Mortality and Deprivation

Torsten Kleinow
joint work with Jie Wen and Andrew J.G. Cairns

Heriot-Watt University, Edinburgh

Actuarial Research Centre, IFoA

IFoA ARC Workshop
2 December 2019
The views expressed in this presentation are those of invited contributors and not necessarily those of the Institute and Faculty of Actuaries (IFoA). The IFoA does not endorse any of the views stated, nor any claims or representations made in this presentation and accept no responsibility or liability to any person for loss or damage suffered as a consequence of their placing reliance upon any view, claim or representation made in this presentation. The information and expressions of opinion contained in this presentation are not intended to be a comprehensive study, nor to provide actuarial advice or advice of any nature and should not be treated as a substitute for specific advice concerning individual situations. On no account may any part of this presentation be reproduced without the written permission of the IFoA.
Crude death rates (males, 2017) by Socio-Economic Group

- number of deaths
- population size
- groups g01 ... g10 refer to deprivation deciles
- roughly linear in age (Gompertz line)
- mortality differentials are decreasing with age
similar shape as male log mortality, but lower level, slightly smaller differences

again, mortality differentials are decreasing with age
Crude death rates (males, age 65) by Socio-Economic Group

- clear differences between groups but some crossovers
- downward trend strongest for least deprived
Crude death rates (females, age 65) by Socio-Economic Group

- similar shape as for males, clear differences, but more crossovers
- again, different trends for different socio-economic groups
Model for the Number of Death in Different Groups

For each period (calendar year) $t$, age $x$ and socio-economic group $i$ we assume for the number of deaths, $D_{xti}$:

$$D_{xti} \sim \text{Poisson} (m_{xti}E_{xti})$$

where

- $E_{xti}$: Central exposure-to-risk (mid-year population estimate)
- $m_{xti}$: force of mortality

Expected number of deaths $E[D_{xti}] = m_{xti}E_{xti}$
For each period (calendar year) \( t \), age \( x \) and socio-economic group \( i \) we assume for the number of deaths, \( D_{xti} \):

\[
D_{xti} \sim \text{Poisson} \left( m_{xti} E_{xti} \right)
\]

where

- \( E_{xti} \): Central exposure-to-risk (mid-year population estimate)
- \( m_{xti} \): force of mortality

Expected number of deaths \( E[D_{xti}] = m_{xti} E_{xti} \)

**Aims:**
- compare different models for the force of mortality \( m_{xti} \).
- identify common and group-specific parameters
Model for the Number of Death in Different Groups

For each period (calendar year) $t$, age $x$ and socio-economic group $i$ we assume for the number of deaths, $D_{xti}$:

$$D_{xti} \sim \text{Poisson}(m_{xti}E_{xti})$$

where

- $E_{xti}$: Central exposure-to-risk (mid-year population estimate)
- $m_{xti}$: force of mortality

Expected number of deaths $E[D_{xti}] = m_{xti}E_{xti}$

Aims:

- compare different models for the force of mortality $m_{xti}$.
- identify common and group-specific parameters

We define socio-economic groups with reference to the Index of Multiple Deprivation for England.
The IMD is a weighted combination of seven indices of deprivation:

- Income (22.5%)
- Employment (22.5%)
- Education (13.5%)
- Health (13.5%)
- Crime (9.3%)
- Barriers to Housing and Services (9.3%)
- Living environment (9.3%)

source: GOV.UK

IMD is calculated for about 33,000 small geographic areas (LSOA), ordered and split into ten deciles.
We consider mortality data in England for the ten IMD deciles (ranked in 2015).

- ages: 40-89, years: 2001-2017
- source: Office for National Statistics
We consider mortality data in England for the ten IMD deciles (ranked in 2015).

- ages: 40-89, years: 2001-2017
- source: Office for National Statistics

Well known fact: Mortality rates are higher in the most deprived areas compared to the least deprived areas
All considered models are variants of group specific Lee-Carter type models with the extension to a second age-period effect by Renshaw & Haberman (2003):

\[
\log m_{x|t} = \alpha_{xi} + \beta_{xi}^1 \kappa_{ti}^1 + \beta_{xi}^2 \kappa_{ti}^2
\]
Models

All considered models are variants of group specific Lee-Carter type models with the extension to a second age-period effect by Renshaw & Haberman (2003):

\[
\log m_{x\tau i} = \alpha_{xi} + \beta_{1x}^{\tau} \kappa_{\tau i}^{1} + \beta_{2x}^{\tau} \kappa_{\tau i}^{2}
\]

Specific versions include models with:

- **common age effect**: \(\alpha_{xi} = \alpha_{x}\)
- **fixed age effects**: constant \(\beta_{1x}^{\tau} = 1\) and linear \(\beta_{2x}^{\tau} = x - \bar{x}\), where \(\bar{x}\) is the mean age in the data set. (Plat, 2009)
- **non-parametric common age effects**: \(\beta_{x}^{k} = \beta_{x}^{k}\) (Kleinow, 2015)

All considered models are variants of group specific Lee-Carter type models with the extension to a second age-period effect by Renshaw & Haberman (2003):

$$\log m_{xti} = \alpha_{xi} + \beta_{xi}^{1} \kappa_{ti}^{1} + \beta_{xi}^{2} \kappa_{ti}^{2}$$

Specific versions include models with:

- **common age effect**: $\alpha_{xi} = \alpha_{x}$
- **fixed age effects**: constant $\beta_{xi}^{1} = 1$ and linear $\beta_{xi}^{2} = x - \bar{x}$, where $\bar{x}$ is the mean age in the data set. (Plat, 2009)
- **non-parametric common age effects**: $\beta_{xi}^{k} = \beta_{x}^{k}$ (Kleinow, 2015)
- **some common period effects**: $\kappa_{ti}^{k} = \kappa_{t}^{k}$ (Li and Lee, 2005 for common $\kappa_{i}^{1}$)
- **added cohort effects**: $\gamma_{ci}$ or $\gamma_{c}$ for cohort $c = t - x$
Models (without cohort effect)

\[
\begin{align*}
    m_1 \quad \log m_{xti} &= \alpha_{xi} + \beta^{1}_{xi} \kappa^{1}_{ti} + \beta^{2}_{xi} \kappa^{2}_{ti} \\
    m_2 \quad \log m_{xti} &= \alpha_{xi} + \beta^{1}_{xi} \kappa^{1}_{ti} + \beta^{2}_{xi} \kappa^{2}_{ti} \\
    m_3 \quad \log m_{xti} &= \alpha_{xi} + \beta^{1}_{x} \kappa^{1}_{t} + \beta^{2}_{xi} \kappa^{2}_{ti} \\
    m_4 \quad \log m_{xti} &= \alpha_{xi} + \beta^{1}_{xi} \kappa^{1}_{ti} \\
    m_5 \quad \log m_{xti} &= \alpha_{xi} + \beta^{1}_{x} \kappa^{1}_{t} + \beta^{2}_{xi} \kappa^{2}_{ti} \\
    m_6 \quad \log m_{xti} &= \alpha_{x} + \beta^{1}_{x} \kappa^{1}_{t} + \beta^{2}_{x} \kappa^{2}_{ti} \\
    m_7 \quad \log m_{xti} &= \alpha_{xi} + \kappa^{1}_{ti} + (x - \bar{x}) \kappa^{2}_{ti} \\
    m_8 \quad \log m_{xti} &= \alpha_{x} + \kappa^{1}_{ti} + (x - \bar{x}) \kappa^{2}_{ti} \\
    m_9 \quad \log m_{xti} &= \alpha_{xi} + \kappa^{1}_{t} + (x - \bar{x}) \kappa^{2}_{ti} \\
    m_{10} \quad \log m_{xti} &= \alpha_{xi} + \kappa^{1}_{ti} + (x - \bar{x}) \kappa^{2}_{ti} \\
    m_{11} \quad \log m_{xti} &= \alpha_{xi} + \kappa^{1}_{t} + (x - \bar{x}) \kappa^{2}_{t} \\
    m_{12} \quad \log m_{xti} &= \kappa^{1}_{ti} + (x - \bar{x}) \kappa^{2}_{ti}
\end{align*}
\]

(Renshaw & Haberman, 2003) (m1 with common $\beta^{2}_{x}$) (Li and Lee, 2005) (Lee and Carter, 1992) (Kleinow, 2015) (m5 with common $\alpha_{x}$) (Plat, 2009) (m7 with common $\alpha_{x}$) (m7 with common $\kappa^{1}_{t}$) (m7 with common $\kappa^{2}_{t}$) (m7 with common $\kappa^{1}_{t}$ and $\kappa^{2}_{t}$) (Cairns et al., 2006)
Models are nested
Estimation and Identifiability

- Maximum Likelihood estimation based on $D_{xti} \sim \text{Poisson}(\mu_{xti} E_{xti})$ is applied to obtain estimated parameter values.
- Most suggested models have some identifiability issues, that is, different parameter values lead to the same fitted mortality rates $m_{xti}$, and, therefore to the same value of the likelihood function.
- To obtain unique parameter values we apply model-specific constraints.
Estimation and Identifiability

- Maximum Likelihood estimation based on $D_{xti} \sim \text{Poisson}(\mu_{xti}E_{xti})$ is applied to obtain estimated parameter values.
- Most suggested models have some identifiability issues, that is, different parameter values lead to the same fitted mortality rates $m_{xti}$, and, therefore to the same value of the likelihood function.
- To obtain unique parameter values we apply model-specific constraints.
- In a first step, models are ranked according to the Bayesian Information Criterion:

$$BIC = k \log(n) - 2 \log(\hat{L})$$

where $k$ represents the degrees of freedom, $n$ is the sample size (number of years $\times$ ages $\times$ groups), and $\hat{L}$ is the likelihood value
- A smaller BIC indicates a better model
### Quantitative Comparison of Models

<table>
<thead>
<tr>
<th>Model</th>
<th>log((\hat{L}))</th>
<th>BIC</th>
<th>log((\hat{L}))</th>
<th>BIC</th>
<th>(\log m_{xi})</th>
</tr>
</thead>
<tbody>
<tr>
<td>m1</td>
<td>-34398.54</td>
<td>85083.16</td>
<td>-35634.78</td>
<td>87555.64</td>
<td>(\alpha_{x_i} + \beta_{x_i}^{1} \kappa_{ti}^{1} + \beta_{x_i}^{2} \kappa_{ti}^{2})</td>
</tr>
<tr>
<td>m2</td>
<td>-34652.44</td>
<td>81600.86</td>
<td>-35900.31</td>
<td>84096.61</td>
<td></td>
</tr>
<tr>
<td>m3</td>
<td>-34848.27</td>
<td>80689.64</td>
<td>-36065.10</td>
<td>83123.31</td>
<td></td>
</tr>
<tr>
<td>m4</td>
<td>-35083.25</td>
<td>80571.50</td>
<td>-36293.44</td>
<td>82991.87</td>
<td></td>
</tr>
<tr>
<td>m5</td>
<td>-35058.84</td>
<td>78423.59</td>
<td>-36242.45</td>
<td>80790.81</td>
<td></td>
</tr>
<tr>
<td>m6</td>
<td><strong>-35336.06</strong></td>
<td><strong>75069.36</strong></td>
<td><strong>-36702.57</strong></td>
<td><strong>77802.39</strong></td>
<td>(\alpha_{x} + \beta_{x}^{1} \kappa_{ti}^{1} + \beta_{x}^{2} \kappa_{ti}^{2})</td>
</tr>
<tr>
<td>m7</td>
<td>-35653.80</td>
<td>78726.80</td>
<td>-37422.19</td>
<td>82263.59</td>
<td></td>
</tr>
<tr>
<td>m8</td>
<td>-37375.07</td>
<td>78260.70</td>
<td>-38213.32</td>
<td>79937.20</td>
<td>(\alpha_{x} + \kappa_{ti}^{1} + (x - \bar{x}) \kappa_{ti}^{2})</td>
</tr>
<tr>
<td>m9</td>
<td>-36104.58</td>
<td>78325.48</td>
<td>-37821.39</td>
<td>81759.10</td>
<td></td>
</tr>
<tr>
<td>m10</td>
<td>-35746.95</td>
<td>77610.23</td>
<td>-37491.71</td>
<td>81099.75</td>
<td></td>
</tr>
<tr>
<td>m11</td>
<td>-36760.83</td>
<td>78335.10</td>
<td>-38171.44</td>
<td>81156.32</td>
<td></td>
</tr>
<tr>
<td>m12</td>
<td>-46822.86</td>
<td>96721.99</td>
<td>-41385.10</td>
<td>85846.45</td>
<td></td>
</tr>
</tbody>
</table>
Parameter estimates - the most general model

\[ m1: \alpha_{xi} + \beta_{xi}^1 \kappa_{ti}^1 + \beta_{xi}^2 \kappa_{ti}^2 \]

- clear differences between socio-economic groups in basic age structure of mortality
Parameter estimates - common age effects

- $m_5: \alpha_{x_i} + \beta_{x}^1 k_{t_i}^1 + \beta_{x}^2 k_{t_i}^2$
- $m_6: \alpha_x + \beta_{x}^1 k_{t_i}^1 + \beta_{x}^2 k_{t_i}^2$
- estimates for $m_5$ are very similar to those of $m_1$
- dashed line is the common age structure $\alpha_x$ in $m_6$
Jie Wen, AJG. Cairns, T. Kleinow: Mortality and Deprivation

Parameter estimates - Plat model

$$m7: \alpha_{xi} + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2$$

$$m8: \alpha_x + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2$$

- again, a very similar shape
- Summary: basic age structure is almost independent of chosen model
Parameter estimates - the most general model

\[ m1: \alpha \xi_i + \beta_1 \kappa_1 t_i + \beta_2 \kappa_2 t_i \]

- no clear differences between groups
- suggests a common parameter
- but not constant as in m7 and m8
Parameter estimates - common age effects

m5/6 – beta1 – Female Population

- m5: $\alpha_x + \beta_x^1 \kappa_1^{t_i} + \beta_x^2 \kappa_2^{t_i}$
- m6: $\alpha_x + \beta_x^1 \kappa_1^{t_i} + \beta_x^2 \kappa_2^{t_i}$
- shape of $\beta^1$ in m5 is similar to m1
- ... but for m6 the shape is very different
- note that $\beta^1$ is constant in m7 and m8
Parameter estimates - the most general model

\[ m1: \alpha_{xi} + \beta_{xi1}^{1} \kappa_{ti}^{1} + \beta_{xi2}^{2} \kappa_{ti}^{2} \]

- clear differences in the trend of mortality between groups
- least deprived show greatest improvements
Parameter estimates - common age effects

\[ m5: \alpha_{xi} + \beta_1^{x1} \kappa_1^{ti} + \beta_2^{x2} \kappa_2^{ti} \]

- again, different trends for different groups
- period effects are very similar to those in m1
- ... this suggests that projections would also look similar
Parameter estimates - common age effects

\[ m6 = \alpha_x \cdot \kappa_1^{t_1} + \beta_1 \cdot \kappa_2^{t_1} \]

- since age effects are now common, the first period effects picks up differences in level of mortality
- we also see different trends
Parameter estimates - Plat model

- \( m7 = \alpha x_i + \kappa_{1i} + (x - \bar{x})\kappa_{2i} \)
- differences in trend are clearly visible
- note that different constraints have been used (compared to \( m5 \) and \( m6 \))
Parameter estimates - Plat model

\[ m8 - \kappa_1 - \text{Female Population} \]

- \( m8: \alpha_x + \kappa_1^{t_i} + (x - \bar{x})\kappa_2^{t_i} \)
- similar to m6; common \( \alpha_x \) leads to different levels
- trends are also different
Define Pearson’s residuals

\[ Z_{xti} = \frac{D_{xti} - E[D_{xti}]}{\text{std}[D_{xti}]} = \frac{D_{xti} - E_{xti} \hat{m}_{xti}}{\sqrt{E_{xti} \hat{m}_{xti}}} \]

where \( \hat{m}_{xti} \) is the fitted death rate at age \( x \) in year \( t \) that we obtain from our various models.
Standardised residuals - common age effects

- m6: $\alpha_x + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2$
- no obvious clusters or pattern
- good fit
Standardised residuals - common age effects

- m6: $\alpha_x + \beta_1 x \kappa_{t_i} + \beta_2 x \kappa_{t_i}$
- no obvious clusters or pattern
- good fit
Standardised residuals - common age effects

m6: \( \alpha_x + \beta_1^1 \kappa_{ti}^1 + \beta_2^2 \kappa_{ti}^2 \)

- no obvious clusters or pattern
- good fit
Standardised residuals - Plat model

\[
m8: \alpha_x + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2
\]

- pattern along the age dimension for group 1 (most deprived)
Standardised residuals - Plat model

m8: \( \alpha_x + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2 \)

- good fit for group 5
Standardised residuals - Plat model

m8: $\alpha_x + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2$

- pattern along the age dimension for group 1 (most deprived) and 10 (least deprived)
Empirical log mortality rates (left) in 2017 (females),

fitted rates from models m6, $\alpha_x + \beta^1_x \kappa^1_{ti} + \beta^2_x \kappa^2_{ti}$ (middle) and

m8, $\alpha_x + \kappa^1_{ti} + (x - \bar{x}) \kappa^2_{ti}$ (right)

fitted rates are similar, but m8 produces smoother rates
Empirical log mortality rates (left) at age 65 (females),

- fitted rates from models $m6$, $\alpha_x + \beta_1^{\kappa_1 t_i} + \beta_2^{\kappa_2 t_i}$ (middle) and
- $m8$, $\alpha_x + \kappa_1^{t_i} + (x - \bar{x})\kappa_2^{t_i}$ (right)

- again, fitted rates look very similar, in particular, similar improvement rates between models

- ... but different improvement rates for different groups
Improvement Rates for Leading Period Effect

- m6: $\alpha_x + \beta_{x1} \kappa_{ti}^1 + \beta_{x2} \kappa_{ti}^2$ (females)
- large differences; from g1 to g10 the improvement rate doubles
- rescaled with common $\beta_{x1}$
- differences are even greater for model m5
Projections are challenging - Model m6

- Projections require assumptions (expert judgement) about differences in level and improvement rates for leading period effect (left)
- and additional assumptions about $\kappa^2$
Projections are challenging - Model m5

- modelling of $\kappa^2$ seems easier, some correlated stationary processes look appropriate
- ... but the main issue about identifying reasonable assumptions (expert judgement) for leading period effects remains
Conclusions

- Age effects are common to all ten socio-economic groups in England (as measured by the Index of Multiple Deprivation).
- ... but the fit is improved if the age effects $\beta^1$ and $\beta^2$ are not constant and linear functions of age.
- However, the fitted rates look very similar for the CAE model and the Plat model with common alpha.
Conclusions

- age effects are common to all ten socio-economic groups in England (as measured by the Index of Multiple Deprivation)
- ... but the fit is improved if the age effects $\beta^1$ and $\beta^2$ are not constant and linear functions of age
- However, the fitted rates look very similar for the CAE model and the Plat model with common alpha
- period effects are different ...
- ... different levels
- ... but also different trends; mortality differentials are increasing (although there is some evidence that improvement rates even in for the least deprived are slowing down)
- Therefore, models with common period effects (in particular $\kappa^1$) are not a good fit
Conclusions

- Age effects are common to all ten socio-economic groups in England (as measured by the Index of Multiple Deprivation).
- ... but the fit is improved if the age effects $\beta_1$ and $\beta_2$ are not constant and linear functions of age.
- However, the fitted rates look very similar for the CAE model and the Plat model with common alpha.
- Period effects are different ...
- ... different levels.
- ... but also different trends; mortality differentials are increasing (although there is some evidence that improvement rates even in for the least deprived are slowing down).
- Therefore, models with common period effects (in particular $\kappa_1$) are not a good fit.
- The challenge is, of course, to project period effects: what assumptions can we make about long term trends?