Elsevier Editorial System(tm) for Insurance: Mathematics and Economics
Manuscript Draft

Manuscript Number: IME-D-09-00123R1

Title: On age-period-cohort parametric mortality rate projections

Article Type: Research Paper

Keywords: Mortality forecasting; age-period-cohort effects; forecast statistics; back-fitting, data truncation

Corresponding Author: Prof S. Haberman,

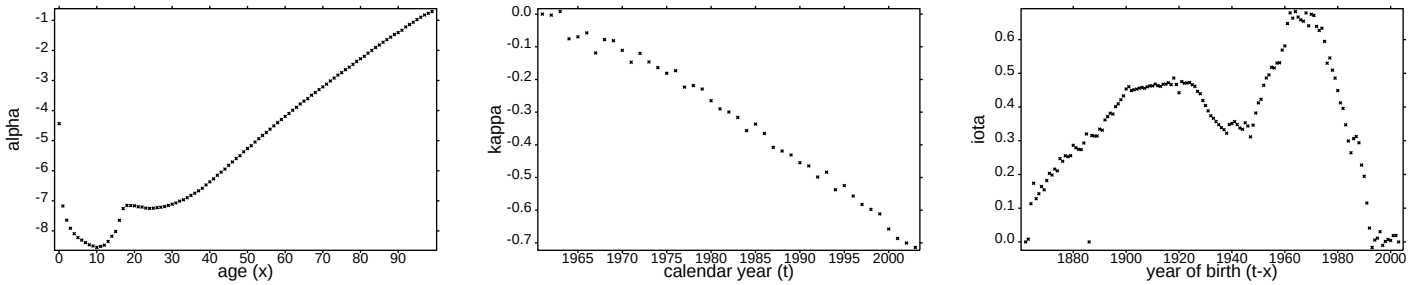
Corresponding Author's Institution: City University

First Author: S. Haberman

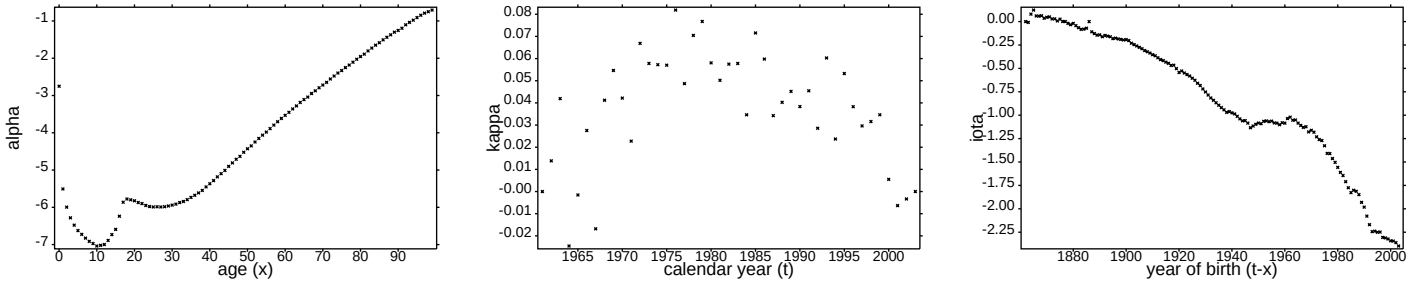
Order of Authors: S. Haberman; Arthur Renshaw, PhD; Arthur E Renshaw, PhD

Abstract: An augmented version of the Lee-Carter modelling approach to mortality forecasting, extended to include an age modulated cohort index in addition to the standard age modulated period index is described and tested for prediction robustness. Life expectancy and annuity value predictions, at pensioner ages and various periods are compared, both with and without the age modulated cohort index, for the England & Wales male mortality experience. The simulation of prediction intervals for these indices of interest is discussed in detail.

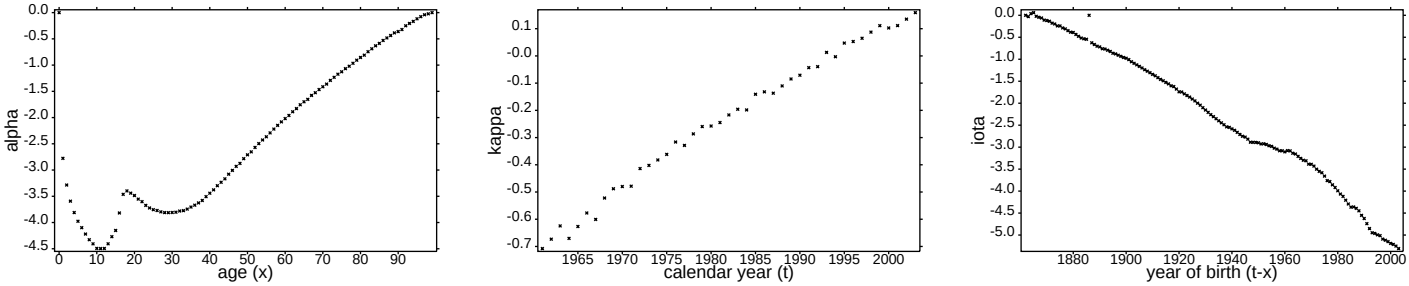
Figure(s)



(i) H0: log{mu(x,t)} = alpha(x) + kappa(t) + iota(t-x)



(ii) H0: log{mu(x,t)} = alpha(x) + iota(t-x) + kappa(t)



(iii) H0: log{mu(x,t)} = kappa(t) + iota(t-x) + alpha(x)

Fig 1. England & Wales 1961-2003 male mortality experience, ages 0-99. Linear structure H0: comprising main age-period-cohort effects fitted in different orders, as specified in the above captions. Main age effect (1st column), main period effect (2nd column), main cohort effect (3rd column). Single stage fitting process, with different model formulae.

Figure(s)

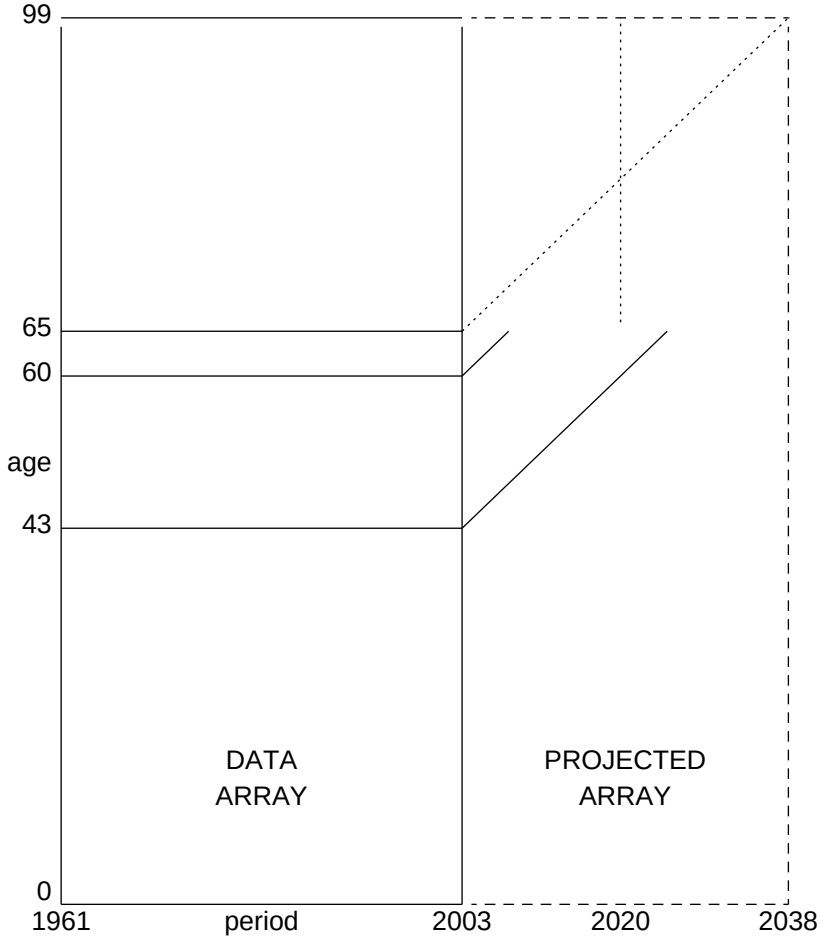
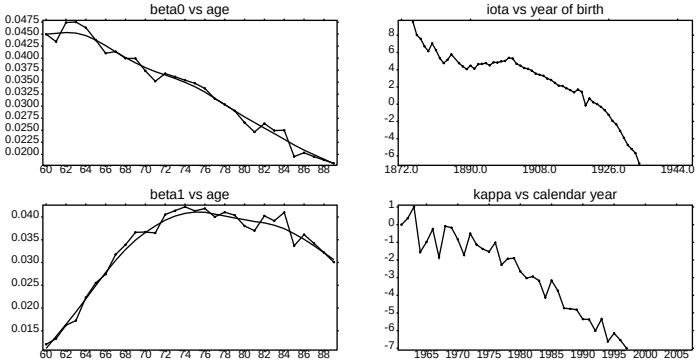
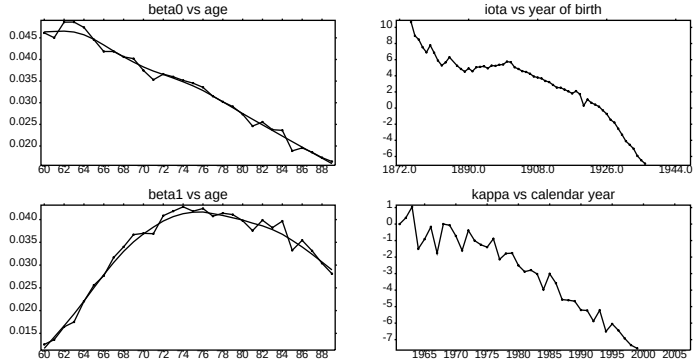


Fig 2. England & Wales 1961-2003 male mortality experience, ages 0-99. Schematic data and projected arrays highlighting projected trajectories (dotted lines) needed when computating statistics of interest (a) by cohort for an individual aged 65 in year 2003, and (b) by period for an individual aged 65 in year 2020. This implies retaining data aged 60 and above for (a), and data aged 43 and above for (b): allowing for the zero weighting of the 5 most recent cohorts in the resulting data array, when truncated. However, truncation by age, involves discarding data contributing directly to relevant cohorts: bounded by the back projection of the (thick) diagonal lines in the projected array.

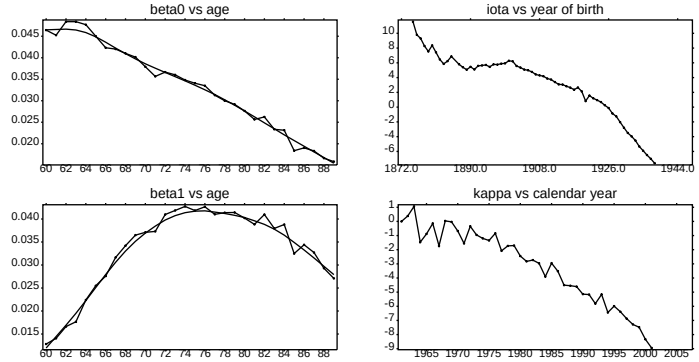
Figure(s)



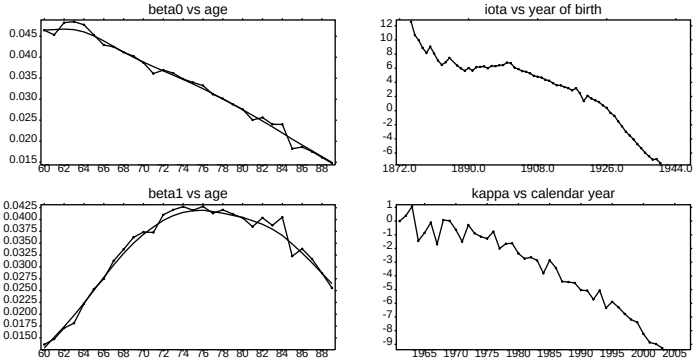
period 1961-1997



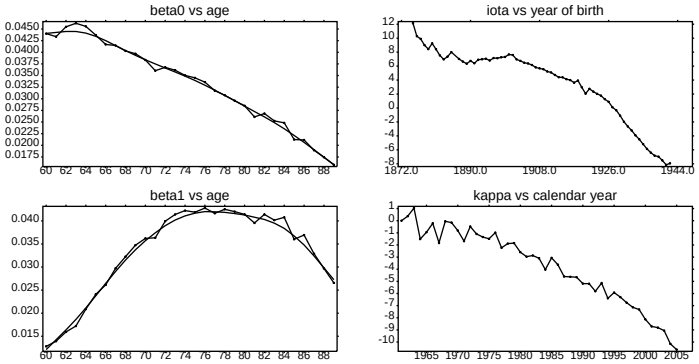
period 1961-1999



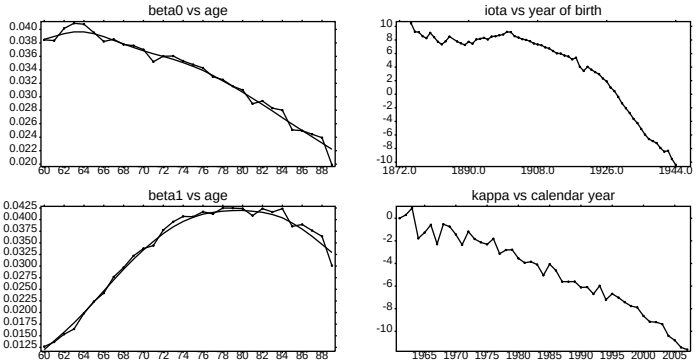
period 1961-2001



period 1961-2003



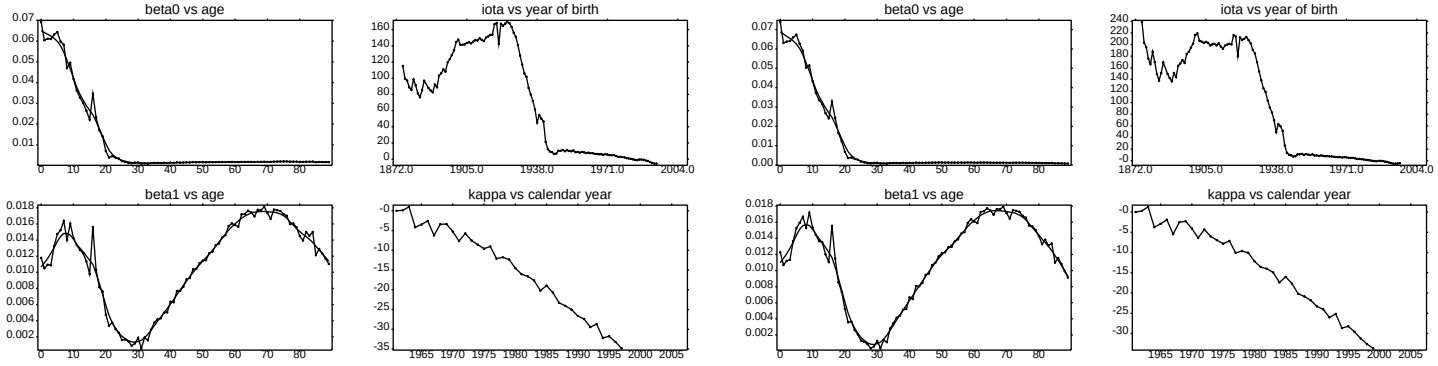
period 1961-2005



period 1961-2007

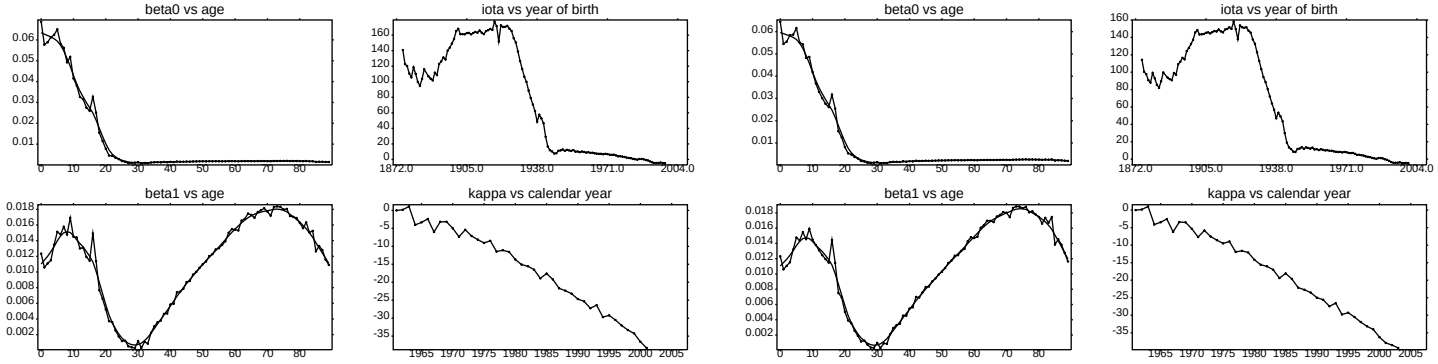
Fig 3. E&W male mortality experience, ages 60-89. APC back-fitting exercise. Reduction factor parameter estimates, with beta parameter smoothing.

Figure(s)



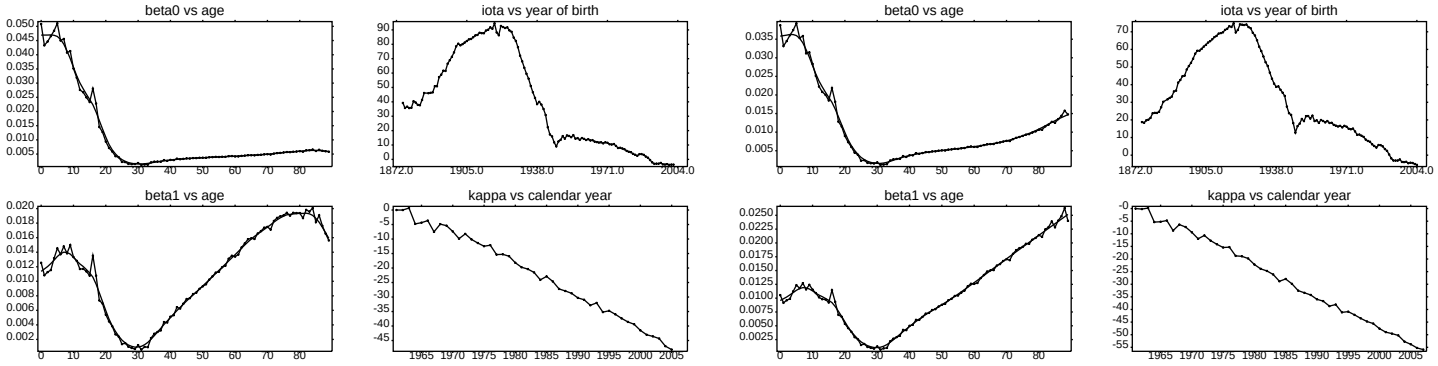
period 1961-1997

period 1961-1999



period 1961-2001

period 1961-2003



period 1961-2005

period 1961-2007

Fig 4. E&W male mortality experience, ages 0-89. APC back-fitting exercise.
Reduction factor parameter patterns, with beta parameter smoothing.

Figure(s)

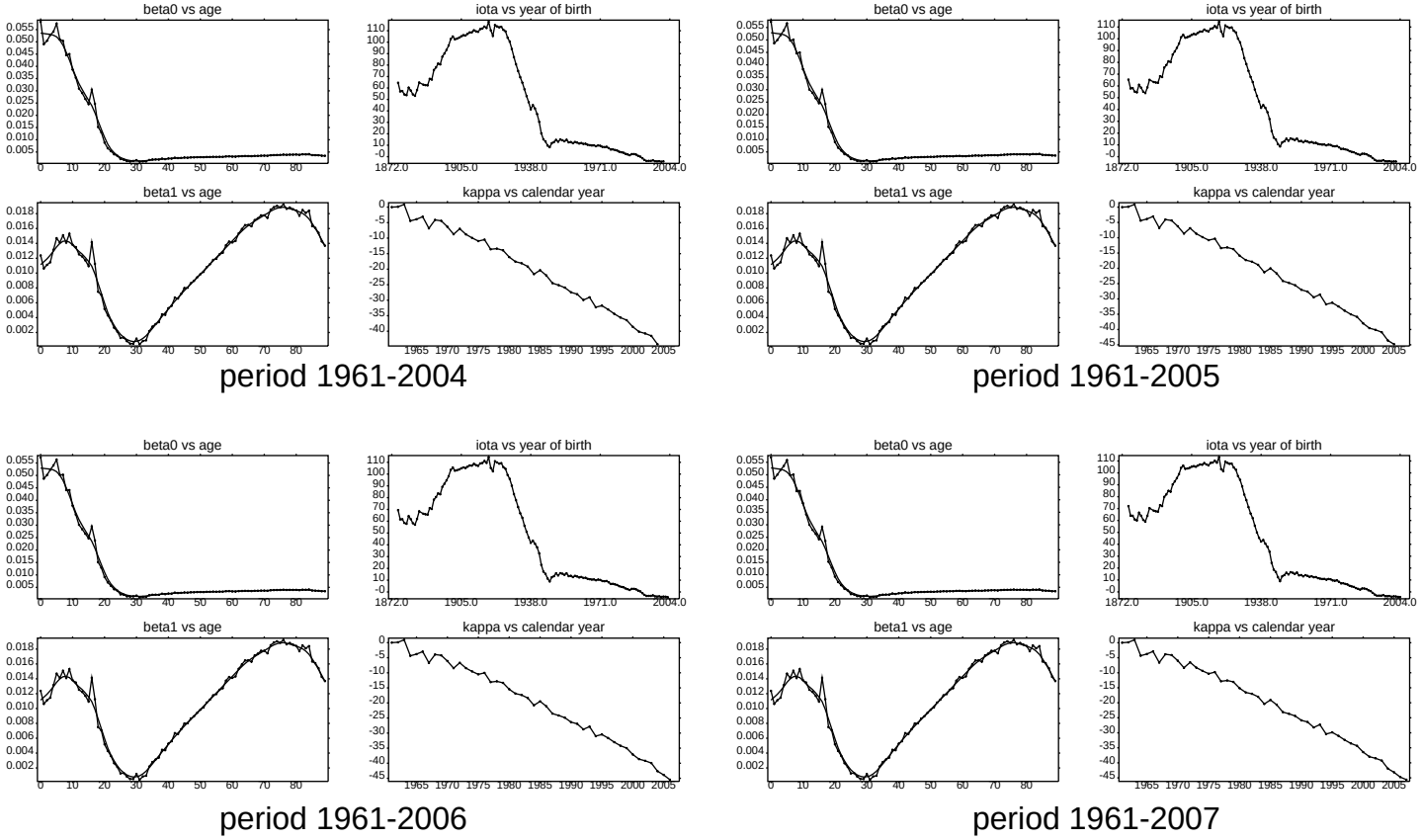
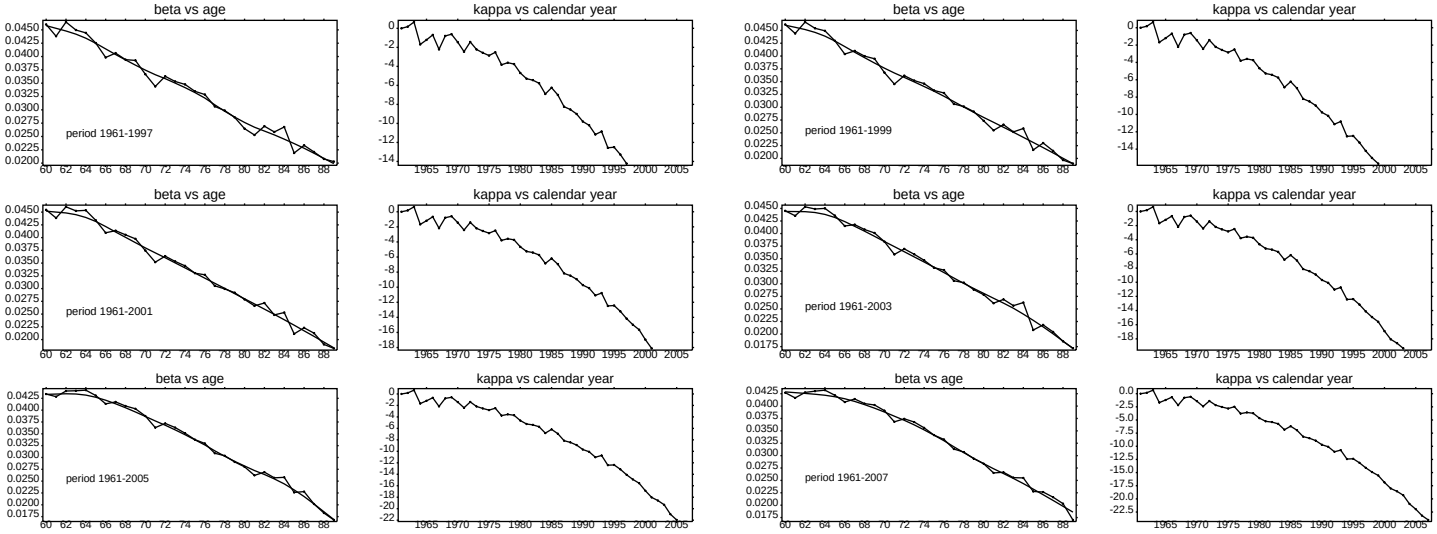
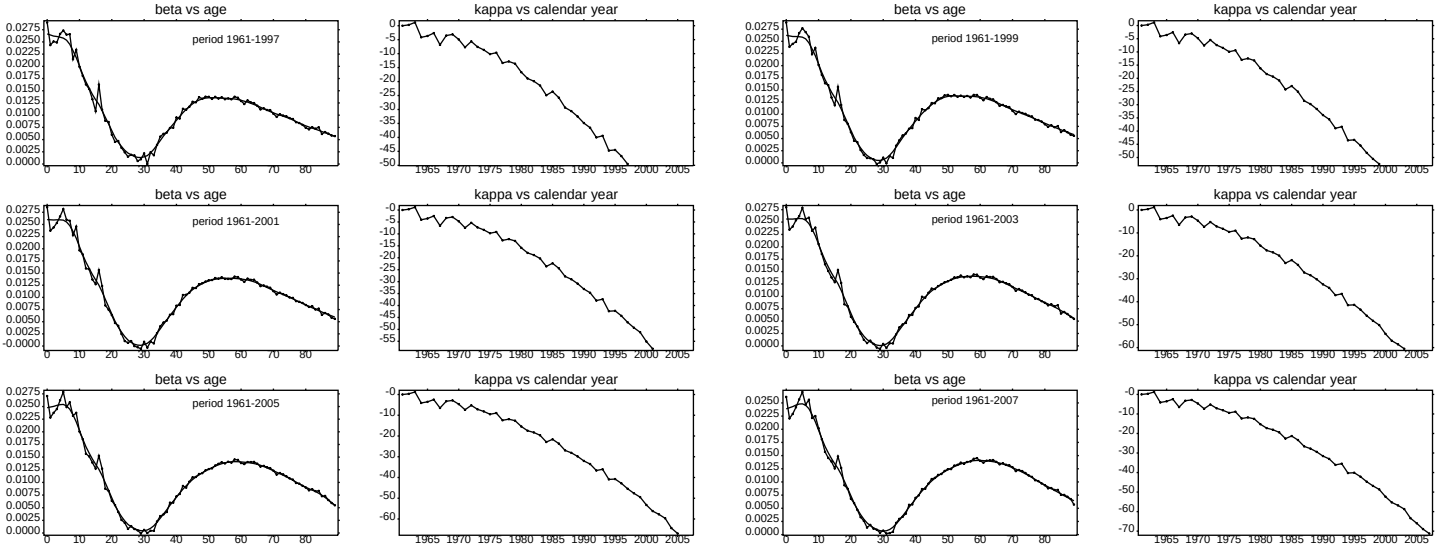


Fig 5. E&W male mortality experience, ages 0-89. APC back-fitting exercise. Reduction factor parameter patterns, with beta parameter smoothing. Patterns in the upper right, and two lower quadrants, are generated by constraining beta1 to the 1961-2004 generated pattern, depicted in the upper left quadrant.

Figure(s)



(i) ages 60-89, sequential periods as stated



(ii) ages 0-89, sequential periods as stated

Fig 6. E&W male mortality experience. LC back-fitting exercise. Reduction factor parameter estimates, with beta parameter smoothing.

Figure(s)

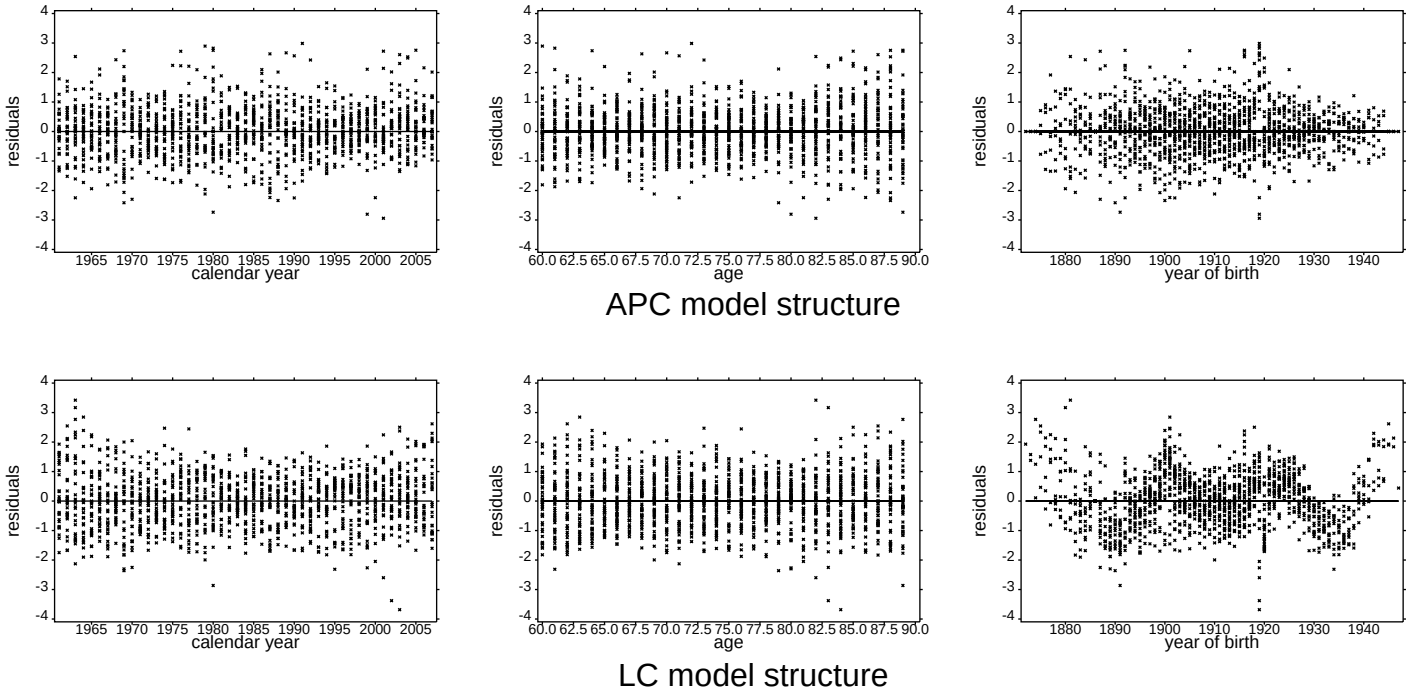
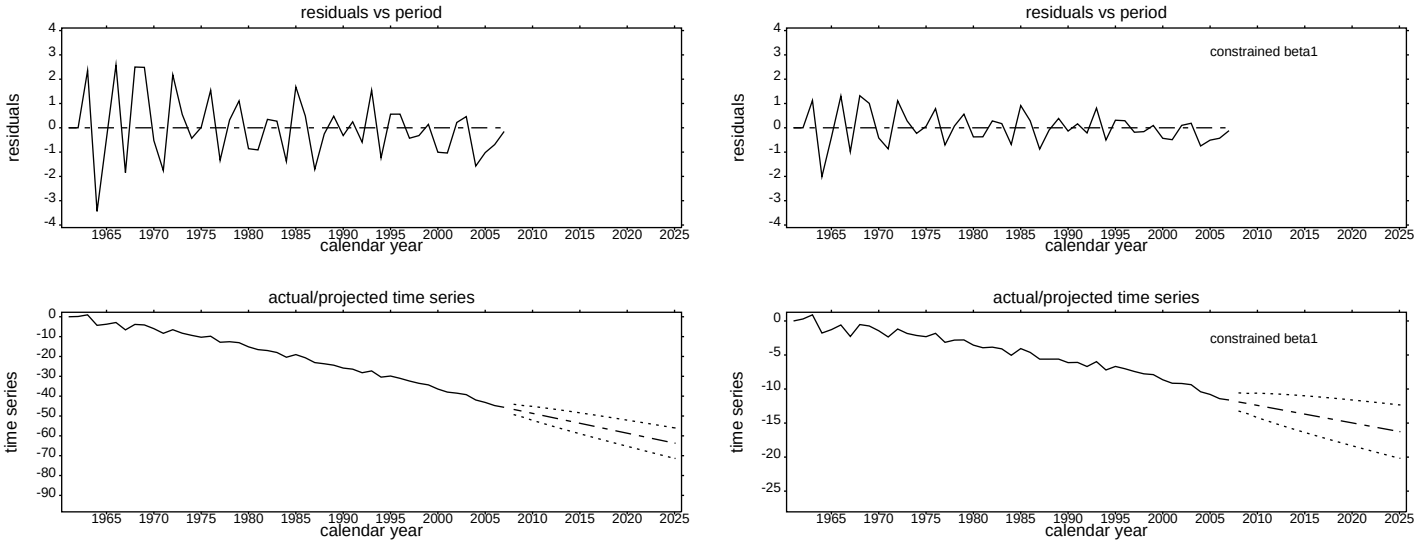


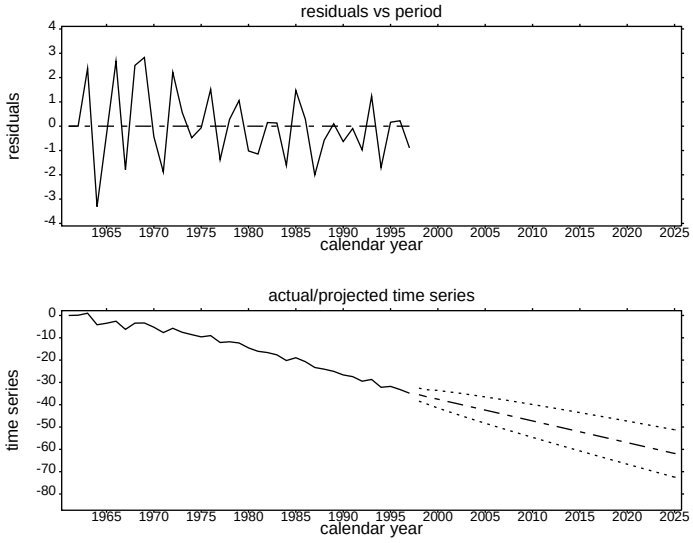
Fig 7. England & Wales 1961-2007 male mortality experience, ages 60-89. Deviance residual plots against: (i) period (left column), (ii) age (centre column), (iii) cohort year-of-birth (right column).

Figure(s)

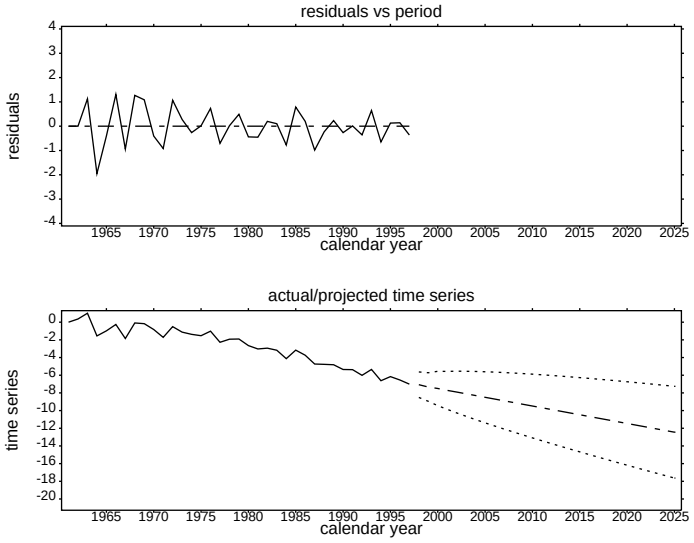


(i) period 1961-2007, (ages 0-89)

(ii) period 1961-2007, (ages 60-89)



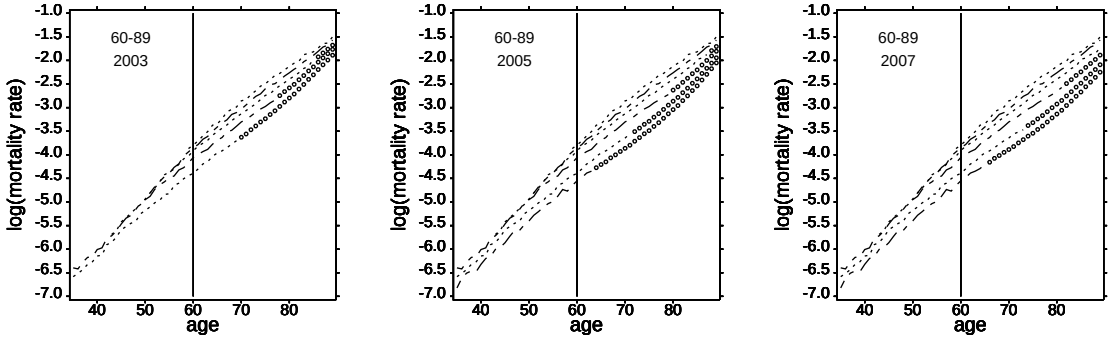
(iii) period 1961-1997, (ages 0-89)



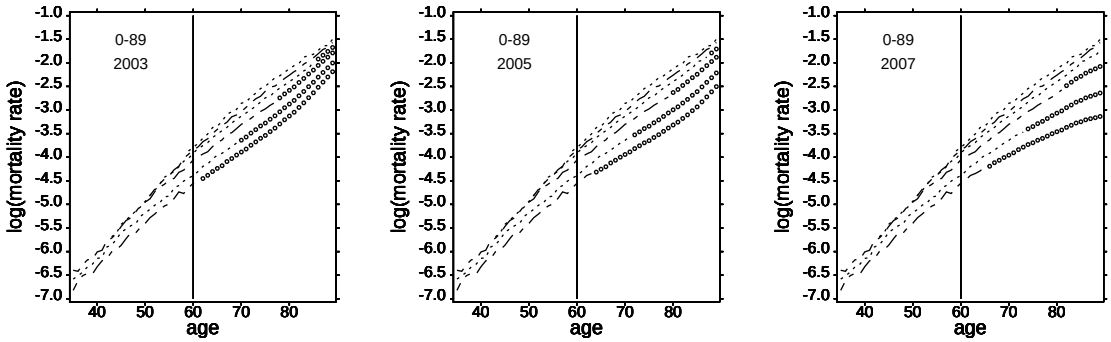
(iv) period 1961-1997, (ages 60-89)

Fig 8. E&W male mortality experience. APC model: period index time series with forecasts and matching residual plot, based on an ARIMA(1,1,0) process. Four cases.

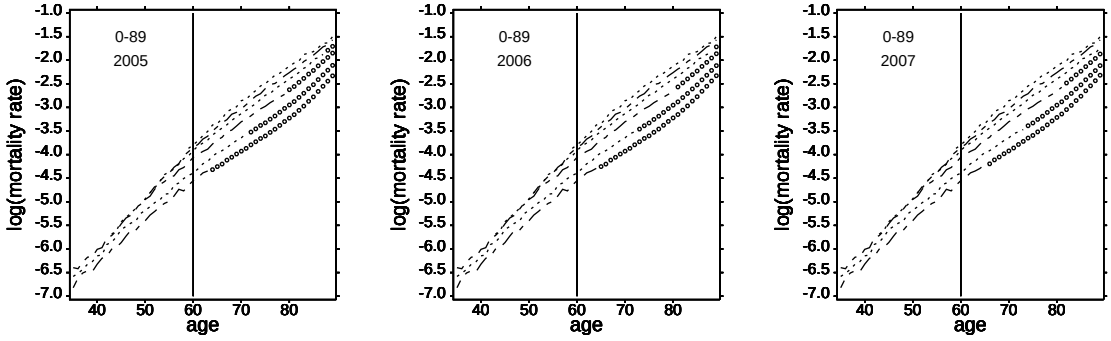
Figure(s)



ages 60-89: periods 2003, 2005, 2007



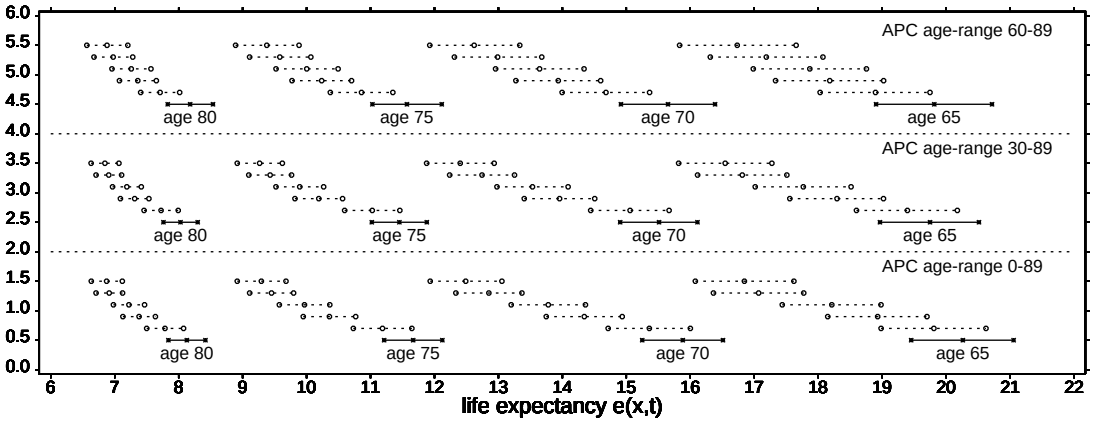
ages 0-89: periods 2003, 2005, 2007



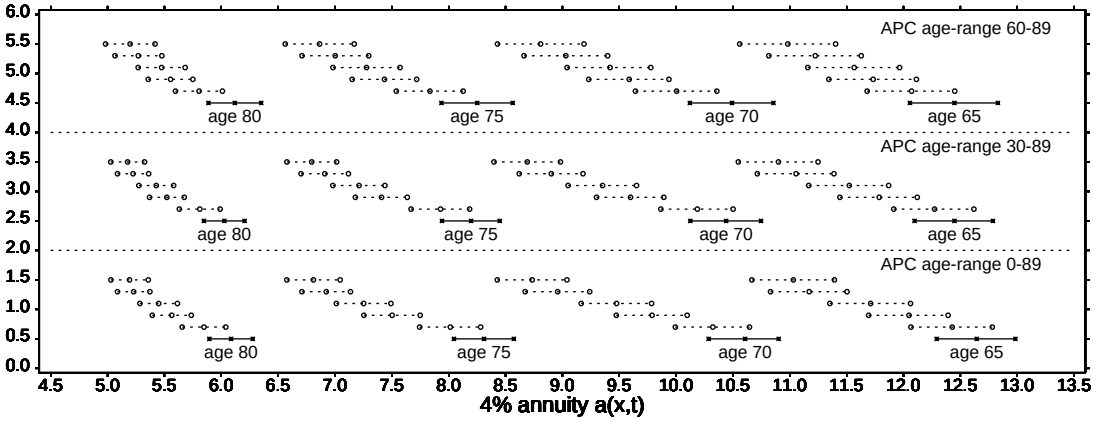
ages 0-89: periods 2005, 2006, 2007 (constrained beta1)

Fig 9. E&W male mortality experience. Empirical log mortality rate profiles with APC forecasts, by individual cohort year-of-birth: 1902, 1910, 1918, 1926, 1934, 1942.

Figure(s)



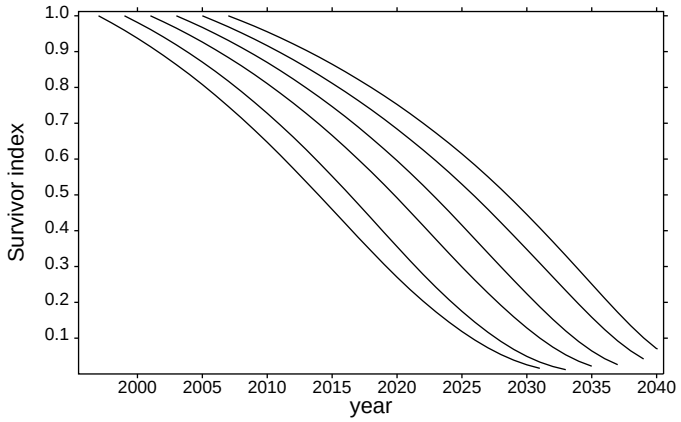
life expectancy predictions, cohort trajectory



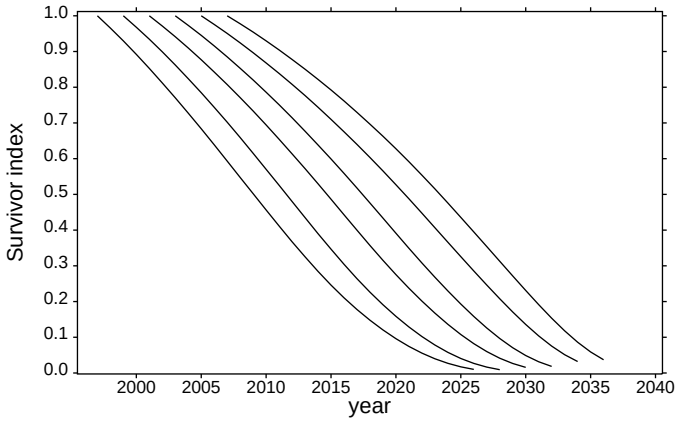
4% annuity predictions, cohort trajectory

Fig 10. E&W male mortality experience, ages 60-89; 30-89; 0-89. APC model. Evolving biennial (1997, 99, 01, 03, 05, 2007) life expectancy and 4% fixed rate annuity predictions, by cohort trajectory: presented in decreasing sequence, for individuals aged 65, 70, 75, 80 respectively. Prediction intervals by bootstrapping the time series prediction error and selecting 2.5, 50, 97.5 percentiles.

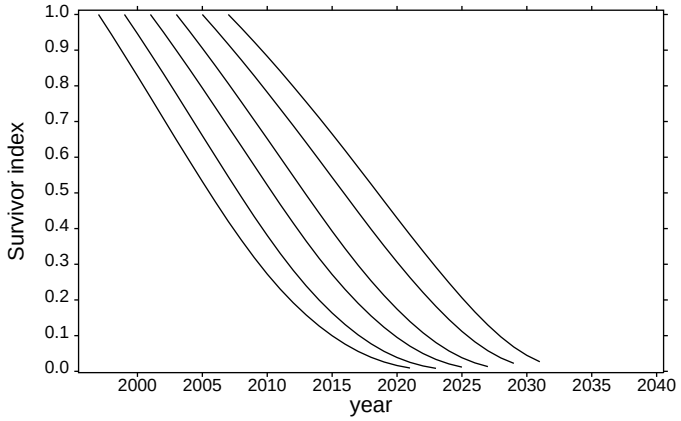
Figure(s)



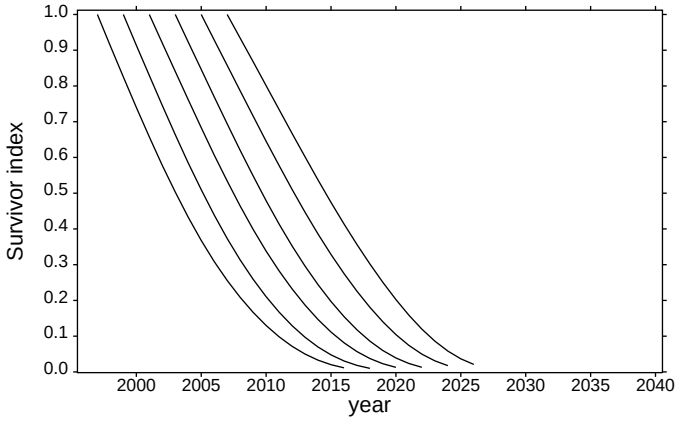
individuals age 65



individuals age 70



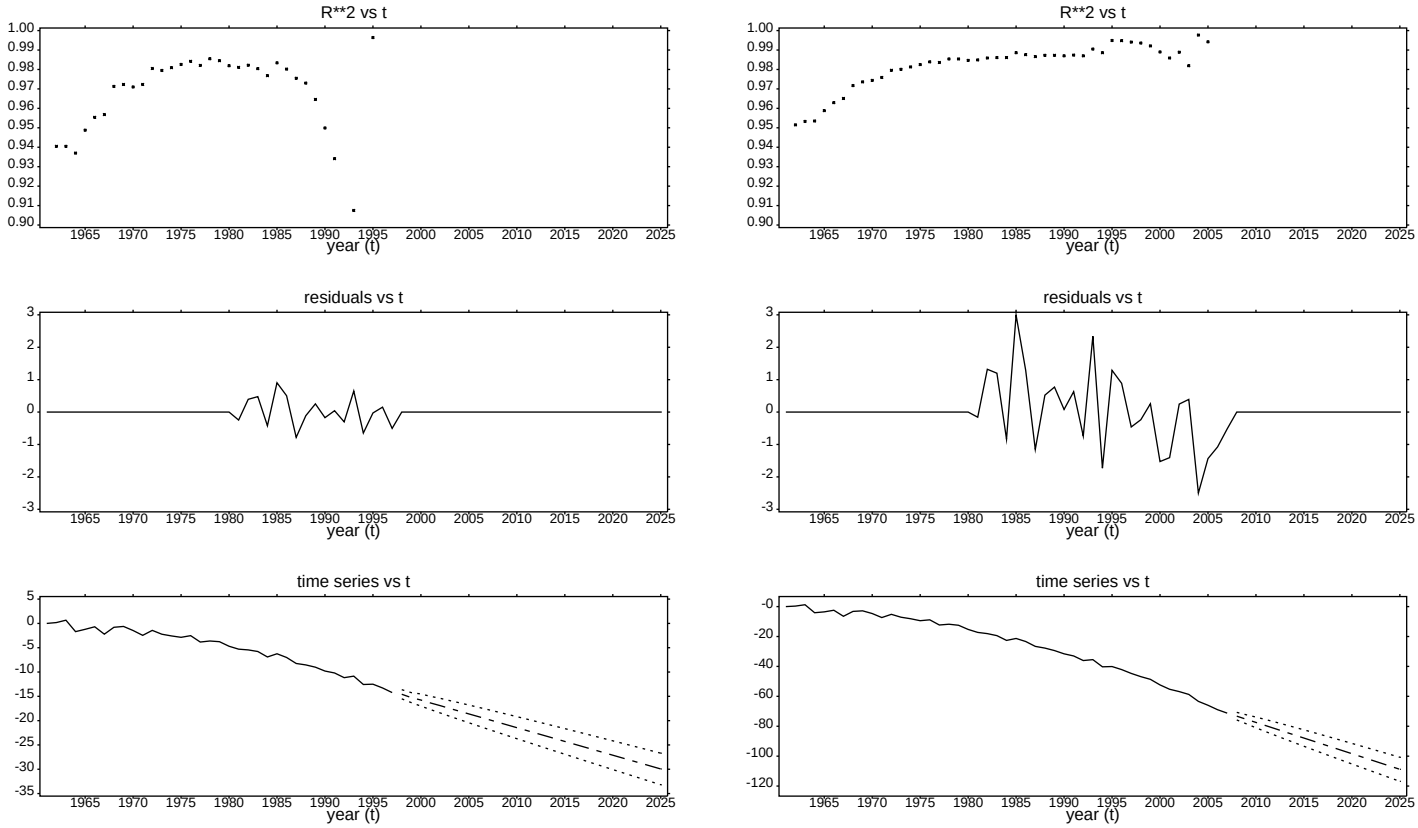
individuals age 75



individuals age 80

Fig 11. E&W male mortality experience (ages 0-89). APC model (with beta1 constrained post 2004). Evolving predicted survivor indices by back-fitting based on the retention of periods 1961 to 1997, 1999, 2001, 2003, 2005, 2007 respectively, depicted in sequence (within each frame): for individuals aged 65, 70, 75, 80.
Simulated prediction intervals omitted.

Figure(s)

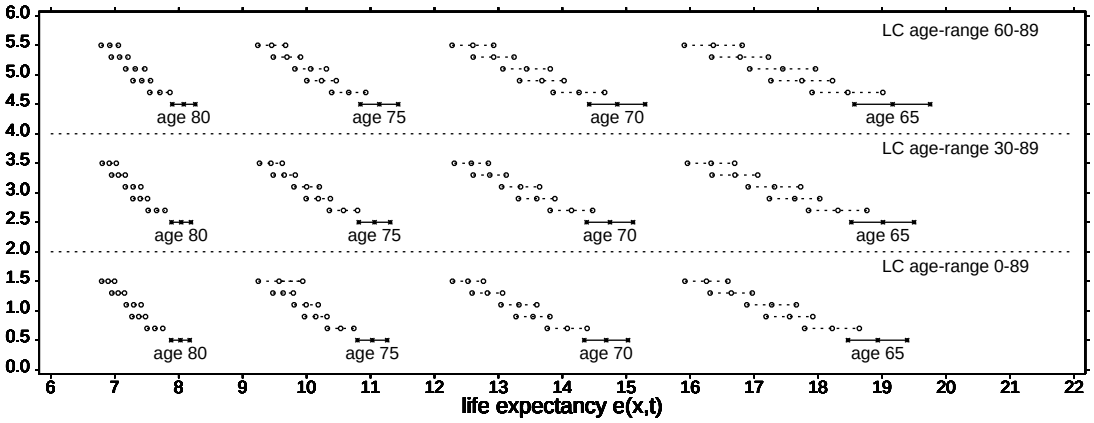


period 1961-1997: ages 60-89

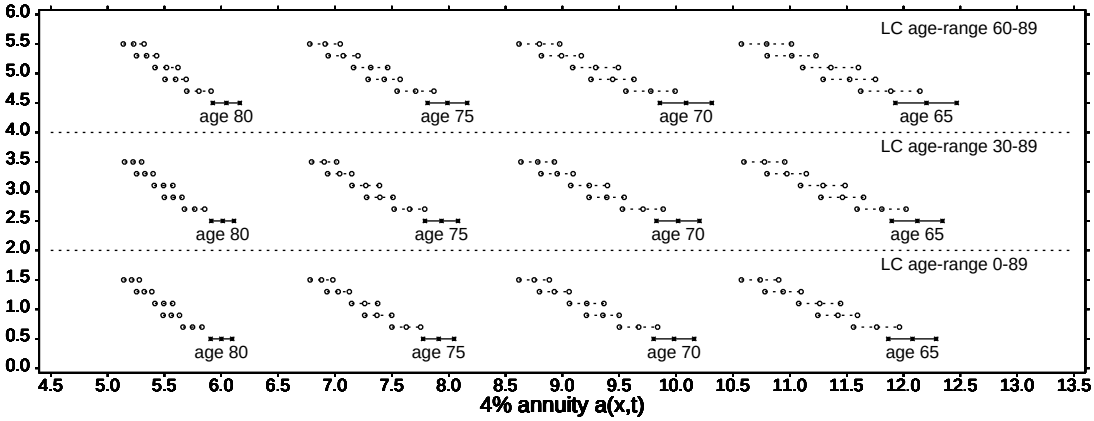
period 1961-2007: ages 0-89

Fig 12. E&W male mortality experience. LC model: period index time series (modelled post 1978), two cases. First row: R**2 linear regression goodness-of-fit statistic. Second row: ARIMA(1,1,0) time series residuals Third row: full time series with predictions and prediction intervals.

Figure(s)



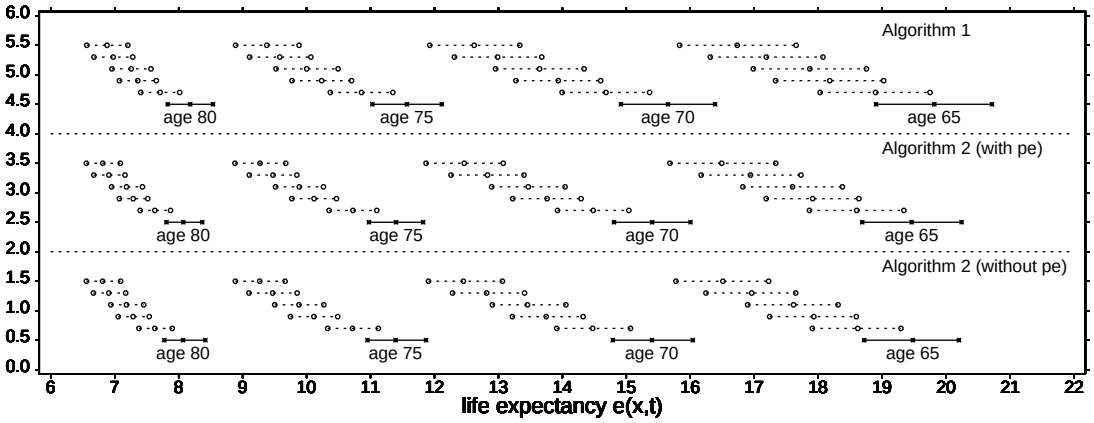
life expectancy predictions, cohort trajectory



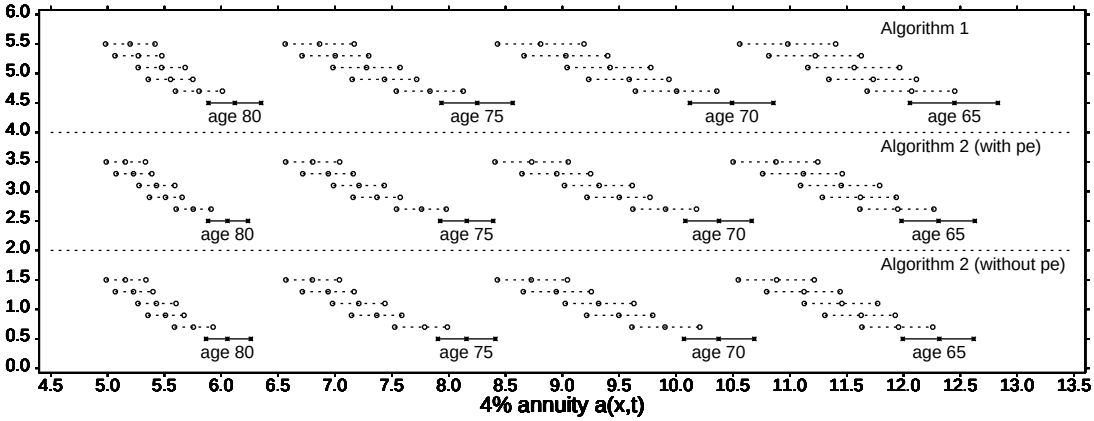
4% annuity predictions, cohort trajectory

Fig 13. E&W male mortality experience, ages 60-89; 30-89; 0-89. LC model. Evolving biennial (1997, 99, 01, 03, 05, 2007) life expectancy and 4% fixed rate annuity predictions, by cohort trajectory: presented in decreasing sequence, for individuals aged 65, 70, 75, 80 respectively. Prediction intervals by bootstrapping the time series prediction error and selecting 2.5, 50, 97.5 percentiles.

Figure(s)



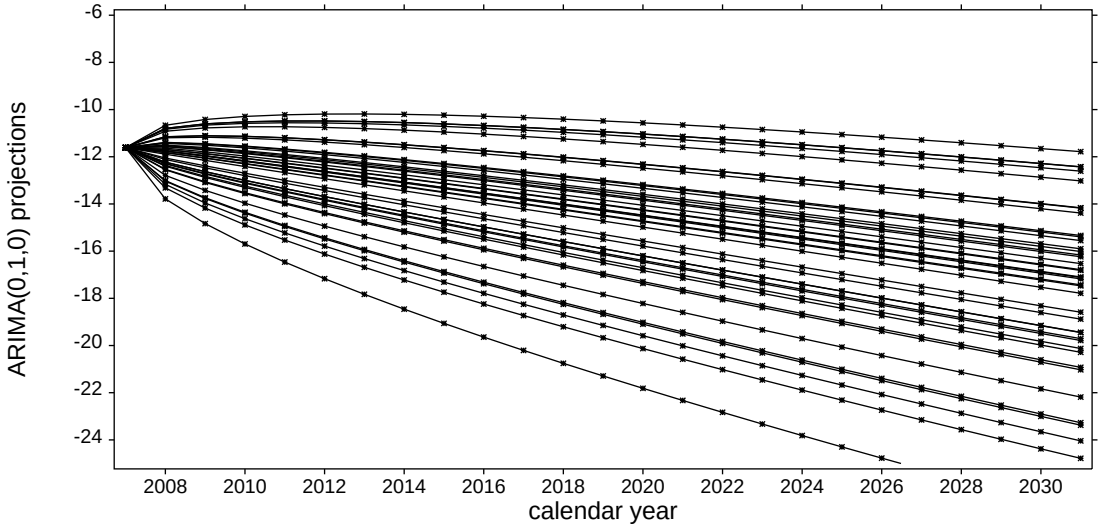
life expectancy predictions, cohort trajectory



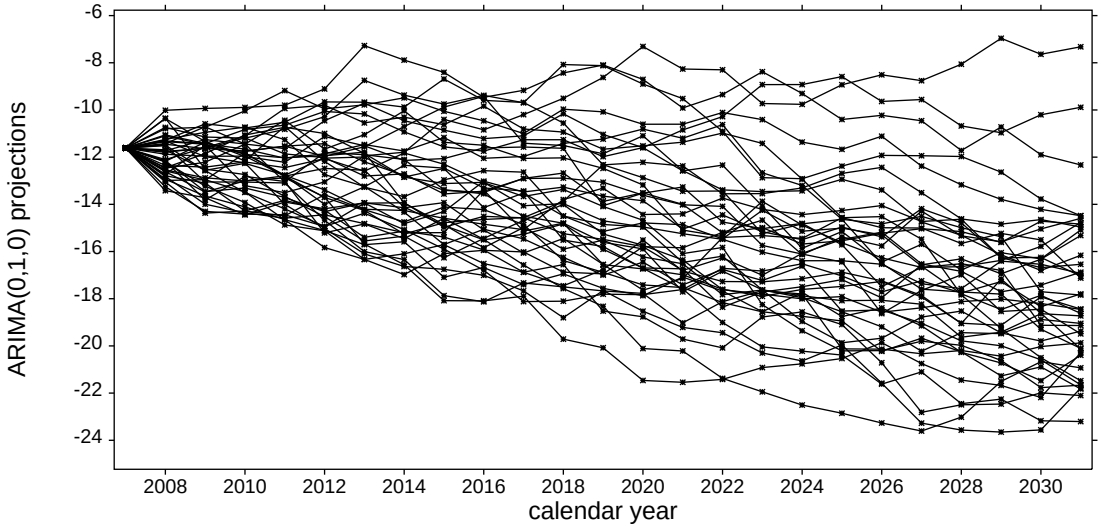
4% annuity predictions, cohort trajectory

Fig 14. E&W male mortality experience, ages 60-89. APC modelling. Evolving biennial (1997, 99, 01, 03, 05, 2007) life expectancy and 4% fixed rate annuity predictions, by cohort trajectory: presented in decreasing sequence, for individuals aged 65, 70, 75, 80 respectively. Prediction intervals by simulation and selecting 2.5, 50, 97.5 percentiles using Algorithm 1 and Algorithm 2, with and without allowance for time series parameter error (pe).

Figure(s)



Algorithm 1: 40 simulations



Algorithm 2 (without Step 1): 40 simulations

Fig 15. E&W male mortality experience, ages 60-89, period 1961-2007. Simulated ARIMA(0,1,0) period times series predictions using Algorithms 1 and 2.

On age-period-cohort parametric mortality rate projections.

Abstract

An augmented version of the Lee-Carter modelling approach to mortality forecasting, extended to include an age modulated cohort index in addition to the standard age modulated period index is described and tested for prediction robustness. Life expectancy and annuity value predictions, at pensioner ages and various periods are compared, both with and without the age modulated cohort index, for the England & Wales male mortality experience. The simulation of prediction intervals for these indices of interest is discussed in detail.

Key words and phrases: Mortality forecasting; age-period-cohort effects; forecast statistics; back-fitting, data truncation

1. Introduction

In Renshaw and Haberman (2006), a method of incorporating cohort effects into the parametric structure of the basic Lee-Carter (*LC*) model for mortality projection was introduced and illustrated (through an age-period-cohort model, abbreviated here to *APC*). In this sequel, we investigate certain wider aspects of this approach to mortality rate projection. It is known that projection results can be sensitive to the length of the historical data period modelled (Denuit and Goderniaux (2005) and Janssen and Kunst (2007)) and we believe that a desirable feature of a mortality projection model is that the parameter estimates should be robust relative to the range of data employed in the estimation – both in terms of the ages and the time periods included. In particular, we ascertain the potential robustness of key indices of interest that depend on predicted mortality rates, by conducting properly constituted back-fitting exercises.

We start from the premise that observed mortality rates covering a continuous period of reasonable duration are required if parameter patterns are to be established, and we then conduct the back-fitting exercise by retaining historical data and stepwise deleting the most recently reported data by individual period. The potential impact on mortality predictions due to data truncation by age, prior to analysis, is also investigated. The simple and transparent nature of mortality predictions under the *LC* modelling framework, generated by treating a single period index as a standard univariate time series model, is also exploited.

The structure of the paper is as follows. Section 2 provides a detailed and systematic presentation of our approach to model fitting, the depiction of issues arising for cohort modelling as a result of possibly arbitrary data truncation by age, a means of 'topping-out' mortality rates in old age and a means by which prediction intervals may be constructed.

A detailed case study, based on the England and Wales 1961-2007 male mortality experience, is then presented in Section 3. Here, the back-fitting exercises, for both *APC* and *LC* modelling, are conducted by deleting data biennially (in the first instance), working backwards from 2007 to 1997, with and without lower data truncation by age. Life expectancy and annuity value predictions are compared for a range of pensioner ages. Results illustrating the different stages are presented pictorially throughout, and

this accounts for the relatively large number of figures. A full discussion of the results follows in Section 4, followed by a summary in Section 5.

2. Poisson Lee-Carter incorporating cohort effects

2.1 Data array

Denote a rectangular mortality data array

$$\left(d_{xt}, e_{xt}, \omega_{xt}\right): \text{ age } x = x_1, x_2, ...x_k, \text{ period } t = t_1, t_2, ..., t_n$$

comprising the reported number of deaths d_{xt}, matching central exposures to the risk of death e_{xt}, and 0-1 weights ω_{xt} to indicate either empty or omitted data cells. It is envisaged that the Lexis diagram underpinning the data array is partitioned into unit square cells of size one year, by single year of age and single calendar year.

2.2 Model details

We target the force of mortality μ_{xt} by modelling the numbers of deaths as independent Poisson responses $D_{xt} \sim Poi(e_{xt}\mu_{xt})$ in combination with the log-link, so that

$$\log E\left(D_{xt}\right) = \log e_{xt} + \log \mu_{xt}$$

together with the parametric structure

$$APC: \log \mu_{xt} = \alpha_x + \beta_x^{(1)}\kappa_t + \beta_x^{(0)}\iota_{t-x}.$$

This generalises the Lee-Carter log-bilinear parametric structure

$$LC: \log \mu_{xt} = \alpha_x + \beta_x\kappa_t$$

by allowing for an age modulated cohort index ι_{t-x} in addition to an age modulated period index κ_t, where α_x is an average of $\log \mu_{xt}$ taken over time *t*.

2.3 Model fitting

Following Renshaw and Haberman (2006), a two-stage fitting strategy is adopted, comprising the estimation of α_x as

$$\hat{\alpha}_x = \sum_t \omega_{xt} \log\left(\frac{d_{xt}}{e_{xt}}\right) = \log \prod_t \left(\frac{d_{xt}}{e_{xt}}\right)^{\omega_{xt}} \tag{1}$$

followed by the implementation of the iterative fitting algorithm

Set starters $\hat{\beta}_x^{(0)}, \hat{\iota}_{t-x}, \hat{\beta}_x^{(1)}, \hat{\kappa}_t$, compute $\hat{d}_{xt}$

↓

Update $\hat{\iota}_{t-x}$, compute $\hat{d}_{xt}$

update $\hat{\beta}_x^{(0)}$, compute $\hat{d}_{xt}$

update $\hat{\kappa}_t$, adjust $\{\hat{\kappa}_t\}$ s.t. $\hat{\kappa}_{t_1} = 0$, compute $\hat{d}_{xt}$

update $\hat{\beta}_x^{(1)}$, adjust $\left\{\hat{\beta}_x^{(1)}\right\}$ s.t. $\hat{\beta}_x^{(1)} > 0 \forall x$, compute $\hat{d}_{xt}$

Compute the model deviance

↓

Repeat the updating cycle, stop when the deviance converges

This is designed to minimise the (conditional) Poisson deviance

$$D = 2\sum_{xt} \omega_{xt} \int_{e_{xt}\mu_{xt}}^{d_{xt}} \frac{d_{xt} - u}{u} du \text{, where } \hat{d}_{xt} = e_{xt} \exp\left(\hat{\alpha}_x + \hat{\beta}_x^{(1)}\hat{\kappa}_t + \hat{\beta}_x^{(0)}\hat{\iota}_{t-x}\right).$$

The first two of the four permitted parameter constraints

$$\kappa_{t_1} = 0,\ \beta_x^{(1)} > 0\ \forall\ x,\ \sum_x \beta_x^{(1)} = \sum_x \beta_x^{(0)} = 1, \tag{2}$$

are used when setting the starting values (Section 2.4) and are subsequently maintained in the iterative fitting algorithm, while the other two constraints are imposed once the algorithm has converged. The second adjustment, ensuring that the age modulating period index $\beta_x^{(1)}$ is consistent in sign, warrants some further detailed discussion (which we defer until Section 4), and may well be found to be redundant in applications. As a potential aid to convergence, we allocate zero weights to the first and last c = 3 (say) cohort years, which occur in opposite diagonal corners of the data array.

2.4 Starting values

Starting values for *APC* are determined by setting $\beta_x^{(0)} = \beta_x^{(1)} = 1, \forall x$ and fitting the Poisson response GLM

$$H_0 : \log \mu_{xt} = \alpha_x + \kappa_t + \iota_{t-x}$$

with $\log \hat{\alpha}_x$ contributing to the offset, and, trivially, the model formula chosen in order to ensure that $\kappa_{t_1} = 0$.

2.5 Why two-stage fitting?

The fitting of model *APC* is complicated by the relationship

year of birth = calendar year – age

which leads to problems of identifiability. Points of concern are best illustrated by fitting the simpler linear model H_0. Refer to Fig 1, in which we have fitted H_0 as a Poisson GLM to the 1961-2003 (ages 0-99) England & Wales male mortality experience, as reported in Renshaw and Haberman (2006), using just three of the many different model formulae possible (design matrices): corresponding to the order in which the three age-period-cohort main effects are specified in the individual captions. This generates three different sets of parameter patterns: only for α_x are the patterns stable and they provide a realistic representation of the main age effects which match the typical shape of a static life table on the log scale. In addition, only the upper α_x profile is correctly scaled.

Faced with the need (in our judgement), to preserve this representation in order to ensure the capture of age effects in a complex modelling situation, we have chosen to adapt the original Lee-Carter (1992) two-stage fitting strategy when fitting *APC*, by conditioning on a predetermined static life table in order to represent the main age-effects. We see no reason to change this approach to fitting *APC* by incorporating iterative adjustments to the main age effects, as suggested by Cairns *et al.* (2009), especially given the wider objective of model projection. In our limited experience of incorporating the Cairns *et al.* (2009) suggestion into the iterative fitting algorithm, we have observed that the already slow rate of convergence is made appreciably slower, and so, we believe that it should be avoided on purely practical grounds. The question as to whether the iterative fitting process then results in attainment of a global optimisation has still to be resolved.

Thus, our approach to fitting *APC* involves 3-stages:

(i) the establishment of a static life table to represent the main age effects;
(ii) conditional on (i), the establishment of main period and cohort effects, through the choice of starting values; and
(iii) conditional on (ii), the establishment of age modulating indices $\beta_x^{(1)}$ and $\beta_x^{(0)}$ using the fitting algorithm.

All aspects of this approach to *APC* model fitting are data driven, rather than being imposed.

2.6 Model dynamics

Mortality rate extrapolation requires the time series forecasts $\left\{\kappa_{t_n+j}: j>0\right\}$ and possibly $\left\{\iota_{t_{n-c}+j-x}: j>0\right\}$, depending on the projected trajectory of interest, so that

$$\mu_{x+j,t_n+j}=\exp\left(\hat{\alpha}_{x+j}\right)F\left(x+j,t_n+j\right),\ F\left(x+j,t_n+j\right)=\exp\left(\hat{\beta}_{x+j}^{(1)}\kappa_{t_n+j}+\hat{\beta}_{x+j}^{(0)}\iota_{t_n+j-x}\right),\ j>0.$$

Here, we have highlighted the decomposition of the forecast mortality rates into the product of a static life table $\exp\left(\hat{\alpha}_x\right)$ and mortality reduction (adjustment) factor *F*. In this approach, we ensure 'continuity' with the latest estimated mortality rates, as opposed to the latest observed mortality rates (as in Renshaw and Haberman (2006)), in the transition to the extrapolated mortality rates. We believe that this feature has a minimal

effect on the indices of interest which are computed subsequently (Section 2.8) and hence it is applied consistently throughout this paper. Typically, we envisage the application of the $ARIMA(p,1,q)$ time series to the main period component: extrapolation of the cohort component is not required for the detailed case study reported in this paper, because we are focusing on back-fitting rather than on cohort-based projections for cohorts reaching age 65 beyond the data range.

2.7 Data truncation

Concerning the use of extrapolated mortality rates in the valuation of fixed rate annuities at typical retirement ages, the question of data truncation by age, prior to modelling, may need to be considered (Cairns *et al.* (2009)). Refer to Fig 2, in which we have depicted typical data and projected arrays, for the England & Wales 1961-2003 male mortality experience (ages 0-99). Here, we highlight the projected trajectories needed when computing statistics of interest (a) by cohort, for an individual aged 65 in year 2003, and (b) by period, for an individual aged 65 in 2020. As illustrated, these require the retention of (a) data aged 60 and above, and (b) data aged 43 and above, allowing for the allocation of zero weights in the last c = 5 cohort years (say). Clearly, data below these age limits contribute directly to the relevant cohort years under age-period-cohort modelling. Thus, we could decide to include all the data available in the modelling exercise or, like Cairns *et al.* (2009), we could argue that our interest in longevity risk means that we should consider a more restricted age range of, say, 60-89. This is a matter of judgement. We follow the former of these approaches on the grounds that the observations at younger ages may improve the model fit and strengthen the parameter estimates for κ_t and ι_{t-x}.

2.8 Indices of interest

We refer to life expectancies and level immediate annuities, computed by the cohort method, thereby allowing for the future evolution of mortality rates. We focus on the most recent period t_n, for which data are available, and so consider individuals aged x at that time. The indices are computed, respectively, as

$$e_x(t_n) = \frac{\sum_{i\geq 0} l_{x+i}(t_n+i)\left\{1-\frac{1}{2}q_{x+i,t_n+i}\right\}}{l_x(t_n)},\ a_x(t_n) = \frac{\sum_{i\geq 1} l_{x+i}(t_n+i)v^i}{l_x(t_n)} = \sum_{i\geq 1} S_{x,t_n}(i)v^i$$

where

$$l_{x+1}(t+1) = (1-q_{xt})l_x(t),\ q_{xt} \doteq 1-\exp(-\mu_{xt})$$

with discount factor v and survivor index (representing the probability of survival from age x to age $x+t$ on the basis of the mortality experience of the cohort aged x in year t_n) $S_{x,t_n}(t): t\geq 0$. In both cases, we truncate the summation over i so that ages beyond a

maximum age of ω are not considered. We note (as mentioned in Section 2.6) that both of these indices do not require the time series modelling of the main cohort index (Fig 2).

2.9 Prediction intervals

The construction of satisfactory prediction intervals for life expectancies involving extrapolated mortality rates, (and, by implication, annuity values), which are computed either by the cohort or period method under the simpler Poisson *LC* modelling framework has proved problematic. See Renshaw and Haberman (2008) for a discussion and investigation of these issues and Denuit *et al.* (2009) for a workable resolution of this problem. This resolution involves the simulation of the forecast error in the associated $ARIMA(p,1,q)$ applied to the main period effect, while ignoring the model fitting error: this latter component is found to be of negligible effect in comparison, a result which is in general agreement with the original findings of Lee and Carter (1992). The simulation strategy that we adopt reads as follows

Algorithm 1

For simulation $m = 1,2,...,M$

1. randomly sample z_m^* from $N(0,1)$.

For $j = 1,2,...,J$

2. set $\kappa_{t_n+j,m}^* = E\left[\kappa_{t_n+j}\right] + \sqrt{Var\left[\kappa_{t_n+j}\right]}.z_m^*$, based on the same *ARIMA* model
3. compute $\mu_{x+j,t_n+j,m}^*$
4. compute the indices of interest.

The construction of equivalent theoretical prediction intervals is also possible (Denuit *et al.* (2009)). This approach readily extends to *APC* modelling for applications which do not require the extrapolation of the cohort index, as is the case here. Further possibilities are discussed and illustrated in Section 4.

2.10 Topping-out by age

Under *APC* (and *LC*) modelling, projected mortality rates

$$\mu_{x+j,t_n+j} : j = 1,2,...,x_k - x \ \left(x < x_k\right),$$

restricted (above) by x_k, which represents the upper limit of the data available, are available for the computation of the indices of interest by the cohort method. In order to implement topping-out by age, when required, projected log mortality rates are extrapolated further along the age axis up to age ω $(> x_k)$, using the following quadratic differencing formula

$$u_j = \log\left(\mu_{x+j,t_n+j}\right) = a + bj + cj\left(j+1\right);\ j = x_k - x - 1, x_k - x, x_k - x + 1, ..., \omega - x.$$

This technique is a variant of the widely used demographic method proposed by Coale and Kisker (1990). We note that the first two terms ensure 'continuity' and, together

with the pre-specification of either $u_{\omega-x}$ or $2c$, the incremental second differences, serve to determine u_j and hence

$$\mu_{x+j,t_n+j} : j = x_k - x + 1, x_k - x + 2, ..., \omega - x .$$

3. Case Study: England & Wales 1961-2007 male mortality experience, ages 0-89

3.1 The data

We focus on the numbers of recorded deaths and matching population sizes exposed to the risk of death, as compiled by the UK Government Actuary's Department (GAD) for the England &Wales male mortality experience. The data are cross-classified by individual calendar year 1961-2007 and age last birthday 0-89. These data are an updated version of the 1961-2003 E&W male mortality experience (ages 0-99), reported and modelled in Renshaw and Haberman (2006): data in the age range 90-99 are no longer readily available from the GAD website.

3.2 Objectives

The primary purpose of this study is to investigate the robustness of the predicted values of certain indices of interest that are dependent on future mortality rates, using the *APC* modelling approach described in Section 2. We carry out this investigation by conducting systematic back-fitting exercises. As a secondary objective, we also investigate the possible effects on these indices of truncating the data set by age from below, prior to modelling. Further, the corresponding predictions using *LC* modelling, with no allowance for cohort effects, are reported for comparison.

Specifically, we focus on the predicted values of life expectancy and present value of a level immediate annuity (based on a 4% interest rate), both computed by the cohort method. Back-fitting involves the retention of data for the periods 1961 to 1997, 1999, 2001, 2003, 2005 and 2007 respectively (abbreviated 97(02)07), while predictions are made for the most recent period available in any retained data, for individuals aged 65, 70, 75, 80 (abbreviated 65(05)80), with and without lower data truncation at age 60 (and 30).

3.3 *APC* back-fitting

The two-stage model fitting strategy (Section 2.3) is used throughout, with the first stage estimates for α_x, obtained by averaging over the requisite period using equation (1). These take the familiar pattern of a static life-table (log scale), with accident 'hump', typically as depicted in the upper left frame in Fig 1: noting that only part of this profile is required when the data are truncated from below by age.

In all cases, zero weights are applied to the $c = 3$ most distant and recent cohorts available for analysis, together with the allocation of a zero weight to the 1886 cohort, for which the population estimates underpinning the calculation of the exposures are suspect (as noted by Renshaw and Haberman (2006)). As an adjunct to modelling, we smooth the fitted age modulating indices $\hat{\beta}_x^{(0)}$, $\hat{\beta}_x^{(1)}$, (using the S-plus super smoother) in order to

avoid localised age induced anomalies when projecting mortality rates (Renshaw and Haberman (2003), Delwarde *et al* (2007)).

The 6 sets of parameter patterns generated by the back-fitting exercise with lower truncation at age 60 are depicted in Fig 3. We particularly draw attention to the stability of these biennial patterns, 97(02)07, coupled with the near linearity of the period index κ_t.

The corresponding 6 sets of parameter patterns generated by the biennial back-fitting exercise 97(02)07, using the full age range 0-89, are depicted in Fig 4. A key feature of note is the progressive straightening of the otherwise consistent pattern in the age modulating period index $\hat{\beta}_x^{(1)}$ at ages 70 and over when the fitting period is extended beyond 2004: back-fitting has been extended to include the cases 1961-2004 (upper left quadrant Fig 5) and 1961-2006 (not shown) and the results confirm the presence of this feature. The consequences for model predictions are discussed shortly. However, this phenomenon is negated if we condition on the established 2004 $\hat{\beta}_x^{(1)}$ pattern (upper left quadrant Fig 5) in the starting values, and suppress their subsequent updating when applying the iterative fitting algorithm (as introduced in Section 2.3). (We note, in passing, that a by-product of this adaptation for model fitting is to speed up the rate of convergence). The results, on constraining $\hat{\beta}_x^{(1)}$ in this way, are depicted in the remaining three quadrants of Fig 5. The linearity of the main period index κ_t, in both Figs 4 & 5, is particularly noteworthy and is in line with the results of Renshaw and Haberman (2006). The distinctive pattern of the cohort index ι_{t-x} and its interpretation (Renshaw and Haberman (2006)) is also of passing interest.

3.4 *LC* back-fitting

The corresponding biennial back-fitting patterns for the case of parameters for the *LC* model, both with and without lower age truncation at 60, are depicted in Fig 6. The stable patterns, together with mild curvature in the period index κ_t are of particular note throughout. This feature is in contrast with the lack of curvature noted in the previous section for the *APC* model.

3.5 Model diagnostics

In our experience, basic displays of deviance residuals, plotted separately against period, age and cohort year-of-birth are both essential and more informative than separate rectangular (age-period) plots of positive and negative residuals. In Fig 7, we present 2 such sets of plots, one each for *APC* and *LC* modelling: for the case of 1961-2007 with age truncation at age 60. These plots are representative of all of the other cases, highlighting the capture of period and age effects under both model structures, together with the ability (upper panel) or inability (lower panel) of the models to capture cohort effects, as the case may be.

3.6 The *APC* period index time series

For *APC* modelling, exploratory time series analysis of the main period indices, leads to the adoption of an $ARIMA(1,1,0)$ in all cases. For this time series

$$y_t = \kappa_t - \kappa_{t-1},\ y_t = \theta + \phi y_{t-1} + \xi_t;\ \xi_t \sim N\left(0, \sigma^2\right)\ i.i.d.\ \ t = t_2, t_3, ..., t_n$$

with forecasts $\kappa_{t_n+j} : j = 1, 2, 3, ...$, where

$$E\left[\kappa_{t_n+j}\right] = \kappa_{t_n} + \sum_{i=1}^{j}\left\{\mu + \phi^i\left(y_{t_n} - \mu\right)\right\},\ \mu = \frac{\theta}{1-\phi}$$

$$Var\left[\kappa_{t_n+j}\right] = \left\{\sum_{i=1}^{j}\left(\sum_{k=0}^{i-1}\phi^k\right)^2\right\}\sigma^2 .$$

Typical details, including residual plots, are illustrated in Fig 8.

3.7 *APC* log mortality rate predictions

The evolutionary aspect of forecast mortality rates generated by *APC* back-fitting, are presented in Fig 9, in the first instance. Here, within each frame, we have plotted the empirical log mortality rates for regularly spaced cohort years-of-birth 1902(08)42 (where appropriate), with the superimposed *APC* projections. The respective frames in the first two rows relate to predictions made in 2003(02)07, both with (first row) and without (second row) prior data truncation at age 60. Specifically, the second row of frames reveals the evolving over-optimistic nature of the post 2003 mortality forecasts induced by the change in shape, in the upper age range, of the age modulating period index $\beta_x^{(1)}$. Modified 2005(01)07 predictions, corrected by constraining the modulating period index $\beta_x^{(1)}$ to 2004 levels, are displayed in the third row of frames.

3.8 *APC* predicted indices of interest

In the respective upper and lower frames in Fig 10, we present the evolving biennial 97(02)07 life expectancy (upper frame) and 4% fixed rate annuity (lower frame) predictions, computed by cohort trajectory, for individuals aged 65(05)80, with and without prior truncation of the data by age. In addition to truncating the data at age 60 (as reported throughout), we have also repeated the analysis on the basis of data truncation at age 30, and report the outcome in Fig 10. Each prediction interval is constructed on the basis of *M* = 5,000 simulations, as described in Section 2.9 using Algorithm 1.

The general alignment of matching predictions made with, and without, lower data truncation, for these data, is noteworthy. Comparing like with like, slightly wider prediction intervals result when lower data truncation at age 60. Further comments follow in Section 4.

3.9 *APC* predicted survival indices

In Fig 11, we present the six evolving biennial predicted survivor indices, within each frame, for individuals aged 65(05)80, based on the full age range: these indices have been used as part of the construction of Figs 9 & 10. Prediction intervals (or fans) are

omitted for clarity. Similar patterns are observed when we use lower data truncation at age 60 (results are not shown).

Here, the tails (last 10 years) of the indices are constructed as described in Section 2.10, using the settings $\omega = 99,\ 2c = 0.01$, throughout. By fixing the incremental second differences, $2c$, in this way, the extrapolated log mortality rates follow from summing the required number of constant differences with $u_{x_k - x} - u_{x_k - x - 1}$ as the starting value, and then combining the outcome with $u_{x_k - x - 1}$ as the initial value before summing a second time.

The topping-out process, typically contributes the following amounts to the overall values of the *APC* predicted indices of interest:

Age	65	70	75	80
Life expectancy	19.81, 18.17	15.65, 14.09	11.57, 10.14	8.18, 6.12
4% annuity	12.45, 11.84	10.49, 9.79	8.25, 7.46	6.12, 5.12

In this table, period 2007 point predictions, subject to prior age truncation at age 60, have been computed with, and without using the extrapolated tail. We note that the contribution from the topping-out process is significant at the oldest ages and that it diminishes, as we would anticipate, as a proportion of the overall measure when we move to younger ages. This process is applied consistently throughout, so as to facilitate the comparisons.

3.10 The *LC* period index time series

Forecasting with the *LC* model is complicated by the lack of linearity in the period index (as shown in Fig 6). While good time series fits are achieved using second order $ARIMA(p,2,q)$ processes, we avoid this possibility, in part, because the associated prediction intervals are excessively wide. Further reasons for not using this process are presented in Section 4. Instead, working backwards, we attempt to choose the optimal set of data to be included in the fitting process so that the resulting period index is effectively linear – thus we curtail the effective portion of the time series at a perceived point of departure from linearity. In order to assist with this somewhat subjective process, we monitor the profile of the R^2 linear regression goodness-of-fit statistics, which have been constructed sequentially working backwards (as proposed by Denuit and Goderniaux (2005)). For the purpose of this analysis, the portion of the *LC* period index post 1978 is modelled as an $ARIMA(1,1,0)$ process. By way of illustration, R^2 profiles, together with the corresponding residual and time series plots, for two cases, are displayed in Fig 12. In order to comply fully with the procedure prescribed in Denuit and Goderniaux (2005), it is then necessary to re-estimate α_x and β_x by fitting the simple Poisson GLM

$$\log \mu_{xt} = \alpha_x + \beta_x (t - \bar{t})$$

to the reduced data set, where $\bar{t}$ denotes the mean of the reduced period.

3.11 *LC* predicted indices of interest

In Fig 13, we present the evolving biennial 97(02)07 life expectancy and 4% fixed rate annuity predictions, computed by the cohort method, for individuals aged 65(05)80, with and without lower data truncation at age 60 and age 30. Direct comparison with Fig 10 is possible, as the same scales are used.

As with Fig 10, the general alignment of the corresponding predictions in Fig 13, made with, and without, lower data truncation, for these data, is noteworthy. As previously, on comparing like with like, we note the slightly wider prediction intervals which result when lower data truncation at age 60 is applied.

Direct comparison of matching elements between Fig 13 and Fig 10 indicates the extent to which life expectancy and annuity predictions are understated on the basis of *LC* modelling, compared with *APC* modelling. This is discussed further in Section 4.

4. Discussion

In this section, we look at a number of issues arising from the above analysis.

4.1 We have conducted equivalent *APC* and *LC* back-fitting exercises on the earlier 1961-2003, ages 0-99, version of the E&W male mortality experience (as analysed in Renshaw and Haberman (2006)). These results are not reported in detail here – however, based on biennial deletions from 1995 to 2003, our findings are generally consistent with the current reported study, but do not show the evolutionary nature of the age modulating period index, post 2004, (at ages 0-89) which is discussed in Section 3.3 above.

4.2 The persistent linear nature of the period index, observed throughout the back-fitting exercises involving biennial 'front-end' 97(02)07 data reductions under *APC* modelling, brings clarity and simplicity to the time series element of this approach to prediction. In a separate exercise, generating 2007 predictions on the basis of biennial 'back-end' 61(02)71 data reductions (age-range 60-89), the persistent linearity of the period index is again observed: generating consistent 2007 predictions in the process, stacked one above the other if depicted as in Fig 10. As a consequence, the need to establish an "optimal fitting period" does not arise. However, it is likely that the linearity of the period index is data specific, although a preliminary investigation indicates that it is also a feature when APC modelling is applied to the corresponding data set for females.

For LC modelling, issues posed by the occurrence of mild curvature in the period index have to be addressed (Section 3.10). Although a more refined treatment of this issue is possible, again involving a subjective element, by establishing a rolling "optimum fitting period", the LC predictions depicted in Fig 13 would have changed very little if this refinement had been fully implemented.

4.3 The relatively close alignment of the predictions under *APC* modelling, with and without prior lower data truncation at ages 60 and 30 (see Fig 10), is reassuring, given the loss of data under truncation that directly contribute to the relevant cohorts. In the back-fitting exercise with prior lower data truncation at age 30, (where the results are not fully

reported), it was found necessary to constrain the age modulating period index $\beta_x^{(1)}$ post 2005 (Fig 10).

4.4 In *APC* and *LC* modelling applied to the case study, the slope of the period index κ_t is negative throughout. In the event that the associated modulating index $\beta_x^{(1)}$ is also negative (at age *x* say), the contribution from the period index to the overall mortality rate represents increasing mortality over the whole of the period concerned. Initially, we have not applied the second of the constraints (2), because mortality deterioration at certain ages could be a realistic possibility. This feature is supported by the data: for example, in the back-fitting exercises, conducted without prior lower age data truncation, negative values of $\beta_x^{(1)}$ result for individuals in their late 20s (Figs 4 & 5). In the event, the constraint ($\beta_x^{(1)} > 0, \forall\, x$) is found to be redundant throughout the *APC* back-fitting exercises reported in Section 3 both with, and without, prior lower data truncation at age 60.

4.5 The above constraint is needed, however, when prior lower data truncation is applied at age 30, when completing the calculations for Fig 10. Here, without applying the constraint, values of the modulating index are consistently negative for ages below the upper 40s. Hence, in this case, the constraint on $\beta_x^{(1)}$ is necessary to prevent the model from inducing a seemingly arbitrary switch across the age range, from improving to deteriorating period contributions to the overall mortality rates.

4.6 In the *LC* and expanded *APC* frameworks, mortality rate projections required for computing the indices of interest by the cohort method are controlled by the product term, $\beta_x\kappa_t$, which comprises a single age modulated period component. Patterns in the time dependent constituent components are data driven (as opposed to imposed), while stability over time in both components is a prerequisite to robust forecasting. We note that the introduction of the additional cohort structure into the *LC* structure is observed to lead to an evolutionary pattern change in the age modulating period index, (for ages 70 and above) after a period of stability, as we change the fitting period. As we have seen, any pattern in the age modulating index, which is observed to become established over a period, or otherwise, may be readily imposed by trivial adjustments to the fitting algorithm.

4.7 As a possible alternative approach to the construction of prediction intervals (Section 2.9), we adapt the algorithm to read as follows

Algorithm 2

For simulation $m = 1, 2, \ldots, M$

1. generate $\theta_m^*, \phi_m^*, \sigma_m^*$

For $j = 0, 1, 2, \ldots, J$

2. randomly sample $z_{j+1,m}^*$ from $N(0,1)$
3. generate $\kappa_{j+1,m}^* = \theta_m^* + \kappa_{j,m}^* + \phi_m^* \left(\kappa_{j,m}^* - \kappa_{j-1,m}^* \right) + \sigma_m^* z_{j+1,m}^*$

4. compute $\mu^{*}_{x+j,t_n+j,m}$

5. compute the indices of interest.

Provision for parameter error in Step 1 may be based on the (marginal) posterior parameter distributions

$$\frac{(n-2)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-2};\ \begin{bmatrix}\theta\\ \phi\end{bmatrix} \sim mulN\left(\begin{bmatrix}\hat{\theta}\\ \hat{\phi}\end{bmatrix},\ \hat{\sigma}^2\begin{bmatrix} n-2 & \sum_{i=t_2}^{t_{n-1}} y_i \\ \sum_{i=t_2}^{t_{n-1}} y_i & \sum_{i=t_2}^{t_{n-1}} y_i^2 \end{bmatrix}^{-1}\right)$$

assuming vague prior knowledge (e.g. Proposition 12.3 pp 356-7 Hamilton (1994)). Starting values $\kappa^{*}_{0,m} = \kappa_{t_n}$, $\kappa^{*}_{-1,m} = \kappa_{t_n-1}$ are required for Step 3. If parameter error were ignored, Step 1 is omitted and we take $\theta^{*}_m = \hat{\theta}$, $\phi^{*}_m = \hat{\phi}$, $\sigma^{*}_m = \hat{\sigma}$ in Step 3. Obviously, the time series reduces to the random walk with drift parameter θ if the auto-regressive parameter ϕ is pre-set to zero. This approach is then consistent with that adopted by Cairns *et al.* (2006), who utilise a bi-variate random walk with drift to model a pair of period indices, controlled by prescribed age modulating functions.

4.8 A comparison of prediction intervals using Algorithm 1 and both versions of Algorithm 2 (*viz.* with and without allowance for the time series parameter error), is made in Fig 14. Here, we reproduce the biennial 97(02)07 life expectancy and 4% fixed rate annuity prediction intervals, by the cohort method, for individuals aged 65(05)80, subject to prior data truncation at age 60 (as in Fig 10), and compare these with the corresponding prediction intervals generated by both versions of Algorithm 2. For the detailed comparison in question, under Algorithm 2, the provision for the time series parameter fitting error is found to generate marginally wider prediction intervals for individuals aged 65, than without this provision. For individuals aged 70(05)80 this effect is negligible. Comparing like with like, Algorithm 1 generates wider prediction intervals than Algorithm 2 throughout. When the comparison exercise is extended to include a version of Algorithm 1 with additional provision for simulating time series parameter error, in parallel with Algorithm 2, it is found to have no material effect. Detailed evidence for this feature is reported in Section 4.9.

4.9 Setting to one side the provision for simulating time series parameter error (Algorithm 2, Step 1), Algorithms 1 & 2 differ essentially in the way that the time series forecasts are simulated: Algorithm 1, Step 2 compared to Algorithm 2, Step3. It is informative and simpler to compare these two steps, in the first instance, by pre-setting $\phi = 0$ and assuming a uni-variate random walk with drift θ. Then, under Algorithm 1, Step 2, the simulated forecast is

$$\kappa^{*}_{t_n+j,m} = \kappa_{t_n} + j\hat{\theta} + \sqrt{j}\hat{\sigma}z^{*}_m,\ \text{where } z^{*}_m \sim N(0,1), \qquad (3)$$

while, using successive substitution, the equivalent simulated forecast under Algorithm 2, Step 3, is

$$\kappa^*_{t_n+j,m} = \kappa_{t_n} + j\hat{\theta} + \hat{\sigma}\sum_{i=1}^{j} z^*_{i,m},\ \text{where } z^*_{i,m} \sim \mathrm{N}(0,1))\ \ i.i.d\ \left(\text{with } \sum_{i=1}^{j} z^*_{i,m} \sim N(0,j)\right),$$

the same as (3), when applied in this form. However Algorithm 2, Step 3 is not applied in this form, but rather in the form described in Algorithm 2. The difference that this makes is illustrated by plotting 40 simulated period time series predictions for the two cases, Fig 15. Thus using Algorithm 1, the familiar ridged structure of an $ARIMA(0,1,0)$ forecast is preserved under simulation, but not under Algorithm 2. Consequently, it would appear that this effect accounts for the reported differences in prediction intervals widths, when comparing the performance of two algorithms. Finally, restoring the provision for simulating parameter error in this case (Algorithm 2 Step 1), requires the posterior distributions

$$\frac{(n-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-1};\ \theta \sim N\left(\hat{\theta},\ \frac{\hat{\sigma}^2}{n-1}\right).$$

As a further illustration of this effect based on $ARIMA(0,1,0)$, life expectancy simulated prediction intervals, using Algorithm 1 and Algorithm 2 without provision for parameter error, for individuals aged 65 in 97(02)07, are tabulated below for comparison:

	2007	2005	2003	2001	1999	1997
A1	(18.35, 21.24)	(17.56, 20.23)	(16.84, 19.51)	(16.51, 19.30)	(15.78, 18.61)	(15.31, 18.26)
A2	(18.53, 20.92)	(17.75, 19.94)	(17.04, 19.22)	(16.72, 19.01)	(15.98, 18.31)	(15.52, 17.95)

Life expectancy 95% prediction intervals, individuals aged 65: without provision for parameter error

When the time series parameter error is allowed for, the results are as follows:

	2007	2005	2003	2001	1999	1997
A1	(18.33, 21.22)	(17.55, 20.22)	(16.83, 19.50)	(16.49, 19.30)	(15.75, 18.61)	(15.28, 18.26)
A2	(18.42, 21.11)	(17.66, 20.11)	(16.93, 19.40)	(16.57, 19.24)	(15.80, 18.56)	(15.29, 18.27)

Life expectancy 95% prediction intervals, individuals aged 65: with provision for parameter error

Again, the allowance for time series error, when applied to Algorithm 1, has no material effect.

For the $ARIMA(1,1,0)$ time series, the corresponding expressions read as follows Algorithm 1, Step 2:

$$\kappa^*_{t_n+j,m} = \kappa_{t_n} + \sum_{i=1}^{j}\left\{\hat{\mu} + \phi^i\left(y_{t_n} - \hat{\mu}\right)\right\} + \hat{\sigma}\sqrt{\sum_{i=1}^{j}\left(\sum_{k=0}^{i-1}\hat{\phi}^k\right)^2}.z^*_m,\ \ z^*_m \sim N(0,1),$$

Algorithm 2, Step 3:

$$\kappa^*_{t_n+j,m} = \kappa_{t_n} + \sum_{i=1}^{j}\left\{\hat{\mu} + \phi^i\left(y_{t_n} - \hat{\mu}\right)\right\} + \hat{\sigma}\sum_{i=1}^{j}\left\{\left(\sum_{k=0}^{i-1}\hat{\phi}^k\right)z^*_{j+1-i,m}\right\},\quad z^*_{j+1-i,m} \sim N(0,1)\ i.i.d.$$

$$\left(\text{with}\ \sum_{i=1}^{j}\left\{\left(\sum_{k=0}^{i-1}\phi^k\right)z^*_{j+1-i,m}\right\} \sim N\left(0, \sum_{i=1}^{j}\left(\sum_{k=0}^{i-1}\phi^k\right)^2\right)\right).$$

4.10 It has not been possible to compare the simulated prediction intervals (Fig 14) with the theoretically constructed prediction intervals (Denuit *et al.* (2009)): these require estimates for $\beta_x^{(1)}$ over the complete relevant age range, including $\beta_x^{(1)}$, $x \geq 90$, which cannot be estimated because of the paucity of data in this region.

4.11 Just as Algorithm 2 (Section 4.7) is a univariate adaptation of the Cairns *et al.* (2006) approach to prediction interval simulation, Algorithm 1 can be adapted to simulate prediction intervals for model structures involving multiple period indices (e.g. Renshaw and Haberman (2003), Cairns *et al* (2006)). Thus, for a vector of period indices $\boldsymbol{\kappa}_t$, modelled as a multivariate random walk with vector of drift parameters $\boldsymbol{\theta}$, variance-covariance matrix $\boldsymbol{\Omega}$ with Cholesky decomposition $\boldsymbol{CC'} = \boldsymbol{\Omega}$, so that

$$\boldsymbol{\kappa}_t = \boldsymbol{\theta} + \boldsymbol{\kappa}_{t-1} + \boldsymbol{\xi}_t,\quad \boldsymbol{\xi}_t \sim mulN\left(\boldsymbol{\theta}, \boldsymbol{\Omega}\right),\quad t = t_2, t_3, ..., t_n,$$

then Algorithm 1, Step 2 reads as follows:

$$\boldsymbol{\kappa}^*_{t_n+j,m} = \boldsymbol{\kappa}_{t_n} + j\boldsymbol{\theta} + \sqrt{j}.\boldsymbol{C}.\boldsymbol{z}^*_m$$

where $\boldsymbol{z}^*_m$ is a vector of randomly sampled standard normal deviates.

4.12 In comparing *APC* and *LC* predictions (Fig 10 with Fig 13), a priori, the capture of demonstrable cohort effects using *APC* (as shown by the residual plots of Fig 7) leads to more favourable predicted mortality outcomes than is otherwise the case using *LC*. As an illustration, in the first two rows of the table below, we tabulate the current 97(02)07 life expectancy point predictions, computed by cohort trajectory, for an individual aged 65, using *APC* and *LC* modelling respectively, subject to data truncation at age 60.

Epoch	2007	2005	2003	2001	1999	1997
APC ARIMA(1,1,0)	19.81	18.90	18.18	17.87	17.19	16.73
LC ARIMA(1,1,0)*	19.16	18.46	17.74	17.45	16.78	16.36
LC ARIMA(0,2,0)	20.01	19.83	18.48	19.57	17.48	18.12

predicted life expectancies, individuals aged 65 * truncated range

There does not appear to be any track record in the literature for modelling the *LC* period index as a second order integrated time series, such as $ARIMA(0,2,0)$, although there is evidence that, for the data set considered here, this process fits the complete time

series well (residual plots omitted). When this model is applied, the resulting life expectancy point predictions are as reported in the third and final row of the above table. Under this approach, the *LC* point predictions are unexpectedly greater than the matching *APC* point predictions, in addition to exhibiting irregularity over time.

4.13 Although the focus in this paper is on 'current' cohort trajectory based predictions at pensioner ages 65(05)80 throughout, this approach to age-period-cohort parameter predictions can be readily extended to individuals at younger ages, provided that the relevant data are available.

5. Summary

In this section, we summarise our conclusions and current views on mortality rate projecting using the *APC* modelling framework.

- The wider aspects of the discussion contained in Lee (2000), relating to *LC* modelling, are also relevant to *APC* modelling.
- The high degree of parameterisation of the model ensures the capture of data driven patterns for age, period, and cohort effects, irrespective of the size of the age range.
- The *APC* model fitting algorithm, perceived as the third stage in the fitting process (Section 2.5), is sufficiently flexible to allow control over both the sign and shape of the period index modulating factor $\beta_x^{(1)}$, as and when required, in order to ensure stability over time.
- Any lack of robustness in the parameter patterns (as the number of parameters is increased), may be attributed to the slow rate of convergence in the fitting algorithm, which itself, is indicative of the determination of a stationary point in a flat region of the likelihood (deviance) surface (Cairns *et al.* (2008)). This remains an issue for further investigation. In Section 3, the change in the shape of the index $\beta_x^{(1)}$ to emerge in the upper limits of the age range is managed by restricting $\beta_x^{(1)}$ to an established pattern, and we note that this results in a faster rate of convergence.
- The modelling of the period index as a second order integrated $ARIMA(p,2,q)$ process is to be avoided for the reason stated in Section 3.10 and the further reasons illustrated in Section 4.12.
- As a basic tenet when comparing indices of interest requiring mortality rate predictions, indices based on models which capture cohort effects in the data, (typically *APC*) should reflect this fact when compared with corresponding indices based on models which fail to capture cohort effects in the data, (typically *LC*). As a manifestation of this, we believe that, for England & Wales, the former set of predicted indices should reflect lower mortality than the latter set.
- There is emerging evidence (Denuit *et al.* (2009)) in support of the original Lee and Carter (1992) (Appendix B) finding, that the prediction error in the period index time series dominates the parametric estimation source of error when fitting

the *LC* structure. The *LC* period index time series was modelled originally as an $ARIMA(0,1,0)$ process but it has been argued in the literature that this result extends to other first order integrated processes. In this study, we have assumed that this result continues to apply on fitting the *APC* structure and time series modelling is restricted to the period index only, in common with the *LC* framework. On this basis, the prediction error in the period index time series, when fully represented, is additionally shown to dominate the parametric estimation source of error in the first order integrated *ARIMA* process (Fig 14, Section 4.9).

- Under the conditions described above (as in the previous bullet point), the construction of upper and lower theoretical prediction intervals is possible in circumstances where the data extend over the complete relevant age range (Denuit (2007), Denuit *et al.* (2009)).

References

Cairns, A.J.G., Blake, D., Dowd, K. 2006. A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. Journal of Risk and Insurance 73, 687-718.

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

Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D., Epstein, D., Khalaf-Allah, M. 2008. Mortality density forecasts: an analysis of six stochastic mortality models. Pensions Institute Discussion Paper PI-0801, Pensions Institute, Cass Business School.

Coale, A., Kisker, E.E. 1990. Defects in data on old age mortality in the United States: new procedures for calculating approximately accurate mortality schedules and life tables at the highest ages. Asian and Pacific Population Forum, 4, 1-31.

Delwarde, A., Denuit, M., Eilers, P. 2007. Smoothing the Lee-Carter and Poisson log-bilinear models for mortality forecasting: a penalised log-likelihood approach. Statistical Modelling 7, 385-401.

Denuit, M. 2007. Distribution of random future life expectancies in log-bilinear mortality projection models. Lifetime Data Analysis, 13, 381-397.

Denuit, M., Goderniaux, A-C., 2005. Closing and projecting life tables using log-linear models. Bulletin of the Swiss Association of Actuaries (1), 29-48.

Denuit, M., Renshaw, A.E., Haberman, S. 2009. Comonotonic approximations to quantiles of life annuity conditional expected present value: extensions to general ARIMA models: simulation methods. Astin Bulletin. [to appear].

Hamilton, J.D. 1994. Times Series Analysis. Princeton University Press, Princeton, New Jersey.

Janssen, F., Kunst, A. 2007. The choice among past trends as a basis for prediction of future trends in old-age mortality. Population Studies 61, 315-326.

Lee, R. 2000. The Lee-Carter method for forecasting mortality, with various extensions and applications. North American Actuarial Journal 4(1) 80-93.

Lee, R., Carter, L. 1992. Modelling and forecasting the time series of US mortality. Journal of the American Statistical Association (with discussion) 87, 659-671.

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

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

Renshaw, A.E., Haberman, S. 2008. On simulation-based approaches to risk measurement in mortality with specific reference to Poisson Lee-Carter modelling. Insurance: Mathematics and Economics 42, 797-816.

*** Revision Checklist**

Ms. Ref. No.: IME-D-09-00123
Title: On age-period-cohort parametric mortality rate projections
Insurance: Mathematics and Economics

We thank the reviewer for his/her helpful and thoughtful comments. Our response to the reviewers' comments (in reverse order) is as follows:

- The details of the 2nd and, additionally, the 8th reference have been updated.
- The paper has been shortened, as suggested, by making reductions to Sections 2.2, 2.3, 2.4 and 2.6, without detriment to the overall narrative. We have also shortened a couple of the displays when presenting equations. However, we believe that it is desirable to retain the display setting out the iterative fitting algorithm (Section 2.3) in order to achieve a self-contained narrative and highlight how the parameter constraints are applied.
- A separate investigation involving a detailed comparison with predictions using the various Cairns *et al* parametric models, and other such models, has reached an advanced stage of preparation. However, because of the size of this project, involving a number of additional considerations not reported here, including the establishment of a basis for conducting such a comparison, it is our intention to present these findings as a separate paper. As a necessary prelude to this work, it was necessary to resolve an apparent anomaly arising from the different Denuit *et al* and Cairns *et al* approaches to the construction of prediction intervals (fan charts), which we address in Section 4 of this paper.
- We are appreciative of the reviewer for drawing our attention to the issue of "optimal fitting period", which has much relevance. In order to deal with this issue more fully and explicitly, we have expanded Section 3.10 and redrafted and expanded Section 4.2. There is also a deletion in Section 4.12 as a consequence.