City Research Online

City St George's, University of London

Citation: Renshaw, A. E. & Haberman, S. (2007). On simulation-based approaches to risk measurement in mortality with specific reference to Poisson Lee-Carter modelling (Actuarial Research Paper No. 181). London, UK: Faculty of Actuarial Science & Insurance, City University London.

This is the unspecified version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/2313/

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

Cass Business School
City of London

Cass means business

Faculty of Actuarial Science and Insurance

On Simulation-based Approaches to Risk Measurement in Mortality with Specific Reference to Poisson Lee-Carter Modelling.

Arthur Renshaw and
Steven Haberman.

Actuarial Research Paper No. 181

May 2007

ISBN 978-1-905752-10-2

Cass Business School
106 Bunhill Row
London EC1Y 8TZ
T +44 (0)20 7040 8470
www.cass.city.ac.uk

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

On simulation-based approaches to risk measurement in mortality with specific reference to Poisson Lee-Carter modelling.

A.E. Renshaw, S. Haberman

Cass Business School, City University, London, EC1Y 8TZ, UK

Abstract

This paper provides a comparative study of simulation strategies for assessing risk in mortality rate predictions and associated estimates of life expectancy and annuity values in both period and cohort frameworks.

Keywords: Poisson modelling; Over-dispersion; Joint modelling; Negative binomial modelling; Mortality projections; Mortality statistics; Simulated risk

1. Introduction

The trend in many countries to reform public sector pension provision and encourage more private sector provision plus the widespread shift within private sector pension provision towards more defined contribution schemes at the expense of defined benefit schemes both mean, that, in the future, we can expect an increased demand for post retirement annuity-type products.

From the viewpoint of insurance companies (and other providers) selling annuity polices involves risk because of their exposure to uncertain future interest rates and mortality rates. In this paper, we consider the second source of risk and the need for effective risk measurement (and risk management) techniques.

Simulation techniques have been suggested in the literature as a means of measuring risk, when modelling dynamic mortality rates and their impact on future predictions of life expectancy and annuity values, because of the general intractability of applying theoretical methods. We give further consideration to three such simulation strategies (denoted A, B, C) which have been applied recently to Poisson Lee-Carter (bilinear) modelling and extrapolation. Strategy A (semi parametric bootstrap) is suggested in Brouhns *et al.* (2002) and illustrated (and compared with Strategy B) in Brouhns *et al.* (2005). Strategy B (parametric Monte-Carlo) is described and illustrated in Brouhns *et al.* (2002). Strategy C (residual bootstrap) is a variant on the method described and illustrated in Koissi *et al.* (2006).

In addition to gaining further insight into the relative merits of these competing simulation strategies, we expand the methodology in order to allow for possible over-dispersion: modelled as a second moment property which, therefore, has implications for risk measurement. The scene is set and illustrated in Section 2 with reference to a simple single parameter model, and the resulting simulations compared with maximum likelihood theory. The strategies are further developed and illustrated in Sections 3 & 4 with reference to the Gompertz law and to Makeham's second law of mortality respectively. Here static age-specific life expectancy and annuity prediction limits, simulated under the different strategies, are compared. Comparison with maximum likelihood theory is again possible. Dynamic Poisson Lee-Carter based predictions are illustrated in Section 5. In addition to representing time as a parameterised factor, we reconsider the possibility of formulating time as a linear

variable (Renshaw and Haberman (2003a & c)) and investigate the implications for risk measurement. Additional consideration is given to aspects of parametric simulation Strategy B in Section 6, to the impact of extra provision for variable dispersion through joint modelling on simulation Strategies A and C in Section 7, and to the impact of switching to negative binomial modelling on simulation Strategies A and C in Section 8. A comparative assessment is made with a non-simulation method of assessing risk in life expectancy predictions (Denuit 2006) in Section 9. Section 10 provides a discussion of the results and presents some concluding comments.

We conventionally refer to interval estimates as confidence intervals (CIs), unless extrapolation is involved in their construction, as with statistics computed by fixed cohort (Sections 5 to 9), when we refer to prediction limits.

2. Single parameter model

We illustrate the methodology in this section by considering a single parameter model in the context of a well-known data set. In an investigation to test whether the impact points of flying-bombs during World War II tended to be grouped in clusters, Clarke (1946) divides an area of 144 sq. km. in south London into $n = 576$ squares of ¼ sq. km., and makes a count of the number of impact points in each square. The (grouped) results are as summarised in the first two rows below:

Number of impacts per square d_i	0	1	2	3	4	7
Number of squares or frequency f_i	229	211	93	35	7	1
Predicted number of squares $\hat{f}_i$	227	211	99	31	7	2
Deviance residual r_i	-1.3655	0.0693	0.9579	1.6962	2.3486	4.0111

Modelling the number of impacts D_i per square i (= 1, 2,…, n) as independent identically distributed Poisson random variables $D_i \sim \text{Poi}(e_i\mu)$, with an assumed single unit of exposure to the risk of impact, the same for all squares $(e_i = 1 \,\forall\, i)$, gives

$$\hat{\mu} = \bar{d} = \sum_i d_i \Big/ n = 0.9323 .$$

The frequency distribution predicted by the fitted model and recorded (to the nearest whole number) in the third row of the above table, is indicative of a good fit.

The associated (classical) inference is standard. Thus, μ is estimated by maximising the log likelihood

$$\log L = \sum_i \left(d_i \log\mu - e_i\mu\right) + \text{const.}$$

so that trivially the fitted values $\hat{d}_i = \hat{\mu} = \bar{d}$, with (minimum) variance

$$-\left(\mathrm{E}\frac{\partial^2 \log L}{\partial \mu^2}\right)^{-1}_{\mu=\hat{\mu}} = \frac{\hat{\mu}}{n}$$, and large sample CI $\hat{\mu} \pm z_{1-\alpha'/2}\sqrt{\frac{\hat{\mu}}{n}}$

where $z_{1-\alpha'/2}$ is the usual standard normal deviate.

Under the alternative canonical parameterisation (log link) $\log\mu = \alpha$, which forms the basis of more complex parameterised structures and the second of the simulation strategies considered below. Hence, $\hat{\alpha} = \log \bar{d}$ with variance

$$-\left(\mathrm{E}\frac{\partial^2 \log L}{\partial \alpha^2}\right)^{-1}_{\alpha=\hat{\alpha}} = \frac{1}{n \exp\hat{\alpha}}$$

which is the inverse $\mathcal{I}^{-1}$ of the information matrix $\mathcal{I}$.

We focus on the construction of CIs for μ by simulation and consider three possible strategies. At each simulation j (= 1, 2, …, N)

A: simulate responses $d_i^{(j)}\ \forall\, i$ by sampling $\mathrm{Poi}(\hat{d}_i)$
and compute $\mu^{(j)}$ by fitting $D_i^{(j)} \sim \mathrm{Poi}(\mu^{(j)})$.

B: randomly generate a standard normal deviate $\varepsilon^{(j)} \sim \mathrm{N}(0,1)$
to simulate the parameter as $\alpha^{(j)} = \hat{\alpha} + \sqrt{\phi}\sqrt{\frac{1}{n\log\hat{\alpha}}}.\varepsilon^{(j)}$
and compute $\mu^{(j)} = \exp\alpha^{(j)}$. Here, ϕ is the optional scale parameter.

C: simulate responses $d_i^{(j)}\ \forall\, i$ by sampling $\{d_i\}$ with replacement
and compute $\mu^{(j)}$ by fitting $D_i^{(j)} \sim \mathrm{Poi}(\mu^{(j)})$.

After N = 5,000 (say) such simulations using each strategy in turn, compute the $\frac{\alpha'}{2}100N$ and $(1-\frac{\alpha'}{2})100N$ order statistics of $\{\mu^{(j)}\}$ to give a two-sided $(1-\alpha')100\%$ CI for μ, together with the median $\tilde{\mu}$ to compare with $\hat{\mu}$.

For comparison purposes, we record both the theoretical based 95% CI for μ, together with the simulated 95% CIs and median based on 5,000 simulations, under each strategy in the table below and note the exceptionally close agreement between corresponding figures.

m.l.e. 95% C.I.	*Simulation Strategy* A	*Simulation Strategy* B	*Simulation Strategy* C
$\hat{\mu} = 0.9323$	$\tilde{\mu} = 0.9323$	$\tilde{\mu} = 0.9325$	$\tilde{\mu} = 0.9323$
(0.853, 1.011)	(0.856, 1.012)	(0.857, 1.016)	(0.854, 1.016)

Here, under simulation Strategy A, $\hat{d}_i = \hat{\mu}$, while as a consequence of the final fitting stage in Strategies A and C, $\mu^{(j)} = \sum_i d_i^{(j)} \Big/ n$.

Under Strategy B, the parameter α is simulated on scaling the standard normal deviate by the standard error of $\hat{\alpha}$, before centring on $\hat{\alpha}$. In addition it is possible to allow for dispersion in the data under this strategy through the incorporation of an optional scale parameter ϕ (>0). The theoretical justification for including ϕ is given in Section 5. For the Poisson distribution (without dispersion), we set $\phi = 1$.

Strategy C is the basic bootstrap for targeting the mean, described in Chapter 2 of Efron and Tibshirani (1993): subject to a different (but equivalent) treatment of the simulated $\{\mu^{(j)}\}$, using order statistics. It is necessary in this paper for us to take a different perspective of this strategy by sampling residuals, in order to be able to manage more general parametric structures. We follow Chapter 9 (and Section 9.4 in particular) of Efron and Tibshirani (1993) on bootstrapping residuals in regression models, making the necessary changes to the subsequent mapping of bootstrap residuals into bootstrap responses. These changes are needed in order to match with the Poisson error assumption, as opposed to the additive error structure, implied by equation (9.26) of Efron and Tibshirani (1993). Then the residual bootstrap version of Strategy C reads as follow:

C: simulate residuals $r_i^{(j)}\ \forall\, i$ by sampling $\{r_i\}$ with replacement, then
map the bootstrap residuals to responses $r_i^{(j)} \mapsto d_i^{(j)}\ \forall\, i$
and compute $\mu^{(j)}$ by fitting $D_i^{(j)} \sim \text{Poi}(\mu^{(j)})$.

Bootstrap deviance residuals are mapped by solving the relationship

$$r_i = \text{sign}\left(d_i - \hat{d}_i\right)\sqrt{2}\sqrt{d_i \log\left(\frac{d_i}{\hat{d}_i}\right) - \left(d_i - \hat{d}_i\right)} \tag{1}$$

for d_i when $r_i = r_i^{(j)}$. Therefore, suppressing the suffix *i* for clarity of notation, and using the prefix * (instead of *j*) to denote bootstrap values, this implies that we require the appropriate root d^* of

$$\text{g}(d) = d \log d - d\left(1 + \hat{a}\right) - \hat{c}^*,\ d \geq 0 \tag{2}$$

where $\hat{a} = \log \hat{d},\ \hat{c}^* = \dfrac{r^{*2}}{2} - \hat{d}$, when mapping $r^* \mapsto d^*$. The derivatives of g

$$\text{g}'(d) = \log d - \hat{a},\quad g''(d) = \frac{1}{d} > 0\ \ \forall\, d > 0$$

imply that $d = \hat{d}$ is a minimum and that the graph of g(*d*) vs. *d* (*d* >0) is concave. There are either one or two roots in this range. The required root is determined by the sign of the residual r^* and lies to the right of the minimum ($d > \hat{d}$) when $r^* > 0$, and to the left of the minimum ($d < \hat{d}$) when $r^* < 0$. The root is readily determined by Newton-Raphson using starting values for *d*, which accord with the above constraints,

within the limits of the domain $\{d : d \geq 0\}$. Note the two special cases $r^* < 0$, $\hat{c}^* > 0$ and $r^* = -\sqrt{2\hat{d}}$, $\hat{c}^* = 0$, each of which implies $d^* = 0$ is the required root.

The process is illustrated in Fig.1, with each frame corresponding to a different entry in the tabulated data above. Restoring the suffix *i*, the bootstrap residual mapping, $r_i^* \mapsto d_i^*$, takes the special form $r_i^* \mapsto d_i$, which involves the mapping of bootstrap residuals to matching observations as a consequence of the simplicity of the model structure based on a single location parameter μ.

3. A two parameter model (Gompertz s law)

In order to take the development further, we take a well-known data set from the mortality literature and apply a model, which can be presented in a linear form. In a study to graduate the UK pensioners' widows' 1979-82 mortality experience Forfar *et al.* (1988) apply Gompertz's law. A short extract of the raw data (d_x, e_x) comprising the respective number of deaths and matching exposures to the risk of death at age *x*, together with the graduated force of mortality $\hat{\mu}_x$, the expected number of deaths $\hat{d}_x$ predicted by the model, in addition to the deviance residuals r_x, is recorded below:

x	e_x	$\hat{\mu}_x$	d_x	$\hat{d}_x$	r_x
20	4.0	.00038219	0	0.00	-0.0553
30	36.0	.00090617	0	0.03	-0.2554
40	115.5	.00214854	0	0.25	-0.7045
50	378.5	.00509423	3	1.93	0.7132
60	1029.0	.01207848	14	12.43	0.4368
70	941.0	.02863823	21	26.95	-1.1925
80	323.5	.06790162	25	21.97	0.6332
90	30.5	.16099560	6	4.91	0.4750
100	1.0	.38172269	0	0.38	-0.8738

The complete table (Forfar *et al.* (1988) Table 15.5 (& 15.1)) is given by individual year of age from 17 to 108, and we note that the bulk of the exposure lies in the age range 45 to 90.

Graduation proceeds by modelling the number of deaths as independent Poisson responses $D_x \sim \text{Poi}(e_x\mu_x)$ where $\log\mu_x = \alpha + \beta x$ (which is a re-parameterisation of Gompertz's law $\mu_x = Bc^x$). The parameters $\boldsymbol{\theta} = (\alpha, \beta)^{\mathrm{T}}$ are estimated first using numerical methods in order to maximise the log likelihood

$$\log L = \sum_x \omega_x \left(d_x \log\mu_x - e_x\mu_x\right) + \text{const} \qquad (3)$$

followed by the computation of

$$\hat{\mu}_x = \exp(\hat{\alpha} + \hat{\beta}x), \ \hat{d}_x = e_x\hat{\mu}_x .$$

Zero/one weights ω_x are included in the formulation so that we can deal with empty/non-empty data cells. The variance-covariance matrix of the parameter estimates is the inverse $\mathcal{I}^{-1}$ of the information matrix

$$\mathcal{I} = \begin{bmatrix} \sum_x \omega_x \hat{d}_x & \sum_x \omega_x x \hat{d}_x \\ \sum_x \omega_x x \hat{d}_x & \sum_x \omega_x x^2 \hat{d}_x \end{bmatrix},$$

comprising the (expectations of the) second order partial derivatives of $-\log L$ with respect to the two parameters.

We focus on the simulation of CIs for the age-specific force of mortality μ_x, life expectancy $e(x)$ and level immediate annuity $a(x)$ with given discount factor v (and fixed rate interest), where

$$\hat{e}(x) = \frac{\sum_{i\geq 0} l_{x+i}(1-\frac{1}{2}\hat{q}_{x+i})}{l_x}, \quad \hat{a}(x) = \frac{\sum_{i\geq 1} l_{x+i} v^i}{l_x}$$

and

$$\hat{q}_x \approx 1-\exp(-\hat{\mu}_x), \quad l_{x+1} = (1-\hat{q}_x) l_x .$$

We use the same three simulation strategies, which are generalised to read as follows: At each simulation j

A: simulate responses $d_x^{(j)}$ by sampling $\text{Poi}(\hat{d}_x), \forall\, x = x_1, x_2, ..., x_n$,
then compute $\mu_x^{(j)}$ by fitting $D_x^{(j)} \sim \text{Poi}(e_x \mu_x^{(j)}) \;\forall\, x$
before computing the statistics of interest.

B: randomly generate a pair of N(0,1) deviates $\boldsymbol{\varepsilon}^{(j)} = (\varepsilon_1^{(j)}, \varepsilon_2^{(j)})^{\mathrm{T}}$
to simulate parameters according to $\boldsymbol{\theta}^{(j)} = \hat{\boldsymbol{\theta}} + \sqrt{\phi}\; \mathcal{C}\boldsymbol{\varepsilon}^{(j)}$
where $\mathcal{C}$ is the Cholesky factorisation ('square root') matrix of the variance-covariance matrix $\mathcal{I}^{-1}$, ϕ is the optional scale parameter
before computing $\mu_x^{(j)}$ (here $\mu_x^{(j)} = \exp(\alpha^{(j)} + \beta^{(j)} x) \;\forall\, x$)
and the statistics of interest.

C: simulate residuals $r_x^{(j)} \;\forall\, x$ by sampling $\{r_x\}$ with replacement,
map the bootstrap residuals to responses $r_x^{(j)} \mapsto d_x^{(j)} \;\forall\, x$
then compute $\mu_x^{(j)}$ by fitting $D_x^{(j)} \sim \text{Poi}(e_x \mu_x^{(j)})$
before computing the statistics of interest.

Subject to centring x at age 70 and scaling (dividing) by 50, the parameter estimates, variance-covariance matrix and Cholesky factorisation are as follows:

$$\hat{\alpha} = -3.5530,\ \hat{\beta} = 4.3166$$

$$\mathcal{I}^{-1} = \begin{bmatrix} 0.001539 & -0.001907 \\ -0.001907 & 0.038653 \end{bmatrix},\ \mathcal{C} = \begin{bmatrix} 0.03923 & 0 \\ -0.04862 & 0.19050 \end{bmatrix}.$$

The simulated median and two-sided 95% (percentile based) CIs for the force of mortality, life expectancy and 4% fixed rate annuity at ages (x = 45, 65, 75), using all three strategies (with N = 5,000) are depicted in Fig 2. In addition, the age-specific force of mortality, life expectancy and annuity value based on model estimates (m.l.e.) are also depicted for comparison. Here, the 95% CI for the force of mortality is based on the close approximation

$$\mathrm{Var}(\hat{\mu}_x) \approx \mathrm{Var}(\hat{\eta}_x)\exp\{2\mathrm{E}(\hat{\eta}_x)\} \text{ where } \hat{\eta}_x = \hat{\alpha} + \hat{\beta}x,$$

and the 95% CIs for life expectancy and annuity values are based on the approximate expressions for the respective variances

$$\mathrm{Var}\{\hat{e}(x)\} \approx \sum_{j\geq 0}\left(\frac{\hat{q}_{x+j}}{\hat{p}_{x+j}e^{i}_{x+j}}\right)\left(\frac{l_{x+j}\hat{e}(x+j)}{l_x}\right)^2 \tag{4}$$

$$\mathrm{Var}\{\hat{a}(x)\} \approx \sum_{j\geq 0}\left(\frac{\hat{q}_{x+j}}{\hat{p}_{x+j}e^{i}_{x+j}}\right)\left(\frac{l_{x+j}v^{j}\hat{a}(x+j)}{l_x}\right)^2 \tag{5}$$

with initial exposures $e^{i}_{x} \approx e_x + \frac{1}{2}d_x$ (e.g. Benjamin and Pollard (1980), Chapter 17).

We note the following:

1. The simulated histograms (not shown) underpinning the CIs are essentially symmetric throughout. In addition, the upper and lower simulated percentiles displayed are effectively the same as the 95% (two-sided) confidence limits calculated using the simulated mean and variance.
2. Within each frame, the degree of vertical alignment of the central measures (medians and maximum likelihood point estimate) indicates the extent of the agreement between the respective first moment estimated targets. In this respect, Strategies A and B are closely aligned with each other and with the maximum likelihood estimate in all cases.
3. The reason for the inconsistent alignment under Strategy C is the presence of bias in the residuals due to the paucity of exposure at ages (with zero deaths) below age 45: this gives rise to a run of small negative residuals which is unrepresentative of the randomly distributed residuals pertaining to the remainder of the age range.
4. Within each frame, the widths of the prediction intervals are essentially the same under Strategies A and B and are in agreement with the width of the m.l.e. CI in the case of the force of mortality (1st column). The m.l.e. CIs are consistently wider than their simulated counterparts for life expectancy (2nd column) and the annuity (3rd column). Under Strategy C the intervals are marginally narrower throughout.

5. Frames depicting the force of mortality (1st column) are not drawn to the same scale for practical reasons. Allowing for the relative magnitude of μ_x, the coefficient of variation (relative error) is observed to decrease with increasing age, on the basis of the following representative (Strategy A) values:

Age	45	65	75
Coefficient of variation	0.114	0.048	0.040

6. Frames depicting life expectancy (2nd column) and the fixed rate annuity using a 4% interest rate (3rd column) are presented on the same respective scales and may be compared column-wise. On this evidence, there is no obvious emerging pattern in the interval widths with age.

Simulation Strategies A and C are conducted with due attention to the vector of weights ω_x, ensuring that any empty data cells are preserved as such, at each stage of the simulation process.

Under Strategy C, the mapping of bootstrap residuals $r_x^{(j)} \mapsto d_x^{(j)}$ is as described in Section 2 with *x* replacing the suffix *i*, resulting in $d_x^{(j)} \neq d_x$ in general. We remark that this mapping is invariant to the scaling of residuals (by ϕ). Thus simulation Strategy B would appear to be the only one of the three strategies considered, which is sensitive to the inclusion of a scale parameter. In this example, we again set $\phi = 1$, but we illustrate and discuss the effects of both estimating and generalising ϕ as a function of age *x*, in later sections.

4. A four parameter non-linear model (Makeham s second law)

In this section, we illustrate the approach for a model, which is non-linear and involves 4 parameters. In a study to graduate the UK male assured lives 1979-82 mortality experience (based on policy counts and policy duration 5+ years), Forfar *et al.* (1988) apply the enhanced Makeham formula

$$\mu_x = \alpha_0 + \alpha_1 x + \exp(\beta_0 + \beta_1 x) .$$

Using the same notation as Section 3, an extract of the data (d_x, e_x) adjusted before modelling to allow for duplicated policies, together with the model estimates, are as follows:

x	e_x	$\hat{\mu}_x$	d_x	$\hat{d}_x$	r_x
15	238.0	.0011916	1.00	0.28	1.0429
25	85112.7	.0006721	61.86	57.20	0.6081
35	363515.5	.0006714	278.06	244.07	2.1280
45	316713.0	.0019716	607.78	624.43	-0.6692
55	259981.3	.0065332	1713.17	1698.51	0.3553
65	73385.6	.0192716	1369.34	1414.26	-1.2007
75	14496.5	.0525103	801.39	761.22	1.4434
85	2981.0	.1371462	386.52	408.84	0.6698

The adjustment to the numbers of deaths and matching exposures by dividing both quantities by the so-called variance ratios based on duplicated policy counts, is as described in Forfar *et al.* (1988) and need not directly concern us here. The complete table (Forfar *et al.* (1988) Table 17.9 with additional reference to Tables 17.6 & 17.8) is given by individual year of age from 10 to 90. We note that there is a relative scarcity of exposure below age 20.

The parameters $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_0, \beta_1)^{\mathrm{T}}$ are estimated by maximising the Poisson log likelihood (3) and $\hat{\mu}_x$, $\hat{d}_x$ and r_x computed subsequently. One possible method of fitting through linearisation is by declaring a GLM with

Poisson responses d_x/e_x, expectation μ_x, weights $\omega_x = e_x$, identity link

and conducting the iterative linear fitting routine:

Set the starter value $\rho_i = \rho_1 = 0.0005$ (say)

↓

Estimate the parameters $(\alpha_0, \alpha_1, \beta, \gamma)$ by fitting

$$\mu_x = \alpha_0 + \alpha_1 x + \beta e^{\rho_i x} + \gamma x e^{\rho_i x}$$

↓

Compute $\rho_{i+1} = \rho_i + \hat{\gamma}/\hat{\beta}$ to update ρ_i

↓

Stop when $\{\rho_i\}$ converges

Here ρ is an auxiliary parameter (assumed known, so that the expression for μ_x is linear in the other four parameters), while the iterative routine is based on a linearisation method for fitting non-linear parameters in the covariates of a GLM (Section 11.4, McCullagh and Nelder (1989); Section 6, Renshaw (1991)). Convergence is rapid with $\hat{\alpha}_0, \hat{\alpha}_1$ as given in the limit, and $\hat{\beta}_0 = \log\hat{\beta}, \hat{\beta}_1 = \rho_i, (\hat{\gamma} = 0)$ in the limit. Details of the symmetric information matrix are as follows:

$$\mathcal{J} = \begin{bmatrix} \sum_x \frac{e_x}{\hat{\mu}_x} & \sum_x \frac{e_x}{\hat{\mu}_x} x & \sum_x \frac{e_x}{\hat{\mu}_x} \exp(\hat{\beta}_0 + \hat{\beta}_1 x) & \sum_x \frac{e_x}{\hat{\mu}_x} x \exp(\hat{\beta}_0 + \hat{\beta}_1 x) \\ & \sum_x \frac{e_x}{\hat{\mu}_x} x^2 & \sum_x \frac{e_x}{\hat{\mu}_x} x \exp(\hat{\beta}_0 + \hat{\beta}_1 x) & \sum_x \frac{e_x}{\hat{\mu}_x} x^2 \exp(\hat{\beta}_0 + \hat{\beta}_1 x) \\ & & \sum_x \frac{e_x}{\hat{\mu}_x} \exp 2(\hat{\beta}_0 + \hat{\beta}_1 x) & \sum_x \frac{e_x}{\hat{\mu}_x} x \exp 2(\hat{\beta}_0 + \hat{\beta}_1 x) \\ & & & \sum_x \frac{e_x}{\hat{\mu}_x} x^2 \exp 2(\hat{\beta}_0 + \hat{\beta}_1 x) \end{bmatrix}$$

Subject to the same centring and scaling of x (Section 3), the parameter estimates and Cholesky factorisation of the variance-covariance matrix $\mathcal{J}^{-1}$ are as follows:

$$\hat{\alpha}_0 = -0.003788,\ \hat{\alpha}_1 = -0.004319,\ \hat{\beta}_0 = -3.329023,\ \hat{\beta}_1 = 4.595709$$

$$\mathcal{C} = \begin{bmatrix} 0.00022455 & 0 & 0 & 0 \\ 0.00024399 & 0.00002625 & 0 & 0 \\ -0.00708817 & -0.00106814 & 0.00476708 & 0 \\ 0.03829598 & -0.01309409 & 0.00927452 & 0.00844291 \end{bmatrix}$$

The equivalent simulations (Section 3), depicting the median and 95% (percentile) CIs for the force of mortality, life expectancy and 4% fixed rate annuity (x = 45, 65, 75), using all three strategies, plus maximum likelihood estimates, (including CIs for life expectancy and annuities using (4) and (5)), are presented in Fig 3. We note the following:

1. As in Section 3, we have not shown any of the simulated histograms, which are symmetrical throughout: a feature reflected in the symmetry of the CIs.
2. Within each frame, the close vertical alignment (including that under Strategy C), of the central measures (medians and maximum likelihood estimate).
3. Within each frame, prediction intervals under Strategies A and B are essentially the same, and marginally narrower than those under Strategy C. The m.l.e. CIs are marginally wider still, in general.
4. The comparative narrowness of the CIs compared with their counterparts in the UK pensioners widow study (Fig 2). This is further reflected in a typical comparison of the coefficient of variation for μ_x under the respective studies:

Age	45	65	75
Male assured lives	0.0092	0.0052	0.0069
Pensioners widows	0.114	0.048	0.040

Force of mortality: coefficient of variation (Strategy A)

5. As before, frames depicting life expectancy (2nd column) and the 4% fixed rate annuity (3rd column) are presented on the same respective scales and can be compared column-wise, with no evidence of an emerging age pattern in the interval widths.

5. Multiple parameter non-linear age-period models

This study features the UK male pensioner 1983-2003 mortality experience illustrated in Fig 4. Here we have plotted the log crude mortality rates (continuous profiles) against period $t = t_1, t_2, ..., t_n$ for selected ages in the range $x = x_1, x_2, ..., x_k$, using data (d_{xt}, e_{xt}) comprising the respective number of deaths and matching exposures to the risk of death at age x in calendar year t. Ages range from 51 to 104 with roughly 95% of the total exposure in the age range 62-89 and 95% of the total deaths in the range 66-93. The rectangular data set is an updated version, with expanded age range, of the UK male pensioner 1983-94 experience modelled previously (Renshaw and Haberman (2003b)). We note that about 5% of the data cells are empty.

The numbers of deaths are modelled as independent Poisson responses $D_{xt} \sim \text{Poi}(e_{xt}\mu_{xt})$ when targeting the force of mortality μ_{xt}. We are interested in the following parametric structures:

$$\text{LC}: \log\mu_{xt} = \alpha_x + \beta_x\kappa_t;\ \sum_x \beta_x = 1,\ \kappa_{t_n} = 0, \tag{6}$$

$$\text{LP}: \log\mu_{xt} = \alpha_x + \beta_x(t_n - t) + \gamma_t(t_n - t);\ \gamma_{t_1} = \gamma_{t_n} = 0, \sum_x \beta_x = 1. \tag{7}$$

The non-linear (LC) structure is that generally attributed to Lee and Carter (1992) (subject to a change in their usual parameter constraints: see Section 6), while the linear Poisson (LP) structure has been suggested and used previously in a similar context by Renshaw and Haberman (2003a). The structures are characterised by writing

$$\mu_{xt} = \exp(\alpha_x)\mathrm{F}(x,t);\ \mathrm{F}(x,t_n) = 1. \tag{8}$$

In the spirit of historical CMI Bureau practice (e.g. CMI (1999)), we interpret (8) as the product of a static life-table $\exp(\alpha_x)$ summarising the over-all main age effects with regard to the calendar year t_n, and a mortality reduction factor *F* encapsulating age-specific dynamic adjustments.

The structures are fitted by maximising the log likelihood expression (3) in which *x* is replaced by *x,t* (or minimising the current model deviance $-2\log(L/L_f)$ where L_f is the likelihood of the full or saturated model: characterised by equating the fitted and actual numbers of deaths). We use an iterative method described in Renshaw and Haberman (2006) when fitting LC and the software package GLIM (Francis *et al.* (1993)) when fitting LP. The model fitting of LC is discussed further in Section 6. For LP model fitting, the constraints $\gamma_{t_1} = \gamma_{t_n} = 0$ apply and the resulting $\hat{\beta}_x$ are scaled (together with the compensating scaling of the multiplicative terms $(t_n - t)$), to comply with the auxiliary constraint $\sum_x \beta_x = 1$. This approach to LP model fitting, in which both components of μ_{xt} (equation (8)) are fitted simultaneously, differs from the two-stage model fitting approach discussed in Renshaw and Haberman (2003a). The fitted log mortality rates $\log\hat{\mu}_{xt}$ under both structures have been superimposed in Fig 4. The underpinning parametric structures are depicted in Fig 5, in which we have superimposed the respective parameter estimates under both structures (LH frames) and plotted their differences in the matching RH frames: the imposition of equivalent parameter constraints under LC and LP modelling provides the basis for parameter comparisons. (We refer loosely to $\kappa_t = t_n - t$ under LP modelling as parameters here).

The extra term involving γ_t under LP modelling ensures that the annual actual and expected total deaths are the same, (in addition to improving the quality of the fit). The estimates of γ_t and $\gamma_t(t_n - t)$ are depicted in Fig 6a, together with the differences between the actual and expected total deaths in each year under both models, in Fig 6b. It is desirable that such differences are zero, as in the construction of static life tables, while special provision for achieving this by adjusting the $\hat{\kappa}_t$s after fitting, is advocated in the original LC model (Lee and Carter (1992)). However, such differences are much smaller in practice under the Poisson formulation of LC,

compared with the original Gaussian formulation, while pattern-free differences are indicative of a good fit. The small reported deviations from zero under LP modelling (RH frame Fig 6b) are directly attributable to the paucity of exposure at ages 103 and particularly 104. Residual plots under LC modelling are similar to those under LP modelling (Fig 7) and hence, have not been presented. Both sets of residual plots are therefore supportive of the respective model structures. In particular, the residual plot against year of birth, together with further diagnostic plots designed to detect residual cohort effects (not shown), fail to identify any obvious residual systematic cohort effect.

Mortality rates are extrapolated according to

$$\dot{\mu}_{x,t_n+s} = \exp(\hat{\alpha}_x)\mathrm{F}(x,t_n+s),\ s>0$$

where

$$\mathrm{LC}:\ \mathrm{F}(x,t_n+s) = \exp\left(\hat{\theta}\hat{\beta}_x s\right),\ s>0;\ \hat{\theta} = -\hat{\kappa}_{t_1}/(t_n - t_1)$$

$$\mathrm{LP}:\ \mathrm{F}(x,t_n+s) = \exp(-\hat{\beta}_x s),\ s>0$$

when forecasts are generated by applying a random walk with drift parameter θ to the time series $\{\hat{\kappa}_t\}$ under LC modelling. (We note that, when the random walk with drift is the preferred time series, selected here by testing for the best ARIMA model for $\{\hat{\kappa}_t\}$, forecasts are generated by extrapolating the straight line joining the two extremes $(t_1, \hat{\kappa}_{t_1})$ and $(t_n, \hat{\kappa}_{t_n})$ of the time series, with $\hat{\kappa}_{t_n} = 0$). Thus, under this special case, it is only necessary to resort to the use of a standard time series package in order to establish the appropriateness of the random walk with drift: refitting the time series to generate forecasts at each simulation is redundant. Under LP modelling: referring to equation (7) and Fig 6a, we extrapolate $\hat{\beta}_x(t_n - t)$ while setting the extrapolation of $\hat{\gamma}_t(t_n - t)$ to zero.

We investigate age-period specific life expectancies $e_x(t)$ and age-period fixed rate annuities $a_x(t)$ with discount factor *v*, under both the cohort method of computing:

$$\hat{e}_x(t) = \frac{\sum_{i\geq 0} l_{x+i}(t+i)\{1-\frac{1}{2}\hat{q}_{x+i}(t+i)\}}{l_x(t)},\ \hat{a}_x(t) = \frac{\sum_{i\geq 1} l_{x+i}(t+i)v^i}{l_x(t)} \qquad (9)$$

where

$$\hat{q}_x(t) \approx 1-\exp(-\hat{\mu}_{xt}),\ l_{x+1}(t+1) = \{1-\hat{q}_x(t)\}l_x(t),$$

and the period method of computing, in which the variation in *t* in the above expressions is suppressed, as in Section 3.

The simulation strategies are as described in Section 3, subject to the replacement of the suffix *x* by the suffix *x,t* and the insertion of an extra step, when

required, involving the extrapolation of $\mu_{xt}^{(j)} \; \forall \, x,t$ with respect to t, prior to computing the statistics of interest. For simulation Strategy B the parametric vector basis $\boldsymbol{\theta}$ comprises $2k + n - 2$ components

$$\text{LC}: \boldsymbol{\theta}^{\mathrm{T}} = \left(\boldsymbol{\alpha}_x^{\mathrm{T}}, \boldsymbol{\beta}_x^{\mathrm{T}}, \boldsymbol{\kappa}_t^{\mathrm{T}}\right) \text{ with } \beta_{x_1} = 1, \kappa_{t_n} = 0 \tag{10}$$

$$\text{LP}: \boldsymbol{\theta}^{\mathrm{T}} = \left(\boldsymbol{\alpha}_x^{\mathrm{T}}, \boldsymbol{\beta}_x^{\mathrm{T}}, \boldsymbol{\gamma}_t^{\mathrm{T}}\right) \text{ with } \gamma_{t_1} = 0, \gamma_{t_n} = 0 \tag{11}$$

and the variance-covariance matrix is computed as the inverse of the information matrix, (based in turn on the second order partial derivatives of $-\log \mathrm{L}$). Under LC modelling, details of the information matrix are given in Appendix A. Under LP modelling, since period effects are treated deterministically, the terms $\kappa_t \equiv t_n - t$ do not feature as explicit components of $\boldsymbol{\theta}$, while the variance-covariance matrix is readily available from standard linear regression packages, such as GLIM. Details of the information matrix are given in Appendix B. It is then necessary to scale the simulated β_x, (coupled with the compensating scaling of the multiplicative κ_t or $(t_n - t)$ terms, as the case may be), in order to comply with the constraint $\sum_x \beta_x = 1$.

The justification for the introduction of dispersion into the Poisson modelling assumption comes from the following:

$$\mathrm{E}(D_{xt}) = e_{xt}\mu_{xt}, \; \mathrm{Var}(D_{xt}) = \phi \frac{\mathrm{V}\{\mathrm{E}(D_{xt})\}}{\omega_{xt}}; \; \mathrm{V}(u) = u \tag{12}$$

with positive scale parameter ϕ (>1 for over-dispersion), zero-one prior weights ω_{xt}, and the characteristic Poisson variance function V(u) = u. The fitted structure minimises the model deviance

$$D = -2\phi \log\left(L / L_f\right) = 2\phi \sum_{xt} \omega_{xt} \int_{\tilde{d}_{xy}}^{d_{xt}} \frac{d_{xt} - u}{\mathrm{V}(u)} \mathrm{d}u \; , \; \tilde{d}_{xt} = e_{xt}\mu_{\mathrm{xt}} \tag{13}$$

resulting only in the scaling of the variance-covariance matrix by the factor ϕ (which is estimated as the model deviance divided by the number of degrees of freedom).

We compare and display the simulated predictions for the UK male pensioners experience, for the force of mortality, life expectancy and a 4% fixed rate life annuity, as follows-

Age (x)	Period (t)	μ_{xt}	$e_x(t)$, $a_x(t)$
65	2003	Fig 8a	Fig 9
75	2003	Fig 8b	Fig 10
65	2012	Fig 8c	Fig 11

For each of the three age-period combinations, the results are summarised by depicting the 2.5, 50, 97.5 percentiles based on 5,000 simulations, under the following cross-classification-

1. Simulation strategy A, B or C (within each frame Figs 8-11)
2. LC or LP modelling (respective LH or RH frames Figs 8-11)

3. Computation by period or cohort (alternative rows Figs 9-11).

Within each frame the percentiles are depicted for each simulation strategy together with the relevant model based estimate (m.l.e). In addition, under simulation strategy B we present the order statistics both with (outer extremes) and without (inner extremes) a scale parameter ϕ. Here $\hat{\phi} = 1.386$ under LC Poisson modelling with dispersion and $\hat{\phi} = 1.373$ under LP modelling with dispersion. Obviously mortality rate extrapolation is redundant under the fixed period computations for 2003. The scales in adjacent frames are chosen consistently throughout to facilitate comparison, with LC modelling on the left and LP modelling on the right. Life expectancy and annuity simulations (Figs 9-11) by period or by cohort are represented in alternative rows of frames and the scales are set consistently throughout all age-period simulation studies for comparison.

Focusing on the results for age 65, period 2003 in the first instance (Figs 8a & 9) and comparing like with like, we note the following:

1. The close vertical alignment of the simulated and estimated central measures throughout.
2. The sighting of this vertical alignment, relative to the abscissa, in adjacent frames, is indicative of the exceptional close agreements between first moment predictions under LC and LP modelling.
3. Under LC modelling (LH frames), Strategy B (with $\phi = 1$) generates marginally greater risk than Strategy A. This runs counter to the findings reported by Brouhns *et al.* (2005) in their case studies.
4. Strategies A and C generate (roughly) the same amount of risk within each frame.
5. Comparing like with like, Strategies A and C generate (roughly) the same amount of risk under LC and LP modelling (adjacent frames).
6. The introduction of the free-standing scale parameter increases the degree of risk by a noteworthy amount, under Strategy B.
7. Computation of life expectancy and annuities by cohort generates greater risk than computation by period, manifested in the respective wider confidence and prediction intervals (comparison of vertically adjacent frames in each part of Fig 9) and reflecting the greater uncertainty due to extrapolation.

This list of findings also applies to each of the other two reported simulation studies: age 75, period 2003 (Figs 8b & 10) and age 65, period 2012 (Figs 8c & 11). On comparing like with like for comparative ages 65, 75; fixed period 2003 (Figs 9 & 10), the confidence and prediction intervals are marginally wider in absolute terms at age 65 compared with age 75. Similarly, on comparing like with like for comparative periods 2003, 2012; fixed age 65, the wider 2012 confidence and prediction intervals are a measure of the increased risk that arises from 9 years of extrapolation generated by these methods. The life expectancy (period) Strategy B prediction interval, age 65, period 2012 (Fig 11, top LH frame), also features in Fig 14 as part of a further comparative study described in Section 9.

6. Simulation Strategy B under LC Poisson modelling: effect of constraints

In this section, we consider the effects on Strategy B of the choice of the identifiability constraints, which are used in the specification of the Lee-Carter model.

The LC non-linear predictor $\eta_{xt} = \alpha_x + \beta_x \kappa_t$ (and hence the fitted numbers of deaths $\hat{d}_{xt}$ and deviance $D(d_{xt}, \hat{d}_{xt})$) is invariant under the two transformations

$$\{\alpha_x, \beta_x, \kappa_t\} \mapsto \{\alpha_x, \beta_x / c, c\kappa_t\}, \ \{\alpha_x, \beta_x, \kappa_t\} \mapsto \{\alpha_x - c\beta_x, \beta_x, \kappa_t + c\}$$

for any non-zero constants *c*. Thus under the iterative fitting process (Renshaw and Haberman (2006)), we adopt the usual LC constraints, which are then mapped accordingly

$$\sum_x \beta_x = 1, \sum_t \kappa_t = 0 \mapsto \sum_x \beta_x = 1, \kappa_{t_n} = 0$$

using the second of these transformations. Alternatively, it is possible to incorporate the right hand set of constraints directly into the core of the fitting algorithm.

The Cholesky decomposition $\mathcal{C}$ of the variance-covariance matrix $\mathcal{I}^{-1}$ proceeds under the premise that the information matrix $\mathcal{I}$ is constructed first on the basis of $2k + n - 2$ free standing parameters, given the need for two constraints to identify the parametric structure. Consequently, given that the choice of constraints is not unique, we investigate whether the choice of constraints has a material effect on the resulting simulations. Recall that for LC Poisson modelling, simulation Strategy B reads as follows:

> Change to a designated set of constraints and construct $\mathcal{I}$ and hence $\mathcal{C}$ accordingly.
> Simulate a set of free standing parameters $\boldsymbol{\theta}^{(j)} = \hat{\boldsymbol{\theta}} + \sqrt{\phi}\, \mathcal{C}\boldsymbol{\varepsilon}^{(j)}$ and expand $\boldsymbol{\theta}^{(j)}$ by inserting the two constrained parameters.
> Change to the constraints $\sum_x \beta_x^{(j)} = 1, \kappa_{t_n}^{(j)} = 0$ before extrapolating and computing the statistics of interest.

We consider the following alternative pairs of constraints:

> Version 1: $\beta_{x_1} = 1, \kappa_{t_n} = 0$ together with Appendix A
> Version 2: $\beta_{x_1} = 1, \kappa_{t_n} = -\sum_{t \neq t_n} \kappa_t$ together with Appendix C
> Version 3: $\beta_{x_1} = 1, \kappa_{t_1} = 0$ together with Appendix A with suitably adjusted domains in the component blocked matrices

and display the simulated predictions based on the UK male pensioners (age 65, year 2003) experience in Fig 12. Version 1 is as described and presented in Section 5, and we have the same scales in the individual frames to match those in the left hand frames in Figs 8a and 9 to facilitate visual comparisons. Version 2 is as described by Brouhns *et al.* (2002). The varying magnitudes of the percentile based confidence and prediction intervals under the different versions are a cause for concern. In addition, we find that a further version based on the constraints

$\beta_{x_1} = 1 - \sum_{x \neq x_1} \beta_x$, $\kappa_{t_n} = 0$, which would thereby eliminate the need to change to a designated set of (different) constraints, generated totally unrealistic intervals.

Such diverse results are directly attributable to the over parameterisation present in the model rather than the non-linearity of the parametric structure. For example, this situation does not arise in the case of Makeham's second law (Section 4) where the structure is non-linear but not over-parameterised. Further, the situation has been found to arise in other examples (details omitted): in the application of simulation Strategy B in a re-analysis of the average cost of claims for privately owned, insured cars, cross-classified by three factors (policyholder's age, car group and vehicle age), with the main effects of all three factors modelled as an over parameterised generalised *linear* model (McCullagh and Nelder (1989) pp296-300).

7. Simulation Strategies A and C: effect of variable scale parameters

Since the constant scale parameter ϕ fails to impact on the CIs generated under simulation Strategies A and C, in this section, we investigate the consequences of reformulating the scale parameter as a function of age (and possibly period) ϕ_{xt}. Under such a formulation, fitting is possible by so-called joint modelling comprising the following:

Stage 1: Model D_{xt} as independent Poisson responses

$$\mathrm{E}(D_{xt}) = e_{xt}\mu_{xt}, \ \mathrm{Var}(D_{xt}) = \phi_{xt}\frac{\mathrm{V}\{\mathrm{E}(D_{xt})\}}{\omega_{xt}}; \ \mathrm{V}(u) = u$$

with variable dispersion ϕ_{xt}, the log link

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

and the LC or LP parametric structures (6) or (7).

We then define $R_{xt} = \omega_{xt}\dfrac{\{D_{xt} - \mathrm{E}(D_{xt})\}^2}{\mathrm{E}(D_{xt})}$, the resulting squared Pearson residuals.

Stage 2: Model R_{xt} as independent gamma responses

$$\mathrm{E}(R_{xt}) = \phi_{xt}, \ \mathrm{Var}(R_{xt}) = \tau\frac{\mathrm{V}\{\mathrm{E}(R_{xt})\}}{\omega_{xt}}; \ \mathrm{V}(u) = u^2$$

with scale parameter τ, the log link

$$\text{LG: } \log\phi_{xt} = \varsigma_x$$

and linear parametric structure in age effects, denoted (LG).

The joint process is implemented by iteratively fitting each stage in sequence, terminating with the convergence of both deviances. Stage 1 residuals form Stage 2 responses while Stage 2 fitted values form the Stage 1 weights ω_{xt}/ϕ_{xt}, where ω_{xt} are the zero-one empty data cell indicators. Typical starting values are $\phi_{xt} = \omega_{xt}$. The choice of Stage 2 gamma distribution is consistent with Stage 1 normally distributed residuals. Alternatively, Stage 1 squared deviance residuals can be used

instead of squared Pearson residuals, leading to the same results, where $\mathrm{E}(R_{xt}) \approx \phi_{xt}$. For primary references and further background details on joint modelling in a mortality setting, we refer the interested reader to Renshaw (1992).

We begin the analysis by refitting the UK male pensioner mortality experience to both the LC/LG and LP/LG joint models. The resulting Stage 1 parameter estimates, including the first moment structures of the joint processes, together with the Stage 1 deviance residual plots, are as depicted in Section 5 (Fig 5, Fig 6a, Fig 7). However, under LP/LG modelling, the so-called Poisson trick no longer applies (due to the variability in ϕ) so that, as with LC/LG modelling, the actual and expected total deaths in each period are not the same (see Fig 13a). (Nevertheless, the term in γ_t, expression (7), is retained to improve the quality of fit).

The Stage 2 parameter estimates $\left(\hat{\phi}_{xt} =\right) \hat{\phi}_x = \exp\hat{\varsigma}_x$, encapsulating the second moment structure of the joint processes, are depicted in Fig 13a. Here, in both frames, we have superimposed the results obtained under LC/LG and LP/LG modelling for comparison. Noting the convex shape of the near identical $\hat{\phi}_x$ age profiles, we observe that under dispersion $(\phi_x < 1))$ occurs at the age extremities, coinciding with the paucity of exposure to the mortality risk.

We repeat the age 65, period 2003, simulation Strategies A and C under both LC/LG and LP/LG joint modelling, displaying the 2.5, 50, 97.5 (N = 5,000) percentiles for life expectancy and 4% fixed rate annuity predictions, computed by period and by cohort (Fig 14). For comparison, the equivalent results from Section 5 (Fig 8), obtained under LC or LP single stage modelling are displayed alongside, using the same abscissae (Fig 8) for consistency. (The third set of intervals, relate to LC or LP negative binomial modelling and are explained separately in the next section).

The consistent small lateral displacements of the medians under joint modelling are due solely to the switch from single to joint modelling and occur to the same extent when joint modelling is conducted with a constant dispersion $(\varsigma_x = \varsigma)$: and we note that the constant dispersion models generate confidence and prediction intervals of similar widths to those generated under single stage modelling (results not depicted). Thus the appreciable increase in interval widths under joint modelling relative to single stage modelling is directly attributable to the second moment age variable dispersion parameters $(\varsigma_x \Leftrightarrow \phi_x)$. When implementing Strategy A, the simulated responses are approximated by sampling the Poisson distribution, which may explain the narrower confidence and prediction intervals, compared with Strategy C. (A way of avoiding this approximation, under suitable conditions, is discussed in Section 10). Skewness is also a noteworthy feature of the intervals, especially under LC/LG joint modelling.

8. Simulation Strategies A and C under negative binomial modelling

Li *et al* (2006), (see also Delwarde *et al* (2007)) present the case for enhancing risk measurement under simulation Strategy A, by switching from the Poisson to a negative binomial response LC model as defined by (12), with no scale factor $(\phi = 1)$ and a characteristic variance function $\mathrm{V}(u) = u + \lambda_x u^2$. Both the first moment non-

linear LC parameters (6) and second moment parameters $\{\lambda_x\}$ are estimated in a single stage: by optimising the negative binomial likelihood.

A key difference between Poisson and negative binomial modelling concerns the nature of their distribution functions. While the former is a member of the exponential family of distributions, the latter is not: consequently the simple relationship $D \propto -2\log(L)$, connecting the deviance D to the likelihood L no longer holds. Thus, for the negative binomial distribution, on evaluating the integral expression (13)

$$D = 2\sum_{x,t} \omega_{xt} \left\{ d_{xt} \log\left(\frac{d_{xt}}{\tilde{d}_{xt}} \right) - \left(d_{xt} + \frac{1}{\lambda_x} \right) \log\left(\frac{1+\lambda_x d_{xt}}{1+\lambda_x \tilde{d}_{xt}} \right) \right\}$$

and, (on adopting expression (22) of Li *et al* (2006)), the kernel of the log likelihood

$$\log L = \sum_{x,t} \omega_{xt} \left\{ \sum_{i=0}^{d_{xt}-1} \log(1+\lambda_x i) + d_{xt} \log \tilde{d}_{xt} - \left(d_{xt} + \frac{1}{\lambda_x} \right) \log\left(1+\lambda_x \tilde{d}_{xt}\right) \right\}$$

where, under LC: $\tilde{d}_{xt} = e_{xt}\mu_{xt} = e_{xt}\exp(\alpha_x + \beta_x \kappa_t)$. On comparing the kernel of D with $\log L$, the expressions differ (apart from their sign) only in respect of the additional first term in $\log L$ involving only the second moment parameters (otherwise $D \propto -2\log L$). Hence, the updating relationships for the first moment parameters used in fitting the LC structure, are derived from either the model deviance or likelihood *viz*

$$\mathrm{u}(\hat{\alpha}_x) = \hat{\alpha}_x + \frac{\sum_t \omega_{xt}\left(d_{xt} - \dot{d}_{xt}\right)}{\sum_t \omega_{xt} \ddot{d}_{xt}}, \quad \mathrm{u}(\hat{\beta}_x) = \hat{\beta}_x + \frac{\sum_t \omega_{xt} \hat{\kappa}_t \left(d_{xt} - \dot{d}_{xt}\right)}{\sum_t \omega_{xt} \hat{\kappa}_t^2 \ddot{d}_{xt}} \tag{14}$$

$$\mathrm{u}(\hat{\kappa}_t) = \hat{\kappa}_t + \frac{\sum_x \omega_{xt} \hat{\beta}_x \left(d_{xt} - \dot{d}_{xt}\right)}{\sum \omega_{xt} \hat{\beta}_x^2 \ddot{d}_{xt}}; \quad \dot{d}_{xt} = \left(\frac{1+\hat{\lambda}_x d_{xt}}{1+\hat{\lambda}_x \hat{d}_{xt}} \right) \hat{d}_{xt}, \quad \ddot{d}_{xt} = \frac{\dot{d}_{xt}}{1+\hat{\lambda}_x \hat{d}_{xt}}$$

reducing to the more familiar updating relationships for Poisson LC (e.g. Table 1 Renshaw and Haberman (2006)) in the limit $\lambda_x = 0 \,\forall\, x$, for which $\ddot{d}_{xt} = \dot{d}_{xt} = \hat{d}_{xt}$).

Model fitting by optimising the deviance D (the GLM approach) is only possible when the set of second moment parameters $\{\lambda_x\}$ are known: a feature we exploit when implementing simulation Strategies A and C. In order to fit the structures in the first instance, we follow Li *et al* (2006) in optimising the likelihood L on the basis of their further updating relationship, adopted to read

$$\mathrm{u}\left(\hat{\lambda}_x\right)=\hat{\lambda}_x+\frac{\sum_t \omega_{xt}\left\{-\sum_{i=0}^{d_{xt}-1}\frac{i}{\left(1+\hat{\lambda}_x i\right)}+\frac{\dot{d}_{xt}}{\hat{\lambda}_x}+\frac{1}{\hat{\lambda}_x^2}\log\left(\frac{\ddot{d}_{xt}}{\dot{d}_{xt}}\right)\right\}}{\sum_t \omega_{xt}\left\{-\sum_{i=0}^{d_{xt}-1}\frac{i^2}{\left(1+\hat{\lambda}_x i\right)^2}-\frac{2d_{xt}}{\hat{\lambda}_x^2}+\frac{2\hat{d}_{xt}\ddot{d}_{xt}}{\hat{\lambda}_x^2\dot{d}_{xt}}+\frac{\hat{d}_{xt}\ddot{d}_{xt}}{\hat{\lambda}_x}+\frac{2}{\hat{\lambda}_x^3}\log\left(\frac{\ddot{d}_{xt}}{\dot{d}_{xt}}\right)\right\}}.$$

Updating relationships (14) with $\hat{\kappa}_t = t_n - t$, together with the relationship

$$\mathrm{u}\left(\hat{\gamma}_t\right)=\hat{\gamma}_t+\frac{\sum_x \omega_{xt}\left(t_n-t\right)\left(d_{xt}-\dot{d}_{xt}\right)}{\sum_x \omega_{xt}\left(t_n-t\right)^2\ddot{d}_{xt}}$$

are used to fit the LP structure in the first instance.

In implementing simulation Strategy A, we use the rejection method (Section 7.3 pp 281-286 Press *et al* (1997)) for generating random deviates from the (fitted) negative binomial probability function (suppressed suffices)

$$\mathrm{p}\left(d\mid\hat{d},\hat{\lambda}\right)=\exp\left\{\log\Gamma\left(d+\hat{\lambda}^{-1}\right)-\log\Gamma\left(\hat{\lambda}^{-1}\right)-\log\Gamma\left(d+1\right)+d\log\left(\frac{\hat{d}}{\hat{d}+\hat{\lambda}^{-1}}\right)-\hat{\lambda}^{-1}\log\left(1+\hat{d}\hat{\lambda}\right)\right\}$$

The implementation of simulation Strategy C is as described in Section 2, subject to the matter of detail. Thus, here, we follow Section 2 and require the appropriate root d^* of

$$\mathrm{g}\left(d\right)=d\left\{\log d-\log\left(1+\hat{\lambda}d\right)\right\}-d\hat{b}-\frac{1}{\hat{\lambda}}\log\left(1+\hat{\lambda}d\right)-\hat{c}^* \qquad (15)$$

where $\hat{b}=\log\hat{d}-\log\left(1+\hat{\lambda}\hat{d}\right)$, $\hat{c}^*=\frac{r^{*2}}{2}-\frac{1}{\hat{\lambda}}\log\left(1+\hat{\lambda}\hat{d}\right)$. When mapping $r^*\mapsto d^*$, we note the derivatives of $g\left(d\right)$:

$$\mathrm{g}'\left(d\right)=\log d\text{ - }\log\left(1+\hat{\lambda}d\right)-\hat{b}\ \Rightarrow\ g'\left(\hat{d}\right)=0$$

$$\mathrm{g}''\left(d\right)=\frac{1}{d\left(1+\hat{\lambda}d\right)}>0\ \ \forall\, d>0,\ \lambda\geq 0$$

which imply that $g\left(d\right)$ is concave with a minimum at $\hat{d}$. We note that the details reduce to those of Section 2 when $\lambda=0$.

The simulation strategies (A and C) are conducted on the basis of updated first moment parameters, optimum deviance, subject to fixed $\hat{\lambda}_x$, resulting in an appreciable reduction in computer time, compared with the updating of first and second moment parameters and the optimisation of the likelihood, at each simulation. (Exercises comparing the two approaches are found to generate identical simulations).

We begin a comparative study with Sections 5 and 7 by refitting the UK male pensioner mortality experience using the respective LC and LP negative binomial model. As anticipated, the patterns in the fitted first moment parameters and deviance residual plots are as depicted in Section 5 (Fig 5, Fig 7) and are hence not reproduced. Both the actual minus expected annual death totals and the second moment parameter estimates $\{\hat{\lambda}_x\}$, superimposed by model structure, LC or LP, are shown in Fig 13b. We note that the concave nature of the pattern in the $\hat{\lambda}_x$s under negative binomial modelling is complementary to the convex nature of the pattern in the ϕ_xs under joint modelling.

The comparative 2.5, 50, 97.5 (N = 5,000) simulated percentiles for (age 65, period 2003) life expectancy and 4% fixed rate annuity predictions, computed by both period and cohort under negative binomial modelling, are displayed alongside their counterparts under (single stage) Poisson modelling and Poisson joint modelling (Fig 14). The consistent small lateral displacements in the medians, in the opposite direction to the displacements under joint modelling, are a measure of the small change in the point predictions brought about by the change in modelling distribution. While the switch from (single stage) Poisson to negative binomial distribution results in consistently wider confidence and prediction intervals under simulation Strategy A, the equivalent intervals are essentially identical under Strategy C: the likely reason for which is discussed in Section 10. Comparing like with like, the widths of the simulated intervals under LC and LP negative binomial modelling are effectively identical. Both the pattern in the $\hat{\lambda}_x$s and the increased width in intervals generated under LC negative binomial Strategy A simulations are consistent with the findings for the Canadian mortality experiences conducted by Li *et al* (2006).

9. Life expectancy: comparison with a non-simulation approach

It is informative to compare, where possible, simulated life expectancy prediction intervals under LC modelling (specifically incorporating extrapolation by random walk with drift), with the theoretical prediction intervals, constructed by computing the order statistics for life expectancy predictions based solely on the extrapolation error in the random walk with drift. This is now possible using formulae due to Denuit (2006). Thus, for life expectancy computations by fixed (extrapolated) period $t_n + s$ ($s > 0$), order statistics are computed as

$$F^{-1}_{e_x(t_n+s)}(p) = 0.5 + \sum_{d\geq 1} \exp\left[-\sum_{j=0}^{d-1} \exp\left\{\hat{\alpha}_{x+j} + \hat{\beta}_{x+j}\left(\hat{\kappa}_{t_n} + s\hat{\theta} + \hat{\sigma}\sqrt{s}\Phi^{-1}(p)\right)\right\}\right], \quad p \in (0,1)$$

where $\hat{\theta}$ and $\hat{\sigma}$ denote the respective estimated drift parameter and standard deviation in the time series $\{\hat{\kappa}_t : t = t_1, ..., t_n\}$ modelled as a random walk with drift, and Φ is the distribution function of the standardised normal variate. Thus, typically, for fixed s (>0), when $p = 0.5$ so that $\Phi^{-1}(0.5) = 0$, the median

$$F^{-1}_{e_x(t_n+s)}(0.5) = 0.5 + \sum_{d\geq 1} \exp\left[-\sum_{j=0}^{d-1} \left\{ \dot{\mu}_{x+j,t_n+s} \right\} \right].$$

Denuit (2006) also provides details of an equivalent formula for the order statistics for life expectancy $e_x(t_n)$, relating to the most recently observed period, and computed along a particular extrapolated cohort.

We compare order statistics computed by these methods for the UK male pensioner experience with the same order statistics generated under (i) (single stage) Poisson based simulation Strategy B (Version 1, including the scale parameter ϕ) and (ii) negative binomial based simulation Strategy A, in Fig 15. Specifically, the 2.5, 50, 97.5 percentiles for life expectancy predictions at age 65, computed by period, for (fixed) projected periods at 4 yearly intervals in the range 2004-2020 are depicted in the upper frame (Fig 15a), and the same percentiles for life expectancy predictions in the period 2003, computed by cohort, for ages at 5 yearly intervals in the range 65-85 in the lower frame (Fig 15b). Comparing like with like, we not the following:

1. The theoretical median is more closely aligned with the simulated negative binomial median than the simulated Poisson median, with consistent staggering of alignments throughout.
2. The theoretical prediction intervals are appreciably wider than both sets of simulated prediction intervals.
3. Whereas the simulated negative binomial prediction intervals are consistently wider than their simulated Poisson counterparts, for fixed age with increasing age (computations by cohort- Fig 15a), the opposite is the case, for fixed age with increasing extrapolated calendar year (computations by period- Fig 15a).

We comment on these results in the final section.

10. Discussion and conclusions

Simulation Strategy C differs from that described and illustrated in Koissi *et al.* (2006) with respect to the method of mapping the bootstrap simulated residuals to responses. Instead of solving equation (1) for $d_i^{(j)}$ when $r_i = r_i^{(j)}$ with fixed $\hat{d}_i \; \forall \, i$ (and $i = (x, t)$), Koissi *et al.* (2006) elect to solve equation (1) for $\hat{d}_i^{(j)}$ with fixed $d_i \; \forall \, i$. We do not follow this approach on the basis that it does not comply with the spirit of Chapter 9, Efron and Tibshirani (1993), and that it fails to simulate any variation when this approach is applied to the simple single parameter model of Section 2.

The use of deviance residuals in simulation Strategy C is appealing because the models are fitted on the basis of deviance optimisation. (Note the parallel relationship between Pearson residuals and OLS fitting). Further, the use of Pearson residuals in this context, as a possible alternative, is potentially problematic. Thus, in the notation of Section 2, with residuals

$$r_i = \frac{d_i - \hat{d}_i}{\sqrt{\mathrm{V}(\hat{d}_i)}}, \; \mathrm{V}(u) = u$$

replacing expression (1), the bootstrap residual mapping $r^* \mapsto d^*$ takes the form

$$d^* = \hat{d} + r^* \sqrt{\hat{d}}$$

which has the potential to generate negative "Poisson" bootstrap responses (*viz* when $\hat{d} < r_{min}^2, r^* = r_{min} < 0$).

A good fit resulting in a set of pattern free random residuals for sampling, repeatedly with replacements, is a basic requirement of Strategy C. Where this is not the case, as in Section 3 (Fig 2), distortions can occur in the simulated histogram of the quantity of interest. We have not however investigated in this paper, the likely magnitude and nature of the distortion, if any, when Strategy C is applied under LC modelling in the presence of residual cohort effects: LC modelling being preferred on the (perceived) over-riding basis of its powerful predictive properties.

Expressions (4) and (5) used in the construction of approximate m.l.e. CIs for (fixed period) life expectancy and fixed rate annuities (Figs 2, 3, 9, 10) are based on the binomial assumption $D_x \sim \text{bin}\left(e_x^i, q_x\right)$ with independence $\forall x$. We make the parallel Poisson assumption $D_x \sim \text{Poi}\left(e_x \mu_x\right)$ to target μ_x with independence $\forall x$ before computing q_x, which does not materially contribute further to the approximate nature of these m.l.e. CIs. Given that expressions (4) and (5) require a knowledge of the exposures, the formulae cannot be applied under dynamic extrapolation (computations by cohort (Figs 9, 10)).

In Section 4, when computing life expectancy and fixed rate annuities we have decided against extrapolating Makeham's second law beyond age 90, given our primary objective of conducting a comparative study of the different simulation strategies. Likewise, given the comparative nature of the aims of this study, we have chosen not to complicate matters by smoothing the parameter sets $\{\alpha_x\}$ and $\{\beta_x\}$ in Sections 5 to 9. However, we do acknowledge that it is important to smooth both parameters sets for practical applications in order to avoid the projection into the future of age-specific irregularities in the $\{\alpha_x\}$ and $\{\beta_x\}$ sequences.

On comparing like with like in the Poisson age-period study of Section 5 (Figs 8 to 11), LC and LP point and interval predictions differ materially only in respect of simulation Strategy B. Under LP modelling the issue of over parameterisation occurs because of the inclusion of the terms in $\gamma_t(t - t_n)$ and is resolved by setting $\gamma_{t_1} = \gamma_{t_n} = 0$: the sole purpose of these terms is to improve the quality of fit (using the so-called Poisson trick- see Renshaw and Haberman (2003a)), while allowing the terms in $\beta_x(t - t_n)$ to capture the main age-specific (linear) period effects. Consequently, the choice of parameter constraints is resolved by this means. Under LC modelling, however, we are uncertain how to resolve this issue of over parameterisation and we illustrate the effects of the constraints in Section 6. Unless and until this issue is resolved, we believe that LC simulations based on Strategy B should be interpreted with caution.

The primary concern (see, for example, Li *et al* (2006) p 14 Section 6) that a constant scale parameter ϕ fails to impact on Strategy A bootstrap intervals, is supported by the findings of Section 5 and extends also to Strategy C. We note that a constant ϕ does impact under parametric Strategy B, but we need a satisfactory resolution of the outstanding issue of over-parameterisation.

We do not however subscribe to the assertion (Li *et al* (2006) p 14 Section 6), requiring internal consistency in the relationship between the probability function of the responses D_{xt} and their first two moments $\mathrm{E}(D_{xt})$ and $\mathrm{Var}(D_{xt})$ as a general principle of modern statistical modelling: Section 7 being a case in point. We quote form the outer cover of McCullagh and Nelder (1989) 'An important feature [of generalised linear and non-linear modelling] is that the principal conclusions depend only on [first and] second moment assumptions as opposed to the complete correctness of an assumed probability function.

Enhanced prediction intervals are obtained under Poisson joint modelling with age-specific scale parameters (Section 7: simulation Strategies A and C) and under negative binomial modelling (Section 8: simulation Strategy A). In contrast with (single stage) Poisson modelling, both of these modelling approaches incorporate specific provision for targeting the second as well as first moment properties, which are subsequently reflected in one or both (joint modelling) of the simulation Strategies A and C.

Both the formation of the joint model (Section 7) and the implementation of simulation Strategy C (throughout) are distribution free. However, simulation Strategy A requires a probability distribution function. In applying Strategy A (Section 7), we approximated by using the Poisson probability function. In the event of complete over-dispersion ($\phi_{xt} \geq 1, \forall x,t$), which is only partially the case (Fig 13, upper right frame), exact responses may be simulated under Strategy A using the two-parameter probability function (suppressed suffices)

$$\mathrm{p}\left(d \mid \hat{d}, \hat{\phi}\right) = \exp\left\{ \log\frac{\hat{d}}{\sqrt{\hat{\phi}}} + (d-1)\log\left(d - \frac{d - \hat{d}}{\sqrt{\hat{\phi}}} \right) - \left(d - \frac{d - \hat{d}}{\sqrt{\hat{\phi}}} \right) - \log\Gamma(d+1) \right\}$$

based on a re-parameterisation of the two-parameter Lagrangian Poisson distribution (e.g. Ter Berg (1996)): reducing to the Poisson probability function when $\phi = 1$. We are unable to assess the extent of the material effect, if any, of basing Strategy A on the Poisson approximation, for the male pensioner experience.

Under simulation Strategy C, the likely explanation as to why the confidence and prediction intervals are not enhanced under negative binomial modelling, compared with (single stage) Poisson modelling, is because of the very small values of λ_x in the body of the data and age range (Fig 13, lower right frame). Thus, in the limit $\lambda \to 0$, for which

$$\lim_{\lambda \to 0} \log(1 + \lambda d) = 0, \ \lim_{\lambda \to 0} \frac{1}{\lambda} \log(1 + \lambda d) = d,$$

expression (15), used to map bootstrap negative binomial deviance residuals to responses, reduces to expression (2), used to map bootstrap Poisson residuals.

In Section 9, we have compared LC simulated prediction intervals for life expectancy predictions, with their theoretical equivalents based on the order statistics of the extrapolation error in the random walk with drift: providing evidence that the theoretical prediction intervals are appreciably wider than any of their simulated counterparts, including negative binomial and joint modelling. Obviously, here, the application of the theoretically based approach (Denuit (2006)), which is restricted to

LC life expectancy predictions with extrapolation confined to the random walk with drift, is less versatile than the simulation approaches. Simulation methods for assessing mortality risk in Poisson LC models are used, in part, in the belief that they capture both model fitting and extrapolation error (Brouhns *et al.* (2002)). Consequently, for life expectancy predictions, it is difficult to reconcile this belief given the narrower prediction intervals generated by these, and like simulation methods, when compared with the wider prediction intervals generated by the non-simulation method predicated solely on extrapolation error.

While the bootstrap is a computer-based method of statistical inference, typically used to simulate confidence intervals, the precedence and basis for using the bootstrap to simulate (LC and LP) prediction intervals, requires further investigation. For example, under LC modelling, on adapting Section 8.5 of Efron and Tibshirani (1993) to simulate the random walk with drift for the time series $\{\hat{\kappa}_t\}$ of Section 5, we compare the theoretical based details of the estimated drift parameter and 2013 predicted time series, with their respective simulated counter-parts (N = 5,000) below:

	Estimated drift parameter 95% confidence interval	*Predicted (2013) time series 95% prediction interval*
Theoretical	$\hat{\theta} = -1.57 \ (-2.60, -0.54)$	$\dot{\kappa}_{2013} = -15.74 \ (-30.30, -1.18)$
Simulated	$\tilde{\theta} = -1.58 \ (-2.60, -0.61)$	$\dot{\kappa}_{2013} = -15.79 \ (-25.95, -6.12)$

Thus, while there is close agreement between the theoretical and simulated estimated drift parameter and its confidence intervals, the theoretical prediction interval for the time series is materially wider that its simulated counter part.

In conclusion, we highlight the following:

- Unless there are compelling reasons for selecting a particular set of constraints when identifying the non-linear (or linear) structure of an over-parameterised model, such as LC, simulation Strategy B should not be used for risk assessment purposes, since different choices of constraints result in widely differing simulated confidence and prediction interval widths.
- Negative binomial and Poisson joint modelling incorporate extra provision for targeting second moment properties, compared to (single stage) Poisson modelling. This is reflected in wider simulated confidence and prediction intervals under Strategy A (both approaches) and Strategy C (joint modelling only), than would otherwise be the case. Strategy C fails to simulate wider intervals under negative binomial modelling for the reason discussed. Further reported case studies would be of interest in this respect.
- LC modelling allows for greater variability in the period component than LP modelling. We have compared confidence and prediction intervals under LC modelling and extrapolation by random walk with drift, with matching intervals under LP modelling and linear extrapolation. For Strategies A and C, in combination with all three modelling approaches (single and joint Poisson, negative binomial), none of the matching simulated prediction intervals are (materially) wider in absolute terms under LC modelling compared with LP modelling: possibly suggesting a failure to capture the full magnitude of the time series forecast error under LC modelling, thereby endorsing the findings of Section 9.

References

Benjamin, B, Pollard, J.H. 1980. The analysis of mortality and other actuarial statistics. Heinemann, London.

Brouhns, N., Denuit, M., Vermunt, J.K., 2002. Measuring the longevity risk in mortality projections. Bulletin of the Swiss Association of Actuaries, 105-130.

Brouhns, N., Denuit, M., van Keilegom, 2005. Bootstrapping the Poisson log-bilinear model for mortality forecasting. Scandinavian Actuarial Journal, 3, 212-224.

Clarke, R.D. (1946). An aoolication of the Poisson distrbution. Journal of the Institute of Actuaries, 72, 481.

CMI Committee, 1999. Standard tables of mortality based on the 1991-1994 experiences. Continuous Mortality Report, 17, 1-227. Institute and Faculty of Actuaries.

Delwarde, A., Denuit, M., Partrat, C. 2007. Negative binomial version of the Lee-Carter model for mortality forecasting.

Denuit, M., 2006. Distribution of the random future life expectancies in log-bilinear mortality projection models. (under review).

Efron, B. Tibshirani, R.J., 1993. An introduction to the bootstrap. Chapman & Hall, New York & London.

Forfar, D.O., McCutcheon, J.J., Wilkie, A.D. 1988. On graduation by mathematical formula. Journal of the Institute of Actuaries, 115, 1-135

Francis, B., Green, M., Payne, C., 1993. The Glim System: Release 4 Manual. Clarendon, Oxford.

Koissi, M-C., Shapiro, A.F., Hognas, G. 2006. Evaluating and extending the Lee-Carter model for mortality forecasting: bootstrap confidence intervals. Insurance: Mathematics and Economics 38, 1-20.

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

Li, S-H, Hardy, M.R., Tan, K.S. 2006. Uncertainty in mortality forecasting: an extension of the classical Lee-Carter approach. University of Waterloo, Ontario N2L 3G1 (email iipr@icarus.math.uwaterloo.ca).

McCullagh, P., Nelder, J.A. 1989. Generalised Linear Models. Chapman and Hall, London & New York.

Press, H.W., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P. 1997. Numerical Recipes in Fortran 77. Cambridge University Press.

Renshaw, A.E. 1991. Actuarial graduation practice and generalised linear and non-lonear models. Journal of the Institue of Actuaries, 118, 295-312.

Renshaw, A.E. 1992. Joint modelling for actuarial graduation and duplicate policies. Journal of the Institue of Actuaries, 119, 69-85.

Renshaw, A.E., Haberman, S., 2003a. Lee-Carter mortality forecasting: a parallel generalised linear modelling approach for England and Wales mortality projections. Applied Statistics, 52, 119-137.

Renshaw, A.E., Haberman, S., 2003b. On the forecasting of mortality reduction factors. Insurance: Mathematics and Economics 33, 255-272.

Renshaw, A.E., Haberman, S., 2003c. 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.

Ter Berg, P. 1996. A log-linear Lagrange Poisson model. Astin Bulletin 26, 123-129.

Appendix A

Details of the information matrix $\mathcal{I}$ under LC modelling, subject to the parameter specification and constraints (10) are as follows:

Matrix	….	*Size*
$\mathcal{I} = \begin{bmatrix} \boldsymbol{m}_{11} & \boldsymbol{m}_{12} & \boldsymbol{m}_{13} \\ \boldsymbol{m}_{12}^{\mathrm{T}} & \boldsymbol{m}_{22} & \boldsymbol{m}_{23} \\ \boldsymbol{m}_{13}^{\mathrm{T}} & \boldsymbol{m}_{23}^{\mathrm{T}} & \boldsymbol{m}_{33} \end{bmatrix}$		(2k+n-2) by (2k+n-2)
$\boldsymbol{m}_{11} = diag\left(\sum_t \hat{d}_{x_i t} \right); i = 1,2,\ldots,k$		(k) by (k)
$\boldsymbol{m}_{22} = diag\left(\sum_t \hat{\kappa}_t^2 \hat{d}_{x_i t} \right); i = 2,\ldots,k$		(k-1) by (k-1)
$\boldsymbol{m}_{33} = diag\left(\sum_x \hat{\beta}_x^2 \hat{d}_{x t_j} \right); j = 1,2,\ldots,n\text{-}1$		(n-1) by (n-1)
$\boldsymbol{m}_{12} = [\, m_{ij} \,]$ with $m_{i,i\text{-}1} = \sum_t \hat{\kappa}_t \hat{d}_{x_i t}; i > 1$, $m_{ij} = 0$; otherwise		(k) by (k-1)
$\boldsymbol{m}_{13} = [\, m_{ij} \,]$ with $m_{ij} = \hat{\beta}_{x_i} \hat{d}_{x_i t_j}; j < n$		(k) by (n-1)
$\boldsymbol{m}_{23} = [\, m_{ij} \,]$ with $m_{ij} = \hat{\beta}_{x_i} \hat{\kappa}_{t_j} \hat{d}_{x_i t_j}; i > 1,\ j < n$		(k-1) by (n-1)

Note $\mathrm{E}(D_{xt}) = \hat{d}_{xt} = e_{xt} \exp\left(\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t\right)$.

Appendix B

Details of the information matrix $\mathcal{I}$ under LP modelling, subject to the parameter specification and constraints (11) are as follows:

Matrix	….	*Size*
$\mathcal{I} = \begin{bmatrix} \boldsymbol{m}_{11} & \boldsymbol{m}_{12} & \boldsymbol{m}_{13} \\ \boldsymbol{m}_{12}^{\mathrm{T}} & \boldsymbol{m}_{22} & \boldsymbol{m}_{23} \\ \boldsymbol{m}_{13}^{\mathrm{T}} & \boldsymbol{m}_{23}^{\mathrm{T}} & \boldsymbol{m}_{33} \end{bmatrix}$		(2k+n-2) by (2k+n-2)
$\boldsymbol{m}_{11} = diag\left(\sum_t \hat{d}_{x_i t} \right); i = 1,2,\ldots,k$		(k) by (k)
$\boldsymbol{m}_{22} = diag\left(\sum_t s_t^2 \hat{d}_{x_i t} \right); i = 1,2,\ldots,k$		(k) by (k)
$\boldsymbol{m}_{33} = diag\left(s_{t_j}^2 \sum_x \hat{d}_{x t_j} \right); j = 2,3,\ldots,n\text{-}1$		(n-2) by (n-2)
$\boldsymbol{m}_{12} = diag\left(\sum_t s_t \hat{d}_{x_i t} \right); i = 1,2,\ldots,k$		(k) by (k)
$\boldsymbol{m}_{13} = [\, m_{ij} \,]$ with $m_{ij} = s_{t_j} \hat{d}_{x_i t_j}; 1 < j < n$		(k) by (n-2)
$\boldsymbol{m}_{23} = [\, m_{ij} \,]$ with $m_{ij} = s_{t_j}^2 \hat{d}_{x_i t_j}; 1 < j < n$		(k) by (n-2)

Note $\mathrm{E}(D_{xt}) = \hat{d}_{xt} = e_{xt} \exp\left(\hat{\alpha}_x + \hat{\beta}_x s_t + \hat{\gamma}_t s_t\right)$ where $s_t = t_n - t$.

Appendix C

Details of the information matrix $\mathcal{I}$ under LC modelling, subject to the constraints $\beta_{x_1} = 1,\ \kappa_{t_n} = -\sum_{t \neq t_n} \kappa_t$ are as follows:

Matrix	….	*Size*
$\mathcal{I} = \begin{bmatrix} \boldsymbol{m}_{11} & \boldsymbol{m}_{12} & \boldsymbol{m}_{13} \\ \boldsymbol{m}_{12}^{\mathrm{T}} & \boldsymbol{m}_{22} & \boldsymbol{m}_{23} \\ \boldsymbol{m}_{13}^{\mathrm{T}} & \boldsymbol{m}_{23}^{\mathrm{T}} & \boldsymbol{m}_{33} \end{bmatrix}$		(2*k*+*n*-2) by (2*k*+*n*-2)
$\boldsymbol{m}_{11} = diag\left(\sum_t \hat{d}_{x_i t} \right);\ i = 1,2,\ldots,k$		(*k*) by (*k*)
$\boldsymbol{m}_{22} = diag\left(\sum_t \hat{\kappa}_t^2 \hat{d}_{x_i t} \right);\ i = 2,\ldots,k$		(*k*-1) by (*k*-1)
$\boldsymbol{m}_{33} = [m_{ij}]$ with $m_{jj} = \sum_x \hat{\beta}_x^2 \left(\hat{d}_{xt_j} + \hat{d}_{xt_n} \right);\ j = 1,2,\ldots,n-1$; $m_{ij} = \sum_x \hat{\beta}_x^2 \hat{d}_{xt_n}\ ;\ i \neq j < n$		(*n*-1) by (*n*-1)
$\boldsymbol{m}_{12} = [m_{ij}]$ with $m_{i,i-1} = \sum_t \hat{\kappa}_t \hat{d}_{x_i t};\ i > 1$; $m_{ij} = 0$; otherwise		(*k*) by (*k*-1)
$\boldsymbol{m}_{13} = [m_{ij}]$ with $m_{ij} = \hat{\beta}_{x_i} \left(\hat{d}_{x_i t_j} - \hat{d}_{x_i t_n} \right);\ j < n$		(*k*) by (*n*-1)
$\boldsymbol{m}_{23} = [m_{ij}]$ with $m_{ij} = \hat{\beta}_{x_i} \left(\hat{\kappa}_{t_j} \hat{d}_{x_i t_j} - \hat{\kappa}_{t_j} \hat{d}_{x_i t_n} \right);\ i > 1,\ j < n$		(*k*-1) by (*n*-1)

Note $\mathrm{E}(D_{xt}) = \hat{d}_{xt} = e_{xt} \exp\left(\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t \right)$.

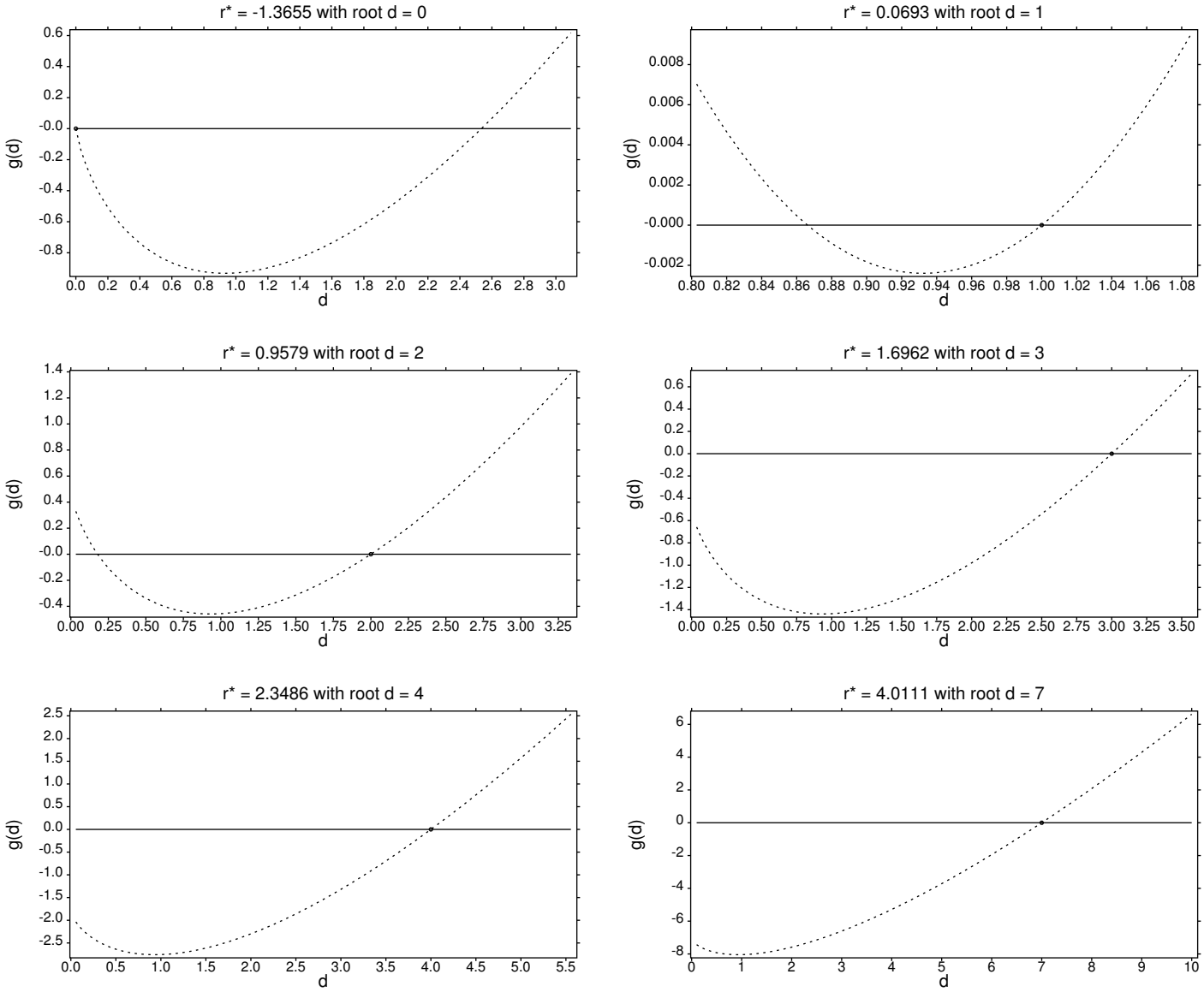


Fig 1. Graphs of g(d) vs d, for each residual r*, mapping r* -> d*, by selecting the relevant root of g(d), conditional on the sign of r*. A special case in which r* -> d with a minimum at d = 0.9323 in each frame.

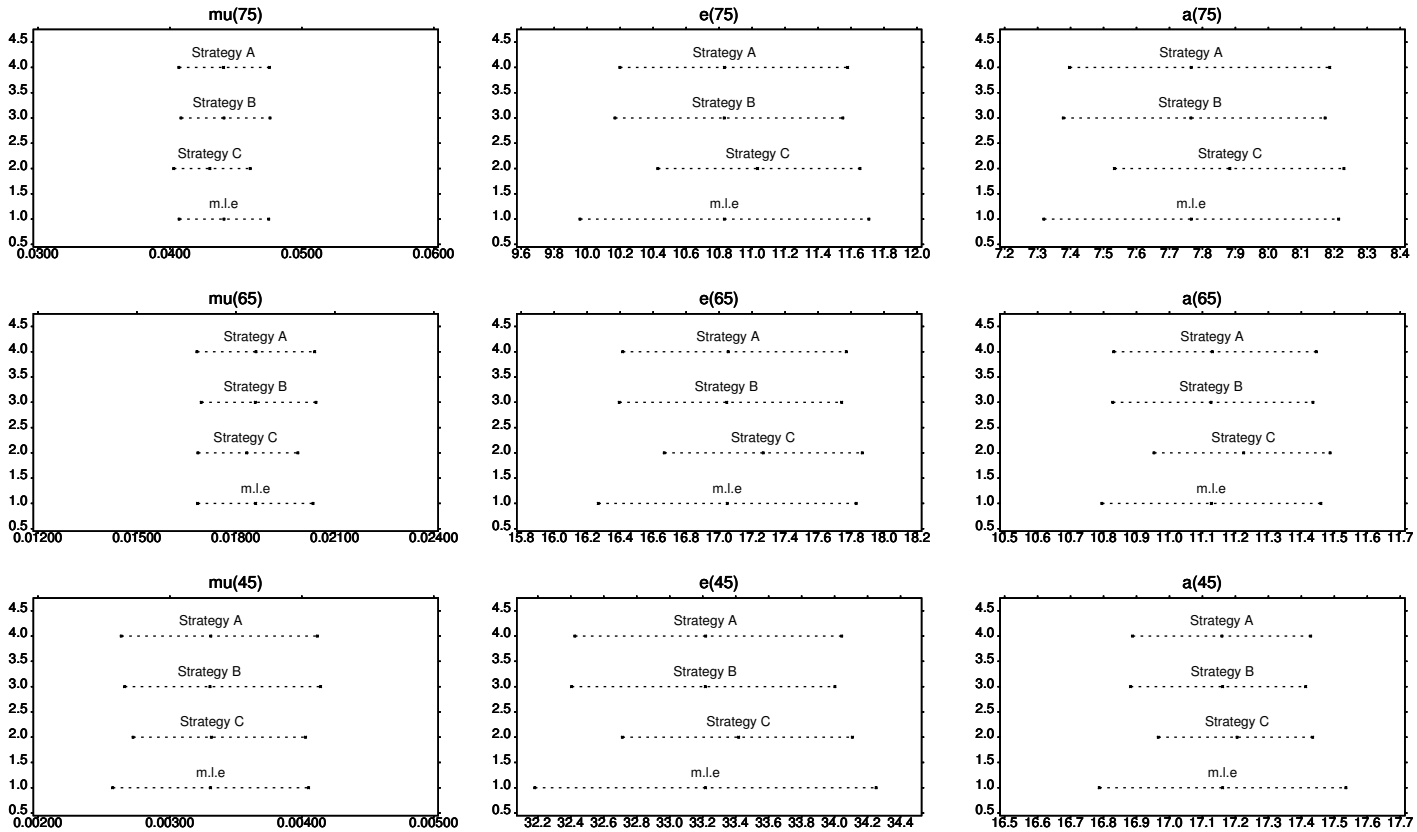


Fig 2. UK pensioners widows 1979-82: simulated 2.5, 50, 97.5 percentiles for each strategy, with maximum likelihood estimates (m.l.e) and 95% CIs. Force of mortality (1st column), life expectancy (2nd column), 4% fixed rate annuity (3rd column); by age 75 (1st row), 65 (2nd row), 45 (3rd row).

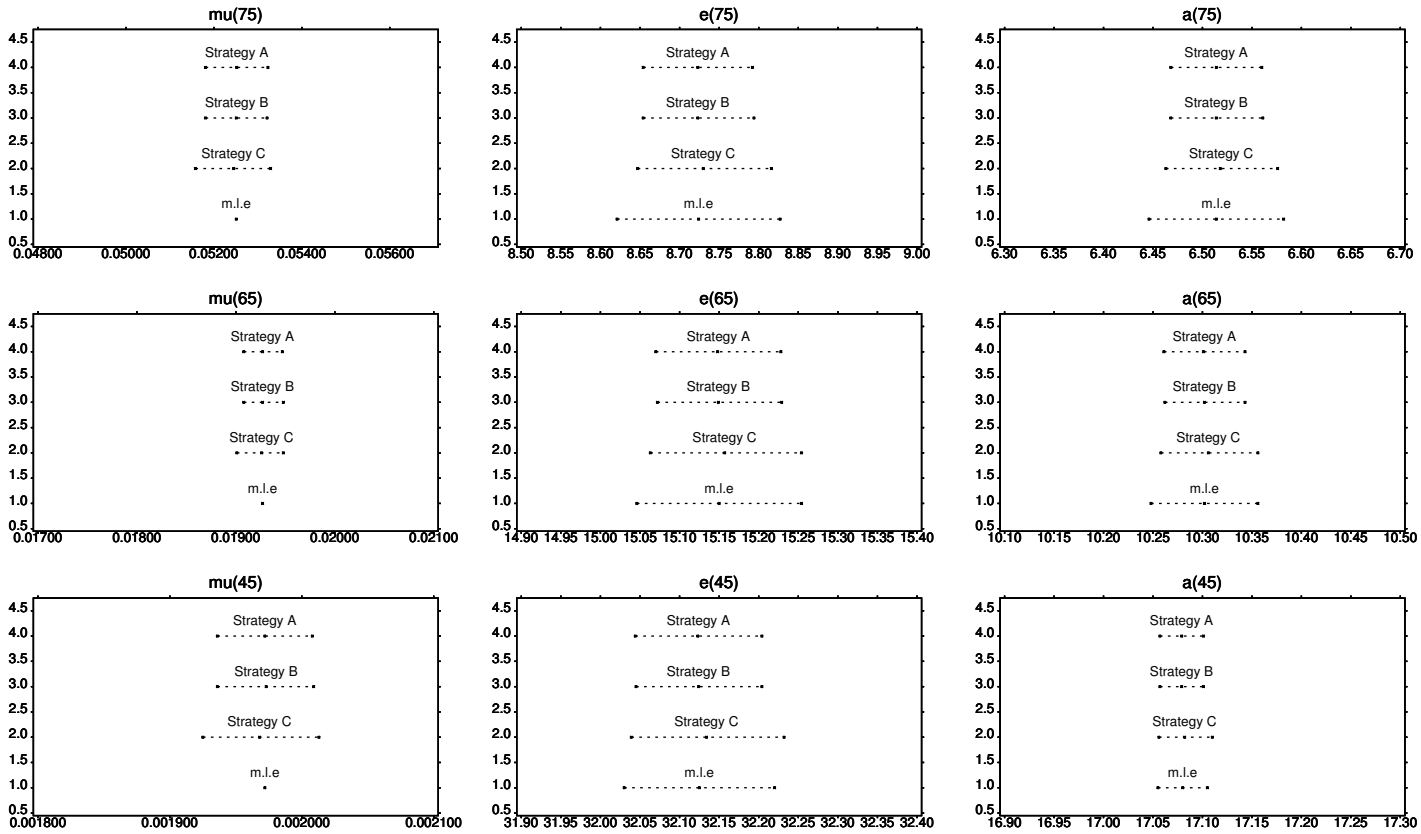


Fig 3. UK male assured lives 1979-82: simulated 2.5, 50, 97.5 percentiles for each strategy, with maximum likelihood estimates (m.l.e) and 95% CIs. Force of mortality (1st column), life expectancy (2nd column), 4% fixed rate annuity (3rd column); by age 75 (1st row), 65 (2nd row), 45 (3rd row).

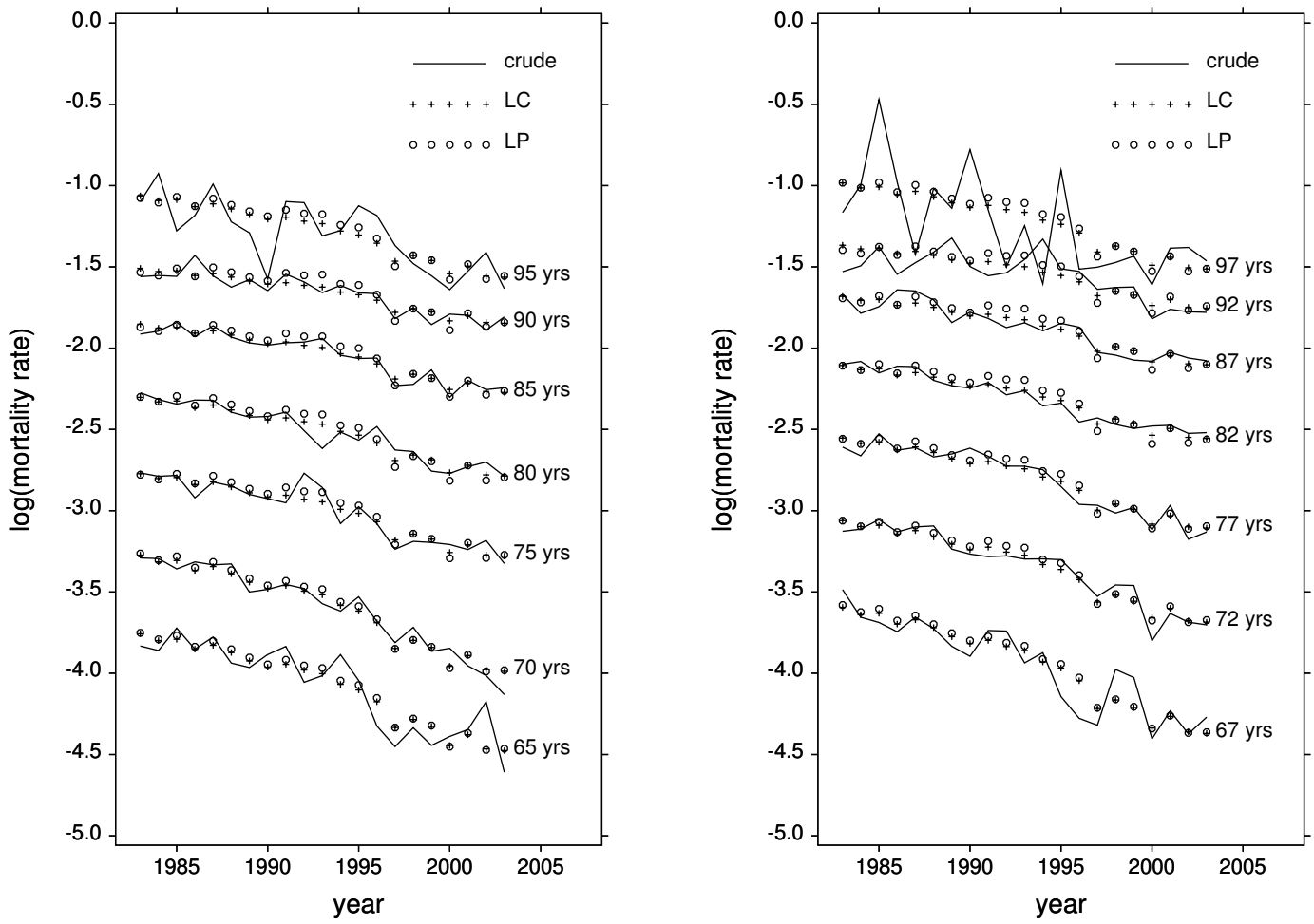


Fig 4. UK male pensioners: crude & fitted log(mortality rates) vs. year by age. Ages 65, 70, 75 ..., 95 left frame, ages 67, 72, 77, ..., 97 right frame.

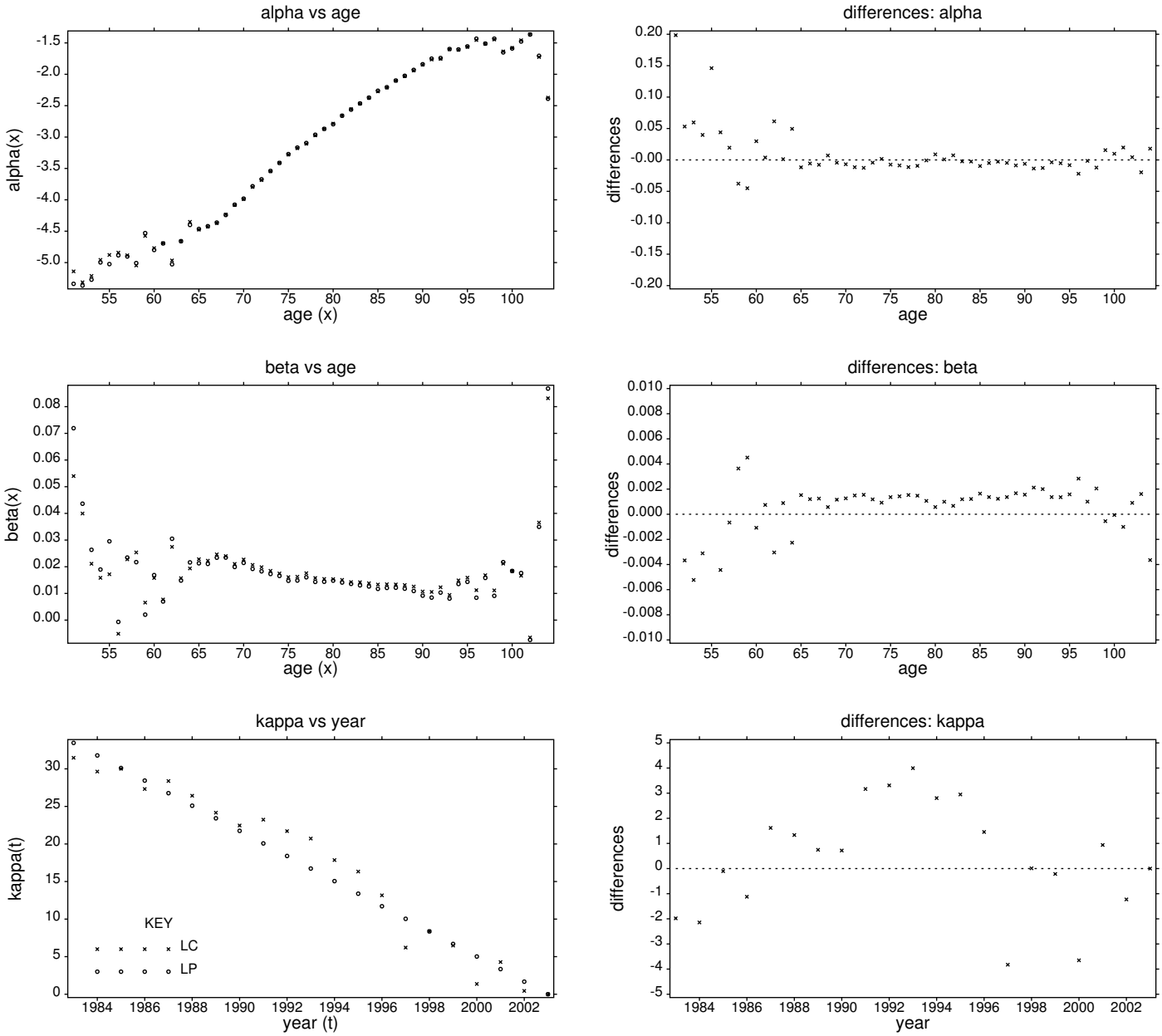
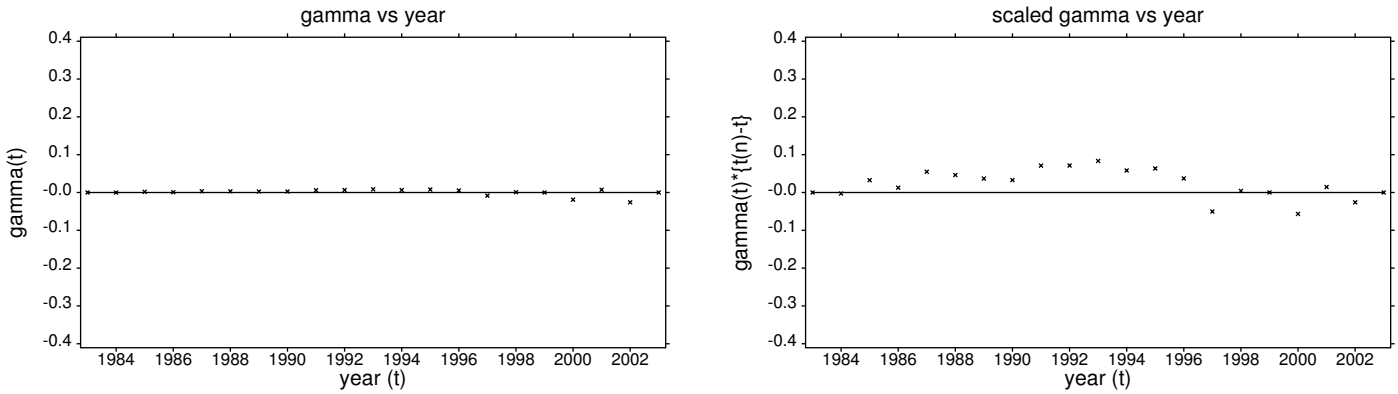
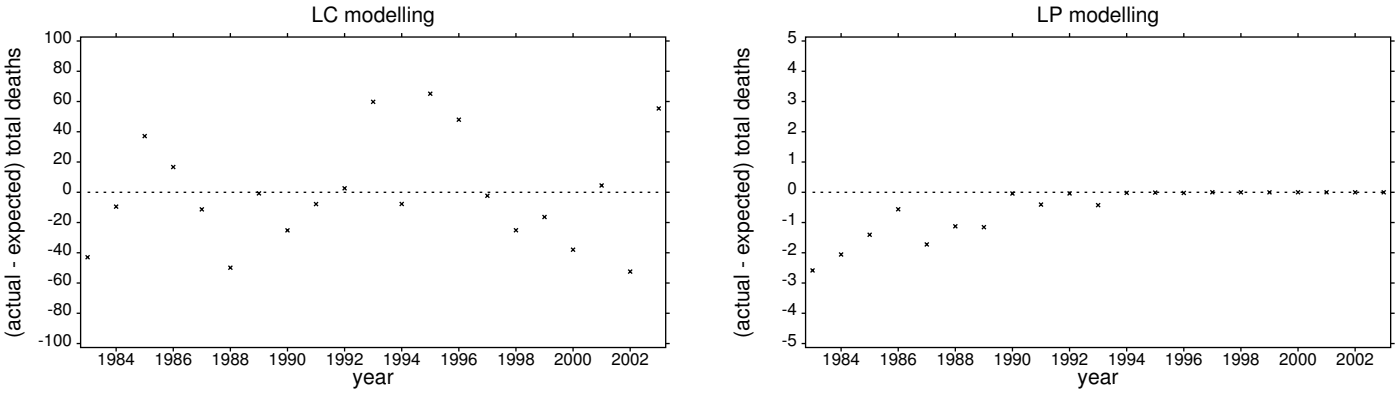


Fig 5. Fitted LC and LP models: superposition of "parameter" estimates.



(a) LP modelling: gamma estimates (left),gamma related estimates (right)



(b) LC and LP modelling: actual minus fitted annual total deaths

Fig 6. LP and LC model fitting: further details.

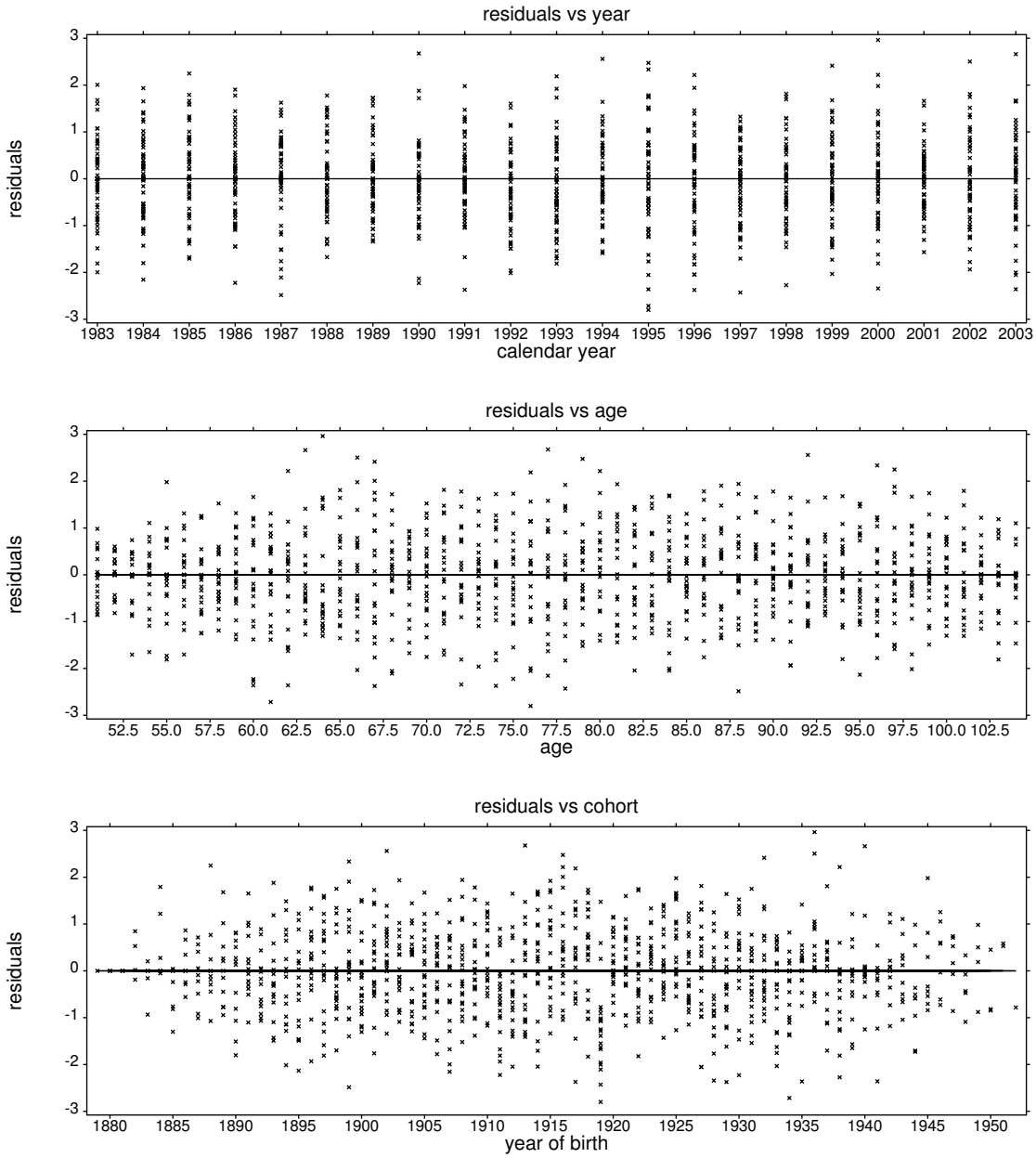
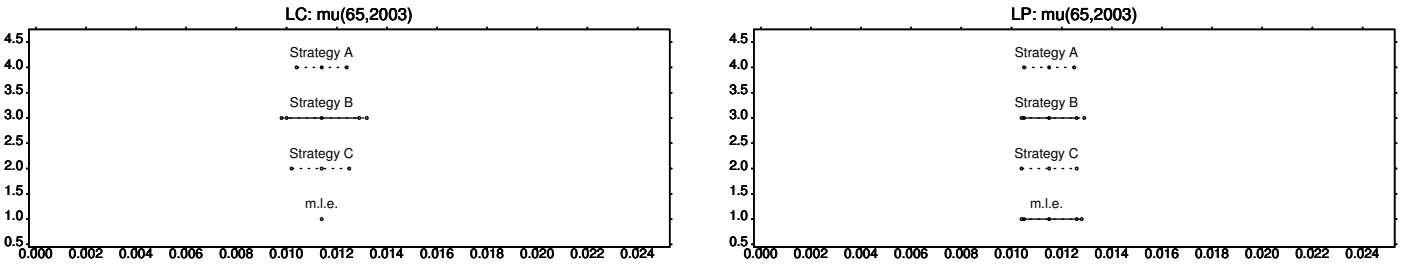
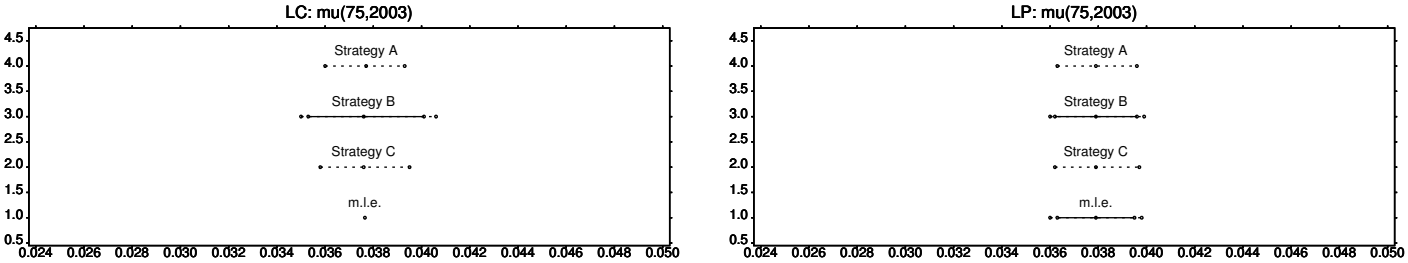


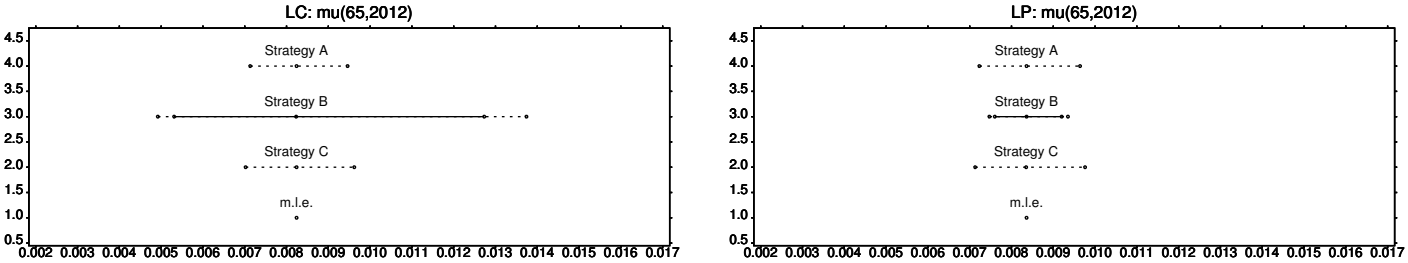
Fig 7. LP modelling: deviance residual plots



(a) Force of mortality- age 65, year 2003

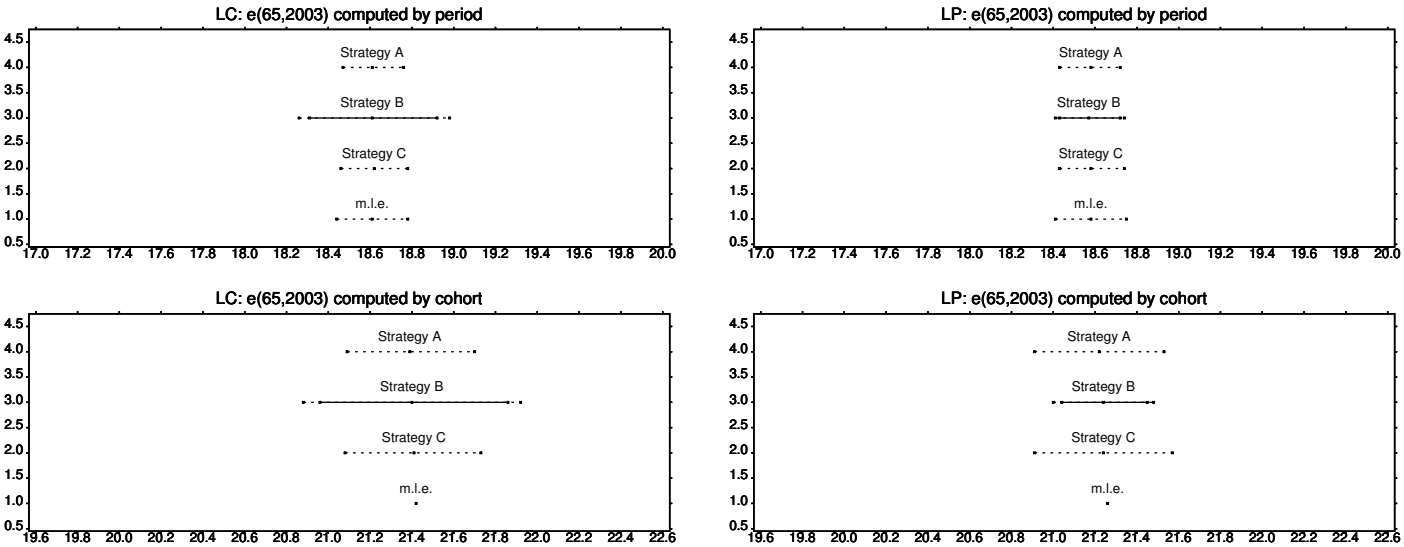


(b) Force of mortality- age 75, year 2003

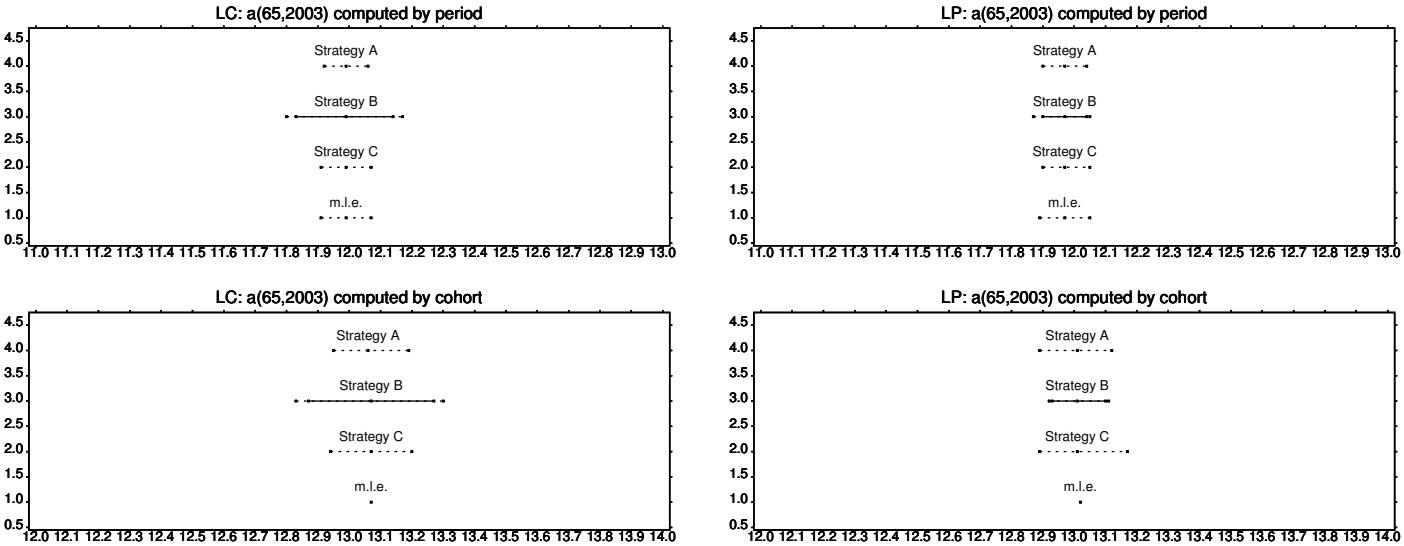


(c) Force of mortality- age 65, year 2012

Fig 8. UK male pensioners: comparison of simulated 2.5, 50, 97.5 percentiles and m.l.e. estimates. Strategy B displays are with/without a scale parameter.

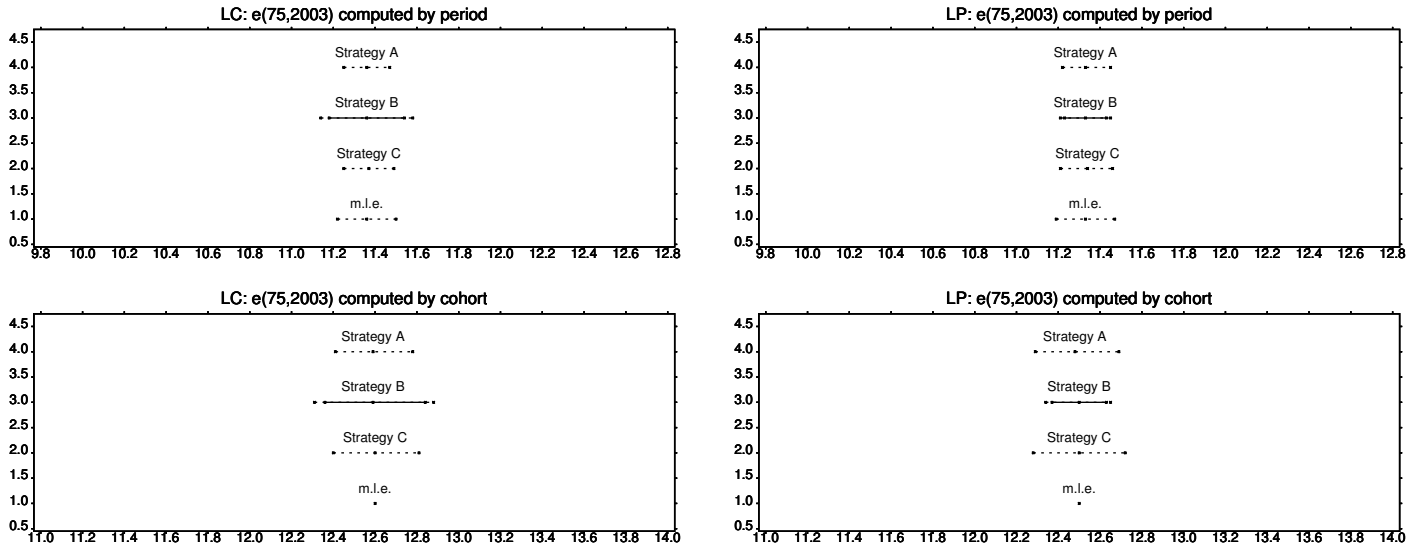


(a) Life expectancy, age 65, year 2003

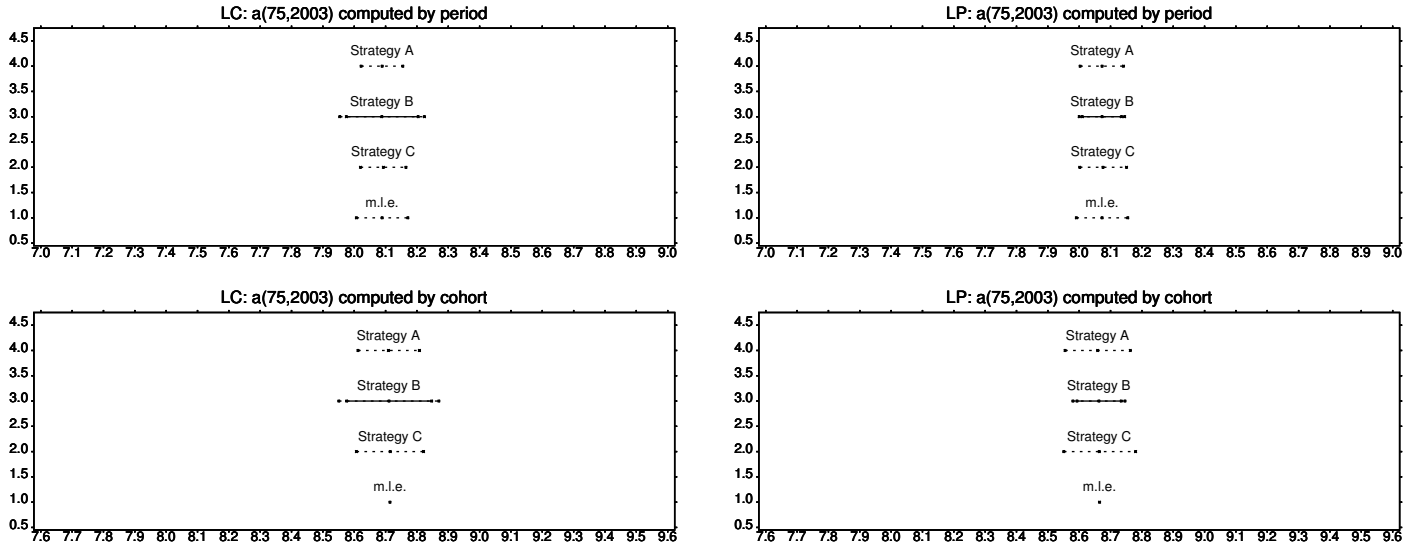


(b) 4 percent fixed rate life annuity, age 65, year 2003

Fig 9. UK male pensioners: comparison of simulated 2.5, 50, 97.5 percentiles and m.l.e. estimates. Strategy B displays are with/without a scale parameter.

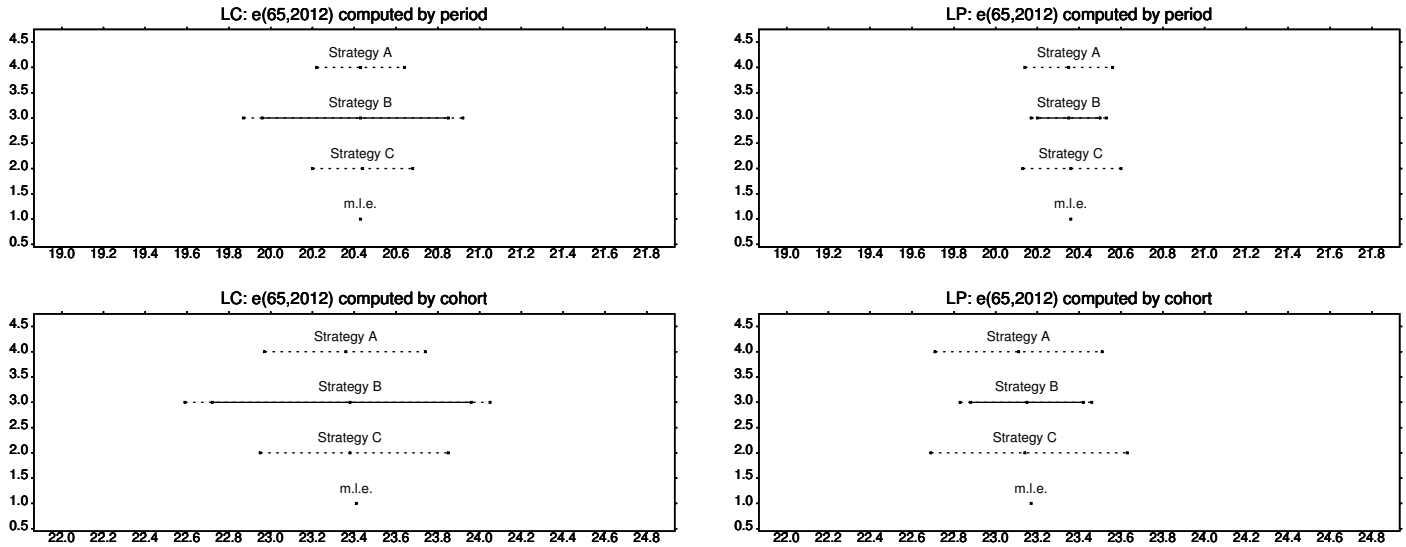


(a) Life expectancy, age 75, year 2003

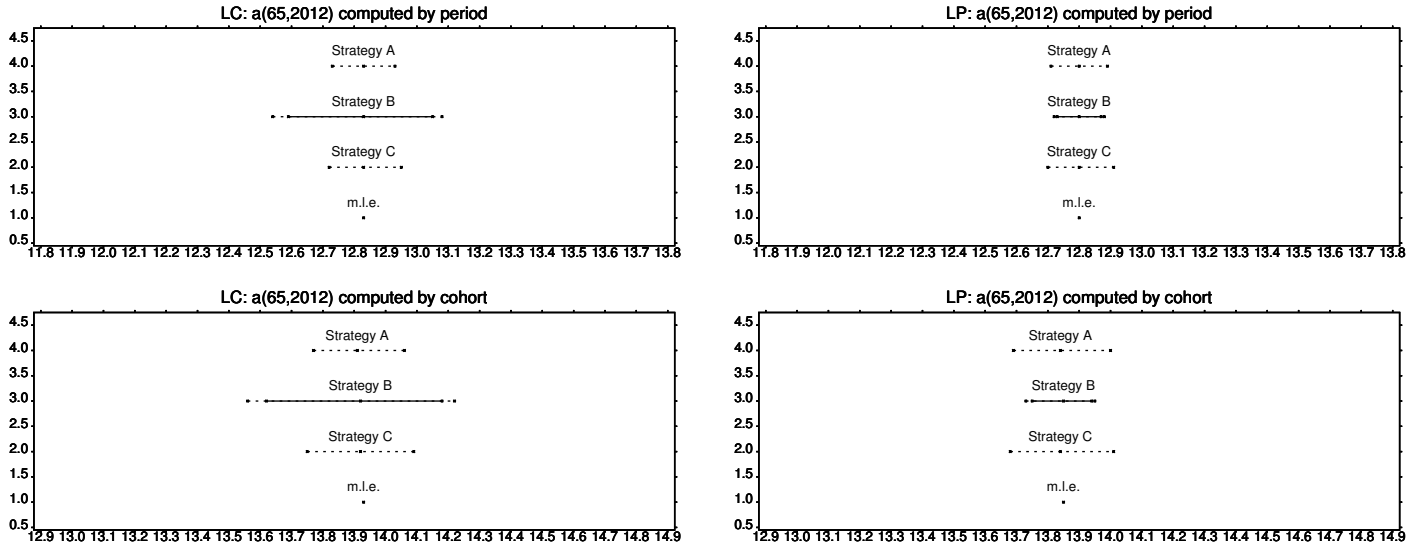


(b) 4 percent fixed rate life annuity, age 75, year 2003

Fig 10. UK male pensioners: comparison of simulated 2.5, 50, 97.5 percentiles and m.l.e. estimates. Strategy B displays are with/without a scale parameter.

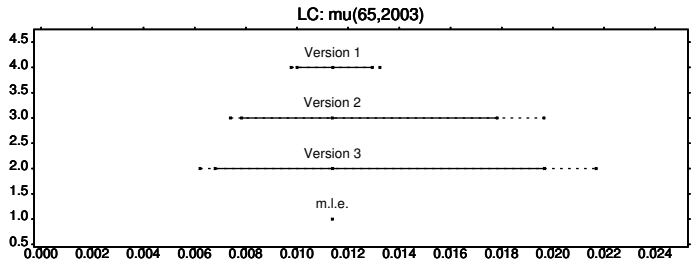


(a) Life expectancy, age 65, year 2012

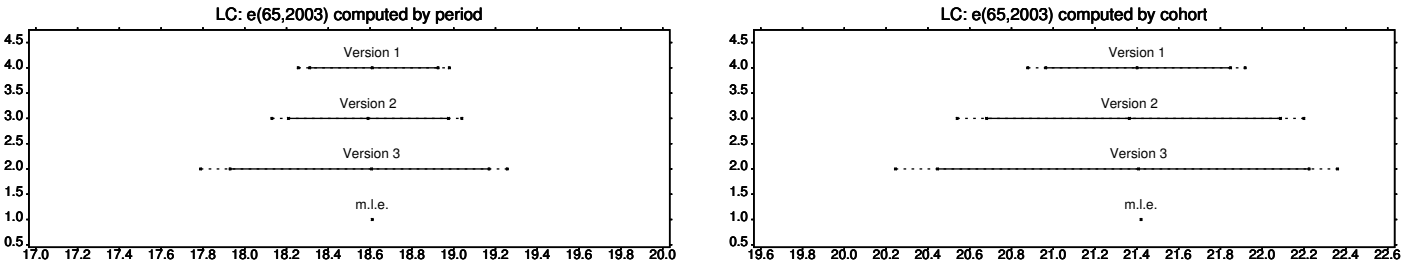


(b) 4 percent fixed rate life annuity, age 65, year 2012

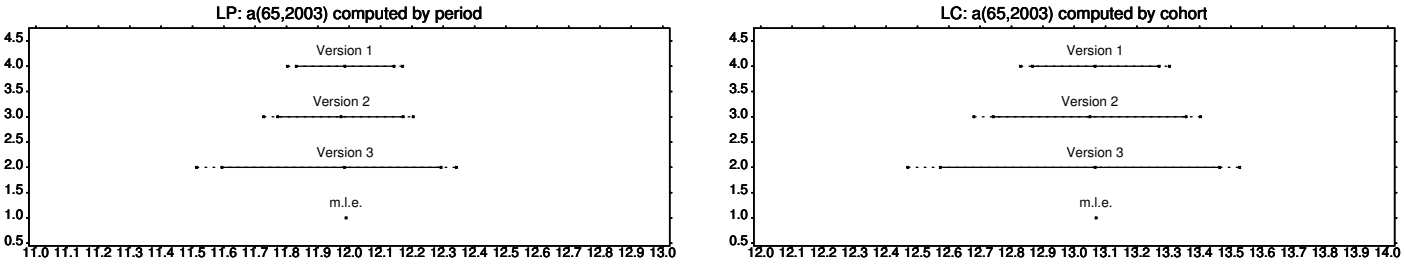
Fig 11. UK male pensioners: comparison of simulated 2.5, 50, 97.5 percentiles and m.l.e. estimates. Strategy B displays are with/without a scale parameter.



(a) Force of mortality, age 65, year 2003

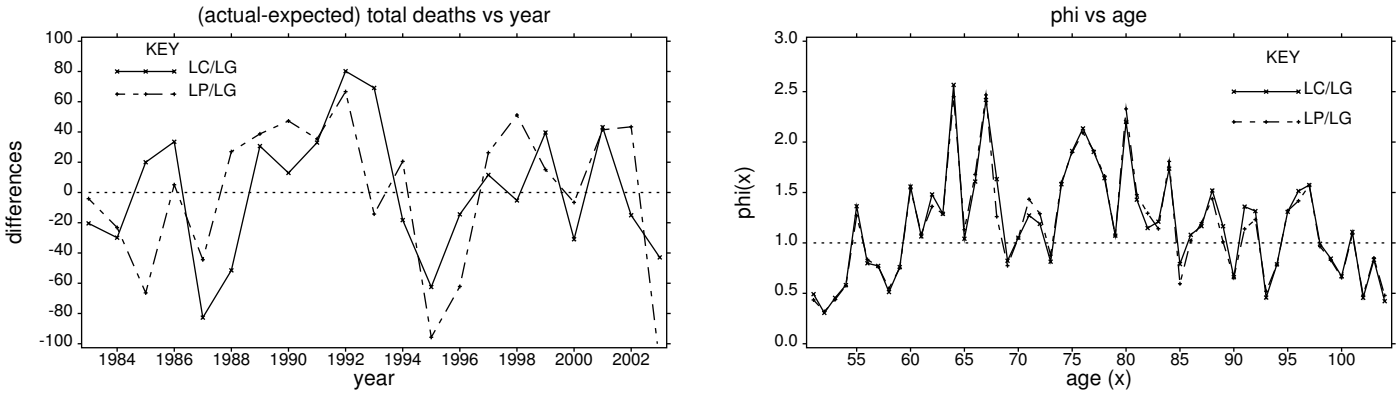


(b) Life expectancy, age 65, year 2003

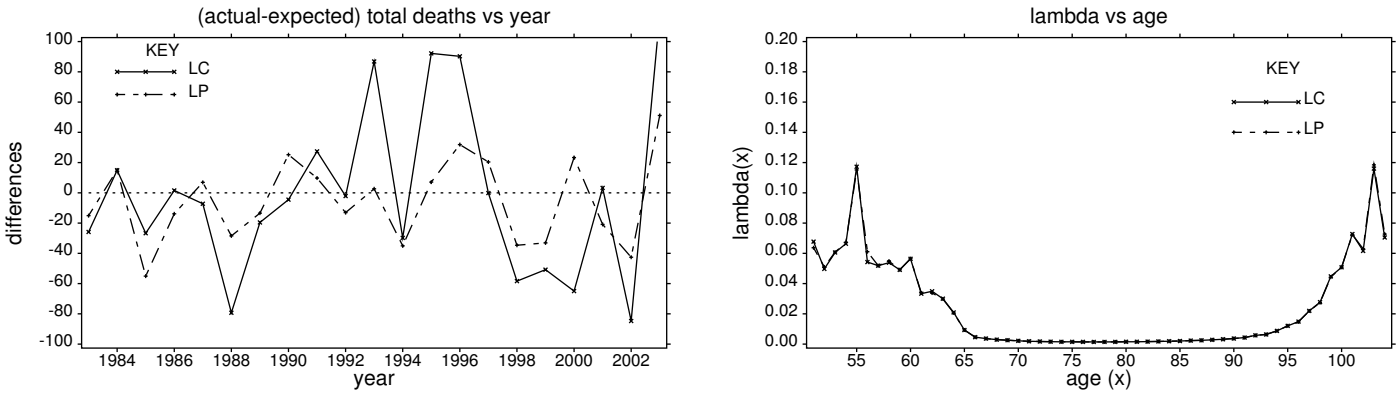


(c) 4 percent fixed rate life annuity, age 65, year 2003

Fig 12. UK male pensioners: comparison of Strategy B simulated 2.5, 50, 97.5 percentiles, with and without a scale parameter, under LC modelling with different versions of the information matrix, (Version 1 as in Figs 8-11).

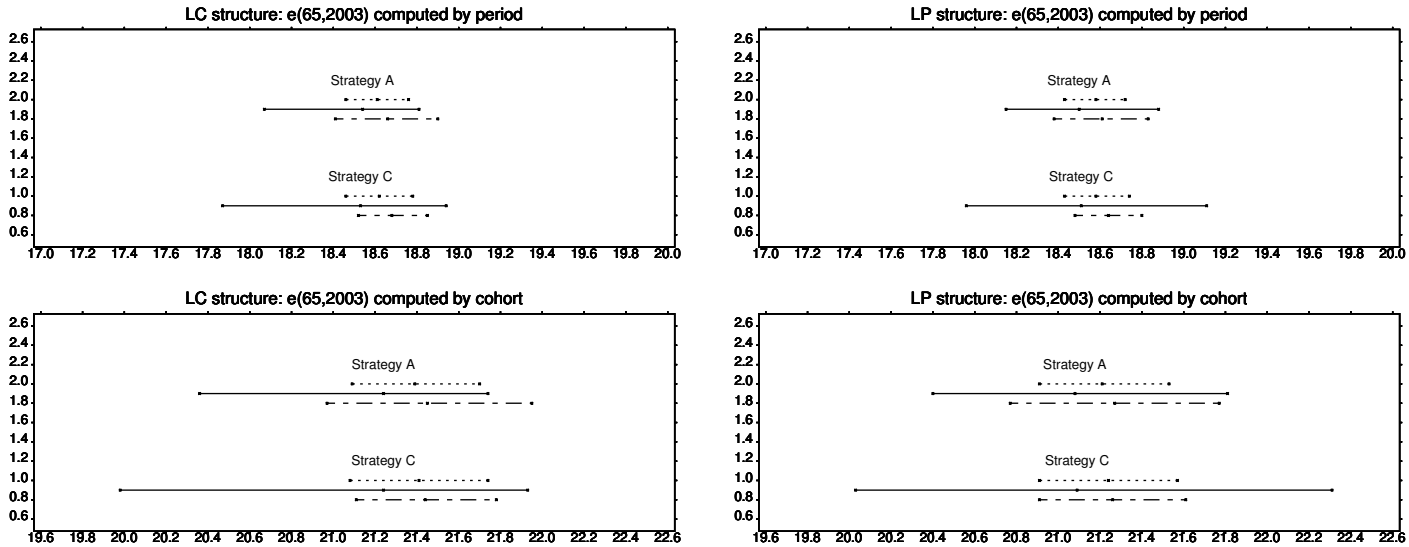


(a) Poisson joint modelling: LC/LG and LP/LG structures, superimposed detail

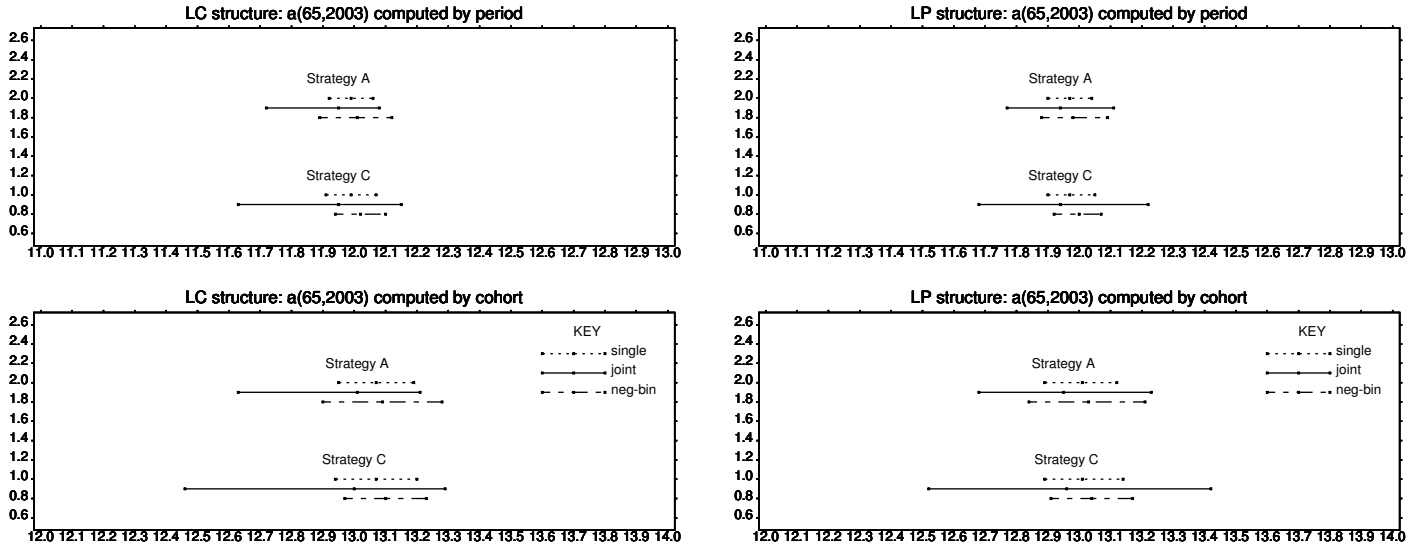


(b) Negative binomial modelling: LC and LP structures, superimposed detail

Fig 13. Left frames: actual minus expected annual total deaths.
Upper right frame: Poisson joint modelling scale parameters.
Lower right frame: negative binomial variance function parameters.

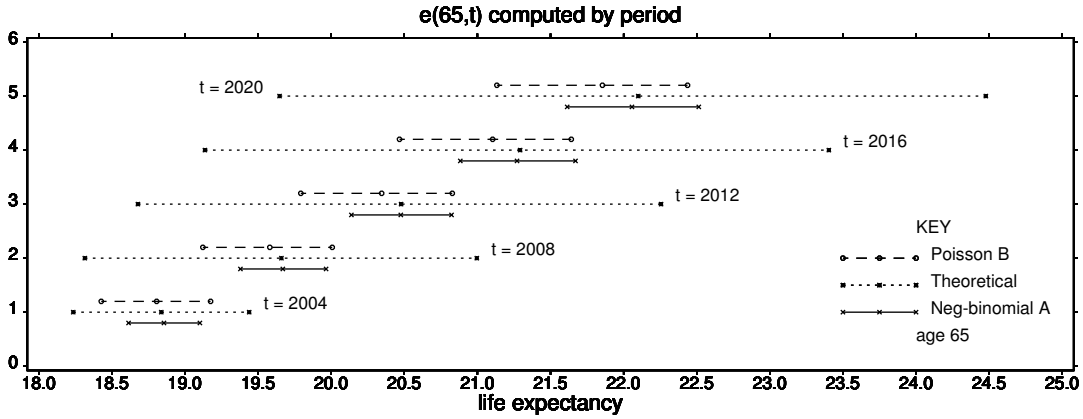


(a) Life expectancy, age 65, year 2003

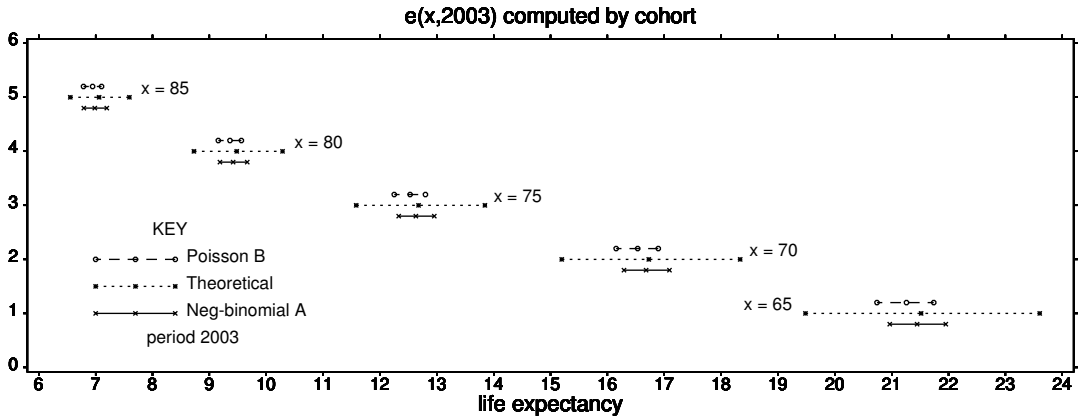


(b) 4 percent fixed rate life annuity, age 65, year 2003

Fig 14. UK male pensioners: comparison of Poisson joint & single models and the negative binomial model; simulated 2.5, 50, 97.5 percentiles.



(a) computations by period, age 65, various periods (t > 2003).



(b) computations by cohort, period 2003, various ages (x).

Fig 15. UK male pensioners: predicted life expectancies. LC structure. Comparison 2.5, 50, 97.5 precentiles: (i) Strategy B, Poisson with scale parameter (version 1). (ii) Strategy A, negative binomial. (iii) Theoretical based counterpart.

Cass Business School
City of London

FACULTY OF ACTUARIAL SCIENCE AND INSURANCE

Actuarial Research Papers since 2001

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

150.	Available August 2004	Finite time Ruin Probabilities for Continuous Claims Severities	D. Dimitrova Z. Ignatov V. Kaishev
151.	August 2004	Application of Stochastic Methods in the Valuation of Social Security Pension Schemes. ISBN 1 901615 72 3	Subramaniam Iyer
152.	October 2003.	Guarantees in with-profit and Unitized with profit Life Insurance Contracts; Fair Valuation Problem in Presence of the Default Option[1]. ISBN 1-901615-73-1	Laura Ballotta Steven Haberman Nan Wang
153.	December 2003	Lee-Carter Mortality Forecasting Incorporating Bivariate Time Series. ISBN 1-901615-75-8	Arthur E. Renshaw Steven Haberman
154.	March 2004.	Operational Risk with Bayesian Networks Modelling. ISBN 1-901615-76-6	Robert G. Cowell Yuen Y, Khuen Richard J. Verrall
155.	March 2004.	The Income Drawdown Option: Quadratic Loss. ISBN 1 901615 7 4	Russell Gerrard Steven Haberman Bjorn Hojgarrd Elena Vigna
156.	April 2004	An International Comparison of Long-Term Care Arrangements. An Investigation into the Equity, Efficiency and sustainability of the Long-Term Care Systems in Germany, Japan, Sweden, the United Kingdom and the United States. ISBN 1 901615 78 2	Martin Karlsson Les Mayhew Robert Plumb Ben D. Rickayzen
157.	June 2004	Alternative Framework for the Fair Valuation of Participating Life Insurance Contracts. ISBN 1 901615-79-0	Laura Ballotta
158.	July 2004.	An Asset Allocation Strategy for a Risk Reserve considering both Risk and Profit. ISBN 1 901615-80-4	Nan Wang
159.	December 2004	Upper and Lower Bounds of Present Value Distributions of Life Insurance Contracts with Disability Related Benefits. ISBN 1 901615-83-9	Jaap Spreeuw
160.	January 2005	Mortality Reduction Factors Incorporating Cohort Effects. ISBN 1 90161584 7	Arthur E. Renshaw Steven Haberman
161.	February 2005	The Management of De-Cumulation Risks in a Defined Contribution Environment. ISBN 1 901615 85 5.	Russell J. Gerrard Steven Haberman Elena Vigna
162.	May 2005	The IASB Insurance Project for Life Insurance Contracts: Impart on Reserving Methods and Solvency Requirements. ISBN 1-901615 86 3.	Laura Ballotta Giorgia Esposito Steven Haberman
163.	September 2005	Asymptotic and Numerical Analysis of the Optimal Investment Strategy for an Insurer. ISBN 1-901615-88-X	Paul Emms Steven Haberman
164.	October 2005.	Modelling the Joint Distribution of Competing Risks Survival Times using Copula Functions. I SBN 1-901615-89-8	Vladimir Kaishev Dimitrina S, Dimitrova Steven Haberman
165.	November 2005.	Excess of Loss Reinsurance Under Joint Survival Optimality. ISBN1-901615-90-1	Vladimir K. Kaishev Dimitrina S. Dimitrova
166.	November 2005.	Lee-Carter Goes Risk-Neutral. An Application to the Italian Annuity Market. ISBN 1-901615-91-X	Enrico Biffis Michel Denuit

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

Statistical Research Papers

Report Number	Date	Publication Title	Author
1.	December 1995.	Some Results on the Derivatives of Matrix Functions. ISBN 1 874 770 83 2	P. Sebastiani
2.	March 1996	Coherent Criteria for Optimal Experimental Design. ISBN 1 874 770 86 7	A.P. Dawid P. Sebastiani
3.	March 1996	Maximum Entropy Sampling and Optimal Bayesian Experimental Design. ISBN 1 874 770 87 5	P. Sebastiani H.P. Wynn
4.	May 1996	A Note on D-optimal Designs for a Logistic Regression Model. ISBN 1 874 770 92 1	P. Sebastiani R. Settimi
5.	August 1996	First-order Optimal Designs for Non Linear Models. ISBN 1 874 770 95 6	P. Sebastiani R. Settimi
6.	September 1996	A Business Process Approach to Maintenance: Measurement, Decision and Control. ISBN 1 874 770 96 4	Martin J. Newby
7.	September 1996.	Moments and Generating Functions for the Absorption Distribution and its Negative Binomial Analogue. ISBN 1 874 770 97 2	Martin J. Newby
8.	November 1996.	Mixture Reduction via Predictive Scores. ISBN 1 874 770 98 0	Robert G. Cowell.
9.	March 1997.	Robust Parameter Learning in Bayesian Networks with Missing Data. ISBN 1 901615 00 6	P.Sebastiani M. Ramoni
10.	March 1997.	Guidelines for Corrective Replacement Based on Low Stochastic Structure Assumptions. ISBN 1 901615 01 4.	M.J. Newby F.P.A. Coolen
11.	March 1997	Approximations for the Absorption Distribution and its Negative Binomial Analogue. ISBN 1 901615 02 2	Martin J. Newby
12.	June 1997	The Use of Exogenous Knowledge to Learn Bayesian Networks from Incomplete Databases. ISBN 1 901615 10 3	M. Ramoni P. Sebastiani
13.	June 1997	Learning Bayesian Networks from Incomplete Databases. ISBN 1 901615 11 1	M. Ramoni P.Sebastiani
14.	June 1997	Risk Based Optimal Designs. ISBN 1 901615 13 8	P.Sebastiani H.P. Wynn
15.	June 1997.	Sampling without Replacement in Junction Trees. ISBN 1 901615 14 6	Robert G. Cowell
16.	July 1997	Optimal Overhaul Intervals with Imperfect Inspection and Repair. ISBN 1 901615 15 4	Richard A. Dagg Martin J. Newby
17.	October 1997	Bayesian Experimental Design and Shannon Information. ISBN 1 901615 17 0	P. Sebastiani. H.P. Wynn
18.	November 1997.	A Characterisation of Phase Type Distributions. ISBN 1 901615 18 9	Linda C. Wolstenholme
19.	December 1997	A Comparison of Models for Probability of Detection (POD) Curves. ISBN 1 901615 21 9	Wolstenholme L.C
20.	February 1999.	Parameter Learning from Incomplete Data Using Maximum Entropy I: Principles. ISBN 1 901615 37 5	Robert G. Cowell

21.	November 1999	Parameter Learning from Incomplete Data Using Maximum Entropy II: Application to Bayesian Networks. ISBN 1 901615 40 5	Robert G. Cowell
22.	March 2001	FINEX : Forensic Identification by Network Expert Systems. ISBN 1 901615 60X	Robert G.Cowell
23.	March 2001.	Wren Learning Bayesian Networks from Data, using Conditional Independence Tests is Equivalant to a Scoring Metric ISBN 1 901615 61 8	Robert G Cowell
24.	August 2004	Automatic, Computer Aided Geometric Design of Free-Knot, Regression Splines. ISBN 1-901615-81-2	Vladimir K Kaishev, Dimitrina S.Dimitrova, Steven Haberman Richard J. Verrall
25.	December 2004	Identification and Separation of DNA Mixtures Using Peak Area Information. ISBN 1-901615-82-0	R.G.Cowell S.L.Lauritzen J Mortera,
26.	November 2005.	The Quest for a Donor : Probability Based Methods Offer Help. ISBN 1-90161592-8	P.F.Mostad T. Egeland., R.G. Cowell V. Bosnes Ø. Braaten
27.	February 2006	Identification and Separation of DNA Mixtures Using Peak Area Information. (Updated Version of Research Report Number 25). ISBN 1-901615-94-4	R.G.Cowell S.L.Lauritzen J Mortera,
28.	October 2006	Geometrically Designed, Variable Knot Regression Splines : Asymptotics and Inference. ISBN 1-905752-02-4	Vladimir K Kaishev, Dimitrina S.Dimitrova, Steven Haberman Richard J. Verrall
29.	October 2006	Geometrically Designed, Variable Knot Regression Splines : Variation Diminishing Optimality of Knots. ISBN 1-905752-03-2	Vladimir K Kaishev, Dimitrina S.Dimitrova, Steven Haberman Richard J. Verrall

Papers can be downloaded from

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

Faculty of Actuarial Science and Insurance

Actuarial Research Club

The support of the corporate members

- CGNU Assurance
- English Matthews Brockman
- Government Actuary's Department

is gratefully acknowledged.

Copyright 2007 © Faculty of Actuarial Science and Insurance,
Cass Business School
106 Bunhill Row,
London EC1Y 8TZ.

ISBN 978-1-905752-10-2