Cause of death specific cohort effects in US mortality

Cristian Redondo Loures, Andrew J. G. Cairns

Heriot-Watt University, Edinburgh

ATRC, Liverpool, June 2019

Outline

- Introduction
- Data sources
- Building crude death rates
- Stochastic modelling
- Conclusions

Actuarial Research Centre
Institute and Faculty of Actuaries

Introduction

- Gain insight into the impact of socioeconomic status on mortality to inform projections.
- Analyse which causes of death affect different groups within a population.
- Use stochastic modelling to understand the drivers of the observed behaviour.

US mortality by single ages and calendar years, independently for each gender, CoD, and education level.

Actuarial Research Centre
Institute and Faculty of Actuaries

Introduction

Education as a covariate in death rates

- Education ⇒ Socioeconomic status.
- "Fixed" through adulthood.
- Data readily available (quality to be assesed though!).

Case and Deaton (2015) found interesting trends in cause of death mortality in the US by education level.

Data sources

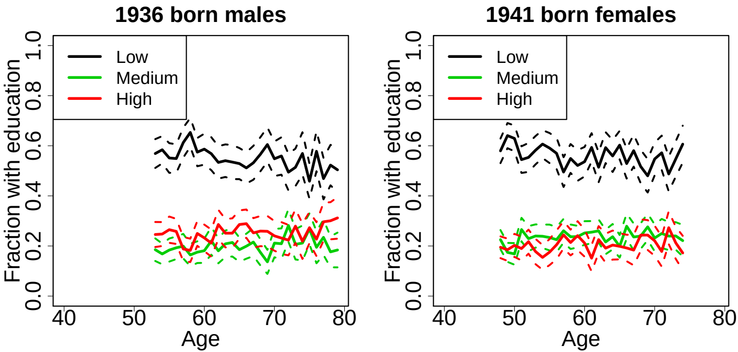
Centers for Disease Control and Prevention (CDC) $\rightarrow$ Data from death certificates, D^C(x,b,e,g).

Human Mortality database (HMD) $\rightarrow$ Total population estimates, $\mathrm{E}_T^C(x,b,g) = \sum_e \mathrm{E}^C(x,b,e,g)$.

Current Population Survey (CPS) $\rightarrow$ Proportion of people with education e, $\mathrm{R}^C(x,b,e,g) = \mathrm{E}^C/\mathrm{E}_T^C$.

Actuarial Research Centre
Institute and Faculty of Actuaries

Ratios of educated people in the CPS are very noisy due to sampling variation.



Education in exposures

Recurrence for the ratios:

$$R^C(x+1,b,e) = \frac{E^C(x+1,b,e)}{E_T^C(x+1,b)} = \frac{E^C(x,b,e) - \Delta(x,b,e)}{E_T^C(x,b) - \Delta_T(x,b)} =$$

$$= \frac{R^C(x,b,e)E_T^C(x,b) - \Delta(x,b,e)}{E_T^C(x,b) - \Delta_T(x,b)}.$$

$\Delta(x,b,e) \rightarrow$ Members lost by cohort born in b at age x with education e.

$\Delta_T(x,b) \rightarrow$ Total members lost by the cohort.

Education in exposures

Estimate initial ratio for each cohort, $R^C(x_0, b, e)$, by least squares:

$$\mathcal{O}_b = \sum_x \left(R^C(x, b, e) - \hat{R}^C(x, b, e) \right)^2$$

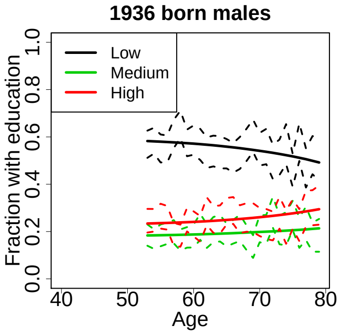
$R^C(x, b, e)$ can be written in terms of $R^C(x_0, b, e)$.
$\hat{R}^C(x, b, e)$ are given by the CPS.

We need to assume the only change in cohort membership Δ is the number of deaths.

Actuarial Research Centre
Institute and Faculty of Actuaries

Education in exposures

Back to the previous cohorts, smoothed ratios:

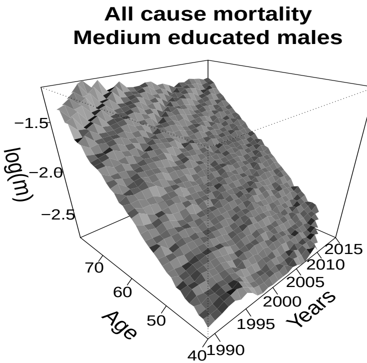


Really powerful method that allows for straightforward extrapolation to ages beyond 79!

Education in exposures

Cohort by cohort smoothing → "Wavy" death rates

Link cohorts by penalising concavity



Actuarial Research Centre
Institute and Faculty of Actuaries

Education in exposures

Concavity penalisation: we want the death rates to be relatively linear in the log scale within calendar years.

$$\mathcal{O}_N = \sum_b \mathcal{O}_b + \sum_x \left[\left(C(x, b+x, e) \right)^2 \right]$$

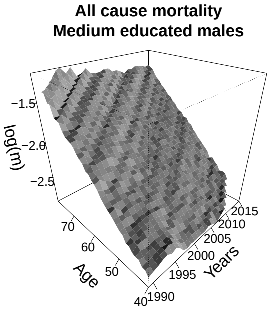
$$C(x,t,e) = \log(m(x,t,e)) - \frac{1}{2}\left(\log(m(x+1,t,e)) + \log(m(x-1,t,e)) \right)$$

Actuarial Research Centre
Institute and Faculty of Actuaries

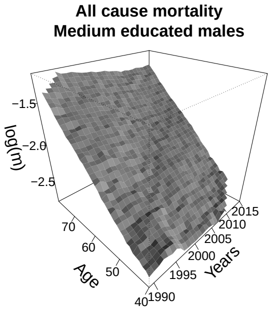
Effect of concavity correction on death rates

Link cohorts by penalising concavity:

No penalty $\Rightarrow$ With penalty

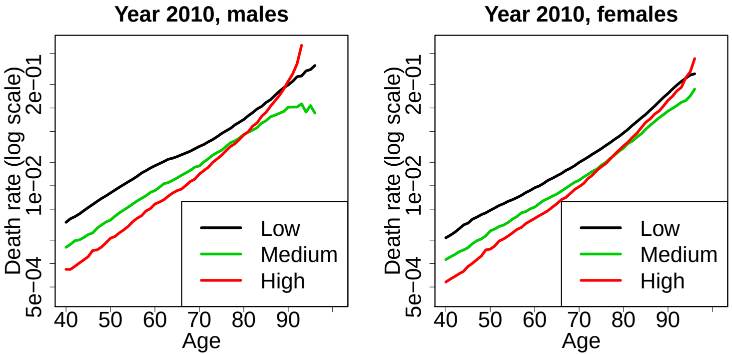


$\Rightarrow$



Education in deaths

All cause mortality



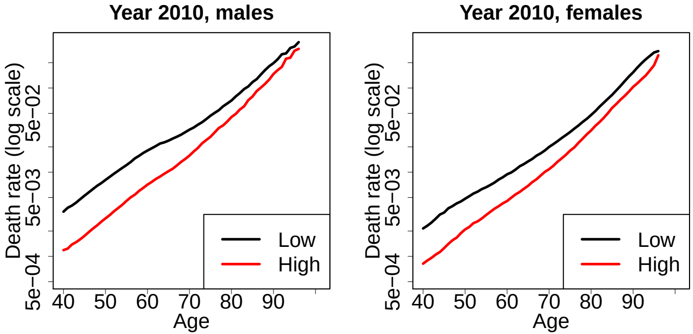
Overestimated death rates for high educated people!

ALL CAUSE MORTALITY

Actuarial Research Centre
Institute and Faculty of Actuaries

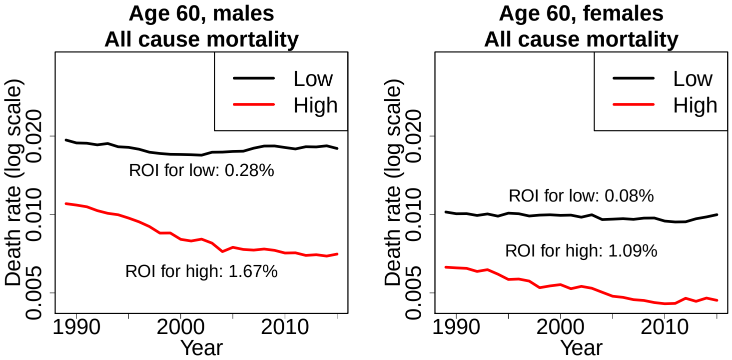
Education in deaths

Merging the medium and high educated groups in a single group solves the issues at high ages.



Initial/final gap males: 2.86 $\rightarrow$ 1.20
Initial/final gap females: 2.62 $\rightarrow$ 1.12

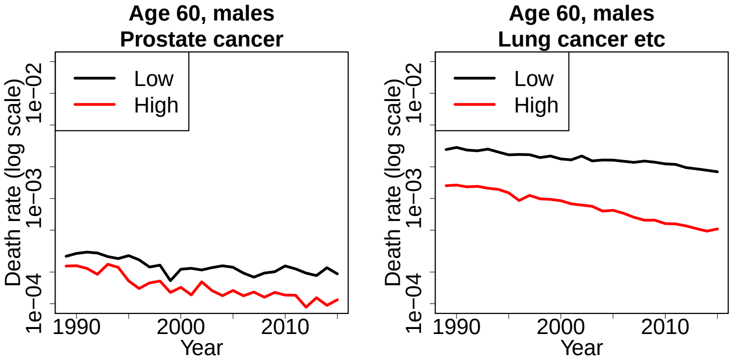
Death rates by education, age 60 by year



Initial/final gap males: 1.75 $\rightarrow$ 2.54
Initial/final gap females: 1.63 $\rightarrow$ 2.13

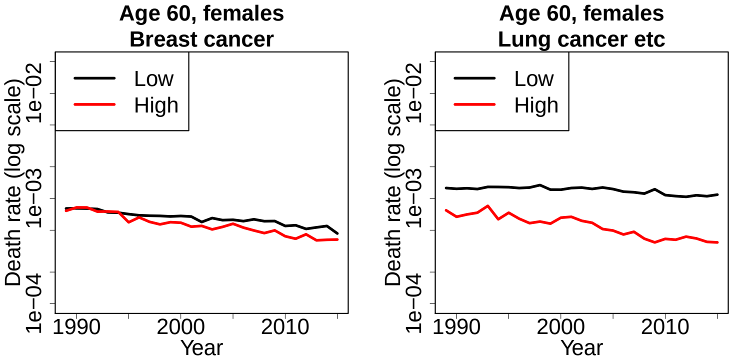
CAUSE OF DEATH MORTALITY

Death rates by education, age 60 by year



Initial→final ratio for prostate cancer: 1.24 → 1.76
Initial→final ratio for lung cancer: 2.21 → 3.49

Death rates by education, age 60 by year



Initial→final ratio for breast cancer: 1.05 → 1.14
Initial→final ratio for lung cancer: 1.63 → 2.84

STOCHASTIC MODELLING

Modelling

We will fit a stochastic model for the mortality rates independently to each gender, education group, and group of related causes of death, for ages 50-75 and years 1989-2015.

Our goal is to identify distinct trends in the cohort effects for different causes of death that point at both the drivers behind mortality for each cause and the changing behaviours of the population analysed.

Modelling

Model for the mortality rates $m(x, t)$:

$$D(x, t) \sim \text{Poisson}(m(x, t)E(x, t)),$$

where

$$\log(m(x, t)) = \alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + \gamma_{t-x}$$

Plat model restricted to high ages.

$\gamma_{t-x} \equiv \gamma_c$ cohort effect, shows how year of birth affects mortality. Gives us information about health behaviours and their evolution.

Modelling

Two approaches:

- Estimate the α's, κ's, and γ's from maximum likelihood, bootstrap to obtain CIs.
- Assume a time series structure for some of the parameters and use Bayesian techniques for the estimation.

For the second approach we take:

$$\gamma_{c+1} = \gamma_c + \mathcal{N}\left(0, \sigma_c^2\right),$$
$$\boldsymbol{\kappa_{t+1}} = \boldsymbol{\kappa_t} + \mathcal{N}\left(\boldsymbol{\mu}, \Sigma\right).$$

Actuarial Research Centre
Institute and Faculty of Actuaries

Bayesian approach

The prior distributions are:

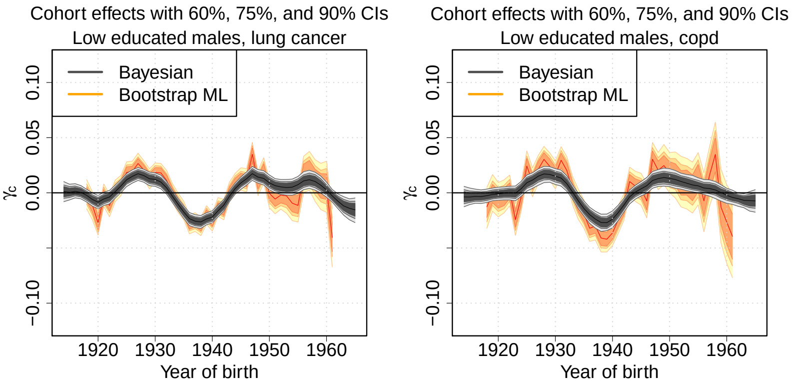
$$\sigma_c^2 \sim \textit{Inv-Gamma}(a, b)$$
$$\Sigma \sim \textit{Inv-Wishart}(\nu, \mathbf{S})$$
$$\mu_1, \mu_2 \sim \mathcal{N}(0, \sigma^2_{\mu_{1,2}})$$

So our log-posterior is:

$$\log\left(P(\alpha_x, \kappa_t^{(1)}, \kappa_t^{(2)}, \gamma_c, \sigma_c, \Sigma, \mu_1, \mu_2 | D, E)\right) \propto$$
$$\ell_P + \ell_{rw\gamma} + \ell_{rw\kappa} +$$
$$\log(\pi(\sigma_c)) + \log(\pi(\Sigma)) + \log(\pi(\mu_1)) + \log(\pi(\mu_2))$$

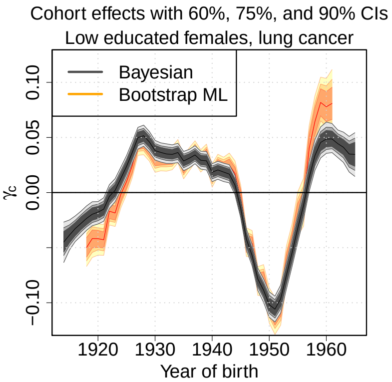
Actuarial Research Centre
Institute and Faculty of Actuaries

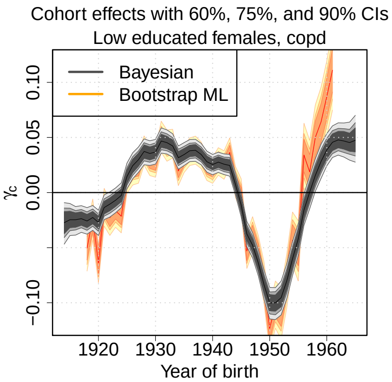
Cohort effects



Actuarial Research Centre
Institute and Faculty of Actuaries

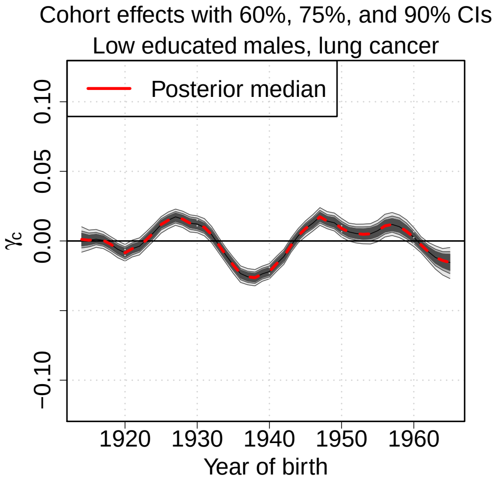
Cohort effects



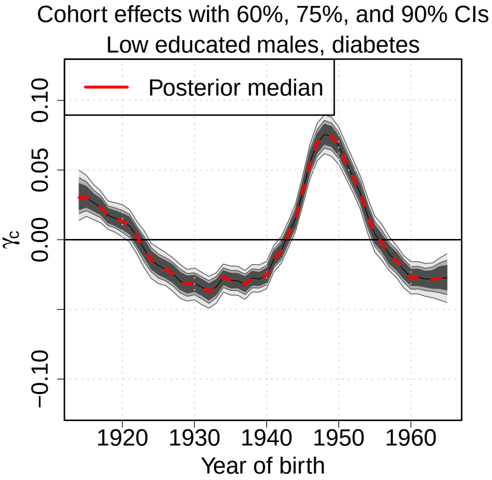


Actuarial Research Centre
Institute and Faculty of Actuaries

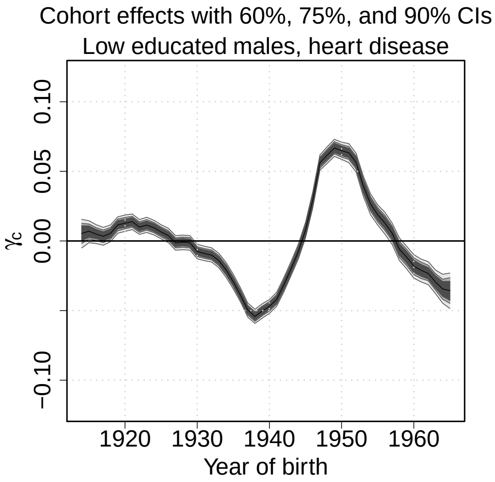
Cohort effects



Cohort effects

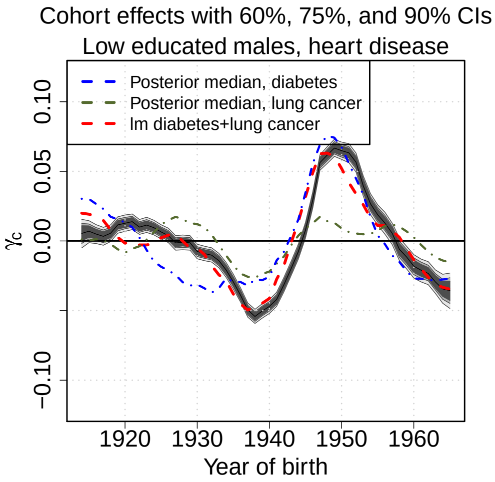


Cohort effects

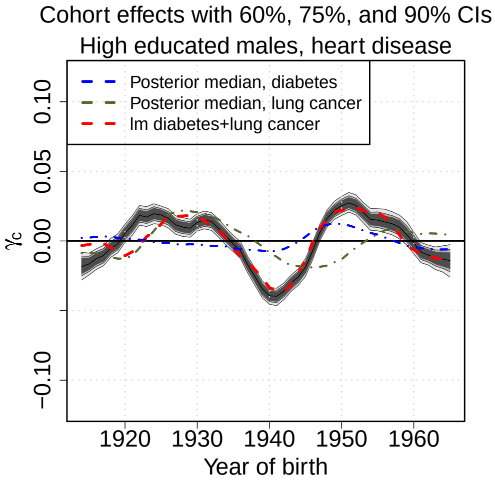


Actuarial Research Centre
Institute and Faculty of Actuaries

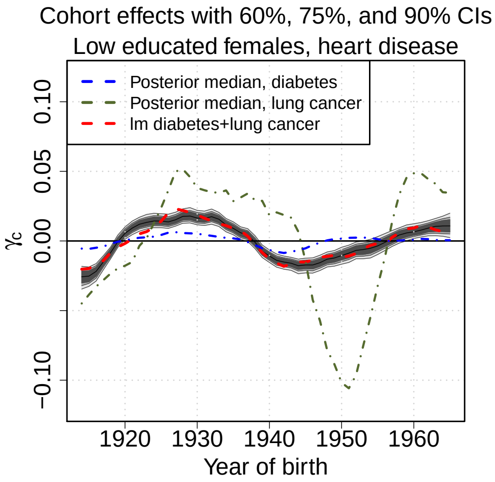
Cohort effects



Cohort effects

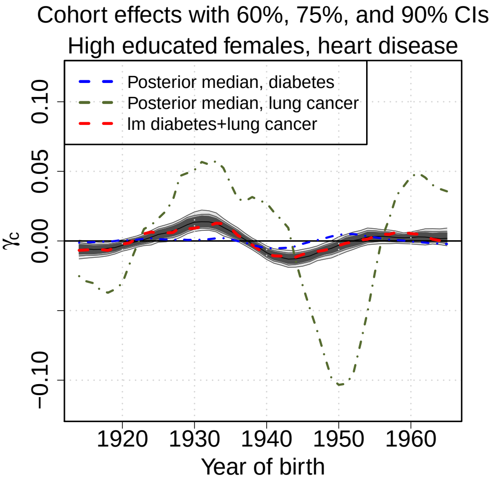


Cohort effects



Actuarial Research Centre
Institute and Faculty of Actuaries

Cohort effects



Conclusions

- From crude rates alone it is clear that educational attainment strongly influences mortality, especially at younger ages.
- Mortality gaps have increased over the last decades.
- The pattern of cohort effects for different causes of death gives us a hint to what lifestyle factors/health behaviours drive overall mortality.
- Males/females and high/low educated people show different cohort effect patterns.

Actuarial Research Centre
Institute and Faculty of Actuaries

Future work

There are several parts of this work that can be extended in the future:

- Systematise the cohort effect analysis.
- Improve the time series part of the model.
- Mortality projections.
- Applications to annuities and life insurance.
- Extension of the analysis to other populations.

Thank You!

Questions?

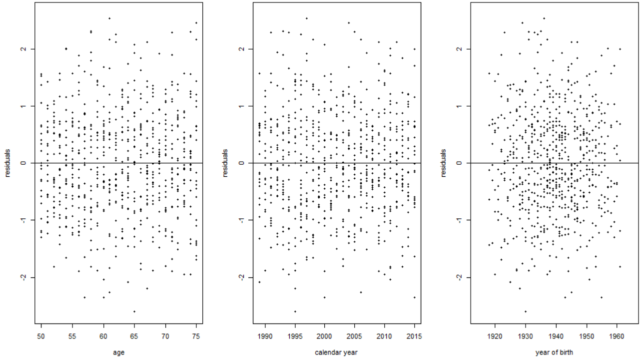
Identifiability constraints for the Plat model:

$$\sum_t \kappa_t^{(1)} = 0, \quad \sum_t \kappa_t^{(2)} = 0, \quad \sum_c \gamma_c = 0,$$

$$\sum_c c\gamma_c = 0, \quad \sum_c c^2\gamma_c = 0.$$

Actuarial Research Centre
Institute and Faculty of Actuaries

Residuals, all cause mortality, low educated males



Actuarial Research Centre
Institute and Faculty of Actuaries

Poisson log-likelihood (both approaches):

$$\ell_P \propto \sum_{x,t} \Big[-E_{xt} \exp(\alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + \gamma_c) + D_{xt}(\alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + \gamma_c) \Big]$$

Parameters easily estimated using R (StMoMo library). Bootstrapping techniques help obtain confidence intervals.

Actuarial Research Centre
Institute and Faculty of Actuaries

Random walk log-likelihood for γ_c:

$$\ell_{rw\gamma} \propto -\frac{1}{2\sigma_c^2}\sum_{c=2}^{n_c}(\gamma_c - \gamma_{c-1})^2 - \frac{n_c - 1}{2}\log\left(\sigma^2\right)$$

Random walk log-likelihood for $\boldsymbol{\kappa}_t$:

$$\ell_{rw\kappa} \propto -\frac{1}{2}\sum_{t=2}^{n_t}\left[(\boldsymbol{\kappa}_d)^T\Sigma^{-1}\boldsymbol{\kappa}_d\right] - \frac{n_t - 1}{2}\log\left(|\Sigma|\right)$$

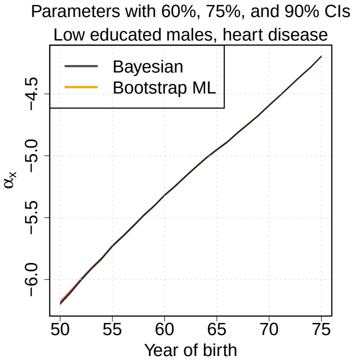
$$\boldsymbol{\kappa}_d = \boldsymbol{\kappa}_t - (\boldsymbol{\kappa}_{t-1} + \boldsymbol{\mu})$$

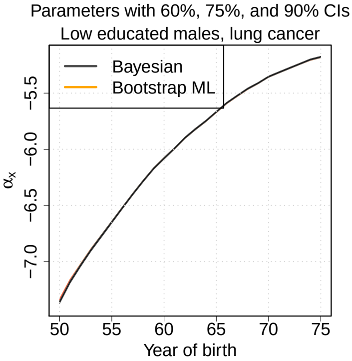
Actuarial Research Centre
Institute and Faculty of Actuaries

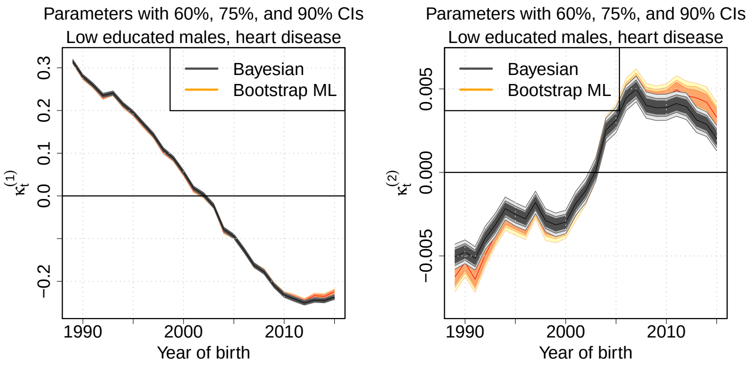
Priors:

$$\sigma_c \sim \textit{Inv-Gamma}(25, 10^{-4})$$
$$\Sigma \sim \textit{Inv-Wishart}\left(10, \begin{pmatrix} 10^{-4} & 0 \\ 0 & 10^{-4} \end{pmatrix}\right)$$
$$\mu_1 \sim \mathcal{N}(0, 0.1^2)$$
$$\mu_2 \sim \mathcal{N}(0, 0.1^2)$$

Actuarial Research Centre
Institute and Faculty of Actuaries


Parameters with 60%, 75%, and 90% CIs
Low educated males, heart disease
Bayesian
Bootstrap ML
−4.5
−5.0
−5.5
−6.0
αx
50
55
60
65
70
75
Year of birth


Parameters with 60%, 75%, and 90% CIs
Low educated males, lung cancer
Bayesian
Bootstrap ML
−5.5
−6.0
−6.5
−7.0
αx
50
55
60
65
70
75
Year of birth



Actuarial Research Centre
Institute and Faculty of Actuaries

List of causes of death:

1-Lung cancer
2-Lifestyle cancers
3-Prostate or breast cancer
4-Other cancers
5-COPD
6-Diabetes
7-Heart disease
8-Cerebrovascular disease
9-Other circulatory disease
10-Dementia and other mental illness
11-Accidental
12-Suicide, poisoning, and cirrhosis
13-Other

Actuarial Research Centre
Institute and Faculty of Actuaries