

Smoothing and Forecasting Mortality Rates with P-splines

Iain Currie
Heriot Watt University

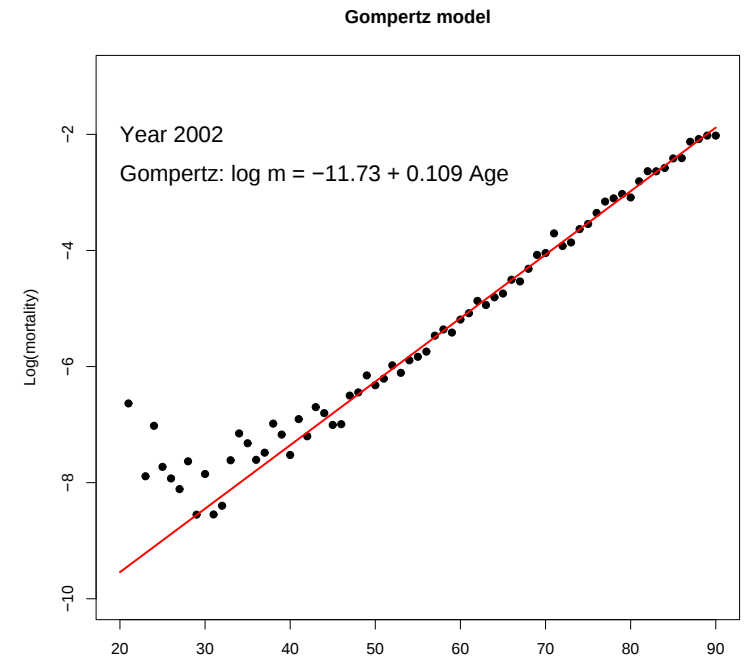
London, April 2007
Manchester, May 2007

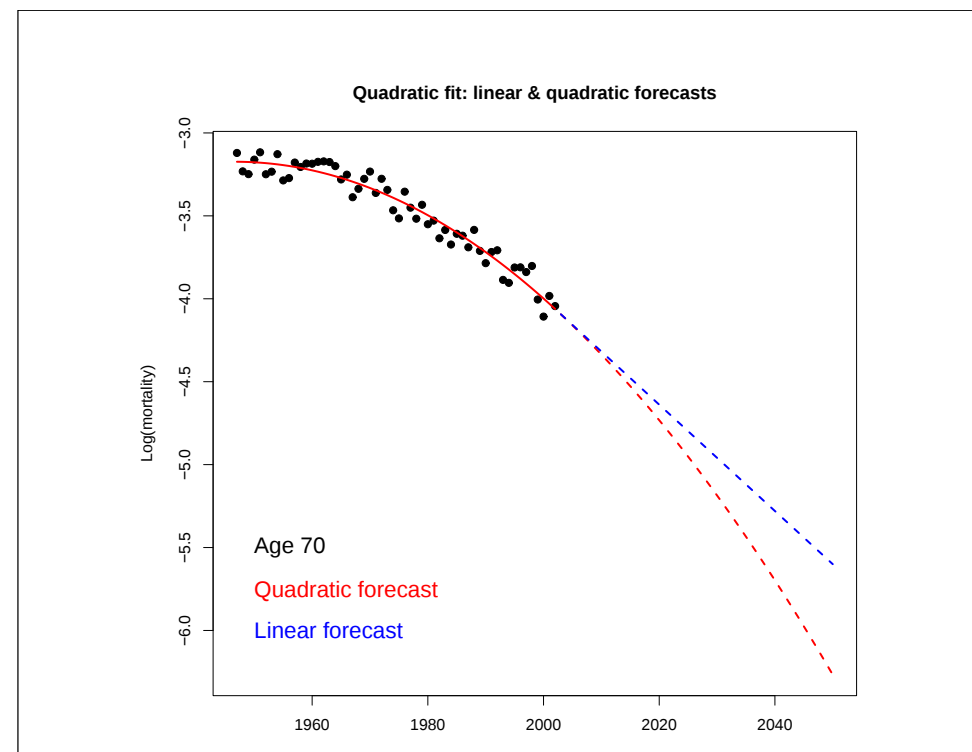
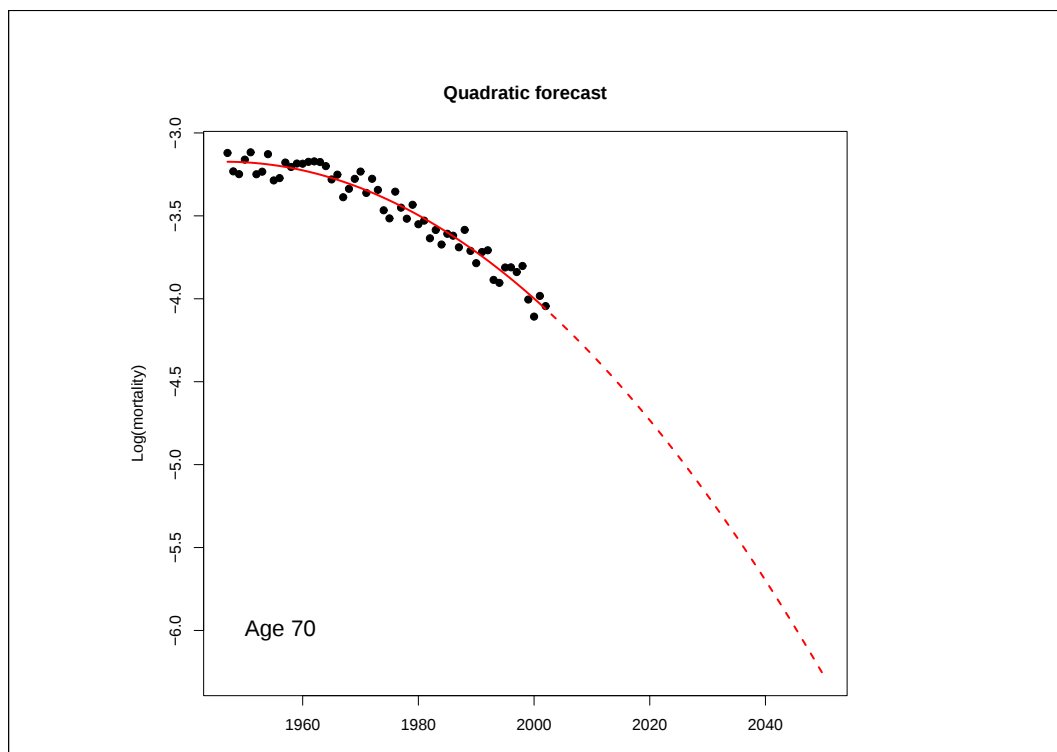
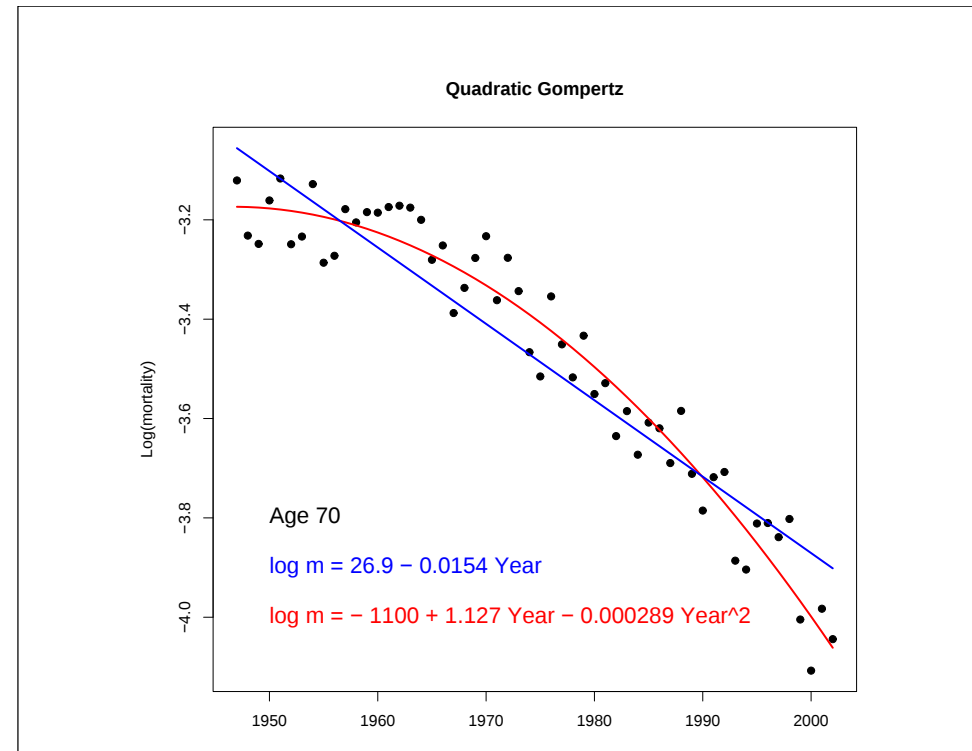
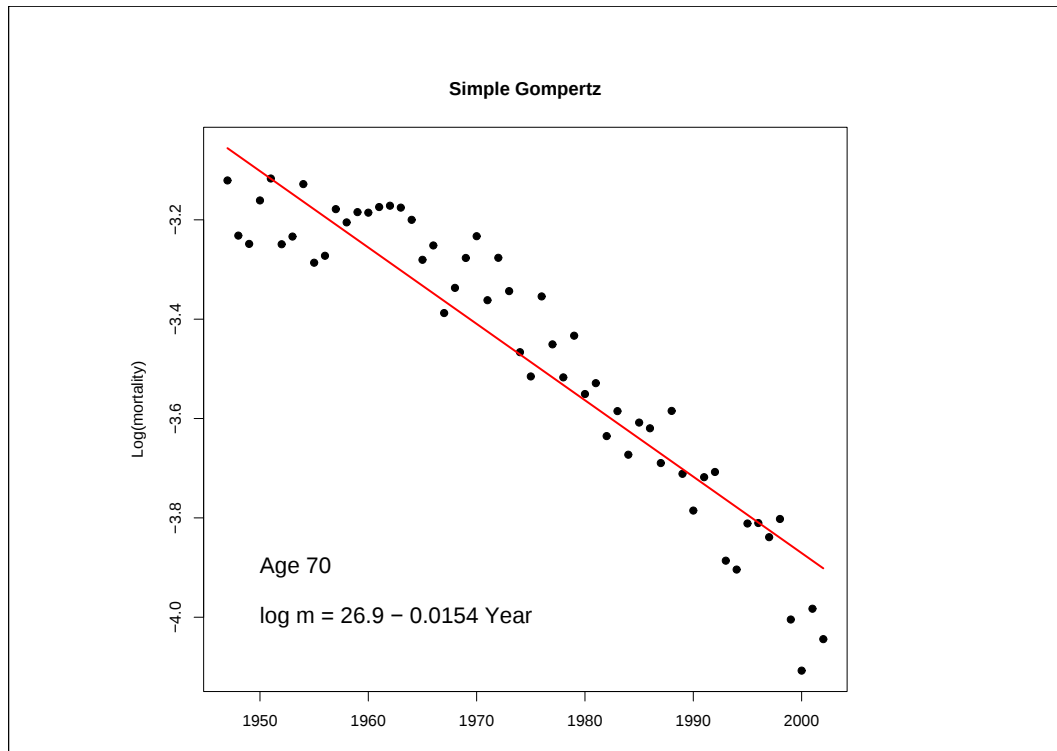
Data and problem

- Data: CMI assured lives
 - Age: 20 to 90
 - Year: 1947 to 2002
- Problem: forecast table to 2046

Plan of talk

- P-splines in 1-dimension
- P-splines in 2-dimensions
 - Lee-Carter model
 - Age-Period-Cohort model
 - 2-d P-splines





Regression basis

In ordinary regression we use basis like

Linear basis: $\{1, x\}$
 Quadratic basis: $\{1, x, x^2\}$
 Polynomial basis: $\{1, x, x^2, \dots, x^p\}$.

Means and fitted values are linear combinations of the basis functions

$$\begin{aligned}\log \mu &= a + bx, & \log \hat{\mu} &= \hat{a} + \hat{b}x \\ \log \mu &= a + bx + cx^2, & \log \hat{\mu} &= \hat{a} + \hat{b}x + \hat{c}x^2 \\ \log \mu &= a + b_1x + \dots + b_px^p, & \log \hat{\mu} &= \hat{a} + \hat{b}_1x + \dots + \hat{b}_px^p\end{aligned}$$

Generalised linear models

Model fitting uses the **Generalised Linear Model** framework.

$$\begin{aligned}d_x &\sim \mathcal{P}(E_x^c \mu_x) \\ \text{and } \log E_x^c \mu_x &= \log E_x^c + a + bx \\ \text{or } \log E_x^c \mu_x &= \log E_x^c + a + bx + cx^2.\end{aligned}$$

B-spline basis

A B-spline regression basis uses **local** basis functions.

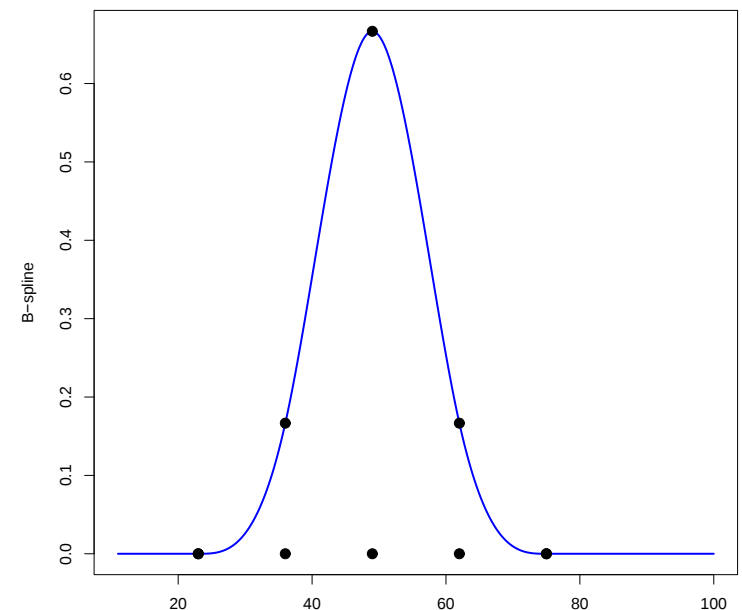
B-spline basis: $\{B_1(x), B_2(x), \dots, B_p(x)\}$

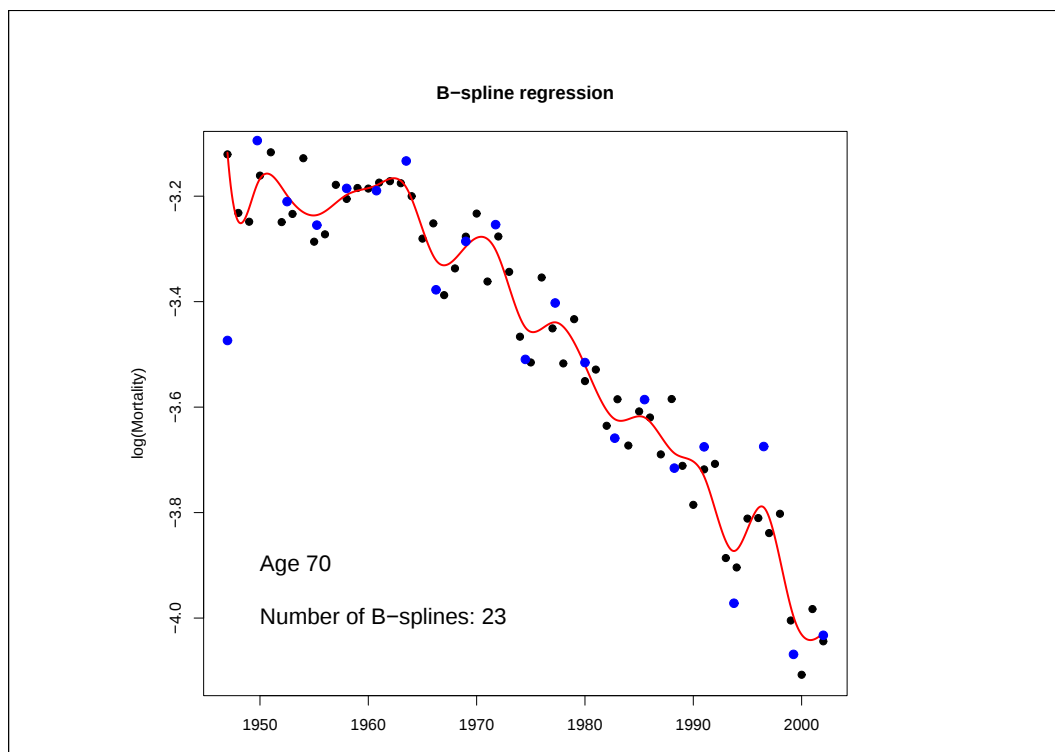
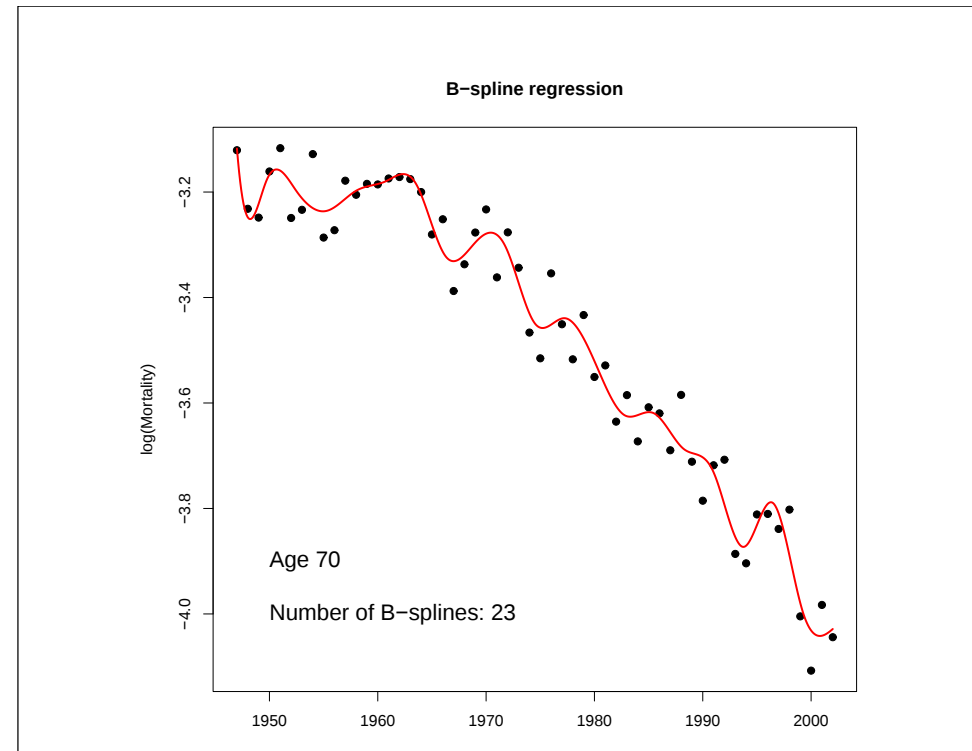
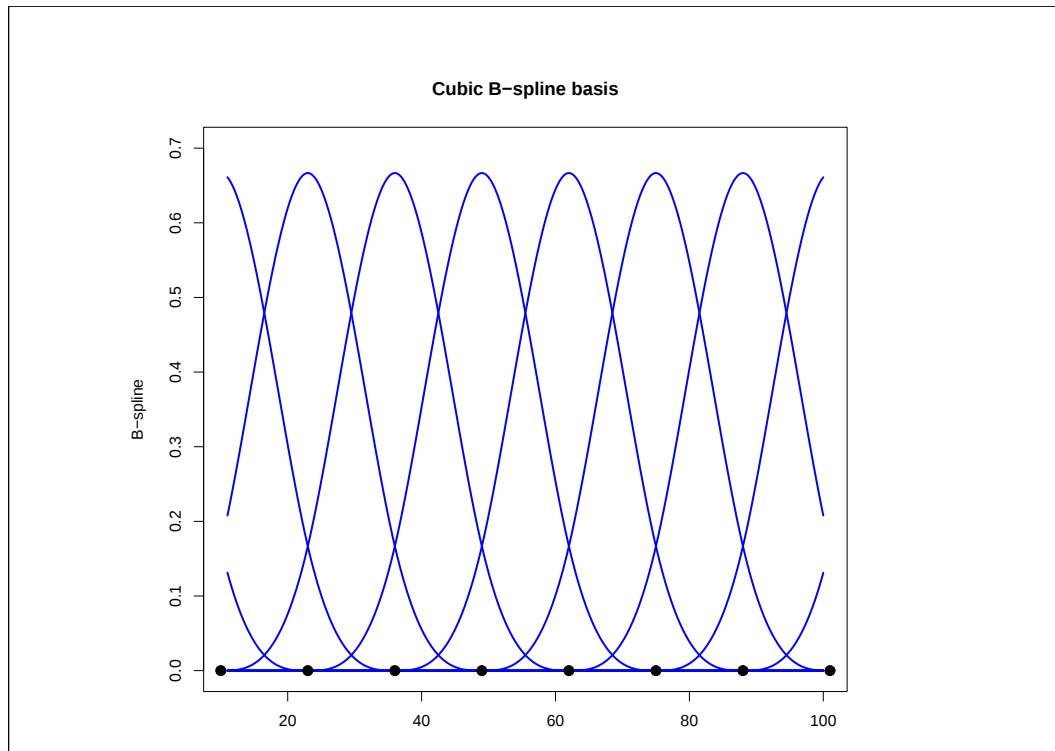
where $B_1(x), B_2(x), \dots, B_p(x)$ are B-splines.

Means and fitted values work the same way

$$\log \mu = \sum_1^p \theta_j B_j(x), \quad \log \hat{\mu} = \sum_1^p \hat{\theta}_j B_j(x)$$

A single cubic B-spline





Penalties

Eilers & Marx (1996) imposed **penalties** on **differences** between adjacent coefficients.

$$\begin{aligned} P_2(\theta) &= (\theta_1 - 2\theta_2 + \theta_3)^2 + \dots + (\theta_{p-2} - 2\theta_{p-1} + \theta_p)^2 \\ &= \theta' D_2' D_2 \theta \end{aligned}$$

where D_2 is a difference matrix.

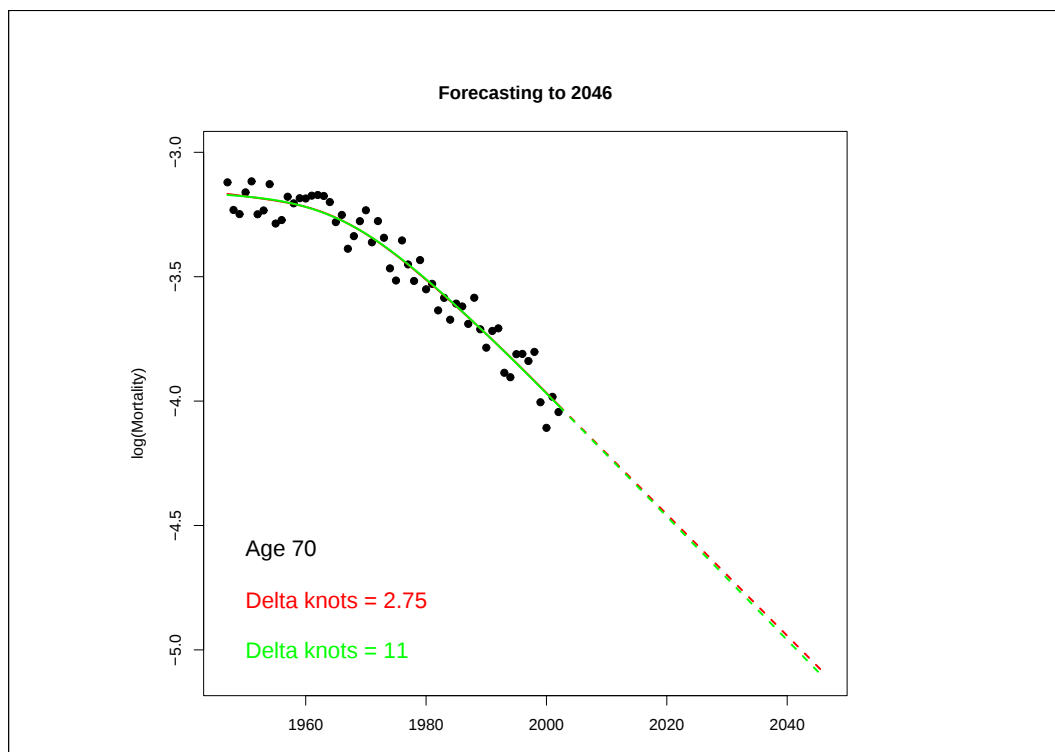
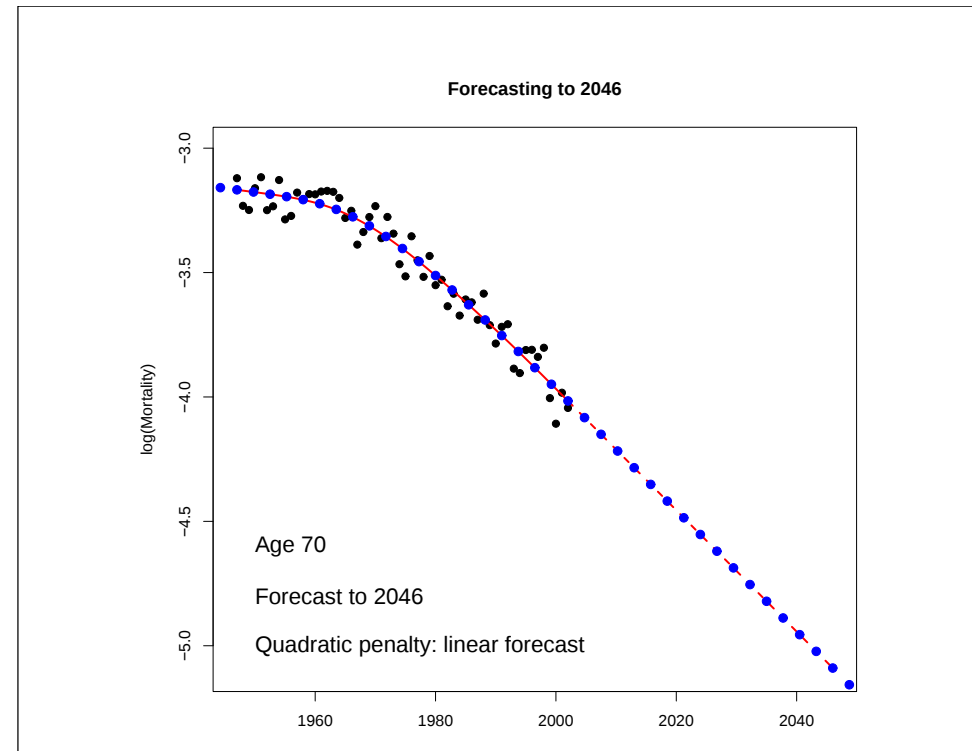
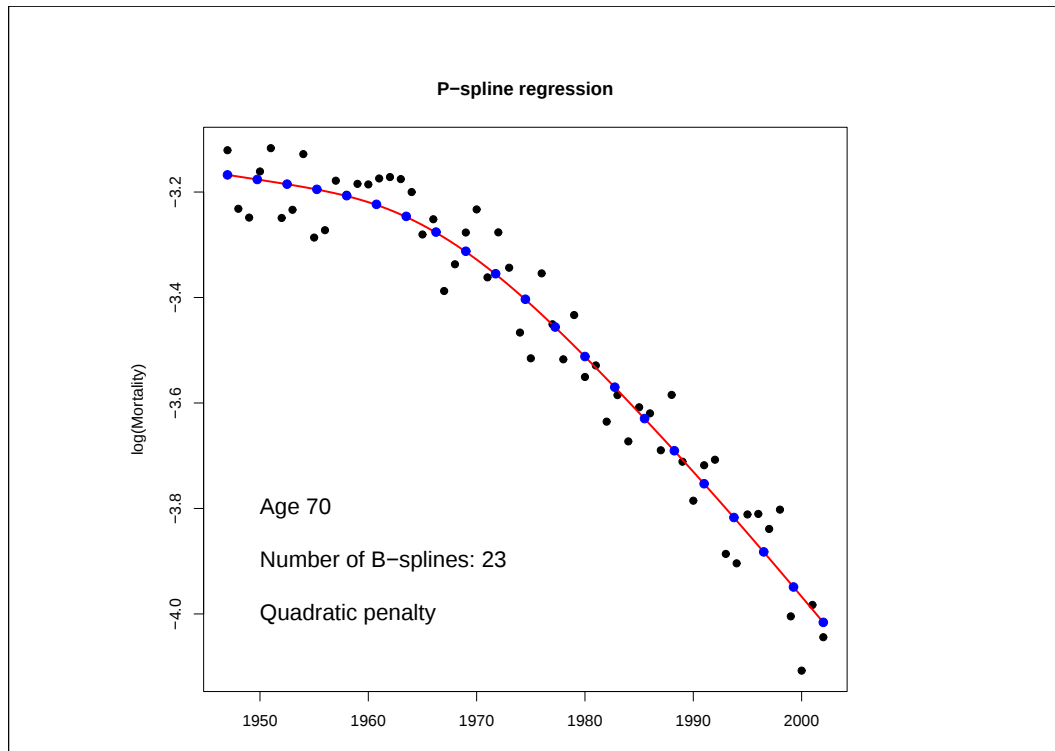
$P_2(\theta)$ is a **roughness penalty**.

Estimation is via **penalised likelihood**

$$PL(\theta) = L(\theta) - \frac{1}{2} \lambda \theta' D_2' D_2 \theta$$

where λ is the **smoothing parameter** which balances **fit** and **smoothness**.

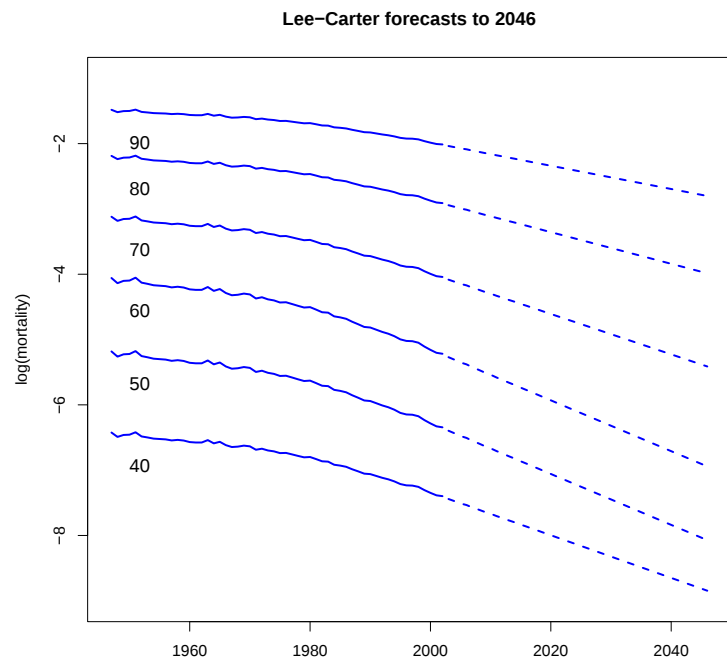
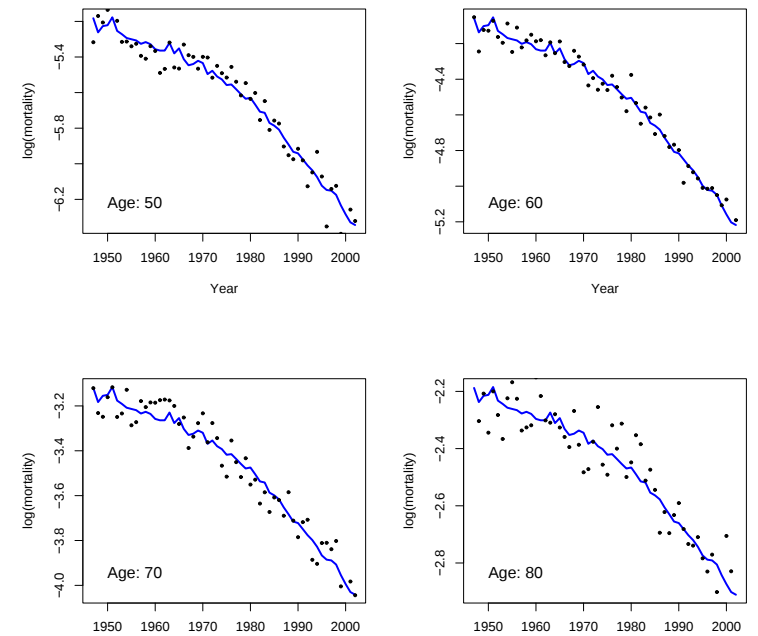
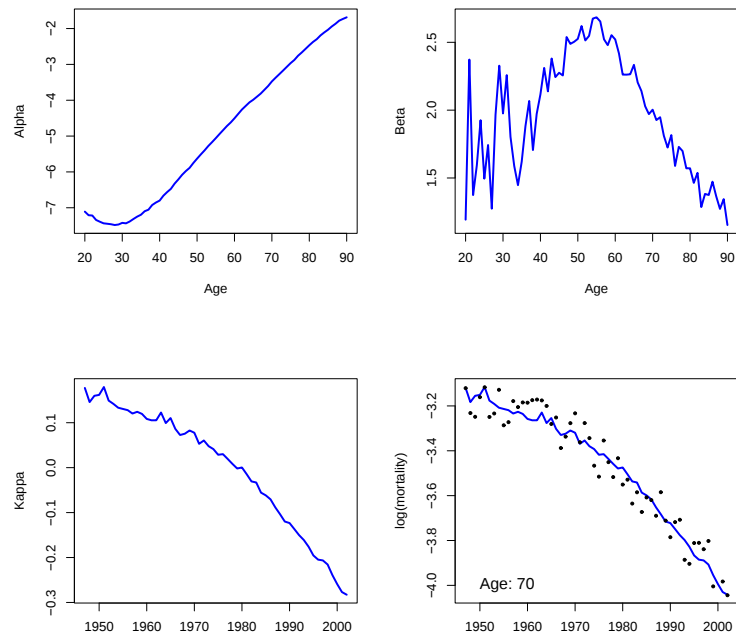
- B-spline regression ($\lambda = 0$)
- Linear regression ($\lambda = \infty$)



Lee-Carter model

Lee & Carter (1992)

$$\log \mu_{ij} = \alpha_i + \beta_i \kappa_j, \quad 20 \leq i \leq 90, \quad 1947 \leq j \leq 2002$$



Discrete Lee-Carter model

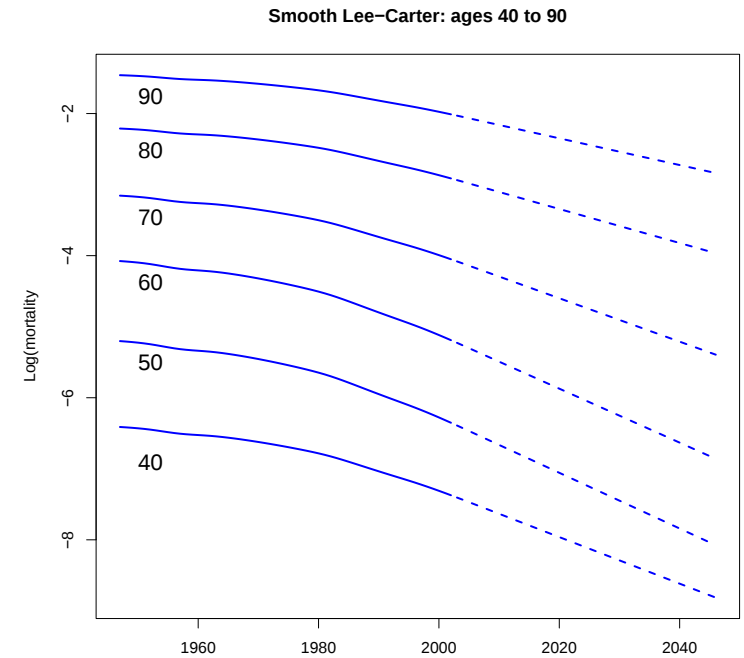
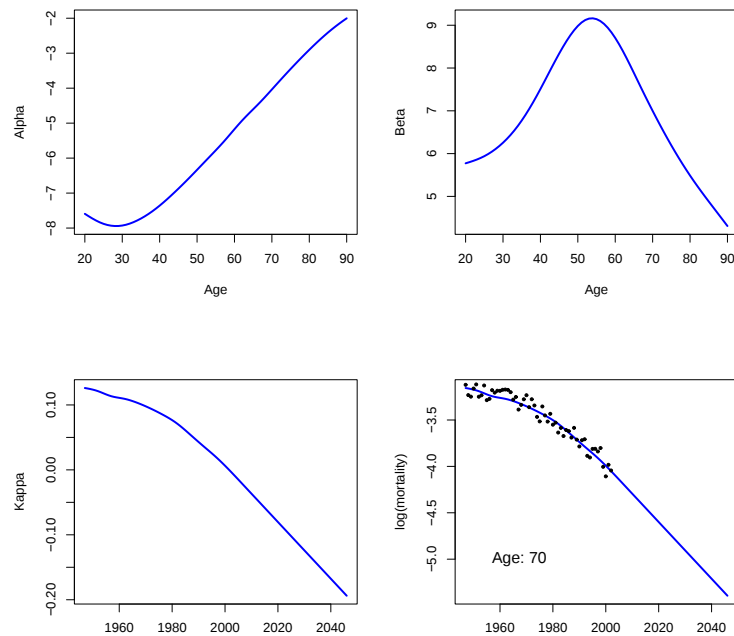
$$\log \mu_{ij} = \alpha_i + \beta_i \kappa_j, \quad 20 \leq i \leq 90, \quad 1947 \leq j \leq 2002$$

or in table form

$$\log M = \alpha 1' + \beta \kappa'$$

Smooth Lee-Carter

$$\alpha \rightarrow B_a a, \quad \beta \rightarrow B_a b, \quad \kappa \rightarrow B_y k$$

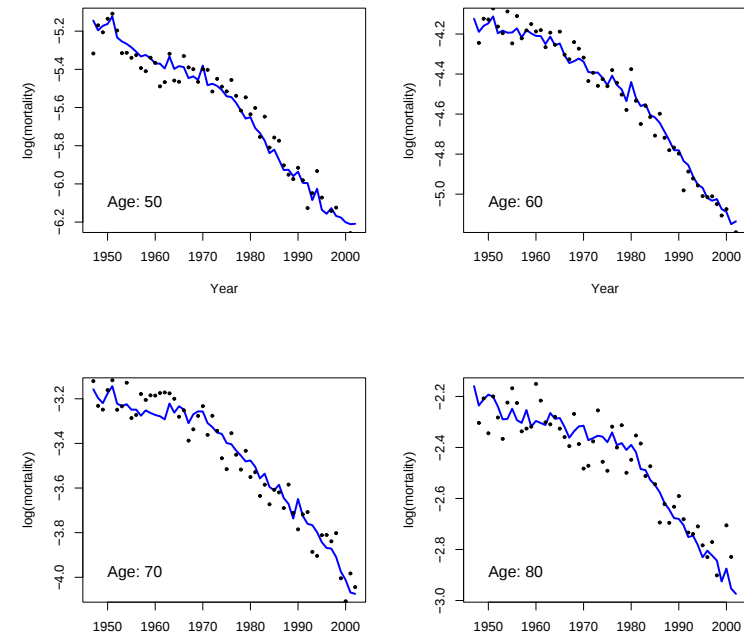


Age-Period-Cohort model

$$\log \mu_{ij} = \alpha_i + \beta_j + \gamma_{i-j}, \quad 20 \leq i \leq 90, \quad 1947 \leq j \leq 2002$$

Parameter redundancy

- $2n_a + 2n_y - 1 = 253$ parameters
- $2n_a + 2n_y - 4 = 250$ free parameters

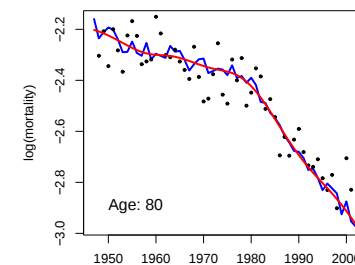
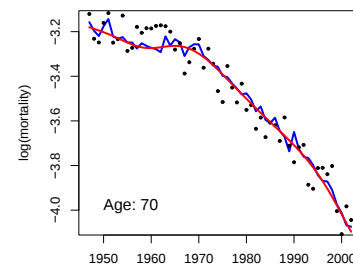
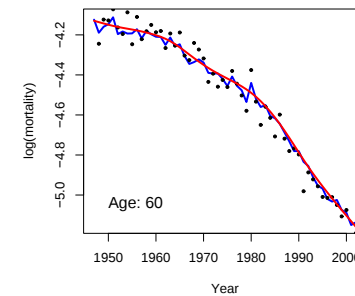
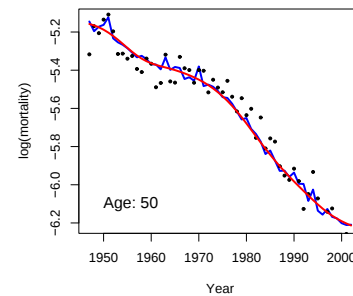


Discrete Age-Period-Cohort model

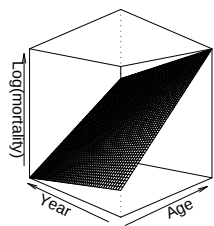
$$\log \mu_{ij} = \alpha_i + \beta_j + \gamma_{i-j}, \quad 20 \leq i \leq 90, 1947 \leq j \leq 2002$$

Smooth Age-Period-Cohort model

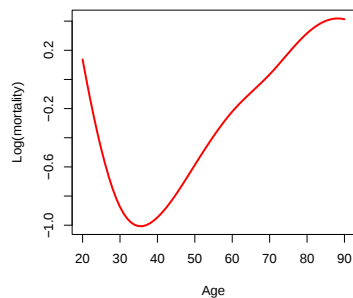
$$\alpha \rightarrow B_a a, \quad \beta \rightarrow B_y b, \quad \gamma \rightarrow B_c c$$



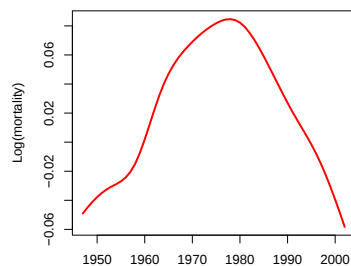
Planar component



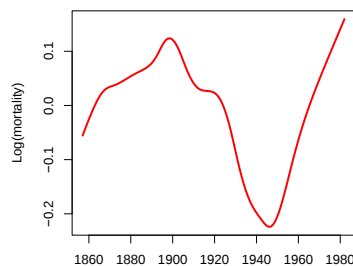
Age component



Year component



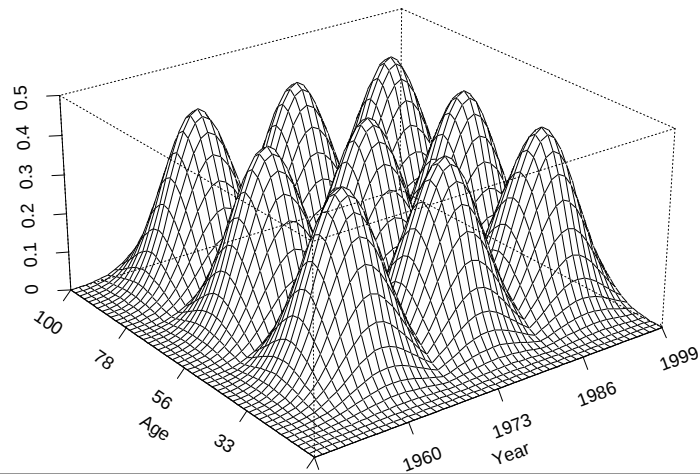
Cohort component



2-d modelling of mortality tables

- Let B_a , 71×13 , be a 1- d B-spline basis for modelling a **single age**.
- Let B_y , 56×10 , be a 1- d B-spline basis for modelling a **single year**.

Can we combine the **marginal** bases B_a and B_y to make a 2-d basis?



2-d modelling of mortality tables

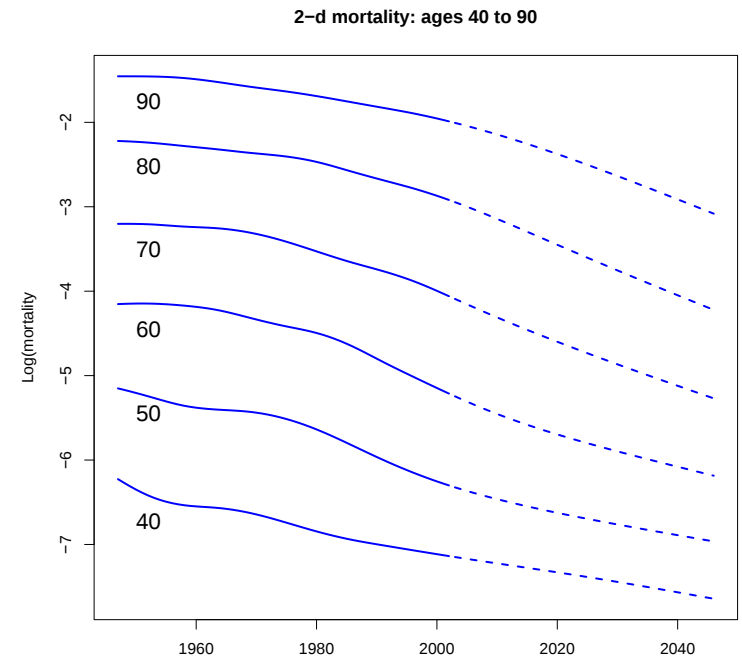
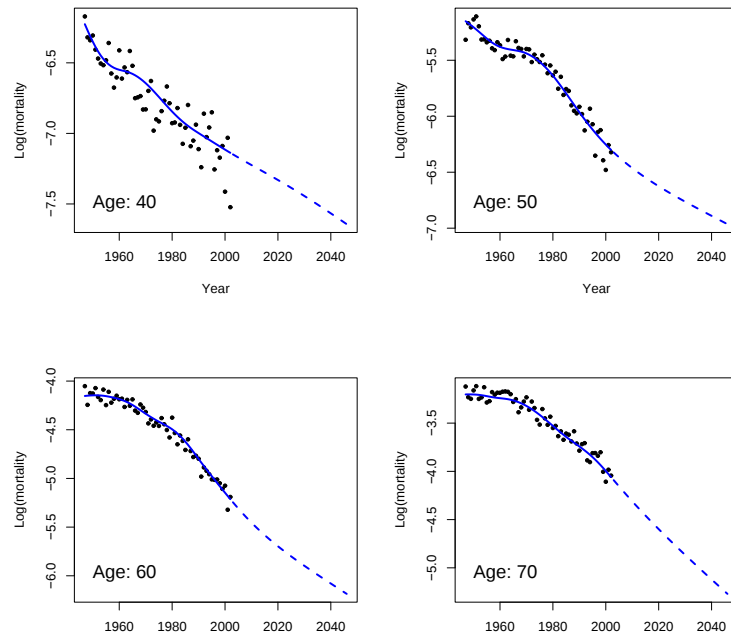
Regression matrix

The Kronecker product organises the multiplication of the marginal bases to give the regression matrix

$$B = B_y \otimes B_a, 3976 \times 130.$$

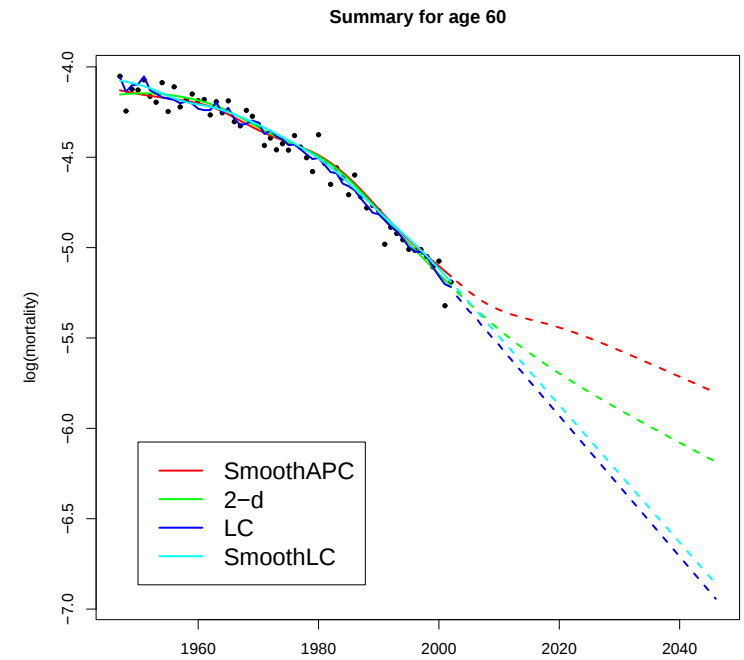
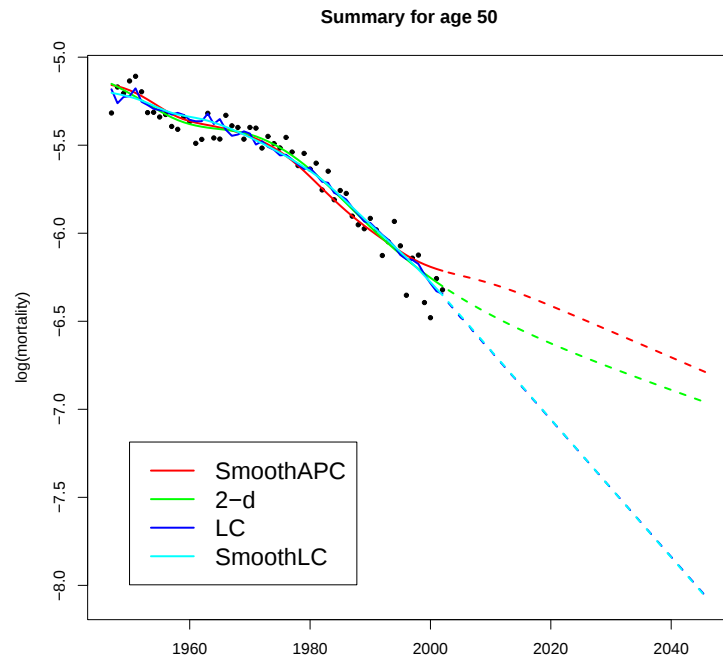
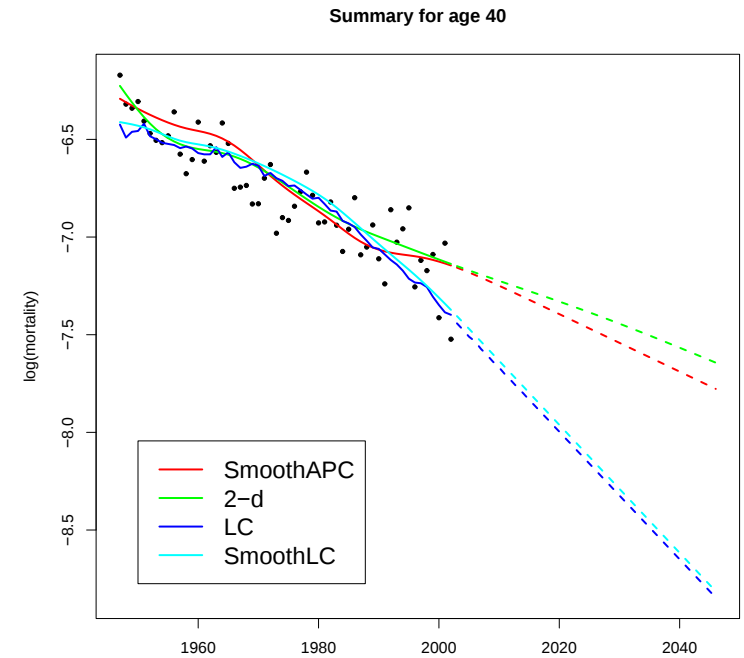
Penalties in 2-d

- Each regression coefficient is associated with the summit of one of the hills.
- Smoothness is ensured by penalizing the coefficients in rows and columns.

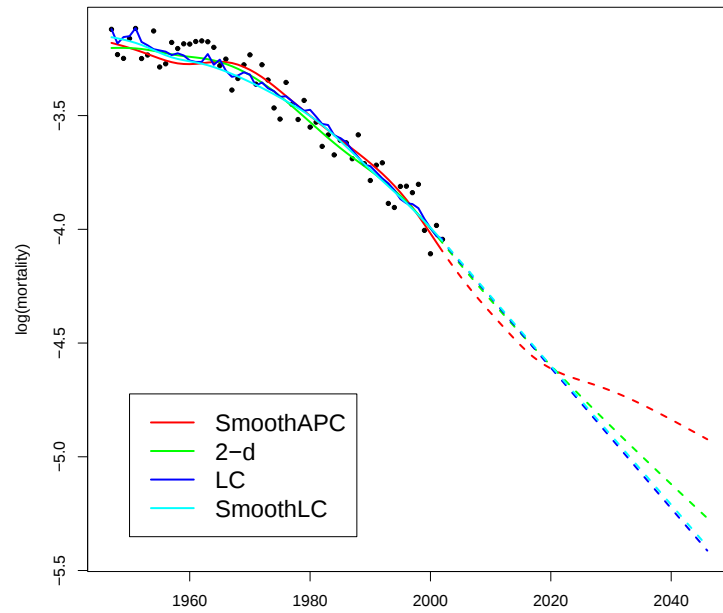


Summary of results

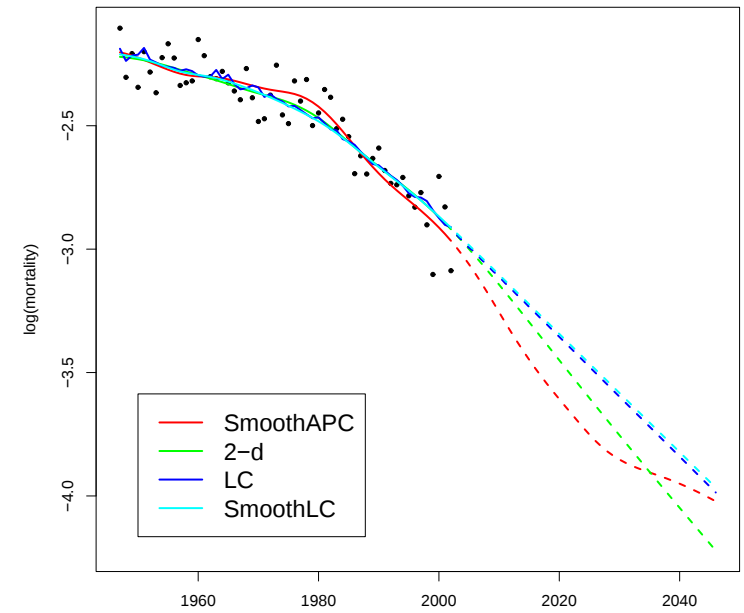
	DEV	TR	BIC	Rank
2-d	6832	63	7357	1
APC smooth	7884	30	8129	2
LC smooth	8568	30	8815	3
APC	7002	248	9057	4
LC	8004	196	9629	5



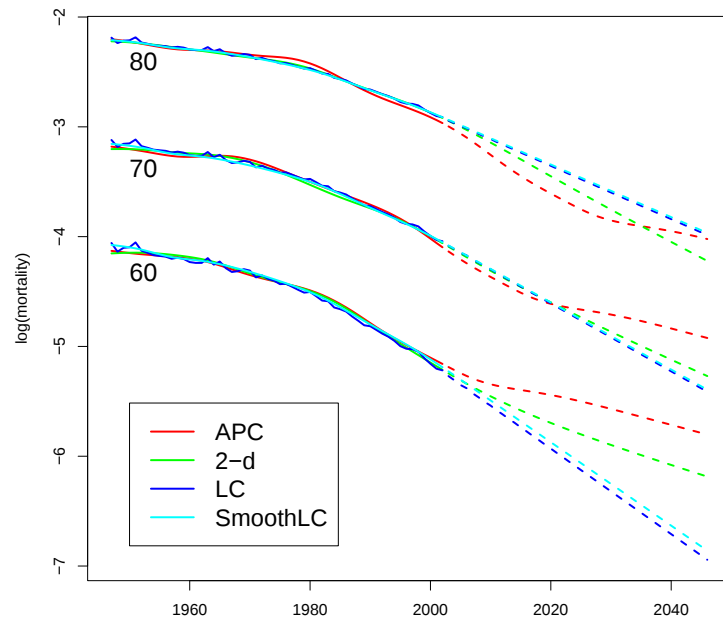
Summary for age 70



Summary for age 80



Summary for ages 60, 70, 80



Conclusions

Forecasting a mortality table depends on

- Model choice
- Forecasting method
- Parameter uncertainty
- Stochastic uncertainty

References

Lee-Carter models

Lee & Carter (1992) J American Statistical Association, 87, 659-675.

Brouhns, Denuit & Vermunt (2002) Insurance: Mathematics & Economics, 31, 373-393.

Age-Period-Cohort models

Clayton & Schliffers (1987) Statistics in Medicine, 6, 449-467.

Clayton & Schliffers (1987) Statistics in Medicine, 6, 469-481.

Penalized spline models

Eilers & Marx (1996) Statistical Science, 11, 758-783.

Currie, Durban & Eilers (2004) Statistical Modelling, 4, 279-298.

Currie, Durban & Eilers (2006) Journal of the Royal Statistical Society, Series B, 68, 259-280.