

THE LEE-CARTER MODEL FOR MORTALITY DYNAMICS: RECENT DEVELOPMENTS

**Steven Haberman and Arthur Renshaw
Cass Business School, City University**

IFID/MITACS – Conference on Financial Engineering for Actuarial Mathematics
9 – 10 November 2008
Fields Institute, University of Toronto

AGENDA

- Introductory Comments
- Nature of Data
- Basic Lee-Carter Model
- Age-Period-Cohort Model
- Risk Measurement
- Back Testing
- Concluding Comments

WHY ARE WE INTERESTED IN MORTALITY?

- Life expectancy is increasing – a good news story
- Future trend is uncertain – leading to “longevity risk”
- Systematic risk for DB pension plans and annuity providers
(noting shift from DB to DC and growing importance of annuities)

WHY ARE WE INTERESTED NOW IN MORTALITY?

- Life expectancy increases have been a long term phenomenon
- Impact on financial institutions hidden by high investment returns.
Since 2000: lower equity returns and low interest rates.

[Note closure of Equitable Life to new business in 2000]

WHY DO WE NEED TO MODEL MORTALITY DYNAMICS?

- Setting prudent reserves for annuity providers and funding strategies for DB pension plans
- Good risk management practice
- Dealing with insurance contracts with guarantees and embedded options e.g. GAOs in UK, Variable Annuities in US
- Development of longevity – linked securities : pricing and hedging.

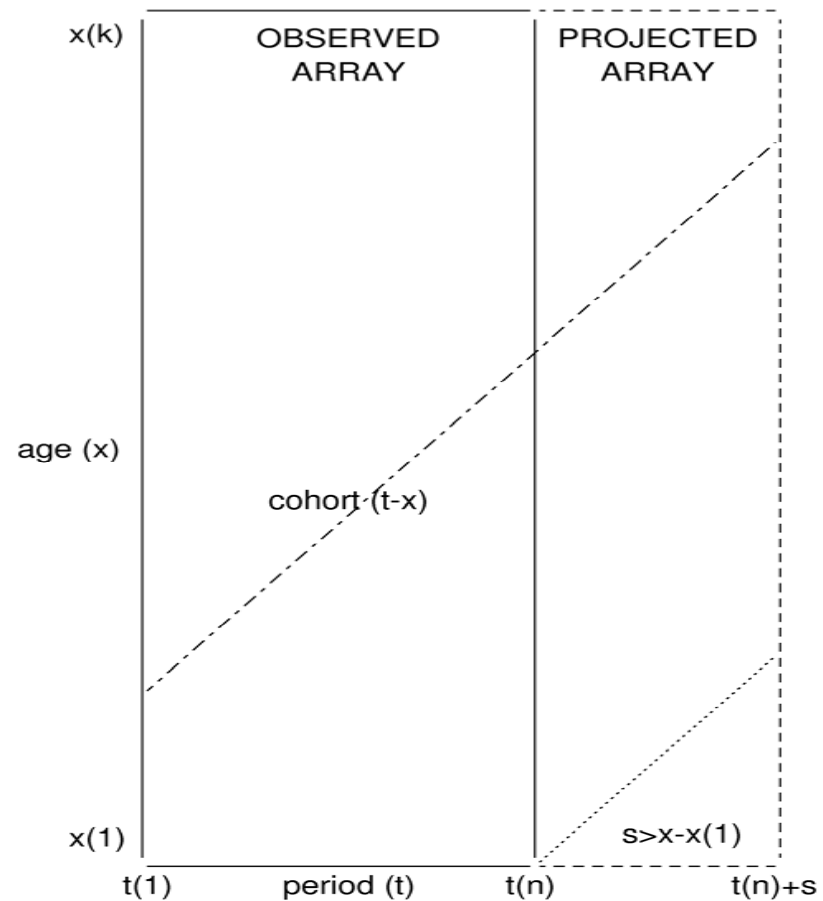
Deterministic models are not fit for purpose.

Hence, we need to discuss stochastic models.

ILLUSTRATION OF DATA CONFIGURATION

Typical rectangular data array and targeted projected array.

Typical observed and projected data arrays



NOTATION

Data: (d_{xt}, e_{xt})

d_{xt} = number of deaths at age x and time t

e_{xt} = matching exposure to risk of death

with empirical central mortality rate

$$\hat{m}_{xt} = d_{xt} / e_{xt} .$$

LEE CARTER MODELS: BASE VERSION

STRUCTURE

One of the benchmark demographic models used for mortality modelling and projections in many countries. Lee and Carter (1992) proposed:

$$\ln m_{xt} = \eta_{xt} + \varepsilon_{xt}, \quad \eta_{xt} = \alpha_x + \beta_x \kappa_t,$$

where the ε_{xt} are IID $N(0, \sigma^2)$ variables.

This is a regression framework with no observable quantities on the RHS.

STRUCTURE (cont)

Structure is invariant under the transformations

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

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

and is made identifiable using the following constraints (which are not unique):

$$\sum_{t=t_1}^{t_n} \kappa_t = 0, \quad \sum_x \beta_x = 1, \quad \text{and which imply the least squares estimator}$$

$$\hat{\alpha}_x = \frac{1}{t_n - t_1 + 1} \sum_{t=t_1}^{t_n} \ln \hat{m}_{xt}$$

INTERPRETATION OF PARAMETERS

α_x : 'average' of $\log m_{xt}$ over time t so that $\exp \alpha_x$ represents the general shape of the age-specific mortality profile.

κ_t : underlying time trend.

β_x : sensitivity of the logarithm of the hazard rate at age x to the time trend represented by κ_t .

ε_{xt} : effects not captured by the model.

FITTING BY SVD

A two-stage estimation process: estimate α_x as above. Estimate κ_t and β_x as the 1st right and 1st left singular vectors in the SVD of the matrix $[\log \hat{m}_{xt} - \hat{\alpha}_x]$.

Thus

$$\log(\hat{m}_{xt}) = \hat{\alpha}_x + s_1 u_1(x) v_1(t) + \sum_{i>1} s_i u_i(x) v_i(t)$$

where

s_i, u_i, v_i = (ordered) singular values and vectors

and

$$\hat{\beta}_x \hat{\kappa}_t = s_1 u_1(x) v_1(t)$$

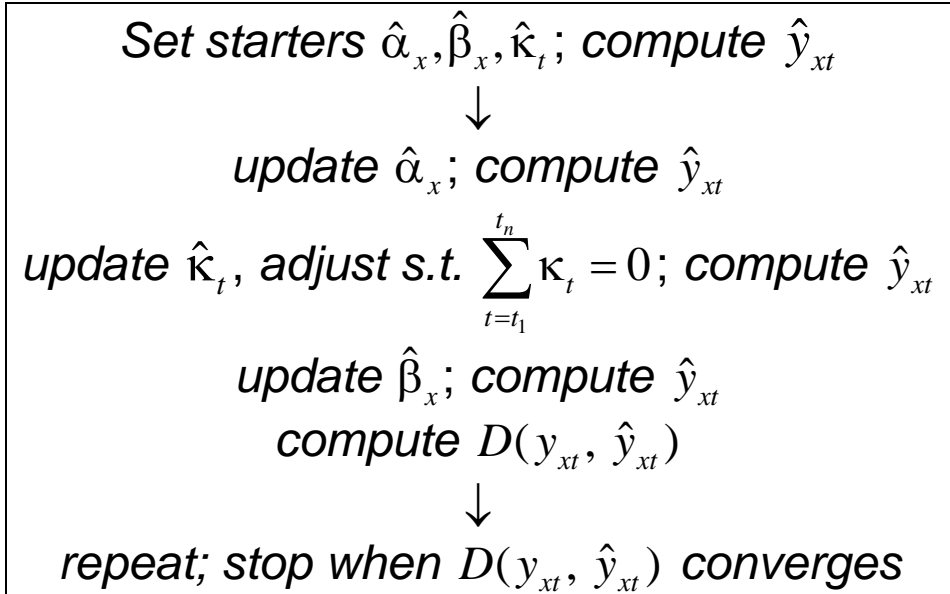
subject to the constraints on κ_t and β_x .

Finally, $\hat{\kappa}_t$ are adjusted so that

$$\sum_{all, x} d_{xt} = \sum_{all, x} \hat{d}_{xt} \quad \forall t. \quad \text{where } \hat{d}_{xt} = e_{xt} \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t).$$

FITTING BY WEIGHTED LEAST SQUARES (GAUSSIAN)

Perform the iterative process



Where $y_{xt} = \log \hat{m}_{xt}$, $\hat{y}_{xt} = \hat{\eta}_{xt}$, $D(y_{xt}, \hat{y}_{xt}) = \sum_{x,t} w_{xt} (y_{xt} - \hat{y}_{xt})^2$

with weights

$$w_{xt} = d_{xt} \text{ (or } = 1 \text{)}.$$

For a typical parameter, we use the updating algorithm:

$$\text{updated}(\theta) = \theta - \frac{\partial D}{\partial \theta} \bigg/ \frac{\partial^2 D}{\partial \theta^2}.$$

POISSON BILINEAR MODEL

$$Y_{xt} = D_{xt}, \quad E(Y_{xt}) = e_{xt} \exp(\alpha_x + \beta_x \kappa_t), \quad \text{Var}(Y_{xt}) = \phi E(Y_{xt})$$

with log-link and non-linear predictor

$$\eta_{xt} = \log e_{xt} + \alpha_x + \beta_x \kappa_t.$$

Perform iterative process with

$$y_{xt} = d_{xt}, \quad \hat{y}_{xt} = \hat{d}_{xt} = e_{xt} \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{\kappa}_t)$$

$$D(d_{xt}, \hat{d}_{xt}) = \sum_{x,t} 2w_{xt} \left\{ d_{xt} \log \left(\frac{d_{xt}}{\hat{d}_{xt}} \right) - (d_{xt} - \hat{d}_{xt}) \right\}$$

with weights

$$w_{xt} = \begin{cases} 1, & e_{xt} > 0 \\ 0, & e_{xt} = 0 \end{cases}$$

LEE CARTER: BINOMIAL

$$\eta_{xt} = \alpha_x + \beta_x \kappa_t \quad \sum_x \beta_x = 1, \kappa_{t_n} = 0$$

and link functions $\eta_{xt} = g(q_{xt})$

Possible choices of g are:

I. complementary log-log link

$$\eta_{xt} = \log \left\{ -\log(1 - q_{xt}) \right\}$$

II. log-odds link

$$\eta_{xt} = \log \left(\frac{q_{xt}}{1 - q_{xt}} \right)$$

III. probit link

$$\eta_{xt} = \Phi^{-1}(q_{xt})$$

DIAGNOSTICS

- 1) Proportion of the total temporal variance explained by the 1st SVD component:

$$\frac{s_1^2}{\sum_{all, i} s_i^2} \times 100\% .$$

(Not a good indicator of goodness of fit.)

- 2) Standardised deviance residuals

$$\hat{\varepsilon}_{xt} = \text{sign}(y_{xt} - \hat{y}_{xt}) \sqrt{\text{dev}(x, t) / \phi}$$

(we could also use standardised SVD residuals)

- 3) Plot differences between actual total and expected total deaths for each time period, t .

PROJECTIONS

Time series (ARIMA)

$$\{\hat{\kappa}_t : t \in [t_1, t_n]\} \mapsto \{\dot{\kappa}_{t_n+s} : s > 0\}.$$

Construct mortality rate projections

$$\dot{m}_{x,t_n+s} = \hat{m}_{xt_n} \exp\{\hat{\beta}_x (\dot{\kappa}_{t_n+s} - \hat{\kappa}_{t_n})\}, s > 0$$

by alignment with the latest available mortality rates.

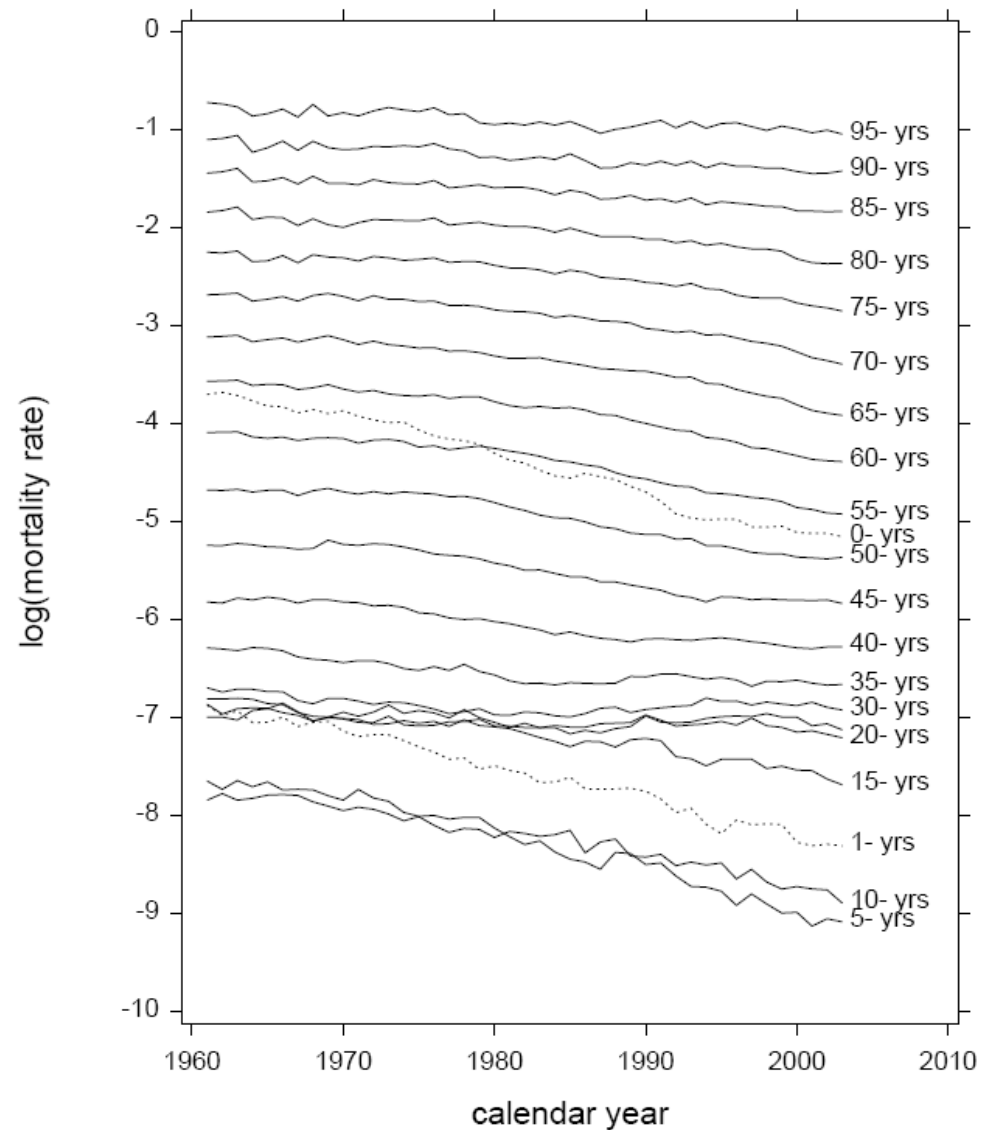
Note

$$F(x, t_n + s) = \exp\{\hat{\beta}_x (\dot{\kappa}_{t_n+s} - \hat{\kappa}_{t_n})\}, s > 0$$

is a mortality reduction factor, as widely used in the UK (e.g. by CMI Bureau) and elsewhere.

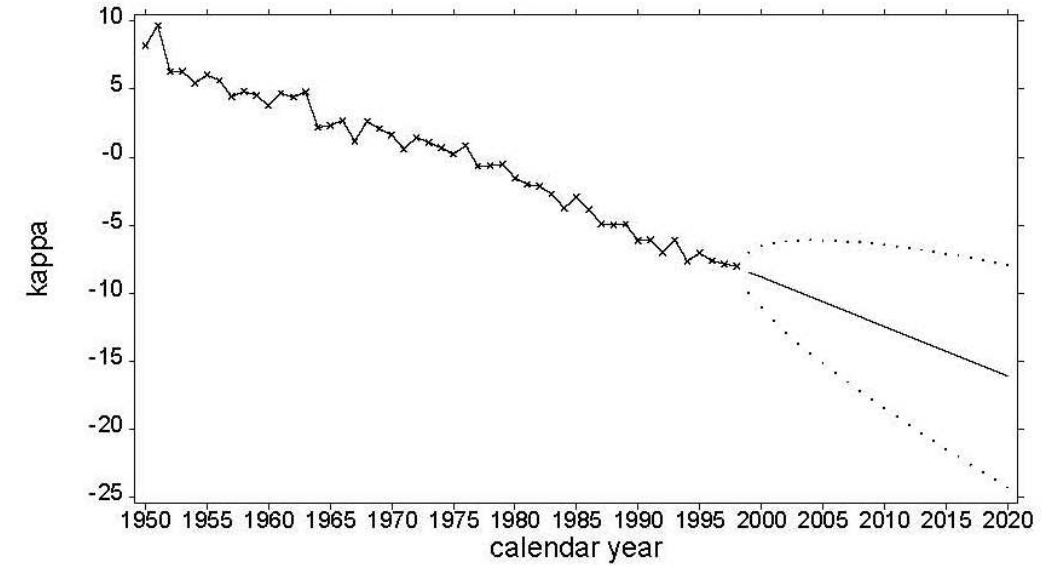
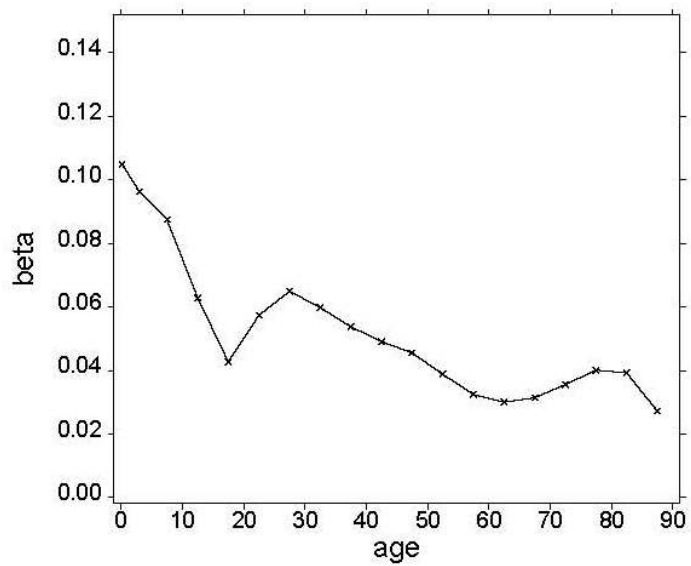
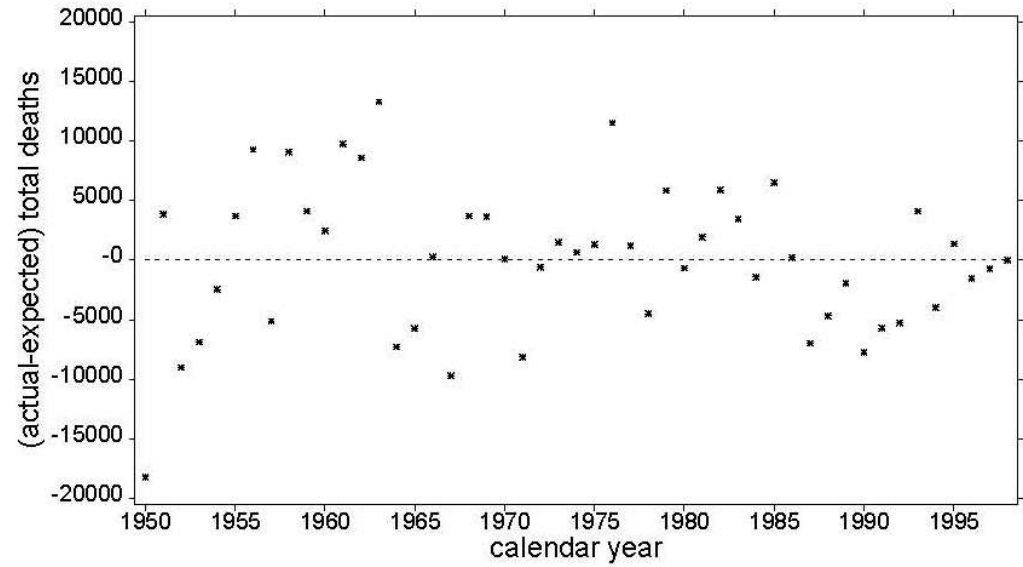
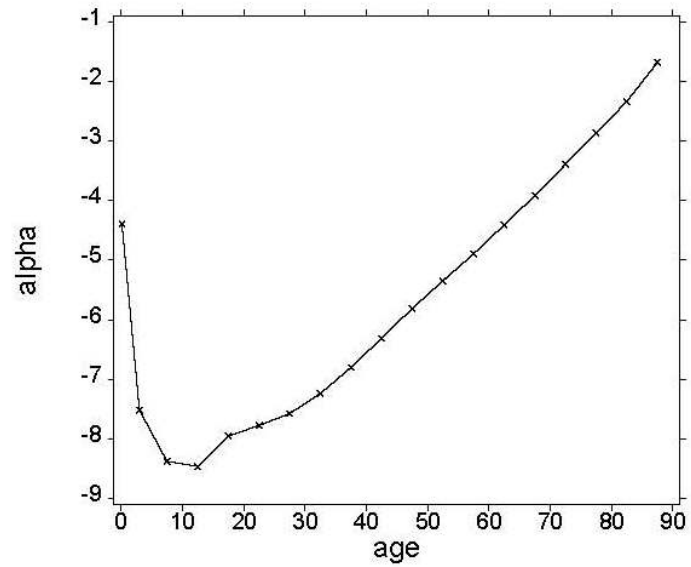
For ARIMA(0,1,0) with drift parameter λ

$$F(x, t_n + s) = \exp(\hat{\beta}_x \hat{\lambda} s), s > 0,$$

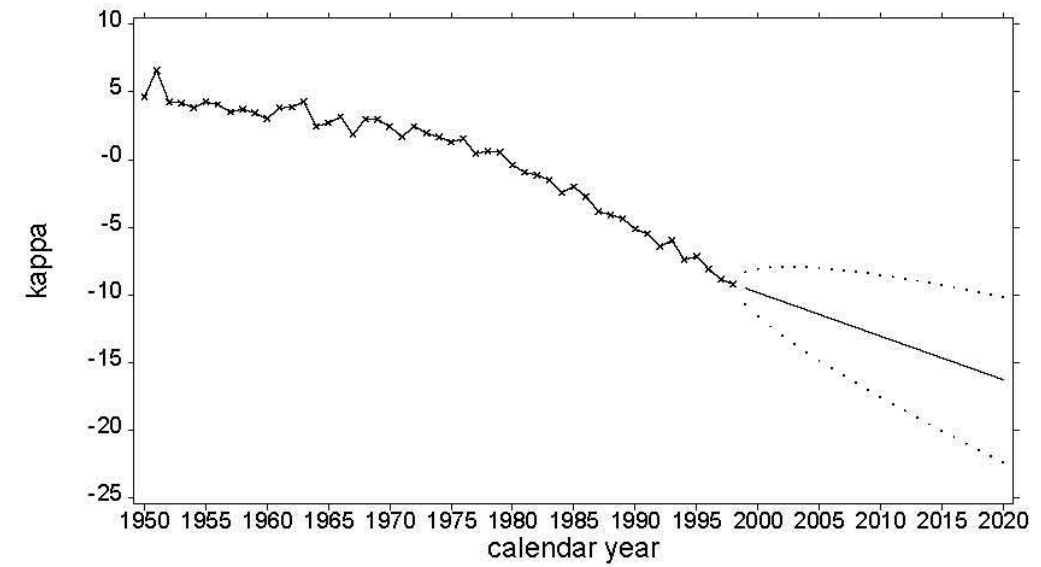
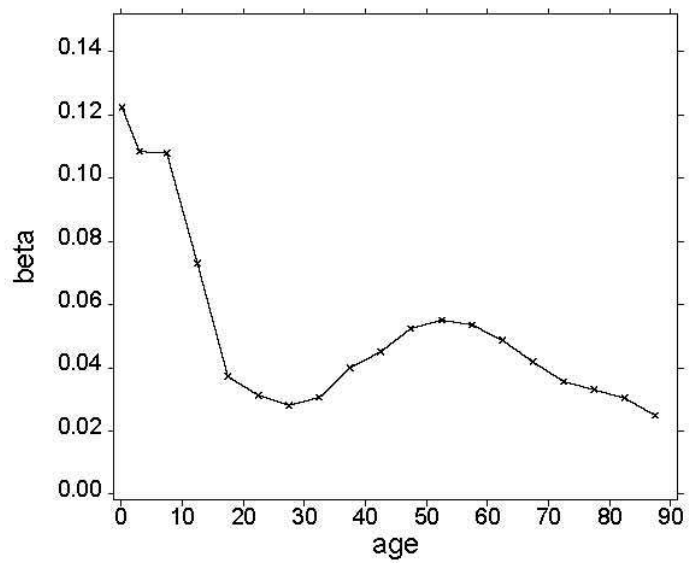
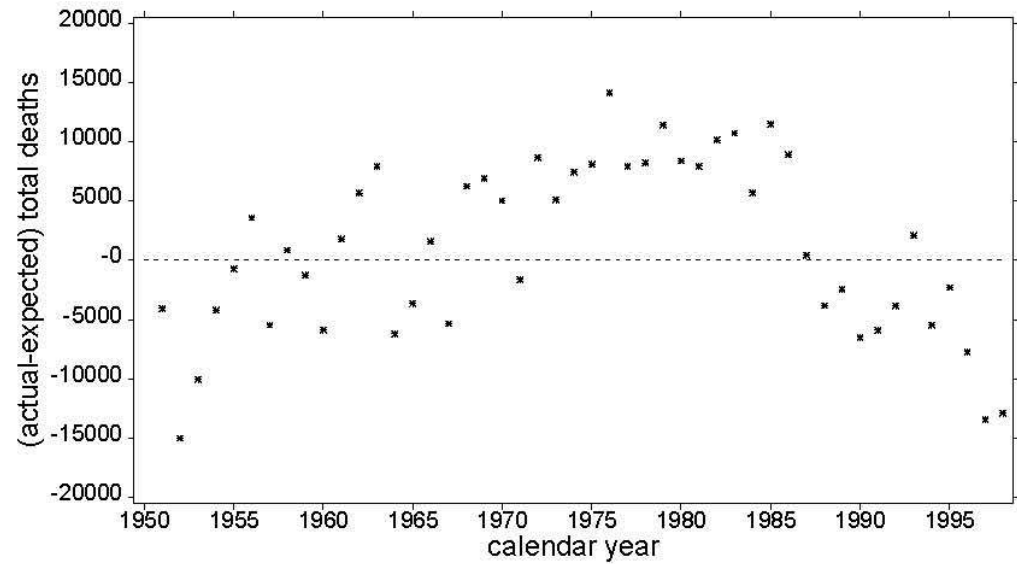
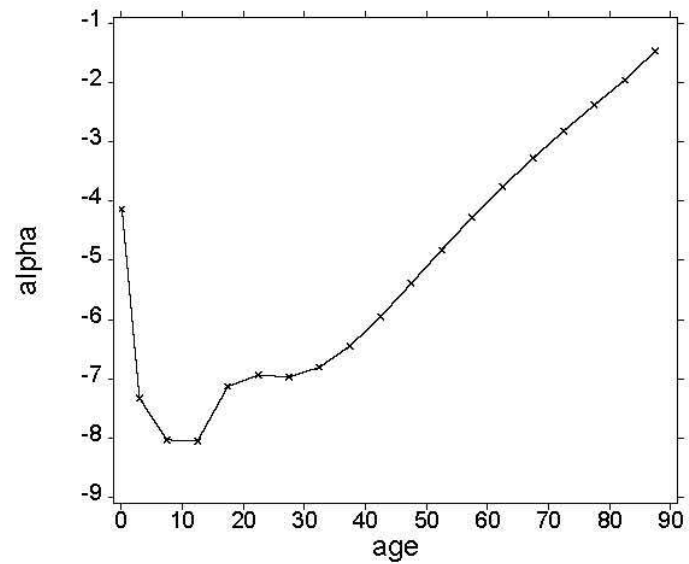


E+W 1961 – 2003 male mortality experience. Log crude mortality rates against period, for grouped age (0-,1-,5-,10-, ..., 90-, 95-).

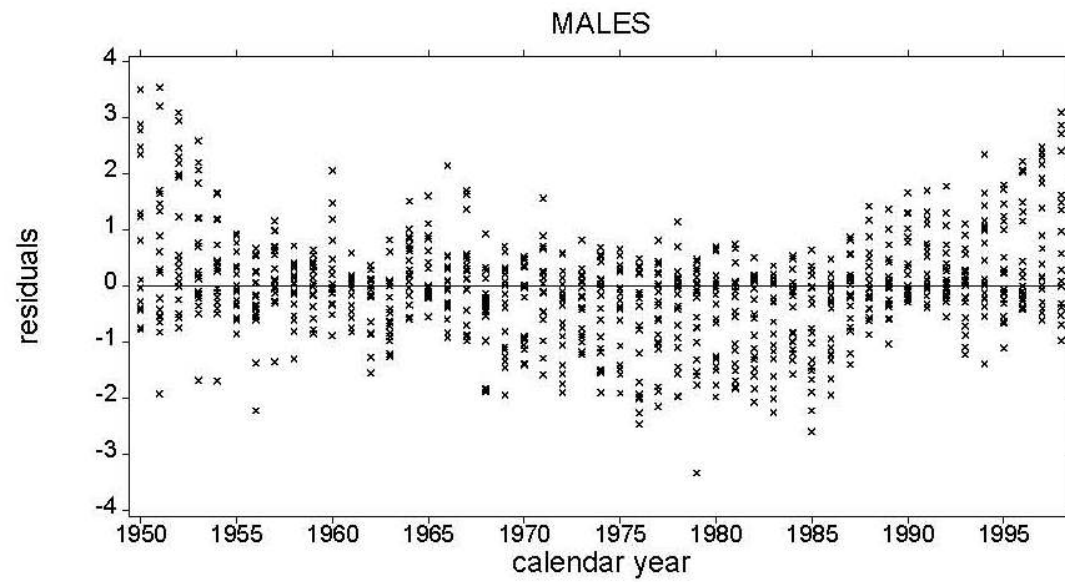
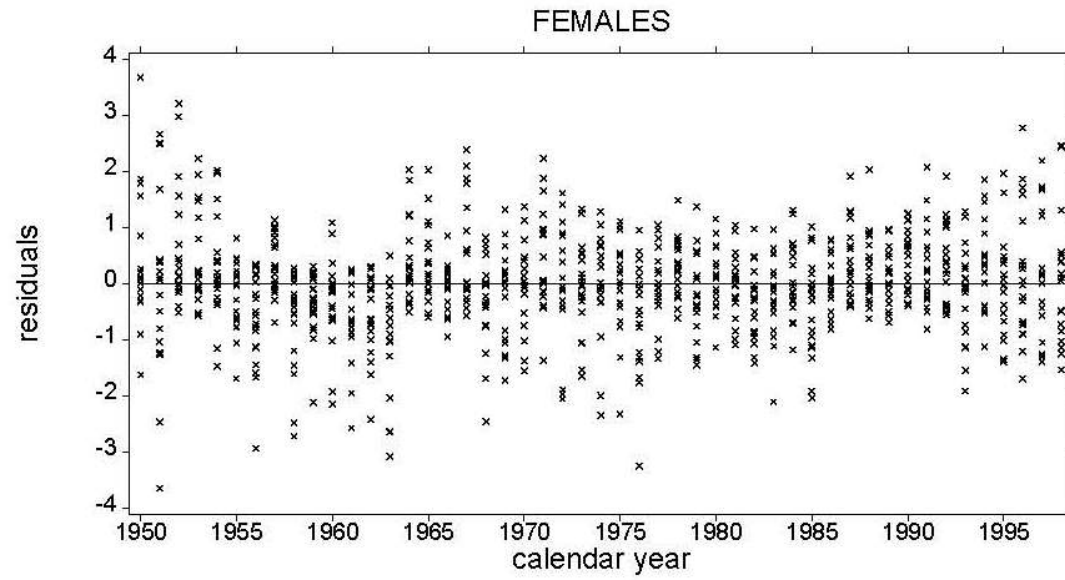
E+W female mortality experience (LC)



E+W male mortality experience (LC)



LC model: E+W mortality experience



INTERIM CONCLUSIONS

- Basic LC model fits England and Wales females mortality experience fairly well; but poor fit for male experience
- Enhancements
 - optimize choice of fitting period
 - add a second factor: $\alpha_x + \beta_x^{(1)} \kappa_t^{(1)} + \beta_x^{(2)} \kappa_t^{(2)}$
 - use a principal components approach
 - allow for a cohort effect

AGE-PERIOD-COHORT MODELS IN LC FRAMEWORK

UK: Strong cohort effect for those born in period 1925 – 1945 (also US, Japan, Germany Sweden)

The Lee-Carter structure may be expanded to incorporate a cohort effect:

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

under the Poisson setting.

The age-cohort substructure

$$\text{AC: } \beta_x^{(1)} = 0$$

is also of interest, while we recall that for the standard model

$$\text{LC: } \beta_x^{(0)} = 0.$$

FITTING AGE-PERIOD-COHORT MODELS

Fitting is problematic because of the relationship
cohort = (period – age) or $z = t - x$
between the three main effects.

We resort to a two-stage fitting strategy, in which α_x is estimated first, according to the original Lee-Carter SVD approach, thus

$$\hat{\alpha}_x = \frac{1}{t_n - t_1 + 1} \sum_{t=t_1}^{t_n} \ln \hat{m}_{xt}$$

The remaining parameters can then be estimated subject to the parameter constraints

$$\sum_x \beta_x^{(0)} = 1, \sum_x \beta_x^{(1)} = 1 \text{ and } \iota_{t_1-x_k} = 0 \text{ or } \kappa_{t_1} = 0.$$

Effective starting values are obtained by setting $\beta_x^{(0)} = \beta_x^{(1)} = 1$ and fitting a restricted version of the model to generate starting values for ι_z and κ_t .

PROJECTIONS

Use separate ARIMA time series

$$\{\hat{\kappa}_t : t \in [t_1, t_n]\} \mapsto \{\dot{\kappa}_{t_n+s} : s > 0\},$$

$$\{\hat{i}_z : z \in [t_1 - x_k, t_n - x_1]\} \mapsto \{i_{t_n-x_1+s} : s > 0\},$$

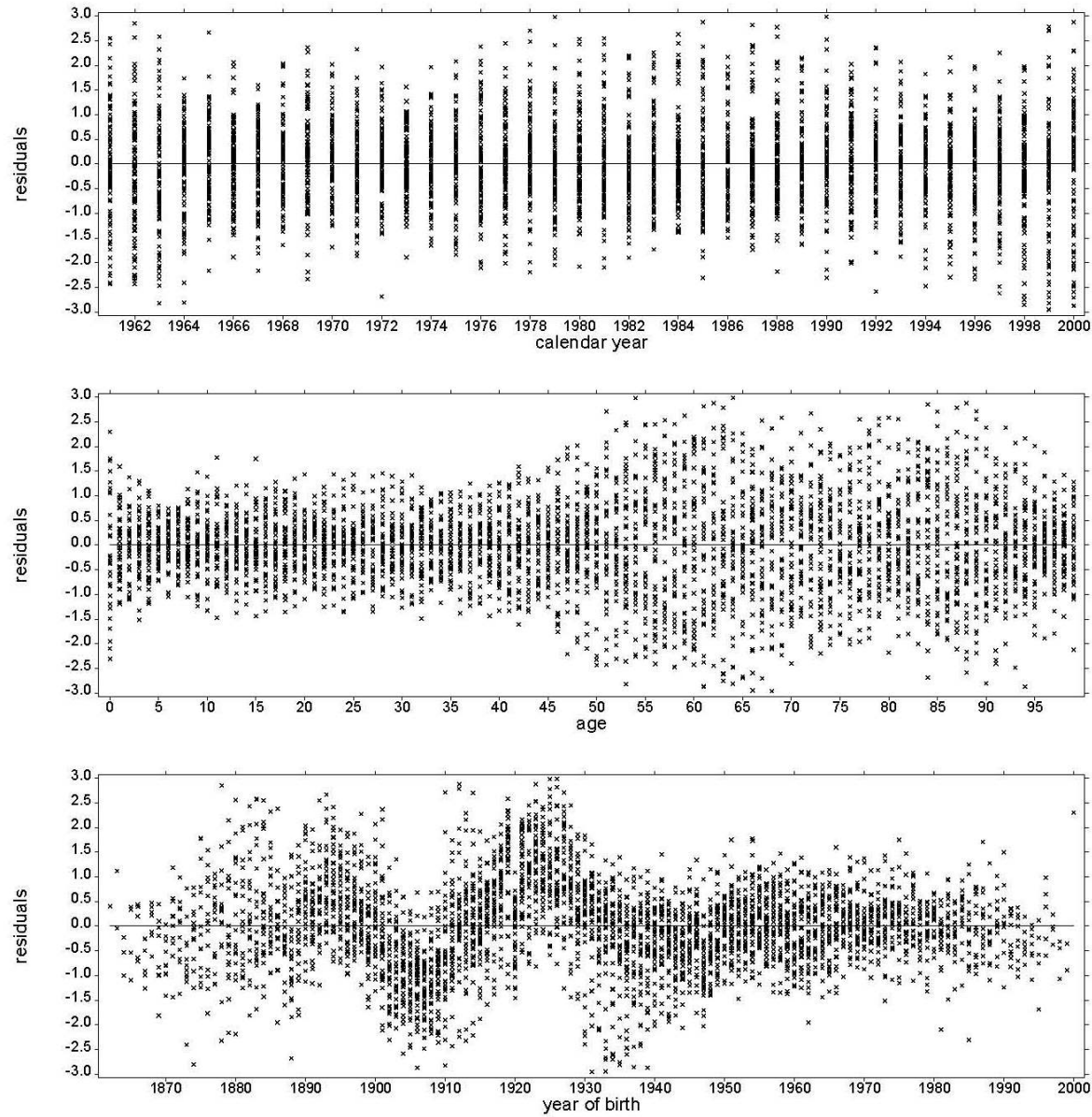
then

$$F(x, t_n + s) = \exp\{\hat{\beta}_x^{(0)}(\tilde{i}_{t_n-x+s} - \hat{i}_{t_n-x}) + \hat{\beta}_x^{(1)}(\dot{\kappa}_{t_n+s} - \hat{\kappa}_{t_n})\}, \quad s > 0$$

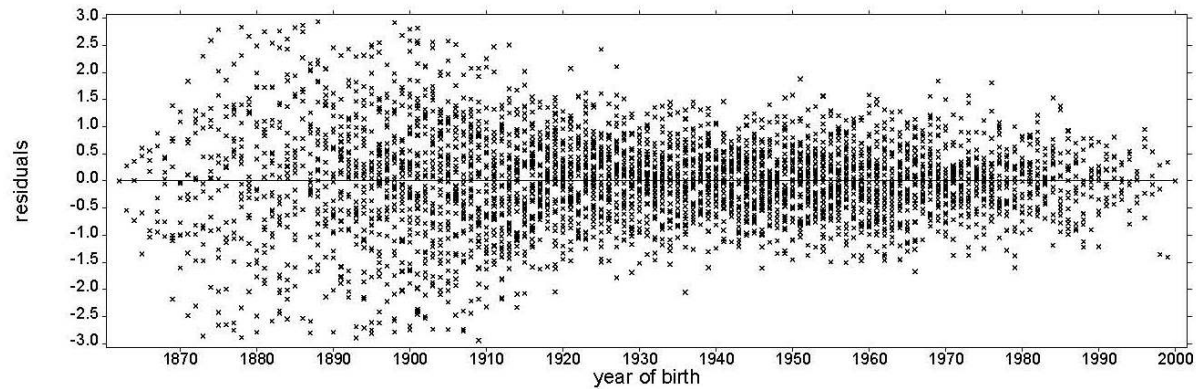
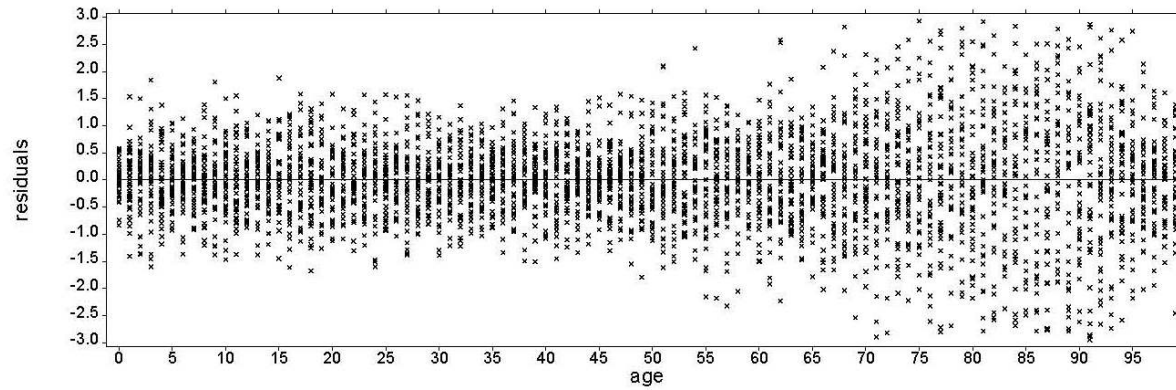
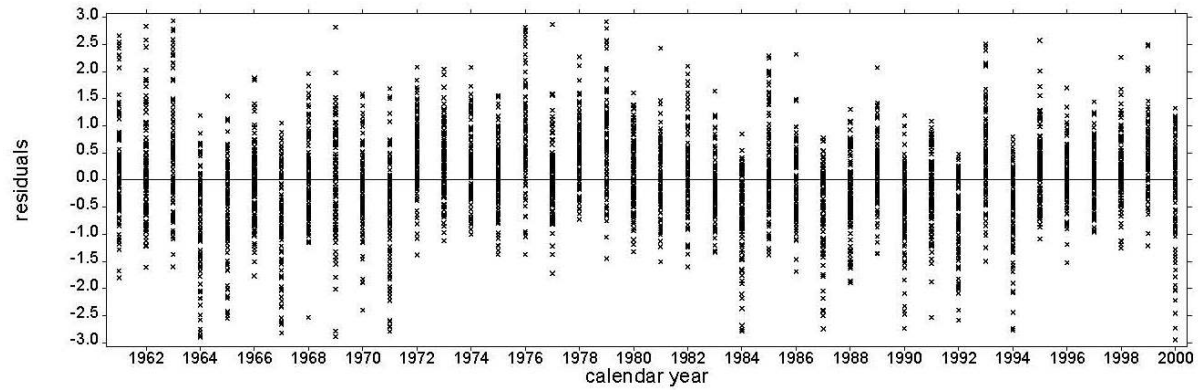
where

$$\tilde{i}_{t_n-x+s} = \begin{cases} \hat{i}_{t_n-x+s}, & s \leq x-x_1 \\ i_{t_n-x+s}, & s > x-x_1 \end{cases}.$$

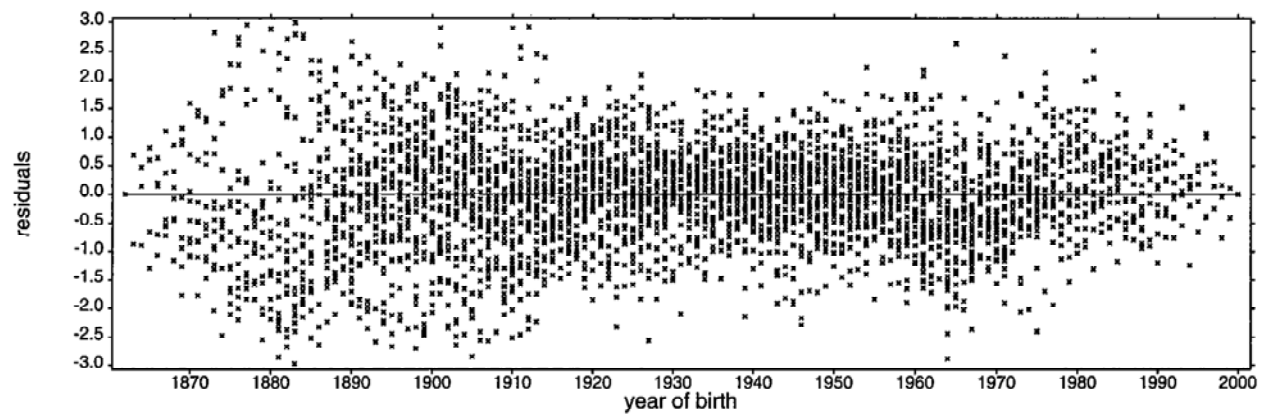
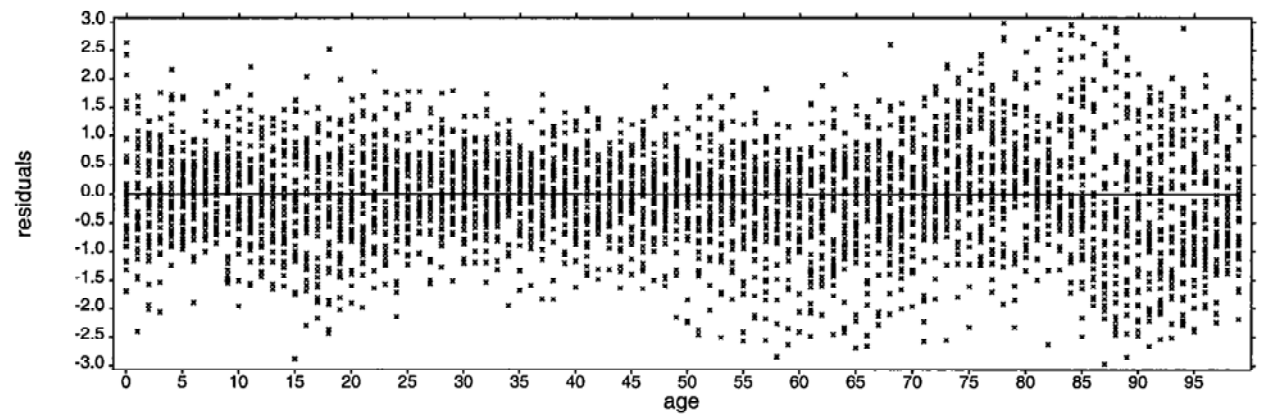
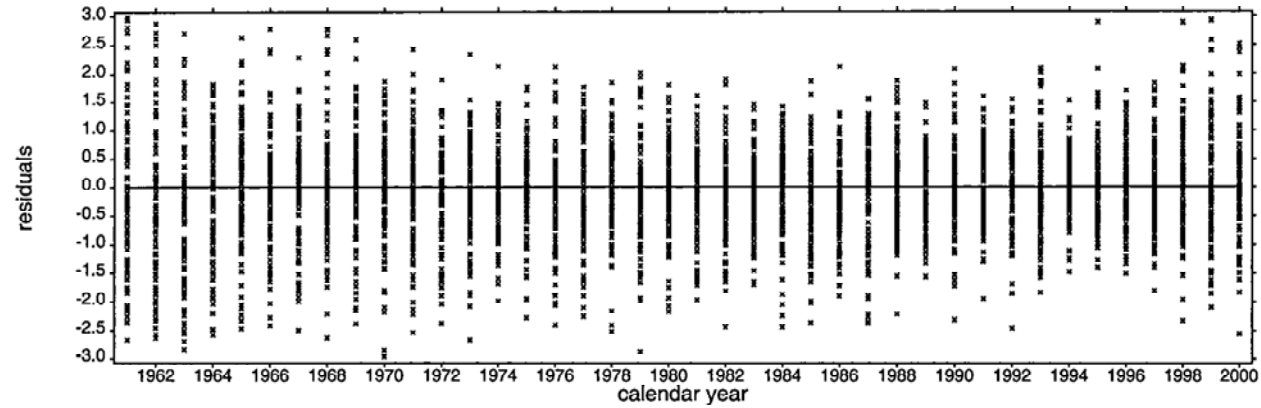
UK female mortality experience (LC) – residual plots



UK female mortality experience (AC) – residual plots



UK female mortality experience (APC) – residual plots



APPLICATIONS

Time Series Forecasts

- For κ_t :

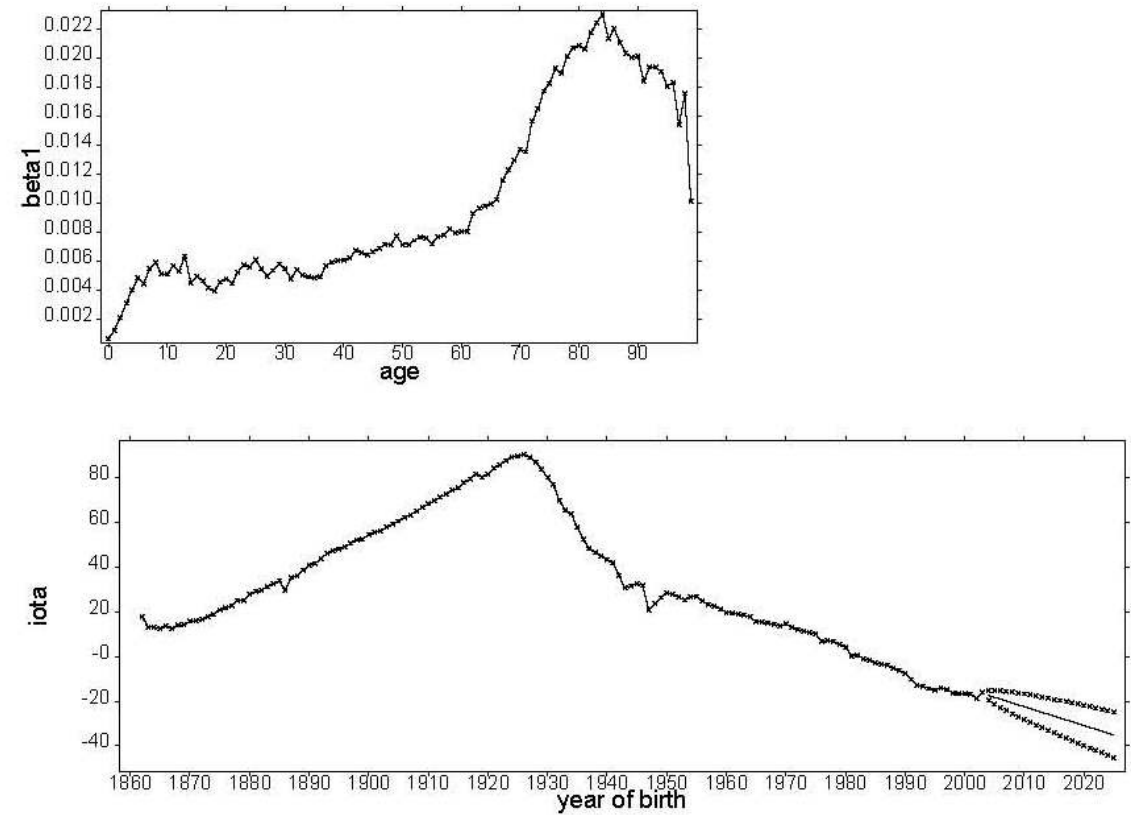
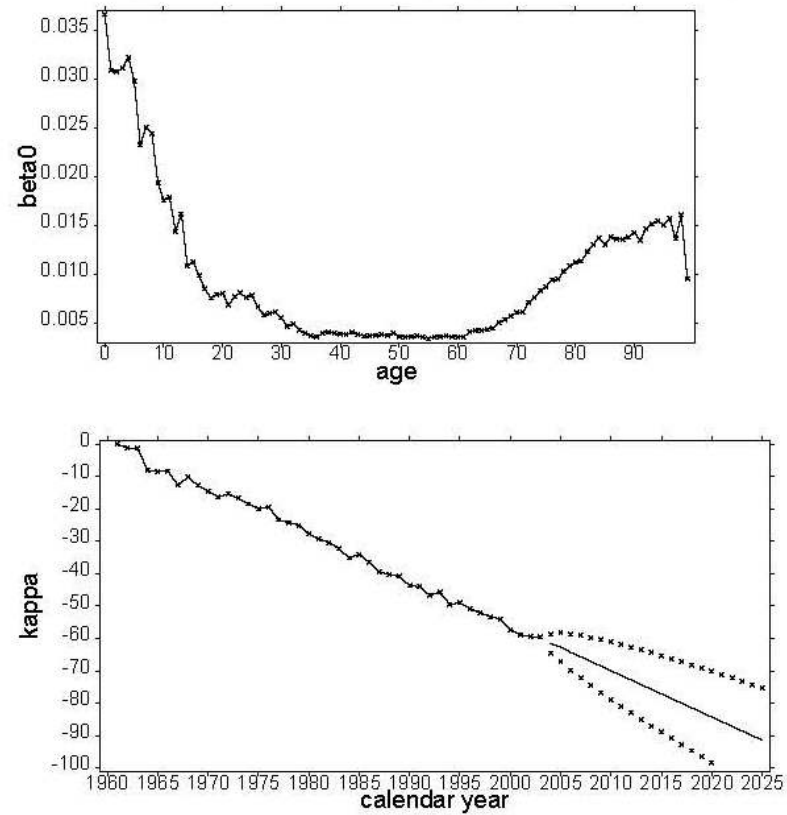
ARIMA (1, 1, 0) for females

ARIMA (2, 1, 0) for males

- For ι_z :

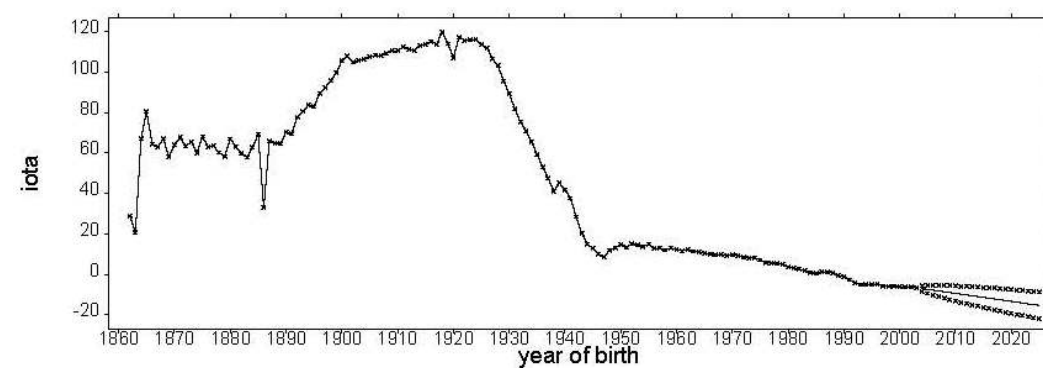
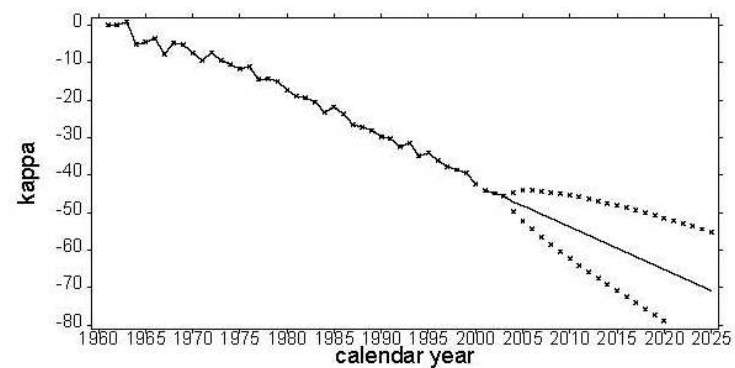
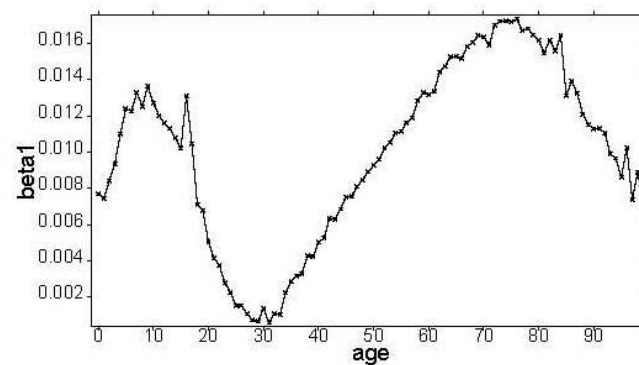
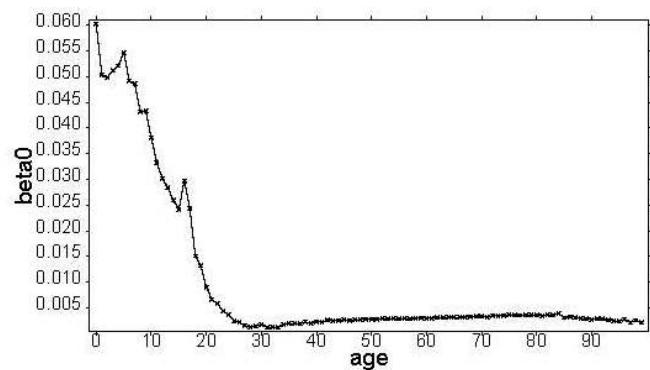
ARIMA (1,1,0) for both genders

(a) E+W female population study



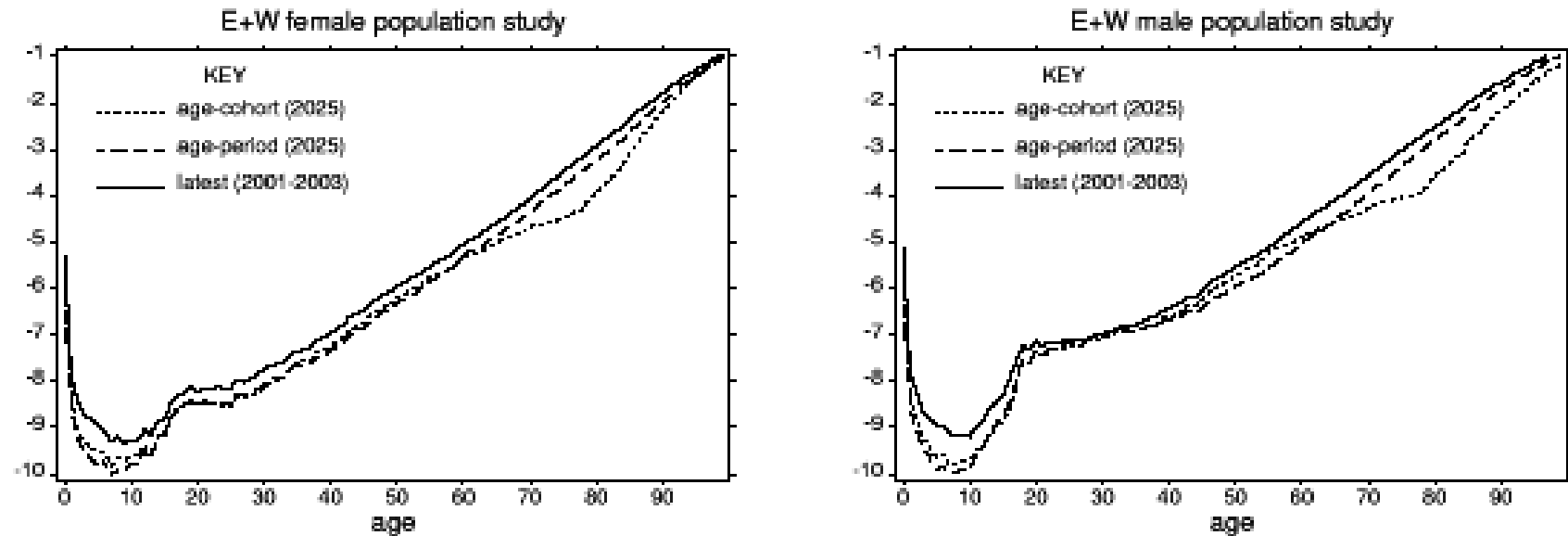
England and Wales population, parameter estimates, APC model:
(a) females; (b) males

(b) E+W male population study

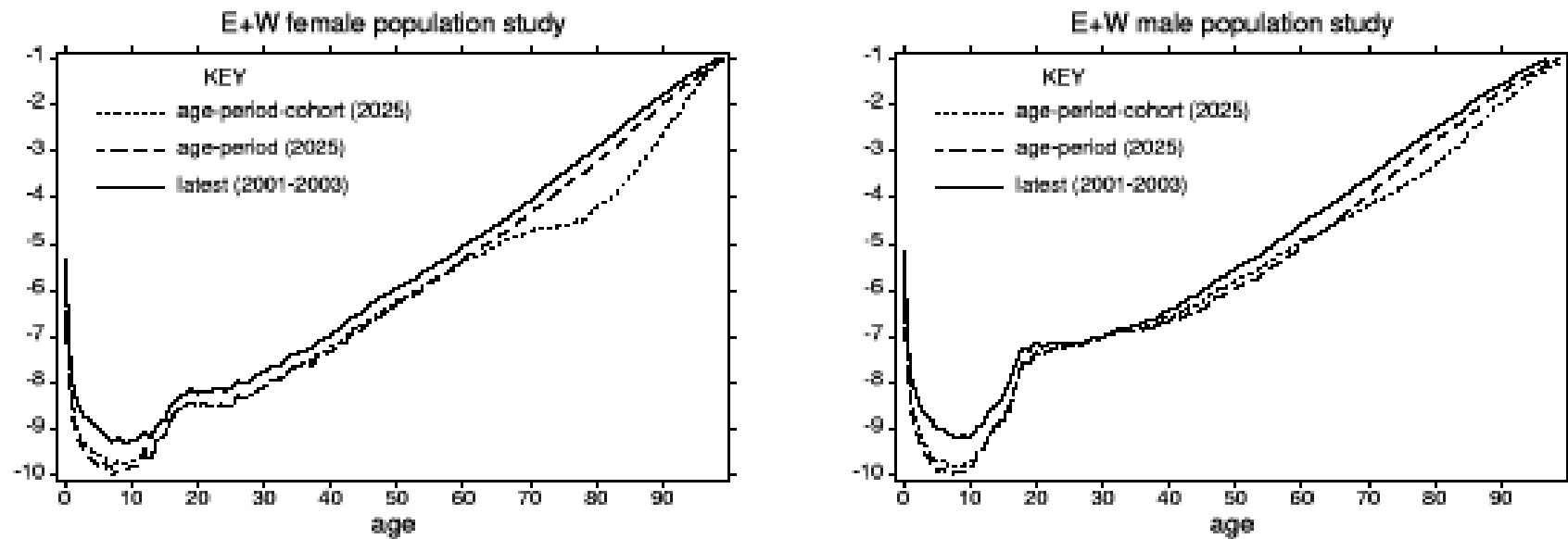


**England and Wales population, parameter estimates, APC model:
(b) females; (b) males**

(a) log(mortality rates): projections by age-period & age-cohort

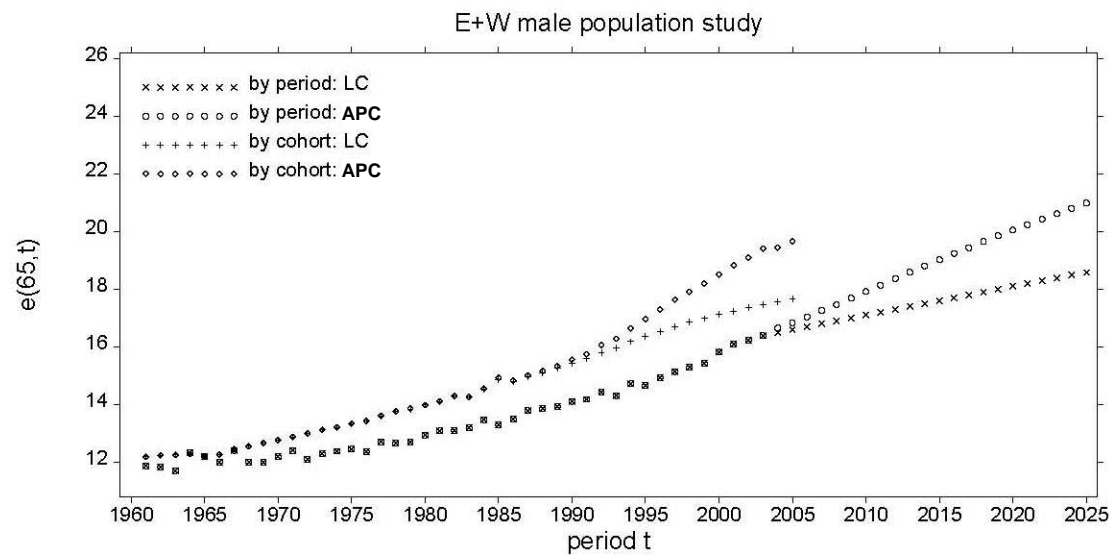
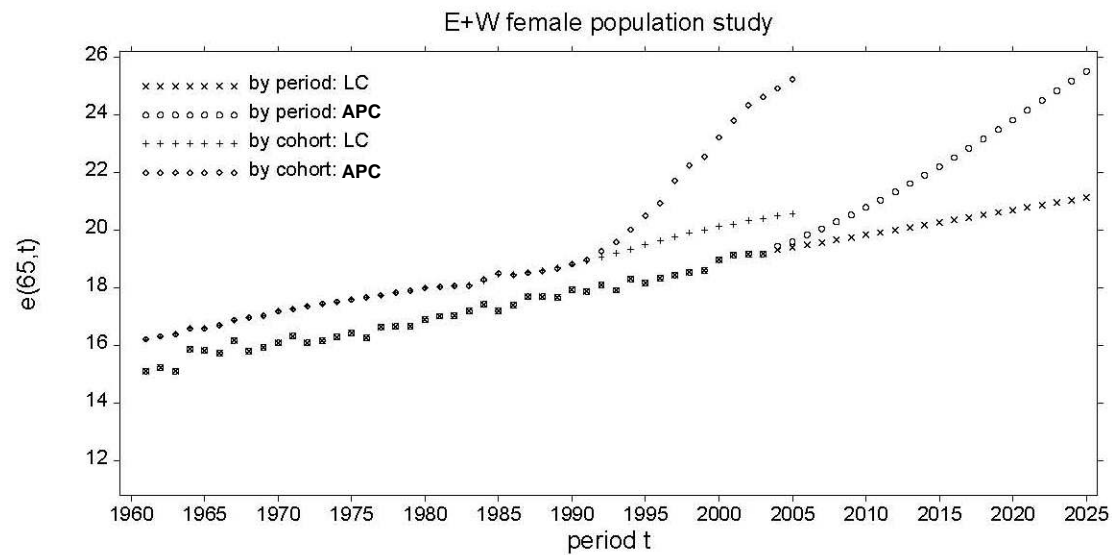


(b) log(mortality rates): projections by age-period & age-period-cohort



Latest and projected $\log \mu_{xt}$ age profiles:

(a) LC and AC modelling; (b) LC and APC modelling



Life expectations at age 65 for a range of periods, computed by period and by cohort under age-period (LC) and age-period-cohort (APC) modelling

RISK MEASUREMENT AND PREDICTION INTERVALS

- uncertainty in projections needs to be quantified i.e. by prediction intervals
- but analytical derivations are difficult
- 2 different sources of uncertainty need to be combined
 - errors in estimation of parameters of Lee Carter model
 - forecast errors in projected ARIMA model
- indices of interest (e.g. hazard rates, annuity values, life expectancies) are complex non linear functions of $\alpha_x, \beta_\chi, \kappa_t$ and ARIMA parameters.

DIFFERENT SIMULATION STRATEGIES

A) Semi-parametric (Poisson) Bootstrap: generates new data sets

Let \hat{d}_x be fitted number of deaths.

Simulate response $d_x^{(j)}$ from Poisson (\hat{d}_x)

Compute $\mu_x^{(j)}$

Fit model: obtain estimates of $\alpha_x^{(j)}, \beta_x^{(j)}, \kappa_t^{(j)}$

Compute $\dot{\kappa}_{t_n+k}^{(j)} = E[\kappa_{t_n+k}^{(j)}] + \sqrt{\text{Var}[\kappa_{t_n+k}^{(j)}]} \cdot z_j^*$

Repeat for $j = 1, \dots, N$

DIFFERENT SIMULATION STRATEGIES (Continued)

C) Residuals Bootstrap: generates new data sets

Let r_x be the deviance residuals

Sample with replacement to get $r_x^{(j)}$

Map from $r_x^{(j)}$ to $d_x^{(j)}$ for each x

Compute $\mu_x^{(j)}$

Fit Model: obtain estimates of $\alpha_x^{(j)}, \beta_x^{(j)}, \kappa_t^{(j)}$

Compute $\dot{\kappa}_{t_n+k}^{(j)} = E\left[\kappa_{t_n+k}^{(j)}\right] + \sqrt{\text{Var}\left[\kappa_{t_n+k}^{(j)}\right]} \cdot z_j^*$

Repeat for $j = 1, \dots, N$

DIFFERENT SIMULATION STRATEGIES (Continued)

- B) Parametric Monte Carlo Simulation: generates new parameter estimates from fitted parameter estimates

Simulate $e^{(j)}$ vector of $N(0,1)$ errors

Let C be the Cholesky factorisation matrix of the variance- covariance matrix (needs to be invertible)

Compute simulated model parameters

$$\theta^{(j)} = \hat{\theta} + \sqrt{\varphi} C e^{(j)}$$

where φ is optional scale parameter

Compute $\dot{K}_{t_n+k}^{(j)}$

Repeat for $j = 1, \dots, N$

JOINT MODELLING

Attempt to model variable dispersion parameter (rather than fixed ϕ)

2 stage process (LG model)

1. Model D_{xt} as independent Poisson response

Define $R_{xt} = \omega_{xt} \frac{\{D_{xt} - E(D_{xt})\}^2}{E(D_{xt})}$ the resulting squared Pearson residuals

2. Define R_{xt} as independent gamma responses

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

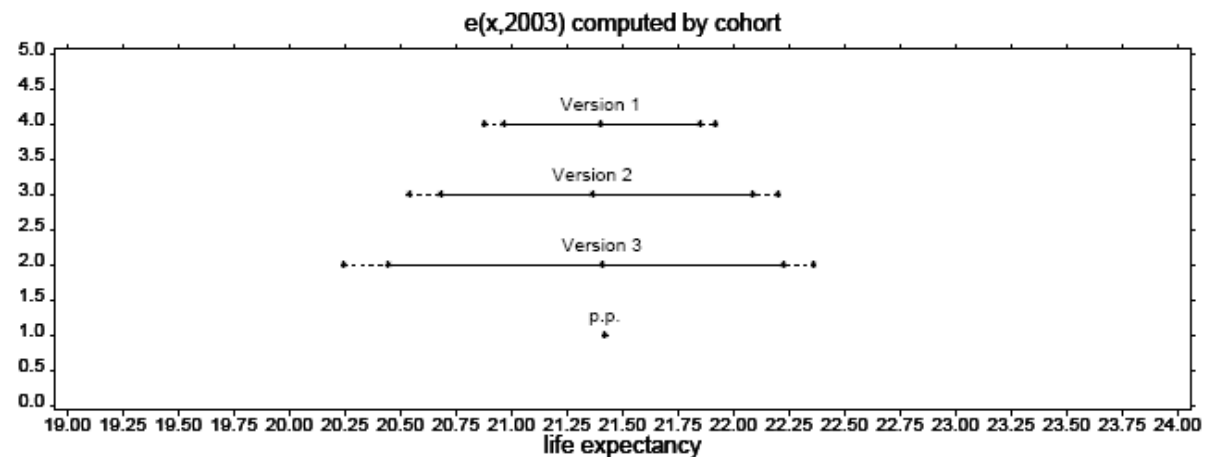
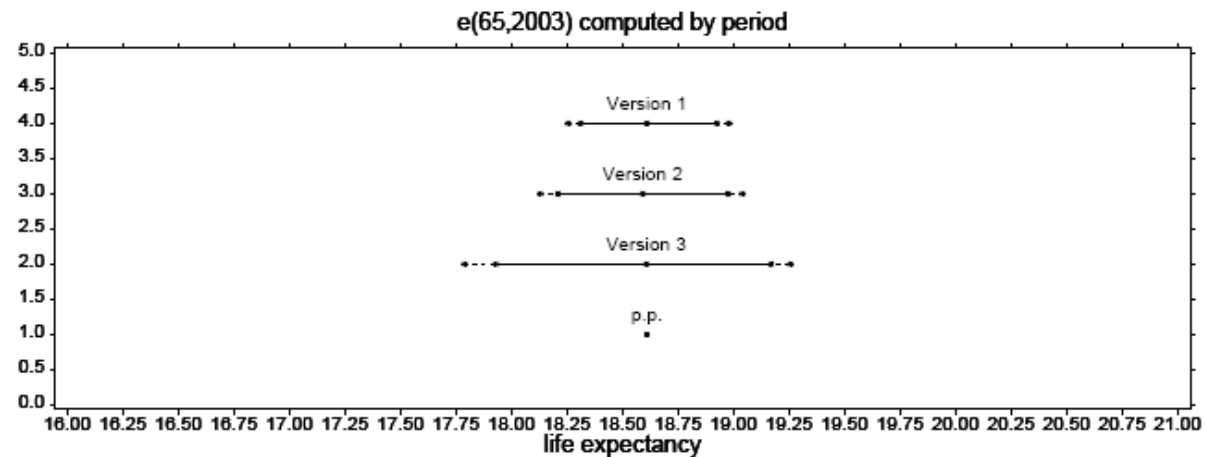
and log link and linear parametric structure in age.

NEGATIVE BINOMIAL MODELLING

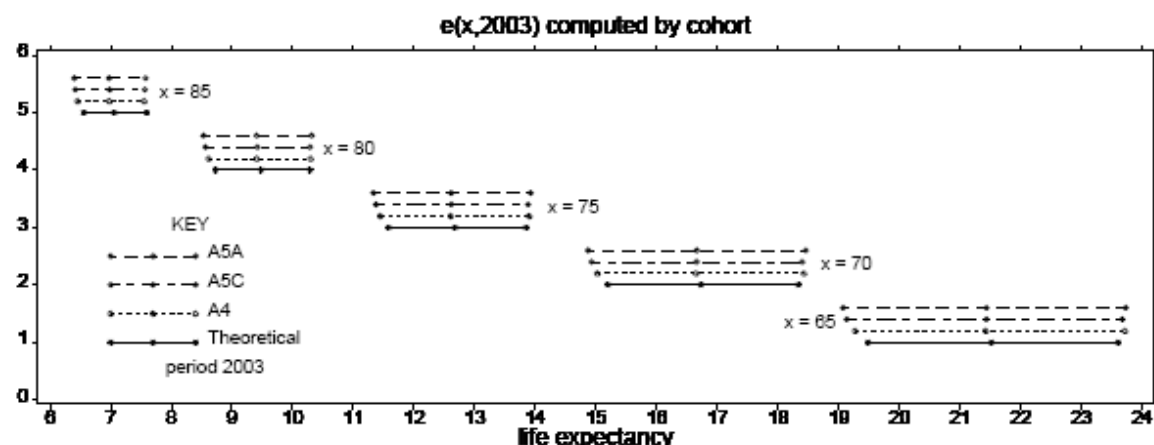
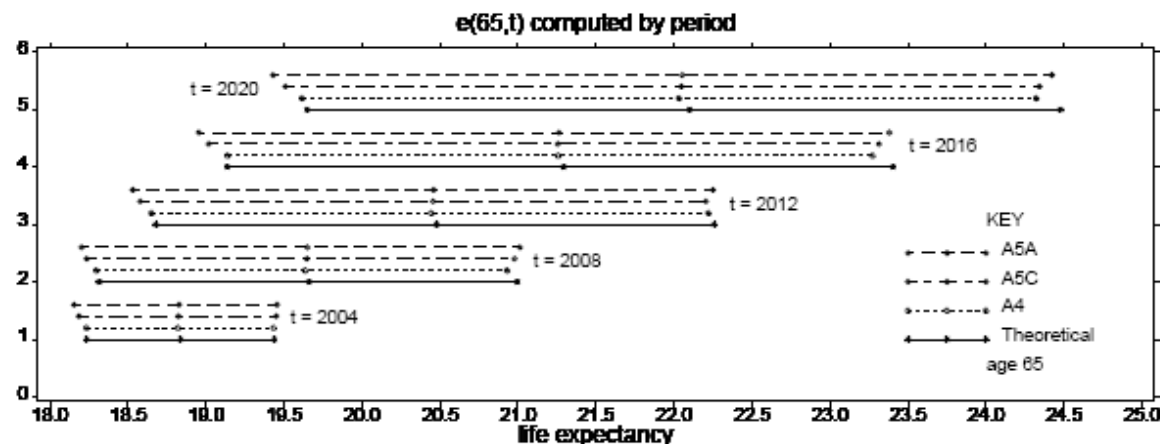
Extend Poisson model (with no scale parameter $\phi = 1$)

Variance function in GLM becomes:

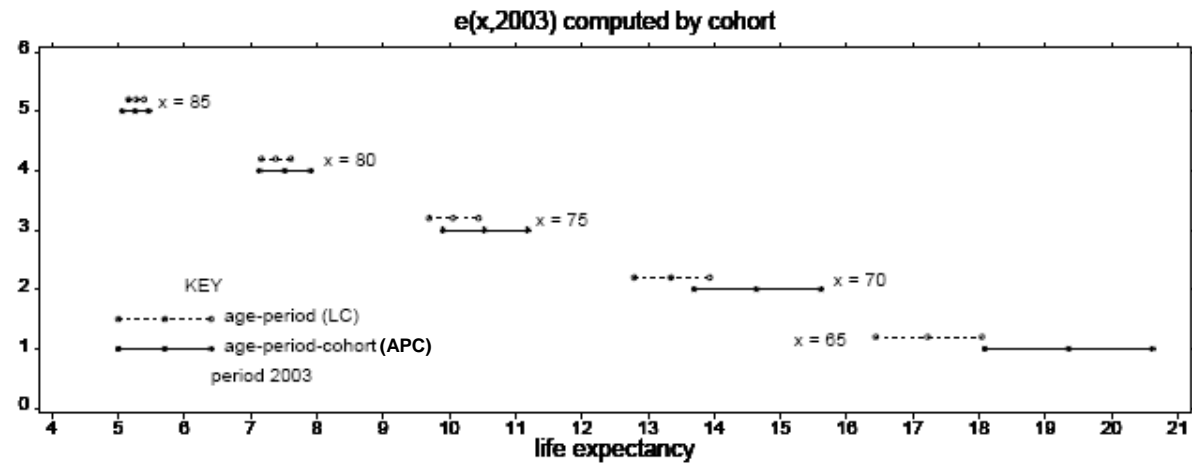
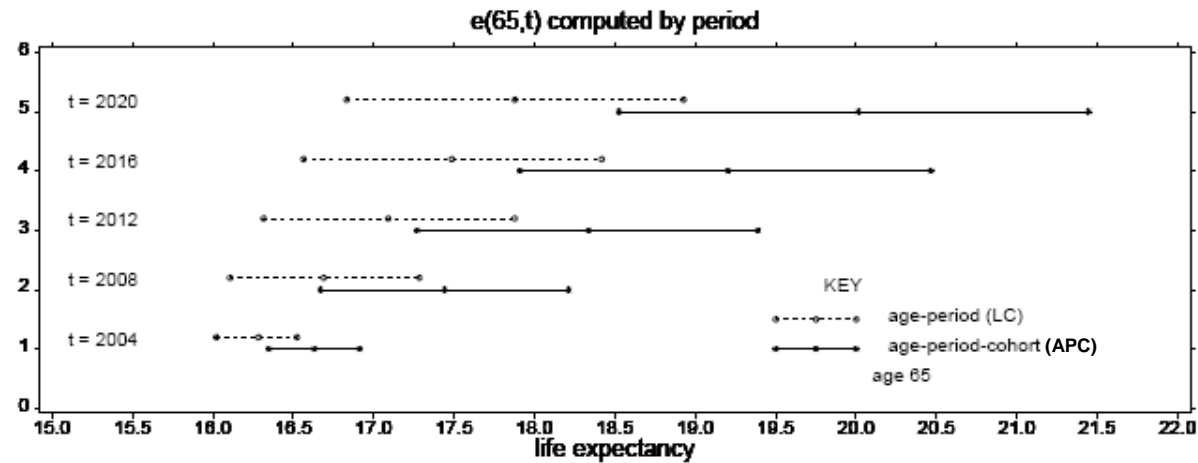
$$V(u) = u + \lambda_x u^2$$



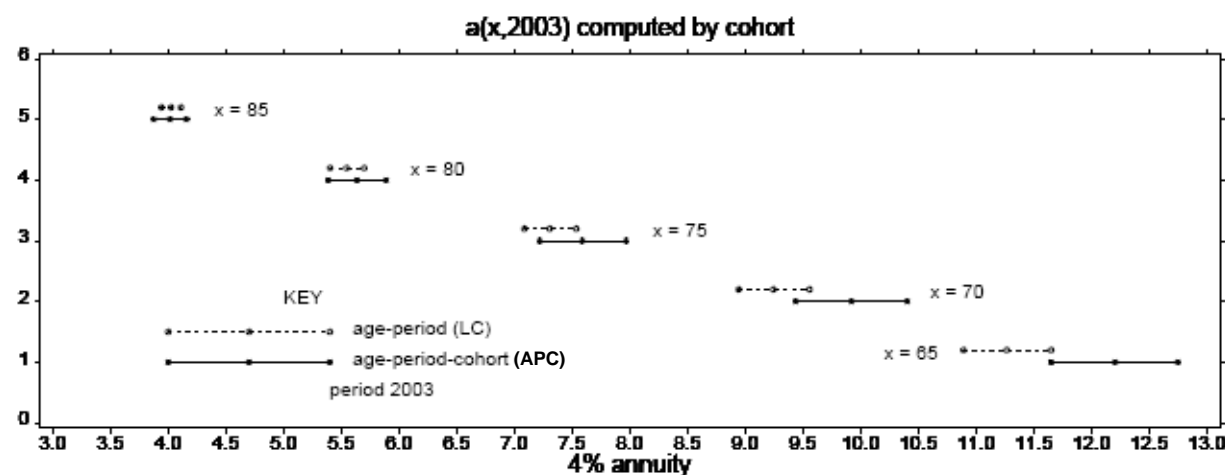
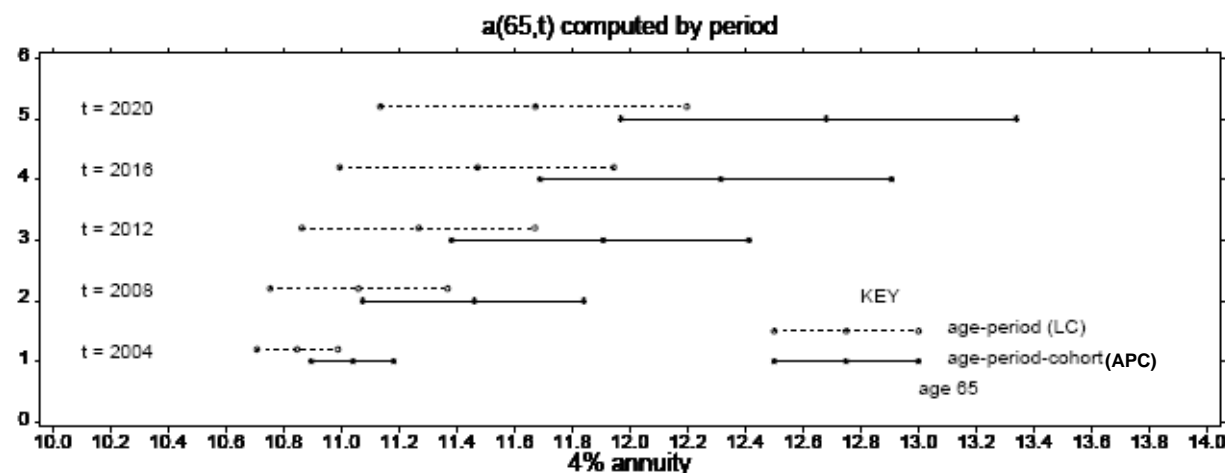
UK 1983-2003 male pensioner mortality experience (ages 51-104).
 Poisson LC: age-period (log-bilinear) fitted parametric structure, with
 and without the inclusion of a free-standing (constant) scale parameter.
 Life expectancy: comparison of 2.5, 50, 97.5 percentile based PIs using
 different versions of simulation Algorithm B, with point predictor (p.p.).



UK 1983-2003 male pensioner mortality experience (ages 51-104).
Poisson LC: age-period (log- bilinear) structure, with period random walk.
Comparison of life expectancy predictions (various age-period start points)
based on 2.5, 50, 97.5 percentiles: (i) By theory. (ii) By bootstrapping
the prediction error in the period component time series (Algorithm A4).
(iii) By bootstrapping the time series prediction error and including
model fitting simulated error (Algorithms A5A or A5C).



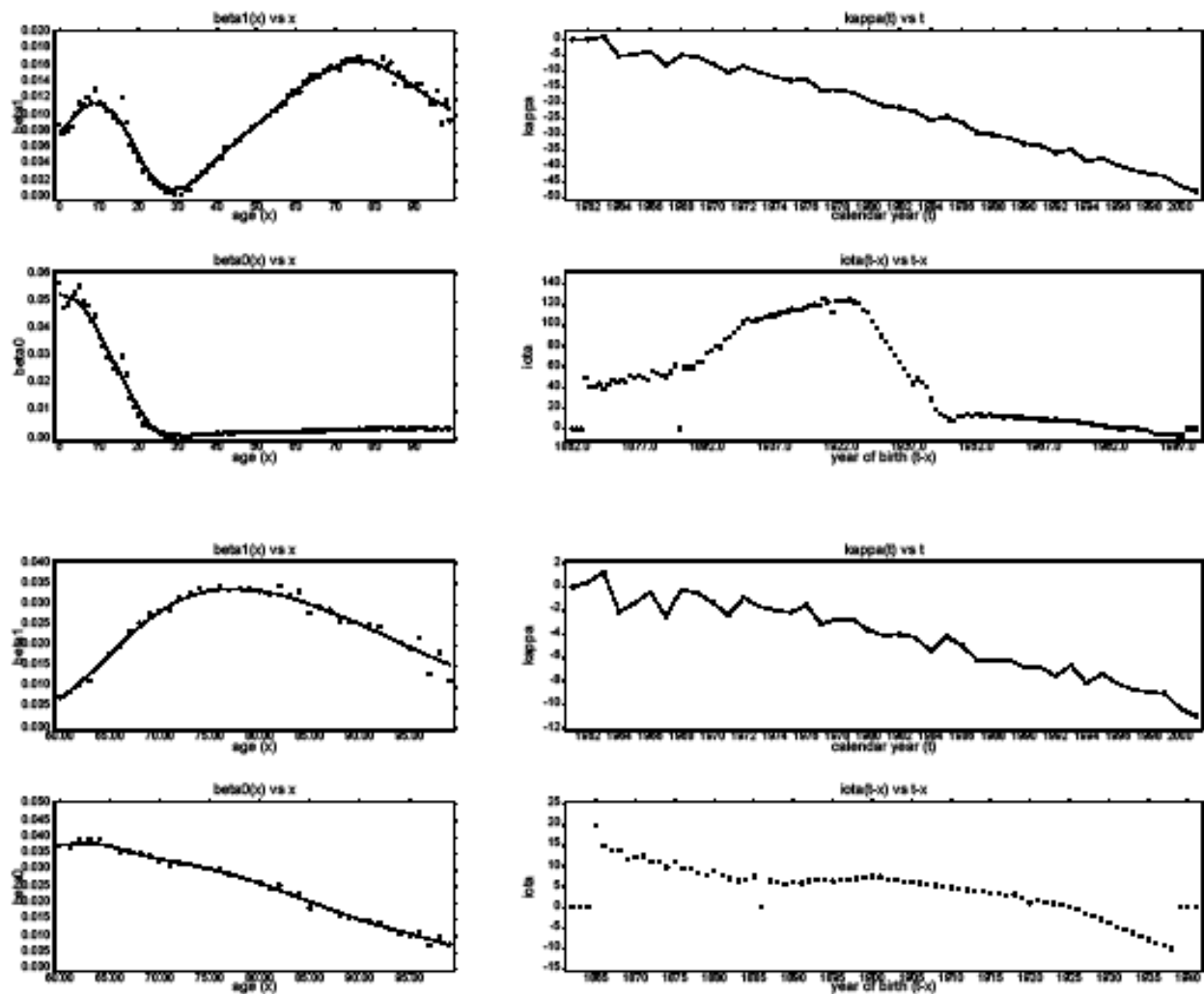
E+W male mortality: comparison life expectancy predictions using (i) age-period-cohort and (ii) age-period Poisson structures. Predictions with intervals by bootstrapping the time series prediction error in the period (and cohort) components, and selecting the resulting 2.5, 50, 97.5 percentiles.



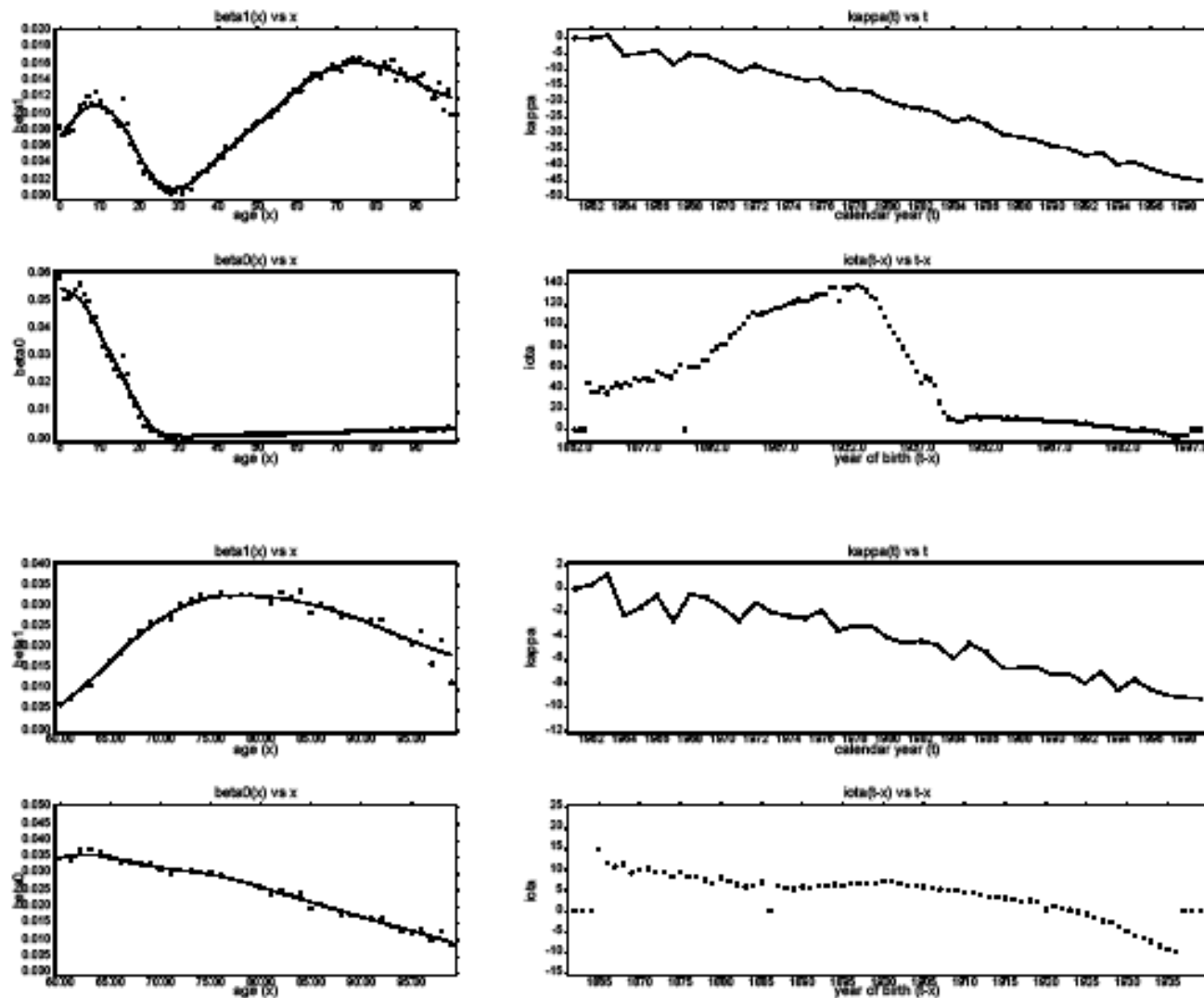
E+W male mortality: comparison 4% fixed rate annuity predictions using (i) age-period-cohort and (ii) age-period Poisson structures. Predictions with intervals by bootstrapping the time series prediction error in the period (and cohort) components, and selecting the resulting 2.5, 50, 97.5 percentiles.

BACK-FITTING EXPERIMENTS

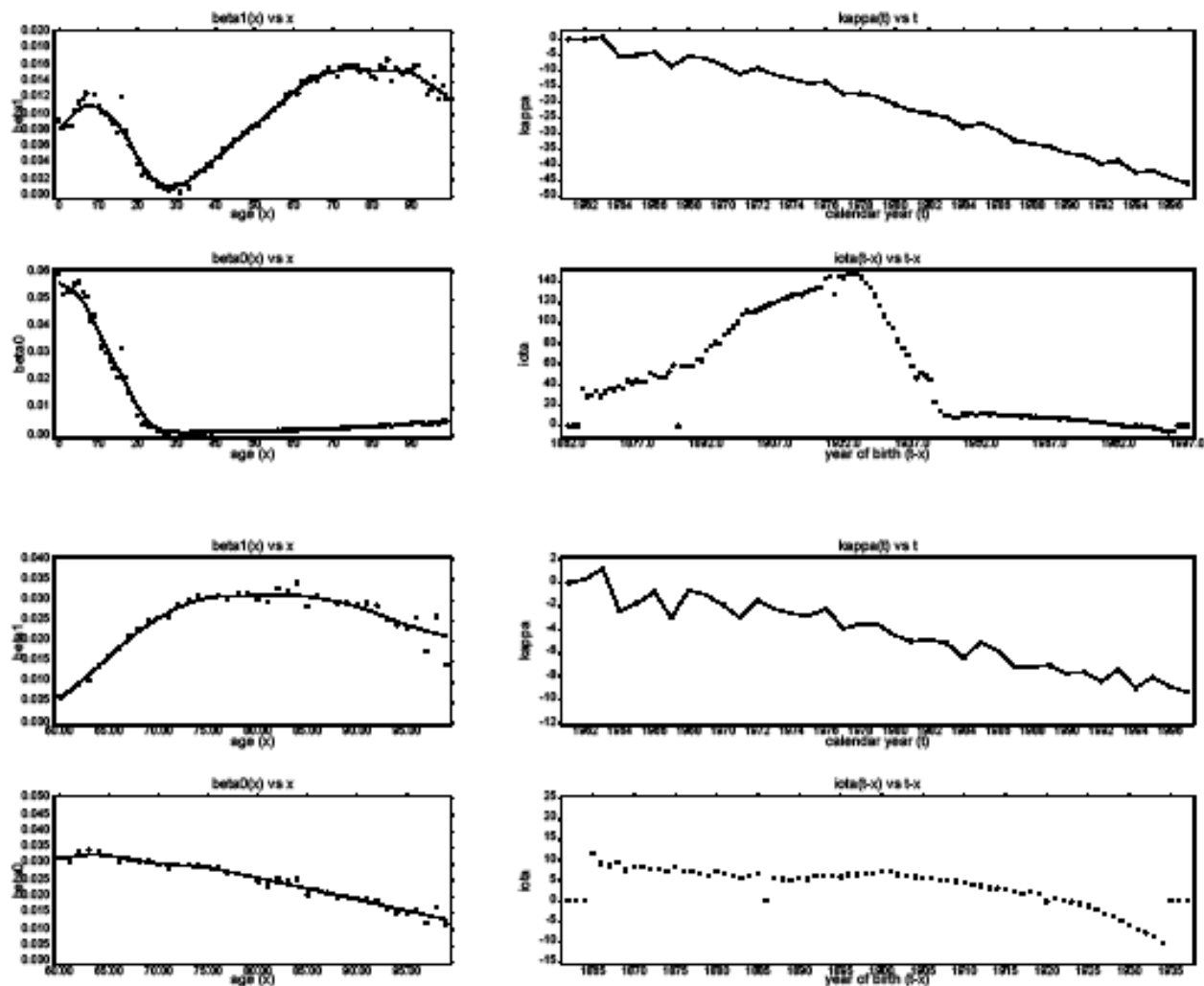
Test stability of model and parameter estimates



E&W male mortality experience. APC Model: mortality reduction factor parameter estimates: period 1961-2001; (i) ages 0-99, (ii) ages 60-99. Left frame beta parameter smoothing using S-Plus 2000 "Supersmooth".



E&W male mortality experience. APC Model: mortality reduction factor parameter estimates: period 1961-1999; (i) ages 0-99, (ii) ages 60-99. Left frame beta parameter smoothing using S-Plus 2000 "Supersmooth".



E&W male mortality experience. APC Model: mortality reduction factor parameter estimates: period 1961-1997; (i) ages 0-99, (ii) ages 60-99. Left frame beta parameter smoothing using S-Plus 2000 "Supersmooth".

FINAL COMMENT

- Other approaches to prediction intervals for Lee-Carter models – Bayesian methods
- Extreme ages – extrapolation methods needed where data are scarce
- Problems with forecasting structural changes
- Time series methods and their application to long forecasting periods

FINAL COMMENT (continued)

- Effect of β_x on smoothness of projected age profiles: need for smoothing of estimates
- Quality of data sources and appropriateness for particular applications: adverse selection and “basis risk”.
- Model error – essential to investigate more than one modelling framework.
- Sources of uncertainty – process, parameter, model, judgement. Not all sources of uncertainty can be quantified.

REFERENCES

- Booth, H, (2006) Demographic Forecasting: 1980 to 2005 in review. *International Journal of Forecasting*, **22**, 547-581.
- Booth, H, Maindonald, J. and Smith, S. (2002) Applying Lee-Carter under conditions of Variable mortality decline. *Population Studies*, **56**, 325-336.
- Brouhns, N., Denuit, M. and Vermunt, J.K. (2002a) A Poisson log-bilinear regression approach to the construction of projected life-tables. *Insurance: Mathematics and Economics* **31**, 373-393.
- Brouhns, N., Denuit, M. and Vermunt, J.K. (2002b) Measuring the longevity risk in mortality projections. *Bulletin of the Swiss Association of Actuaries*, 105-130.
- Brouhns, N., Denuit, M. and van Keilegom, I. (2005). Bootstrapping the Poisson log-bilinear model for mortality forecasting. *Scandinavian Actuarial Journal*, **3**, 212 – 224.
- Cossette, H., Delwarde, D., Denuit, M., Guillot, F. and Marceau, E. (2007). Pension plan valuation and mortality projection : a case study with mortality data. *North American Actuarial Journal*, 11(2), 1 – 34.
- Davison, A. and Hinkley, D. (2006). Bootstrap methods and their applications. Cambridge University Press.
- Delwarde, A., Denuit, M. and Partrat, C. (2007). Negative binomial version of the Lee-Carter model for mortality forecasting. *Applied Stochastic Models in Business and Industry* **23**, 385 – 401.
- Denuit, M. (2007). Distribution of the random future life expectancies in log-bilinear mortality projection models. *Lifetime Data Analysis* **13**, 381-397.

REFERENCES (continued)

Denuit, M., Haberman, S. and Renshaw, A. (2009). A note on the distribution of random future life expectancies in Poisson log-bilinear mortality projections. In preparation.

Haberman, S. and Renshaw, A. (2008). On simulation-based approaches to risk measurement in mortality with specific reference to binomial Lee-Carter modelling. Presented to Society of Actuaries' Living to 100 Symposium. Orlando, Florida, USA.

Haberman, S and Renshaw, A. (2008). Mortality, longevity and experiments with the Lee-Carter model. *Lifetime Data Analysis*, **14**, 286-315

Haberman, S and Renshaw, A. (2008). Measurement of longevity risk using bootstrapping for Lee-Carter and generalised linear Poisson models of mortality. *Methodology and Computing in Applied Probability*, To Appear.

Hatzopoulos, P. and Haberman, S. (2009) A parameterized approach to modelling and forecasting mortality. *Insurance: Mathematics and Economics*, To appear.

Koissi, M-C., Shapiro, A.F. and 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. (2000) The Lee-Carter method of forecasting mortality, with various extensions and applications (with discussion). *North American Actuarial Journal* **4 (1)**, 80-93.

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

Li, S-H, Hardy, M and Tan, K. (2006). Uncertainty in mortality forecasting: an extension to the classical Lee-Carter approach. University of Waterloo working paper.

REFERENCES (continued)

Pitacco, E. Denuit, M., Haberman, S and Olivieri A-M (2008). Modelling longevity dynamics for pensions and annuity business. Oxford University Press. In preparation.

Renshaw, A.E. (1992). Joint modeling for actuarial graduation and duplicate policies. *Journal of Institute of Actuaries* **119**, 69 – 85.

Renshaw, A.E. and 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. and Haberman, S. (2003b) On the forecasting of mortality reduction factors. *Insurance: Mathematics and Economics* **32**, 379-401.

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

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

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

Tuljapurkar, S., Li, N. and Boe, C. (2000) A universal pattern of mortality decline in the G7 countries. *Nature*, **405**, 789-792.

Wilmoth, J.R. (1993) Computational methods for fitting and extrapolating the Lee-Carter model of mortality change. *Technical Report. Dept. of Demography, University of California, Berkeley*.