An Extension of the Lee-Carter Model for Mortality Projections

By

Udi Makov

,Shaul Bar-Lev, Yaser Awad ,

Extensions 1:

Bayesian Trend choice

The Lee-Carter model :

$$y_{x,t} = \alpha_x + \beta_x k_t + \varepsilon_{x,t}, \quad \varepsilon_{x,t} \rightarrow F(o, \sigma_\varepsilon^2),$$

$$k_t = k_{t-1} + \theta + w_t, \quad w_t \xrightarrow{iid} N(o, \sigma_w^2),$$

$y_{x,t}$: the logarithm of the central death rate for age x in year t.

α_x : general shape across age of the mortality schedule.

β_x : indicates the tendency of the logarithm of the force of mortality at age x when the general level of mortality (k_t) changes.

k_t : is an index that describes the variation in the level of mortality at time t.

Estimation Methodology

The methodology is based on the following steps:

1. Fixing appropriate prior distributions for the unknown parameters

2. Using the Kalman Filter algorithm to estimate the predicted values of the time series process, including the variance of the prediction

3. Using the Gibbs sampler to estimate the posterior distributions of the unknown parameters, including the parameters of the unobserved time series index, whose values were estimated by the Kalman Filter algorithm.

The Lee Carter model as a state-space model

Equations below are describing the state process model and the observation model as follow:

$$\begin{cases} k_t = k_{t-1} + \theta + w_t & \text{, } \textbf{state equation} \\ y_t = \alpha + \beta k_t + \nu_t & \text{, } \textbf{measurement equation} \end{cases}$$

We assumed the process noise is $w_t \overset{iid}{\sim} N(0, \sigma_w^2)$

and the measurement noise is $\nu_t \overset{iid}{\sim} N_p(0, \sigma_\varepsilon^2 I)$

General Presentation of time trend model

$$\begin{cases} k_t = \rho k_{t-1} + \theta + \eta_t - \rho \eta_{t-1} + w_t - \varphi w_{t-1} \\ \eta_t = \gamma_1 + \gamma_2 t \end{cases}$$

In vectorial presentation

$$(k - X\gamma) = P(k - X\gamma) + \theta + (I - \Psi)w$$

where

$$k = (k_{\min}, k_{\min+1}, \cdots, k_{\max}) \quad \theta = (\theta, \theta, \cdots, \theta)$$

$$w = (w_{\min}, w_{\min+1}, \cdots, w_{\max})$$

$$P = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ \rho & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \rho & 0 \end{bmatrix}, \quad X = \begin{bmatrix} 1 & t_{\min} \\ \vdots & \vdots \\ 1 & t_{\max} \end{bmatrix}, \quad \gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix}, \quad \Psi = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ \varphi & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \varphi & 0 \end{bmatrix}$$

where

$$k \sim N_T(X\gamma + \theta, \sigma_k^2(UQ^{-1}))$$

$$Q = (I = P)'(I - P) = \begin{bmatrix} 1+\rho^2 & -\rho & \cdots & \cdots & 0 \\ -\rho & 1+\rho^2 & -\rho & \ddots & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & 0 \\ \vdots & \ddots & -\rho & 1+\rho^2 & -\rho \\ 0 & \cdots & 0 & -\rho & 1 \end{bmatrix} \quad U = \begin{bmatrix} 1+\varphi^2 & -\varphi & \cdots & \cdots & 0 \\ -\varphi & 1+\varphi^2 & -\varphi & \ddots & 0 \\ 0 & -\varphi & 1+\varphi^2 & -\varphi & 0 \\ \vdots & \ddots & -\varphi & 1+\varphi^2 & -\varphi \\ 0 & \cdots & 0 & -\varphi & 1 \end{bmatrix}$$

Extensions

$\text{Model 1}: \rho = 1, \gamma_1 = \gamma_2 = 0, \varphi = 0 \Rightarrow \text{ARIMA}(0,1,0)$

(Lee & Carter, 1992; Pedroza, 2006)

$$k_t = k_{t-1} + \theta + w_t, \quad \text{with } w_t \sim N(0, \sigma_w^2)$$

Then the resulting distribution is:

$k \sim N_T(\theta, \sigma_w^2 Q_{\rho=1}^{-1})$

Model 2 $: \theta = 0, \varphi = 0 \Rightarrow$ ARIMA(1,0,0) with linear trend (Czado et al., 2006)

$$\begin{cases} k_t = \rho k_{t-1} + \eta_t - \rho \eta_{t-1} + w_t \\ \eta_t = \gamma_1 + \gamma_2 t \end{cases}$$

Then the resulting distribution is:

$$k \sim N_T(X\gamma, \sigma_k^2 Q^{-1})$$

Model 3: $\rho = 1, \gamma_1 = \gamma_2 = 0 \Rightarrow \text{ARIMA}(0,1,1)$

Simple exponential smoothing with growth

$$k_t = k_{t-1} + \theta + w_t - \varphi w_{t-1}$$

Then the resulting distribution is:

$$k \sim N_T(\theta, \sigma_w^2 U Q_{\rho=1}^{-1})$$

Model 4: $\gamma_1 = \gamma_2 = 0, \varphi = 0 \Rightarrow \mathrm{ARIMA}(1,0,0)$

As case 2 without linear trend

$$k_t = \rho k_{t-1} + \theta + w_t$$

Then the resulting distribution is:

$$k \sim N_T(\theta, \sigma_w^2 Q^{-1})$$

Model 5: $\gamma_1 = \gamma_2 = 0 \Rightarrow \text{ARIMA}(1,0,1)$

$$k_t = \rho k_{t-1} + \theta + w_t - \varphi w_{t-1}$$

Then the resulting distribution is:

$$k \sim N_T(\theta, \sigma_w^2 Q^{-1} U)$$

Model Choice

Model 1: $\rho = 1, \gamma_1 = \gamma_2 = 0, \phi = 0 \Rightarrow ARIMA(0,1,0)$

(Lee & Carter, 1992; Pedroza, 2006)

Model 2: $\theta = 0, \varphi = 0 \Rightarrow ARIMA(1,0,0)$ *with linear trend, (Czado et al., 2005)*

Model 3: $\rho = 1, \gamma_1 = \gamma_2 = 0 \Rightarrow ARIMA(0,1,1)$

Simple exponential smoothing with growth

Model 4: $\gamma_1 = \gamma_2 = 0, \varphi = 0 \Rightarrow ARIMA(1,0,0)$

As case 2 without linear trend

Model 5: $\gamma_1 = \gamma_2 = 0 \Rightarrow ARIMA(1,0,1)$

Some data consideration

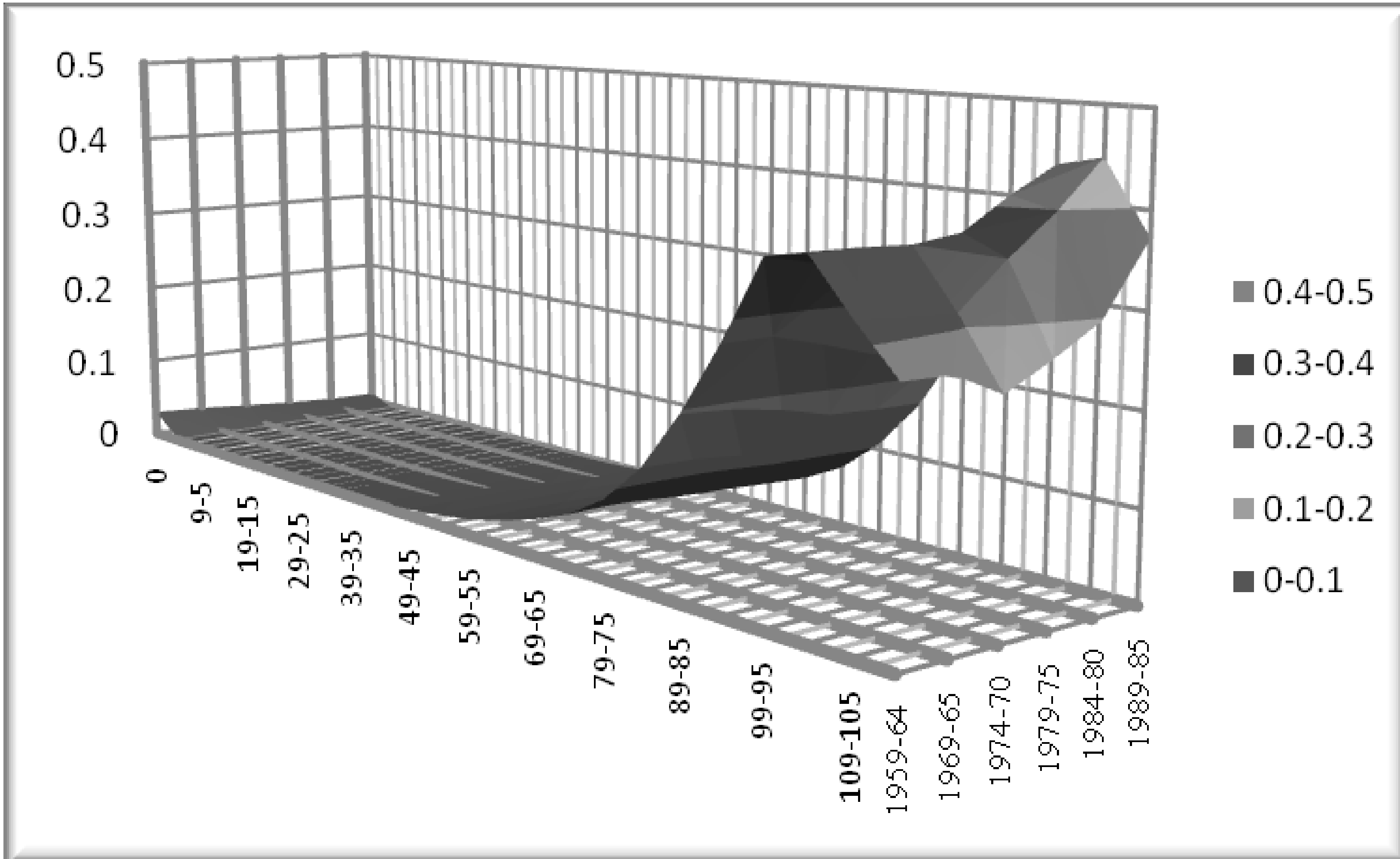
We analyze the data for U.S and Ireland male. The data obtained from the Human Mortality Database .

The data consists of annual age-specific death rates for years 1959-1999.

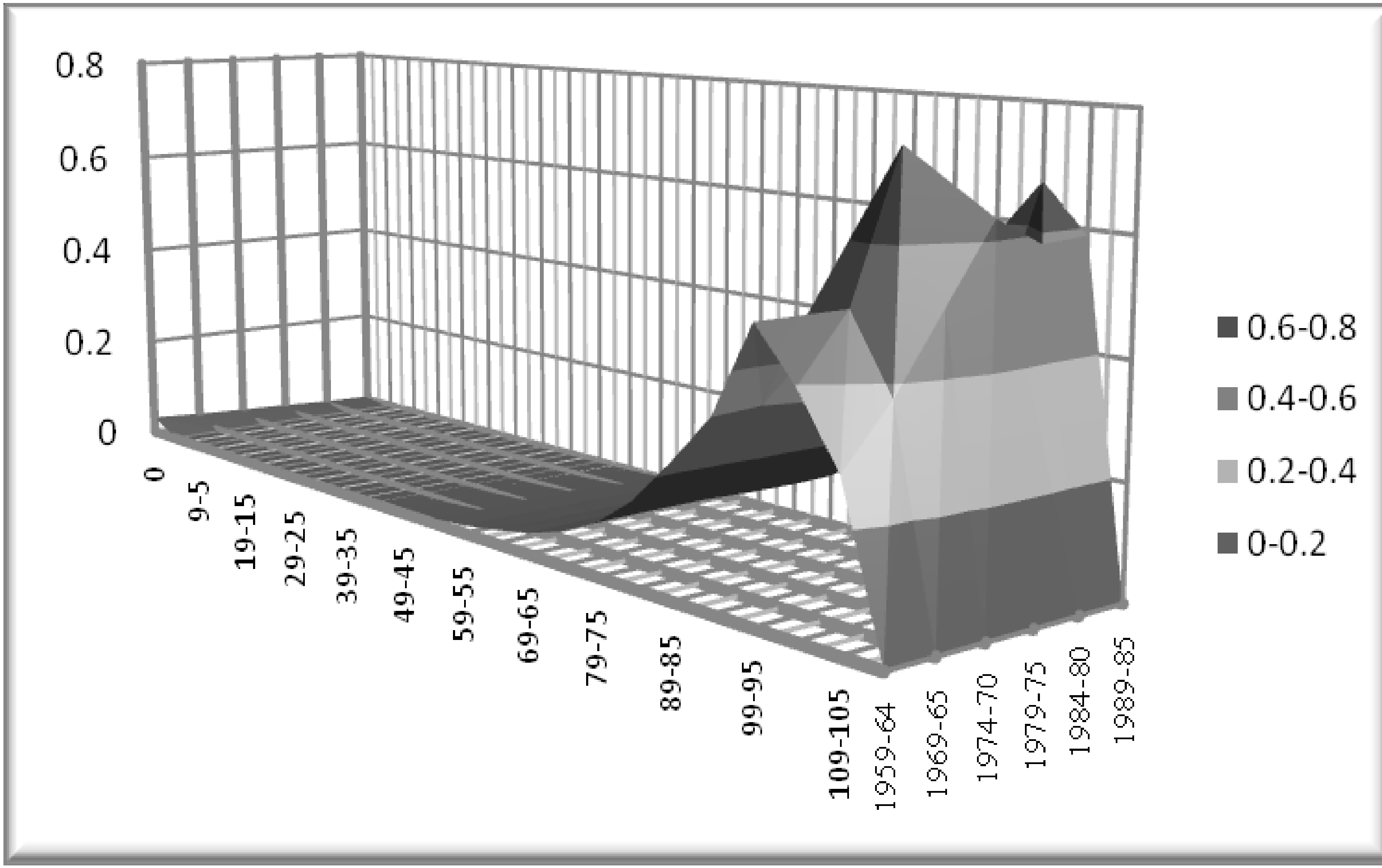
The age group for U.S are 0,1-4,5-9,…,105-109,110+.
The age group for Ireland are 0,1-4,5-9,…,105-109.

Therefore; for U.S (P=24, T=6) , for Ireland (P=23, T=6), that is 1958-1989 grouped by 5 years.

The distribution of the mortality rates by age groups and time U.S Male



The distribution of the mortality rates by age groups and time Ireland Male



Prior distribution

$$\alpha \sim N_p(a_\alpha, \sigma_\alpha^2 I),\ \sigma_\alpha^{-2} \sim Gamma(a_x, b_x)$$

$$\beta \sim N_p(0, \sigma_\beta^2 I),\ \sigma_\beta^{-2} \sim Gamma(a_\beta, b_\beta)$$

$$\rho \sim N(0, \sigma_\rho^{-2})\ truncated\ to\ the\ \mathrm{int}erval\ (-1,1)$$

$$\sigma_{\mathbf{w}}^{\mathbf{-2}} \sim Gamma(a_w, b_w),\ \sigma_\varepsilon^{-2} \sim Gamma(a_\varepsilon, b_\varepsilon)$$

$$\begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} \sim N_2\left(\begin{pmatrix} \gamma_{01} \\ \gamma_{02} \end{pmatrix}, \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix} \right)$$

Bayesian Model Choice

For models comparison we estimated the *Deviance information criterion* – DIC for normal likelihood. smaller values for “better” models. (Spiegelhalter, et al. 2002).

$$DIC = D(\overline{\Theta}) + 2pD$$

where :

$$D(\overline{\Theta}) = -2\log(f(y/\overline{\Theta}))$$

$p =$ the effective number of parameters

$\overline{\Theta} =$ the sample average of the simulated values of Θ

We used the data of 1958-1989 to estimate the models parameters and the DIC index.

Bayesian Model Choice

The DIC indices for all five models using U.S male data are as follows:

DIC[model 1:ARIMA(0,1,0)]= -149.25
DIC[model 2:ARIMA(1,0,0)]= -154.26
DIC[model 3:ARIMA(0,1,1)]= -146.55
DIC[model 4:ARIMA(1,0,0)]= -108.96
DIC[model 5:ARIMA(1,0,1)]= -104.96

The DIC indices for all five models using Irish male data are as follows:

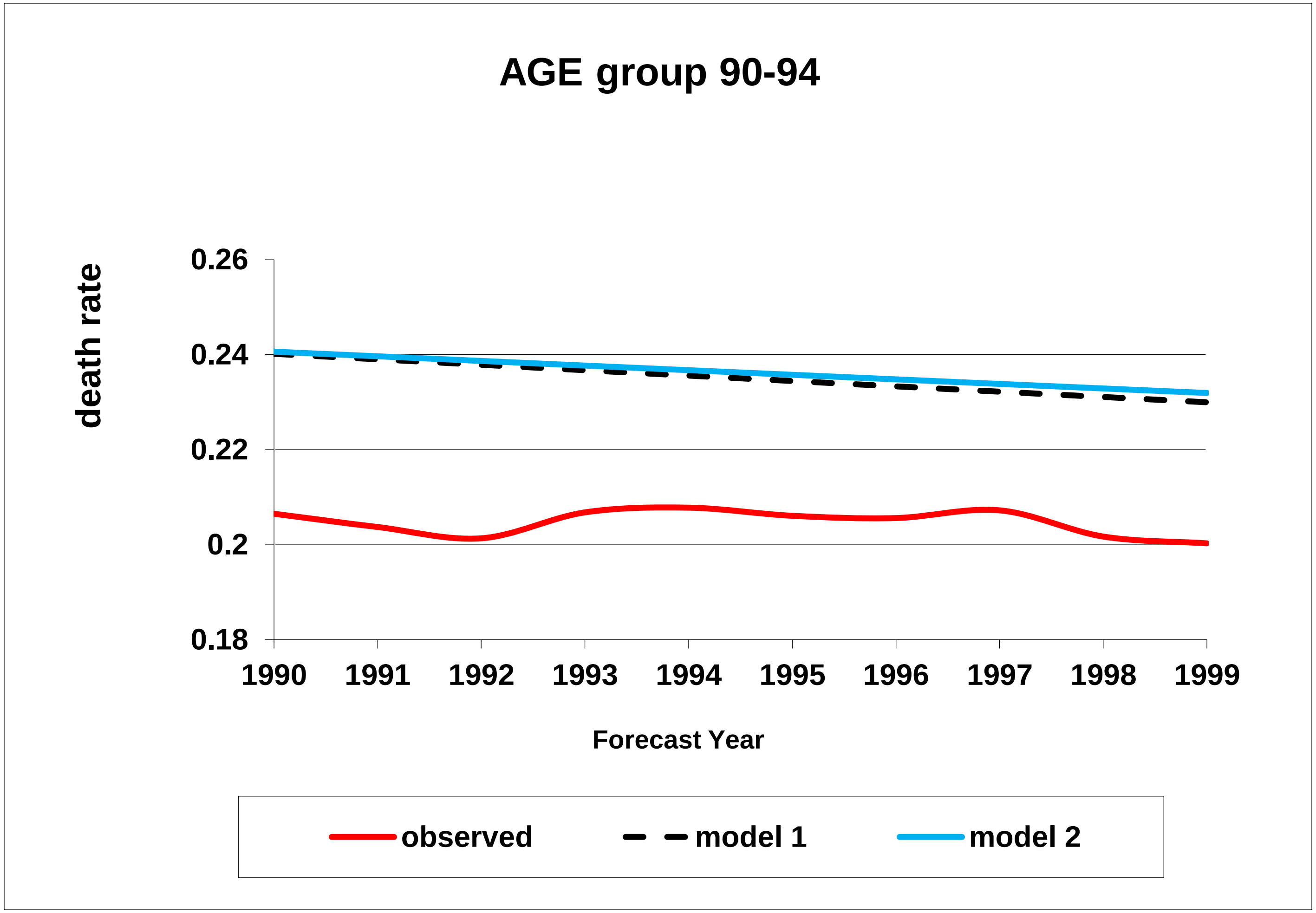
DIC[model 1:ARIMA(0,1,0)]= -130.07
DIC[model 2: ARIMA(1,0,0)]= -141.77
DIC[model 3: ARIMA(0,1,1)]= -123.91
DIC[model 4: ARIMA(1,0,0)]= -88.78
DIC[model 5: ARIMA(1,0,1)]= -87.35

We used the data of 1958-1989 to estimate the models parameters.

Then we used the data from 1990-1999 to assess the predictive quality of the models and we compared between them using also graphical presentation

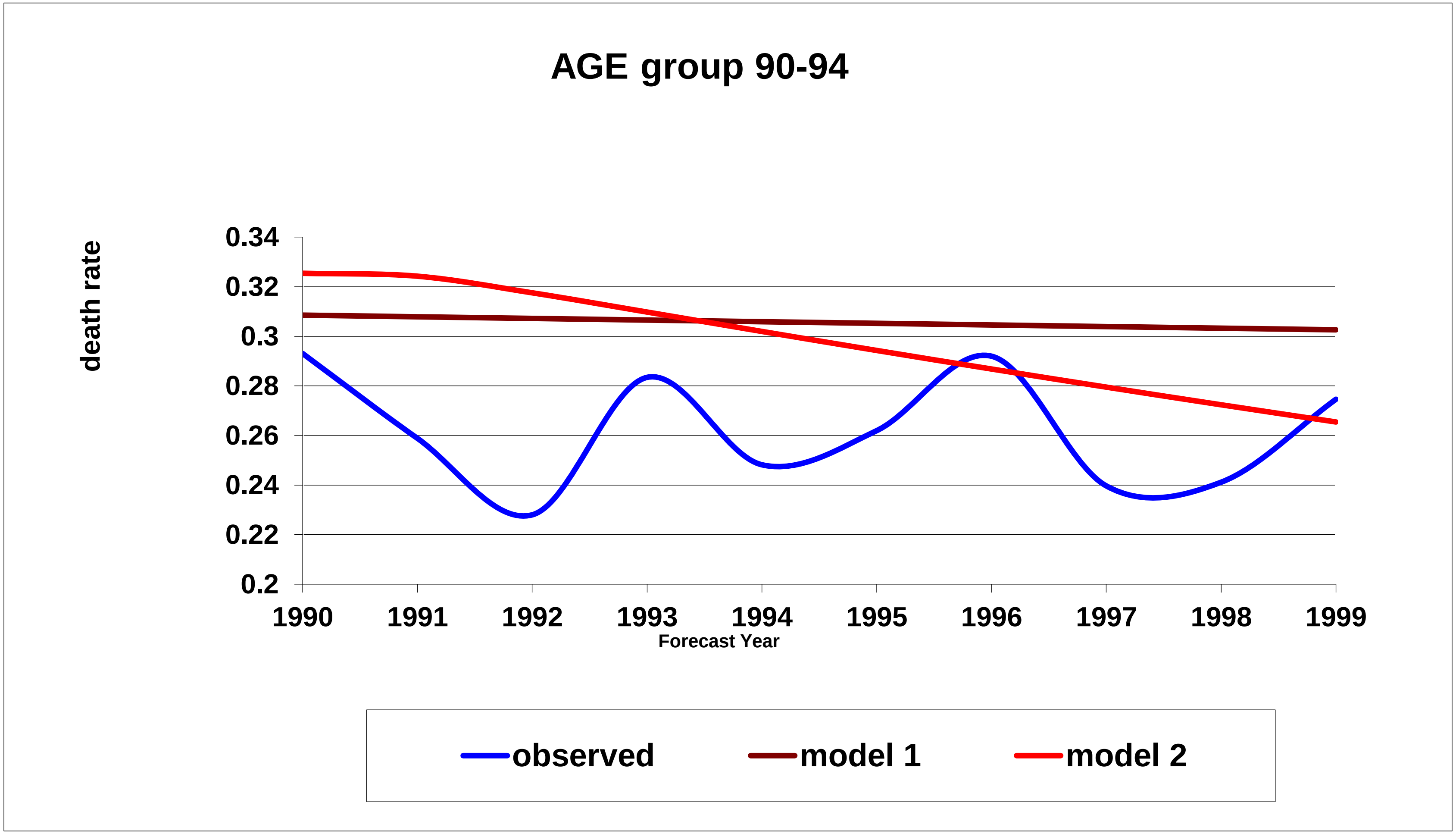
Comparison between model 1 and model 2

U.S Male



Comparison between model 1 and model 2

Ireland Male



Extensions 2:

Mortality Projections Simultaneous for Multiple populations

$$
\begin{tikzcd}
 & & & & & D & & & \\
 & & & & & \mu \arrow[u] & & & \\
 & & \kappa \arrow[urrr] & & & \beta \arrow[u] & & \alpha \arrow[ull] & \\
 & \gamma \arrow[ur] & \rho \arrow[u] & \sigma^2_\kappa \arrow[ul] & & \sigma^2_\beta \arrow[u] & & e \arrow[u] & \\
\gamma_0 \arrow[ur] & \Sigma_0 \arrow[u] & \sigma^2_\rho \arrow[u] & a_\kappa \arrow[u] & b_\kappa \arrow[ul] & a_\beta \arrow[u] & b_\beta \arrow[ul] & a_z \arrow[u] & b_z \arrow[ul]
\end{tikzcd}
$$

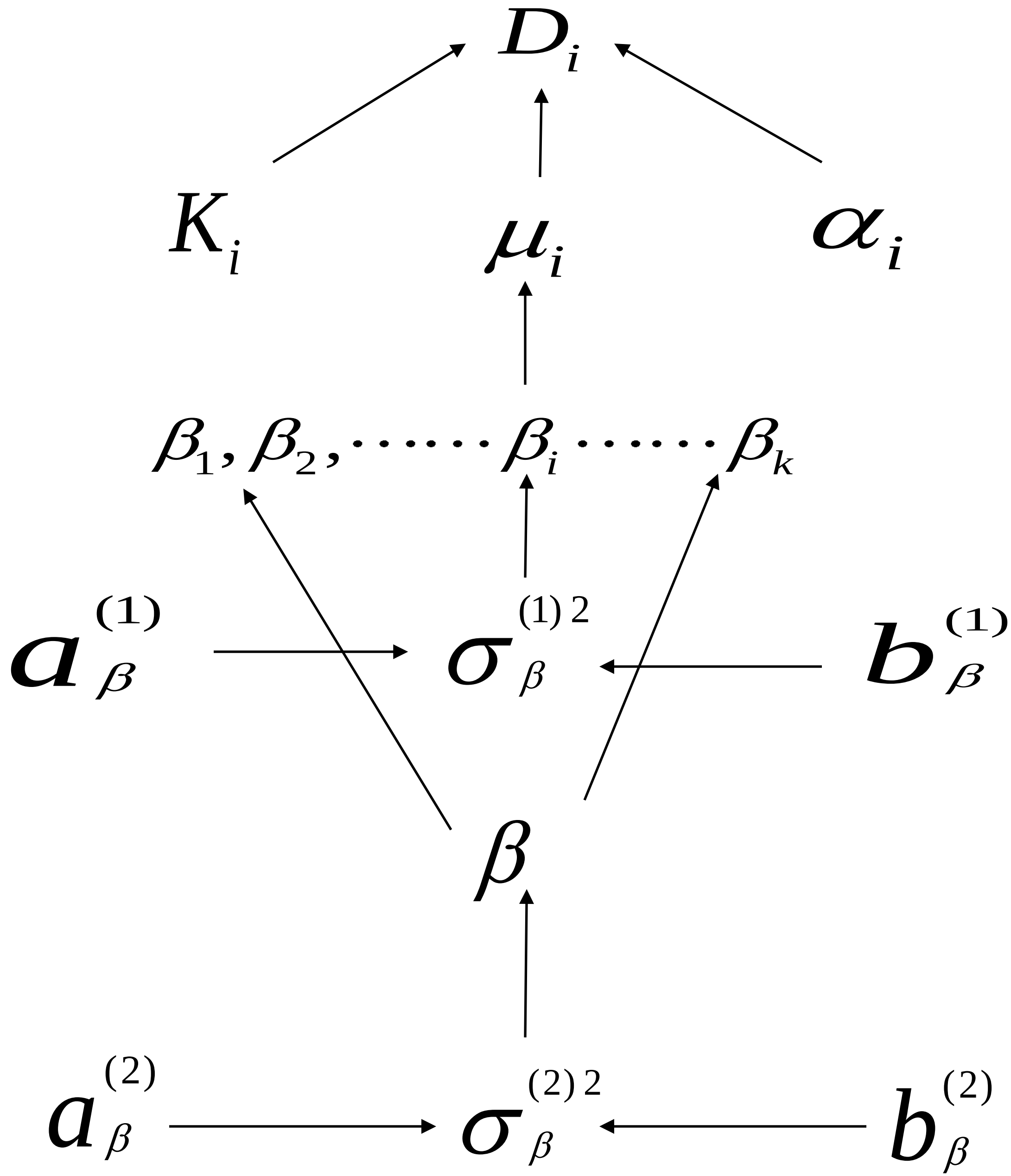
Exchangeable sequence of random variables

Formally, an exchangeable sequence of random variables is a finite or infinite sequence

X1, X2, X3, ... of random variables such that for any finite permutation σ of the indices 1, 2, 3, ..., i.e. any permutation σ that leaves all but finitely many indices fixed, the joint probability distribution of the permuted sequence

$$X_{\sigma(1)}, X_{\sigma(2)}, X_{\sigma(3)}, \ldots$$

is the same as the joint probability distribution of the original sequence.


D_i
K_i
μ_i
α_i
$\beta_1, \beta_2, \cdots\cdots \beta_i \cdots\cdots \beta_k$
$a_\beta^{(1)}$
$\sigma_\beta^{(1)\,2}$
$b_\beta^{(1)}$
β
$a_\beta^{(2)}$
$\sigma_\beta^{(2)\,2}$
$b_\beta^{(2)}$

Exchangeability between $\theta's$ parameters:

Suppose we have mortality projection for m countries and we believe the time trends parameters belong to a joint distribution

Suppose $\underline{\theta} = (\theta_1, \theta_2, \cdots, \theta_m)$ are i.i.d variables, with

$\pi(\theta_j / b, \eta) \rightarrow N(b, \eta), \forall\, j, \eta > 0$

The joint distribution of $\theta's$ is: $\pi(\underline{\theta} / b, \eta) = \prod_{j=1}^{m} \pi(\theta_j / b, \eta)$

Suppose $b \sim N(\mu_b, \sigma_b^2)$ and $\eta \sim IG(a_\eta, c_\eta)$, where $\mu_b, \sigma_b^2, a_\eta, c_\eta$ assumed known (maybe chosen to reflect vague prior information)

REPLACE: Sampling from θ

$$(\theta / K, k_0, \sigma_w^2) \rightarrow N\left(\frac{k_n - k_0}{T}; \frac{\sigma_w^2}{T}\right)$$

By the following posterior distributions, for m=2:

$$(\theta_1 / \Psi^*, b, \eta) \rightarrow N\left(\left[\frac{k_{T,1} - k_{0,1}}{T}\eta + b\frac{\sigma^2_{w,1}}{T}\right] \Big/ \left[\frac{\sigma^2_{w,1}}{T} + \eta\right]; \left[\frac{\sigma^2_{w,1}}{T}\eta\right] \Big/ \left[\frac{\sigma^2_{w,1}}{T} + \eta\right]\right)$$

$$(\theta_2 / \Psi^*, b, \eta) \rightarrow N\left(\left[\frac{k_{T,2} - k_{0,2}}{T}\eta + b\frac{\sigma^2_{w,2}}{T}\right] \Big/ \left[\frac{\sigma^2_{w,2}}{T} + \eta\right]; \left[\frac{\sigma^2_{w,2}}{T}\eta\right] \Big/ \left[\frac{\sigma^2_{w,2}}{T} + \eta\right]\right)$$

Where $\{K^1, \cdots, K^m, k_{0,1}, \cdots, k_{0,m}, \sigma^2_{w,1}, \cdots, \sigma^2_{w,m}\} \equiv \Psi^*$

The rest of the posterior distributions should stay the same as before for each country separately.

Similarly; we assumed Exchangeability on other parameters of the model: Alpha and Beta

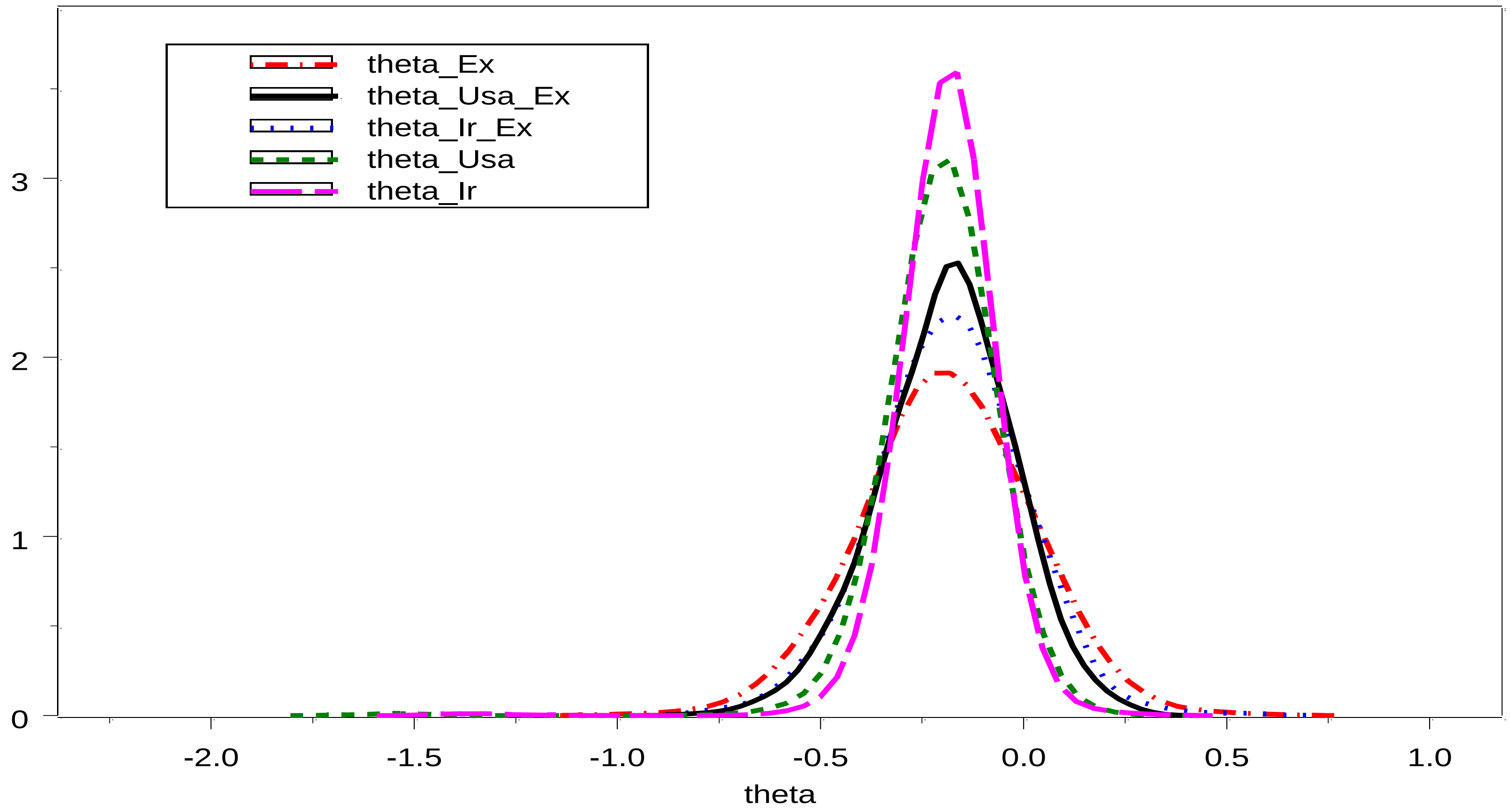
Validation

- We used the data from U.S and Ireland male of 1958-1989 to estimate the models parameters

- Then we used the data from 1990-1999 to assess the predictive quality of the models and we compared between them using also graphical presentation

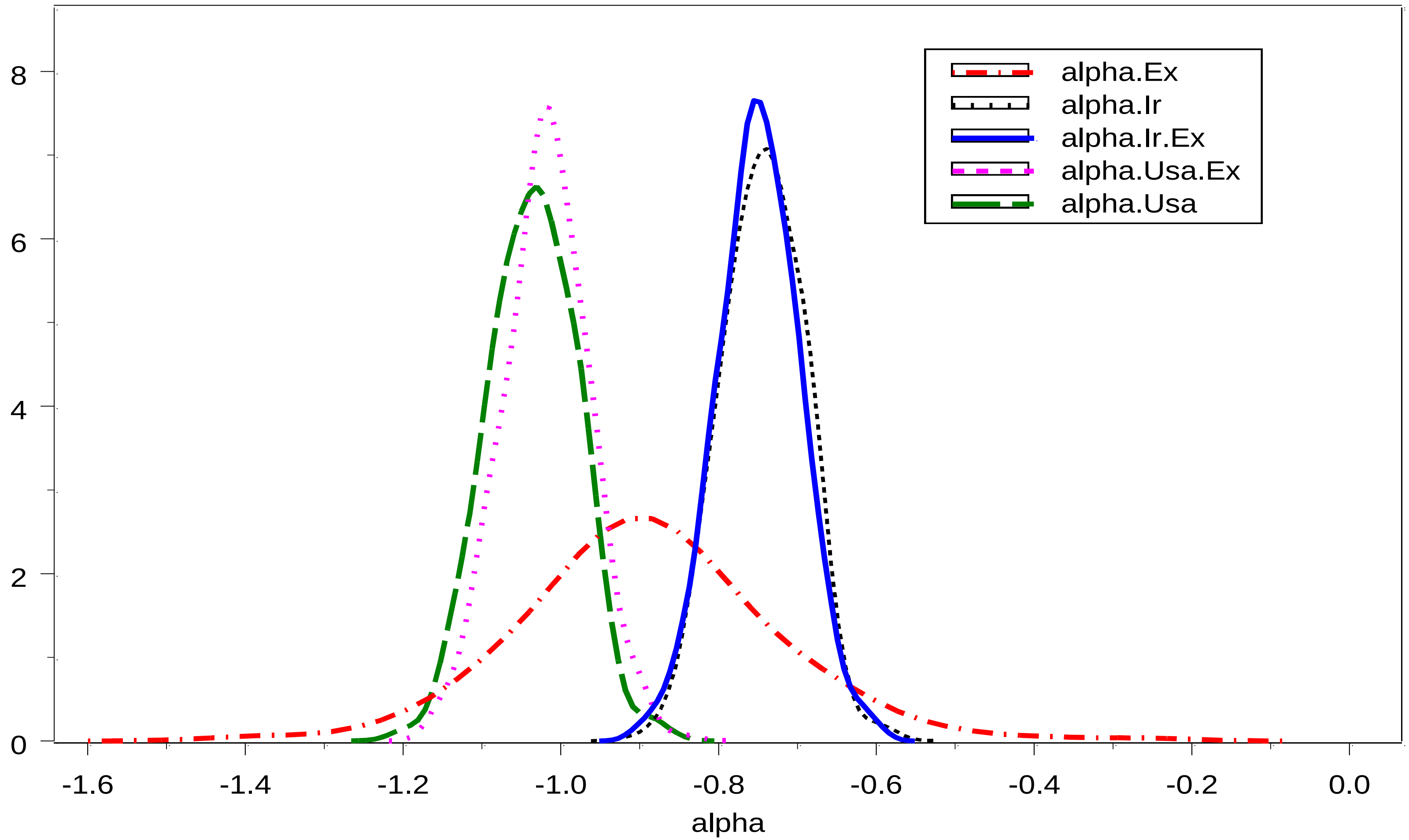
Exchangeability w.r.t Ɵ

Figure 1: Posterior dis. of theta : USA and Ireland



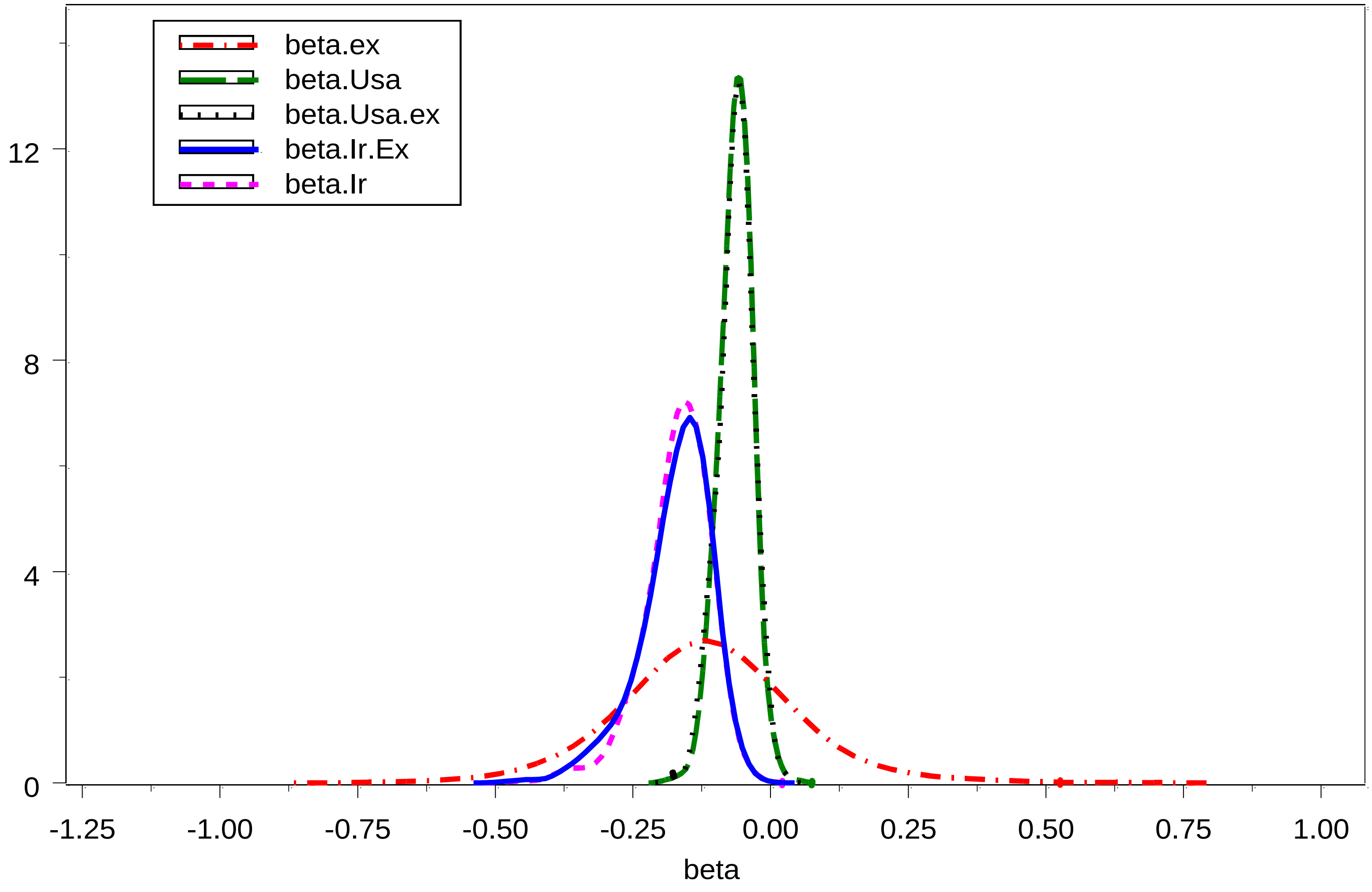
Exchangeability w.r.t α

Figure 3: Posterior dis. of alpha, age-group 105-109: Usa and Ireland



Exchangeability w.r.t β

Figure 5: Posterior dis. of beta, age-group 105-109: Usa and Ireland



Impact on α & β

beta_Usa_Ex	beta_Usa	alpha_Usa_Ex	alpha_Usa	beta_ir_Ex	beta_ir	alpha_Ir_Ex	alpha_Ir	age
0.151	0.145	-3.995	-3.991	0.314	0.313	-4.034	-4.032	**0**
0.103	0.102	-7.124	-7.121	0.199	0.200	-7.140	-7.141	**1-4**
0.109	0.108	-7.801	-7.807	0.152	0.152	-7.816	-7.822	**5-9**
0.078	0.076	-7.736	-7.730	0.079	0.079	-7.961	-7.969	**10-14**
0.045	0.047	-2.911	-2.917	0.007	0.011	-2.787	-2.778	**70-74**
0.039	0.043	-2.533	-2.544	0.005	0.007	-2.357	-2.352	**75-79**
0.031	0.032	-2.121	-2.128	0.012	0.009	-1.930	-1.922	**80-84**
0.028	0.028	-1.721	-1.730	0.022	0.022	-1.507	-1.492	**85-89**
0.023	0.021	-1.361	-1.371	0.014	0.010	-1.163	-1.156	**90-94**
0.007	0.008	-1.041	-1.063	0.006	0.008	-0.748	-0.735	**95-99**
-0.030	-0.029	-0.931	-0.937	-0.062	-0.063	-0.661	-0.644	**100-104**
-0.064	-0.063	**-1.019**	**-1.032**	**-0.173**	**-0.168**	-0.749	-0.740	**105-109**
-0.062	-0.060	-1.231	-1.258	-0.114	-0.112	-0.704	-0.683	**110+**

Impact on α & β

alpha_Usa_Ex	alpha_Usa	beta_ir_Ex	beta_ir	age
-1.019	-1.032	-0.173	-0.168	105-109

Impact on k(t)

k(t)-Ir-Ex	k(t)-Ir	k(t)-Usa-Ex	k(t)-Usa	year=t
1.277	1.259	0.683	0.698	**60-64**
0.973	0.979	1.303	1.315	**65-69**
0.425	0.432	1.161	1.173	**70-74**
-0.074	-0.067	0.116	0.119	**75-79**
-0.833	-0.835	-1.134	-1.145	**80-84**
-1.768	-1.768	-2.130	-2.160	**85-89**
-0.194	**-0.189**	**-0.187**	**-0.204**	**theta**

Impact on Θ

k(t)-Ir-Ex	k(t)-Ir	k(t)-Usa-Ex	k(t)-Usa	
-0.194	-0.189	-0.187	-0.204	theta

Impact on RMSE

Mortality at age 105	
0.2755	USA
0.2705	USA Ex
1.2750	Ireland
1.0401	Ireland Ex

Improvement 1.8%

Improvement 18.4%

Thank you