^{(2)} + \beta_x^{(3)}\kappa_t^{(3)} + \beta_x^{(4)}\gamma_{t-x}^{(4)}$$

The age effects are assumed to be of the form:

$$\beta_x^{(1)} = 1$$
$$\beta_x^{(2)} = (x - \bar{x})$$
$$\beta_x^{(3)} = ((x - \bar{x})^2 - \hat{\sigma}_x^2)$$
$$\beta_x^{(4)} = 1$$

Where $\hat{\sigma}_x^2$ is the mean of $(x - \bar{x})^2$. Thus:

$$\text{logit}(m_{t,x}) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \kappa_t^{(3)}((x - \bar{x})^2 - \hat{\sigma}_x^2) + \gamma_{t-x}^{(4)}$$

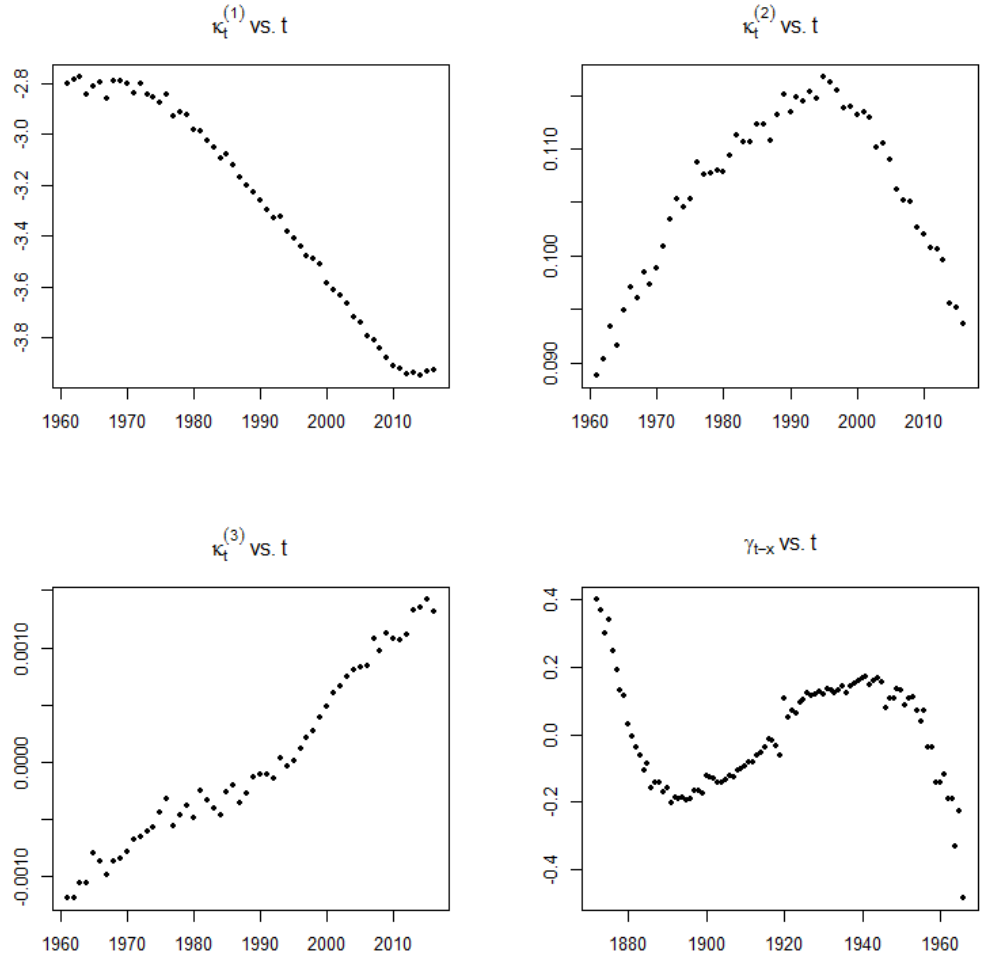
The only new assumption compared to Model D is that we now assume a non-linear relationship between $logit(m_{t,x})$ and age.

When this model was first developed the inspiration behind adding a quadratic term for age came from possible curvature observed in the $logit(q_{t,x})$ vs. age plot from US data [4]. As seen in Figure 6 there appears to be a linear relationship between these two variables for our Irish males data meaning the quadratic term in this model may not be necessary and that this assumption does not fully agree with our data.

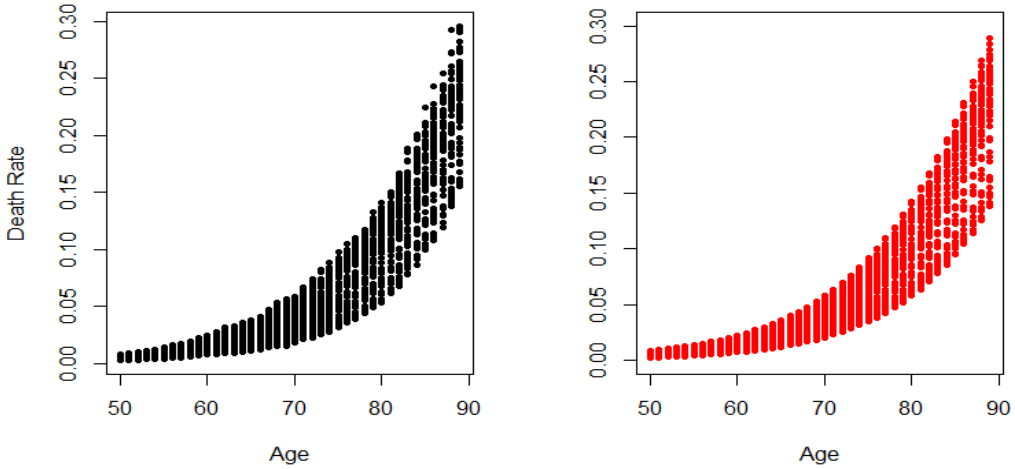
Plotted below in Figure 13 are the parameter estimates for model E as well as the fitted values $\hat{m}_x$ obtained from fitting model E alongside the observed data.

Taking a look at the parameter estimates for Model E in Figure 13 (a) we see that $\kappa_t^{(1)}$ follows a pattern that we would expect to see. Due to the addition of the quadratic age term,$((x - \bar{x})^2 - \hat{\sigma}_x^2)$, the dynamics of mortality over time become highly complex and interpreting the parameters $\kappa_t^{(3)}$,$\kappa_t^{(3)}$ and $\gamma_{t-x}^{(4)}$ are difficult to interpret.

Visually the model seems to display an adequate fit to our data with perhaps some slight under-estimation of death rates for people aged 85+ in more recent years.



(a) Parameter estimates from Model E



(b) Observed (left) and fitted (right) values of m_x from Model E across years 1961-2016

Figure 13: Parameter estimates and fitted values from Model E

3.6 Model F

This model is an extension of the Lee and Carter (1992)[12] model that includes a cohort effect as follows:

$$log(m_{t,x}) = \beta_x^{(1)} + \beta_x^{(2)}\kappa_t^{(2)} + \beta_x^{(3)}\gamma_{t-x}^{(3)}$$

Where we assume

$$\beta_x^{(3)} = 1$$

Giving us the model

$$log(m_{t,x}) = \beta_x^{(1)} + \beta_x^{(2)}\kappa_t^{(2)} + \gamma_{t-x}^{(3)}$$

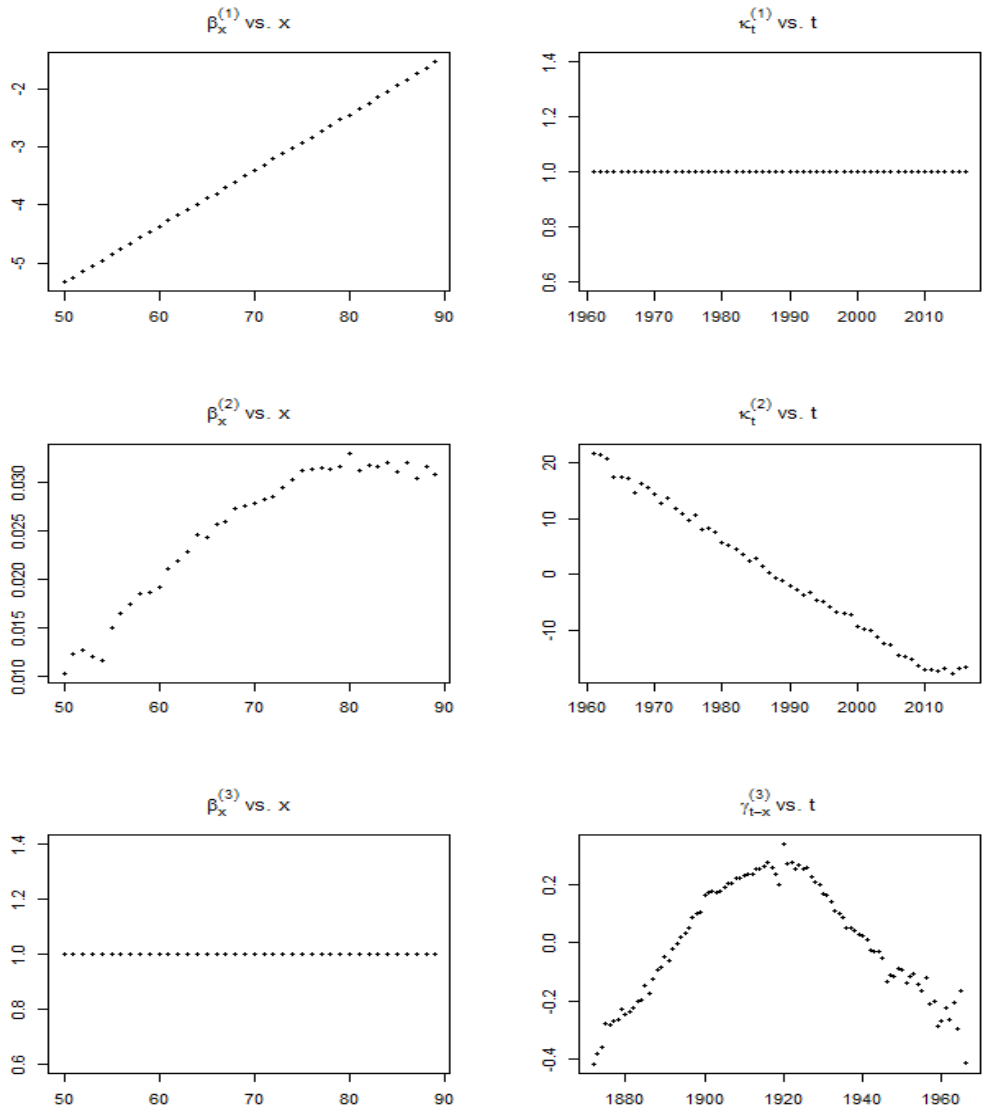
So the Lee-Carter model is a special case where $\beta_x^{(3)}$ and $\gamma_{t-x}^{(3)}$ are set to zero.

This model builds upon the advantages of the Lee-Carter model and allows us to take a different approach to modelling death rates compared to models C, D and E.

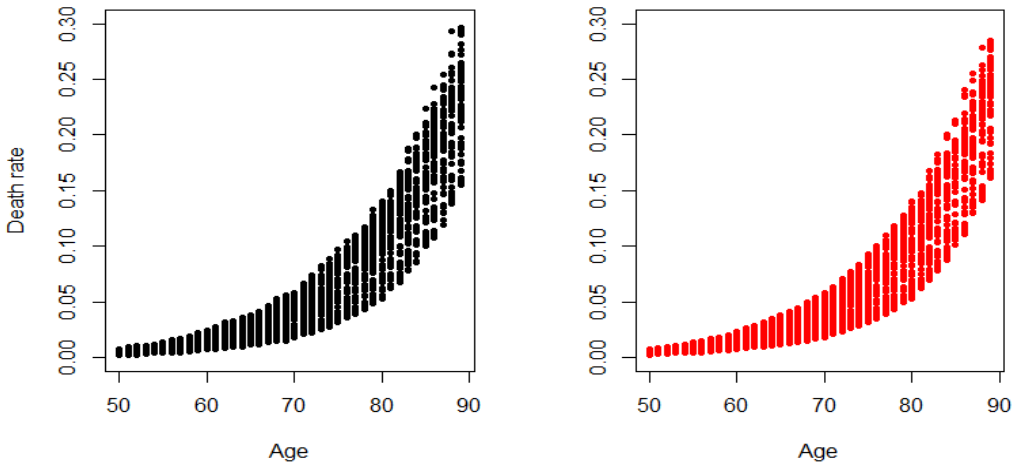
One of the key features of this model that differs from the previous three is that now we are using Poisson regression. As discussed earlier this may pose issues due to over-dispersion of deaths at higher ages. Another distinct feature of this model is that we now have an independent age effect $\beta_x^{(1)}$ rather than an independent period effect, this model focuses more on describing the progression of the relationship between age and death rates over time rather than the trend of general mortality improvement over time. Also notice that now we do not assume any parametric form of the period effects. In other words, this model does not assume any smoothness in the relationship between age and death rates. Depending on one's beliefs about the underlying randomness in age effect one might favor a model that aligns with their beliefs. For example, if one believes that should be an underlying smoothness between ages and that there should be randomness from one year to the next and between cohorts, then models D or E might be favoured. If one believes that there is underlying randomness in all of these effects, more weight might be placed on Model F.

Plotted below are the parameter estimates and fitted values from model F.

Looking at Figure 14 (b) we can see that model F provides a reasonably good fit to our data. What is interesting about the parameter estimates is that when looking at $\beta_x^{(1)}$, the model fits an almost perfectly linear relationship between age and log death rates without any influence of smoothing. All this provides is further evidence that the relationship between age and death rates is exponential. We also see that, although the age effect $\beta_x^{(2)}$ is increasing, $\kappa_t^{(2)}$ has been decreasing on a much larger scale dominating the influence of this term. Since $\kappa_t^{(1)}$ has been muted in this model the effect of decreasing general mortality has been shifted to this term which worsens the model's ability to explain the relationship between death rates and time since the relationship is being described by a single parameter.



(a) Parameter estimates from Model F



(b) Observed (left) and fitted (right) values of m_x from Model F across years 1961-2016

Figure 14: Parameter estimates and fitted values from Model F

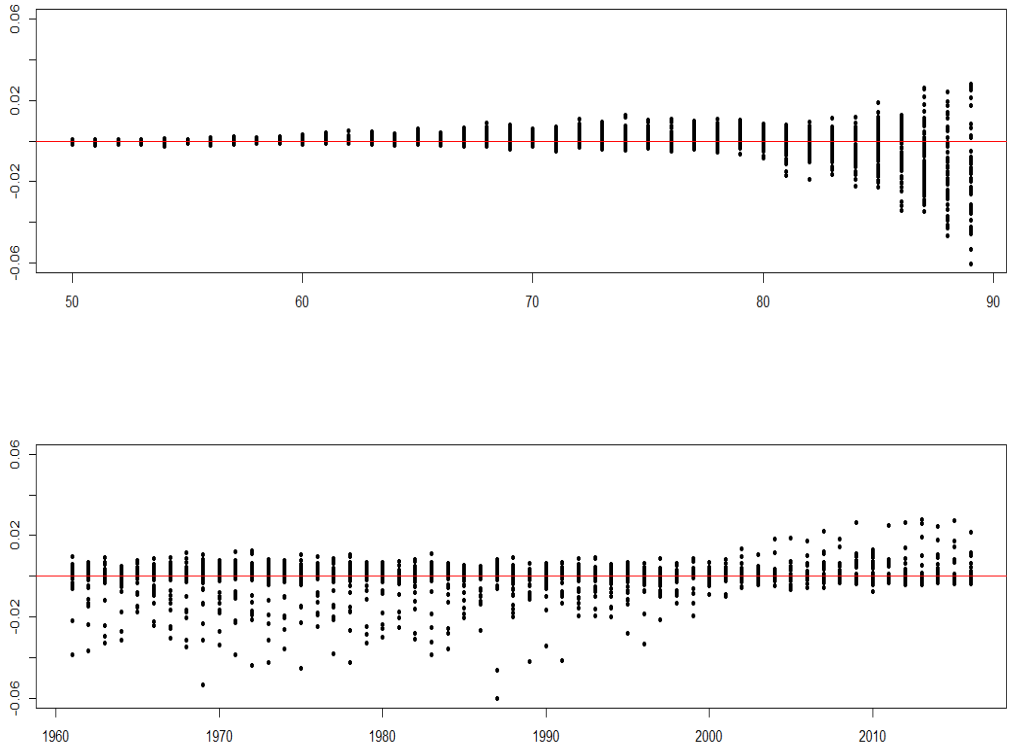
3.7 Comparisons

The aim of this section is to compare the goodness of fit and forecast reliability of each model based on a chosen set of quantitative and qualitative criteria and select the optimal model(s) to use in estimating longevity risk.

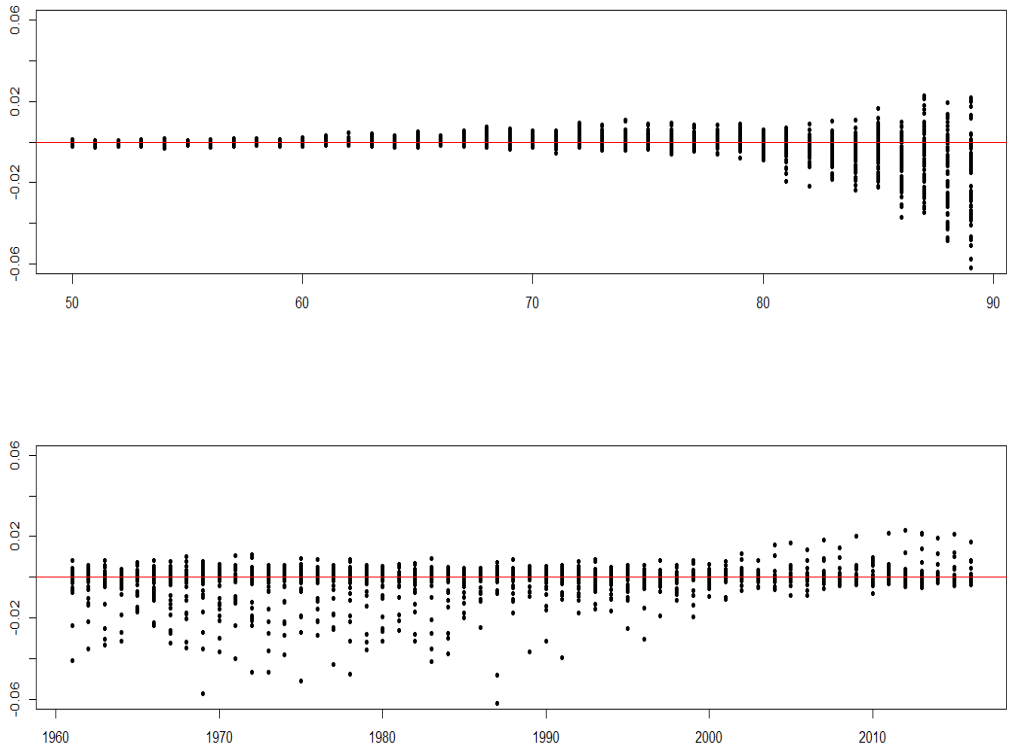
3.7.1 Goodness of fit

So far we have taken a crude look at how the fitted values of our models match up against the observed data. For a more accurate evaluation, we analyze the residuals and standardized residuals of each model, similar to analysis performed in previous studies [6]. Plotted below are the residuals of each model relative to the variables Age and Year from our data.

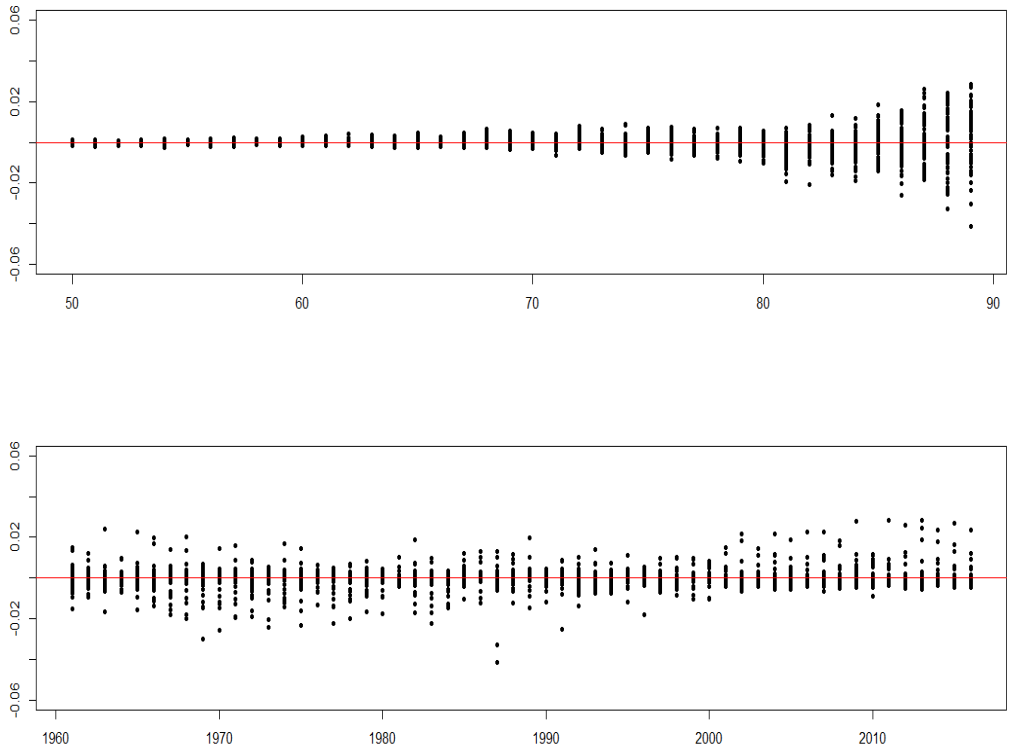
These plots show that the accuracy of all six models with respect to age tends to deteriorate after age 80. Models A and B seem to have the widest range of errors with respect to age. This is due to the nature of mortality becoming more unpredictable at older ages. With respect to the variable year, all models tend to become more accurate as time goes on. This seems to suggest that mortality improvement has become more predictable in recent years but this is most likely due to improvement in the quality of data over time rather than an underlying feature of mortality improvement.



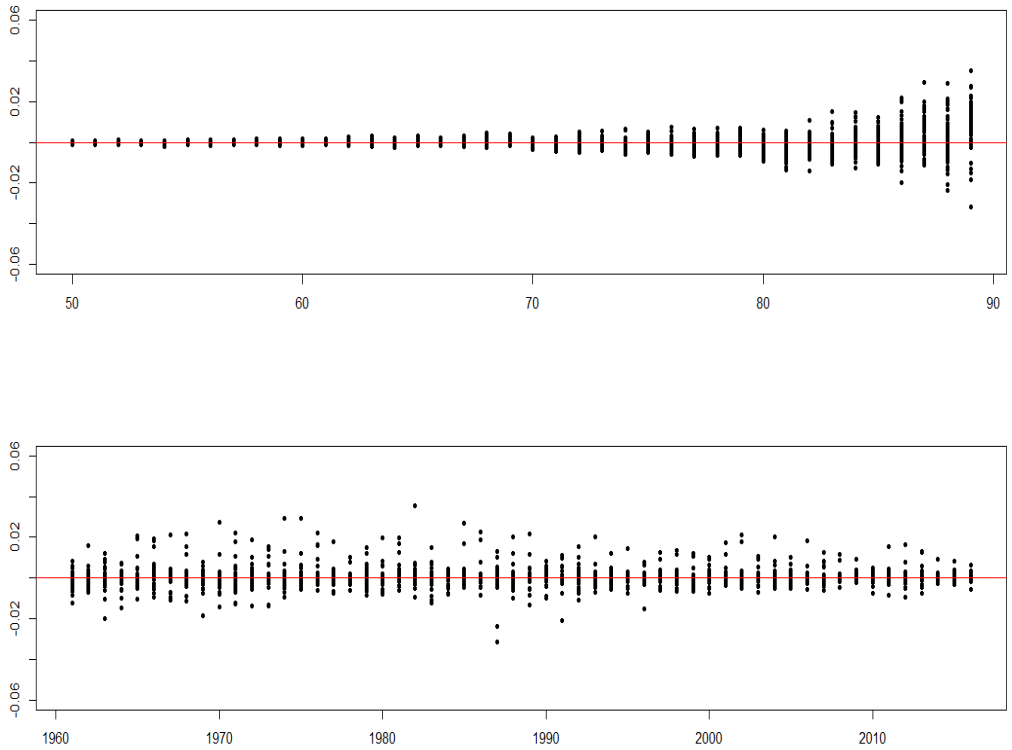
(a) Model A residuals



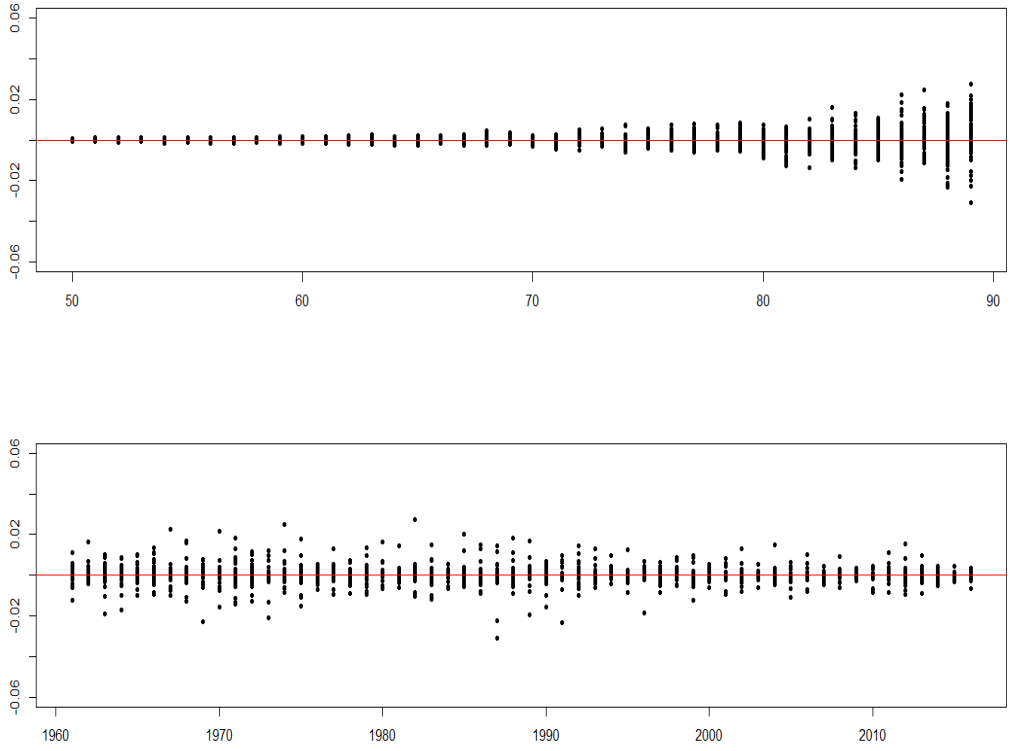
(b) Model B residuals



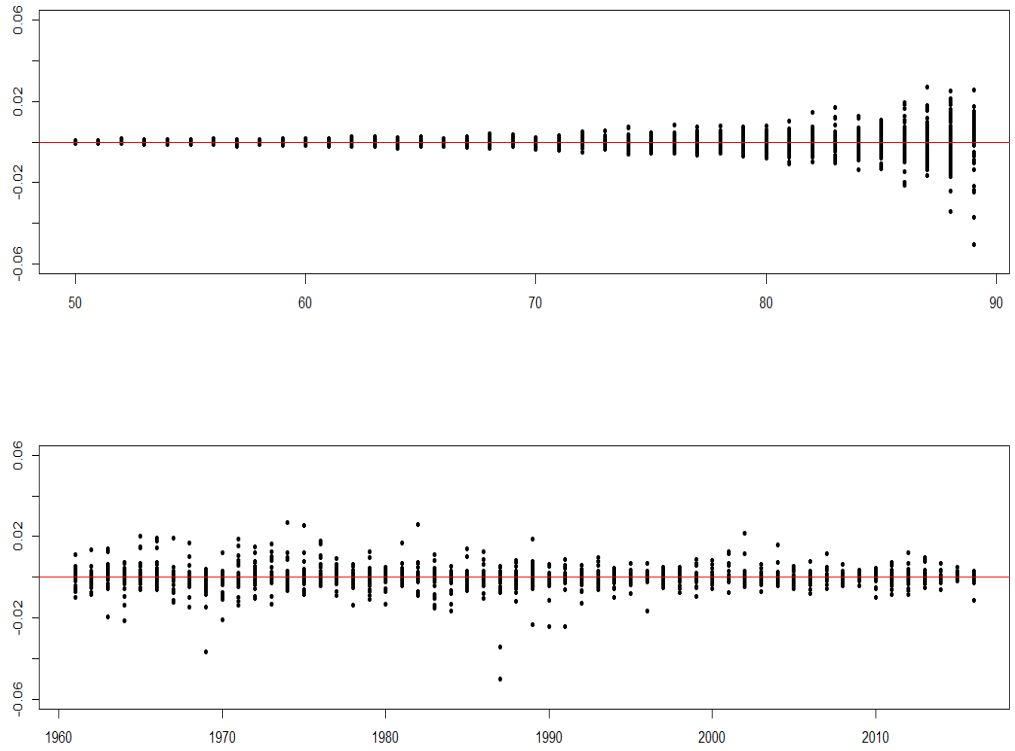
(c) Model C residuals



(d) Model D residuals



(e) Model E residuals



(f) Model F residuals

To quantify the quality of fit of each model we consider three metrics, mean average deviation, mean squared error and mean average percentage error defined as:

$$MAD = \frac{1}{n}\sum_{i=1}^{n}|m_i - \hat{m}_i|$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(m_i - \hat{m}_i)^2$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}|\frac{m_i - \hat{m}_i}{m_i}|$$

Tabulated below are the MAD, MSE and MAP values for each model with respect to their residuals.

	MAD	MSE	MAPE
Model A	0.00376	5.3×10^{-5}	6.08%
Model B	0.00355	5.34×10^{-5}	6.2%
Model C	0.00274	2.47×10^{-5}	5.38%
Model D	0.0022	1.81×10^{-5}	3.9%
Model E	0.002	1.37×10^{-5}	3.57%
Model F	0.0021	1.64×10^{-5}	3.67%

Table 1: MAD, MSE and MAPE values for the residuals of each model

Table 1 tells us that models D, E and F seem to be providing the best fit to our data with model E showing Model E marginally ranking as the best fit with respect to these three metrics.

Next we look at the standardized residuals of each of the three models, defined as:

$$r_{x,t} = \frac{m_{x,t} - \hat{m}_{x,t}}{\sqrt{\hat{m}_{x,t}/E_{x,t}}}$$

By construction, standardized residuals should approximately follow a standard normal distribution, so we expect to find mean ≈ 0, standard deviation ≈ 1, skewness ≈ 0 and kurtosis ≈ 3.

	Mean	Std.Dev	Skew	Kurt
Model A	0.1949	1.6685	0.3335	-0.1289
Model B	-0.0029	0.0175	-1.3793	5.5066
Model C	-0.0056	1.4356	0.3875	0.3714
Model D	0.0072	1.0397	0.139	0.0445
Model E	0.0009	0.9591	0.1115	0.0254
Model F	-0.0003	0.9804	0.1047	0.0311

Table 2: Main statistics for the standardized residuals of each model.

Looking at table 2, for models A and B these statistics are not what we would hope to see. What this tells us is that the amount of error in models A and B is not consistent across the full range of observed data, again highlighting the inaccuracy of models A and B for older ages. The standardized residuals of models D, E and F seem to closest to being normally distributed. To assess this more rigorously we carry out a Shapiro-Wilk normality test and find that only models A, B and C show highly significant departures from normality, indicating that models C−F seem to be giving more consistent fits across the entire data set.

3.7.2 forecast reliability

With each model, we can obtain predictions of future death rates. Since we will use these predictions to evaluate the amount of longevity risk faced by our SIPP product it is important that we assess the reliability of these predictions.

To gain a clearer idea of how the forecasts for each model compare to one another we can perform a number of "backtests", similar to those done by Maccheroni (2017)[13], by splitting our data into two period subsets and fitting each model to the subset of older data and use the parameter estimates to forecast death rates for the subset of more recent data. We can then compare the forecasts of each model to experienced data.

Our first test consists of training each model on a subset of our data and using these fits to obtain point estimate forecasts for the remaining data and measuring the residual error. For the short term test, we split the data 75:25 and train each model on the first forty-two years (1961-2002) and forecast death rates for the remaining fourteen years (2003-2016). For the medium-term test, the data was split 50:50, each model was fitted to the data from 1961-1988 and forecasts were obtained for 1989-2016. Measures of point estimate short-term and medium-term forecast error for each model are tabulated below.

	Short Term		
	MAD	MSE	MAPE
Model A	0.00526	6.81×10^{-5}	12.37%
Model B	0.00244	2.26×10^{-5}	9.05%
Model C	0.00548	6.7×10^{-5}	13.91%
Model D	0.00321	2.13×10^{-5}	10.32%
Model E	0.00709	1.93×10^{-4}	11.68%
Model F	0.00211	1.29×10^{-5}	6.87%

Table 3: Three metrics, MAD, MSE and MAPE, summarizing the short term forecast error of each of the models

	Medium Term		
	MAD	MSE	MAPE
Model A	0.00966	2.49×10^{-4}	19.53%
Model B	0.0042	4.63×10^{-5}	15.46%
Model C	0.00818	1.52×10^{-4}	25.88%
Model D	0.00362	4.12×10^{-5}	11.91%
Model E	0.00729	2.09×10^{-4}	12.94%
Model F	0.00311	3.09×10^{-5}	9.98%

Table 4: Three metrics, MAD, MSE and MAPE, summarizing the medium term forecast error of each of the three models

	MAD	MSE	MAPE
Model A	83.66%	266.07%	57.83%
Model B	72.43%	104.35%	70.73%
Model C	49.19%	127.65%	86.01%
Model D	12.84%	93.12%	15.45%
Model E	3.99%	8.01%	10.81%
Model F	47.76%	139.79%	45.27%

Table 5: Percentage increases in MAD, MSE and MAPE for each model when going from short term to medium term forecasting tests

Tables 3 and 4 summarize the accuracy of the predictions of each model while table 5 gives us an idea of how sensitive each model is with respect to the amount of data given. Although high forecast accuracy is desirable, a lot of consideration should be given to the results of table 5 as well. Initially, each model is used to obtain predictions for 14 years of death rates using 42 years of data. Then, we perform the same test except each model is now used to predict 28 years of death rates given only 28 years of data. So the results in Table 5 are a good measure of a model's stability when used to forecast over a longer period and when given less data.

Looking at each table we see that models A and C seem to perform the worst in terms of both accuracy and stability. Model B provides reasonably accurate short-term forecasts which is unexpected as earlier results suggested that it does not fit our data well, although it is quite inconsistent when given less data and used to predict for longer periods, as shown in Table 5. Model D seems to have the optimal balance between forecast accuracy and stability. Model E appears to be the least sensitive to changes in the data by a significant margin but displays poor forecasting accuracy. Model F ranks the highest in terms of accuracy, as shown in tables 3 and 4, but is quite sensitive to changes in data.

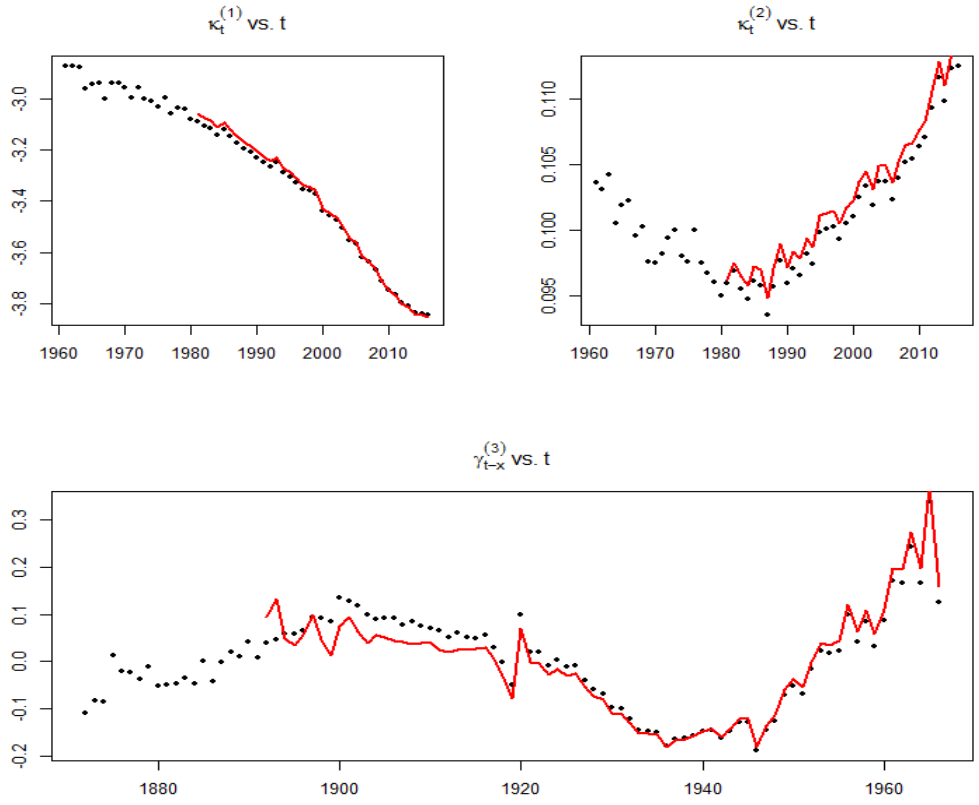
The main message to take from the above analysis is that models A, B and C are objectively worse than models D, E and F in terms of their goodness of fit and the reliability of their forecasts. Also note that models A,B and C do not incorporate a cohort effect while models D, E and F do. So without explicitly testing for the significance of the cohort effect, we have found evidence that strongly suggests that its inclusion is necessary.

When deciding which model to use to obtain forecasts from the remaining three there is some subjectivity involved. Depending on what qualities were most desirable to the reader one might favour a particular model over another. For example, if one required a model with high forecasting accuracy with little regard for stability over longer periods one might favour Model F. To further compare models D, E and F we rank each based on their BIC values and consider the robustness of their parameter estimates when given different periods of data.

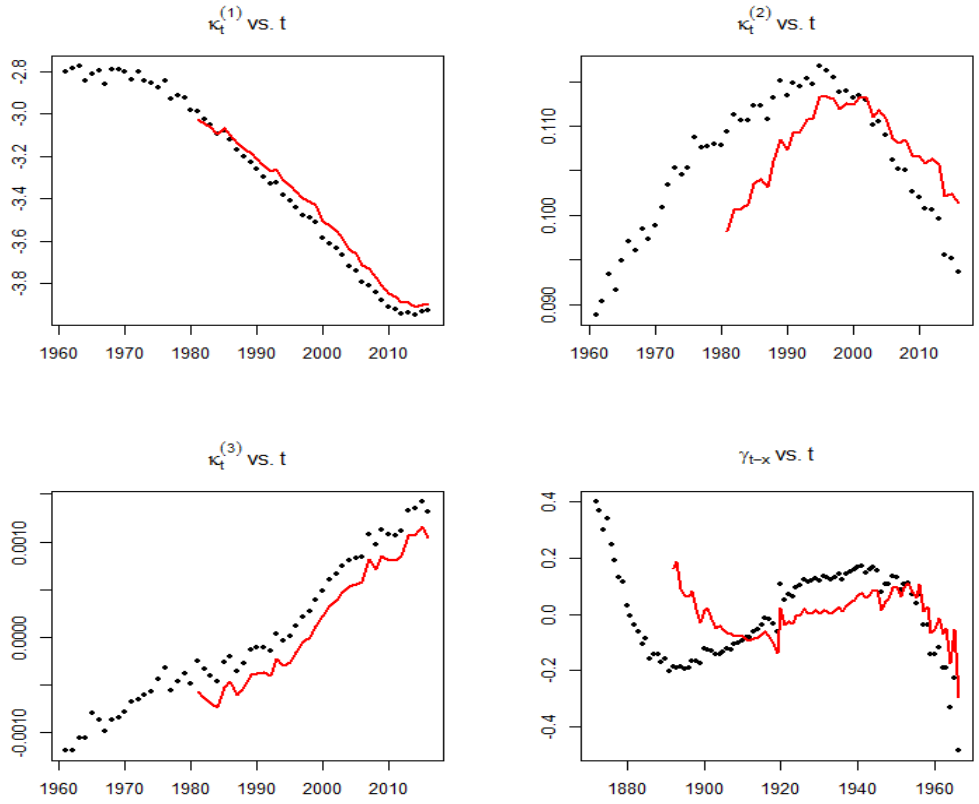
	BIC	
Model D	22091.43	(2)
Model E	22123.4	(3)
Model F	21969.76	(1)

Table 6: BIC measures for models D,E and F

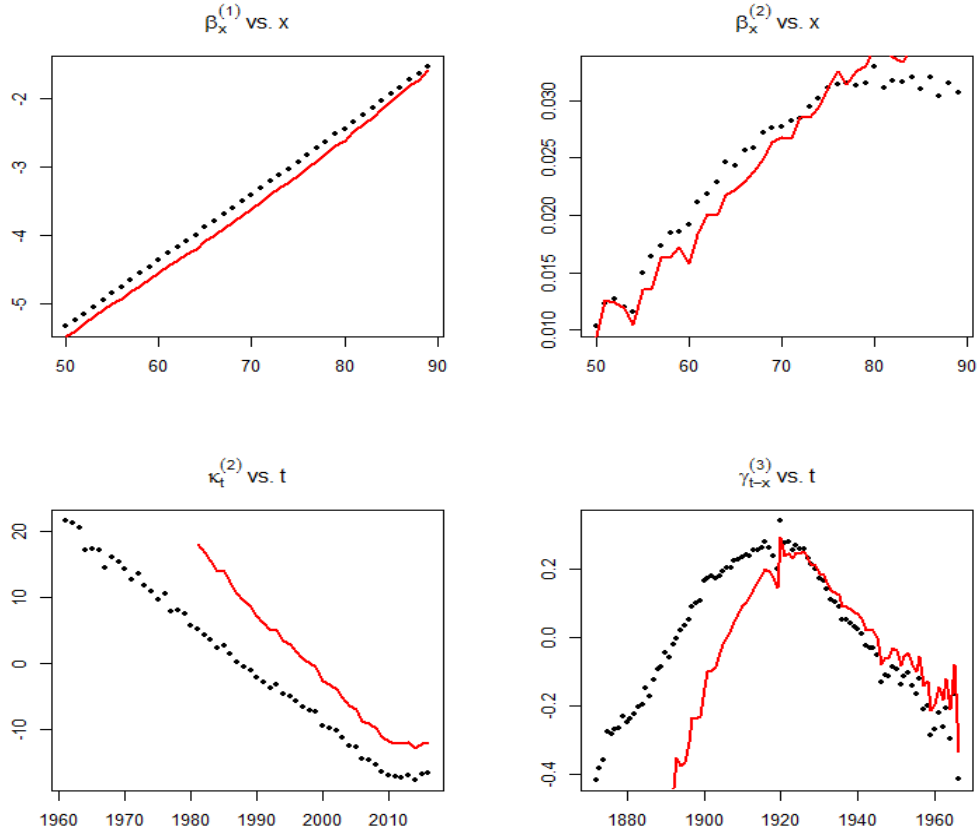
Below are plots of the parameter estimates for models D, E and F given the full range of data as well as red lines showing the parameter estimates of each model when only data from 1981 to 2016 is given. What these plots tell us is that the parameter estimates of Model D seem to be the most robust when given different periods of data which is a desirable quality when considering to forecast with a particular model. We can also see that Model E seems to be more sensitive to changes in the data it is provided with than table 5 would suggest.



(a) Parameter estimates for Model D using Irish Males data from 1961 to 2016 (dots) and 1981 to 2016 (lines)



(b) Parameter estimates for Model E using Irish Males data from 1961 to 2016 (dots) and 1981 to 2016 (lines)



(c) Parameter estimates for Model F using Irish Males data from 1961 to 2016 (dots) and 1981 to 2016 (lines)

4 Forecasting

In order to estimate the longevity risk faced by the SIPP product we first need projections of future death rates. In this section, we take a look at death rate forecasts over the next 35 years according to three different models, detail how they are obtained and compare each forecast. Given the results from the previous section, we choose to forecast using only models D, E and F. The forecasting of death rates requires the modelling of the period and cohort effects using time series techniques. In the package StMoMo, The forecasting of each model is implemented using the function `forecast`.

For each model the period effects $\kappa_t^{(i)}$ are projected using the standard approach ([3], [11]) using a random walk with drift. Specifically it is assumed that:

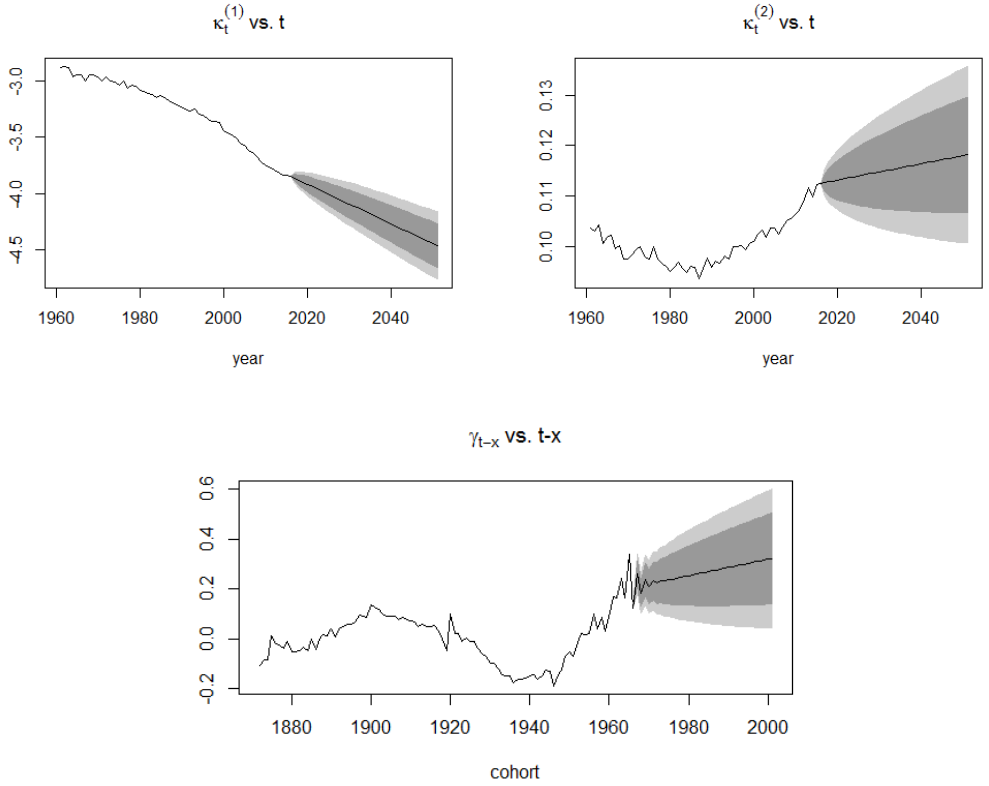
$$\kappa_t^{(i)} = \delta + \kappa_{t-1}^{(i)} + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma)$$

where δ is the drift parameter and σ is the variance of the white noise ϵ_t.

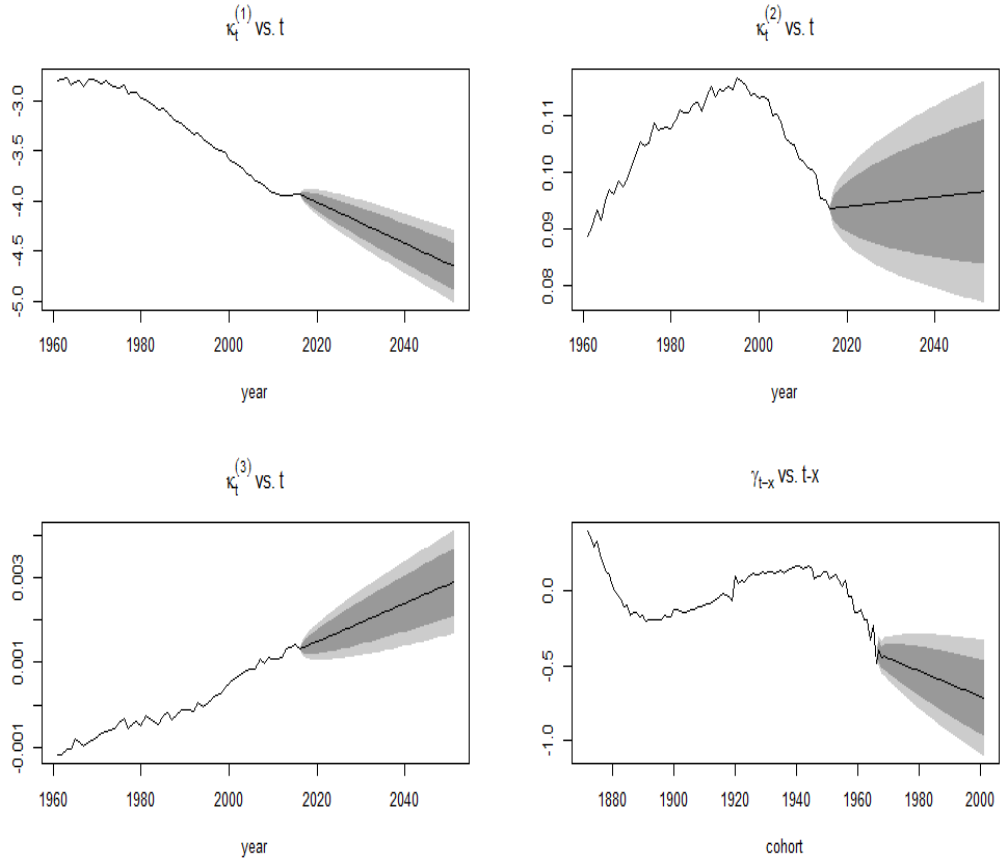
As pointed out by Currie (2016)[9], one of the main challenges when forecasting stochastic mortality models is specifying the dynamics of the cohort effect. The function `forecast` follows previous studies ([15], [5]) and it is assumed that the cohort effect, γ_{t-x}, follows a univariate ARIMA process which is independent of the period effect, κ_t. By default, the cohort effect is projected using an ARIMA(1,1,0) process with drift:

$$\Delta\gamma_{t-x} = \delta + \theta_1 \Delta\gamma_{t-x-1} + \epsilon_{t-x}$$

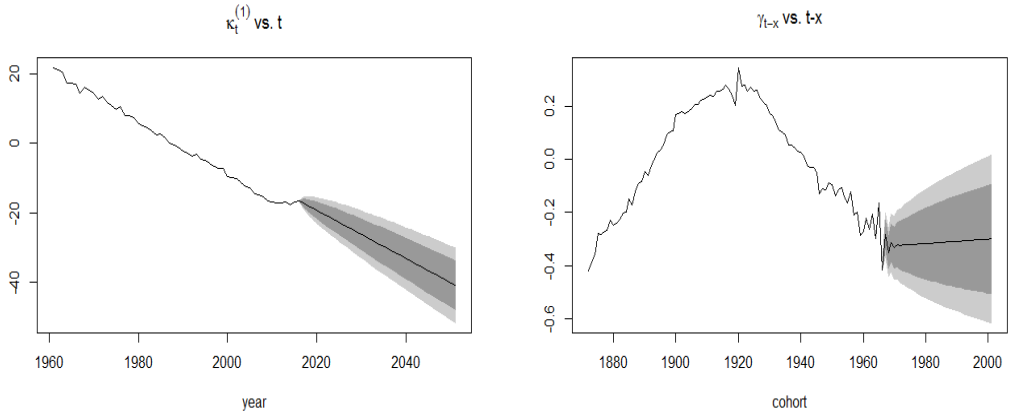
Plotted below are the projected period and cohort effects for each model.



(a) Parameter Forecasts for the next 35 years from Model D



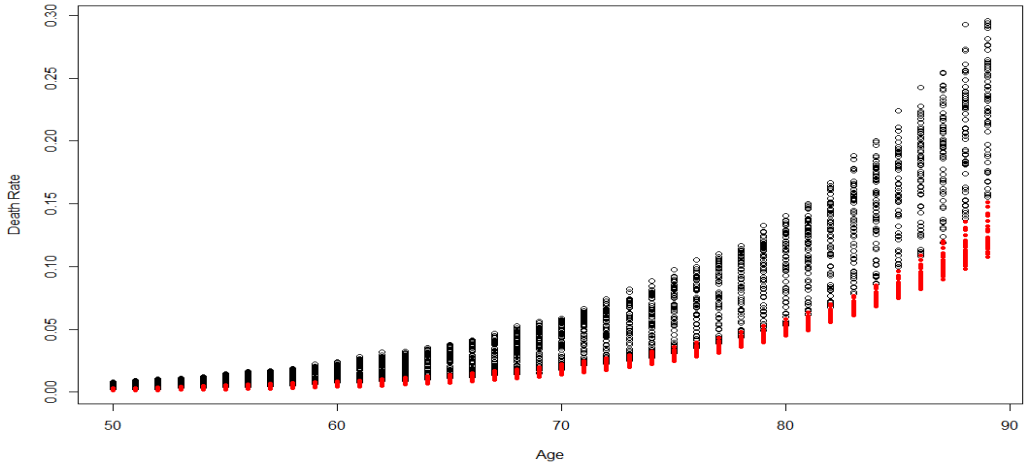
(b) Parameter Forecasts for the next 35 years from Model E



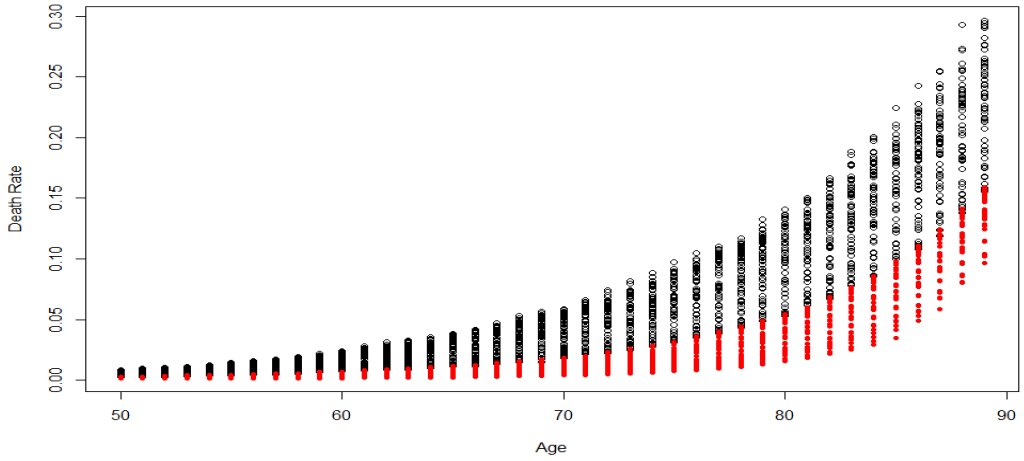
(c) Parameter Forecasts for the next 35 years from Model F

Figure 17: forecast of the period and cohort effects of models D,E and F applied to the Irish Males population for ages 50-89 and the period 1961-2051. Solid lines represent the central forecast, dark grey fans represent the 80% prediction intervals and light grey fans represent the 95% prediction intervals.

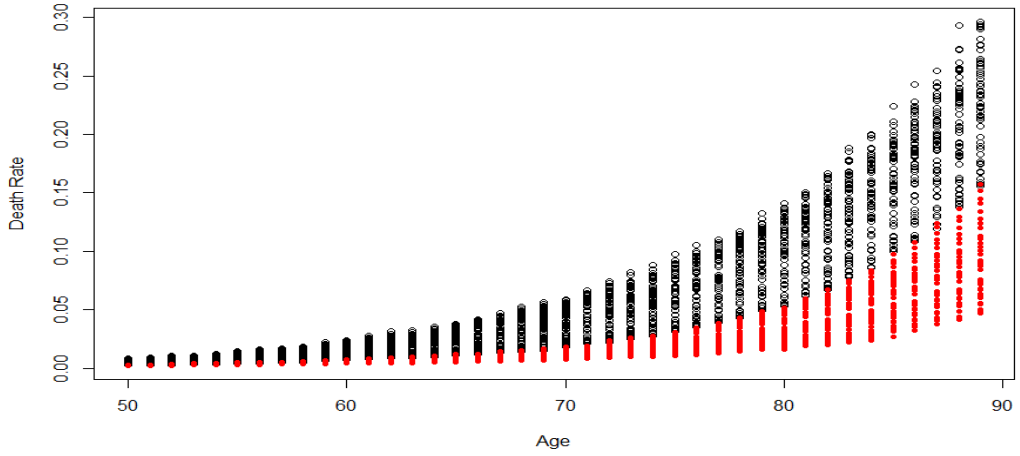
Next, we look at the mean death rate forecasts produced by each model. It is important to remember that each of these models provides a distribution of forecasts which gives a more complete picture of their projections, but looking at mean forecasts will give us an idea of how each projection differs.



(a) Model D



(b) Model E



(c) Model F

Figure 18: Mean death rate forecasts for the next 35 years from models D,E and F with respect to age. The observed death rates for Irish Males from 1961-2016 are plotted with black circles (hollow) and the mean forecasts are plotted using red dots (solid).

From Figure 18 we can see stark differences between the projections of each model. Model D looks to be the most conservative in its projections predicting slight improvement, more so at older ages, but nothing like what has been observed over the last sixty years. Model F on the other hand predicts extremely high rates of improvement at all ages with death rates for 89-year-old males projected to be less than 5% in 2050 compared to the current estimates of roughly 16%. These projections would be viewed as quite implausible by most standards in relation to future mortality but it is important to keep an open mind and to remember that the need for this study stems from the fact that past projections have consistently underestimated improvements in mortality. Model E provides a more balanced forecast predicting fairly significant improvements in mortality while retaining the relationship between age and death rates, i.e. death rates still tend to start increasing quite sharply at ages 80+. Depending on one's personal expectations for future mortality one might favour one model's forecast over another. For example, while most would disregard the projections of Model F some believe that there is no theoretical limit to human age and would consider these projections more plausible than those of models D and E.

5 Longevity Risk

In this section, we estimate the longevity risk associated with a specific life annuity product known as the SIPP with respect to models D, E and F and assess the impact of this risk on the product.

The SIPP is tailored for people who expect to retire at 65. Customers can purchase the product at age 65 where they would immediately begin to receive annual payments in arrears of €10,000 for a term of 25 years, this is referred to as option 1. Alternatively, customers can avail of option 2 and purchase the product at age 55, 10 years prior to their expected retirement, whereupon reaching age 65 they will begin to receive annual payments in arrears of €10,000 for a term of 25 years. Both options must be purchased by a single lump-sum payment. The price of this product is dependent on the expected net present value of the future payments which is calculated using an assumed set of mortality rates for the customer base and a chosen fixed interest rate. The interest rate may vary but remains fixed for the term of the payments meaning we can calculate a price for a variety of different interest rates but once an interest rate is chosen the only remaining variable is the set of mortality rates.

There is debate surrounding life-related insurance and annuity products as to whether or not future mortality improvement should be taken into account when pricing such products. Some may argue that the observed trend of mortality improvement shows no signs of stopping and that pricing without taking this trend into account can leave companies exposed to the risk of the net present value of their payments being higher than expected. Others argue that there is no underlying physical law that mortality follows (or at least not that we know of) hence we cannot make predictions that are accurate enough to be considered in pricing and if mortality improvement is to be taken into account we raise a new set of issues. Firstly, it must be assumed that there is no chance of life expectancy ever-receding at any point in the future which some may consider a bold assumption. Companies also face the question of "Whose projections do we accept?", as we have shown there are numerous valid models that can give vastly different projections. Another issue raised by incorporating mortality improvement into pricing is that prices of life insurance products would increase thus becoming less competitive. It could also be argued that since most mortality models give projections for the general population that these projections are not entirely relevant to the insured population. For example, recent estimates from the Central Statistics Office suggest that nearly 73% of all deaths in Ireland are caused by three main contributors: cancer; circulatory diseases (including heart disease) and respiratory diseases. If more advanced treatments were to be developed resulting in these contributors becoming highly curable or preventable, which is the main focus of a lot of current research, then we would see a massive increase in general life expectancy. But the insured population undergoes screening processes meaning they are generally healthier and would not experience the same levels of mortality improvement as the general population if such treatments were to be developed.

There are valid arguments on either side but, for most companies, the decision ultimately depends on the industry standard and how competitive they require their prices to be. In the UK the Continuous Mortality Investigation (CMI) provides a standard model and estimates for future mortality improvement. The CMI model is deterministic and depends on two factors, age and rate of mortality improvement. The rate of mortality improvement is divided into short term and long term rates. Short-term rates are estimated based on historical data and long-term rates are assumption based. The CMI model is updated annually and aims to provide companies with quality, impartial analysis of mortality improvement but the degree to which companies choose to incorporate this analysis into their underwriting processes varies depending on the company.

For the sake of impartiality, we provide the expected present value of benefits and associated longevity risk for each option of the SIPP with and without taking mortality improvement into account.

	Option 1	Option 2
Current	€171706.41	€159006.59
Model D	€182399.70	€162187.80
Model E	€185114.30	€193253.10
Model F	€193749.10	€193018.10

Table 7: Expected present value of benefits of options 1 & 2 of the SIPP according to current mortality rates and projected mortality rates from models D, E and F at 0% interest.

In Table 7 the "Current" estimates refer to the net present value of the benefits calculated using the most recent observed mortality data, i.e. assuming mortality will remain at its current level and not taking into account any potential mortality improvement. This table shows us that the EPV of benefits predicted by each of our three models is much higher than what would be calculated by just using current mortality rates and also provides more insight into the differences between the forecasts of each model. The most interesting prediction of the three here is model E. Since the structure of benefit payments is identical for both options but option 2 has a 10 year deferral period, we would expect the EPV of benefits of option 2 to be lower than those of option 1 but model E predicts otherwise. This model predicts mortality rates for the 55-year-old group to improve so greatly that they are expected to live to an older age, on average, than the group who are already age 65.

For each of the models, 1000 sample paths were simulated for future mortality rates. Each sample path gives a different set of mortality rates and thus a different EPV of benefits. These values can then be collected to form a distribution. The mean of the forecast distribution for each model is given in Table 7.

Given that we have a forecast distribution for each model we can then calculate the longevity risk according to each model using the definition:

$$LR = \frac{\text{99.5\% quantile of the distribution of price}}{\text{Mean price}}$$

This is best illustrated using an example, below is the distribution of EPV of benefits for option 1 obtained from 1000 simulations of Model D.

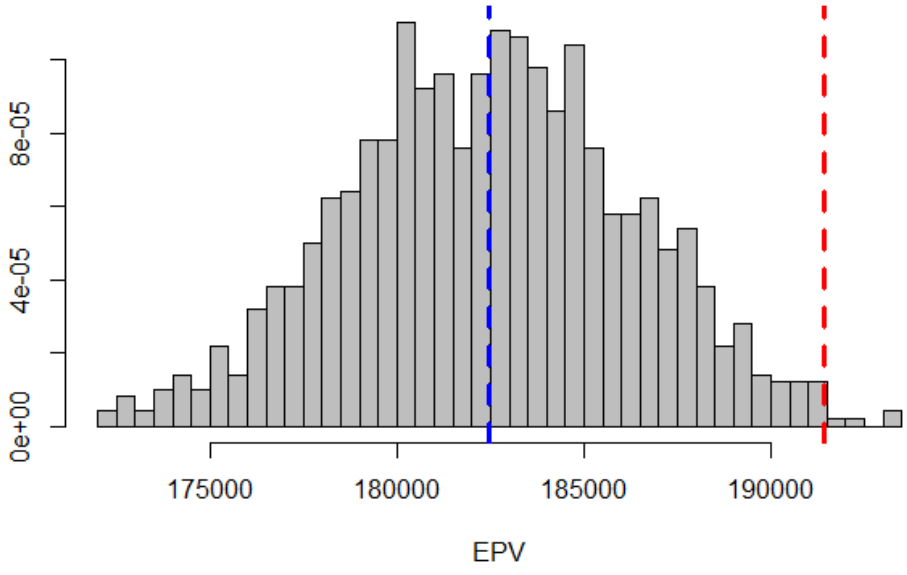


Figure 19: A histogram displaying the distribution of the EPV of benefits obtained from Model D at a 0% interest rate. The mean (blue dotted line) and 99.5% quantile (red dotted line) are also shown.

The blue dotted line indicates the mean of the distribution and the red dotted line indicates the 99.5% quantile, in this case, this value is €191429.90. What this means is that the probability of the EPV of benefits being higher than €191429.90 is 0.5%, according to Model D. So in this case the longevity risk associated with this annuity is

$$\frac{191429.9}{182399.7} \approx 1.04951$$

Meaning that if we consider 99.5% of possible outcomes then the EPV of benefits, per policy, that is experienced could be up to 5% larger than we anticipate. When dealing with thousands of policies, this poses a serious financial risk. This example also shows why longevity risk is systematic and cannot be diversified away, if mortality rates were to improve at a certain rate such that the average present value of benefits was, say, 2% higher than what was expected then it would not matter how many policies were sold as the cost to the insurer would be higher **per policy** and so the more policies that are sold, the higher the cost of longevity risk becomes.

Deciding upon a method to manage longevity risk is a difficult task for insurers. Traditional methods include product design, underwriting, natural hedging and reinsurance. One of the more popular methods of longevity risk management is through longevity bonds which pay a coupon that is linked to the survivorship of a cohort (in our example this would be 65-year-old males), although there is no clear consensus on which method is most effective. Comparing and recommending different methods of longevity risk management is beyond the scope of this project but for more detail into these methods see references [1]; [7]; [8]; [14].

Tabulated below are the mean and longevity risk associated with the forecast distribution of each model with respect to options 1 2 of the SIPP for various interest rates.

Model D						
	EPV of Benefits			Longevity Risk		
	i=0%	i=0.5%	i=2%	i=0%	i=0.5%	i=2%
Option 1	€182399.70	€171856.14	€144846.40	1.04920	1.04729	1.04189
Option 2	€162187.80	€145552.20	€106263.60	1.10747	1.10362	1.09281

Model E						
	EPV of Benefits			Longevity Risk		
	i=0%	i=0.5%	i=2%	i=0%	i=0.5%	i=2%
Option 1	€185114.30	€174536.27	€147121.90	1.07162	1.06848	1.05941
Option 2	€193253.10	€172702.90	€124484.90	1.0795	1.07623	1.06692

Model F						
	EPV of Benefits			Longevity Risk		
	i=0%	i=0.5%	i=2%	i=0%	i=0.5%	i=2%
Option 1	€193749.10	€182120.20	€152540.30	1.03727	1.03581	1.03166
Option 2	€193018.10	€172522.30	€123970.10	1.05648	1.05479	1.04995

Looking at these tables we can identify two trends in longevity risk common to all three models. The first is that option 2 carries more risk than option 1. This is to be expected as, in general, the younger the cohort the less certainty we can have regarding their future mortality rates and so there is a higher probability of experiencing mortality rates that deviate from what we expect. Interestingly, the sensitivity of longevity risk to changes in cohort age seems to vary between models. Another trend that can be seen from the tables is that longevity risk decreases as interest rates increase. This is due to the fact that as interest rates increase the present value of future benefits decreases so there is less real cost to the insurer if the survival time of the cohort is longer than the estimated life expectancy.

Given this information, once a mortality model is chosen, the price of either product can be adjusted to a desirable level and the longevity risk associated with this price can be calculated. For example, if an insurer were to assume the projections of model D and an

interest rate of 0.5% then the price of option 1 could be adjusted to, say, €165,000 as to only partially account for mortality improvement and increase competitiveness:

$$LR^* = 1.04729 \times \frac{171856.14}{165000} = 1.09081$$

Increasing longevity risk by about 4.5%. Alternatively, the price could be raised to a level such that longevity risk is negligible, say, €179000:

$$LR^* = 1.04729 \times \frac{171856.14}{179000} = 1.00549$$

This effectively mutes the risk but this price might be deemed non-competitive and raise more issues than it solves. Overall we can see how this framework might allow us to find a balance between price and longevity risk.

6 Conclusions

We have attempted to explain mortality improvements for Irish males aged 50-89 using six different mortality models in order to assess the longevity risk associated with the SIPP annuity product. With regards to model selection, what we've shown is that models that take into account period and cohort effects perform significantly better at explaining trends in our data and providing reliable forecasts than models that only consider the effect of age on mortality rates. Of the models that include both period and cohort effects each have different strengths. For example, the generalised Lee-Carter model, Model F, allows for greater flexibility in the period effects, $\beta_x^{(i)}$, while the models from the CBD family, models D and E, impose smoothness in the age effect as an assumption but allows for richer period effects. To some extent, it is up to the reader to decide what weights to place on different selection criteria.

If the reader takes into account the robustness of the parameter estimates, then Model D is preferred. This model fits our data set well and the stability of the parameter estimates over time enables one to place some degree of trust in its projections of mortality rates. The lack of robustness in the other models means that we cannot wholly rely on projections produced by them.

Model D shows us that:

- Mortality rates have been improving over time at all ages: the 'intercept' period term $\kappa_t^{(1)}$ has been declining over time so that the upward-sloping plot of the logit of mortality rates against age has been shifting downwards over time (See Figure 12 (a)).
- These improvements have been greater at lower ages than at higher ages: the 'slope' period term $\kappa_t^{(2)}$ has been increasing over time so that the plot of the logit of mortality rates against age has been steepening as it shifts downwards over time (See Figure 12 (a)). This feature has also been noted by Cairns et al [4].
- The changes over time in $\kappa_t^{(1)}$ have been approximately linear but the estimates for $\kappa_t^{(2)}$ display slightly different behaviour, decreasing in the initial years of the data period but beginning to increase from around 1985 onwards.
- The cohort effect, $\gamma_{t-x}^{(3)}$, plays a significant role in understanding the trend mortality improvement over time.

Other models also display these features.

We have also seen the benefits of projecting using stochastic mortality models in terms of estimating longevity risk and can conclude that taking a stochastic rather than a deterministic approach is more appropriate when modelling mortality as it allows us to obtain a forecast distribution rather than a single forecast.

With regards to the longevity risk associated with life annuities, our analysis provides a rough view of the longevity risk associated with a specific annuity product and shows that younger cohorts carry more risk and that this risk diminishes for higher interest rates. We have also shown why longevity risk is systematic and cannot be dealt with by simply selling a higher number of policies and provided a framework under which insurers can adjust prices to account for mortality improvement while maintaining a healthy balance between competitiveness and risk. These are important results to consider when designing and pricing survival linked financial products. All this being said, it is important to remember that there is still a lot of uncertainty regarding model selection and optimal methods of longevity risk management which makes quantifying and accounting for this risk difficult for insurers. This uncertainty generates a great need for further research.

7 Appendix

For the sake of clarity about how the values for longevity risk in section 6 were calculated the R code is included with step by step commentary. Here we calculate longevity risk for option 1 of the SIPP using model D.

```
library(StMoMo)

###Choose number of simulations
nsim=1000

###perform simulations
sim.d1 = simulate(model.d,nsim=nsim,h=35)

###Convert death rates to mortality rates
qx.rates.d1 = matrix(1-exp(-sim.d1$rates),nrow=40,ncol=35*nsim)

###Obtain survival rates
px.rates.d1 = matrix(1,nrow=40,ncol=35*nsim) - qx.rates.d1

###Convert mortality rates to deferred mortality rates
kqx.rates.d1 = matrix(NA,nrow=40,ncol=35*nsim)
for(i in 1:nsim){
 sim.mat = qx.rates.d1[,(1+35*(i-1)):(35+35*(i-1))]
  for(j in 1:35){
   sim.vec=sim.mat[,j]
    for(k in 2:40){
     kqx.rates.d1[1,(j + 35*(i-1))] = sim.vec[1]
     kqx.rates.d1[k,(j + 35*(i-1))] = sim.vec[k]*prod(px.rates.d1[(1:k-1),(j + 35*(i-1))])

  }
 }
}


###set interest rate
int.rate.d1 = 0.02

###Obtain EPV of benefits
benefit.d1 = numeric(26)
for(i in 2:26){
  benefit.d1[1]=0
  benefit.d1[i] = 10000*(1+int.rate.d1)^(-i)
}

total.benefit.d1 = numeric(26)
for(i in 1:26){
 total.benefit.d1[i]=sum(benefit.d1[1:i])
}


###Obtain the necessary deferred mortality rates from each simulation
kqx.price.d1 = matrix(NA,nrow=26,ncol=nsim)
for(i in 1:nsim){
 block = kqx.rates.d1[,(1+35*(i-1)):(35+35*(i-1))]
  for(j in 1:25){
    kqx.price.d1[j,i]=block[(15+j),j]
     for(k in 1:nsim){
       kqx.price.d1[26,k] = 1 - sum(kqx.price.d1[1:25,k])
     }
  }

}

###Calculate the EPV of benefits for each simulation
ev.d1 = matrix(NA,nrow=26,ncol=nsim)
for(i in 1:nsim){
  ev.d1[,i]=kqx.price.d1[,i]*total.benefit.d1
}

price.d1 = numeric(nsim)
for(i in 1:nsim){
 price.d1[i]=sum(ev.d1[,i])
}


###Visualize the distribution of EPV of benefits
par(mfrow=c(1,1))
hist(price.d1,breaks=40,col=8,freq=F,main="",ylab = "",xlab="EPV")
abline(v=mean(price.d1),col=4,lty=2,lwd=3)
abline(v=quantile(price.d1,0.995),col=2,lty=2,lwd=3)

###Calculate Longevity Risk
LRD1 = quantile(price.d1,0.995)/mean(price.d1)
LRD1
```

References

[1] David Blake et al. "Longevity bonds: financial engineering, valuation, and hedging". In: *Journal of Risk and Insurance* 73.4 (2006), pp. 647–672.

[2] Natacha Brouhns, Michel Denuit, and Jeroen K Vermunt. "A Poisson log-bilinear regression approach to the construction of projected lifetables". In: *Insurance: Mathematics and economics* 31.3 (2002), pp. 373–393.

[3] Andrew JG Cairns, David Blake, and Kevin Dowd. "A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration". In: *Journal of Risk and Insurance* 73.4 (2006), pp. 687–718.

[4] Andrew JG Cairns et al. "A quantitative comparison of stochastic mortality models using data from England and Wales and the United States". In: *North American Actuarial Journal* 13.1 (2009), pp. 1–35.

[5] Andrew JG Cairns et al. "Mortality density forecasts: An analysis of six stochastic mortality models". In: *Insurance: Mathematics and Economics* 48.3 (2011), pp. 355–367.

[6] Maria Francesca Carfora, Luisa Cutillo, and Albina Orlando. "A quantitative comparison of stochastic mortality models on Italian population data". In: *Computational Statistics & Data Analysis* 112 (2017), pp. 198–214.

[7] Thomas Crawford, Richard de Haan, and Chad Runchey. "Longevity risk quantification and management: a review of relevant literature". In: *Society of Actuaries* 73 (2008).

[8] Jiajia Cui. "Longevity risk pricing". In: *Available at SSRN 972038* (2007).

[9] Iain D Currie. "On fitting generalized linear and non-linear models of mortality". In: *Scandinavian Actuarial Journal* 2016.4 (2016), pp. 356–383.

[10] Benjamin Gompertz. "XXIV. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. FRS &c". In: *Philosophical transactions of the Royal Society of London* 115 (1825), pp. 513–583.

[11] Steven Haberman and Arthur Renshaw. "A comparative study of parametric mortality projection models". In: *Insurance: Mathematics and Economics* 48.1 (2011), pp. 35–55.

[12] Ronald D Lee and Lawrence R Carter. "Modeling and forecasting US mortality". In: *Journal of the American statistical association* 87.419 (1992), pp. 659–671.

[13] Carlo Maccheroni and Samuel Nocito. "Backtesting the lee–carter and the cairns–blake–dowd stochastic mortality models on italian death rates". In: *Risks* 5.3 (2017), p. 34.

[14] Andrew Ngai and Michael Sherris. "Longevity risk management for life and variable annuities: The effectiveness of static hedging using longevity bonds and derivatives". In: *Insurance: Mathematics and Economics* 49.1 (2011), pp. 100–114.

[15] Arthur E Renshaw and Steven Haberman. "A cohort-based extension to the Lee–Carter model for mortality reduction factors". In: *Insurance: Mathematics and economics* 38.3 (2006), pp. 556–570.

[16] Arthur E Renshaw and Steven Haberman. "Lee–Carter mortality forecasting with age-specific enhancement". In: *Insurance: Mathematics and Economics* 33.2 (2003), pp. 255–272.

[17] *The World Bank life expectancy data*. URL: `https://data.worldbank.org/indicator/SP.DYN.LE00.IN?locations=IE`. 2020.

[18] Andrés Villegas, Vladimir K Kaishev, and Pietro Millossovich. "StMoMo: An R package for stochastic mortality modelling". In: *7th Australasian Actuarial Education and Research Symposium*. 2015.

[19] Richard C Willets. "The cohort effect: insights and explanations". In: *British Actuarial Journal* (2004), pp. 833–898.