A Bayesian Model for Small Population Bias and Sampling Effects in Stochastic Mortality Modelling

Dr. Liang Chen, FIA, FSAI, CERA, AFHEA
Prof. Andrew J.G. Cairns, Dr. Torsten Kleinow

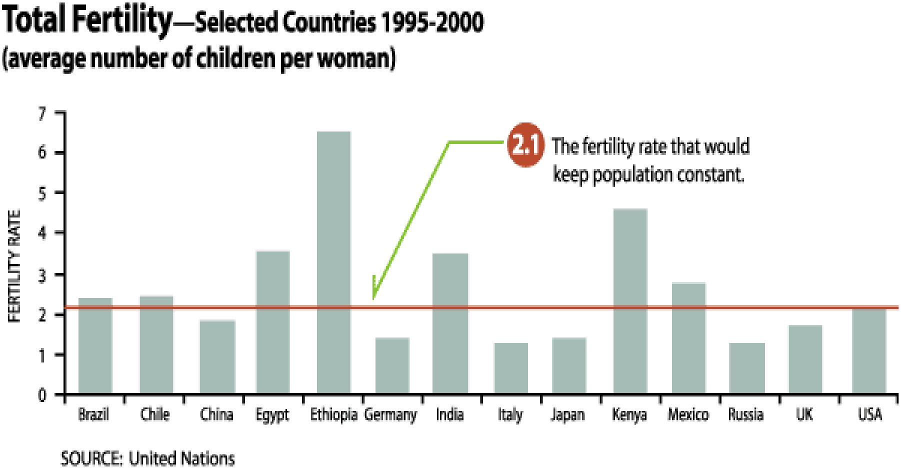
Heriot-Watt University, Edinburgh

University College Cork, Ireland

Funded by Actuarial Research Centre, IFoA

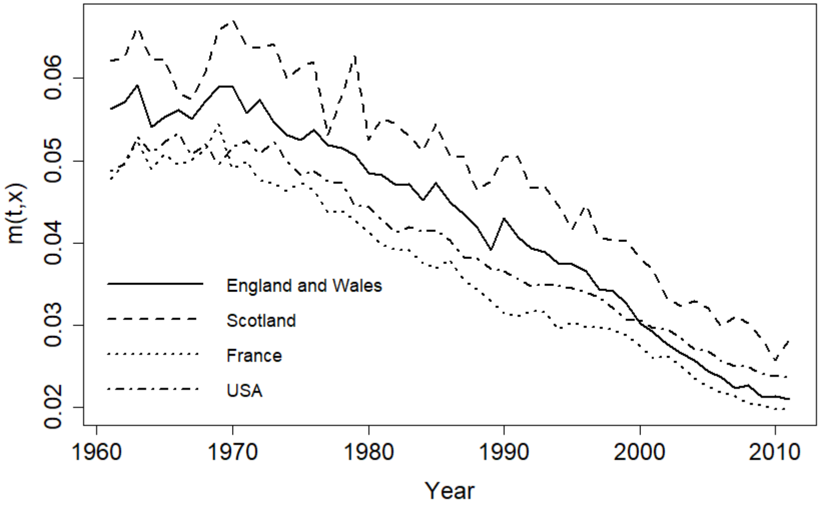
ATRC, Liverpool, 27 June 2019

Fertility

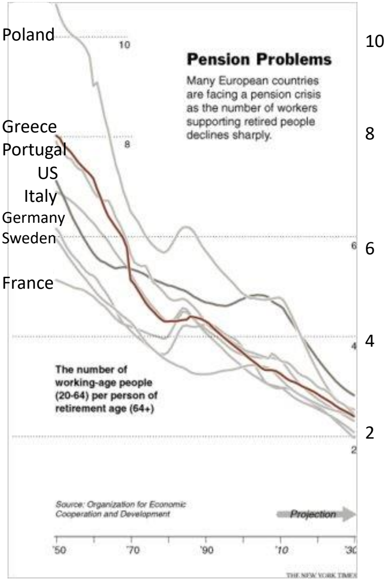


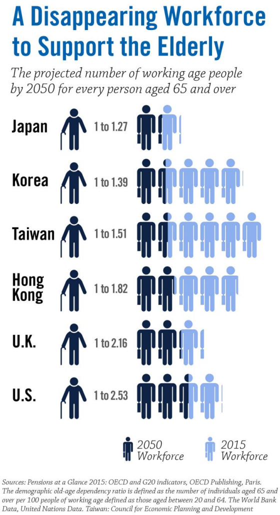
Total Fertility—Selected Countries 1995-2000
(average number of children per woman)
FERTILITY RATE
2.1 The fertility rate that would keep population constant.
Brazil
Chile
China
Egypt
Ethiopia
Germany
India
Italy
Japan
Kenya
Mexico
Russia
UK
USA
SOURCE: United Nations

Mortality Improvement at age 70 (HMD)



Effect of Ageing





Measure for Mortality

- $m_c(t,x)$, the age-specific crude death rate at age x, year t, More specifically

$$m_c(t,x) = \frac{\text{Number of deaths during calendar year t, age x last birthday}}{\text{Average population during calendar year t aged x last birthday}}.$$

- $m(t,x)$, the underlying death rate, which is equal to the expected deaths divided by the exposure. More specifically

$$m(t,x) = \frac{D(t,x)}{E(t,x)}$$

- $q(t,x)$, the mortality rate, which is the probability that an individual aged exactly x at exact time t will die between t and $t+1$.
- $\mu(t,x)$, the force of mortality.

Why analyse small population

- Experiencing faster mortality improvement, lower interest rate, more pressure on pension funds.
- Most pension schemes are less than 1% of national population.
- Significantly more variability exhibited for mortality rates of small population
- Stochastic models might poorly fit small populations

Motivation

For small population:

- Greater sampling variation of deaths causes increased uncertainty of parameter estimates and high levels of uncertainty on projected mortality rates.
- Diverge between future realized rates and projections, future sampling variation, uncertain projection.

Stochastic Model and Data

- Stochastic Model:

$$\begin{aligned} D(t,x)|\theta_1 &\sim \text{Pois}(m(\theta_1,t,x)E(t,x)) \\ m(\theta_1,t,x) &= -\log(1-q(\theta_1,t,x)) \\ \text{logit } q(\theta_1,t,x) &= \kappa_t^{(1)} + \kappa_t^{(2)}(x-\bar{x}) + \kappa_t^{(3)}((x-\bar{x})^2 - \hat{\sigma}_x^2) + \gamma_{t-x}^{(4)} \end{aligned}$$

- Data: Benchmark exposure $E_0(t,x)$ and corresponding deaths count $D_0(t,x)$ of the males in England and Wales (EW) in the HMD database, during year 1961 to 2011, aged 50-89 last birthday.

- Simulation Method (Chen, Cairns and Kleinow 2017)
  - Estimate θ_1 for benchmark population, denoted as $\hat{\theta}_{1,0}$
  - Construct small population $E_w(t,x) = wE_0(t,x)$ for $w = 1, 0.1, 0.01, 0.001,$
  - (Re-) Simulate $D_w(t,x)|\hat{\theta}_{1,0} \sim \text{Pois}(m(\hat{\theta}_{1,0},t,x)wE_0(t,x))$
  - Estimate θ_1 for $D_w(t,x)$, denoted as $\hat{\theta}_1^w$.

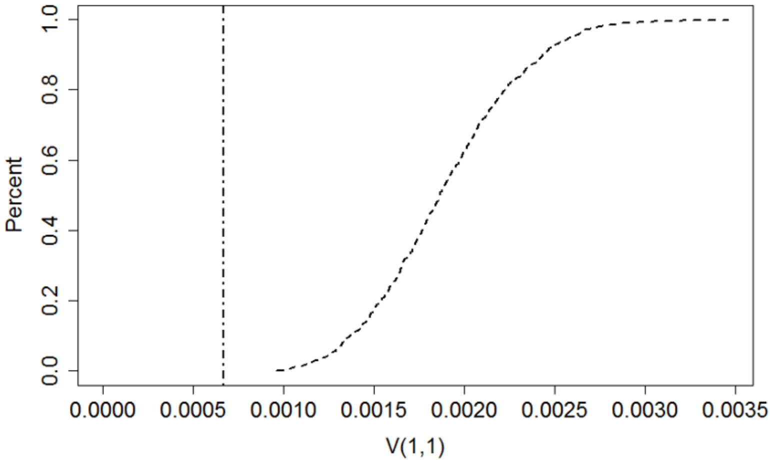
Actuarial Research Centre
Institute and Faculty of Actuaries

Two-Stage and Bayesian Approaches

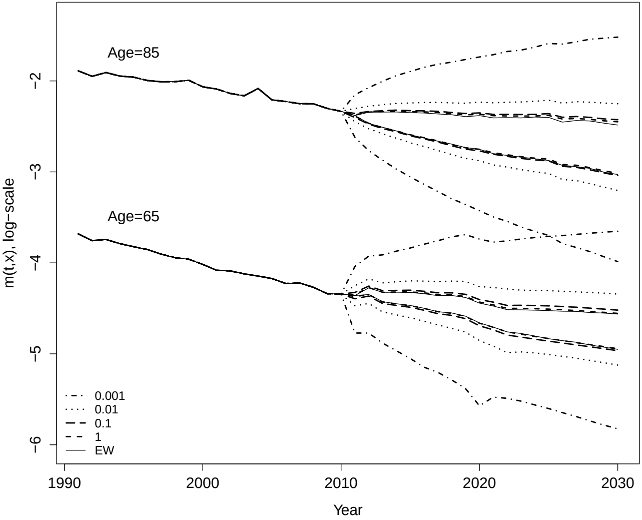
- Two-stage approach leads to biased estimates of volatility for small populations
  - Large sampling variation affects latent parameter estimation, with significant noise obscuring the true signal (Cairns et al. 2011)
  - Result in non-negligible bias to the parameter estimation of the projecting model (Chen, Cairns and Kleinow 2017)
  - Over-fit the short cohorts (Cairns et al. 2009)

Actuarial Research Centre
Institute and Faculty of Actuaries

$\hat{V}_\epsilon^{w}(1,1)$ given $\hat{\theta}_1^{w}$ verse $\hat{V}_\epsilon^{\text{EW}}(1,1)$, $w = 0.01$



Projected Mortality Rate



Two-Stage and Bayesian Approaches

- Two-stage approach leads to biased estimates of volatility for small populations
  - Large sampling variation affects latent parameter estimation, with significant noise obscuring the true signal (Cairns et al. 2011)
  - Result in non-negligible bias to the parameter estimation of the projecting model (Chen, Cairns and Kleinow 2016)
  - Over-fit the short cohorts (Cairns et al. 2009)
- Bayesian approach offers a way to avoid or reduce this bias by:
  - Combining the Poisson likelihood with the projecting time series models
  - The estimated latent parameters are restricted to be more like proposed time series models when projecting models dominate while modelling small populations.
  - Using more informative prior distribution with the knowledge of the larger benchmark population.
  - Better estimation for short cohorts.

Prior Distributions

- $(\kappa_{t_1}^{(1)}, \kappa_{t_1}^{(2)}, \kappa_{t_1}^{(3)}) \propto 1$,

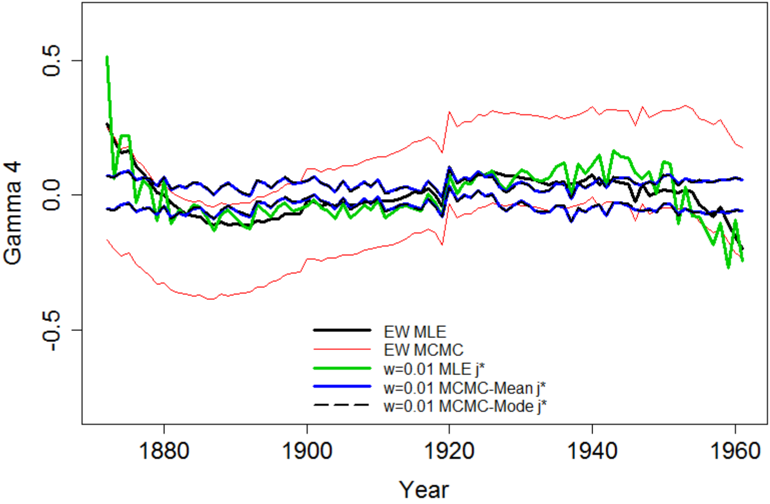
- $\boldsymbol{\kappa}_t = \boldsymbol{\kappa}_{t-1} + \boldsymbol{\mu} + \boldsymbol{\epsilon}_t$ for $t \leq t_2$,
  - $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu) \propto 1$,
  - $\boldsymbol{\epsilon}_t \sim MVN(0, \boldsymbol{V}_\epsilon)$, i.i.d there dimensional multi-variate normal distribution independent of t,

- $\boldsymbol{V}_\epsilon \sim InverseWishart(\nu, \boldsymbol{\Sigma})$
  - MCMC-Mean: Fix the mean of the prior to $\hat{\boldsymbol{V}}_\epsilon^{\text{EW}}$
  - MCMC-Mode: Fix the mode of the prior to $\hat{\boldsymbol{V}}_\epsilon^{\text{EW}}$

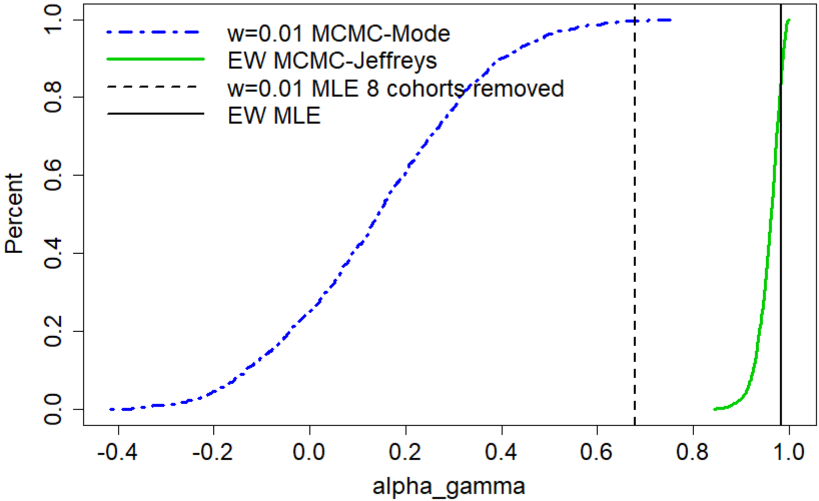
- $\gamma_c^{(4)} = \alpha_\gamma \gamma_{c-1}^{(4)} + \epsilon_c$ for $c > t_1 - x_{n_a}$,
  - i.i.d $\epsilon_c \sim N(0, \sigma_\gamma^2)$,
  - $\alpha_\gamma \propto (1 - \alpha_\gamma^2)^g$ for $|\alpha_\gamma| < 1$,
  - $\sigma_\gamma^2 \sim$ Inverse Gamma (a_γ, b_γ)

- $\gamma_{c_1}^{(4)} \sim N(0, \frac{\sigma_\gamma^2}{1-\alpha_\gamma^2})$

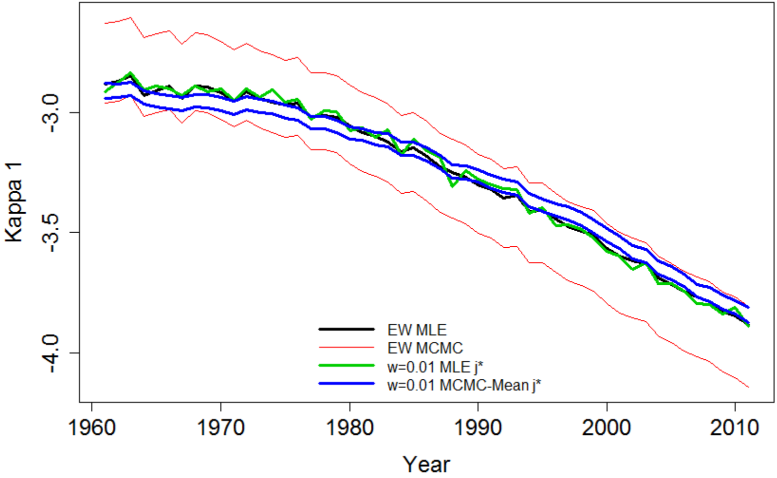
Results: γ MCMC



Results: α_γ MCMC

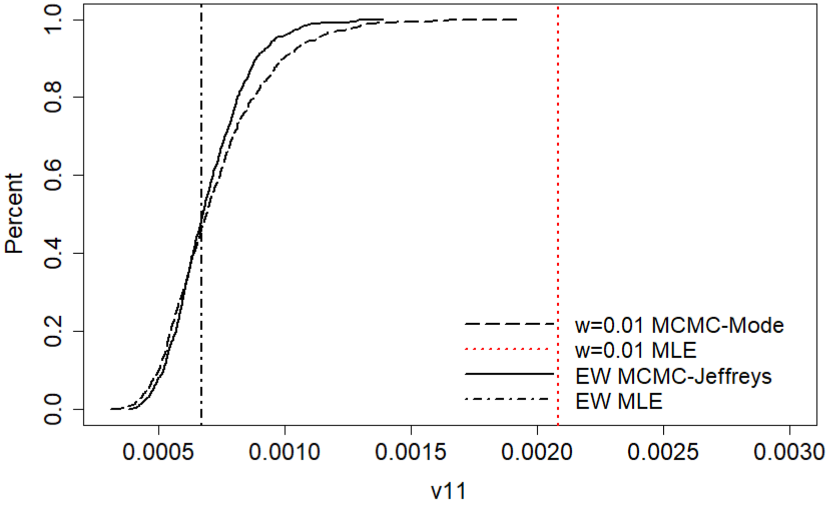


Results $\boldsymbol{\kappa}^{(1)}$

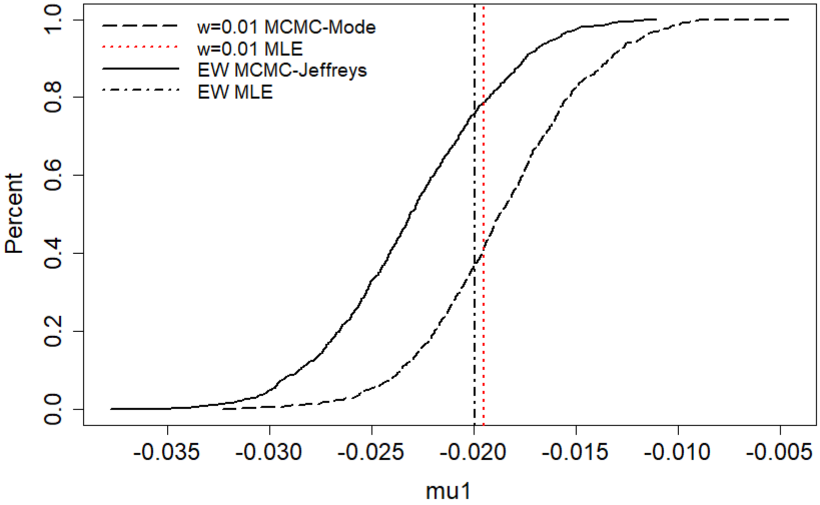


Actuarial Research Centre
Institute and Faculty of Actuaries

Results: $\boldsymbol{V}_\epsilon(1,1)$ MCMC



Results: $\mu(1,1)$ MCMC



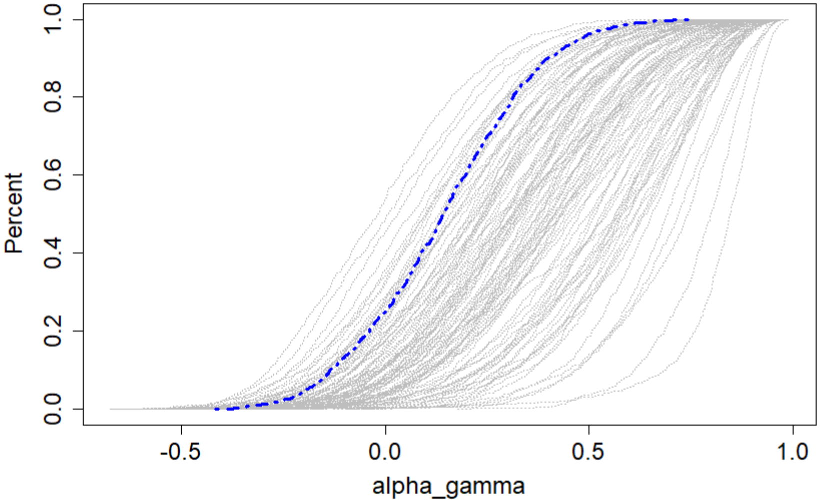
Longevity Risk of A Temporary Annuity

	$i =$ 4%		$i =$ 2%		$i =$ 0%	
	Mean	LR (%)	Mean	LR (%)	Mean	LR (%)
EW-MCMC	12.2631	5.27	14.8394	6.28	18.3365	7.47
w-MCMC	12.1220	5.72	14.6420	6.76	18.0556	7.95
EW-MLE	12.2166	4.24	14.7720	5.04	18.2371	5.98
w-MLE	12.2052	5.12	14.7441	6.08	18.1805	7.09

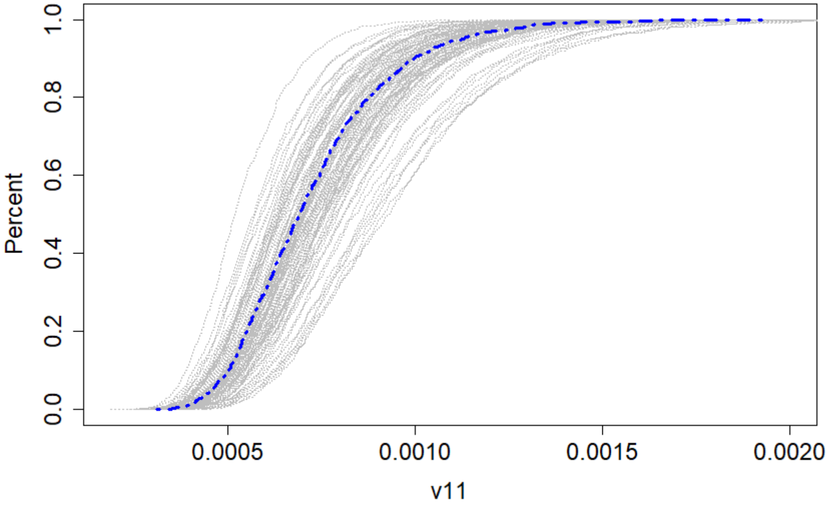
	$i =$4%		$i =$2%		$i =$ 0%	
	Mean	LR (%)	Mean	LR (%)	Mean	LR (%)
EW-MCMC	8.2744	7.02	12.2519	8.14	18.6117	9.45
w-MCMC	7.9292	8.23	11.6832	9.45	17.6509	10.87
EW-MLE	8.1928	5.50	12.1150	6.35	18.3759	7.32
w-MLE	8.2539	7.57	12.2039	8.80	18.5059	10.17

Actuarial Research Centre
Institute and Faculty of Actuaries

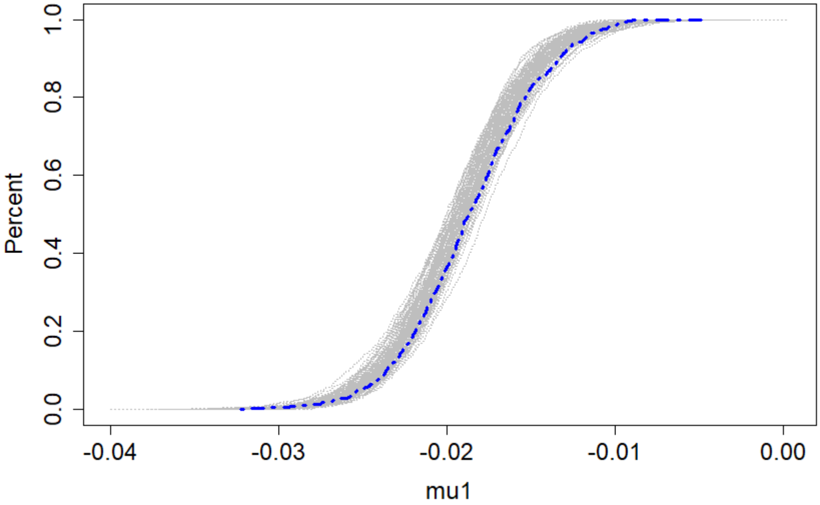
Sensitivity Test: α_γ MCMC



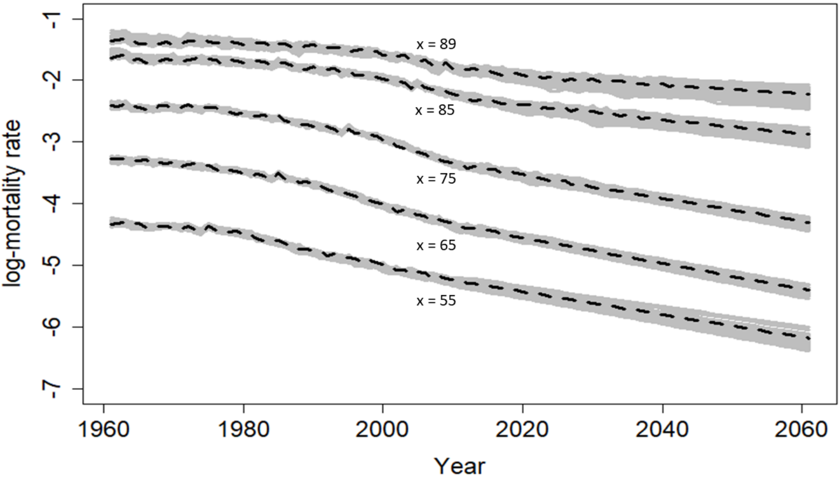
Sensitivity Test: $\boldsymbol{V}_{\epsilon}(1,1)$ MCMC



Sensitivity Test: μ_1 MCMC



Sensitivity Test: $m(t, x)$ MCMC



Summary

- We have demonstrated to the users of the stochastic mortality models how the information of a larger population could be embedded for parameter estimation and forecasts by a Bayesian model.
- Studied to what extent the parameter estimation could be improved compared with the two-stage approach and the financial implication.
- The users should be informed how the importance of the prior information takes over the parameter estimation of a much smaller population and in what way the sampling variation affects the parameter estimation and mortality forecasts.

To summarize part of our findings:

- We find that our Bayesian model and methodology of using the information of large referencing population provide an improved estimation for the volatility of small population.
- The (central) projections of small populations are not "significantly" different from the "true" projections (of the larger reference population).
- When the population is small, the prior distributions, in particular the time series prior for the latent parameters, dominate the likelihood.

Thank You!

Questions?