New Directions in the Modelling of Longevity Risk

Andrew Cairns

Heriot-Watt University,

and The Maxwell Institute, Edinburgh

Joint work (in progress!) with:

George Mavros, Torsten Kleinow, George Streftaris

Life Convention, Brussels, 5 November 2012

Plan

- Motivation
- Genealogy
- New directions in modelling
- Numerical illustrations – single population models
- Remarks on multiple populations

Motivation

- Application focus:
  - risk measurement and management of longevity risk
  - multiple populations
  - life insurance diversification benefits
  - basis risk in standardised longevity contracts
- industry requires robust models

Development of New Models

- Many new stochastic mortality models since Lee-Carter
- Are they fit for purpose?
- Are they robust?

GENEALOGY: 1st GENERATION MODELS

Currie/Richards (M4)
2-D P-splines
2002, ...

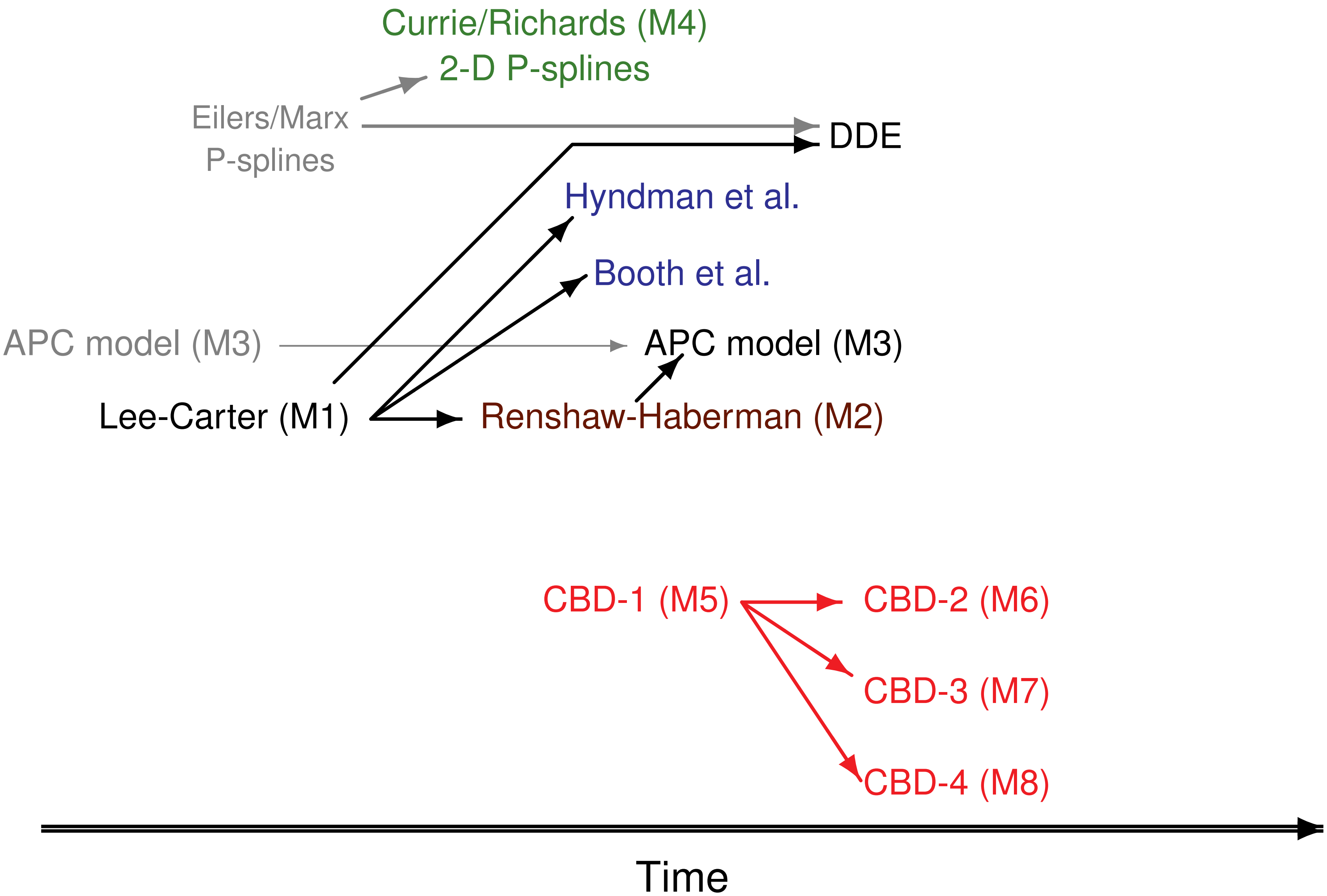
Eilers/Marx
P-splines

Lee-Carter (M1)
1992

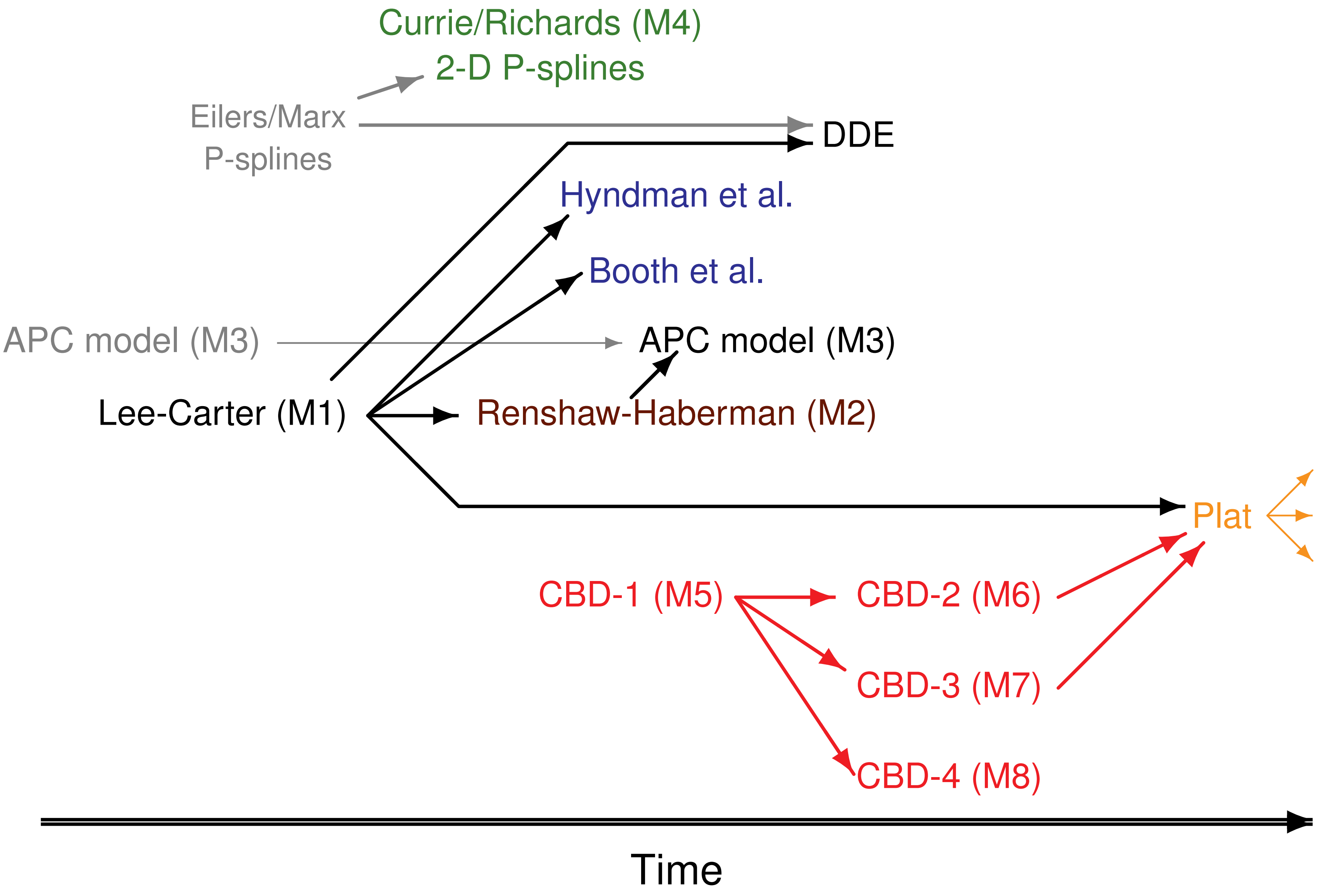
CBD-1 (M5)
2006

Time

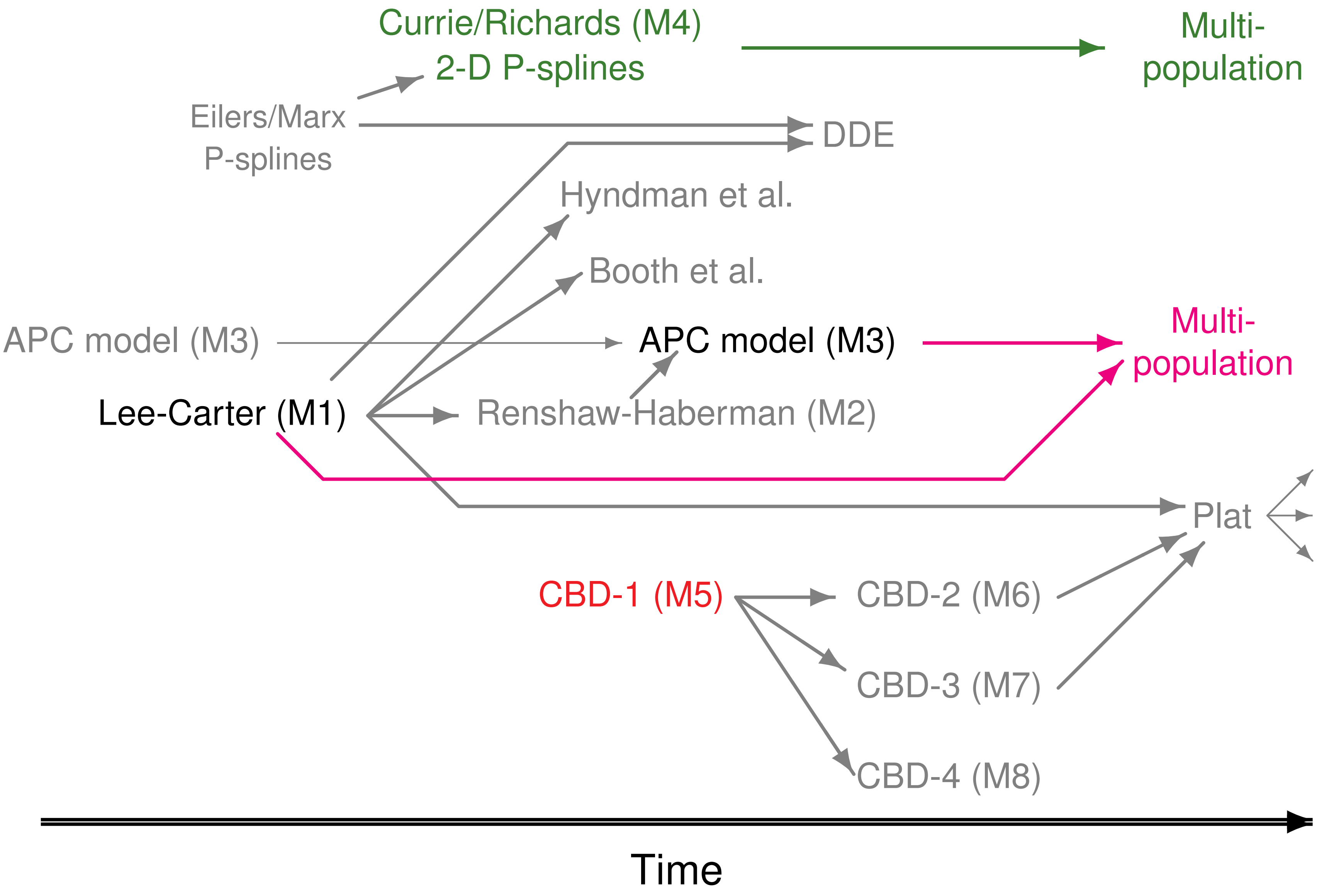
Improvements + more complexity



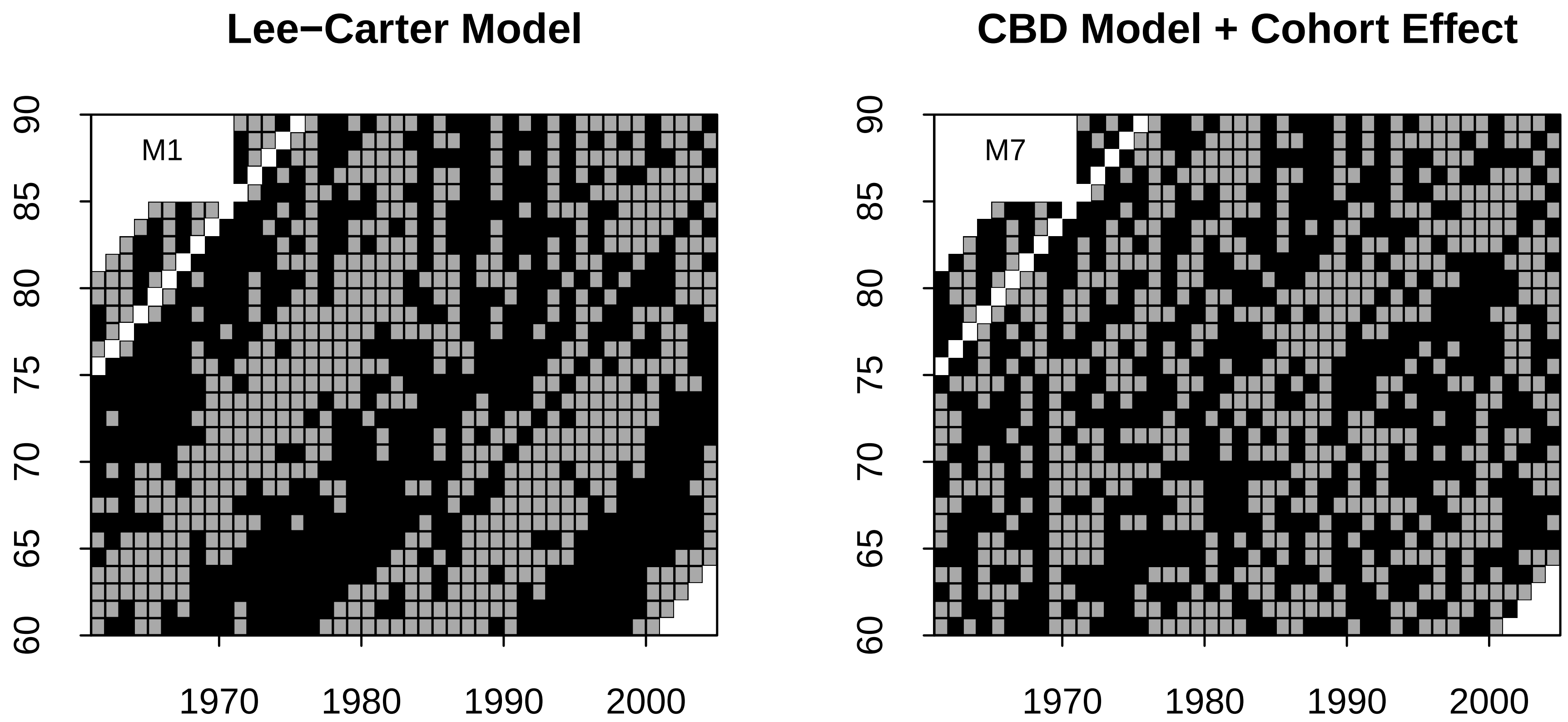
More improvements + even more complexity



Multiple population modelling



Why do we need complexity?

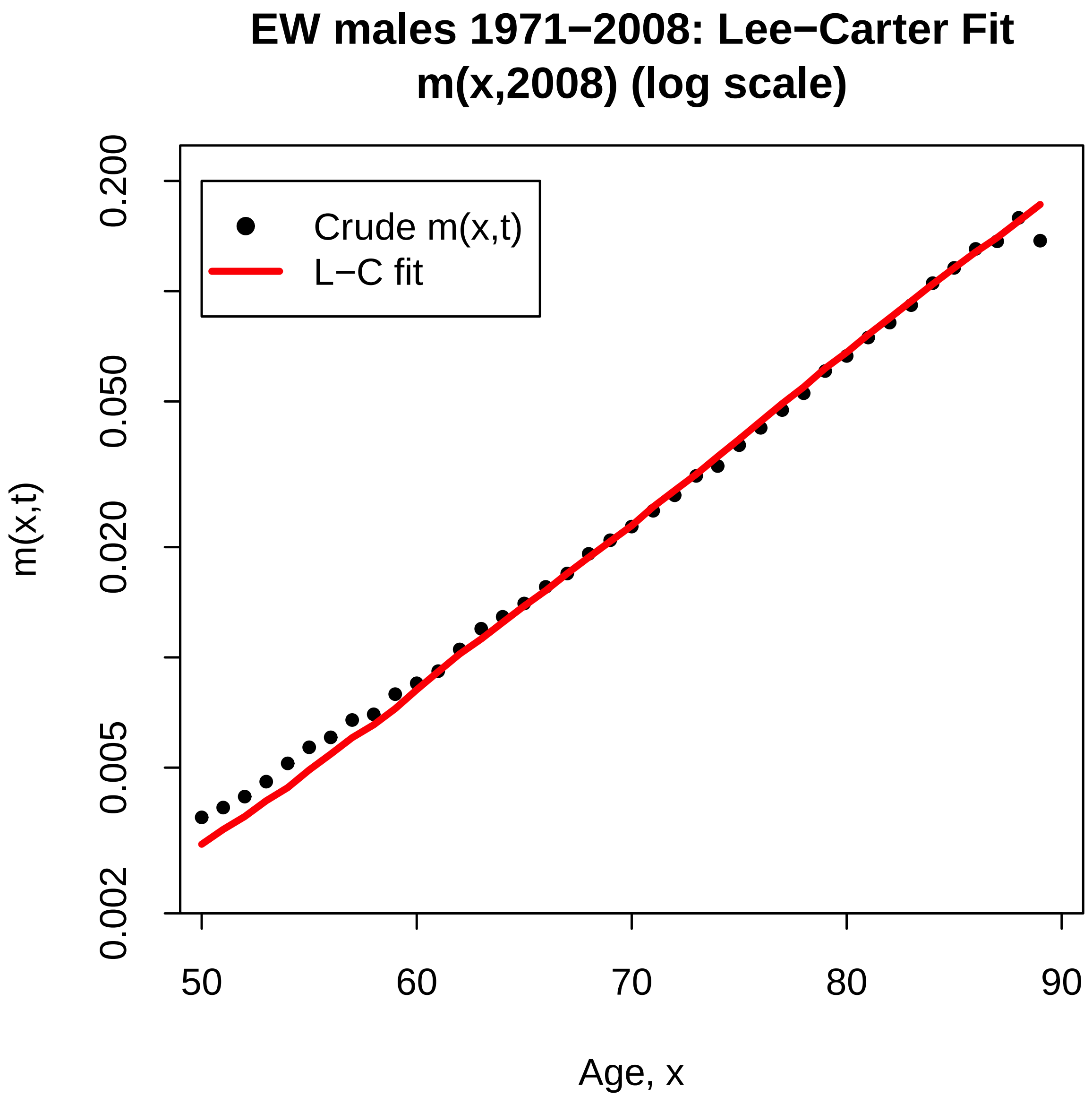


Black $\Rightarrow$ model *over*-estimates $m(x,t)$ death rate

Gray $\Rightarrow$ model *under*-estimates $m(x,t)$ death rate

LC: non-random clusters $+$ errors are too big

Need for complexity: accurate base table for forecasts



Issues on complexity

- More complex $\Rightarrow$ More random processes
- More random processes $\Rightarrow$

  MUCH more difficult to model multiple populations
- Excessive complexity $\Rightarrow$

  potentially less robust forecasts

A Possible Way Forward

Single-population models

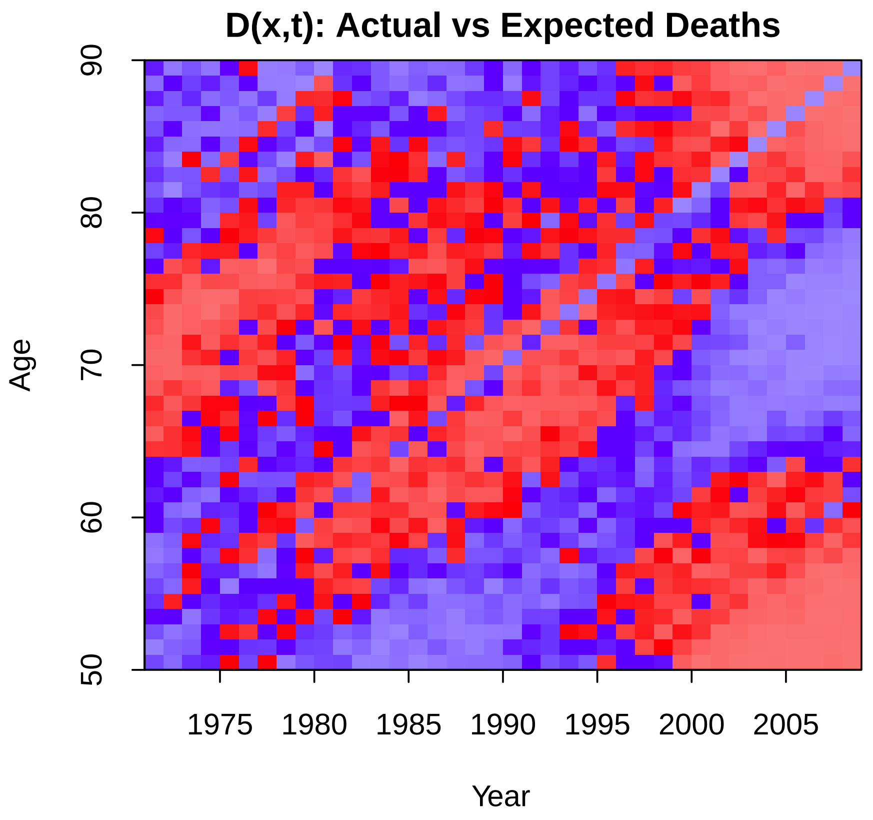
- Focus on small number of key drivers

  $\Rightarrow$ much easier to extend to multi-populations

- Focus on greater robustness of forecasts

Case Study: CBD/Plat Revisited

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



Red $\Rightarrow$ actual deaths $>$ expected deaths

CBD/Plat Revisited: Key Idea: Possible responses

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

Add:

- Cohort effect, $\gamma(t-x)$
- Extra age-period effects
- Do something new ......

Key Idea: CBD/Plat Revisited

Underlying $\log\ m(x,t) =$

- $\beta(x) + \kappa_1(t) + \kappa_2(t)(x - \bar{x})$: two key drivers

  PLUS

  $R(x,t)$ *Residuals*

- Assume: vector $R(t) \rightarrow R(t+1)$ mean reverting process

  $\Rightarrow$ long term risk depends on two key drivers

Specific Model

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

- $\big(\kappa_1(t), \kappa_2(t)\big)$: bivariate random walk
- $R(t) =$ ($n_x \times 1$ vector) VAR(2), reverting to 0

$$R(t) = AR(t-1) + BR(t-2) + Z(t)$$

- $Z(x,t) \sim$ i.i.d.$N(0, \sigma_Z^2)$
- $A = A_1 + A_2$ and $B = -A_2 A_1$

VAR matrices A_1 and A_2

$$A_i = \begin{pmatrix} a_i & 0 & 0 & \cdots & & & & \\ c_i & d_i & 0 & 0 & \cdots & & & \\ d_i/2 & c_i & d_i/2 & 0 & 0 & \cdots & & \\ 0 & d_i/2 & c_i & d_i/2 & 0 & 0 & \cdots & \\ 0 & 0 & d_i/2 & c_i & d_i/2 & 0 & 0 & \cdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \end{pmatrix}$$

a_i =AR terms for new members;

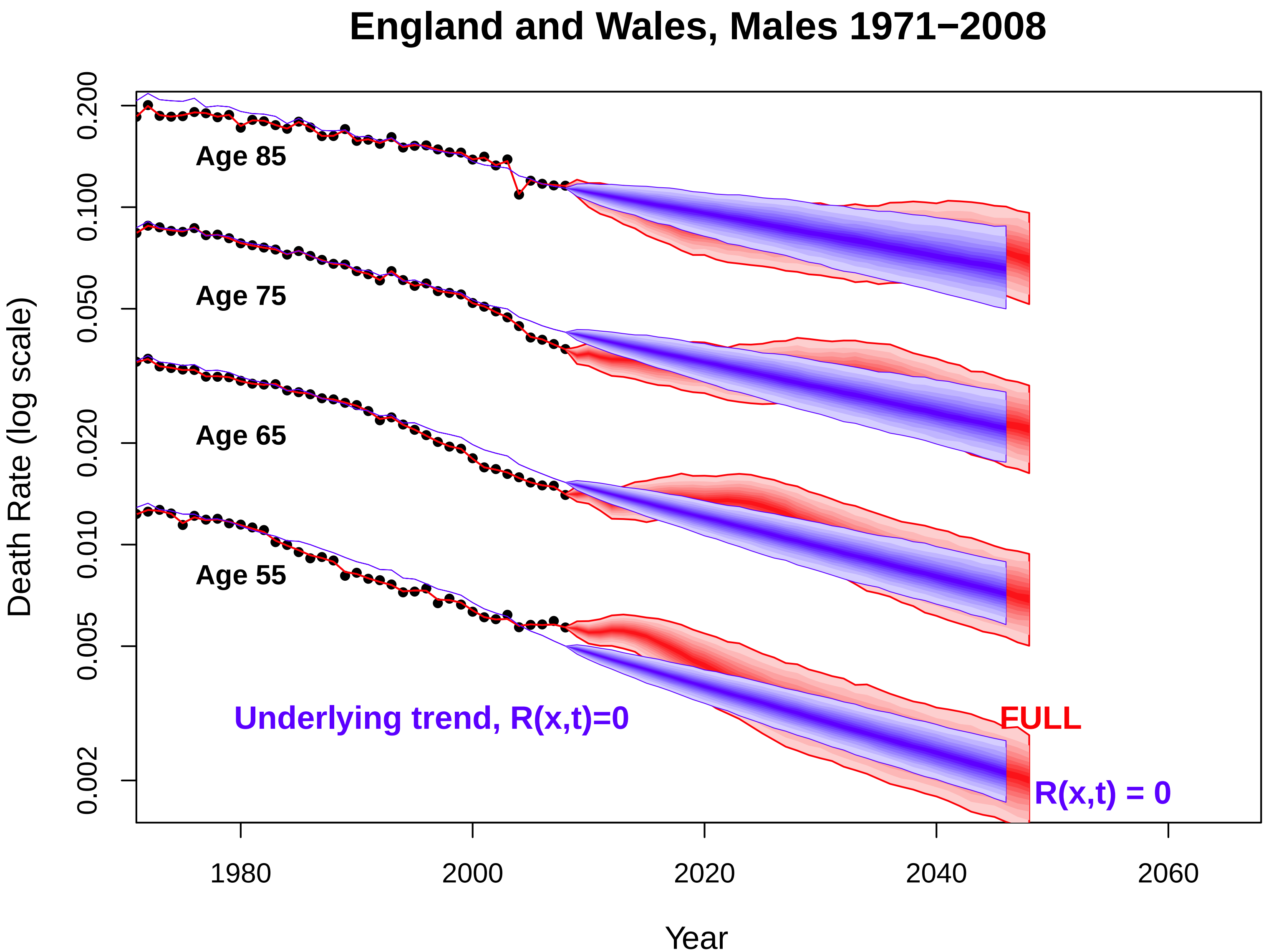
c_i = cohort persistence;

d_i = diffusion coeff.

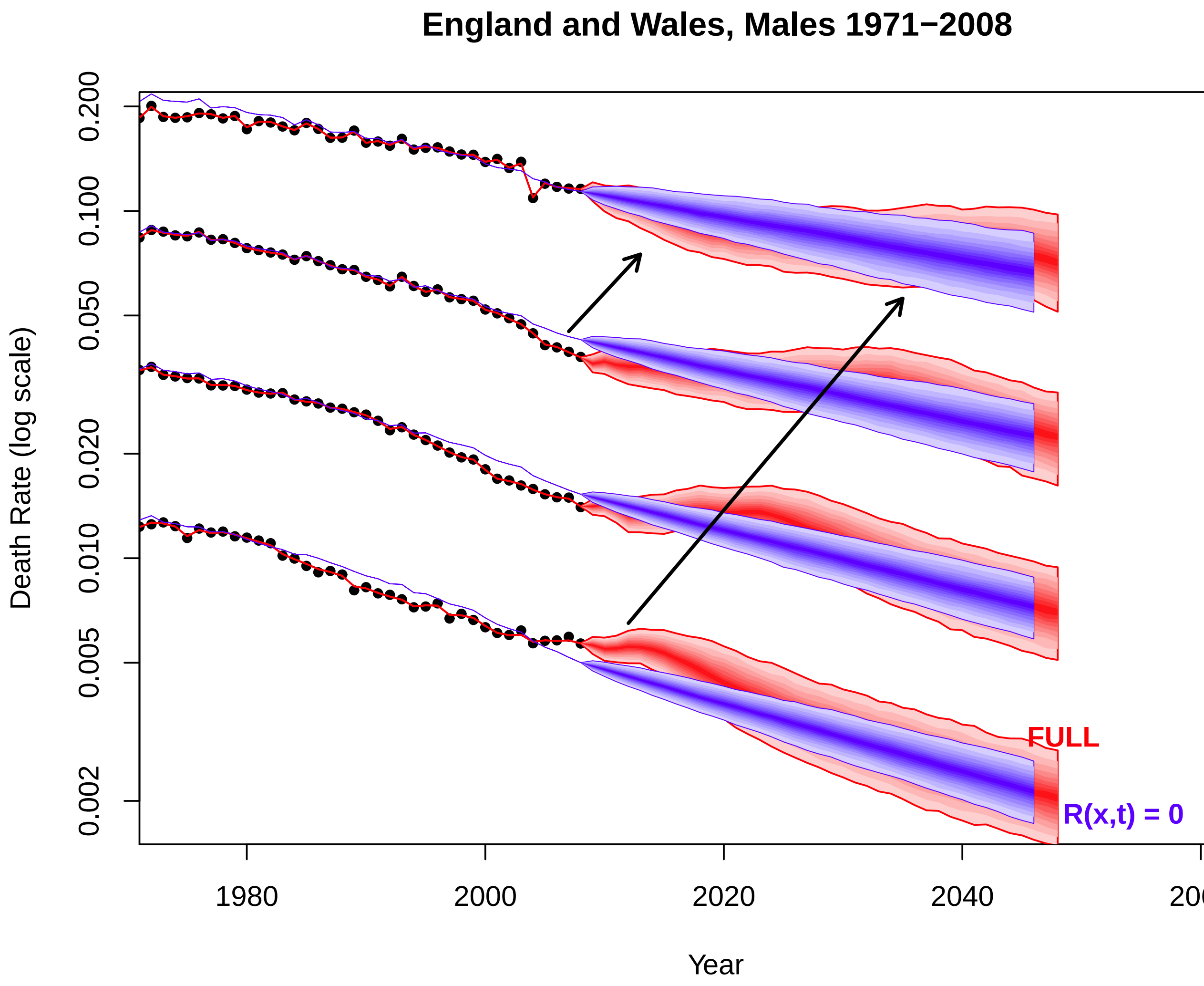
Further details

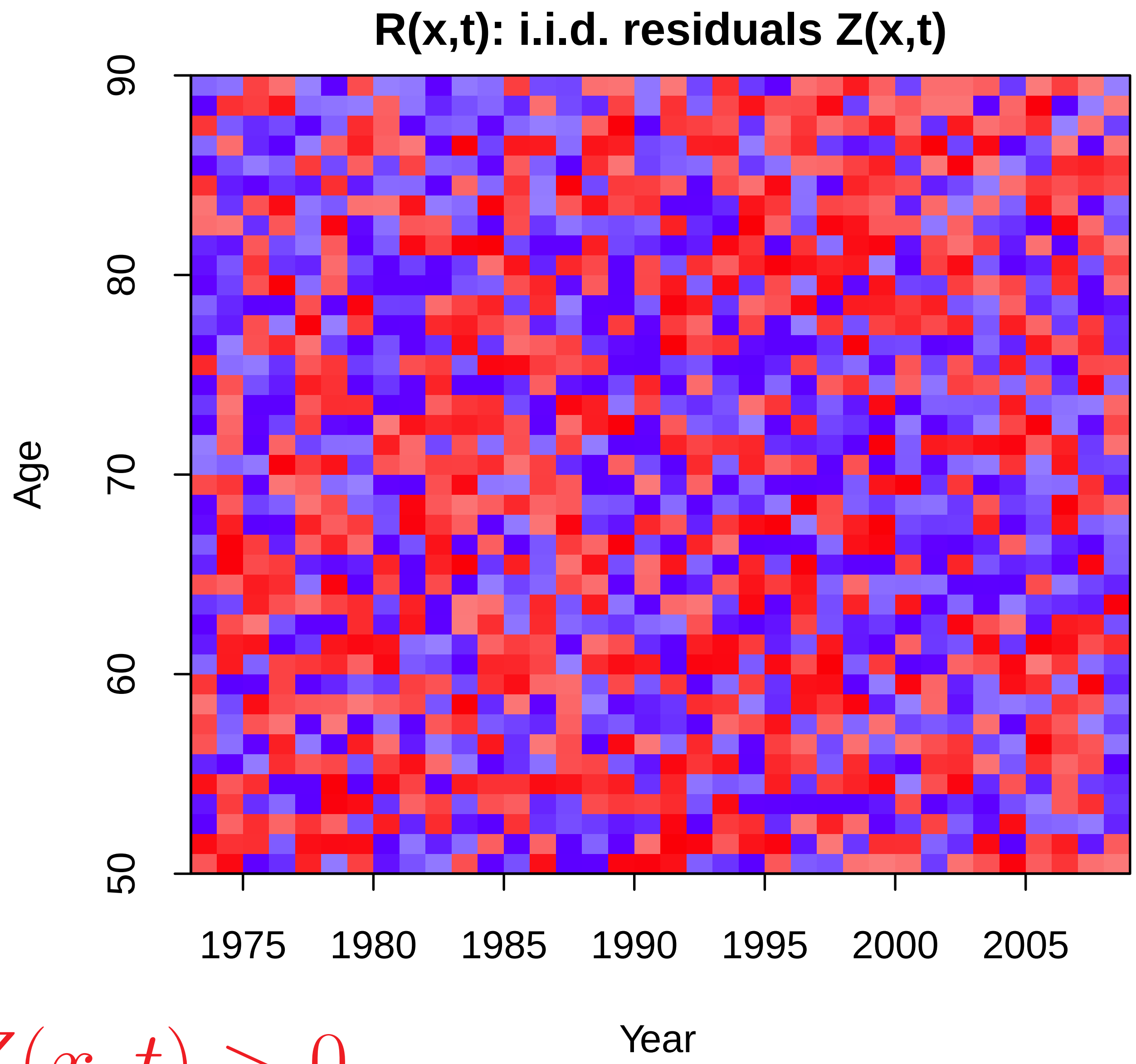
- Deaths: $D(x,t) \sim$ Poisson $(m(x,t)E(x,t))$
- Bayesian approach:

  posterior density $=$ likelihood $\times$ prior
- Upcoming results: mode of posterior density
- Further work: Bayesian parameter uncertainty

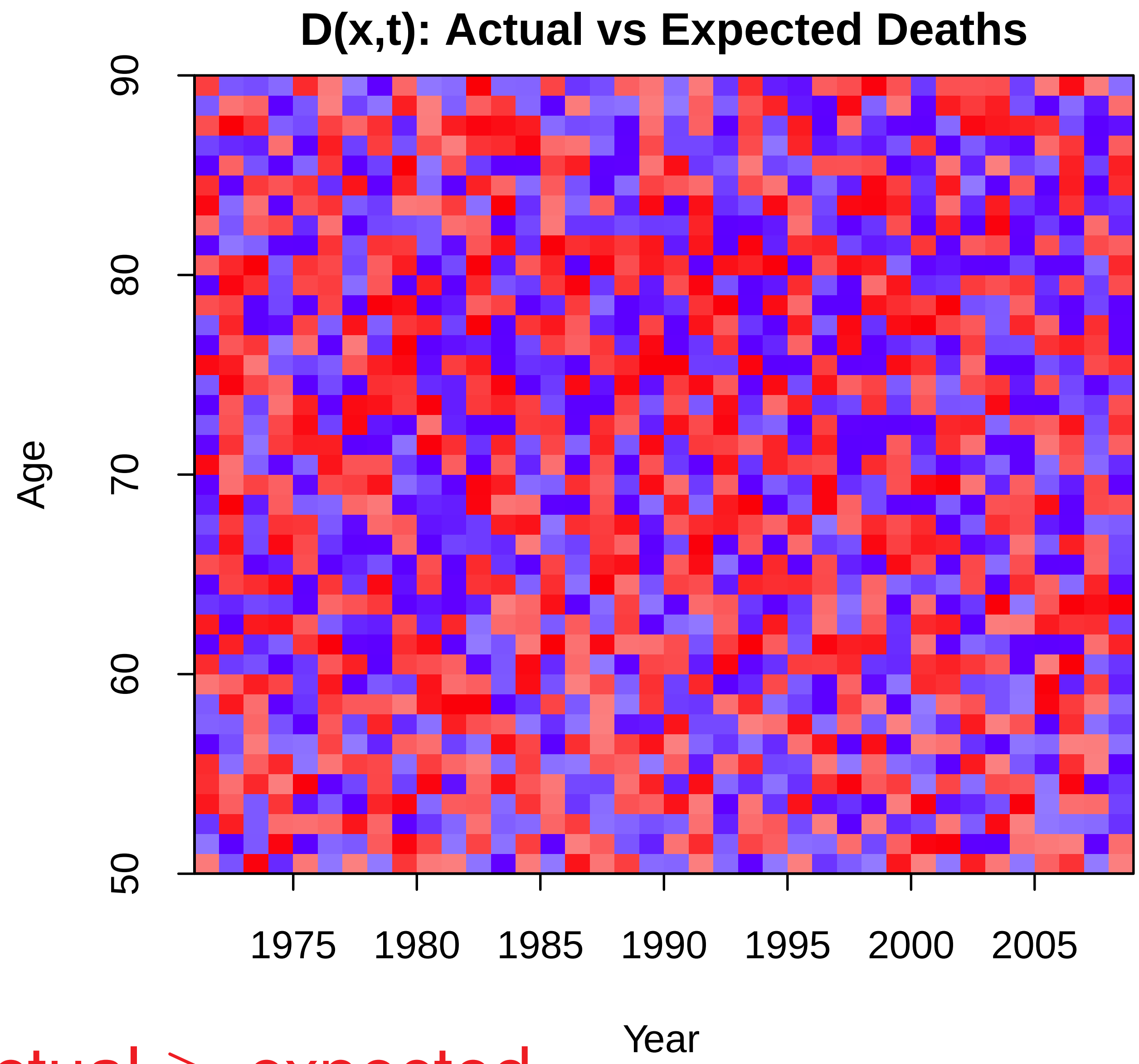

England and Wales, Males 1971−2008
Death Rate (log scale)
0.200
0.100
0.050
0.020
0.010
0.005
0.002
Age 85
Age 75
Age 65
Age 55
Underlying trend, R(x,t)=0
FULL
R(x,t) = 0
1980
2000
2020
2040
2060
Year

Cohort-type effects





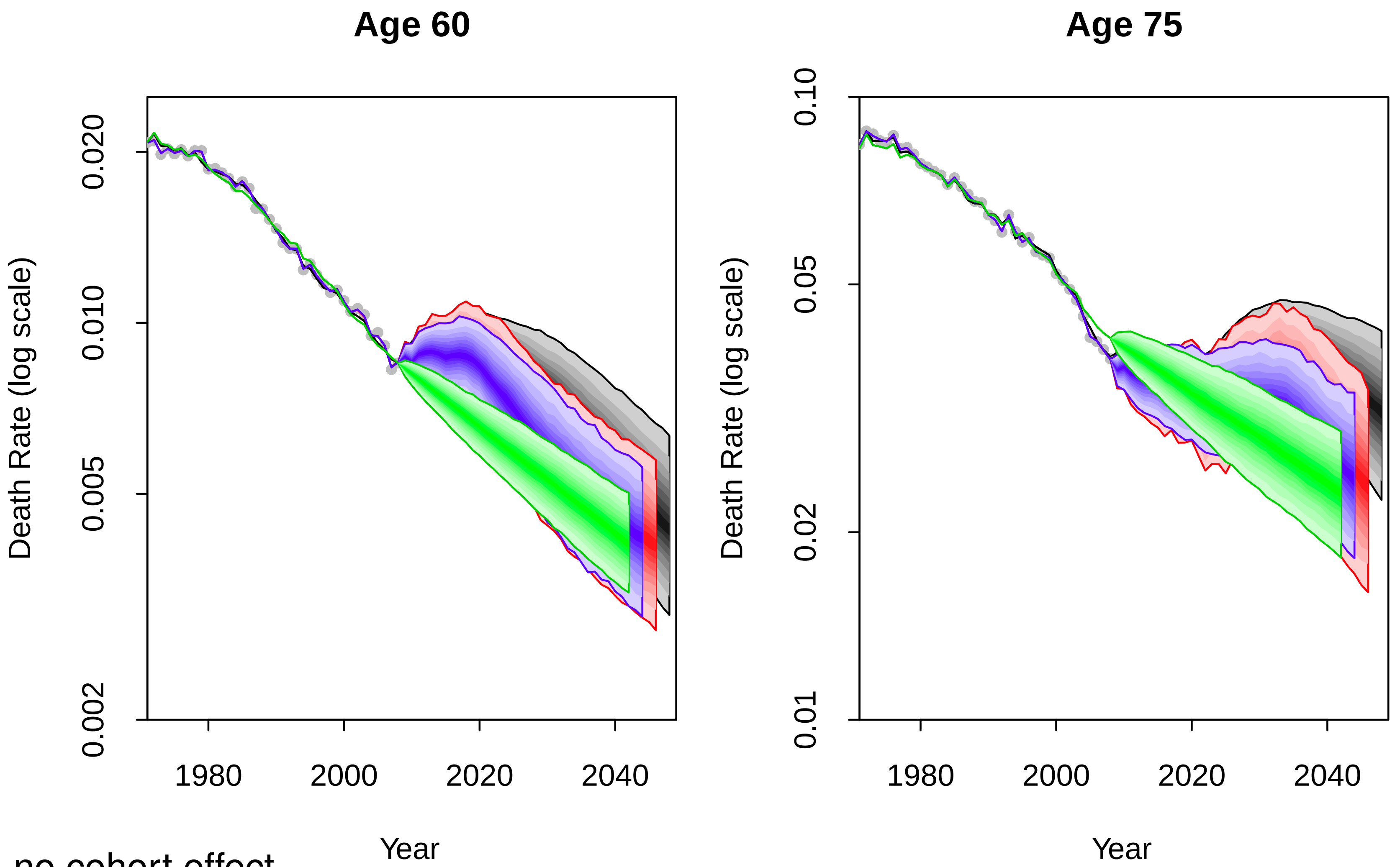
Red $\Rightarrow Z(x,t) > 0$



Red $\Rightarrow$ actual $>$ expected

Comparison with related models

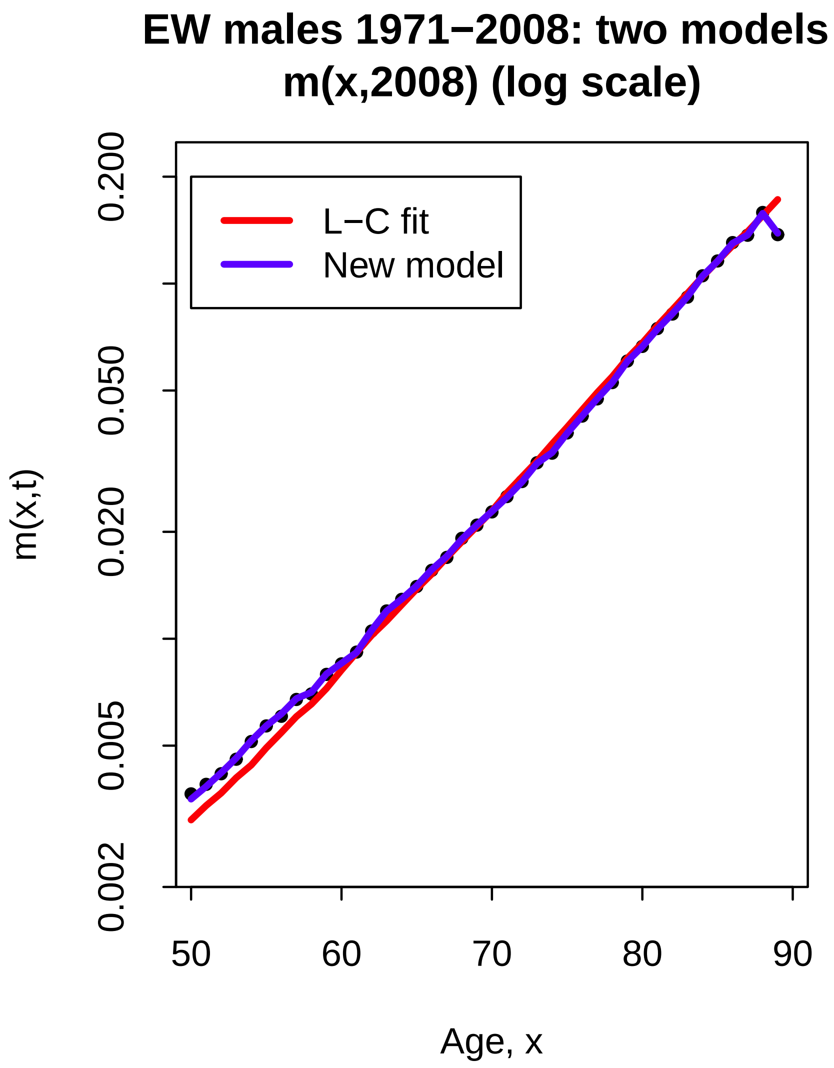
$$\log m(x,t) = \beta(x) + \kappa_1(t) + \kappa_2(t)(x - \bar{x})$$
$$\log m(x,t) = \beta(x) + \kappa_1(t) + \kappa_2(t)(x - \bar{x}) + \gamma(t - x)$$
$$\log m(x,t) = \beta(x) + \kappa_1(t) + \kappa_2(t)(x - \bar{x}) + R(x,t)$$ $R(t) = AR(t-1) + BR(t-2) + Z(t)$ (A, B as specified earlier)
$$\log m(x,t) = \beta(x) + \kappa_1(t) + \kappa_2(t)(x - \bar{x}) + R(x,t)$$ $R(t) = AR(t-1) + BR(t-2) + Z(t)$ (simplified A, B)

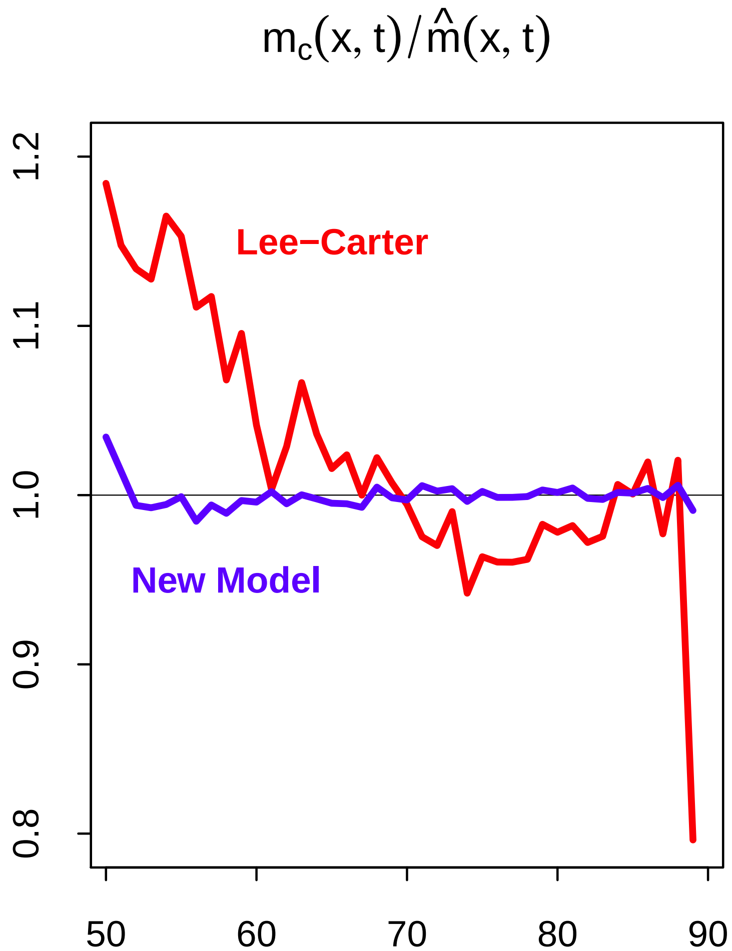


Green: no cohort effect

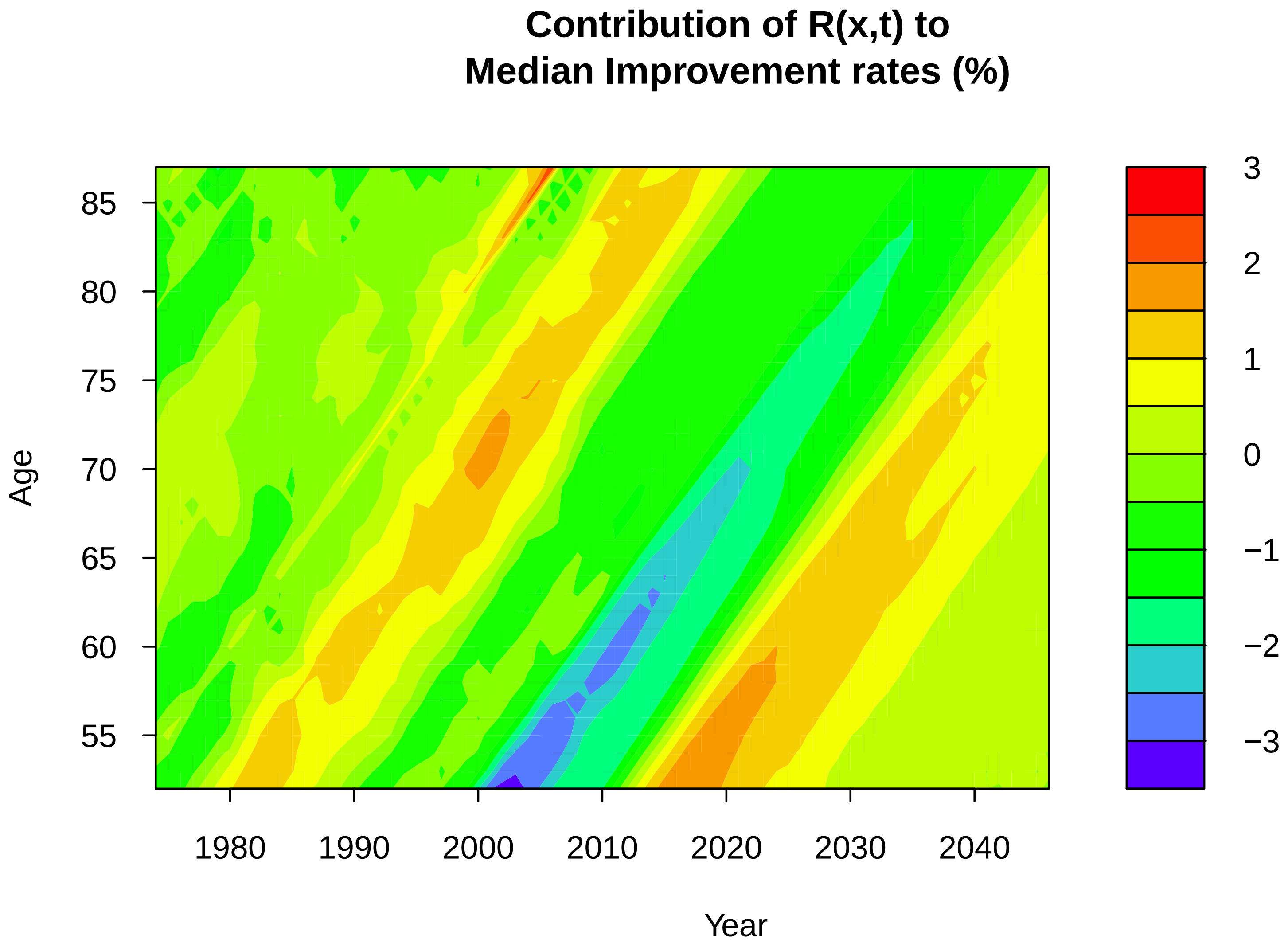
Red, Blue, Gray: similar trends but also differences

Base Table Accuracy

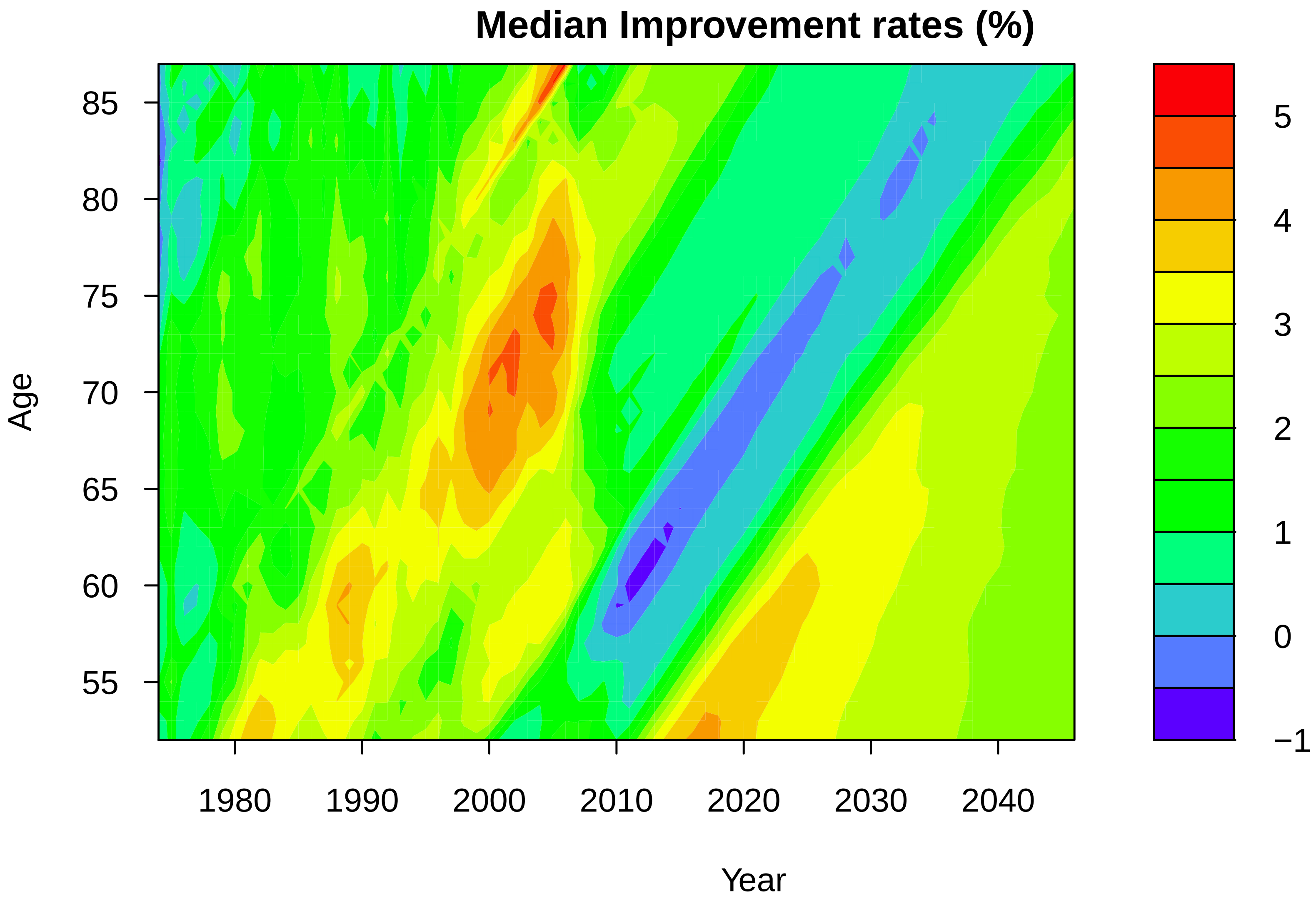




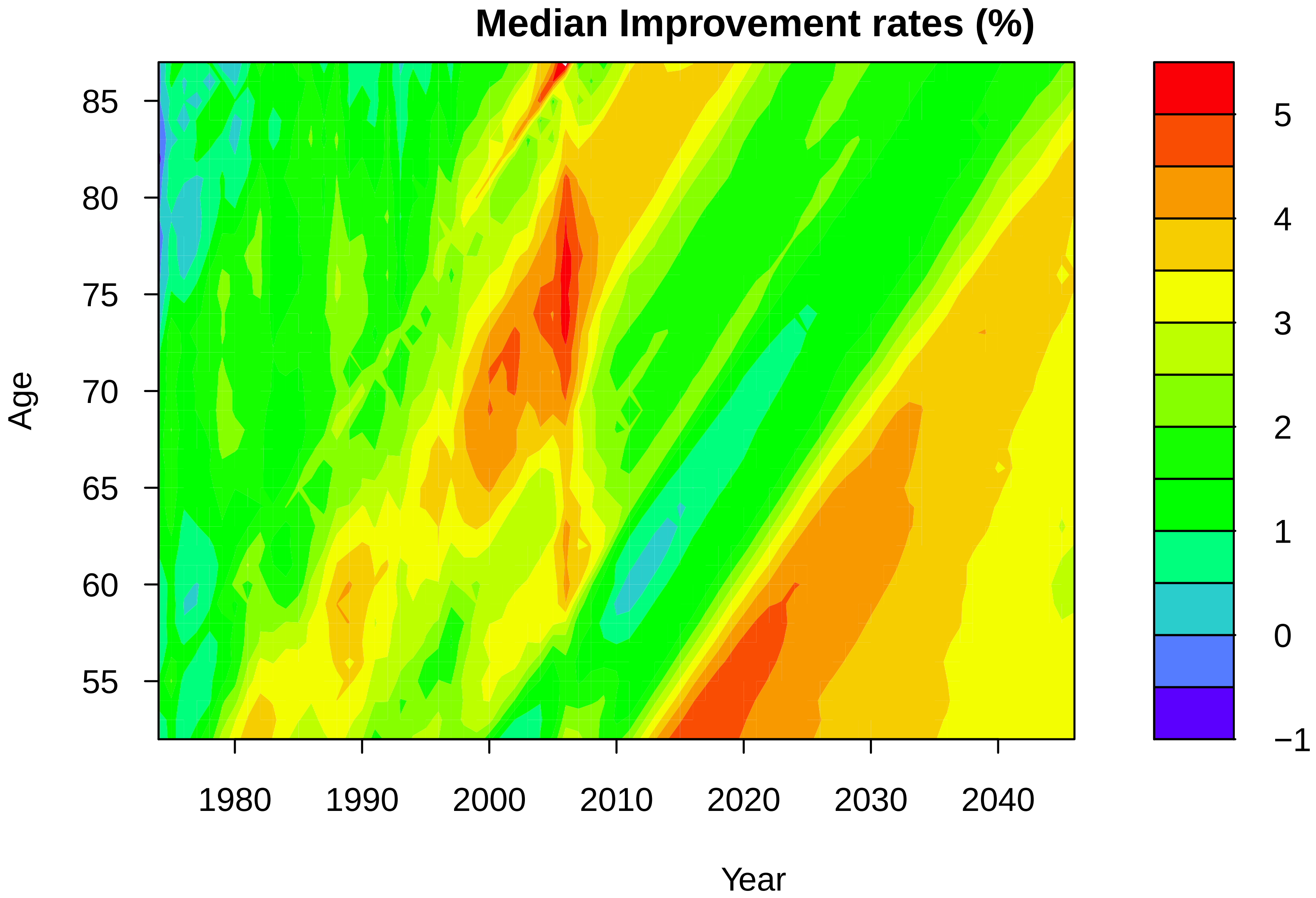
Contribution of $R(x,t)$ to improvement rates



Median total improvement rates



Median total improvement rates: Adjusted $\kappa(t)$



Multiple populations: some thoughts

- Aim for a parsimonious structure
- Items to deal with:
  - Population, P, specific $\kappa_i^{(P)}(t)$
  - Population, P, specific $R^{(P)}(x, t)$

Multiple populations: possible structures

Mortality – version 1:

- Population, P, specific $\kappa_i^{(P)}(t)$ correlated
- $R^{(P)}(x,t)$: assume independent

Mortality – version 2:

- All populations have the same $\kappa_i(t)$
- $R^{(P)}(x,t)$: assume independent
- Greater role for $R^{(P)}(x,t)$ as country specific effect

Conclusions

- New model
  - focus on a small number of core period effects
  - adds alternative $R(x,t)$ to popular cohort effects, $\gamma(t-x)$
- Model risk more evident in the mean reverting $R(x,t)$
- But general framework should prove to be more robust:

  long term underlying trends ($\kappa(t)$) are reasonably consistent

Further work

- Bayesian parameter uncertainty
- Multiple populations: focus on underlying $\kappa(t)$

  $\Rightarrow$ less complexity

Questions

W: www.ma.hw.ac.uk/∼andrewc

E: A.J.G.Cairns@hw.ac.uk