MORTALITY, LONGEVITY AND EXPERIMENTS WITH THE LEE-CARTER MODEL

Professor Steven Haberman
Cass Business School, City University
AGENDA

- Mortality Trends
- Data and Notation
- Lee-Carter Model: Base
- Lee-Carter Model: Advanced
- Lee-Carter Model: Cohort Effects
- Lee-Carter Model: Risk Measurement
- Concluding Comments
MORTALITY TRENDS

Changes in the shape of mortality curves:

(i) downward trends in mortality rates – at young and old ages;

(ii) an increasing concentration of deaths around the average age at death. This is also described as the **rectangularization** of the survival function;

(iii) the average age at death increasing over time. This is also described as **expansion** of the survival function.

But (ii) changes when we focus on oldest ages: see later.
MALE ENGLISH LIFE TABLE  Based on a birth cohort of 100,000
## DISTRIBUTION OF AGES AT DEATH CONDITIONAL ON REACHING AGE 65 (ENGLAND AND WALES)

<table>
<thead>
<tr>
<th>Year</th>
<th>Males Median</th>
<th>Males IQR</th>
<th>Females Median</th>
<th>Females IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1851</td>
<td>75.0</td>
<td>70.1 – 80.1</td>
<td>75.8</td>
<td>70.6 – 81.5</td>
</tr>
<tr>
<td>1871</td>
<td>74.8</td>
<td>70.0 – 80.3</td>
<td>75.7</td>
<td>70.6 – 81.4</td>
</tr>
<tr>
<td>1891</td>
<td>74.6</td>
<td>69.8 – 80.0</td>
<td>75.6</td>
<td>70.5 – 81.2</td>
</tr>
<tr>
<td>1911</td>
<td>75.3</td>
<td>70.4 – 80.8</td>
<td>76.9</td>
<td>71.5 – 82.4</td>
</tr>
<tr>
<td>1931</td>
<td>75.8</td>
<td>70.8 – 81.1</td>
<td>77.8</td>
<td>72.4 – 83.3</td>
</tr>
<tr>
<td>1951</td>
<td>76.3</td>
<td>71.1 – 81.7</td>
<td>79.3</td>
<td>73.7 – 84.7</td>
</tr>
<tr>
<td>1971</td>
<td>76.6</td>
<td>71.2 – 82.4</td>
<td>81.2</td>
<td>75.1 – 86.9</td>
</tr>
<tr>
<td>1991</td>
<td>79.0</td>
<td>72.0 – 85.0</td>
<td>83.5</td>
<td>76.7 – 89.5</td>
</tr>
</tbody>
</table>
IMPLICATIONS FOR ANNUITIES AND PENSIONS

- Future mortality trend needs to be identified and incorporated in present value calculations (e.g. prices and reserves).

- Projections involve risk (for the insurer) which needs to be measured and managed.
AGENDA

• Mortality Trends

• Data and Notation

• Lee-Carter Model: Base

• Lee-Carter Model: Advanced

• Lee-Carter Model: Cohort Effects

• Lee-Carter Model: Risk Measurement

• Concluding Comments
ILLUSTRATION OF DATA CONFIGURATION

Typical rectangular data array and targeted projected array.
**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

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

Random variables:
- \( D_{xt} \) = number of deaths at age \( x \) and time \( t \)
- \( Y_{xt} \) = response (generic)
AGENDA

- Mortality Trends
- Data and Notation
- Lee-Carter Model: Base
- Lee-Carter Model: Advanced
- Lee-Carter Model: Cohort Effects
- Lee-Carter Model: Risk Measurement
- Concluding Comments
3. LEE CARTER MODELS: BASE VERSION

STRUCTURE

One of the standard benchmark models used in many countries, e.g. US Bureau of Census. Lee and Carter (1992) proposed:

\[
\log m_{xt} = \eta_{xt} + \varepsilon_{xt}, \quad \eta_{xt} = \alpha_x + \beta_x \kappa_t,
\]

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

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

$$\sum_{t=t_1}^{t_n} \kappa_t = 0, \sum_{x} \beta_x = 1,$$

and which imply the least squares estimator

$$\hat{\alpha}_x = \log \prod_{t=t_1}^{t_n} \hat{m}_{xt}^{1/n}.$$
\( \alpha_x \): ‘average’ of \( \log m_{xt} \) over time \( t \) so that \( \exp \alpha_x \) represents the general shape of the age-specific mortality profile.

\( \kappa_t \): underlying time trend.

\( \beta_x \): sensitivity of the logarithm of the hazard rate at age \( x \) to the time trend represented by \( \kappa_t \).

\( \varepsilon_{xt} \): effects not captured by the model.
FITTING BY SVD

(Lee and Carter (1992))

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

Thus

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

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 $\kappa_t$ and $\beta_x$.

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

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

(Lee and Carter (1992)); Haberman and Russolillo (2005)

Use least squares regression:

- estimate $\hat{\alpha}_x$ as above
- estimate $\kappa_t$ as the sum over age of $[\log \hat{m}_{xt} - \hat{\alpha}_x ]$
- regress $[\log \hat{m}_{xt} - \hat{\alpha}_x ]$ on $\kappa_t$ without a constant term to estimate $\beta_x$
- adjust $\hat{\kappa}_t$ as above to get equality of actual and expected deaths.
Fitting by Weighted Least Squares (Gaussian)

(Wilmoth (1993))
Perform the iterative process

\[
\hat{\alpha}_x, \hat{\beta}_x, \hat{\kappa}_t; \text{compute } \hat{y}_{xt} \downarrow \\
\text{update } \hat{\alpha}_x; \text{compute } \hat{y}_{xt} \\
\text{update } \hat{\kappa}_t, \text{ adjust s.t. } \sum_{t=t_i}^{t_n} \kappa_i = 0; \text{compute } \hat{y}_{xt} \\
\text{update } \hat{\beta}_x; \text{compute } \hat{y}_{xt} \\
\text{compute } D(y_{xt}, \hat{y}_{xt}) \downarrow \\
\text{repeat; stop when } D(y_{xt}, \hat{y}_{xt}) \text{ converges}
\]

Where

\[
y_{xt} = \log \hat{m}_{xt}, \hat{y}_{xt} = \hat{n}_{xt}, D(y_{xt}, \hat{y}_{xt}) = \sum_{x,t} w_{xt} (y_{xt} - \hat{y}_{xt})^2
\]

with weights

\[
w_{xt} = d_{xt} \quad \text{(or } = 1).
\]

For a typical parameter, we use the updating algorithm:

\[
\text{updated } (\theta) = \theta - \frac{\partial D}{\partial \theta} \left/ \frac{\partial^2 D}{\partial \theta^2} \right.
\]
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)

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

and

(Lee and Carter (1992))

Originally used extrapolations

\[ m_{x,t_n+s} = \exp(\alpha_x + \beta_x \hat{\kappa}_{t_n+s}), \ s > 0 \]

\[ m_{x,t_n+s} = \exp(\alpha_x + \beta_x \hat{\kappa}_{t_n}) \exp\{\beta_x (\hat{\kappa}_{t_n+s} - \hat{\kappa}_{t_n})\}, \ s > 0 \]

\[ m_{x,t_n+s} = \left[ \prod_{t=t_1}^{t_n} \hat{m}^{1/n} \exp(\beta_x \hat{\kappa}_{t_n}) \right] \exp\{\beta_x (\hat{\kappa}_{t_n+s} - \hat{\kappa}_{t_n})\}, \ s > 0. \]
Construct mortality rate projections

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

by alignment with the latest available mortality rates.

Note

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

is a mortality reduction factor, as widely used in the UK and elsewhere.

For ARIMA(0,1,0) with drift parameter \( \lambda \)

\[ F(x, t_n + s) = \exp(\hat{\beta}_x \hat{\lambda}s), \ s > 0, \]
POISSON SETTING

(Brouhns et al (2002a))

Define

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

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

\[ w_{xt} = \begin{cases} 
1, & e_{xt} > 0 \\
0, & e_{xt} = 0 
\end{cases} \]
APPLICATIONS

England and Wales mortality experience by gender, with grouped ages

LC: Original Lee-Carter model (Gaussian)
GW: Gaussian bilinear version of Lee-Carter with weights $w_{xt} = d_{xt}$
PB: Poisson bilinear version of Lee-Carter

Plots show

- estimates
- residual plots
- comparative log mortality reduction factor projections for 2020 – using ARIMA (1,1,0) time series models
- age-specific log mortality rate profiles.
E+W female mortality experience (LC)

- \[ \alpha \]
- \[ \beta \]
- \[ \kappa \]

**Graphs**

1. **Alpha** vs. age
   - Y-axis: \(-9\) to \(-1\)
   - X-axis: 0 to 90

2. **Beta** vs. age
   - Y-axis: \(0.00\) to \(0.14\)
   - X-axis: 0 to 90

3. **Kappa** vs. calendar year
   - Y-axis: \(-25\) to \(10\)
   - X-axis: 1950 to 2020

4. **(actual-expected) total deaths**
   - Y-axis: \(-20000\) to \(20000\)
   - X-axis: 1950 to 2000

**Legend**

- Solid line: actual deaths
- Dotted line: expected deaths
E+W male mortality experience (LC)

- Graphs showing trends in mortality rates over age and calendar year.
AGENDA

• Mortality Trends
• Data and Notation
• Lee-Carter Model: Base
  • Lee-Carter Model: Advanced
• Lee-Carter Model: Cohort Effects
• Lee-Carter Model: Risk Measurement
• Concluding Comments
4. LEE-CARTER MODELS: ADVANCED

DOUBLE BILINEAR STRUCTURES

(e.g. Booth et al. (2002), Renshaw & Haberman (2003c & d))

The LC model structure may be expanded to

\[ \eta_{xt} = \alpha_x + \sum_{i=1}^{2} \beta_x^{(i)} \kappa_t^{(i)} \]

under the Gaussian setting, or equivalently

\[ \eta_{xt} = \log e_{xt} + \alpha_x + \sum_{i=1}^{2} \beta_x^{(i)} \kappa_t^{(i)} \]

under the Poisson setting.
DOUBLE BILINEAR STRUCTURES (continued)

Fitting is possible by expanding the previous methods: either by incorporating the 2\textsuperscript{nd} SVD components into the model structure, or through the inclusion of additional stages in the core of the iterative fitting methods.

Univariate or multivariate time series

\[
\{(\hat{\kappa}_t^{(1)}, \hat{\kappa}_t^{(2)}) : t \in [t_1, t_n]\} \mapsto \{(\hat{\kappa}_{t_n+s}^{(1)}, \hat{\kappa}_{t_n+s}^{(2)}): s > 0\}
\]

are then applied, and the mortality reduction factor

\[
F(x, t_n + s) = \exp \sum_{i=1}^{2} \beta_x^{(i)} (\hat{\kappa}_{t_n+s}^{(i)} - \hat{\kappa}_{t_n}^{(i)}), \ s > 0
\]

used to project mortality rates.
AGENDA

- Mortality Trends
- Data and Notation
- Lee-Carter Model: Base
- Lee-Carter Model: Advanced
- Lee-Carter Model: Cohort Effects
- Lee-Carter Model: Risk Measurement
- Concluding Comments
6. MODELLING COHORT EFFECTS IN LC FRAMEWORK
AGE-PERIOD-COHORT MODELS

Renshaw & Haberman (2006)
The Lee-Carter structure may be expanded to

\[ M : \eta_{xt} = \alpha_x + \beta_{x}^{(0)} \nu_{t-x} + \beta_{x}^{(1)} \kappa_t \]

under the Gaussian setting, or equivalently,

\[ M : \eta_{xt} = \log e_{xt} + \alpha_x + \beta_{x}^{(0)} \nu_{t-x} + \beta_{x}^{(1)} \kappa_t \]

under the Poisson setting.

The age-cohort substructure
AC: \( \beta_{x}^{(1)} = 0 \)
is also of interest, while we recall that for the standard model
LC: \( \beta_{x}^{(0)} = 0 \).
FITTING AGE-COHORT MODELS (AC)

Note the similarity with LC structure, with $\kappa_t$ replacing $\kappa_t$. Thus

$$AC : \eta_{xt} = \alpha_x + \beta_x t_{t-x}$$

or

$$AC : \eta_{xt} = \log e_{xt} + \alpha_x + \beta_x t_{t-x}$$

subject to the constraints

$$t_{t-k} = 0, \sum_x {\beta_x} = 1.$$

Fitting then follows by adjusting the core of the iterative fitting method in terms of $t_{t-x}$.
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 \( \alpha_x \) is estimated first, according to the original Lee-Carter SVD approach, thus

\[
\hat{\alpha}_x = \log \prod_{t=t_1}^{t_n} \hat{m}_{xt}^{1/n}.
\]

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

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

Effective starting values are obtained by setting \( \beta_x^{(0)} = \beta_x^{(1)} = 1 \) and fitting \( H_0 \) to generate starting values for \( t_z \) and \( \kappa_t \).
PROJECTIONS

Use separate ARIMA time series

\[ \{\hat{K}_t : t \in [t_1, t_n]\} \rightarrow \{\hat{K}_{t_n+s} : s > 0\}, \]

\[ \{\hat{t}_z : z \in [t_1 - x_k, t_n - x_1]\} \rightarrow \{\hat{i}_{t_n-x_1+s} : s > 0\}, \]

then

\[ F(x, t_n + s) = \exp\{\hat{\beta}_x^{(0)} (\hat{t}_{t_n-x+s} - \hat{t}_{t_n-x}) + \hat{\beta}_x^{(1)} (\hat{K}_{t_n+s} - \hat{K}_{t_n})\}, \ s > 0 \]

where

\[ \hat{t}_{t_n-x+s} = \hat{t}_{t_n-x+s}, \ s \leq x-x_1 \]

\[ \hat{i}_{t_n-x+s} = \hat{i}_{t_n-x+s}, \ s > x-x_1. \]
APPLICATIONS

Data set
- UK mortality experience by gender (and single year of age)

Error structure
- Poisson setting

Residual plots
- Age period LC structure fails to capture cohort effects (ripple effect for YoB prior to 1950s)
- Age cohort AC structure shows a ripple period effect
- Note features of M structure.
UK female mortality experience (LC) - residual plots
UK female mortality experience (AC)- residual plots

+ve residuals

-ve residuals

age

year

0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95


0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95

UK female mortality experience (M)- residual plots

- Residuals vs. calendar year
- Residuals vs. age
- Residuals vs. year of birth
Model fitting

- Estimates are conditional on order of specification of 3 main effects

\[ H_0 : \eta_{xt} = \log e_{xt} + \alpha_x + \iota_{t-x} + \kappa_t \]

- Need to condition on one of main effects, e.g. age effects
UK female mortality experience (H(0))

(a) linear predictor: \eta(x,t) = \text{offset} + \alpha(x) + \iota(t-x) + \kappa(t)

(b) linear predictor: \eta(x,t) = \text{offset} + \kappa(t) + \alpha(x) + \iota(t-x)

(c) linear predictor: \eta(x,t) = \text{offset} + \iota(t-x) + \kappa(t) + \alpha(x)
APPLICATIONS (continued)

Time Series Forecasts

- For $\kappa_t : y_t = a_0 + a_1 t + \sum_{i=1}^{p} \phi_i y_{t-i}$ with $p=2$
  
  $$= \kappa_t - \kappa_{t-1}$$

($a_1 \neq 0$ for males; $a_1 = 0$ for females)

- For $\iota_z$ ARIMA (1,1,0)
(a) log(mortality rates): projections by age-period & age-cohort

E+W female population study

KEY
- age-cohort (2025)
- age-period (2025)

E+W male population study

KEY
- age-cohort (2025)
- age-period (2025)

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

E+W female population study

KEY
- age-period-cohort (2025)
- age-period (2025)

E+W male population study

KEY
- age-period-cohort (2025)
- age-period (2025)

Latest and projected $\log \mu_{xt}$ age profiles:
(a) LC and AC modelling; (b) LC and M modelling
England and Wales population, parameter estimates, model M: (a) females; (b) males
Life expectancies at age 65 for a range of periods, computed by period and by cohort under age-period (LC) and age-period-cohort (M) modelling.
AGENDA

- Mortality Trends
- Data and Notation
- Lee-Carter Model: Base
- Lee-Carter Model: Advanced
- Lee-Carter Model: Cohort Effects
- Lee-Carter Model: Risk Measurement
- Concluding Comments
RISK MEASUREMENT AND PREDICTION INTERVALS

- uncertainty in projections needs to be quantified i.e. by prediction intervals
- but analytical derivations are impossible
- 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 $\alpha_x, \beta_x, \kappa_t$ and ARIMA parameters.
DIFFERENT SIMULATION STRATEGIES


A) 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.

Repeat for $j = 1, ..., N$
B) Simulate $e^{(j)}$ vector of $N(0,1)$ deviates

Let $C$ be the Cholesky factorisation matrix of the variance-covariance matrix

Compute simulated model parameters

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

where $\varphi$ is optional scale parameter

Repeat for $j = 1, ..., N$
C) 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.

Repeat for $j = 1, \ldots, N$
UK MALE PENSIONERS: COMPARISON OF SIMULATED 2.5, 50, 97.5 PERCENTILES AND M.L.E. ESTIMATES

(a) Life expectancy, age 65, year 2003

(b) 4 percent fixed rate life annuity, age 65, year 2003
UK MALE PENSIONERS: COMPARISON OF SIMULATED 2.5, 50, 97.5 PERCENTILES AND M.L.E. ESTIMATES

(a) Life expectancy, age 75, year 2003

(b) 4 percent fixed rate life annuity, age 75, year 2003
UK MALE PENSIONERS: COMPARISON OF SIMULATED 2.5, 50, 97.5 PERCENTILES AND M.L.E. ESTIMATES

(a) Life expectancy, age 65, year 2020

(b) 4 percent fixed rate life annuity, age 65, year 2020
AGENDA

• Mortality Trends
• Data and Notation
• Lee-Carter Model: Base
• Lee-Carter Model: Advanced
• Lee-Carter Model: Cohort Effects
• Lee-Carter Model: Risk Measurement

• Concluding Comments
• Prediction intervals for Lee-Carter models – Bayesian methods.

• Model error – essential to investigate more than one modelling framework.

• Extreme ages – extrapolation methods needed.

• Forecasting structural changes.

• Time series methods and long forecasting periods.
• Effect of $\beta_x$ on smoothness of projected age profiles.

• Performance of forecasting methods.

• Quality and appropriateness of data.

• Sources of error – process, parameter, model, judgement.
REFERENCES


REFERENCES (continued)


REFERENCES (continued)


