The Age-Period-Cohort (APC) Model

Iain Currie

Heriot-Watt University

&

LONGEVITAS

London
19th April 2012

Identifiability

The Lee-Carter model

$$\log(\mu_{i,j}) = \alpha_i + \beta_i \kappa_j$$

is not identifiable since, for example,

- Location: $\kappa_j \longrightarrow \kappa_j - \delta_1$ and $\alpha_i \longrightarrow \alpha_i + \beta_i \delta_1$ leaves the value of $\log(\mu_{i,j})$ unchanged. Similarly,
- Scale: $\kappa_j \longrightarrow \delta_2 \kappa_j$ and $\beta_i \longrightarrow \beta_i / \delta_2$ also leaves $\log(\mu_{i,j})$ unchanged.

We need two identifiability constraints to fix a particular set of parameters. But how to choose them??

Choosing identifiability constraints

There are two issues

- Fitting: All choices give the same fitted values of $\log(\mu_{i,j})$. The choice of constraints is irrelevant.
- Forecasting: Forecasting with the Lee-Carter model depends on forecasting the estimated values of $\boldsymbol{\kappa}$.

The scale constraint, $\sum \beta_i = 1$, has no impact on the forecast but the location constraint most certainly does. We must be sure that our choice, $\sum \kappa_j = 0$, is a good one.

Good identifiability constraints

Here are three reasons why $\sum \kappa_j = 0$ is a good choice:

- Interpretability
- Stability
- Canonical correlation

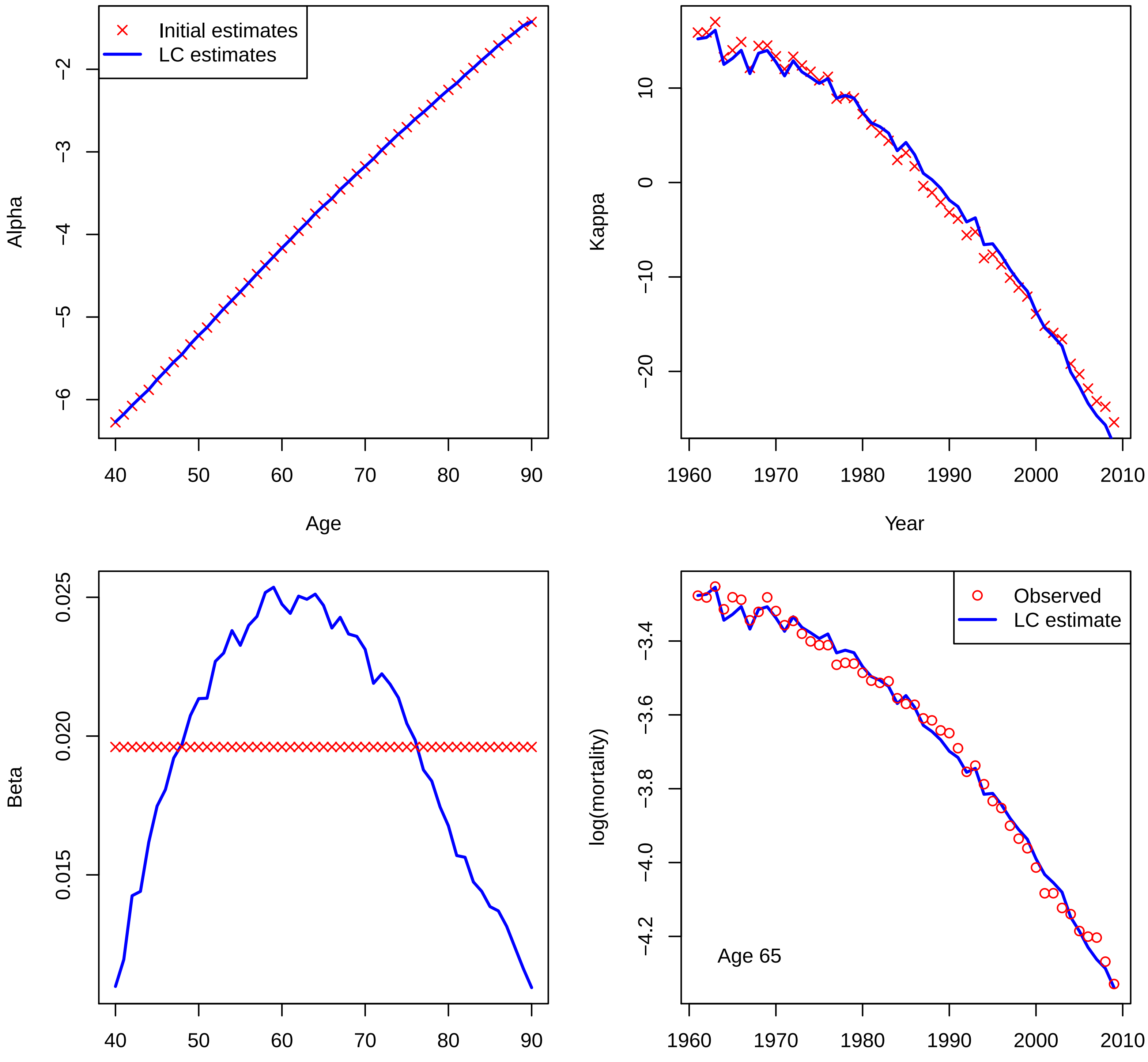
Interpretability

The age parameter $\hat{\alpha}_i$ is roughly average mortality at age i for all ages.

$$\begin{aligned}\sum_j \log(\hat{\mu}_{i,j}) &= \sum_j (\hat{\alpha}_i + \hat{\beta}_i \hat{\kappa}_j) \\ &= n_y \hat{\alpha}_i + \hat{\beta}_i \sum \hat{\kappa}_j \\ &= n_y \hat{\alpha}_i\end{aligned}$$

where n_y is the number of years.

Similary, the period parameter $\hat{\kappa}_j$ is roughly average mortality (centered) in period j in all periods. (next slide).



Initial estimates
LC estimates
Alpha
Age
Kappa
Year
Beta
Observed
LC estimate
log(mortality)
Age 65

Stability

Introduction of β in the LC model has little effect on the initial estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\kappa}$ (previous slide).

Canonical correlation

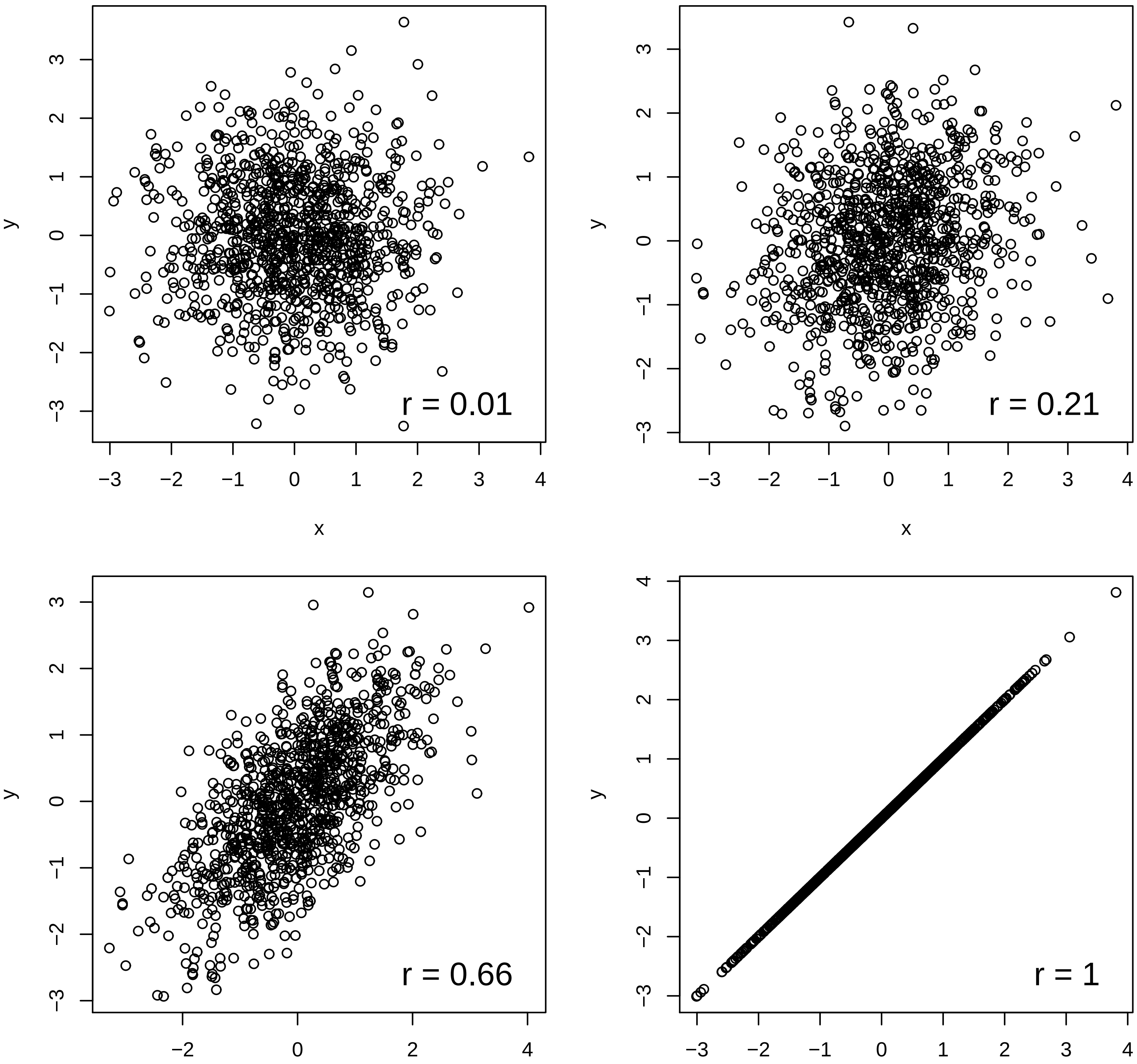
Canonical correlation measures the strength of the linear relation between two random vectors. We seek vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ such that the correlation between $\boldsymbol{a}'\hat{\boldsymbol{\alpha}}$ and $\boldsymbol{b}'\hat{\boldsymbol{\kappa}}$ is maximized. This maximized correlation is the first canonical correlation, $r(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\kappa}})$.

We find

$$r(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\kappa}}) = 0.21,$$

a reassuringly low number.

Reference: Mardia, Kent & Bibby (1979) Chap 10.


r = 0.01
r = 0.21
r = 0.66
r = 1
x
y

Summary for LC model

- Interpretability: $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\kappa}}$ are approximately row and columns means of data matrix.
- Stability: Introduction of $\hat{\boldsymbol{\beta}}$ has little effect on $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\kappa}}$.
- Canonical correlation: $r(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\kappa}}) = 0.21$ is low.

Age-Period-Cohort model

The APC model is defined

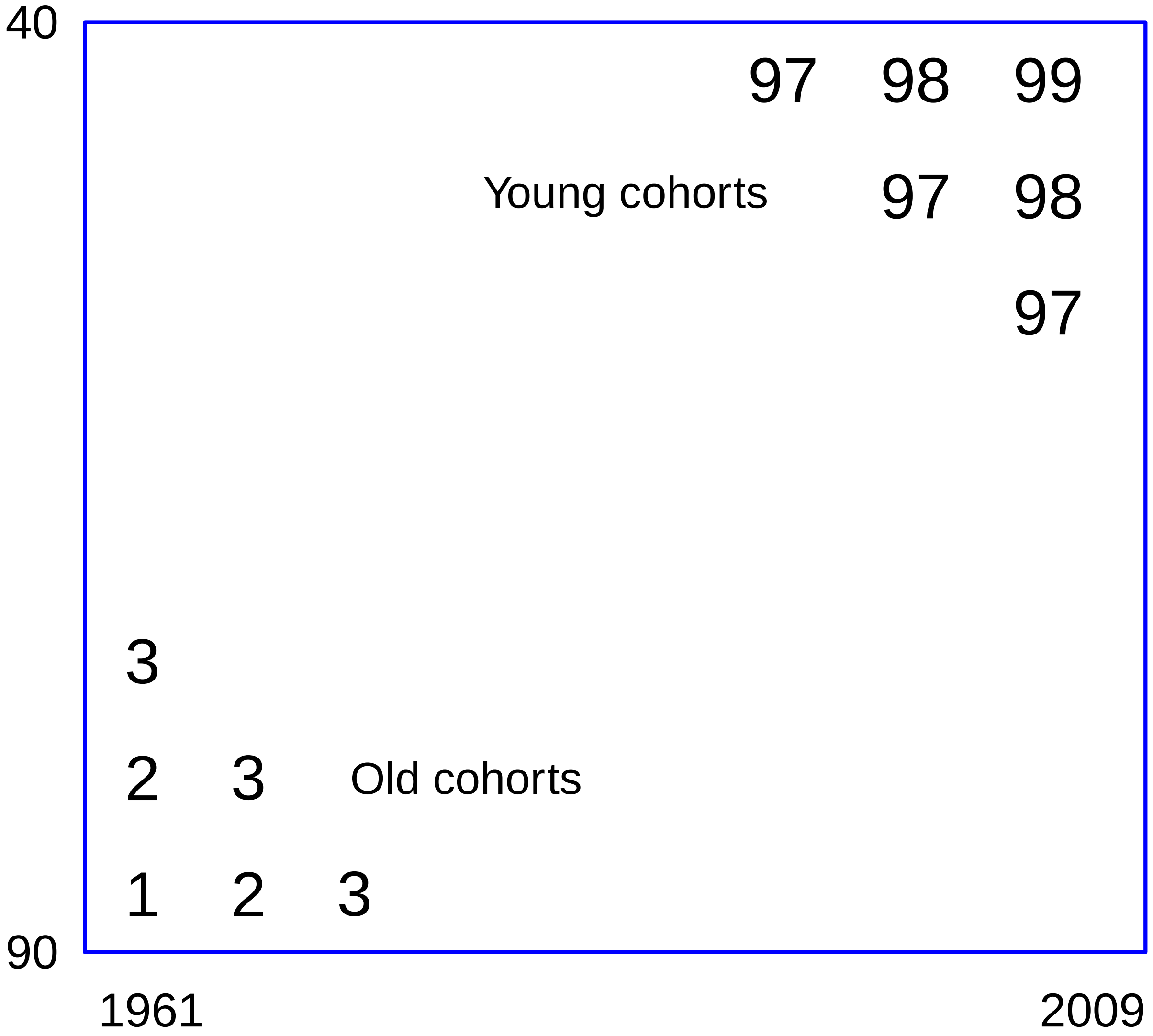
$$\log(\mu_{i,j}) = \alpha_i + \kappa_j + \gamma_{j-i}.$$

- We have a generalized linear model or GLM.
- However, the model is not identifiable and needs three location constraints to fix coefficients.
- The choice of constraints is critical since the idea is to forecast $\hat{\boldsymbol{\kappa}}$ and $\hat{\boldsymbol{\gamma}}$ with $\boldsymbol{\alpha}$ fixed at $\hat{\boldsymbol{\alpha}}$.

Notation

- Let $n_x = 51$ be the number of ages, $n_y = 49$ be the number of years and $n_c = n_x + n_y - 1 = 99$ be the number of cohorts.
- Number the cohorts from 1 (oldest) to n_c (youngest). Let w_c be the number of times cohort c appears. Thus $w_1 = 1$, $w_2 = 2$, $w_3 = 3$, $\ldots$, $w_{n_c-1} = 2$, $w_{n_c} = 1$

UK male mortality data

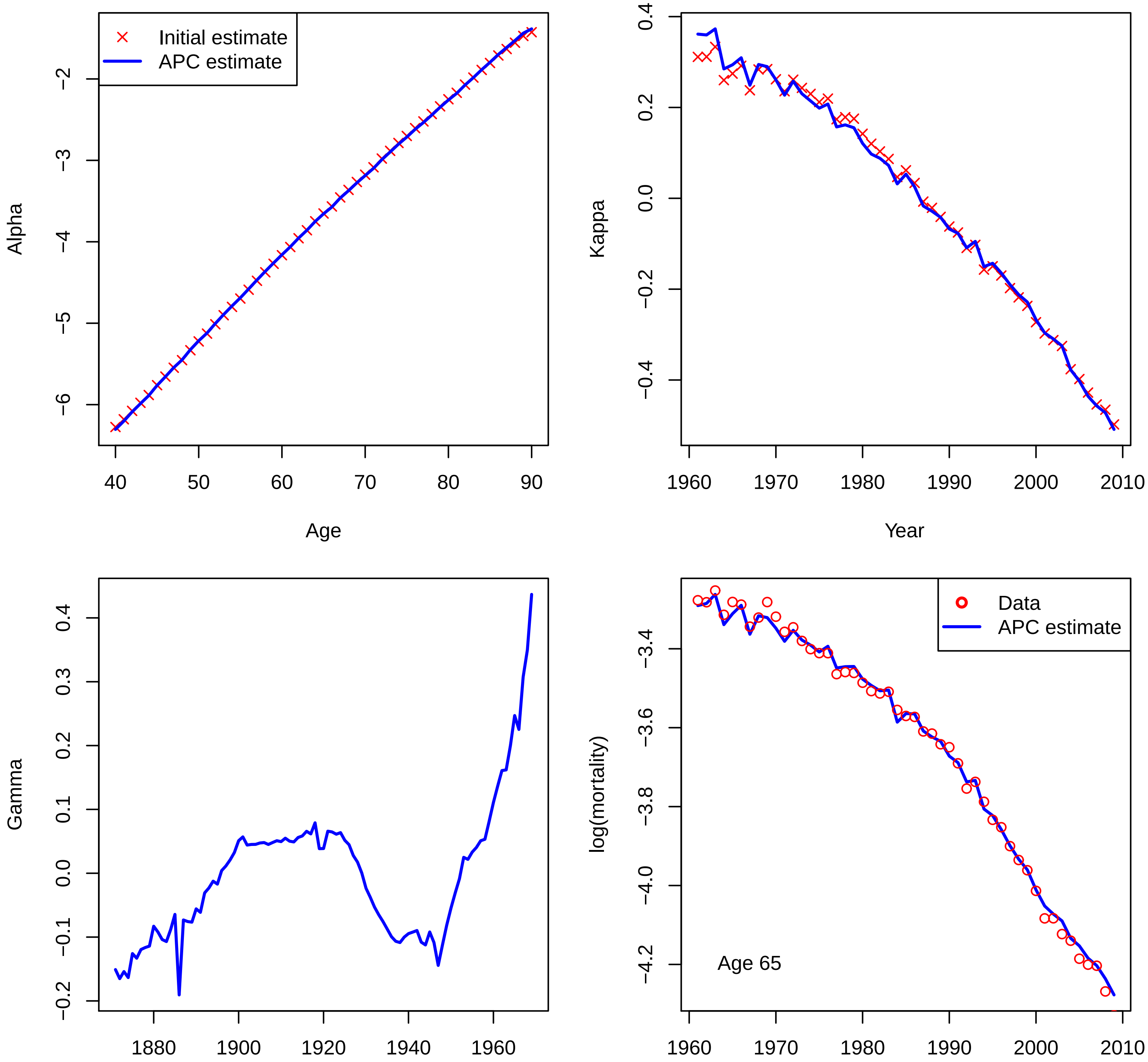


Cairns et al (2009) suggest constraints

$$\sum \kappa_j = 0, \quad \sum w_c \gamma_c = 0, \quad \sum c w_c \gamma_c = 0,$$

with rationale

- first constraint, on average the period term is zero (as in the Lee-Carter model);
- second constraint, on average the cohort term is zero (over the whole table);
- third constraint, the cohort term has slope zero (over the whole table).



Initial estimate
APC estimate
Alpha
Age
Kappa
Year
Gamma
Data
APC estimate
log(mortality)
Age 65

Comments

- Note the large values of the recent γ_c. These only make sense when linked with the large negative κ_j values. There is a strong link between $\hat{\boldsymbol{\kappa}}$ and $\hat{\boldsymbol{\gamma}}$.
- Canonical correlations: We find

$$r(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\kappa}}) = 0.44, \quad r(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\gamma}}) = 0.56, \quad r(\hat{\boldsymbol{\kappa}}, \hat{\boldsymbol{\gamma}}) = 0.66.$$

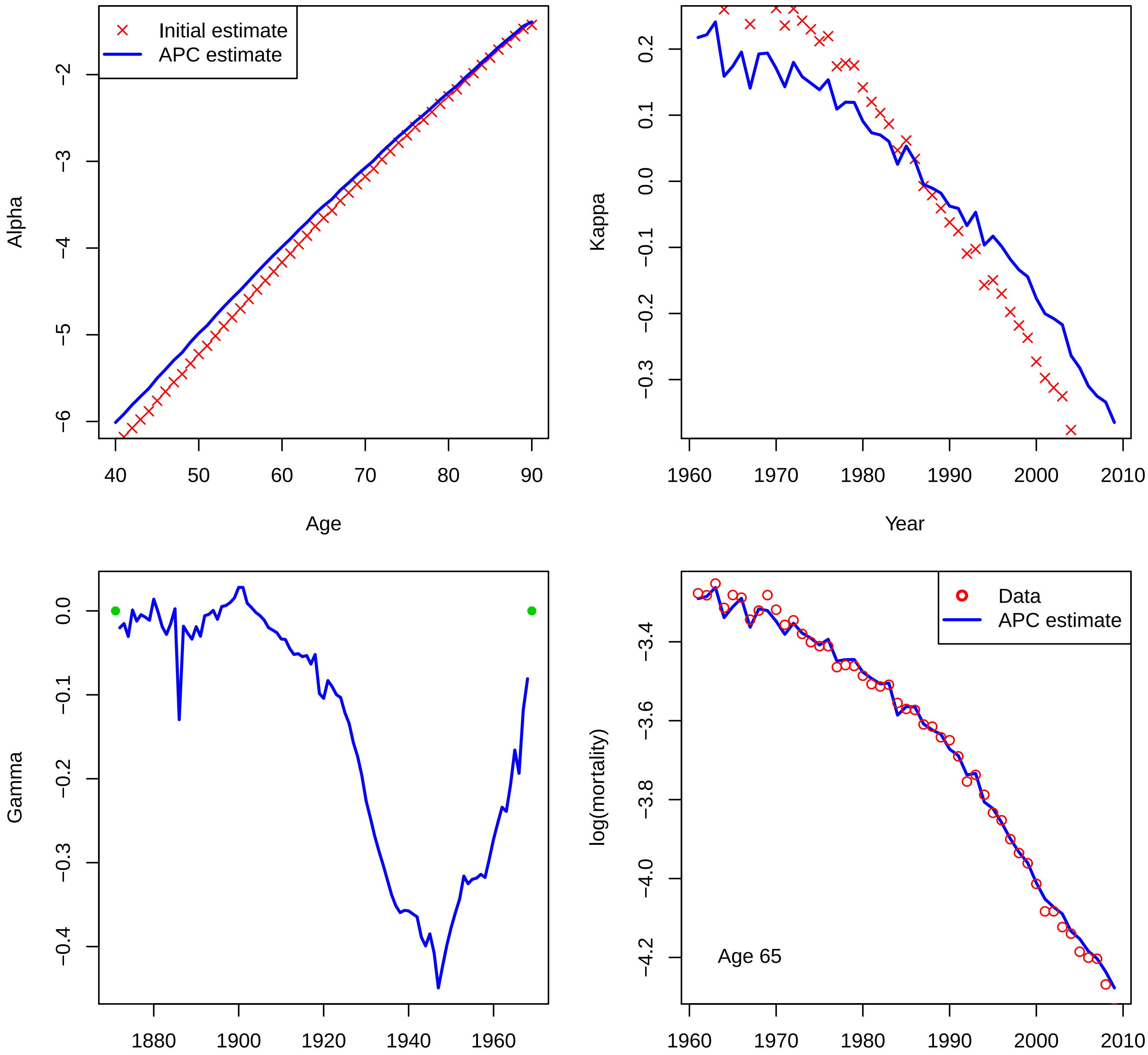
- The standard assumption is to assume that $\hat{\boldsymbol{\kappa}}$ and $\hat{\boldsymbol{\gamma}}$ can be forecast independently (Renshaw & Haberman (2006), Cairns *et al* (2011)) but
- $\hat{\boldsymbol{\kappa}}$ and $\hat{\boldsymbol{\gamma}}$ independent $\Rightarrow r(\hat{\boldsymbol{\kappa}}, \hat{\boldsymbol{\gamma}}) = 0.$

Other constraints

Cairns et al (2009) noted that old and young γ_c were poorly or very poorly determined. To avoid distortion they suggested dropping these parameters. Actually they suggested dropping the old and young data but we prefer to drop the problem cohort parameters.

$$\sum \kappa_j = 0, \quad \gamma_1 = \gamma_{n_c} = 0.$$

The model is now identifiable!



Initial estimate
APC estimate
Alpha
Age
Kappa
Year
Gamma
Data
APC estimate
log(mortality)
Age 65

Comments

- The fitted values with the constraints

  (a) $\sum \kappa_j = 0, \quad \sum w_c\gamma_c = 0, \quad \sum c w_c\gamma_c = 0$,

  (b) $\sum \kappa_j = 0, \quad \gamma_1 = \gamma_{n_c} = 0$

  are identical.

- Forecasting with

  $\sum \kappa_j = 0, \quad \gamma_1 = \gamma_{n_c} = 0$ is out of the question.

- Canonical correlations: We find

$$r(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\kappa}}) = 1.00, \quad r(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\gamma}}) = 1.00, \quad r(\hat{\boldsymbol{\kappa}}, \hat{\boldsymbol{\gamma}}) = 1.00.$$

Yet more constraints

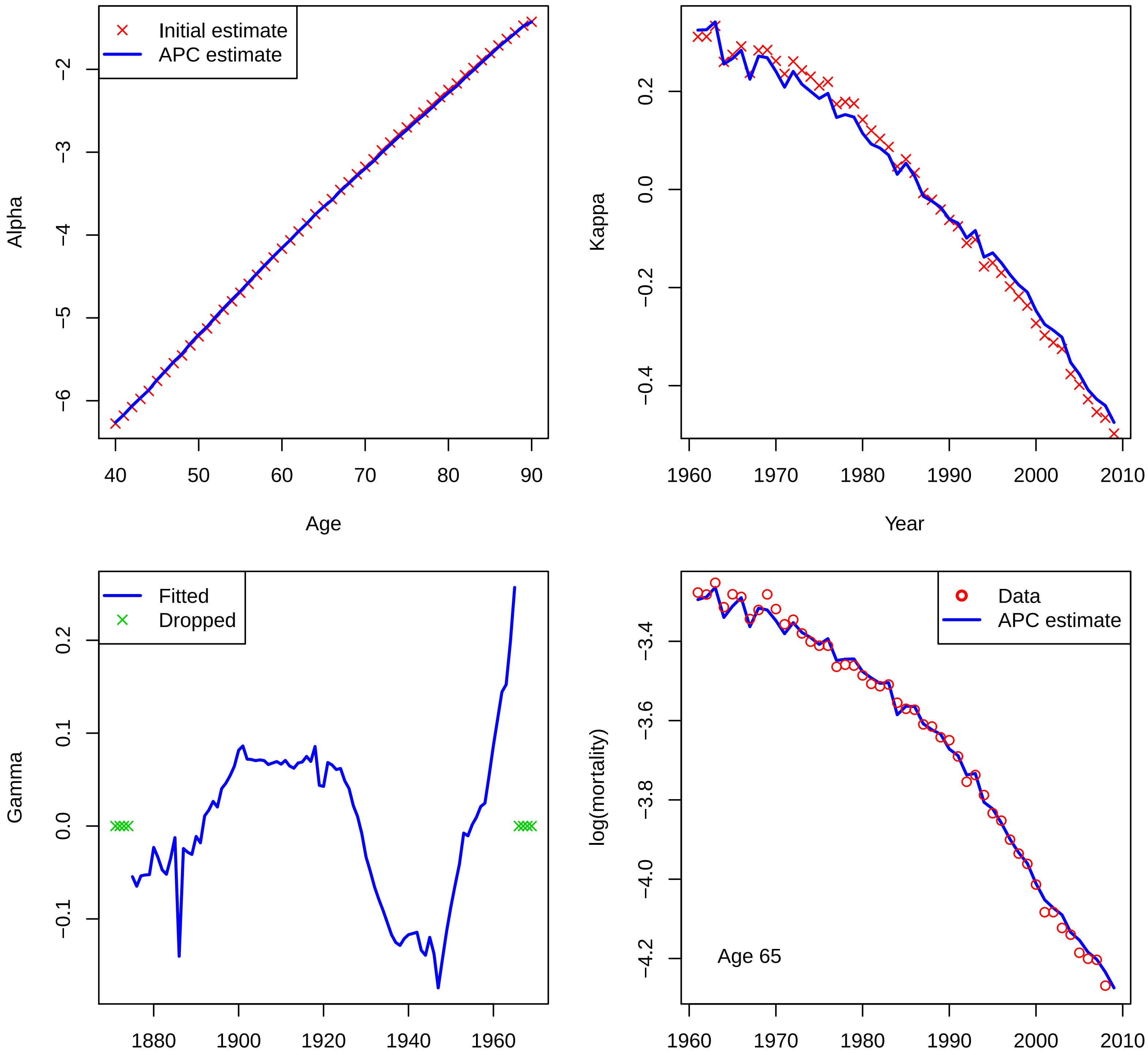
Cairns et al (2009) suggested for their model M7

$$\sum \gamma_c = \sum c\gamma_c = \sum c^2\gamma_c = 0.$$

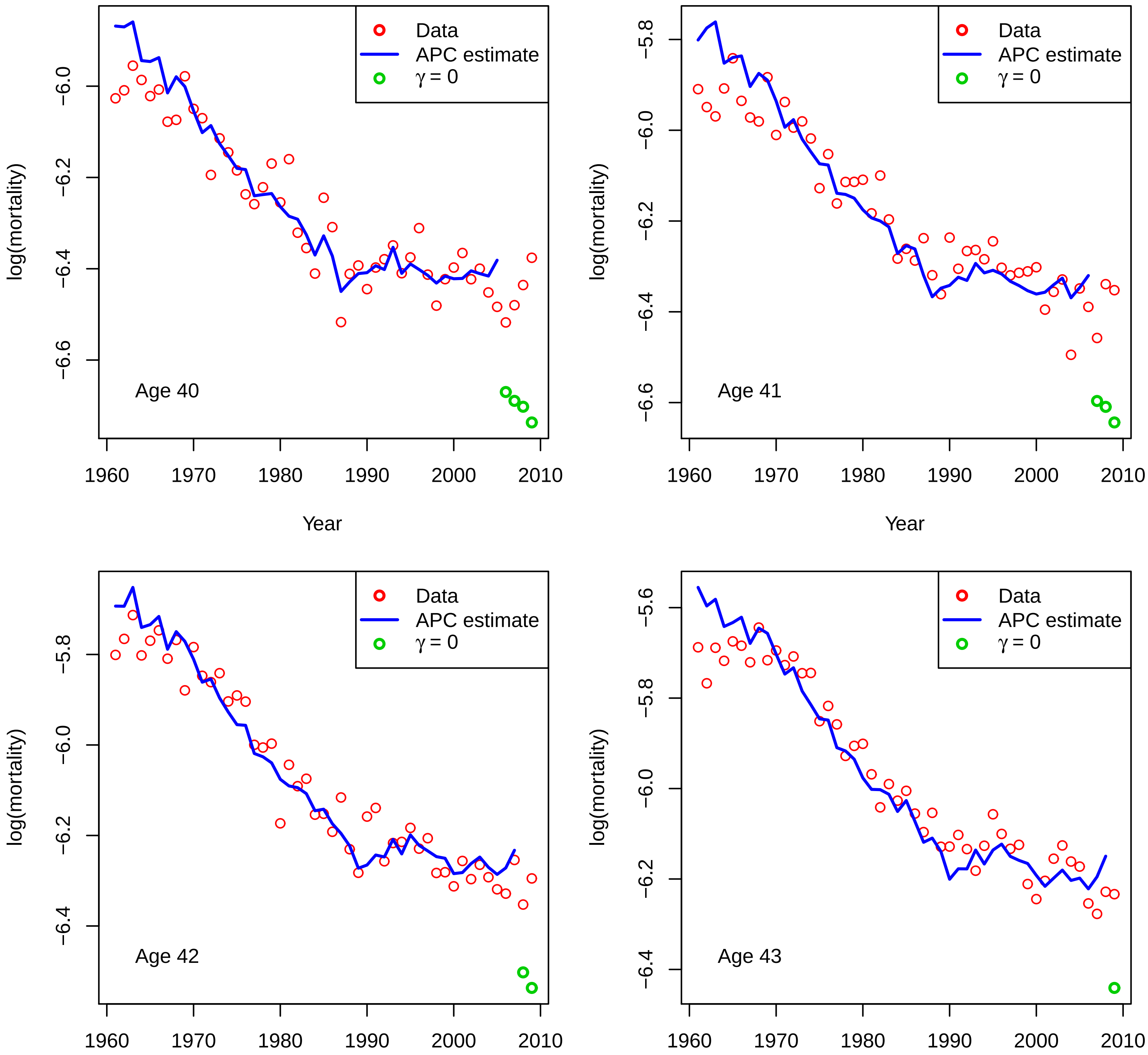
This has the interesting property that in the quadratic regression of the $\hat{\gamma}$ on c all three coefficients are zero. This has a certain appeal when it comes to forecasting. We add the constraints which centre the period effect and constrain the four oldest and youngest cohort coefficients to be zero.

$$\sum \kappa_j = 0, \quad \gamma_c = 0, \quad c = 1, \ldots, 4, n_c - 3, \ldots, n_c.$$

We have a total of twelve (12) constraints. The idea is to sacrifice fit in return for coefficients with better forecast properties.



Initial estimate
APC estimate
Alpha
Age
Kappa
Year
Fitted
Dropped
Gamma
Data
APC estimate
log(mortality)
Age 65



Data
APC estimate
γ = 0
Age 40
Age 41
Age 42
Age 43
log(mortality)
Year
1960
1970
1980
1990
2000
2010
−6.0
−6.2
−6.4
−6.6
−5.8
−5.6

Comments

- Forecasting with these twelve constraints looks problematic.
- Canonical correlations: We find

$$r(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\kappa}}) = 0.43, \quad r(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\gamma}}) = 0.63, \quad r(\hat{\boldsymbol{\kappa}}, \hat{\boldsymbol{\gamma}}) = 0.66.$$

The introduction of further constraints has raised the canonical correlations slightly overall.

Summary

- Identification constraints have no effect on the fitted values of $\log \hat{\mu}_{i,j}$; these are model invariants.
- The fitted coefficients are constraint dependent.
- Additional constraints result in a less good fit but do not appear to help forecasting.
- In the APC model canonical correlations between the three sets of fitted parameters make the assumption of independence of $\hat{\boldsymbol{\kappa}}$ and $\hat{\gamma}$ questionable.
- See Clayton & Schifflers (1987) for a careful discussion on what may and may not be inferred from the APC model.
- Problems of identifiability and constraints apply to other models with cohort effects, eg, Renshaw & Haberman (2006), M6, M7 & M8 in Cairns et al (2009).

References

- Cairns, Blake, Dowd, Coughlan, Epstein, Ong & Balevich (2009). North American Actuarial Journal.
- Clayton & Schifflers (1987). Statistics in Medicine.
- Currie (2012). IWSM 2012 (available on Longevitas website).
- Mardia, Kent & Bibby (1979). Multivariate Analysis: Academic Press.
- Renshaw & Haberman (2006). Insurance: Mathematics & Economics.
- Age-Period-Cohort models (APC). Projections Toolkit Technical Guide, Chapter 13.