Modelling Socio-Economic Differences in the Mortality of Danish Males Using a New Affluence Index

Andrew J.G. Cairns

Heriot-Watt University, Edinburgh

Director, Actuarial Research Centre, IFoA

Joint work with D. Blake, K. Dowd, M. Kallestrup-Lamb, C. Rosenskjold

LoLitA Closing Conference, Paris, January 2018

Outline

- Introduction and motivation for multi-population modelling
- Constructing a new Danish dataset
- Modelling Danish sub-population mortality
- Forecast correlations and population basis risk
- Concluding remarks

1. Motivation: Multi-Population Stochastic Mortality

- Focus: pension plan or annuity provider
- Concerned about *longevity risk*:
  the risk that in aggregate people live longer than anticipated.
  $\Rightarrow$ *insufficient cash* to pay promised pensions
- Risk measurement $\longrightarrow$ risk management

Actuarial Research Centre
Institute and Faculty of Actuaries

Multi-Population Questions and Challenges

Why model multiple populations?

- Different socio-economic groups
- Males and females
- Different countries
- etc.

We seek to understand and model the differences:

- How big are the differences?
- Are the differences changing over time?

Multi-Population Questions and Challenges (cont.)

Challenges:

- Data availability
- Data quality and depth
- Multi-population modelling requires
  - (fairly) simple single-population models
  - simple dependencies between populations
  - BUT also plausible correlation structure between groups, ages and through time

Multi-Population Questions and Challenges (cont.)

Applications:

- Risk management e.g.
  - Pension plan: reduce exposure to longevity risk
  - (Re)insurer: reduce exposure and concentration
    $\Rightarrow$ release economic/regulatory capital
    $\Rightarrow$ increased capacity
- Compare risk management options
  - Customised versus index-based hedges
  - Base table; central forecast; risk
  - Index-based $\Rightarrow$ population basis risk

Actuarial Research Centre
Institute and Faculty of Actuaries

2. A New Case Study and a New Model

- Sub-populations differ from national population
  - socio-economic factors
  - other factors
- Denmark: National Register Database
  - High quality data on ALL residents
  - 1981-2012 available
  - Can subdivide population using covariates on the database

Danish Data

- *What can we learn from Danish data that will inform us about other populations?*

- Key covariates (amongst others):
  - Wealth
  - Income

Actuarial Research Centre
Institute and Faculty of Actuaries

Problem

- High income $\Rightarrow$ "affluent" *and low mortality*
  BUT
- Low income $\not\Rightarrow$ not affluent, high mortality
- High wealth $\Rightarrow$ "affluent" *and low mortality*
  BUT
- Low wealth $\not\Rightarrow$ not affluent, high mortality

Empirical solution: use a combination

- Affluence, $A =$ wealth $+K\times$ income
- $K = 15$ seems to work well *statistically* as a predictor
- Low affluence, A, predicts poor mortality

Actuarial Research Centre
Institute and Faculty of Actuaries

Subdividing Data (after much experimentation!)

- Males resident in Denmark for the previous 12 months
- Divide population in year t
  - into 10 equal sized Groups (approx)
  - using *affluence*, A
- Individuals can change groups up to age 67
- Group allocations are locked down at age 67

  (better than not locking down at age 67)

Subdivided Data

- Ages 55-94;   Years 1985-2012
- Exposures $E^{(i)}(t, x)$ for groups $i = 1, \ldots, 10$

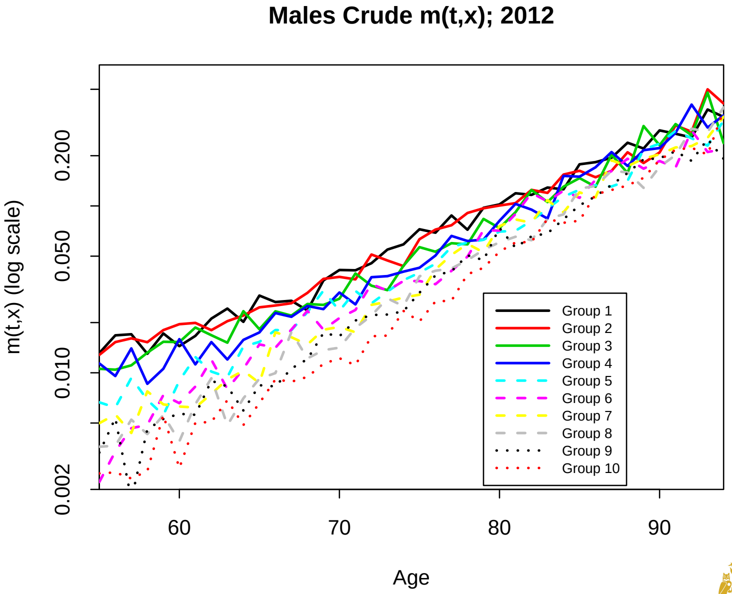
  range from over 4250 down to 13
- Deaths $D^{(i)}(t, x)$

  range from 151 down to 4
- Crude age-specific death rates

  $\hat{m}^{(i)}(t, x) = D^{(i)}(t, x)/E^{(i)}(t, x)$

  Population i, Year t, Age x
- Small groups $\Rightarrow$ Poisson risk is important

Crude death rates 2012



Actuarial Research Centre
Institute and Faculty of Actuaries

Modelling the underlying death rates, $m^{(k)}(t,x)$

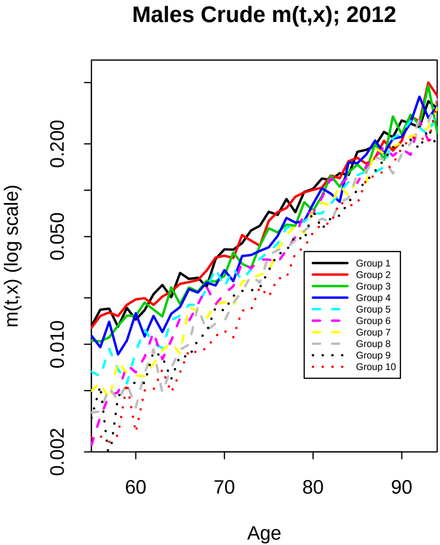
$m^{(k)}(t,x)$ = pop. k death rate in year t at age x
Population k, year t, age x

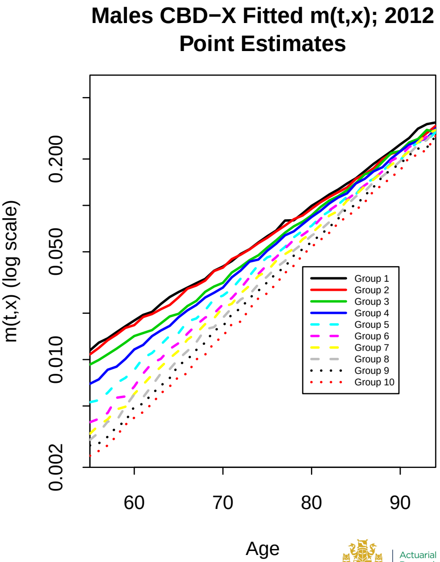
$$\log m^{(k)}(t,x) = \beta^{(k)}(x) + \kappa_1^{(k)}(t) + \kappa_2^{(k)}(t)(x-\bar{x})$$

(Extended CBD with a non-parametric base table, $\beta^{(k)}(x)$)

- 10 groups, $k = 1,\ldots,10$ (low to high affluence)
- 28 years, $t = 1985,\ldots,2012$
- 40 ages, $x = 55,\ldots,94$

Model-Inferred Underlying Death Rates 2012





Actuarial Research Centre
Institute and Faculty of Actuaries

Modelling the death rates, $m_k(t, x)$

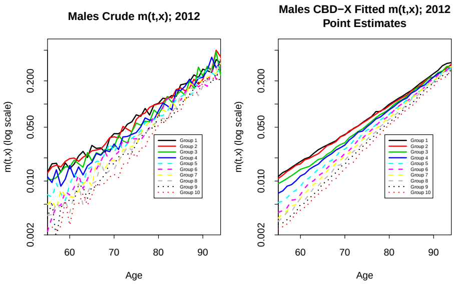
$$\log m^{(k)}(t,x) = \beta^{(k)}(x) + \kappa_1^{(k)}(t) + \kappa_2^{(k)}(t)(x - \bar{x})$$

- Model fits the 10 groups well without a cohort effect
- Non-parametric $\beta^{(k)}(x)$ is essential to preserve group rankings
  - Rankings are evident in crude data
  - *"Bio-demographic reasonableness"*:

more affluent $\Rightarrow$ healthier

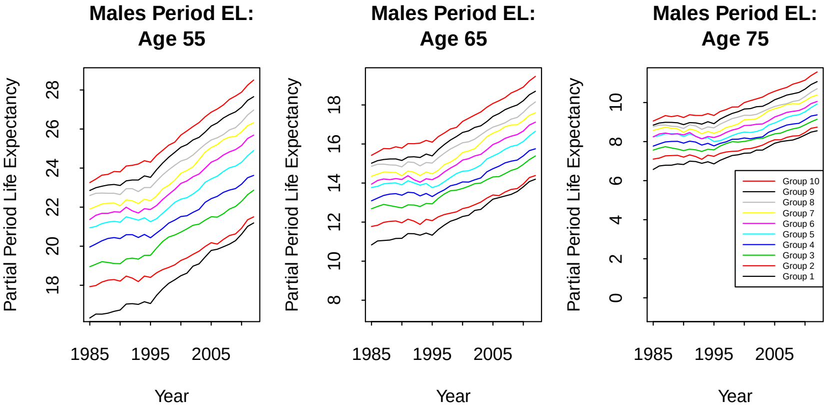
Actuarial Research Centre
Institute and Faculty of Actuaries

Model-Inferred Underlying Death Rates 2012



- Gap reduces from over $5\times$ to $1.3\times$
- Or $+14$ years difference for Group $1 \to 10$, age 55; $+9$ at 67.
- Convergence $\Rightarrow$ way ahead for modelling very high ages???

Partial Period Life Expectancy for Groups 1-10



Actuarial Research Centre
Institute and Faculty of Actuaries

Time series modelling

- $t \to t+1$: Allow for correlation
  - between $\kappa_1^{(k)}(t+1)$ and $\kappa_2^{(k)}(t+1)$
  - between groups $k = 1, \ldots, 10$

- Medium/long term, $t \to t+T$:
  group specific period effects *gravitate* towards the national trend (*coherence*)
  $\Rightarrow$ Bio-demographical reasonableness:
  groups should not diverge

Actuarial Research Centre
Institute and Faculty of Actuaries

A specific model

$$
\begin{aligned}
\kappa_1^{(i)}(t) &= \kappa_1^{(i)}(t-1) + \mu_1 + Z_{1i}(t) && \text{(random walk)} \\
&\quad -\psi\left(\kappa_1^{(i)}(t-1) - \bar{\kappa}_1(t-1)\right) && \text{(gravity between groups)} \\
\kappa_2^{(i)}(t) &= \kappa_2^{(i)}(t-1) + \mu_2 + Z_{2i}(t) \\
&\quad -\psi\left(\kappa_2^{(i)}(t-1) - \bar{\kappa}_2(t-1)\right)
\end{aligned}
$$

where

$$
\bar{\kappa}_1(t) = \frac{1}{n}\sum_{i=1}^{n} \kappa_1^{(i)}(t) \quad \text{and} \quad \bar{\kappa}_2(t) = \frac{1}{n}\sum_{i=1}^{n} \kappa_2^{(i)}(t)
$$

Actuarial Research Centre
Institute and Faculty of Actuaries

A specific model

$$\begin{aligned}
\kappa_1^{(i)}(t) &= \kappa_1^{(i)}(t-1) + \mu_1 + Z_{1i}(t) - \psi\left(\kappa_1^{(i)}(t-1) - \bar{\kappa}_1(t-1)\right) \\
\kappa_2^{(i)}(t) &= \kappa_2^{(i)}(t-1) + \mu_2 + Z_{2i}(t) - \psi\left(\kappa_2^{(i)}(t-1) - \bar{\kappa}_2(t-1)\right)
\end{aligned}$$

Model structure $\Rightarrow$

- $(\bar{\kappa}_1(t), \bar{\kappa}_2(t)) \sim$ bivariate random walk
- Each $\kappa_1^{(i)}(t) - \bar{\kappa}_1(t) \sim AR(1)$ reverting to 0
- Each $\kappa_2^{(i)}(t) - \bar{\kappa}_2(t) \sim AR(1)$ reverting to 0
- $\beta^{(i)}(x)$ vs $\beta^{(j)}(x) \Rightarrow$ intrinsic group differences

Actuarial Research Centre
Institute and Faculty of Actuaries

Non-trivial correlation structure: between different ages and groups

$$
\begin{aligned}
\kappa_1^{(i)}(t) &= \kappa_1^{(i)}(t-1) + \mu_1 + Z_{1i}(t) - \psi\left(\kappa_1^{(i)}(t-1) - \bar{\kappa}_1(t-1)\right) \\
\kappa_2^{(i)}(t) &= \kappa_2^{(i)}(t-1) + \mu_2 + Z_{2i}(t) - \psi\left(\kappa_2^{(i)}(t-1) - \bar{\kappa}_2(t-1)\right)
\end{aligned}
$$

The $Z_{ki}(t)$ are iid multivariate normal, mean 0 and

$$
Cov(Z_{ki}, Z_{lj}) = \begin{cases} v_{kl} & \text{for } i = j \\ \rho v_{kl} & \text{for } i \neq j \end{cases}
$$

$\rho =$ cond. correlation between $\kappa_1^{(i)}(t)$ and $\kappa_1^{(j)}(t)$ etc.

Actuarial Research Centre
Institute and Faculty of Actuaries

Comments

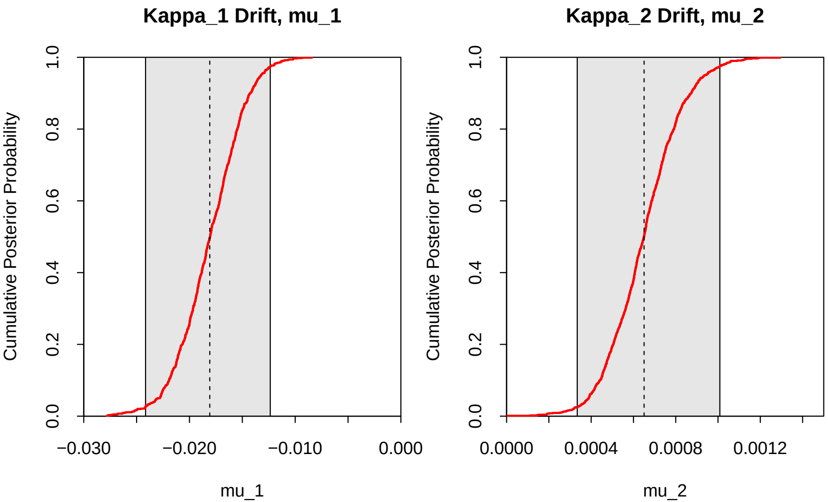
- Model is very simple
  - One gravity parameter, $0 < \psi < 1$
  - One between-group correlation parameter, $0 < \rho < 1$
- Many generalisations are possible
- But more parameters + more complex computing
- This simple model seems to fit quite well.
- Nevertheless $\Rightarrow$ potential for further work

Bayesian implementation: Prior distributions

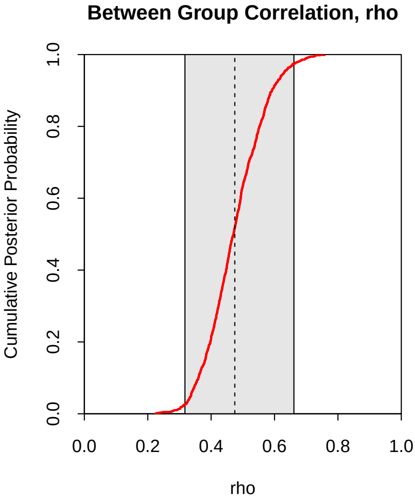
- As uninformative as possible
- $\mu_1,\ \mu_2 \sim$ improper uniform prior
- $\{v_{ij}\} \sim$ Inverse Wishart
- $\rho \sim \text{Beta}(2, 2)$
- $\psi \sim \text{Beta}(2, 2)$

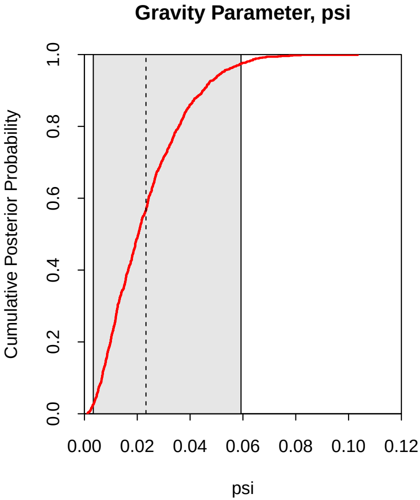
State variables and process parameters estimated using MCMC (Gibbs + Metropolis-Hastings)

Posterior Distributions and 95% Credibility Intervals



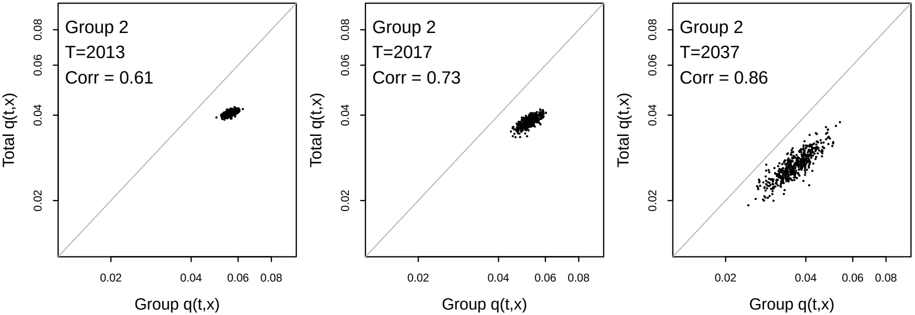
Posterior Distributions and 95% Credibility Intervals





Actuarial Research Centre
Institute and Faculty of Actuaries

Simulated Group versus Population Mortality, $q(t, x)$



As T increases: $+1$ year; $+5$ years; $+25$years

- Scatterplots become more dispersed
- Shift down and to the left
- Correlation increasess

Actuarial Research Centre
Institute and Faculty of Actuaries

Forecast Correlations

Deciles are quite narrow subgroups

More diversified e.g.

- Blue collar pension plan

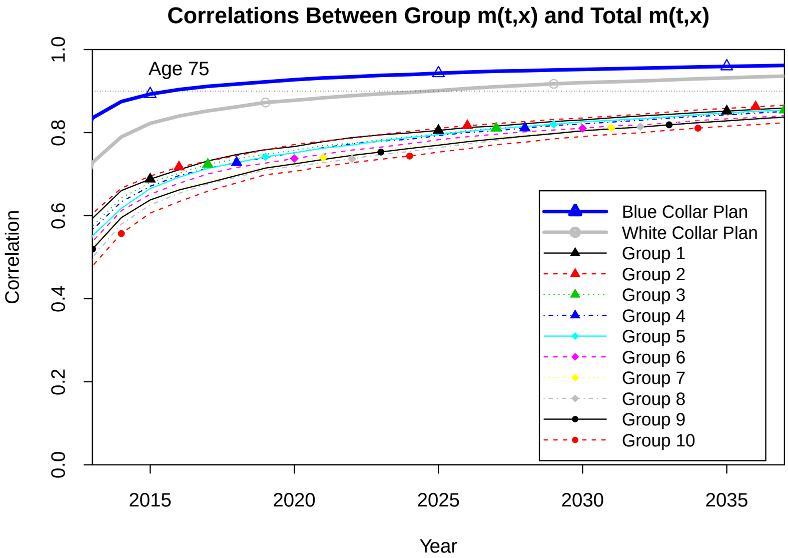
  $\Rightarrow$ equal proportions of groups 2, 3, 4
- White collar pension plan

  $\Rightarrow$ equal proportions of groups 8, 9, 10
- Mixed plan

  $\Rightarrow$ proportions $(0, 0, 1, 2, \ldots, 7, 8)/36$ (e.g. amounts)

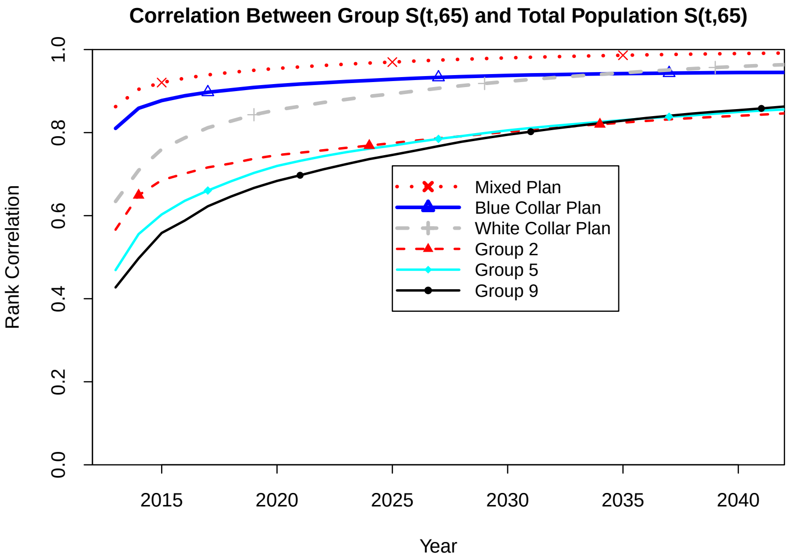
Actuarial Research Centre
Institute and Faculty of Actuaries

Forecast Correlations: Mortality Rates at Age 75



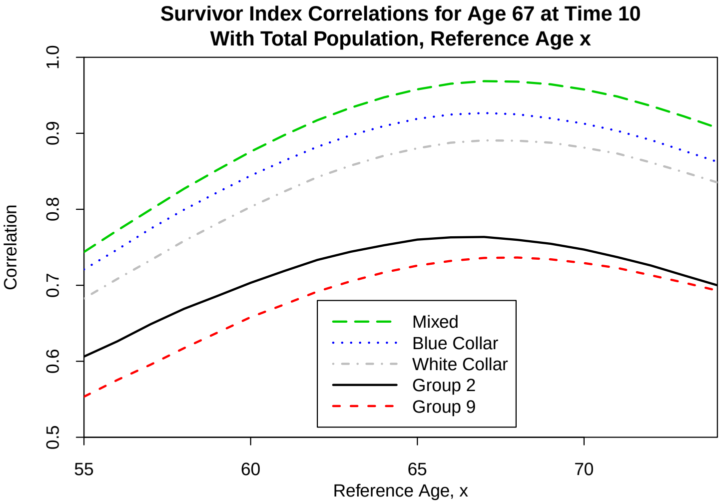
Actuarial Research Centre
Institute and Faculty of Actuaries

Forecast Correlations: Cohort Survivorship from 65



Forecast Correlations: Cohort Survivorship

Different reference ages: $cor\left(S_X(10, 67),\ S_{TOT}(10, x)\right)$



Modelling Conclusions

- Development of a new multi-population dataset for Denmark

  strong bio-demographically reasonable group rankings based on a new measure of affluence
- Unlike multi-country data

  a priori ranking of affluence-related groups
- Proposal for a simple new multi-population model
- Mortality rates converge at high ages
- Strong correlations over medium to long term

  even allowing for parameter uncertainty
- Correlations depend strongly on diversity of sub-population

Actuarial Research Centre
Institute and Faculty of Actuaries

3. Postscript 1: Education as an Alternative Covariate

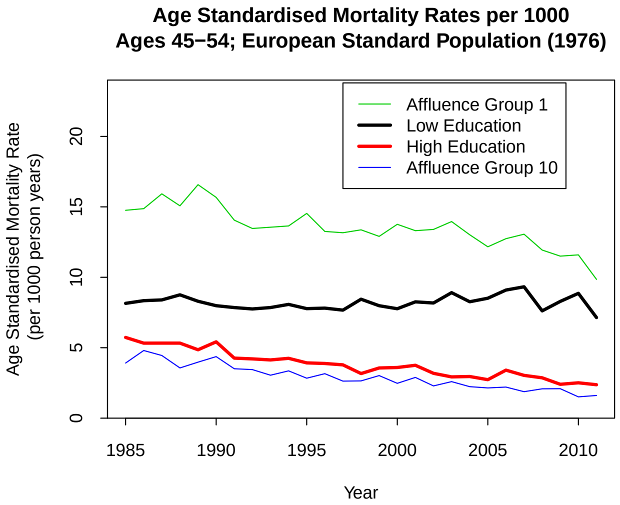
- Level of Educational Attainment also known to be a good predictor
  - Various US studies
  - Mackenbach et al. (2003) including Denmark: Std. Mortality Rates
  - Brønnum-Hansen and Baadsgaard (2012) Denmark: $LE(x = 30)$
- As close as possible on a *like for like* basis:

  Crude death rates; age 30+; matching years.

  Affluence $\Rightarrow$
  - Wider spread of SMR's than M. et al. (2003)
  - Wider spread of $LE(30)$ than BHB (2012)
- Issue: "grade inflation" distorts results
- More to be done.

Actuarial Research Centre
Institute and Faculty of Actuaries

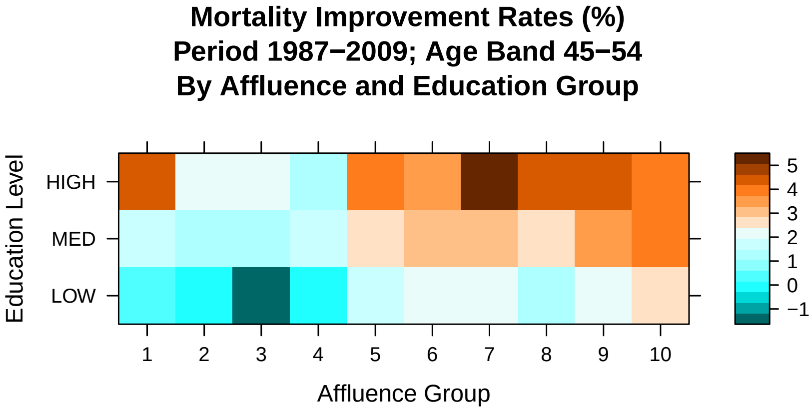
3. Postscript: Education as an Alternative Covariate



3. Postscript: Education as an Alternative Covariate

Dig a bit deeper:

Affluence + Education: average ASMR's over 5 years



4. Summary

- Danish data allows insight into relative mortality dynamics between socio-economic sub-populations
- Applications
    - Improved risk measurement
    - Assess effectiveness of longevity hedges
- Education as an alternative covariate
- Health inequalities
- Work in progress:
    - Canada, USA, England
    - Hedging: Regulatory and economic capital relief

Working paper available on website.

E: A.J.G.Cairns@hw.ac.uk W: www.macs.hw.ac.uk/∼andrewc

Actuarial Research Centre
Institute and Faculty of Actuaries

Thank You!

Questions?

E: A.J.G.Cairns@hw.ac.uk W: www.macs.hw.ac.uk/~andrewc

HERIOT WATT UNIVERSITY

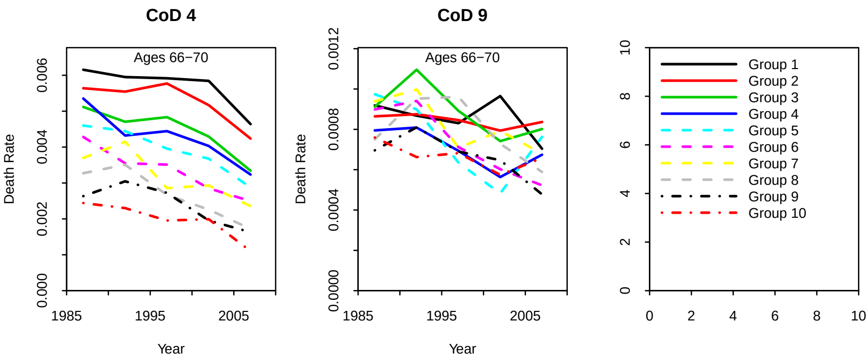
Actuarial Research Centre
Institute and Faculty of Actuaries

Postscript 2: Cause of Death Data – Health Inequalities

- Deaths subdivided into 29 CoD groups
- Compare affluence groups
- Biggest differences at younger age groups e.g. 51-55
- Causes of death linked to lifestyle
  $\Rightarrow$ some CoD death rates are up to 20$\times$ higher for low affluence groups
- Growing gaps: liver diseases; diabetes
- *Almost all CoD groups have a strong statistically significant difference*

Actuarial Research Centre
Institute and Faculty of Actuaries

Denmark: Cause of Death Data – Health Inequalities



- 5 × 5 ages and years
- CoD4: Lung cancer and related cancers
- CoD9: Cancer of lymphatic or blood-forming tissues

Denmark: Cause of Death Data – Health Inequalities

- Some causes of death have no obvious link to lifestyle/affluence/education
- Possible explanations (a very non-expert view)
  - *onset* is not dependent on lifestyle/affluence/education
  - BUT less affluent/educated $\Rightarrow$

    ??? later diagnosis

    ??? engage less well with treatment process

    ??? lower quality housing

Actuarial Research Centre
Institute and Faculty of Actuaries