ARC Centre of Excellence in Population Ageing Research

Working Paper 2012/11

Rethinking Age-Period-Cohort Mortality Trend Models

Daniel H. Alai and Michael Sherris

Daniel H. Alai is an Associate Investigator at the Centre of Excellence in Population Ageing Research (CEPAR), Australian School of Business, UNSW, Sydney 2052 Australia; E-mail: daniel.alai@unsw.edu.au

Michael Sherris is Professor in the School of Risk and Actuarial Studies and Chief Investigator at the Centre of Excellence in Population Ageing Research (CEPAR), Australian School of Business at the University of New South Wales; E-mail: m.sherris@unsw.edu.au

Rethinking Age-Period-Cohort Mortality Trend Models

Daniel H. Alai[1] Michael Sherris[2]
CEPAR, Risk and Actuarial Studies, Australian School of Business
UNSW, Sydney NSW 2052, Australia

Abstract

Longevity risk arising from uncertain mortality improvement is one of the major risks facing annuity providers and pension funds. In this paper we show how applying trend models from non-life claims reserving to age-period-cohort mortality trends provides new insight in estimating mortality improvement and quantifying its uncertainty. Age, period, and cohort trends are modelled with distinct effects for each age, calendar year, and birth year in a generalized linear models framework. The effects are distinct in the sense that they are not conjoined with age coefficients, borrowing from regression terminology, we denote them as main effects. Mortality models in this framework for age-period, age-cohort, and age-period-cohort effects are assessed using national population mortality data from Norway and Australia to show the relative significance of cohort effects as compared to period effects. Results are compared with the traditional Lee-Carter model. The bilinear period effect in the Lee-Carter model is shown to resemble a main cohort effect in these trend models. However the approach avoids the limitations of the Lee-Carter model when forecasting with the age-cohort trend model.

Keywords: Mortality Modelling, Age-Period-Cohort Models, Generalized Linear Models, Lee-Carter Models
JEL classifications: G22, G23, C51, C18

[1] daniel.alai@unsw.edu.au
[2] m.sherris@unsw.edu.au

1 Introduction

Quantifying trends and uncertainty in risks is one of the most important factors for assessing the financial significance of future liabilities for insurance companies. In non-life insurance, outstanding liabilities for existing policies are usually the largest item on the liability side of the balance sheet. Similarly in life insurance and pension funds, future benefit payments, contingent upon mortality experience, represent the largest liability. The uncertainty in future mortality improvement has a significant impact on the capital of life insurers under recent risk-based capital requirements and on the liabilities of defined benefit pension schemes under market value based accounting requirements.

Non-life and life insurance risk modelling traditionally used different modelling approaches to quantify trends and uncertainty given the differing nature of the risks involved. The underlying risks of non-life insurance and life insurance companies are driven by different risk factors. The former is concerned with uncertain trends in future claims from different accident years, and the latter with uncertain trends in claims from improvements in underlying mortality rates. Although the nature of the two types of liabilities are different, the underlying approach to modelling trends and uncertainty has been shown to be similar when expressed in a statistical framework. This has been recognised and applied in particular by Venter (2008) and Gluck and Venter (2009). The approach developed uses generalized linear models (GLM), a framework previously used for trend modelling by Haberman and Renshaw (1996) and Renshaw *et al.* (1996). The aim of this paper is to exploit the insight that non-life trend models provide in mortality risks as well as forecasting not provided by currently used models. This paper builds a bridge between the two forms of insurance modelling.

There are many different approaches to modelling trends and uncertainty in mortality. Most stochastic mortality models include period effects across ages and increasingly cohort effects have been recognised resulting in age-period-cohort models. A seminal model for studying mortality trends and forecasting was developed by Lee and Carter

(1992). Subsequent GLM formulations were developed by Brouhns *et al.* (2002) and Renshaw and Haberman (2003a). The Lee-Carter family of models has also been extended by many authors to capture additional features in the data. The most commonly included feature is a cohort effect although there are many models accounting for different age, cohort and period effects. These are summarized and compared in Cairns *et al.* (2009) and Haberman and Renshaw (2011). Cohort effects are usually added after improvement trends are modelled using period effects. However, since these improvement trends clearly differ across ages, the period effects are modelled with a bilinear term in age and time. The introduction of these cohort effects confounds the period effects in these models. The precise definition of a cohort effect is a topic of discussion (e.g. Willets (1999, 2004) and Murphy (2009)). One of the aims of this paper is to clearly show the importance of main cohort effects and to demonstrate the relationship between these and bilinear terms where period effects vary by age. Bilinear effects correspond to interaction effects using regression terminology.

The model includes age, period, and cohort effects without *interacting* age with either period or cohort effects. In the trend modelling framework we formulate and assess age-period, age-cohort, and age-period-cohort models and apply them to national population data. The age-period and age-cohort models considered have been largely overlooked in the mortality modelling literature. In the GLM framework, the age, period, and cohort effects and confidence intervals are readily estimated.

We find that the age-cohort model performs well in capturing trends in mortality data and we show empirically that it performs much better than an age-period model. We also show that the impact of a bilinear period effect resembles that of a main cohort effect. Consequently, the traditional Lee-Carter model has features similar to our age-cohort model rather than our age-period model. Although we show empirically that main cohort effects capture trends more effectively than main period effects, this does not imply that bilinear cohort effects are to be preferred over bilinear period effects. A comparison between bilinear effects is discussed in Renshaw and Haberman (2006). The advantages of main effects are that they are more intuitive to model and interpret and are very useful when forecasting.

Finally, we demonstrate the benefits of the age-cohort model by forecasting cohort life expectancies. In contrast to period life expectancies, cohort life expectancies have practical and intuitive interpretations. The trend models allow these to be estimated naturally and efficiently compared to the traditional period models, even in the case where the traditional models are modified to include cohort effects.

Organization of the paper. Section 2 introduces the notation and discusses the structure of the data in *triangle* form. We outline rudimentary models for insurance data and their connection to GLMs in Section 3. In Section 4 we fit and assess the age, period, and cohort mortality trends for Norway and Australia using GLMs. The impact of data period on estimating the cohort effects and improving model estimation is studied in Section 5. Section 6 is devoted to fitting the Lee-Carter model and showing the similarities between the bilinear period and main cohort effects. In Section 7 we apply the trend models to forecast cohort life expectancy as well as quantify its uncertainty to show the practical benefits of the approach. Section 8 concludes the paper.

2 Data and Notation

We apply the structure traditionally used in non-life insurance in a life insurance context. Let row index i represent the year of birth (or cohort) and column index j, the age at death. The resulting diagonal, $k = i + j$, represents the calendar year of death (or period). The year of birth and age at death are equivalent to the accident year and development year, respectively, in the non-life setting. For a consistent presentation of the data, we set the lowest value of birth year equal to zero. Mortality data is typically recorded by calendar year of death; as in the left table in Figure 1. A consequence of setting the lowest value of birth year equal to zero is that the lowest value of calendar year of death takes value J, the maximum age.

Let $D_{i,j}$ denote the number of people born in year i having died at age j, and hence in calendar year k. Furthermore, let $E_{i,j}$ be a deterministic exposure measure, for example the number of people born in year i having attained age j.

At time I, which represents the most recent calendar year (and birth year), we have observed deaths for $\mathcal{D}_I = \{D_{i,j},\ J \leq k \leq I\}$. In order to apply the non-life trend models to mortality data we transform it to *triangle*, or in most cases trapezoid, form by substituting birth year for calendar year as one of the axes. This is shown in the right table in Figure 1. Both tables in Figure 1 distinctly present the same data; the formulation on the right is similar to the way non-life insurance data is presented. Since practically all recent stochastic mortality models are based on period effects using life table data, this transformation is a distinct and significant shift in perspective that we show to enhance our understanding of mortality trends. It is also a more natural way to consider trends when it is the development of mortality by cohort that is of most importance.

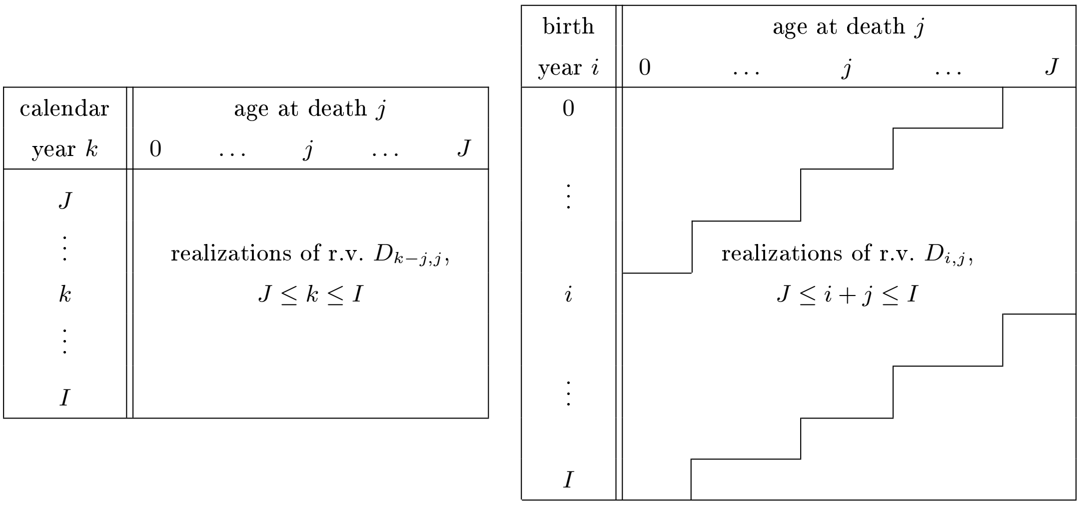

Figure 1: Transforming life insurance data, the mortality triangle.

Remark 2.1 In the formulation of the mortality triangle, the interpretation of one of the indices is not precise. The age at death, usually last birthday, is precise but the calendar year of death is not. An age at death of value j could have occurred in years $i + j - 1$ to $i + j + 1$.

3 Generalized Linear Models in Insurance

In this section we provide a brief introduction to GLMs and outline the well known Lee-Carter model used for mortality data. For an excellent overview of GLMs, see McCullagh and Nelder (1989), and for its applications to insurance see Haberman and Renshaw (1996) and England and Verrall (2002).

3.1 Generalized Linear Models

GLMs encompass a broad class of regression models including standard linear regression with a normally distributed response, logistic regression and log-linear Poisson regression. We study the Poisson distribution with mean μ; a member of the exponential family with $\theta = \ln(\mu)$, $\phi = 1$ (although we relax this assumption below to allow for overdisperion), and $b(\theta) = \exp(\theta)$. We make use of the canonical link function for the Poisson distribution, thereby simplifying the GLM framework to Poisson regression. The log-link function implies that the impact of covariates on the response is multiplicative in nature (additive on the log-response). With the exception of ϕ, likelihood methods are used to estimate the model parameters.

Overdispersion

A distribution where $\phi = 1$, such as the Poisson, is referred to as *overdispersed* when ϕ is not *restricted* to unity. That is, when the variability of the data is greater than that anticipated by the distribution, it is referred to as overdispersed. Overdispersion can be a sign of incorrect model assumptions. However, where the systematic component and the mean-variance relationship of the data are appropriate, it is usual to allow an estimate of ϕ to differ from unity.

Quasi-likelihood is used for regression parameter estimation since the overdispersed distribution is, strictly speaking, no longer a valid probability distribution. The dispersion parameter is typically estimated using Pearson residuals, given by

$$\widehat{\phi} = \frac{1}{d} \sum_{i+j \leq I} \frac{(X_{i,j} - \widehat{\mu}_{i,j})^2}{V(\widehat{\mu}_{i,j})},$$

where d is the degrees of freedom of the model (number of observations less number of unknown parameters) and $V(\cdot)$ is the variance function of the underlying distribution. In the case of the Poisson distribution $V(\mu) = \mu$.

Allowing for overdispersion provides added flexibility but is only relevant for estimating second moments. The dispersion parameter plays no role in the estimation of regression coefficients. However, it directly impacts the standard deviation of the estimators in the form of a scalar multiplier.

3.2 The Lee-Carter Model

One of the most cited models for mortality modelling and forecasting was proposed by Lee and Carter (1992). They observed that mortality improvements varied considerably among different age groups. They intended to develop a parsimonious model that would specify a time-varying index to account for these mortality improvements. Their model has received a lot of attention in the literature and has been used as the foundation of many subsequent models; Lee (2000), Cairns *et al.* (2006, 2009) and Renshaw and Haberman (2003b, 2006).

The mortality rate is defined as $m_{i,j} = D_{i,j}/E_{i,j}$, where, depending on the measure of exposure, m is either a crude or central mortality rate. Ignoring the error term, the Lee-Carter model specifies the following structure for the log mortality rate:

$$\ln(m_{i,j}) = \alpha_{1,j} + \alpha_{2,j}\kappa_k,$$

Due to the multiplicative, or bilinear, structure of the second term on the right-hand side in the above equation, the parameters $\alpha_{2,j}$ and κ_k can only be determined up to a constant factor. To ensure identifiability, the constraints $\sum_j \alpha_{2,j} = 1$ and $\sum_k \kappa_k = 0$ are imposed. These constraints imply that $\alpha_{1,j}$ can be taken as the average, over time, of the $\ln(m_{i,j})$. This is typically done and establishes a link to GLM that we investigate below.

The α parameters can be viewed as regression parameters with respect to age. The parameter κ is the time-varying parameter discussed above, known as the mortality index. If κ had been a known value, then the structure would be linear with respect to the parameters and a straightforward application of GLM could be used to fit the model.

3.3 The Link between Lee-Carter and GLM

The Lee-Carter model is, strictly speaking, not a GLM. However, if we set

$$\alpha_{1,j} = \frac{1}{I-J+1} \sum_{k=J}^{I} \ln(m_{k-j,j}),$$

as suggested in Renshaw and Haberman (2000), we can then reformulate the model as

$$\ln(m^*_{i,j}) = \alpha_{2,j}\kappa_k,$$

where,

$$m^*_{i,j} = m_{i,j} \left(\prod_{k=J}^{I} m_{k-j,j} \right)^{\frac{-1}{I-J+1}}.$$

The $m^*_{i,j}$ can be interpreted as a normalization of the mortality rates with respect to the geometric mean for each age j; normalized in the sense that the geometric average of $m^*_{i,j}$ with respect to i, for each age j, is equal to one. In addition define $D^*_{i,j} = m^*_{i,j} E_{i,j}$ to be the resulting *normalized* deaths. The motivation of this normalization is to model the response in a GLM framework. Any distribution from the exponential family can be assumed in conjunction with the log-log link function. This link function implies that the impact of covariates on both the response as well as the log-response are multiplicative in nature (additive on the log-log-response). The error structure is likewise affected. The normalized deaths are not used in our approach.

4 Mortality Trends and GLM models

An important application of mortality modelling is to estimate trends in order to understand potential future mortality improvements. In this section, we study age-period, age-cohort, and age-period-cohort models and estimate the models using national population mortality data from Norway and Australia. We use Norway because it has a long history of mortality data and Australia because this provides a good contrast, being a younger country with a diverse population. Norway also has a relatively smaller population compared to Australia. The data was obtained from the Human Mortality Database (2011). The Norwegian and Australian population mortality data dates back to calendar years 1846 and 1922, respectively. Although the reliability and relevance of the most distant data can be called into question, we found no issues when incorporating the most distant calendar year data. That is, the inclusion of the distant calendar year data did not noticably impact the estimation or uncertainty of the age trend or the recent calendar year trend. However, we did find issues when incorporating the most distant cohort year data. We comment on these issues below.

The interpretation of main effects in these trend model is important. Compared to bilinear effects, they are very easy to explain. A main period effect is one that impacts all ages to the same extent for the period under consideration. Examples of main period effects could include natural disasters and health pandemics that impact all ages. Main period effects are difficult to support in practice, even the 1918 influenza pandemic was not a true main period effect as age-groups were not uniformly affected.

A main cohort effect impacts all members of a particular cohort from birth until death. The main cohort effect captures the mortality improvement experienced by that particular cohort. This definition differs slightly from the one provided in Willets (1999) and Murphy (2009), where cohort effects are essentially described as *delayed* period effects. Main cohort effects can also be updated over time for incomplete cohorts but we do not consider this at present.

4.1 Age-Period Models

We commence with a model that includes distinct age and period effects. This model is a special case of the Lee-Carter model.

Model Assumptions 4.1 (Log Mortality Model: Age-Period)

- Deaths $D_{i,j}$ are independent (overdispersed) Poisson distributed given known exposure units $E_{i,j}$.

- The regression formula is given by

$$\eta_{i,j} = \ln E_{i,j} + \beta_0 + \beta_{2,j} + \beta_{3,i+j},$$

  where $\beta_{2,0} = \beta_{3,0} = 0$.

- The link function is given by $g(\mu) = \ln(\mu)$.

Figure 2 shows the estimated regression coefficients $\widehat{\beta}_{2,j}$, the age coefficients that describe the age trend, and $\widehat{\beta}_{3,i+j}$, the calendar year coefficients that describe the period trend. The dispersion parameter was estimated using Pearson residuals and was found to be 33.65 for Norway and 24.62 for Australia. The dispersion parameter estimate was used to obtain the confidence intervals shown in Figure 2. The 95% confidence intervals appear to be very tight, especially for the age coefficients.

There are clear trends in these two sets of coefficients. For both countries, the age effects are similar. The shape is well recognized and proportional to that of the average log mortality rates by age. This age pattern of mortality can be fitted using the eight-parameter model proposed by Heligman and Pollard (1980). The shape of the curve is more linear in the Australian data than in the Norwegian for ages beyond thirty. This suggests that Australian data conforms more to an exponential hazard assumption for older ages. As expected, both countries portray a decay of the age trend at the extremely old ages and exhibit high variability in this region due to small numbers of deaths and exposures in the data. The period trend is downward sloping showing the general mortality improvement in both countries, particularly since the 1970's in Australia, which agrees with the findings of Booth *et al.* (2006).

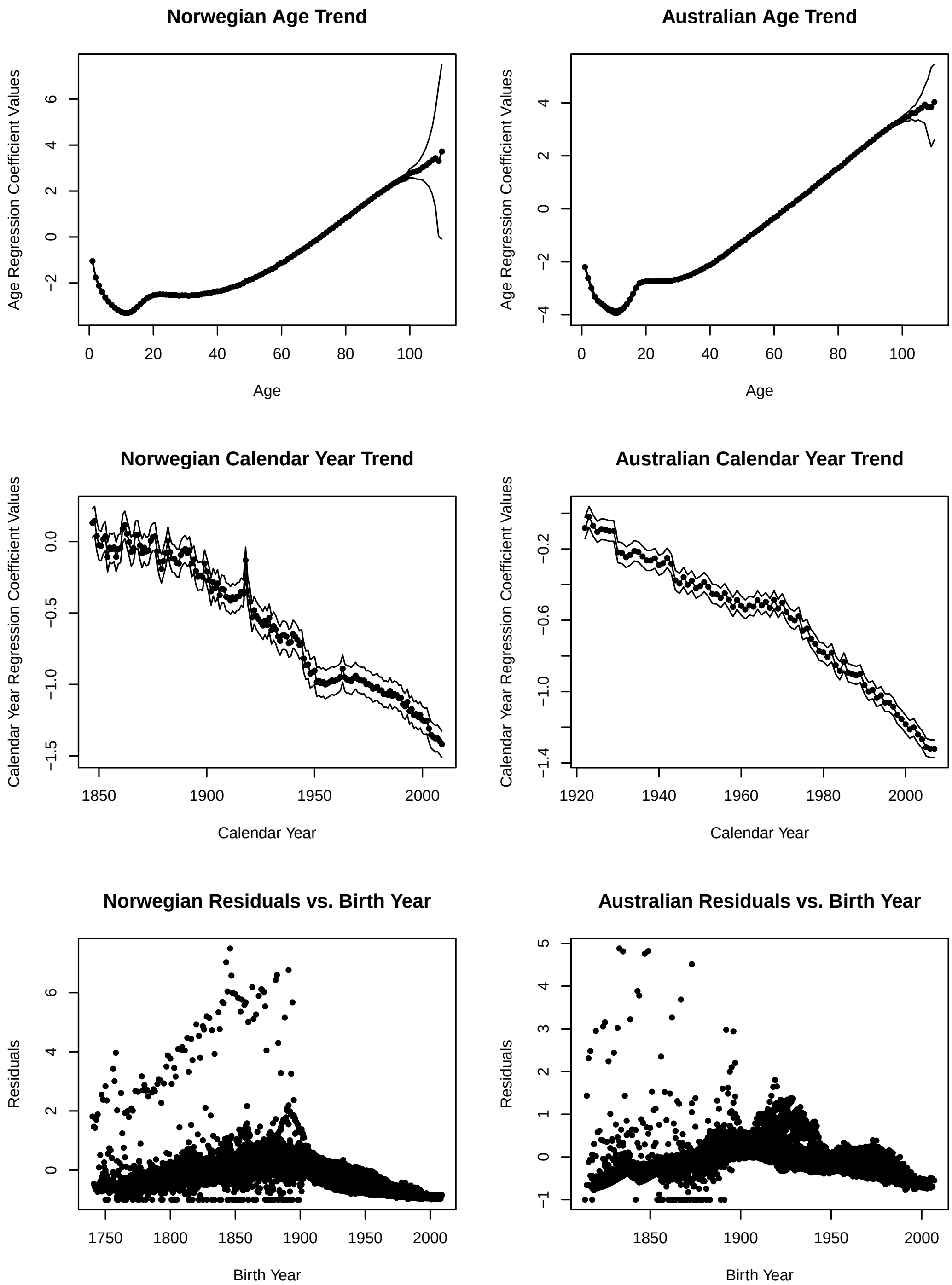

Figure 2: Norwegian and Australian age trends, calendar year trends, and residuals from fitted GLM Model 4.1.

Residuals are plotted versus birth year to ascertain whether there is a discernable pattern suggestive of a need for cohort effects. Figure 2 shows that the residuals are not well behaved. There is an interesting shift in behaviour, common to both countries, that occurs at the turn of the 20^{th} century, namely, that we no longer observe large residuals. This phenomenon is explained by studying the residual plot versus age (not included in the paper), which shows that *all* large residuals result from the model's inability to capture trends at the extremely old ages (centenarians). Since, as of 2008, there exist no centenarians born after 1908, we consequently do not observe large residuals for those later birth years.

Remarks 4.2:

- The residuals we present are the *working* residuals, those from the final iteration of the iteratively reweighted least squares fit. They are not for use in diagnostic checking but serve our purpose of identifying potentially omitted trends.
- The working residuals are not the residuals from a standard regression and are bounded from below by negative unity. This is a direct consequence of the fact that we have a non-negative response; observations of value zero yield working residuals of value negative unity.

4.2 Age-Cohort Models

The next model includes both age and birth year, or cohort, effects. As far as we can determine, it has not been considered in the actuarial mortality literature to date. We fit both Norwegian and Australian population mortality data dating back to calendar years 1846 and 1922, respectively. We omit the ten most distant years of cohort data, due to their instability, to ensuring reliable estimation. Omitting more than ten years of data leads to interesting results that we explore further in Section 5 below.

Model Assumptions 4.3 (Log Mortality Model: Age-Cohort)

- Deaths $D_{i,j}$ are independent (overdispersed) Poisson distributed given known exposure units $E_{i,j}$.
- The regression formula is given by

$$\eta_{i,j} = \ln E_{i,j} + \beta_0 + \beta_{1,i} + \beta_{2,j},$$

  where $\beta_{1,0} = \beta_{2,0} = 0$.
- The link function is given by $g(\mu) = \ln(\mu)$.

Remark 4.4 This model conforms exactly to a GLM version of the classic chain ladder method used in non-life insurance modelling.

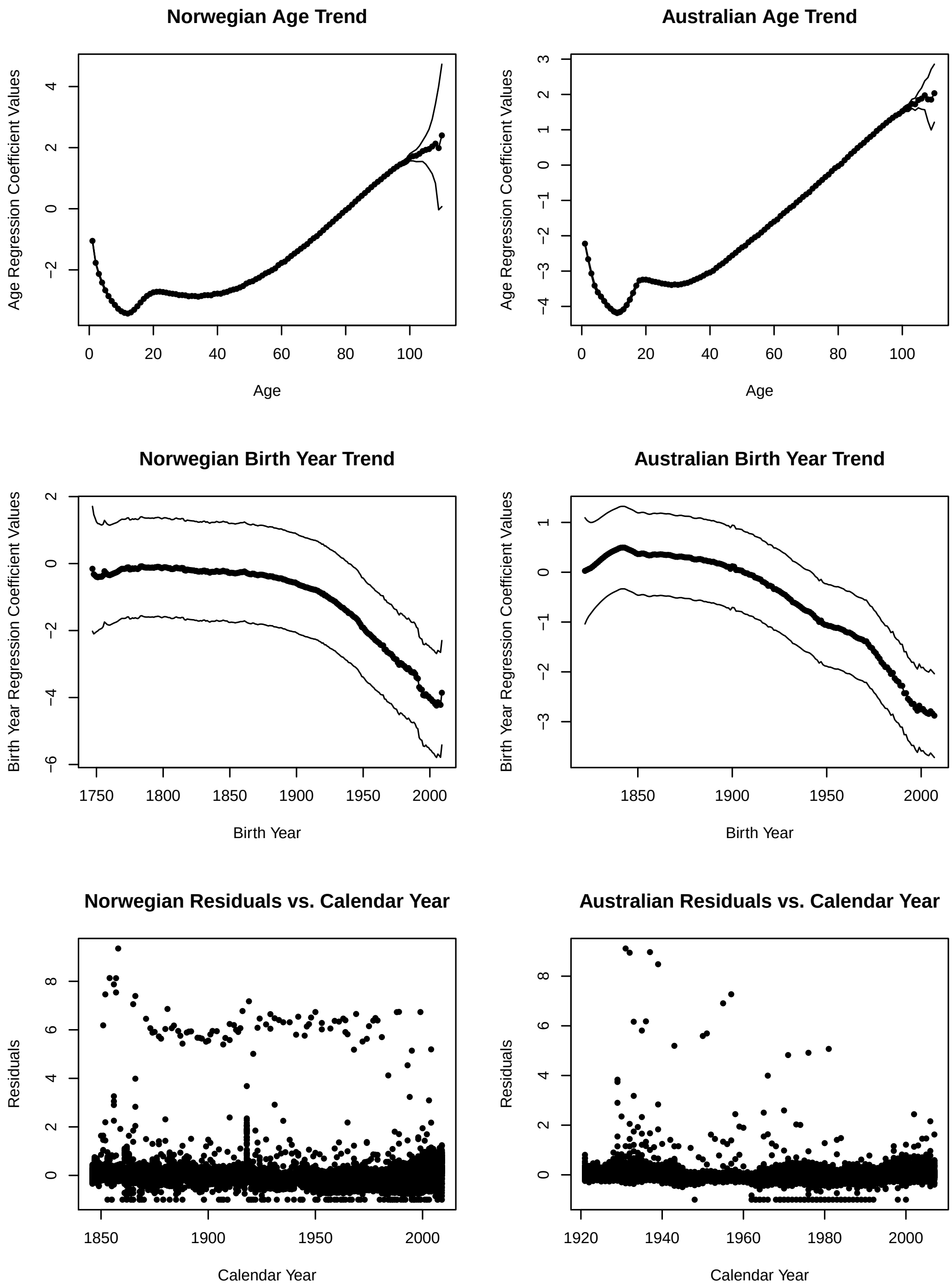

Figure 3: Norwegian and Australian age trends, birth year trends, and residuals from fitted GLM Model 4.3.

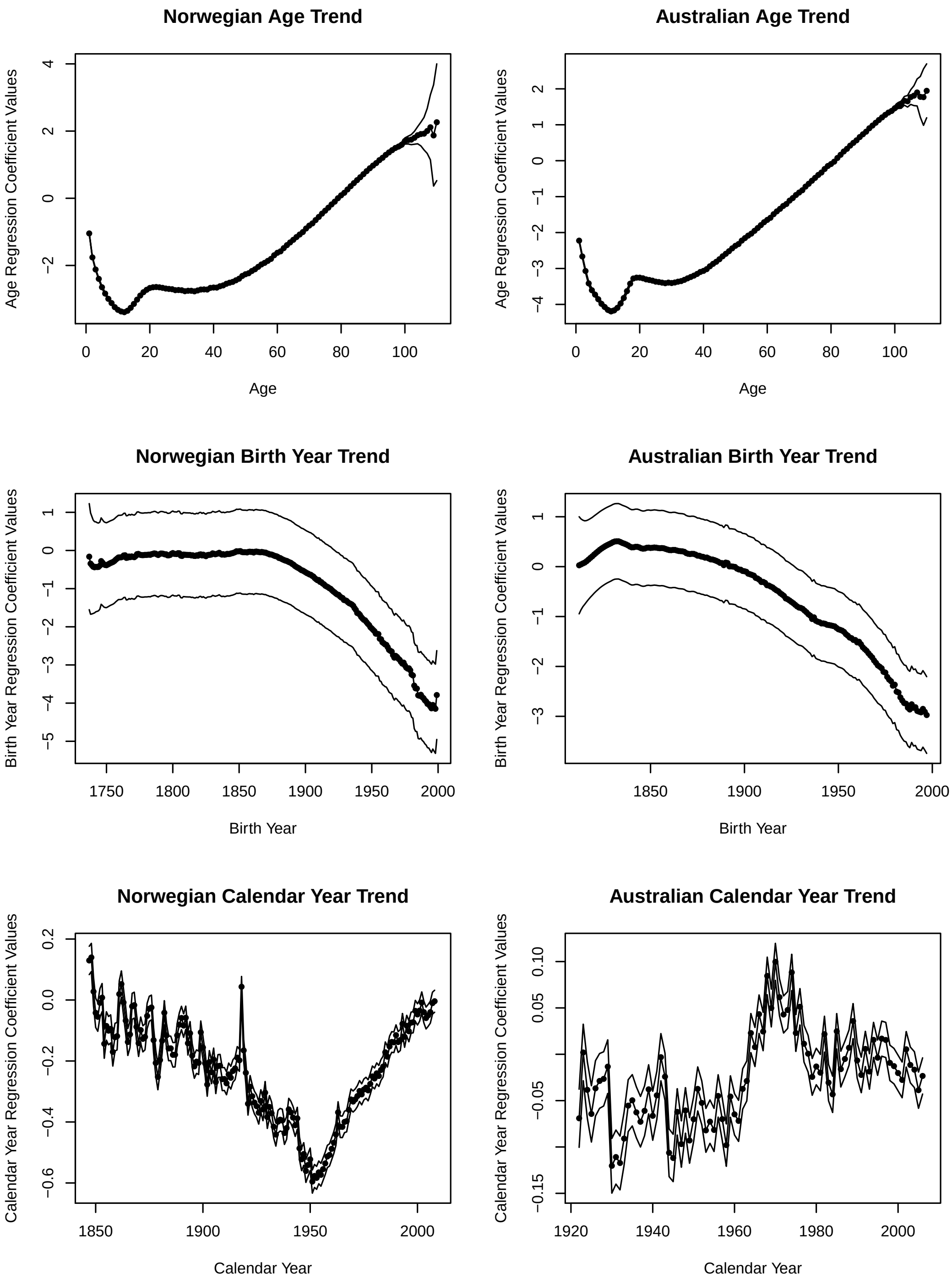

Figure 4: Norwegian and Australian age, birth and calendar year trends from fitted GLM Model 4.5.

Figure 3 shows the estimated regression coefficients $\widehat{\beta}_{2,j}$, the age trend, and $\widehat{\beta}_{1,i}$, the birth year coefficients that describe the cohort trend. The dispersion parameter estimates were found to be 12.66 for Norway and 8.10 for Australia; both values substantially lower than in the age-period model.

The age trend has the same shape as fitted by the age-period model. Unlike the period trend in the age-period model, the cohort trend is smooth. The standard deviations of the birth year coefficients are relatively large. A lack of data and an increased number of parameters jointly contribute to this phenomenon. We only have data for the very old in the most distant cohorts and data for the very young in the most recent cohorts. Furthermore, we have replaced X calendar year coefficients with, in our case, $X + 100$ birth year coefficients. Note that the number of extra parameters is the maximum age ($J = 110$) minus the number of the omitted cohort years (10).

The residuals are plotted versus calendar year in order to analyze the impact of omitting period effects. The results are seen to be much improved over the age-period model, especially for the Australian data, and do not suggest that calendar year coefficients are required for each period. The large residuals are due to the model's inability to handle the centenarians rather than any misspecification with respect to either birth year or calendar year assumptions.

4.3 Age-Period-Cohort Models

Finally a model for age, birth year, and calendar year of death effects, usually referred in mortality modelling as an age-period-cohort model, is considered. This model has previously been considered in Currie (2006) and Renshaw and Haberman (2006). We fit both Norwegian and Australian population mortality data. Furthermore, as in the age-cohort model, we omit the ten most distant years of cohort data in order to ensure efficient estimation whilst retaining the vast majority of the data.

Model Assumptions 4.5 (Log Mortality Model: Age-Period-Cohort)

- Deaths $D_{i,j}$ are independent (overdispersed) Poisson distributed given known exposure units $E_{i,j}$.

- The regression formula is given by

$$\eta_{i,j} = \ln E_{i,j} + \beta_0 + \beta_{1,i} + \beta_{2,j} + \beta_{3,i+j},$$

  where $\beta_{1,0} = \beta_{2,0} = \beta_{3,0} = 0$.

- The link function is given by $g(\mu) = \ln(\mu)$.

Figure 4 shows the estimated regression coefficients $\widehat{\beta}_{2,j}$, the age trend, the $\widehat{\beta}_{1,i}$, the cohort trend, and the $\widehat{\beta}_{3,i+j}$, the period trend. The dispersion parameter estimates were found to be 7.09 for Norway and 6.76 for Australia; both values not significantly lower than in the age-cohort model.

The age and cohort trends in this *full* model are almost identical to that in the age-cohort model. The period trend, however, has changed drastically compared to the age-period model. The period trends for the two countries are not readily interpreted and are not comparable. The Australian period trend, especially, is very inconsistent. This is not surprising considering what was observed in the residuals of the age-period and age-cohort models. The residuals in the age-period model suggested factors were missing, those in the age-cohort model did not. This model could be seen as being overparameterized. Consequently, its results should be considered with some caution.

We have assessed the age-period and age-cohort models by studying their residuals with respect to the omitted trend. Likelihood based goodness-of-fit statistics are not available due to our allowance for overdispersion. Alternatives include quasi-likelihood and residual deviance statistics. We show residual deviance statistics in Table 1, where we observe that the age-cohort models have a much better fit than the age-period models. The age-period-cohort models have the lowest residual deviance, which is to be expected. As deviance statistics are only comparable on the same dataset, we note that the results of Table 1 are all based on data that omits the most distant ten cohorts.

	Norway		Australia	
	Residual Deviance	Number of Parameters	Residual Deviance	Number of Parameters
Intercept	11,733,398	1	21,911,465	1
Age-Period	663,279	274	224,996	197
Age-Cohort	219,706	374	73,409	297
Age-Period-Cohort	118,859	537	60,675	383

Table 1: Analysis of deviance.

5 The Impact of Cohort Data Retention

We modify the mortality triangle presented in Figure 1 in order to understand the impact of the most distant cohorts on parameter uncertainty. These cohorts represent the older ages in the most distant calendar years.

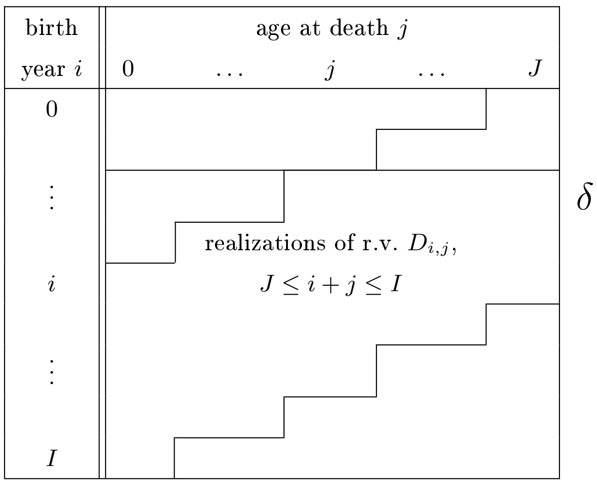

Figure 5: Data retention, the mortality triangle.

In Figure 5 a threshold level δ is included to represent data retention; data below the line is retained and used in the model. As seen in Section 4, for models that include cohort effects, we omit the first ten cohorts in order to obtain tractable results (i.e. $\delta = 10$). This is due to a significant lack of information on these most distant cohorts. In these cohorts, the only available information comes from the extremely old ages, where the deaths and exposure data is already scarce and unstable.

The retention level δ has a significant impact on the standard deviations (or confidence intervals) of parameter estimators, but does not significantly impact their respective expected values. There is a natural trade-off present. We reduce the uncertainty in the estimation of the retained cohort coefficients by excluding the most distant, and inherently unstable, cohorts. However, when we exclude these cohorts, we lose data for the older ages. As a result, the uncertainty in the estimation of the older age coefficients increases. Figure 6 shows this using the Norwegian dataset and the fit for the age-cohort model using $\delta = 50$ and $\delta = 110$.

The results show a significant improvement of the confidence intervals of the birth year coefficients and a near negligible increase in the confidence intervals of the older age coefficients. Furthermore, as we exclude the most distant cohorts, we no longer observe large residuals for corresponding lagged calendar years in the residual plot. Although it is discernable in Section 4, the elimination of these old age residuals leaves a clear picture of potentially missing effects. It is evident, for example, that the age-cohort model is unable to capture the influenza pandemic of 1918.

6 Mortality Modelling using Lee-Carter

In this section we fit Norwegian and Australian population mortality data using the Lee-Carter model. The Lee-Carter model was chosen as a basis for comparison due to its prominence and usage in the literature. Figure 7 shows the results of the Lee-Carter fit. Age and period effects are comparable to results from the fitted GLM models of Section 4.

In order to compare bilinear with main effects, we investigate the *implied* cohort effects present in the Lee-Carter model. The age-interaction parameter in the Lee-Carter model designates the extent to which, on average, a mortality improvement affects each age. Since a unique age in a specific calendar year is synonymous with a unique cohort, the combination of the mortality index with the age interaction, thus far

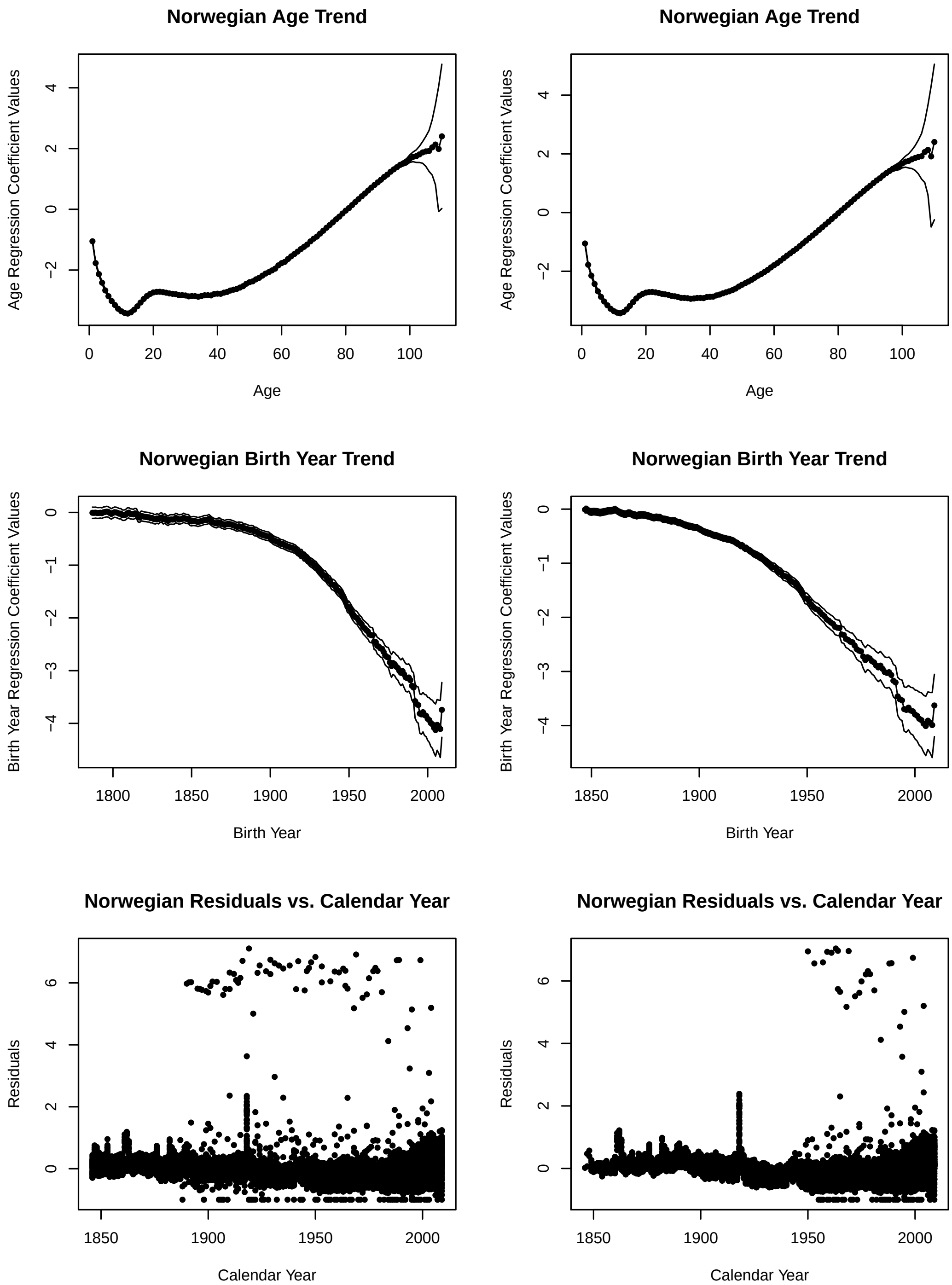

Figure 6: Norwegian age trend, birth year trend and residuals for $\delta = 50$ and $\delta = 110$.

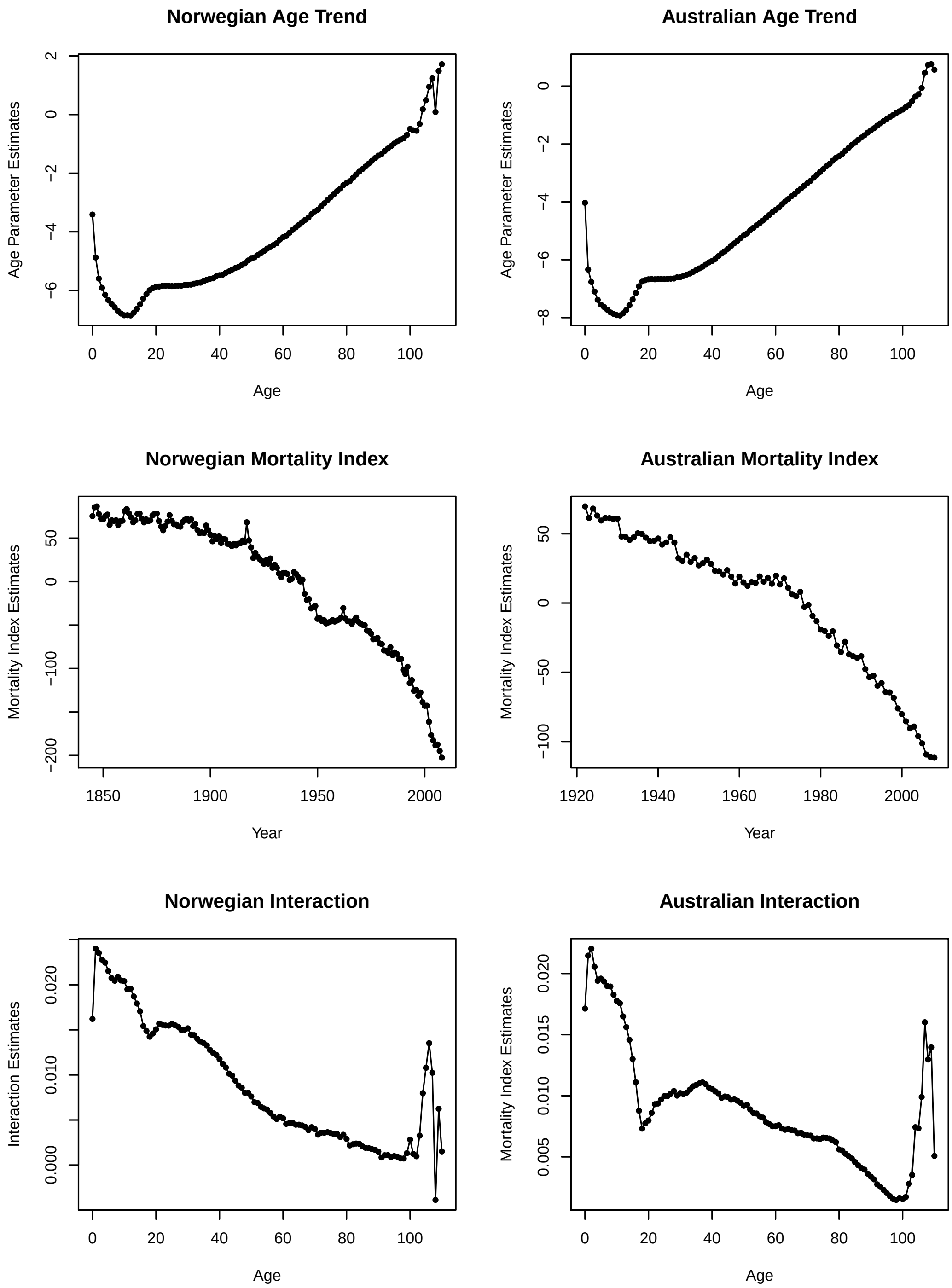

Figure 7: Norwegian and Australian estimated Lee-Carter model parameters.

referred to as a bilinear period effect, has the potential to capture the corresponding main cohort effect, and vice versa. This type of replacement was discussed in Cairns *et al.* (2011), where it was suggested that cohort effects could be omitted in lieu of well-chosen age and period effects. Our approach is to assess if main cohort effects can replace bilinear period effects since these are more natural to model and estimate.

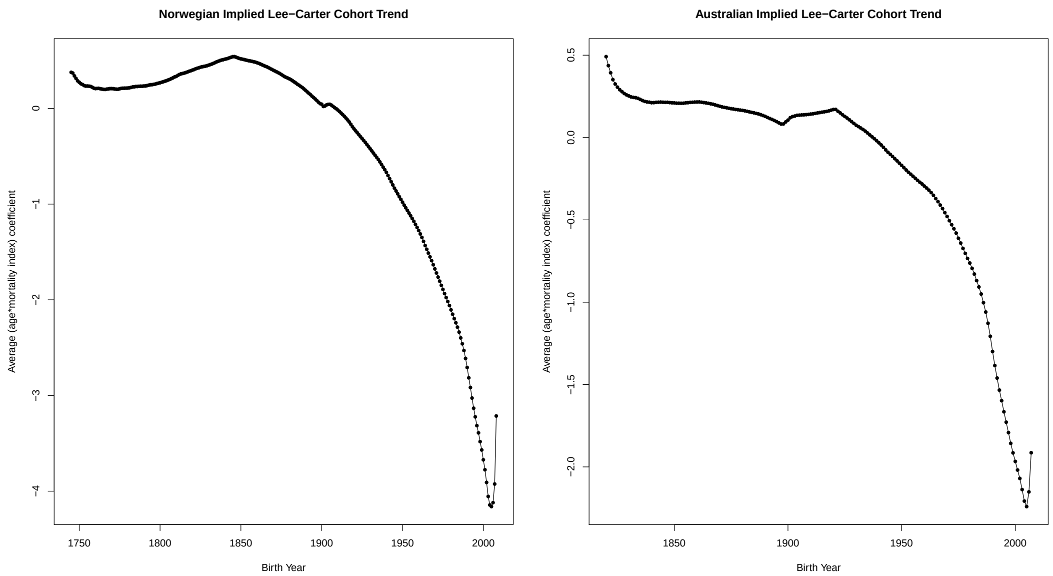

Figure 8: Norwegian and Australian implied cohort trends for the Lee-Carter model.

Define $\gamma_{i,j}$ to be the average of $\alpha_{2,j}\kappa_k$, as defined in Section 3.2, over time. We refer to $\gamma_{i,j}$ as an implied cohort effect. Figure 8 shows the implied cohort trend in the Lee-Carter model. It is very similar to the main cohort trend from our age-cohort model and thus provides evidence that supports the replacement argument.

7 Forecasting using the Age-Cohort Model

In this section we forecast life expectancies for our age-cohort model with $\delta = 50$, that is, where the fifty most distant cohorts are omitted from the dataset. Due to the model assumptions, forecasting in the age-cohort model does not require any form of time series extrapolation. This is in contrast to forecasting with period effects; Kuang *et al.* (2008a,b) provide identifiability and forecasting solutions under this scenario. Using the age and cohort trends, we populate the mortality triangle, shown in Figure 1, with estimates for the unobserved death rates. Note that exposure data does not play a role in obtaining the estimated death rates but will come to play when determining the uncertainty around these rates. Having observed and fitted death rates for all cohorts and ages, a cohort life expectancy table is then produced. The change in perspective provided by the trend models described in Section 2 *naturally* results in cohort life expectancies. Cohort life expectancies are much more meaningful than period life expectancies in quantifying future longevity. Figure 9 shows the expected lifetime at birth plotted against year of birth for Norway and Australia.

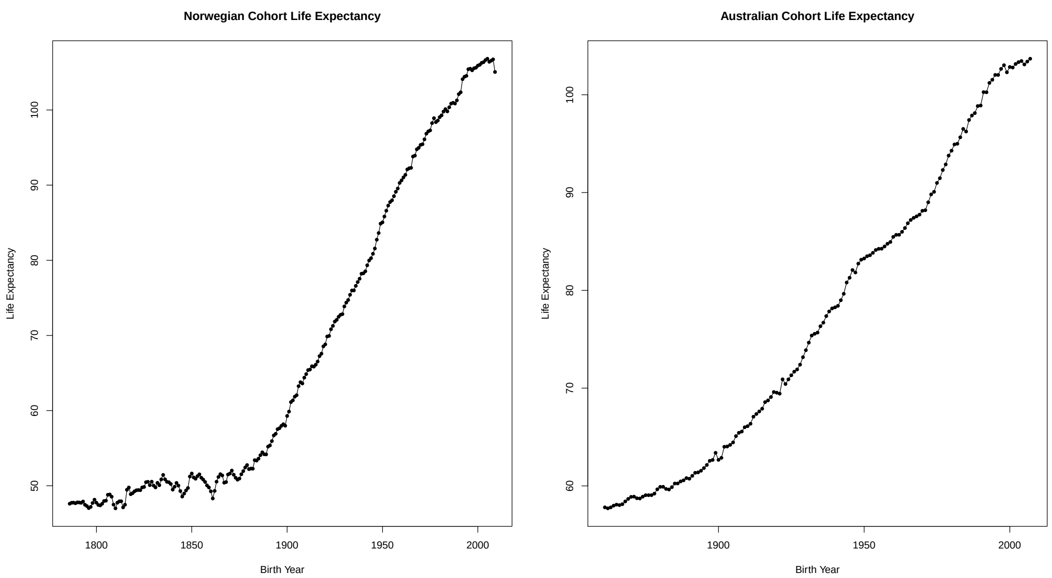

Figure 9: Norwegian and Australian life expectancy at birth.

The uncertainty of these estimated life expectancies is obtained using simulation. First, as a means of addressing parameter uncertainty, we simulate the necessary regression coefficients from a multivariate normal distribution. Recall that the regression coefficients are maximum likelihood estimators; consequently, they are asymptotically normally distributed with a covariance matrix obtained via the inverse of the Fisher-information. Next, with the obtained simulated regression coefficients, we simulate deaths using the overdispersed Poisson assumption. Simulation from the Poisson distribution addresses the process uncertainty inherent in our stochastic model. In order to simulate deaths, we require a measure of exposure; we utilize the average exposure by age of the observed data. In other words, for the simulated data, the cohorts are assumed to have the same exposure levels varying by age.

Figure 10 shows the results of 1,000 simulations. This figure shows the uncertainty in the forecasted life expectancies increases as less information is known about the cohort. Higher levels of both process and parameter uncertainty contribute to this phenomenon.

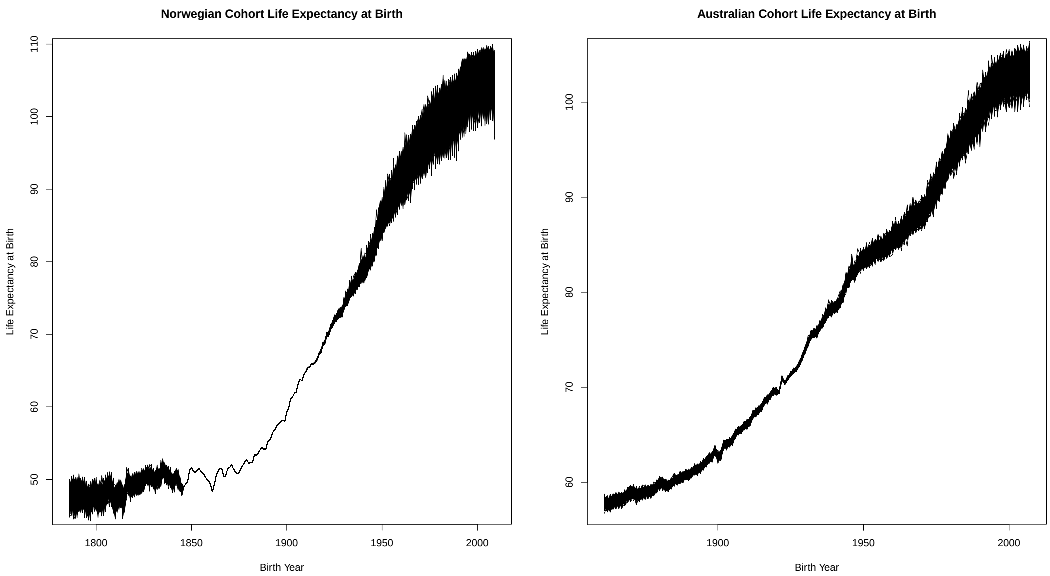

Figure 10: Norwegian and Australian life expectancy at birth for 1,000 simulated scenarios.

8 Conclusions

This paper has shown how trend modelling methods commonly used in non-life insurance to estimate trends in claims experience can enhance the modelling of mortality data. Although doing this is not novel, by treating mortality data in a structure that is similar to that used in non-life insurance, the different perspective to that traditionally used based on life table period data produces models that are relatively easy to estimate, have more direct interpretation and a more natural application to forecasting mortality.

Underlying periodic mortality drivers typically affect a subset of ages and can have immediate and/or delayed consequences; examples include anti-smoking campaigns and new drugs to treat chronic diseases common at older ages. Many stochastic mortality models, motivated by Lee and Carter (1992), include period effects with age interaction to capture time trends varying by age and require additional cohort effects to capture missing trends. We model this using a mortality effect specific to each cohort, namely, the main cohort effect. This approach greatly simplifies model estimation and parameter uncertainty calculations as well as having improved forecasting implications. The model was naturally applied to forecasts of cohort life expectancies.

In conclusion, there is a caution in utilizing the commonly used bilinear period effect with respect to forecasting. Namely, it supposes that for each age, the share of future mortality improvements, given by projections of the mortality index, coincide with the share of past mortality improvements. In other words, age groups that have seen the greatest mortality improvement in the past are forecasted to experience the greatest mortality improvement in the future. This pitfall has been widely recognized and is addressed in a variety of ways; see e.g. Booth *et al.* (2002) for an Australian context. Such issues are fortunately not inherent in our age-cohort model.

Acknowledgements

The authors would like to acknowledge the financial support of ARC Linkage Grant Project LP0883398 Managing Risk with Insurance and Superannuation as Individuals Age with industry partners PwC, APRA and the World Bank as well as the support of the Australian Research Council Centre of Excellence in Population Ageing Research (project number CE110001029). The authors are also indebted to the comments of an anonymous reviewer that led to improvements in both the clarity and structure of the paper.

References

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

Booth, H., Hyndman, R. J., Tickle, L., and de Jong, P. (2006). Lee-Carter mortality forecasting: a multi-country comparison of variants and extensions. *Demographic Research*, **15**, 289–310.

Brouhns, N., Denuit, M., and Vermunt, J. K. (2002). A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics*, **31**(3), 373–393.

Cairns, A. J. G., Blake, D., and Dowd, K. (2006). A two-factor model for stochastic mortality: Theory and calibration. *J. Risk and Insurance*, **73**(4), 687–718.

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

Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., and Khalaf-Allah, M. (2011). Mortality density forecasts: an analysis of six stochastic mortality models. *Insurance: Mathematics and Economics*, **48**(3), 335–367.

Currie, I. D. (2006). Smoothing and forecasting mortality rates with p-splines. Presented at the Institute of Actuaries.

England, P. D. and Verrall, R. J. (2002). Stochastic claims reserving in general insurance. *British Actuarial J.*, **8**(3), 443–518.

Gluck, S. M. and Venter, G. G. (2009). Stochastic trend models in casualty and life insurance. Enterprise Risk Management Symposium.

Haberman, S. and Renshaw, A. E. (1996). Generalized linear models and actuarial science. *The Statistician*, **45**(4), 407–436.

Haberman, S. and Renshaw, A. E. (2011). A comparative study of parametric mortality projection models. *Insurance: Mathematics and Economics*, **48**(1), 35–55.

Heligman, L. and Pollard, J. H. (1980). The age pattern of mortality. *J. Institute of Actuaries*, **107**, 49–80.

Human Mortality Database (2011). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded in 03/2011).

Kuang, D., Nielsen, B., and Nielsen, J. P. (2008a). Forecasting with the age-period-cohort model and the extended chain-ladder model. *Biometrika*, **95**(4), 987–991.

Kuang, D., Nielsen, B., and Nielsen, J. P. (2008b). Identification of the age-period-cohort model and the extended chain-ladder model. *Biometrika*, **95**(4), 979–986.

Lee, R. D. (2000). The Lee-Carter method of forecasting mortality, with various extensions and applications (with discussion). *North American Actuarial J.*, **4**(1), 80–93.

Lee, R. D. and Carter, L. R. (1992). Modeling and forecasting U.S. mortality. *J. American Statistical Association*, **87**, 659–671.

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*. Chapman & Hall, London, 2nd edition.

Murphy, M. (2009). The 'Golden Generations' in historical context. *British Actuarial J.*, **15**(Supplement), 151–184.

Renshaw, A. E. and Haberman, S. (2000). Modelling for mortality reduction factors. Actuarial Research Paper, v. 127, City University, London.

Renshaw, A. E. and Haberman, S. (2003a). Lee-Carter mortality forecasting: A parallel generalized linear modelling approach for England and Wales mortality projections. *J. Royal Statistical Soc. C*, **52**(1), 119–137.

Renshaw, A. E. and Haberman, S. (2003b). Lee-Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics*, **33**(2), 255–272.

Renshaw, A. E. and Haberman, S. (2006). A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, **38**(3), 556–570.

Renshaw, A. E., Haberman, S., and Hatzopoulos, P. (1996). The modelling of recent mortality trends in United Kingdom male assured lives. *British Actuarial J.*, **2**(2), 449–477.

Venter, G. G. (2008). Triangles in life and casualty. AFIR Colloquium papers.

Willets, R. (1999). Mortality in the next millenium. Presented to the Staple Inn Actuarial Society.

Willets, R. C. (2004). The cohort effect: insights and explanations. *British Actuarial J.*, **10**(4), 833–877.