

Mortality and Deprivation

Torsten Kleinow

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

Heriot-Watt University, Edinburgh

Actuarial Research Centre, IFoA

Life Conference - Liverpool - November 2018



Actuarial
Research Centre
Institute and Faculty
of Actuaries

- Index of Multiple Deprivation (IMD)
- Observed mortality by age and year
- Simultaneously modelling mortality in all deciles:
 - choosing common parameters
 - comparing the goodness of fit



Index of Multiple Deprivation

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



Index of Multiple Deprivation

The IMD is a weighted combination of seven indices of deprivation:

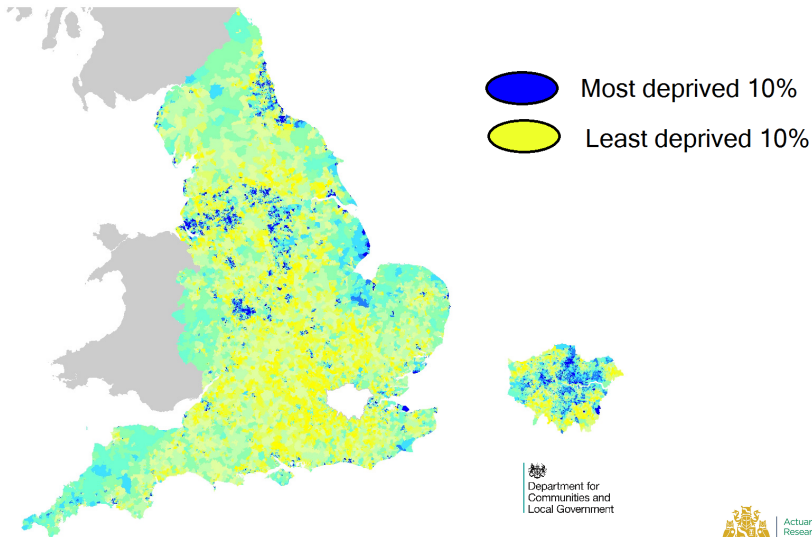
- Income (22.5%)
- Employment (22.5%)
- Education (13.5%)
- Health (13.5%)
- Crime (9.3%)
- Barriers to Housing and Services (9.3%)
- Living environment (9.3%)

source: GOV.UK

- just over 30,000 LSOAs (Lower Layer Super Output Area) in England
- ordered and split into ten deciles:
10% most deprived, ..., 10% least deprived



Index of Multiple Deprivation (IMD) areas



- Index of Multiple Deprivation (IMD)
- **Observed mortality by age and year**
- Simultaneously modelling mortality in all deciles



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



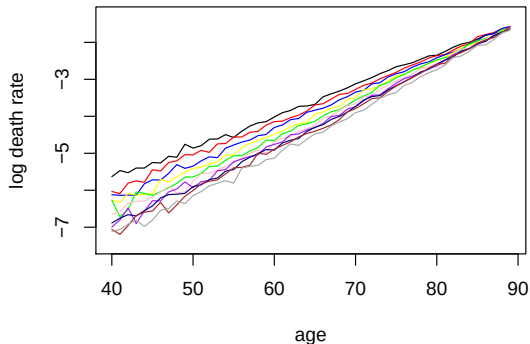
- We consider mortality data for males in England for the ten IMD deciles (2015).
- ages: 40-89, years: 2001-2015
- source: Office for National Statistics

There are more deaths per 1,000 lives in the most deprived areas compared to the least deprived areas.



Mortality rates by IMD decile in 2001

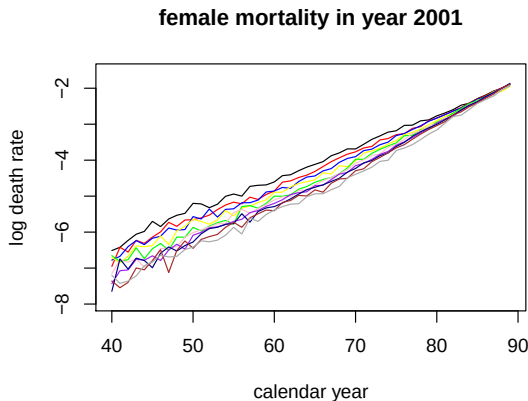
male mortality in year 2001



- roughly linear in age (Gompertz line)
- mortality differentials are decreasing with age



Mortality rates by IMD decile in 2001

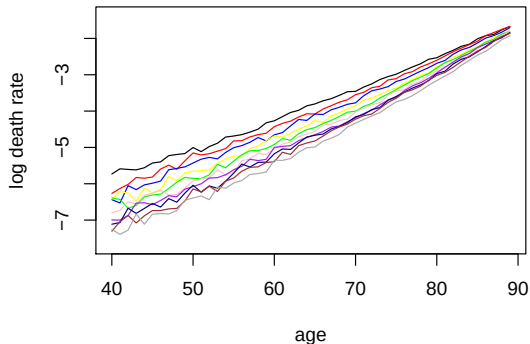


- similar shape as male log mortality, but lower level, slightly smaller differences
- again, mortality differentials are decreasing with age



Mortality rates by IMD decile in 2001

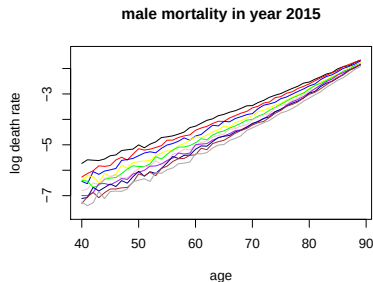
male mortality in year 2015



- similar shape as in 2001
- differences at high ages are larger



Mortality rates by IMD decile



- downward shift from 2001 to 2015
- differences between most deprived and least deprived have increased since 2001
- higher differences at high ages



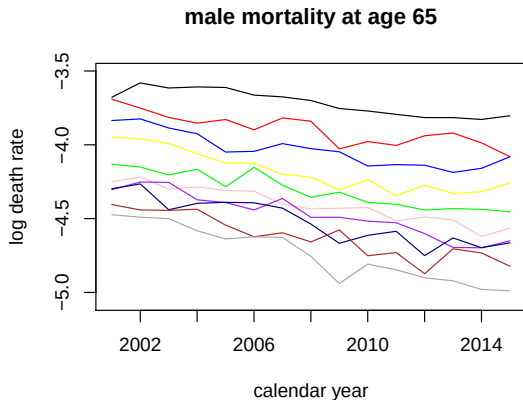
Number of deaths per 1,000 lives by IMD decile - Change over time

	Deaths per 1,000 lives		ratio
	most deprived	least deprived	
2001	25.3	11.4	2.219
2005	27.0	9.7	2.784
2010	23.0	8.2	2.805
2015	22.3	6.8	3.279

Males aged 65



Mortality rates by IMD decile for males aged 65



- downward trend strongest for least deprived



- Index of Multiple Deprivation (IMD)
- Observed mortality by age and year
- **Simultaneously modelling mortality in all deciles**



Stochastic Models for the Number of Deaths by Decile

Moving away from the number of deaths per 1,000 lives, we consider a stochastic mortality model:

For each period (calendar year) t , age x and IMD decile i we define

D_{xti} : Number of deaths,

E_{xti} : Central exposure-to-risk

m_{xti} : force of mortality



Stochastic Models for the Number of Deaths by Decile

Moving away from the number of deaths per 1,000 lives, we consider a stochastic mortality model:

For each period (calendar year) t , age x and IMD decile i we define

D_{xti} : Number of deaths,

E_{xti} : Central exposure-to-risk

m_{xti} : force of mortality

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

So, expected number of deaths, $E[D_{xti}] = m_{xti}E_{xti}$



Stochastic Models for the Number of Deaths by Decile

Moving away from the number of deaths per 1,000 lives, we consider a stochastic mortality model:

For each period (calendar year) t , age x and IMD decile i we define

D_{xti} : Number of deaths,

E_{xti} : Central exposure-to-risk

m_{xti} : force of mortality

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

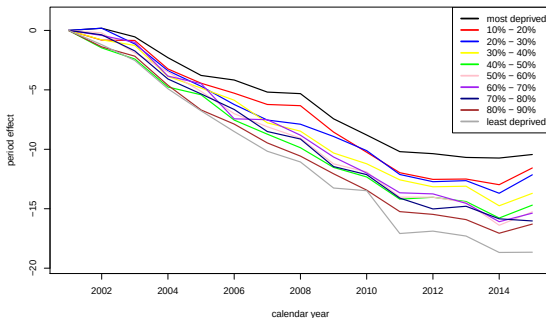
So, expected number of deaths, $E[D_{xti}] = m_{xti}E_{xti}$

Aim: to compare different models for the force of mortality m_{xti} .



Lee-Carter model fitted to individual IMD decile

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



- downward trend strongest for least deprived
- no improvements for most deprived since 2011
- slowdown of improvements for least deprived since 2011



Focus of our research:

- Stochastic models that describe mortality experiences in all socio-economic groups simultaneously.
- Model uncertainty addressed by comparing a wide variety of models (Goodness of fit, robustness, ...)
- Leading to projections, and more importantly, mortality scenario generation allowing us
 - to put probabilities on certain scenarios and ...
 - then use those for Value at Risk calculations, annuity pricing, etc.



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):

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

where $c = t - x$ is the cohort (year of birth).



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):

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

where $c = t - x$ is the cohort (year of birth).

Specific versions include models with:

common age effect : $\alpha_{xi} = \alpha_x$

non-parametric common age effects : $\beta_{xi}^k = \beta_x^k$ (Kleinow, 2015)

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)

common period effects : $\kappa_{ti}^k = \kappa_t^k$ (Li and Lee, 2005)

and variations with and without cohort effects.



Some research questions

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

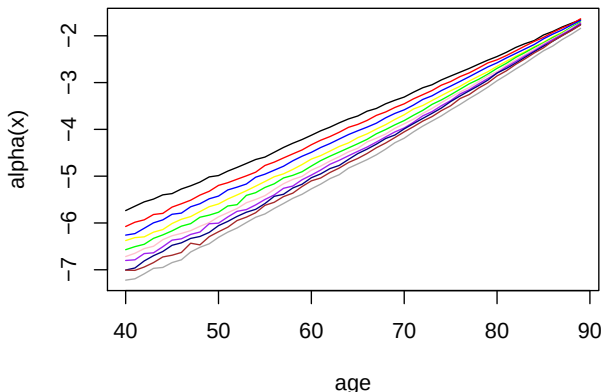
- What parameters should be chosen to be group specific and which parameters are common?
- Should age-effects be estimated?
- Should we include cohort effects (common or group specific)?
- What parameters show the greatest differences between IMD groups?



Parameter estimates - m12 - the most general model

$$\log m_{xti} = \alpha_{xi} + \beta_{xi}^1 \kappa_{ti}^1 + \beta_{xi}^2 \kappa_{ti}^2 \quad (\text{Renshaw\&Haberman, 2003})$$

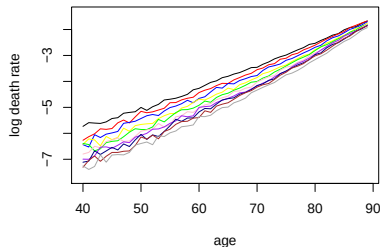
MLE estimated $\alpha(x)$ – m12



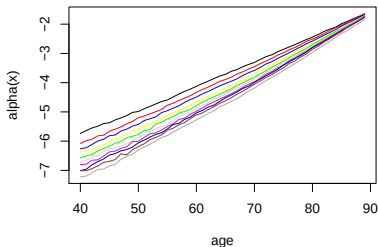
Parameter estimates - m12 - the most general model

$$\log m_{xti} = \alpha_{xi} + \beta_{xi}^1 \kappa_{ti}^1 + \beta_{xi}^2 \kappa_{ti}^2 \quad (\text{Renshaw\&Haberman, 2003})$$

male mortality in year 2015



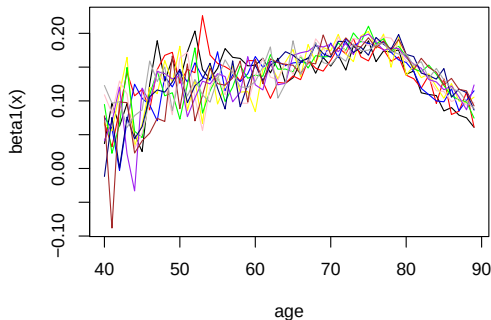
MLE estimated alpha(x) - m12



Parameter estimates - m12 - the most general model

$$\log m_{x_{ti}} = \alpha_{xi} + \beta_{xi}^1 \kappa_{ti}^1 + \beta_{xi}^2 \kappa_{ti}^2 \text{ (Renshaw\&Haberman, 2003)}$$

MLE estimated beta1(x) – m12



general shape seems
to be the same for all
deciles

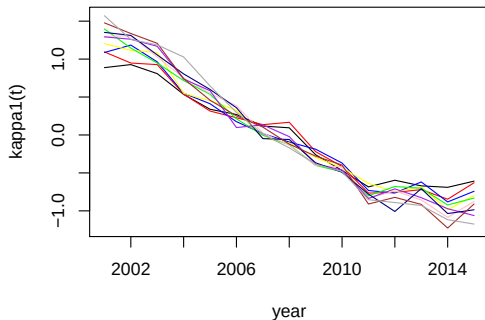
differences might not
be important



Parameter estimates - m12 - the most general model

$$\log m_{x_{ti}} = \alpha_{xi} + \beta_{xi}^1 \kappa_{ti}^1 + \beta_{xi}^2 \kappa_{ti}^2 \text{ (Renshaw\&Haberman, 2003)}$$

MLE estimated kappa1(t) – m12



strongest slope for
least deprived decile
(grey)

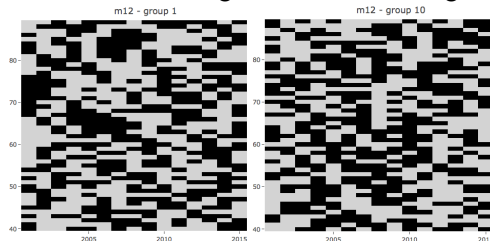
weakest improvements for most
deprived decile
(black)



Residuals - m12 - the most general model

$$\log m_{xti} = \alpha_{xi} + \beta_{xi}^1 \kappa_{ti}^1 + \beta_{xi}^2 \kappa_{ti}^2 \quad (\text{Renshaw\&Haberman, 2003})$$

Residuals = Fitted log rates - observed log rates



black shows positive residuals (fitted rates are too high)
grey shows negative residuals (fitted rates are too low)

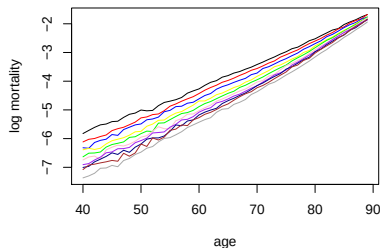
no obvious pattern - good



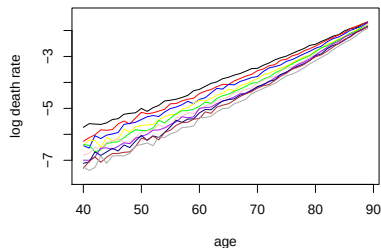
Parameter estimates - m12 - the most general model

$$\log m_{xti} = \alpha_{xi} + \beta_{xi}^1 \kappa_{ti}^1 + \beta_{xi}^2 \kappa_{ti}^2 \quad (\text{Renshaw\&Haberman, 2003})$$

Fitted mortality in 2015 - m12



male mortality in year 2015



Parameter estimates - m12 - the most general model

Bayesian Information Criterion: $k \log n - 2 \log(L)$

	Log-likelihood	parameters	constraints	d.o.f.	BIC
m12	-30,131.47	1800	40	1760	75,966.82



Parameter estimates - m12 - the most general model

Bayesian Information Criterion: $k \log n - 2 \log(L)$

	Log-likelihood	parameters	constraints	d.o.f.	BIC
m12	-30,131.47	1800	40	1760	75,966.82

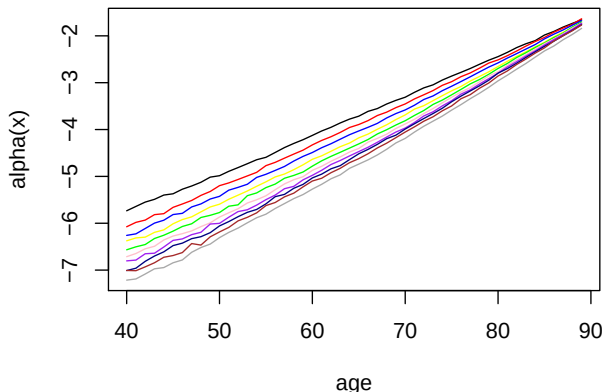
- good fit
- very large number of parameters



Parameter estimates - m2 - common beta

$$\log m_{xti} = \alpha_{xi} + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2 \text{ (Kleinow, 2015)}$$

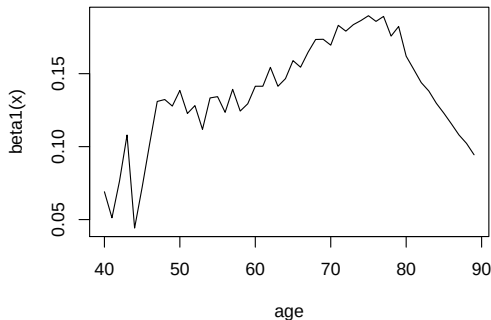
MLE estimated $\alpha(x)$ – m2



Parameter estimates - m2 - common beta

$$\log m_{xti} = \alpha_{xi} + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2 \text{ (Kleinow, 2015)}$$

MLE estimated beta1(x) – m2



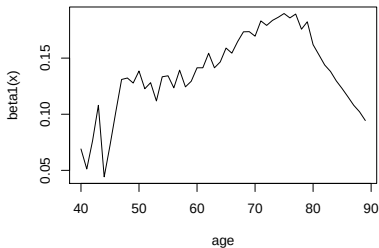
general shape is similar to shape of individual β^1 in previous model



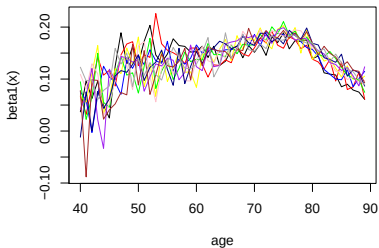
Parameter estimates - m2 - common beta

$$\log m_{xti} = \alpha_{xi} + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2 \quad (\text{Kleinow, 2015})$$

MLE estimated beta1(x) - m2



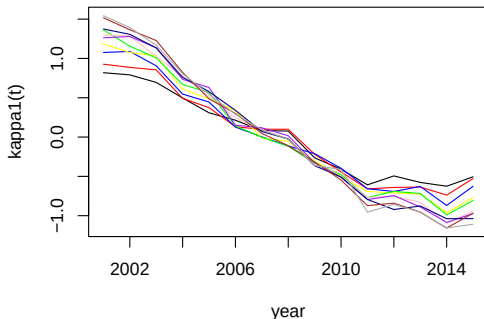
MLE estimated beta1(x) - m12



Parameter estimates - m2 - common beta

$$\log m_{xti} = \alpha_{xi} + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2 \text{ (Kleinow, 2015)}$$

MLE estimated kappa1(t) – m2



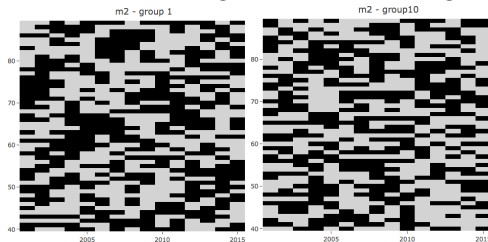
similar to period effects in previous (more general) model

clearly different improvements for different IMD deciles

Parameter estimates - m2 - common beta

$$\log m_{xti} = \alpha_{xi} + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2 \text{ (Kleinow, 2015)}$$

Residuals = Fitted log rates - observed log rates



black shows positive residuals (fitted rates are too high)

grey shows negative residuals (fitted rates are too low)

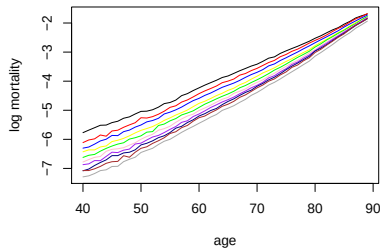
no obvious pattern - age effects seem to be common to all deciles



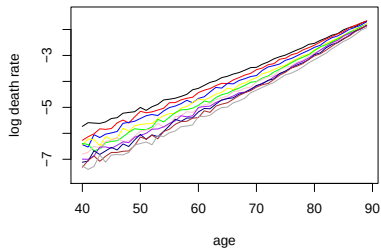
Parameter estimates - m2 - common beta

$$\log m_{xti} = \alpha_{xi} + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2 \quad (\text{Kleinow, 2015})$$

Fitted mortality in 2015 - m2



male mortality in year 2015



Parameter estimates - m2 - common beta

Bayesian Information Criterion: $k \log n - 2 \log(L)$

	Log-likelihood	parameters	constraints	d.o.f.	BIC
m2	-30,591.20	900	22	878	69,016.49
m12	-30,131.47	1800	40	1760	75,966.82



Parameter estimates - m2 - common beta

Bayesian Information Criterion: $k \log n - 2 \log(L)$

	Log-likelihood	parameters	constraints	d.o.f.	BIC
m2	-30,591.20	900	22	878	69,016.49
m12	-30,131.47	1800	40	1760	75,966.82

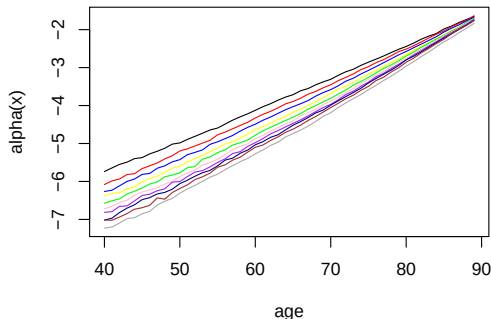
- reasonable fit
- smaller number of parameters than m12
- BIC improved



Parameter estimates - m1 - constant & linear beta

$$\log m_{xti} = \alpha_{xi} + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2 \text{ (Plat, 2009)}$$

MLE estimated alpha(x) – m1



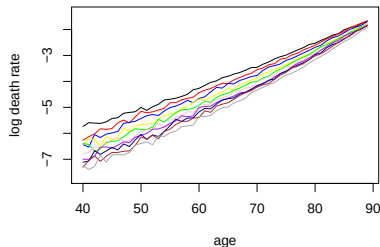
The first age effect is almost identical to α in the other two models



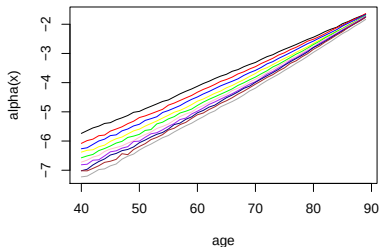
Parameter estimates - m1

$$\log m_{xti} = \alpha_{xi} + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2 \text{ (Plat, 2009)}$$

male mortality in year 2015



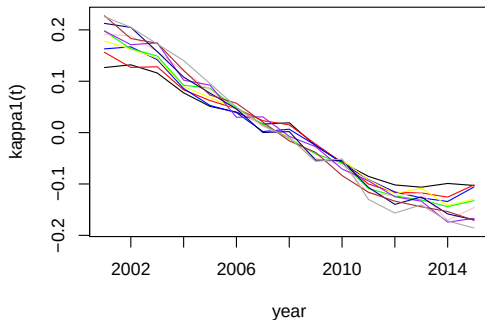
MLE estimated $\alpha(x)$ - m1



Parameter estimates - m1

$$\log m_{xti} = \alpha_{xi} + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2 \text{ (Plat, 2009)}$$

MLE estimated kappa1(t) - m1



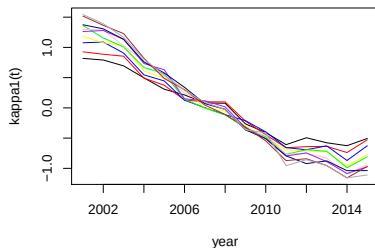
The first period effect is also similar to the period effects in the other two models



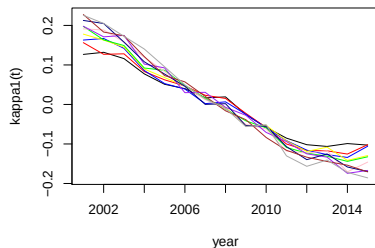
Parameter estimates - m1

$$\log m_{xti} = \alpha_{xi} + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2 \text{ (Plat, 2009)}$$

MLE estimated kappa1(t) - m2

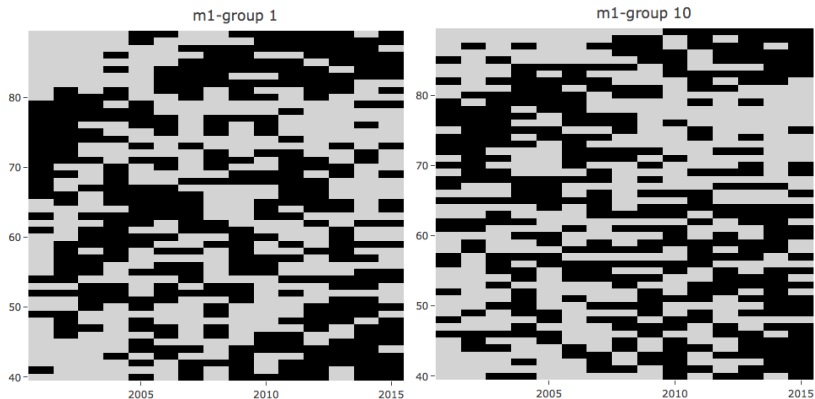


MLE estimated kappa1(t) - m1



Parameter estimates - m1

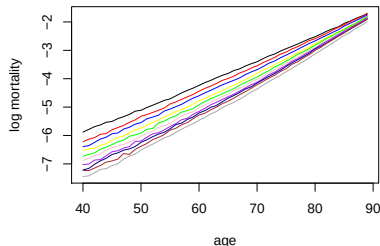
$$\log m_{xti} = \alpha_{xi} + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2 \text{ (Plat, 2009)}$$



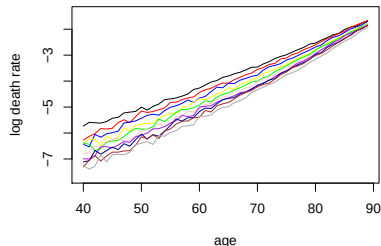
Parameter estimates - m1

$$\log m_{xti} = \alpha_{xi} + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2 \text{ (Plat, 2009)}$$

Fitted mortality in 2015 - m1



male mortality in year 2015



Fitted mortality rates are now very smooth



Parameter estimates - m1

Bayesian Information Criterion: $k \log n - 2 \log(L)$

	Log-likelihood	parameters	constraints	d.o.f.	BIC
m1	-31,403.22	800	20	780	69,766.10
m2	-30,591.20	900	22	878	69,016.49
m12	-30,131.47	1800	40	1760	75,966.82



Parameter estimates - m1

Bayesian Information Criterion: $k \log n - 2 \log(L)$

	Log-likelihood	parameters	constraints	d.o.f.	BIC
m1	-31,403.22	800	20	780	69,766.10
m2	-30,591.20	900	22	878	69,016.49
m12	-30,131.47	1800	40	1760	75,966.82

- model does not fit as well as m2 and m12 (lower likelihood)
- number of parameters reduced further
- but BIC does not improve



Next step

$$\begin{aligned} \text{m1:} \quad \log m_{xti} &= \alpha_{xi} + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2 && (\text{Plat, 2009}) \\ \text{m2:} \quad \log m_{xti} &= \alpha_{xi} + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2 && (\text{Kleinow, 2015}) \\ \text{m12:} \quad \log m_{xti} &= \alpha_{xi} + \beta_{xi}^1 \kappa_{ti}^1 + \beta_{xi}^2 \kappa_{ti}^2 && (\text{Renshaw\&Haberman, 2003}) \end{aligned}$$



Next step

$$\begin{aligned} \text{m1:} \quad \log m_{xti} &= \alpha_{xi} + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2 && (\text{Plat, 2009}) \\ \text{m2:} \quad \log m_{xti} &= \alpha_{xi} + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2 && (\text{Kleinow, 2015}) \\ \text{m12:} \quad \log m_{xti} &= \alpha_{xi} + \beta_{xi}^1 \kappa_{ti}^1 + \beta_{xi}^2 \kappa_{ti}^2 && (\text{Renshaw\&Haberman, 2003}) \end{aligned}$$

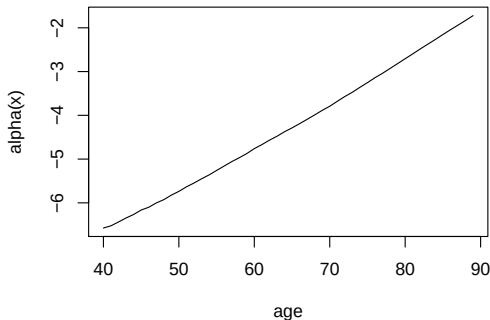
Try a common alpha for m1 and m2.



Parameter estimates - m14 - common alpha and beta

$$\log m_{xti} = \alpha_x + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2$$

MLE estimated alpha(x) - m14



The main age effect is now very smooth

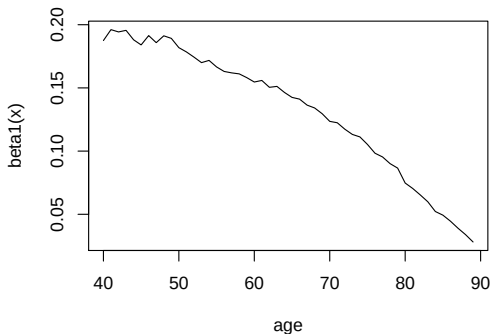
Decile specific noise is averaged out



Parameter estimates - m14 - common alpha and beta

$$\log m_{xti} = \alpha_x + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2$$

MLE estimated beta1(x) - m14



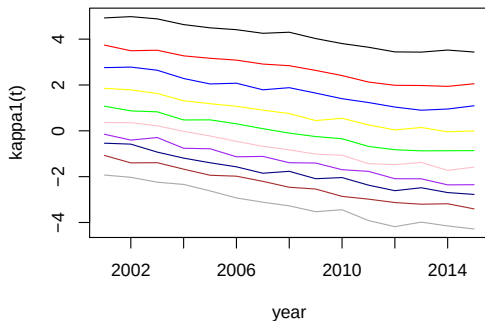
Different shape compared to common β in previous model



Parameter estimates - m14 - common alpha and beta

$$\log m_{xti} = \alpha_x + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2$$

MLE estimated kappa1(t) – m14

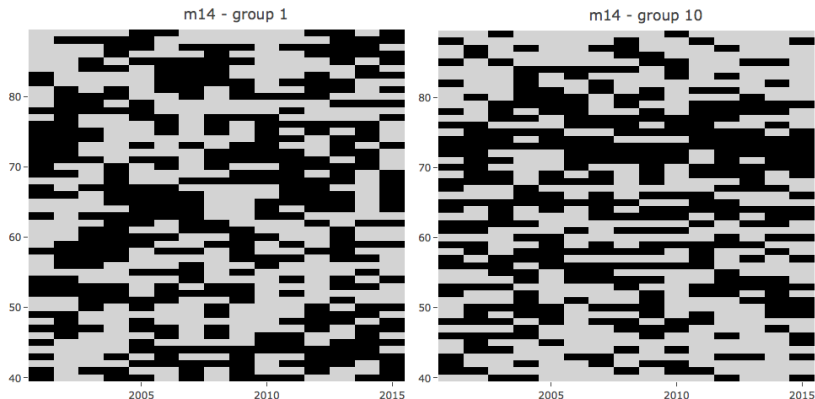


The period effect κ is now showing differences in the level of mortality (since α are common)



Parameter estimates - m14 - common alpha and beta

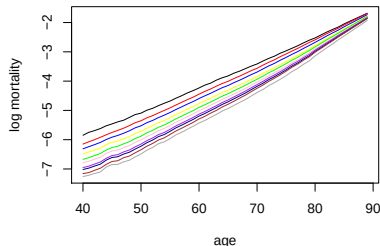
$$\log m_{xti} = \alpha_x + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2$$



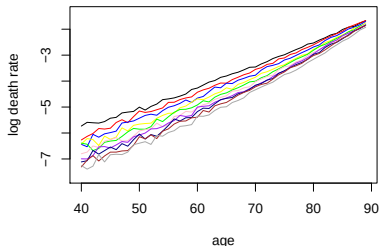
Parameter estimates - m14 - common alpha and beta

$$\log m_{xti} = \alpha_x + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2$$

Fitted mortality in 2015 - m14



male mortality in year 2015



Fitted rates are now very smooth



Parameter estimates - m14 - common alpha and beta

$$\log m_{xti} = \alpha_x + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2$$

Bayesian Information Criterion: $k \log n - 2 \log(L)$

	Log-likelihood	parameters	constraints	d.o.f.	BIC
m1	-31,403.22	800	20	780	69,766.10
m2	-30,591.20	900	22	878	69,016.49
m12	-30,131.47	1800	40	1760	75,966.82
m14	-30,852.96	450	4	446	65,685.42



Parameter estimates - m14 - common alpha and beta

$$\log m_{xti} = \alpha_x + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2$$

Bayesian Information Criterion: $k \log n - 2 \log(L)$

	Log-likelihood	parameters	constraints	d.o.f.	BIC
m1	-31,403.22	800	20	780	69,766.10
m2	-30,591.20	900	22	878	69,016.49
m12	-30,131.47	1800	40	1760	75,966.82
m14	-30,852.96	450	4	446	65,685.42

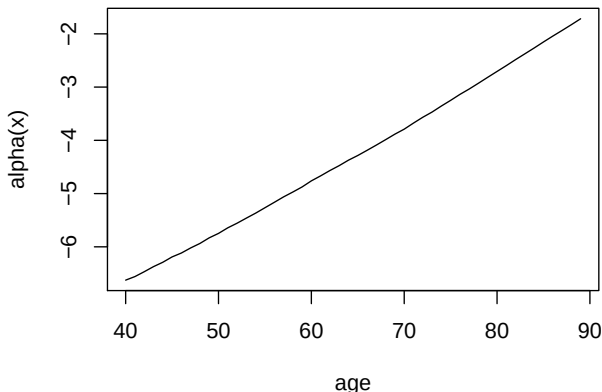
- higher likelihood than m1 with fewer parameters
- best BIC



Parameter estimates - m6 - M1 with common alpha

$$\log m_{xti} = \alpha_x + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2$$

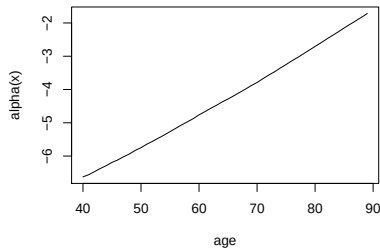
MLE estimated alpha(x) – m6



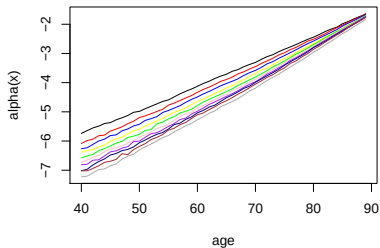
Parameter estimates - m6 - M1 with common alpha

$$\log m_{xti} = \alpha_x + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2$$

MLE estimated alpha(x) - m6



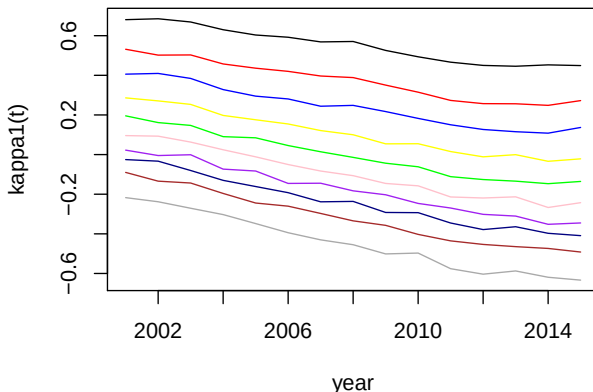
MLE estimated alpha(x) - m1



Parameter estimates - m6 - M1 with common alpha

$$\log m_{xti} = \alpha_x + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2$$

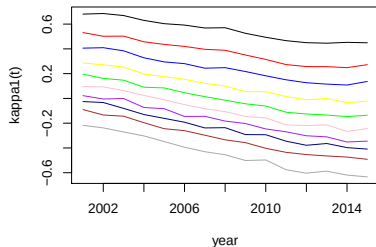
MLE estimated kappa1(t) - m6



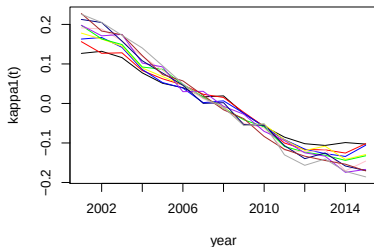
Parameter estimates - m6 - M1 with common alpha

$$\log m_{xti} = \alpha_x + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2$$

MLE estimated kappa1(t) - m6



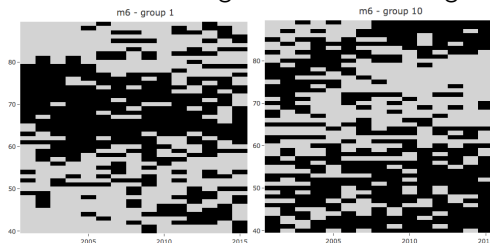
MLE estimated kappa1(t) - m1



Parameter estimates - m6 - M1 with common alpha

$$\log m_{xti} = \alpha_x + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2$$

Residuals = Fitted log rates - observed log rates



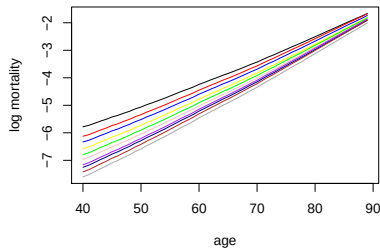
black shows positive residuals (fitted rates are too high)
grey shows negative residuals (fitted rates are too low)
seems to be a pattern (age direction)



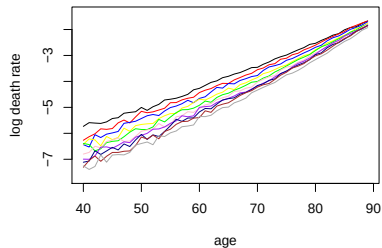
Parameter estimates - m6 - M1 with common alpha

$$\log m_{xti} = \alpha_x + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2$$

Fitted mortality in 2015 - m6



male mortality in year 2015



Parameter estimates - m6 - M1 with common alpha

Bayesian Information Criterion: $k \log n - 2 \log(L)$

	Log-likelihood	parameters	constraints	d.o.f.	BIC
m1	-31,403.22	800	20	780	69,766.10
m2	-30,591.20	900	22	878	69,016.49
m6	-32,052.10	350	2	348	67,209.29
m12	-30,131.47	1800	40	1760	75,966.82
m14	-30,852.96	450	4	446	65,685.42



Parameter estimates - m6 - M1 with common alpha

Bayesian Information Criterion: $k \log n - 2 \log(L)$

	Log-likelihood	parameters	constraints	d.o.f.	BIC
m1	-31,403.22	800	20	780	69,766.10
m2	-30,591.20	900	22	878	69,016.49
m6	-32,052.10	350	2	348	67,209.29
m12	-30,131.47	1800	40	1760	75,966.82
m14	-30,852.96	450	4	446	65,685.42

- smallest number of paramters
- second best model in terms of BIC
- better BIC than m1



Our conclusions so far...

- There are clear differences between the mortality rates for the ten IMD deciles.
- The improvement rates (from 2001 – 2015) are also different.



Our conclusions so far...

- There are clear differences between the mortality rates for the ten IMD deciles.
- The improvement rates (from 2001 – 2015) are also different.
- Models with common age effects seem to perform better than models with group specific age effects.



Our conclusions so far...

- There are clear differences between the mortality rates for the ten IMD deciles.
- The improvement rates (from 2001 – 2015) are also different.
- Models with common age effects seem to perform better than models with group specific age effects.
- For a wider age range, models with common non-parametric age effects (Kleinow (2015) + common α) produce a good fit in terms of BIC and heatmaps.
- However, for a narrower age range (65-89), models with constant/linear β 's, (Plat (2009) + common α) are better.



Our conclusions so far...

- There are clear differences between the mortality rates for the ten IMD deciles.
- The improvement rates (from 2001 – 2015) are also different.
- Models with common age effects seem to perform better than models with group specific age effects.
- For a wider age range, models with common non-parametric age effects (Kleinow (2015) + common α) produce a good fit in terms of BIC and heatmaps.
- However, for a narrower age range (65-89), models with constant/linear β 's, (Plat (2009) + common α) are better.
- Cohort effects do not improve the fit of those models.
- If a cohort effect is included it should be a common cohort effect.



Thank You!

Questions and Comments

