MULTI-POPULATION MORTALITY MODELLING

Andrew Cairns

Heriot-Watt University, Scotland

and

The Maxwell Institute, Edinburgh

Joint work with

David Blake, Kevin Dowd, Malene Kallestrup-Lamb and Carsten Rosenskjold

Plan

- Motivation and challenges
- Danish males data

  10 sub-populations grouped by wealth

- An extended CBD multi-population model
- Bayesian implementation and results

Motivation for stochastic mortality modelling

- Life expectancy is increasing/mortality falling

  $\Rightarrow$potential impact on

  - pension plan finances; costs $\nearrow$
  - life insurance premiums and reserves

- Past patterns $\Rightarrow$ future improvements uncertain

- Need good stochastic models for

  - central forecasts
  - assessment of uncertainty around central trend
  - development of risk management strategies

Motivation for multi-population modelling

A: Risk assessment

- Multi-country (e.g. consistent demographic projections)
- Males/Females (e.g. consistent demographic projections)
- Socio-economic subgroups (e.g. blue or white collar)
- Smokers/Non-smokers
- Annuities/Life insurance
- Limited data $\Rightarrow$ learn from other populations

Motivation for two-population modelling

B: Risk management for pension plans and insurers

- Retain systematic mortality risk; versus:
- 'Over-the-counter' deals (e.g. longevity swap)
  - own experience $\Rightarrow$ 100% risk reduction
  - potentially expensive
- Standardised mortality-linked securities
  - linked to national mortality index
  - $< 100\%$ risk reduction
  - less expensive
  - potential secondary market

Two or more populations

- Linked in some way
- But not identical
- Desire for consistent forecasts
  - distributions
  - individual future scenarios

Key hypothesis

- $m^{(k)}(t,x) =$ pop. k death rate in year t at age x
- Hypothesis (e.g. Li and Lee, 2005):

  For each age x, and for any two populations j and k

  $$\frac{m^{(j)}(t,x)}{m^{(k)}(t,x)} \text{ does not diverge over time}$$

- Hypothesis $\Rightarrow$ Consequences depend on your choice of stochastic mortality model

Challenges

- Data availability
- Data quality and depth
- Model complexity
  - single population models can be complex
  - 2-population versions are more complex
  - multi-pop ......
- Multi-population modelling requires
  - (fairly) simple single-population models
  - simple dependencies between populations

A New Case Study and a New Model

- Sub-populations differ from national population
  - socio-economic factors
  - geographical variation
  - other factors
- Denmark
  - High quality data on ALL residents
  - 1981-2005 available
  - Can subdivide population using covariates on the database

Danish Data

- Key covariates
  - Net assets
  - Net income

Problem

- High income $\Rightarrow$ "wealthy" *and healthy* BUT
- Low income $\not\Rightarrow$ not wealthy, poor health
- High assets $\Rightarrow$ "wealthy" *and healthy* BUT
- Low assets $\not\Rightarrow$ not wealthy, poor health

Solution: use a combination

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

Subdividing Data

- Males resident in Denmark for the previous 12 months
- Divide population in year t
  - into 10 equal sized Groups (approx)
  - using wealth in year $t - 1$
- Individuals can change groups up to age 67
- Group is locked down at age 67

(better than not locking down at age 67)

Subdivided Data

- Exposures $E^{(i)}(t,x)$ for groups $i = 1, \ldots, 10$

  range from over 4000 down to 20

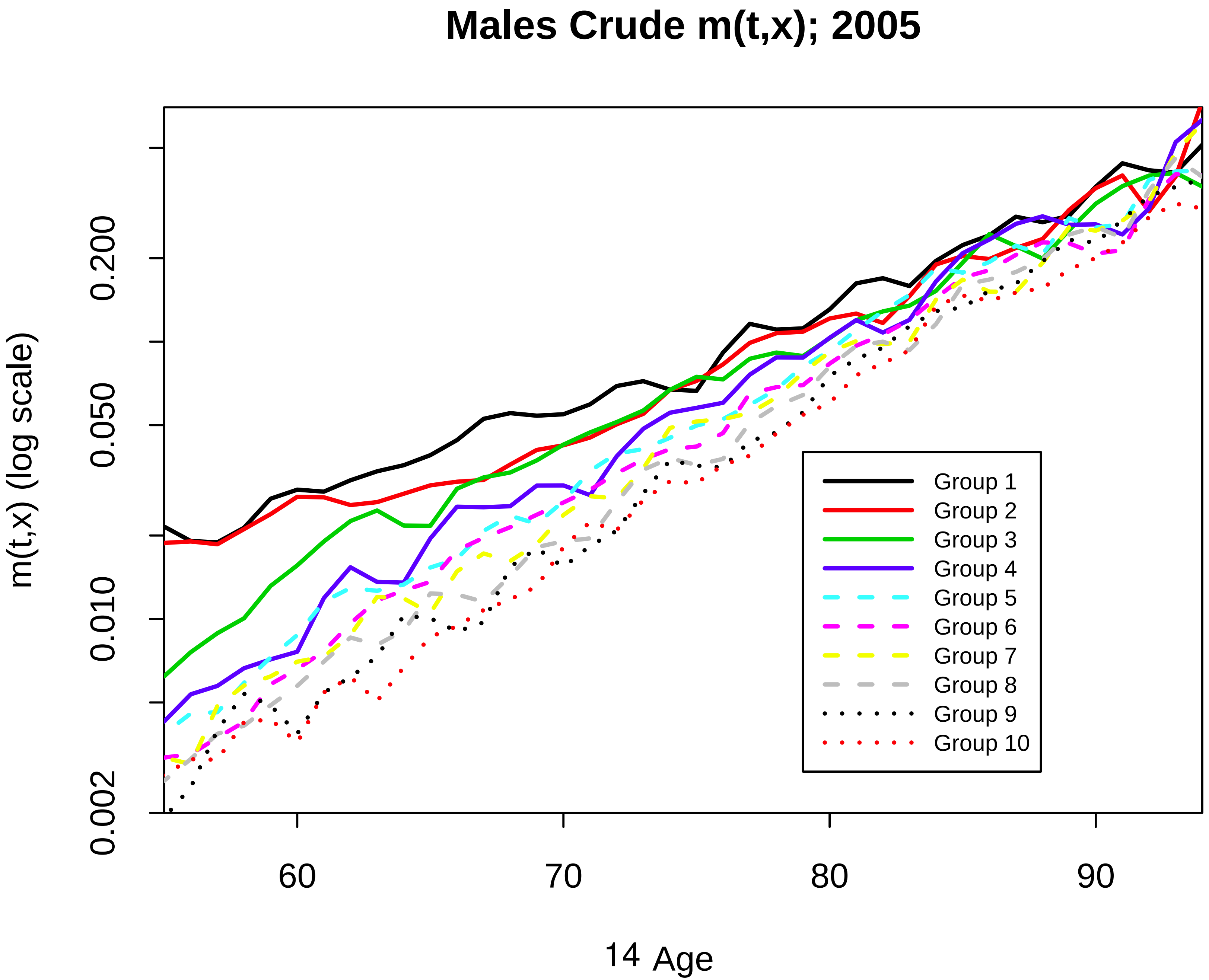
- Deaths $D^{(i)}(t,x)$

  range from 150 down to 6

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

- Small groups $\Rightarrow$ Poisson risk is important

Crude death rates 2005



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

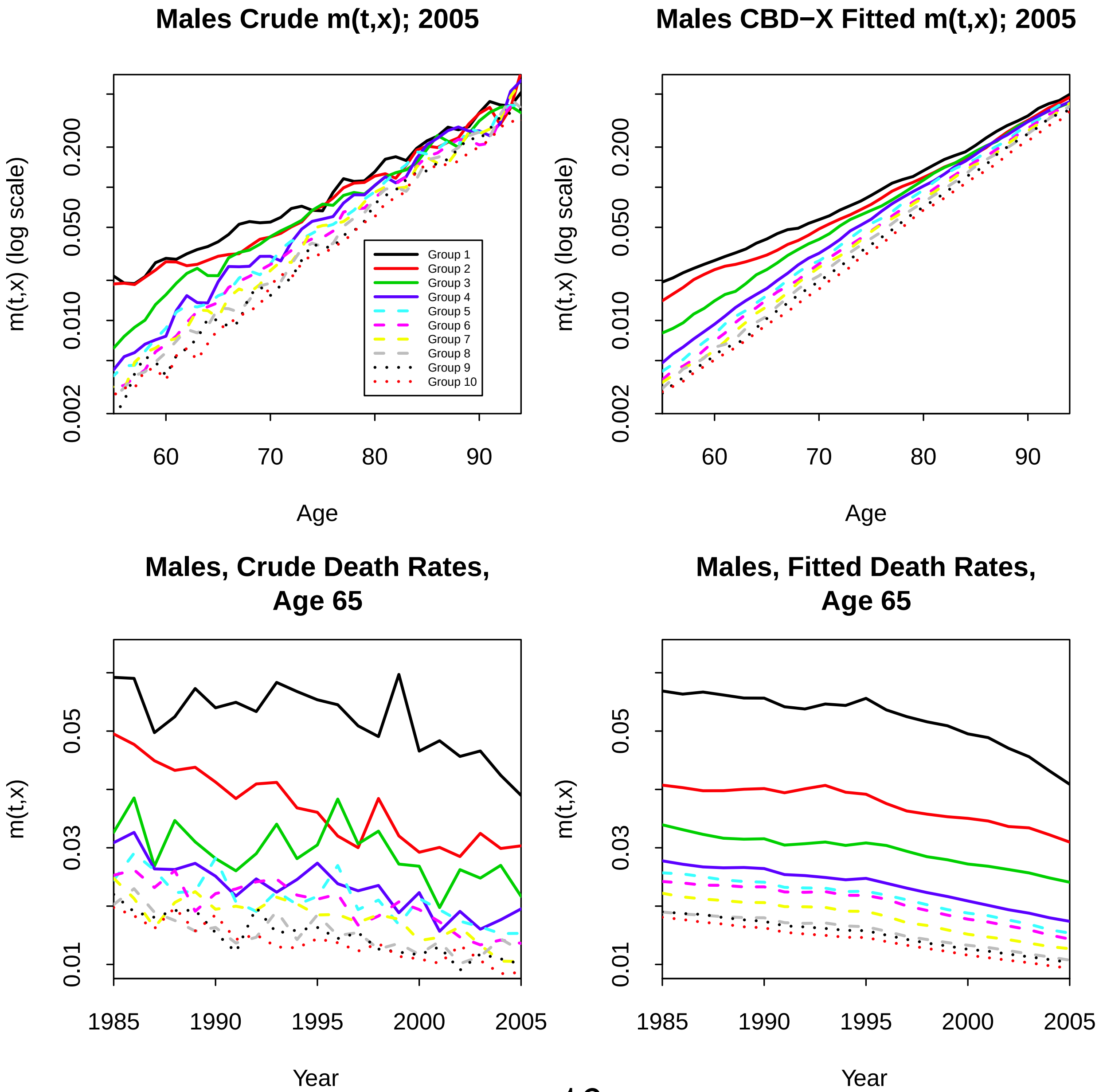
Population k, year t, age x

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

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

- 10 groups, $k = 1, \ldots, 10$ (low to high wealth)
- 21 years, $t = 1985, \ldots, 2005$
- 40 ages, $x = 55, \ldots, 94$

Model-Inferred Underlying Death Rates 2005



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
  - *"Biological reasonableness"*: wealthier $\Rightarrow$ healthier

Bayesian modelling

- Combines
  - conditional Poisson likelihood
  - time series likelihood for the $\kappa_j^{(k)}(t)$
  - (uninformative) prior distributions for process parameters
- Output *posterior distribution* for
  - $\beta^{(k)}(x),\ \kappa_1^{(k)}(t),\ \kappa_2^{(k)}(t)$ latent state variables
  - time series process parameters

Time series modelling

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

  groups should not diverge
- Sufficient that we have mean reversion in $\kappa_1^{(j)}(t) - \kappa_1^{(k)}(t)$ and $\kappa_2^{(j)}(t) - \kappa_2^{(k)}(t)$

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

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

A specific model

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

- $(\bar{\kappa}_1(t), \bar{\kappa}_2(t)) \sim$ 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

A specific model

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

The $Z_{i,j}$ are 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.

Comments

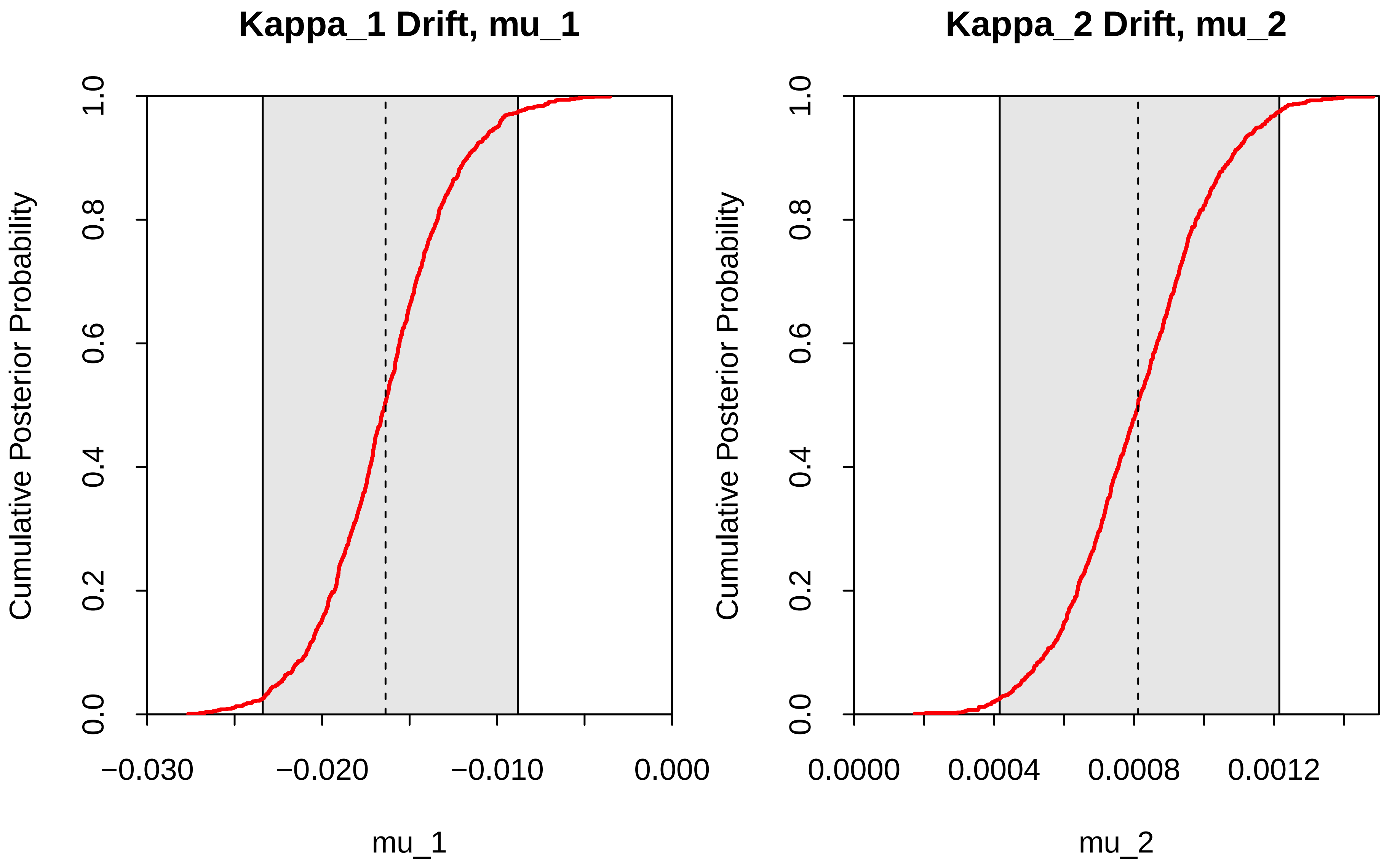
- Model is very simple
  - One gravity parameter, $0 < \psi < 1$
  - One between-group correlation parameter, $0 < \rho < 1$
- Many generalisations are possible
- But more parameters $+$ more complex computing
- This simple model seems to fit quite well.
- Nevertheless $\Rightarrow$ work in progress

Prior distributions

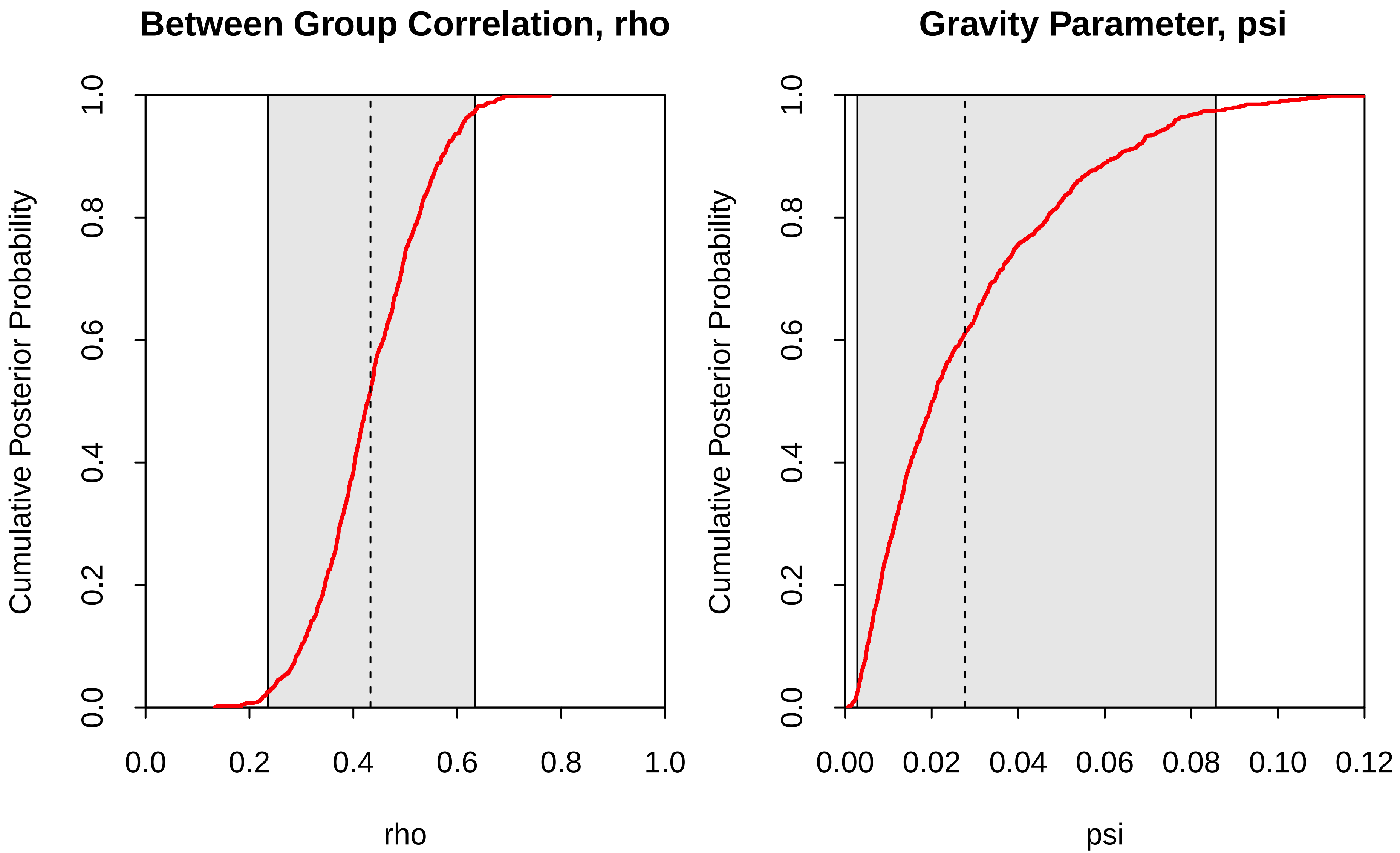
- As uninformative as possible
- $\mu_1, \ \mu_2 \sim$ improper uniform prior
- $\{v_{ij}\} \sim$ Inverse Wishart
- $\rho \sim$ Beta$(2, 2)$
- $\psi \sim$ Beta$(2, 2)$

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

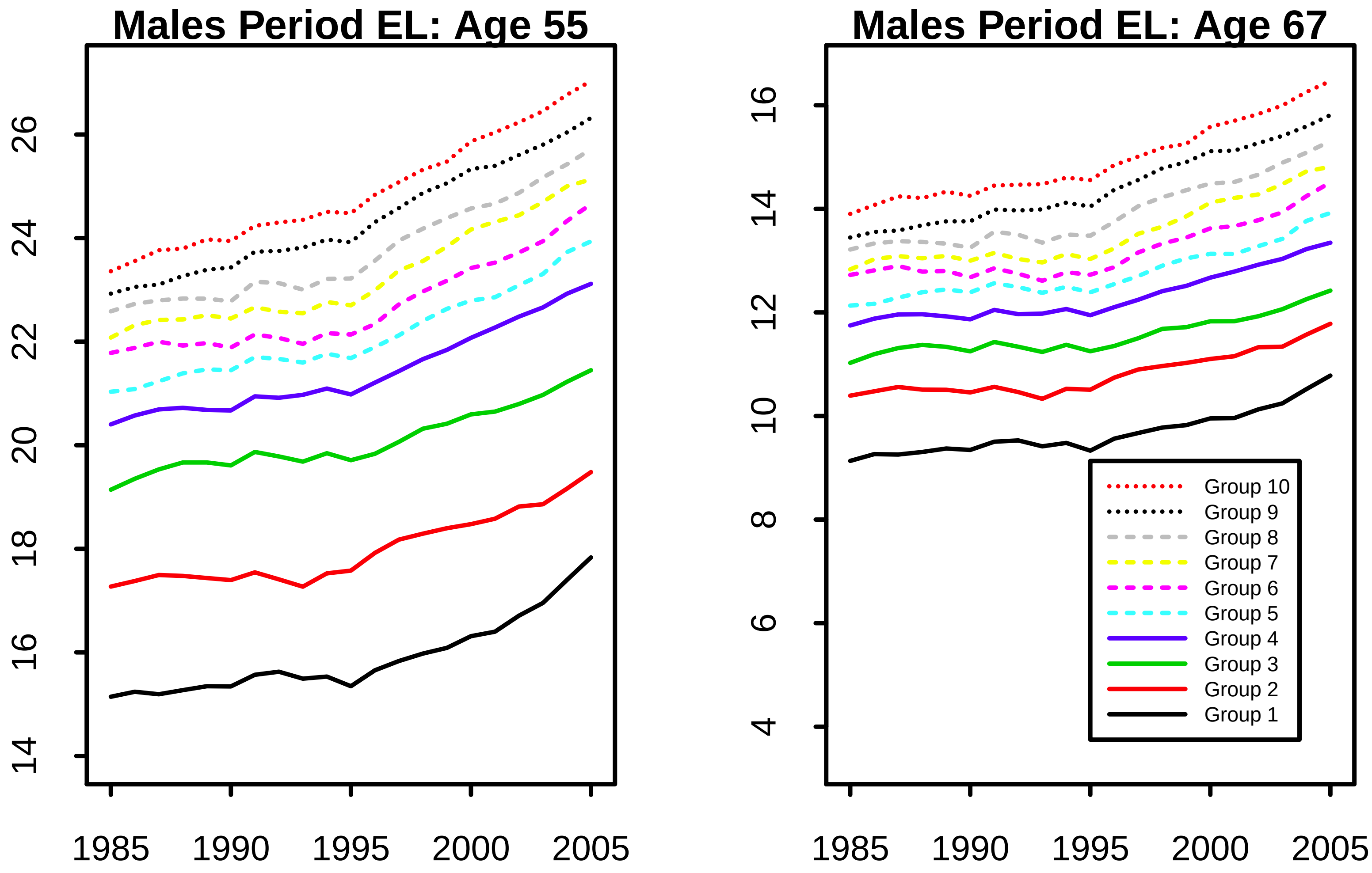
Posterior Distributions



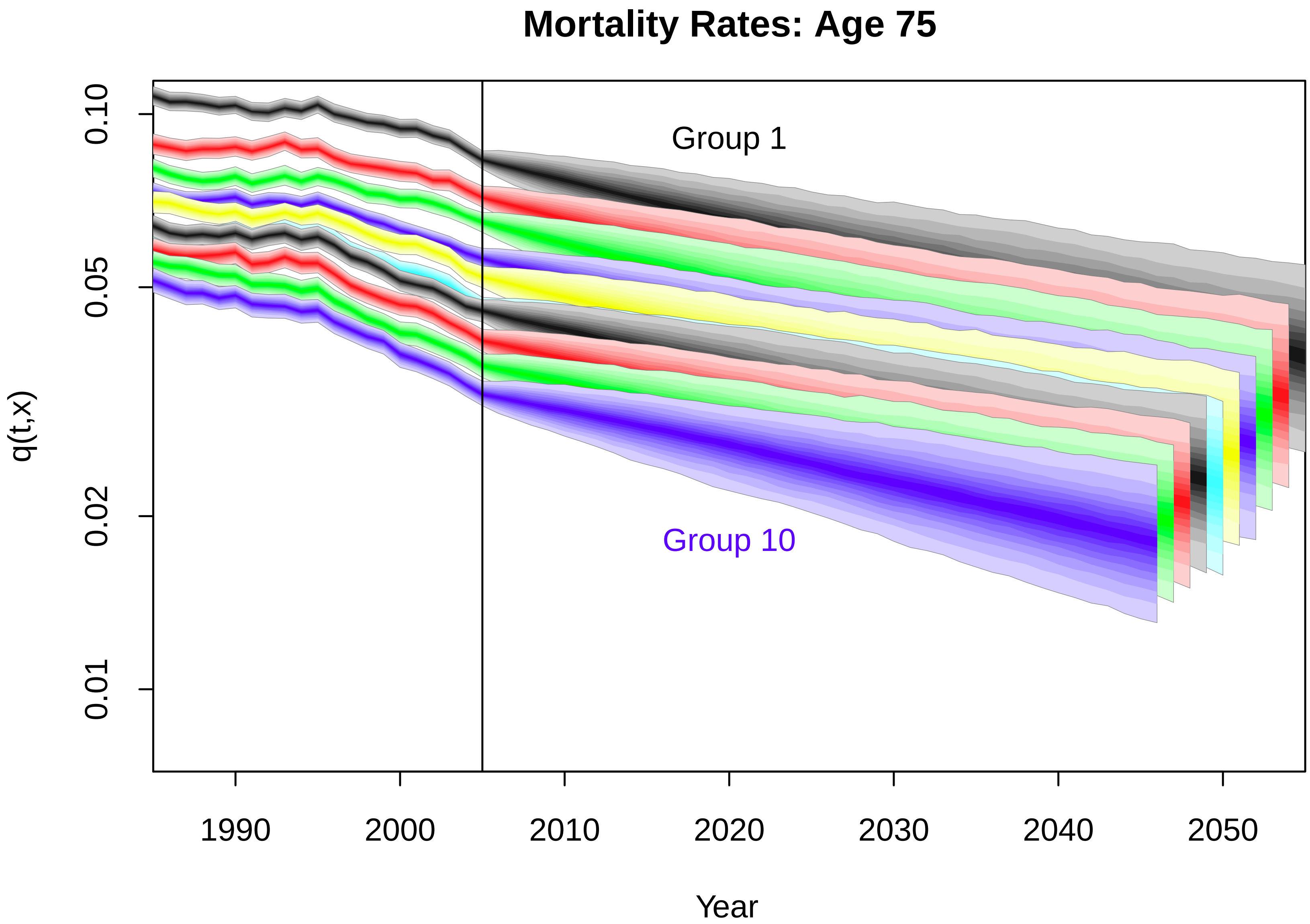
Posterior Distributions



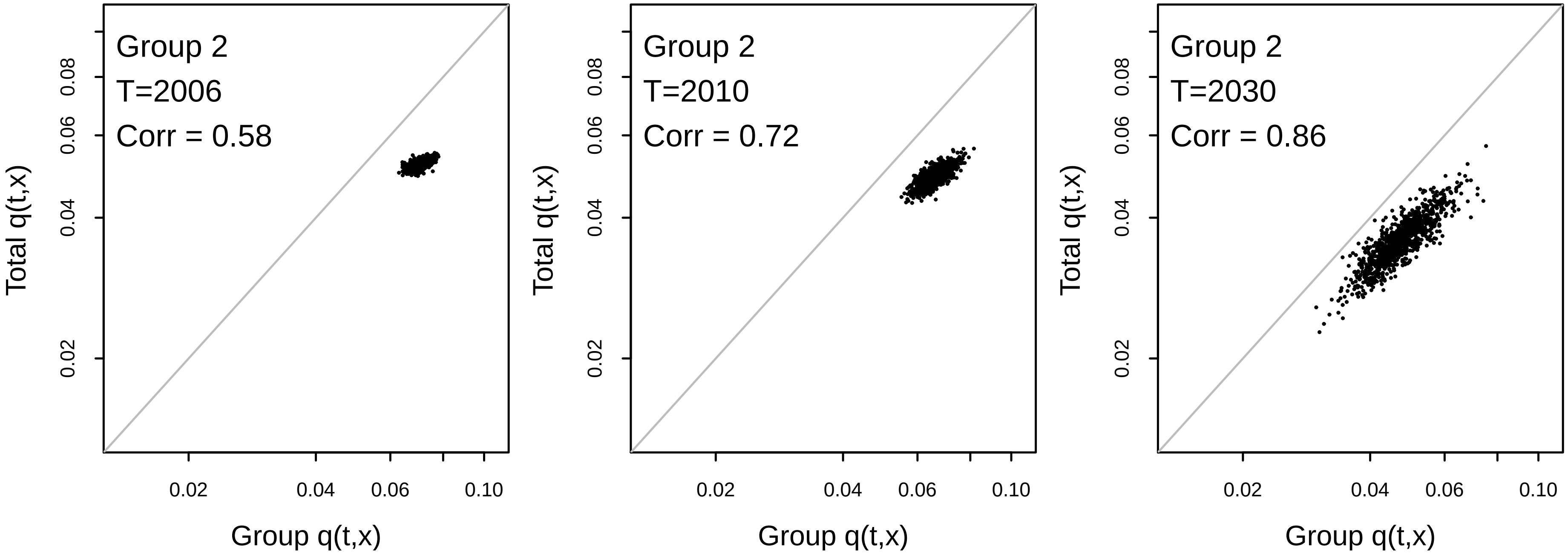
Life Expectancy for Groups 1 to 10



Mortality Fan Charts Including Parameter Uncertainty



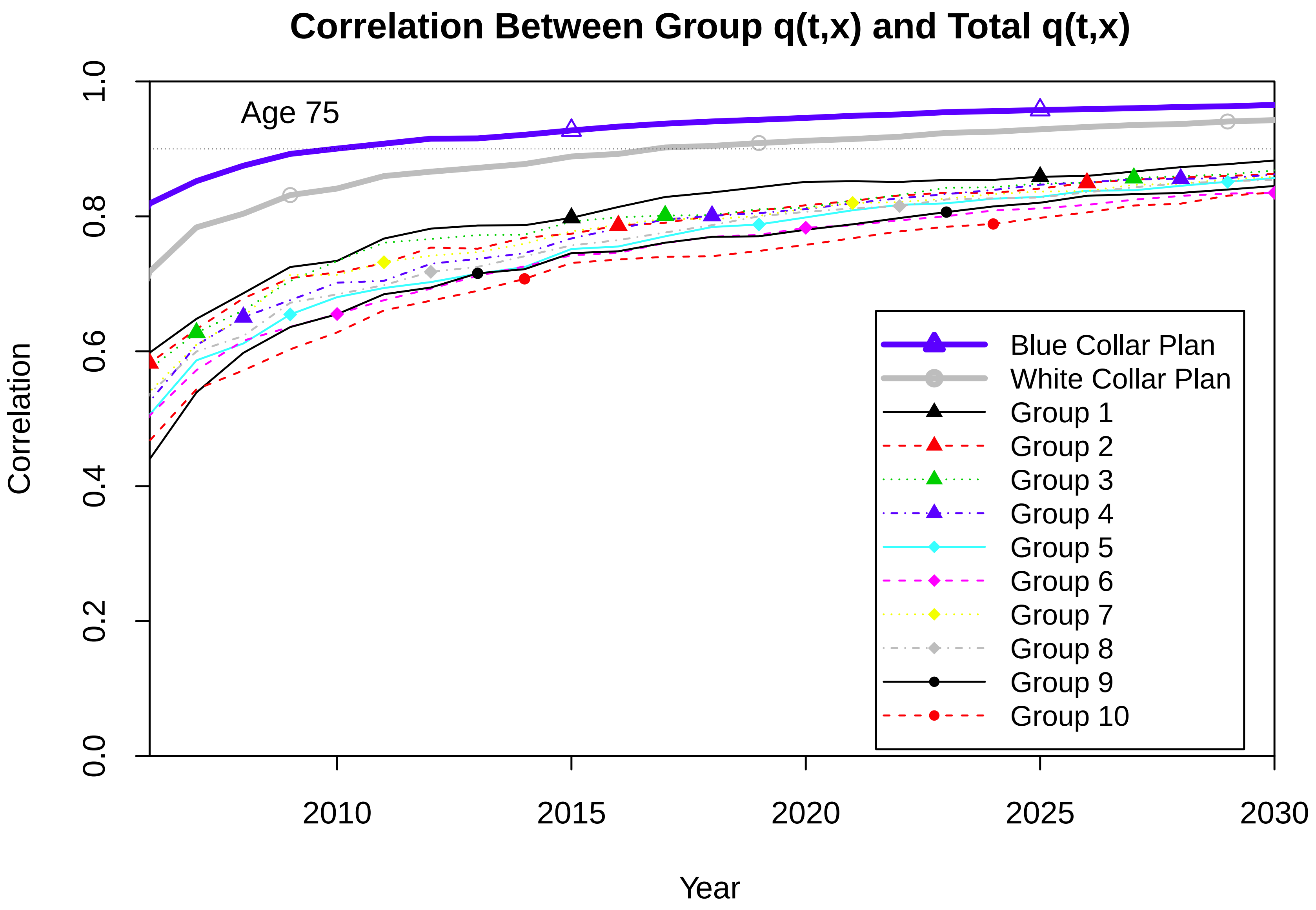
Simulated Group versus Population Mortality



As T increases

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

Forecast Correlations



Conclusions

- Development of a new multi-population dataset for Denmark

  strong biologically reasonable group rankings

- Unlike multi-country data

  a priori ranking of wealth-related groups

- Proposal for a simple new multi-population model

- Next steps:
  - Females data
  - More general correlation and gravity structures