A General Procedure for Building Mortality Models

Andrew Hunt and David Blake

andrew.hunt.1@cass.city.ac.uk

Longevity 8 – Waterloo, Canada

8 September 2012

Motivation

- Recently, there has been a proliferation of new mortality models
- Some of these models are "black-box algorithms" such as PCA
  - Involving terms that lack "demographic significance"
- Others have functional terms added without justification and require a priori assumptions about what the model should look like
- We introduce a general procedure which provides structure to the model building process
- This requires an explicit "toolkit" of functions

Age/Period/Cohort Modelling Framework

$$\eta_{x,t} = \alpha_x + \sum_{j=1}^{n} \beta_x^j \kappa_t^j + \beta_x^0 \gamma_{t-x}$$

Link function between model and data; typically

$$\eta_{x,t} = \ln\left(\hat{\mu}_{x,t}\right) = \ln\left(\frac{D_{x,t}}{E_{x,t}^c}\right)$$

$$\eta_{x,t} = \text{logit}\left(\hat{q}_{x,t}\right) = \ln\left(\frac{D_{x,t}}{E_{x,t}^0}\right) - \ln\left(1 - \frac{D_{x,t}}{E_{x,t}^0}\right)$$

Cohort parameters – "lifelong" factors depending upon year of birth

Often set to 1 for simplicity and robustness

Static life table – baseline mortality rates at each age

Period factors governing trends in time at same age range

Age functions – which ages mortality is changing at – in non-parametric form. Smooth age functions $f^j(x;\theta)$ reduce number of parameters to be estimated

Examples of Models in this Form

- Lee and Carter (1992) $$\ln\left(\mu_{x,t}\right)=\alpha_x+\beta_x\kappa_t$$
- Renshaw and Haberman (2006) $$\ln\left(\mu_{x,t}\right)=\alpha_x+\beta_x^1\kappa_t+\beta_x^0\gamma_{t-x}$$
  - (modified version of this model without age function on the γ's found to be more robust)
- Cairns, Blake and Dowd (2006) $$\text{logit}\left(q_{x,t}\right)=\kappa_t^1+(x-\overline{x})\kappa_t^2$$
- Plat (2009) $$\ln\left(\mu_{x,t}\right)=\alpha_x+\kappa_t^1+(x-\overline{x})\kappa_t^2+(x-\overline{x})^+\kappa_t^3+\gamma_{t-x}$$
- O'Hare and Li (2012) $$\ln\left(\mu_{x,t}\right)=\alpha_x+\kappa_t^1+(x-\overline{x})\kappa_t^2+\left((x-\overline{x})^++\left((x-\overline{x})^+\right)^2\right)\kappa_t^3+\gamma_{t-x}$$

Model Selection Criteria

- Adequacy –
  - There should be a sufficient number of terms to capture all significant structure in, and provide a good fit to the data
- Parsimony –
  - Have the smallest number of terms and free parameters necessary (trade off with the adequacy of the model)
- Demographic Significance –
  - Models should be biologically reasonable
  - Terms allow identification with underlying biological and socio-economic processes occurring in the population
- Completeness –
  - Models should span entire age range and not be limited to a subset of ages by construction
  - Models should include allowance for cohort effects and be able to separate these from age/period terms

General Model Building Procedure

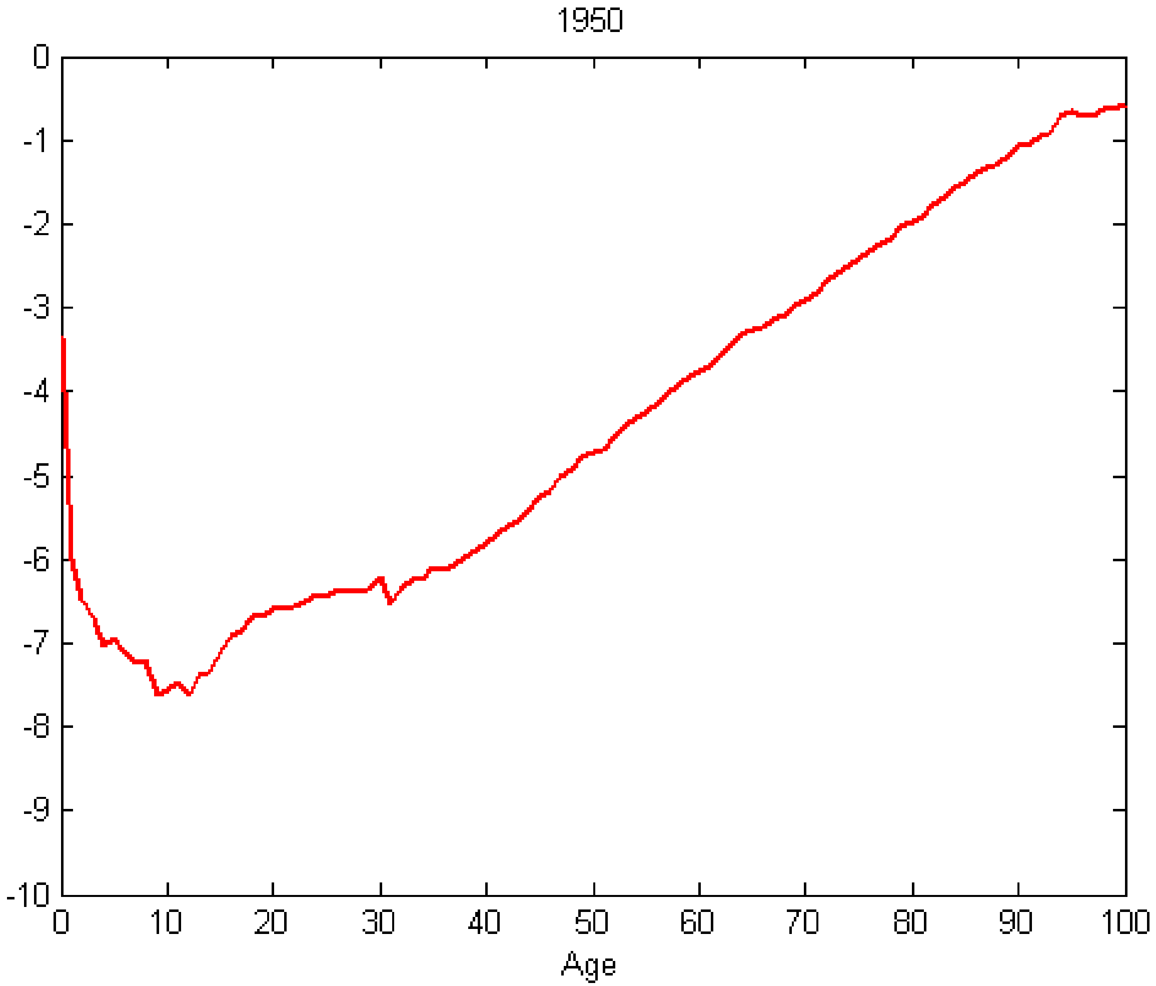
Important to Avoid Over-Parameterisation

- Major risk of procedure is overfitting the model, i.e. adding too many age/period terms.
- We try to avoid this by –
  - Measuring goodness of fit using a metric that penalises the number of parameters used
    - e.g. the Bayes Information Criterion
  - Applying subjective as well as statistical tests on whether new terms are needed – they must be demographically significant as well as statistically significant
  - Requiring age functions to be smooth over a range of adjacent ages to avoid trying to fit statistical noise

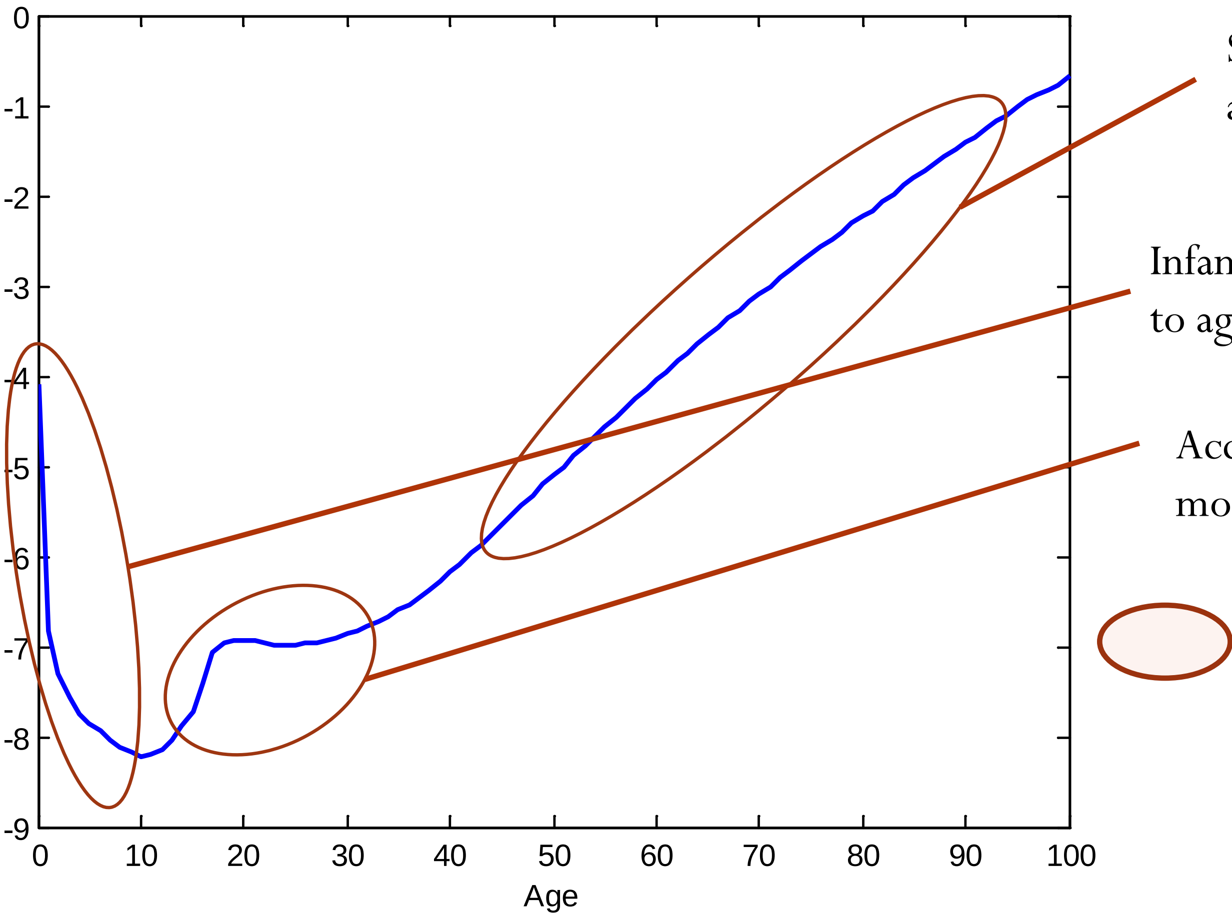
Application

- We apply the general procedure to data for men in the UK from ages 0-100 and years 1950-2009
- Death counts and central exposures to risk from the Human Mortality Database
- We test the final model and its residuals and compare it with existing model fitting procedures
- Need access to a suitable "toolkit" of age functions which we can use to build an appropriate model

UK Male Data



Stage 0 - Static Life Table

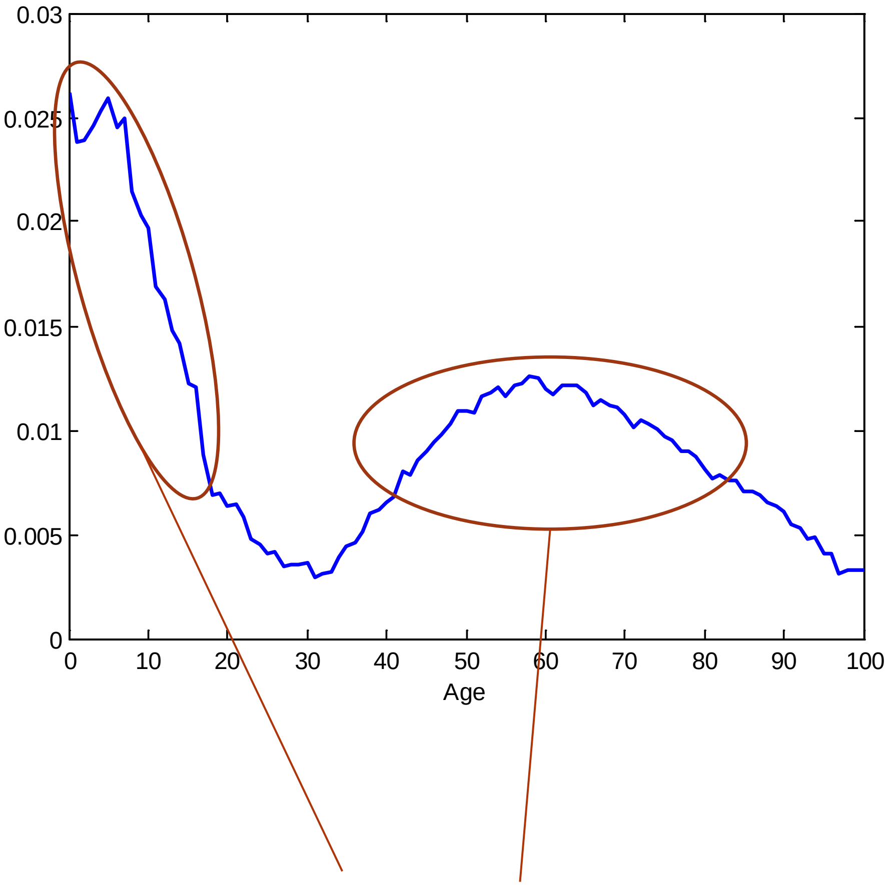


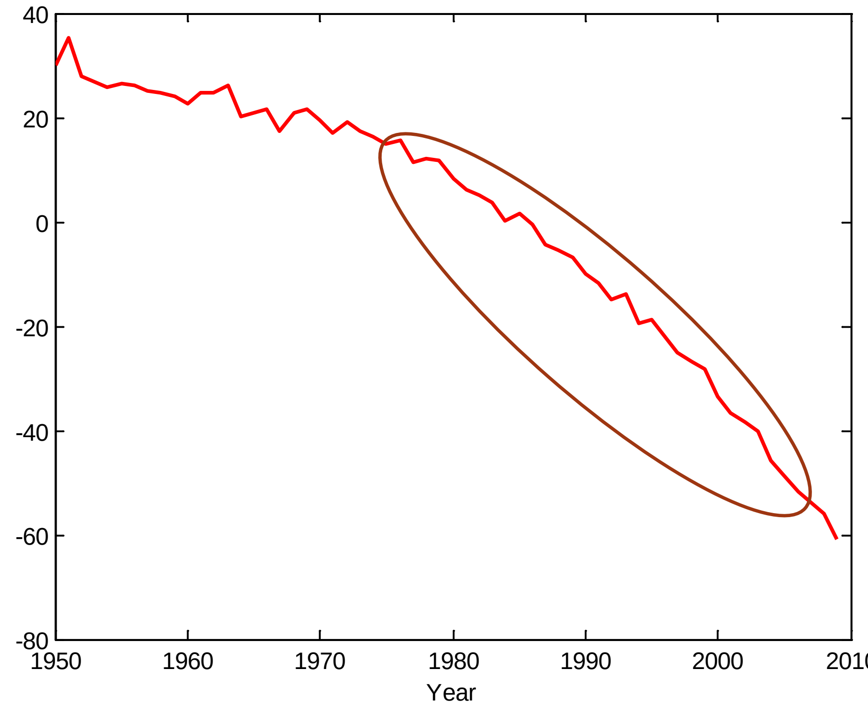
Senescent mortality – approx linear from age 50

Infant mortality – sharp falling off to age 10

Accident hump – increased mortality rates for men aged 16-25

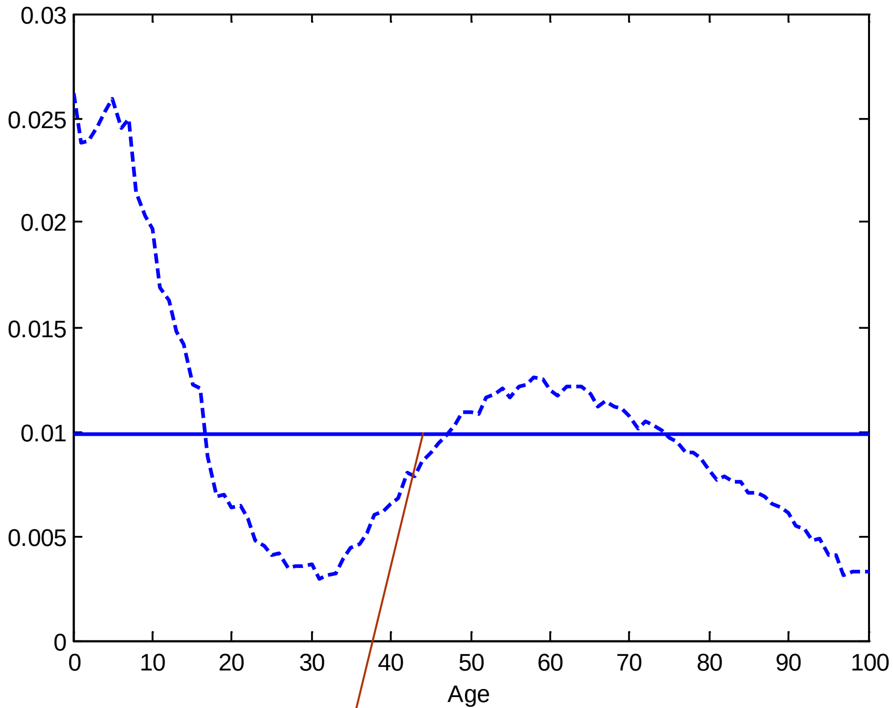
Stage 1 - Add Non-Parametric Age/Period Complementary Pair to Find 1st Dominant Trend

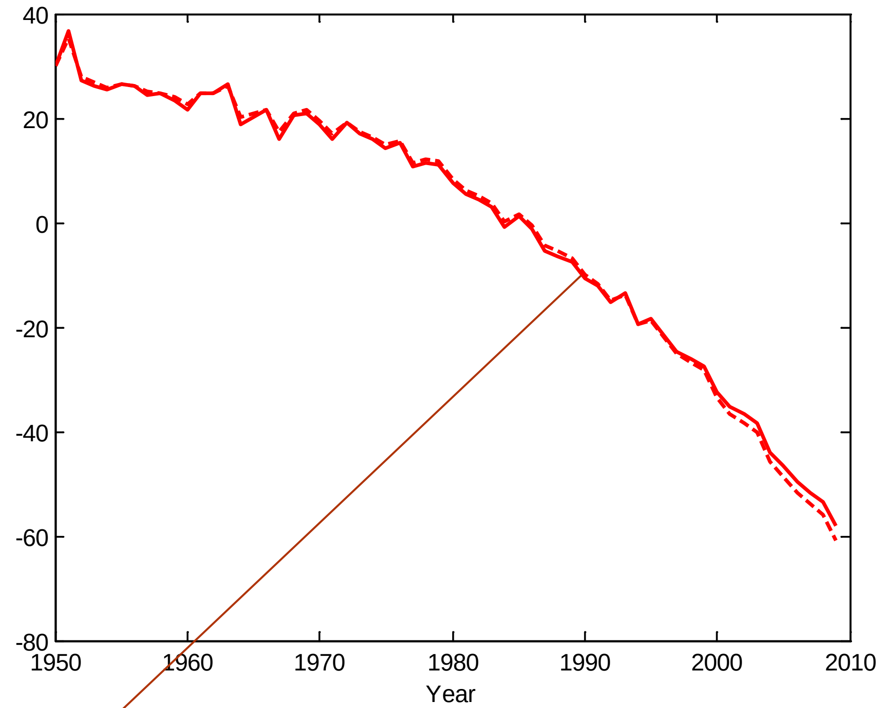




Combines features from across the age range with different underlying causes but which are correlated

Stage 1 – Find Simplest Parametric Term That Does the Same Job

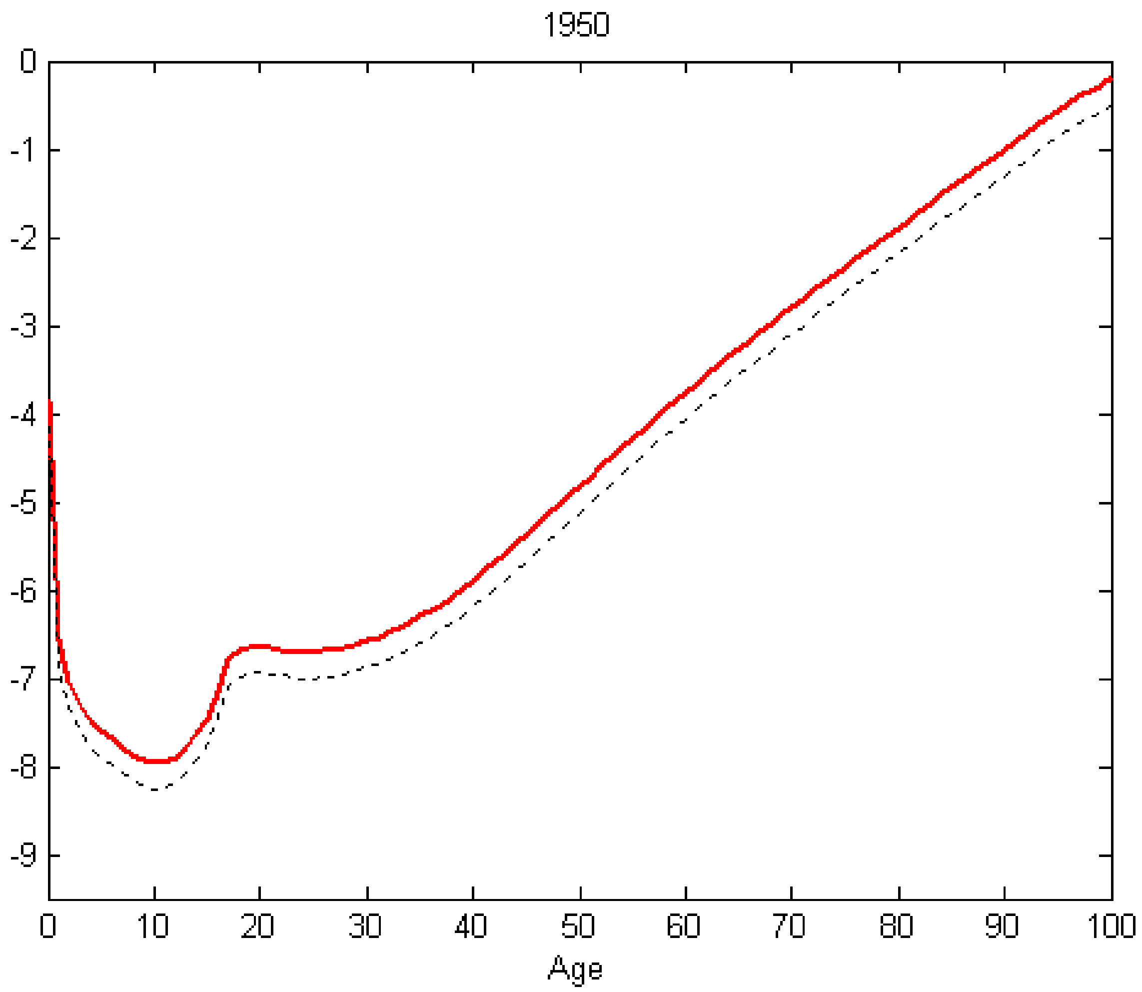




Constant age function – changes in overall level of mortality due to accidents, hygiene, etc

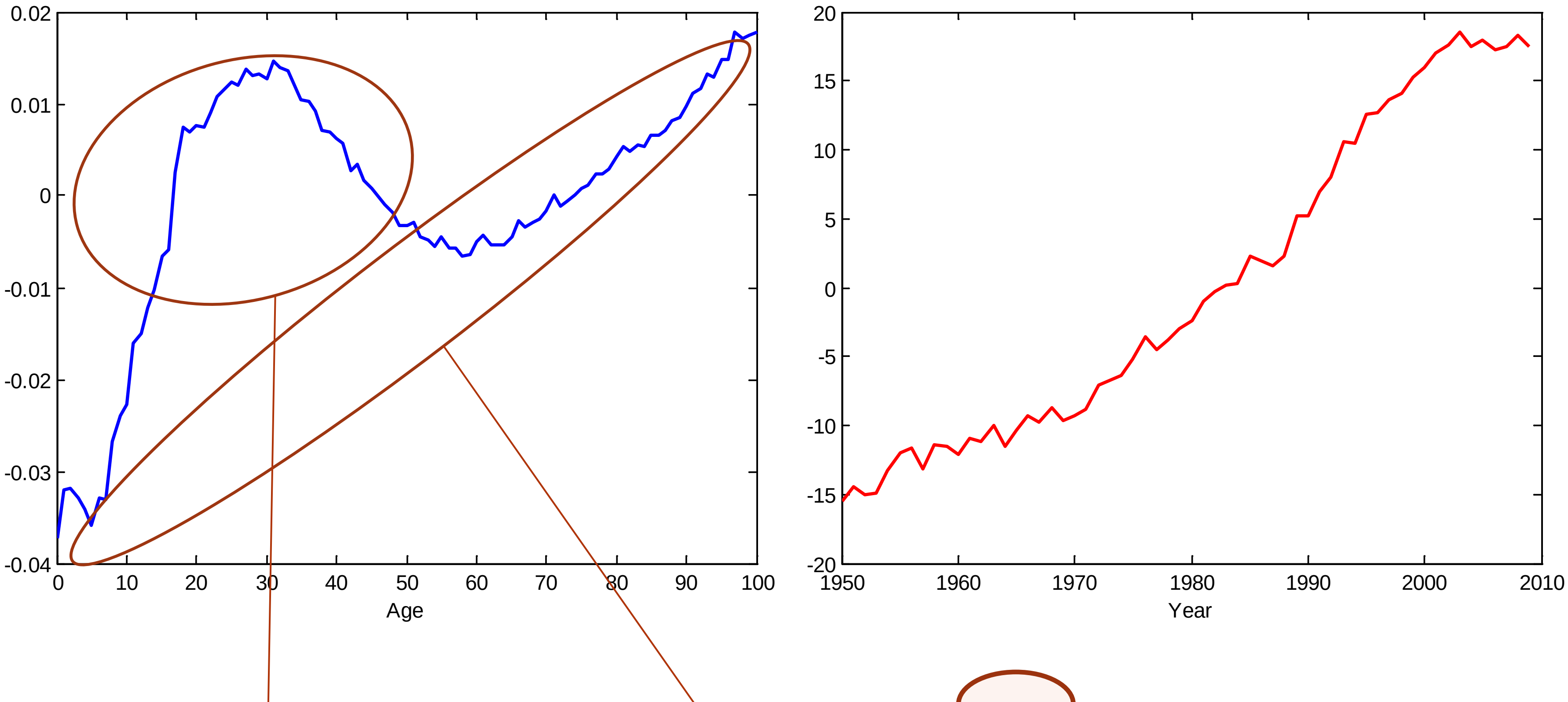
Finds same trend in the data – general fall in mortality rate with time

Stage 1 – Effect of 1st Term



Stage 1 - Test Goodness of Fit

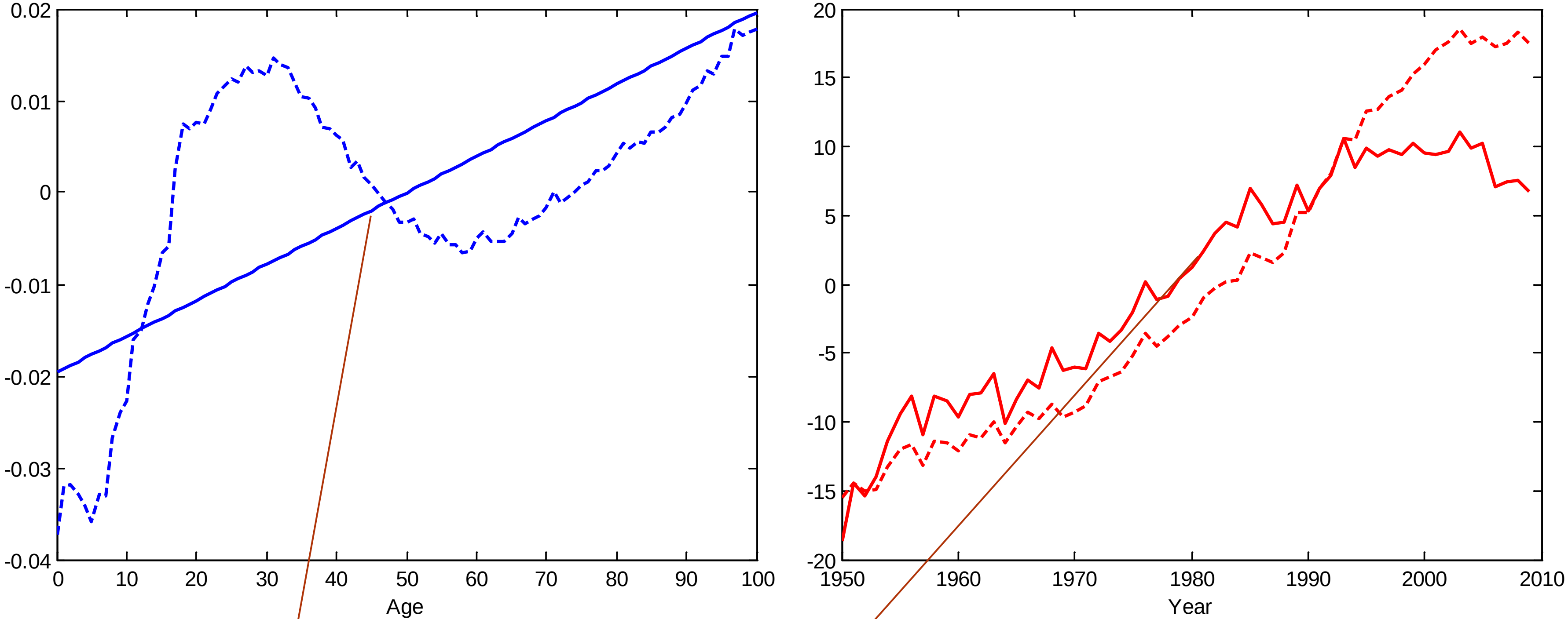
Stage 2 - Add Non-Parametric Age/Period Complementary Pair to Find 2nd Dominant Trend



“Hump” for low ages?

Linear trend across entire age range?

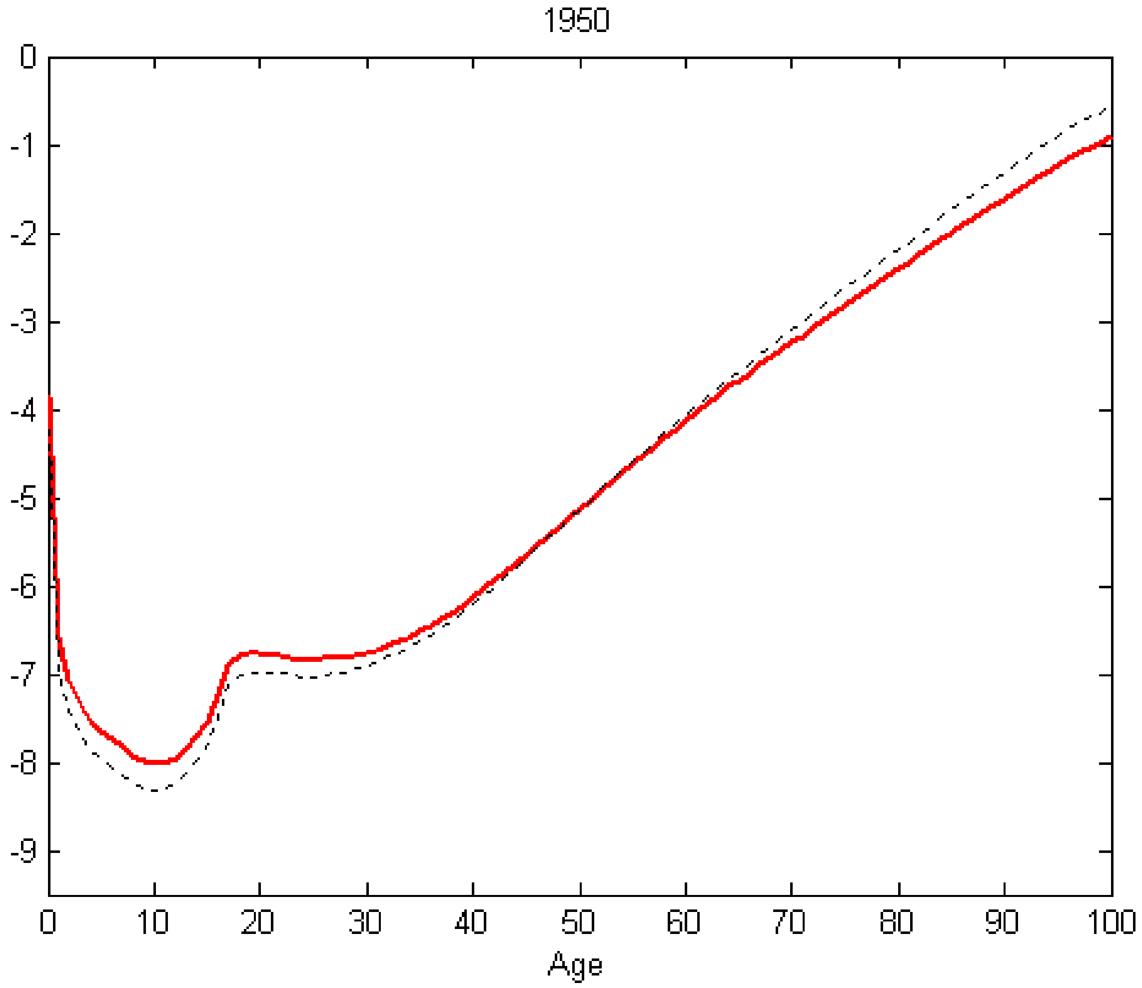
Stage 2 – Find Simplest Parametric Term That Does the Same Job



Linear age function – governs changes in the Gompertz slope for given year

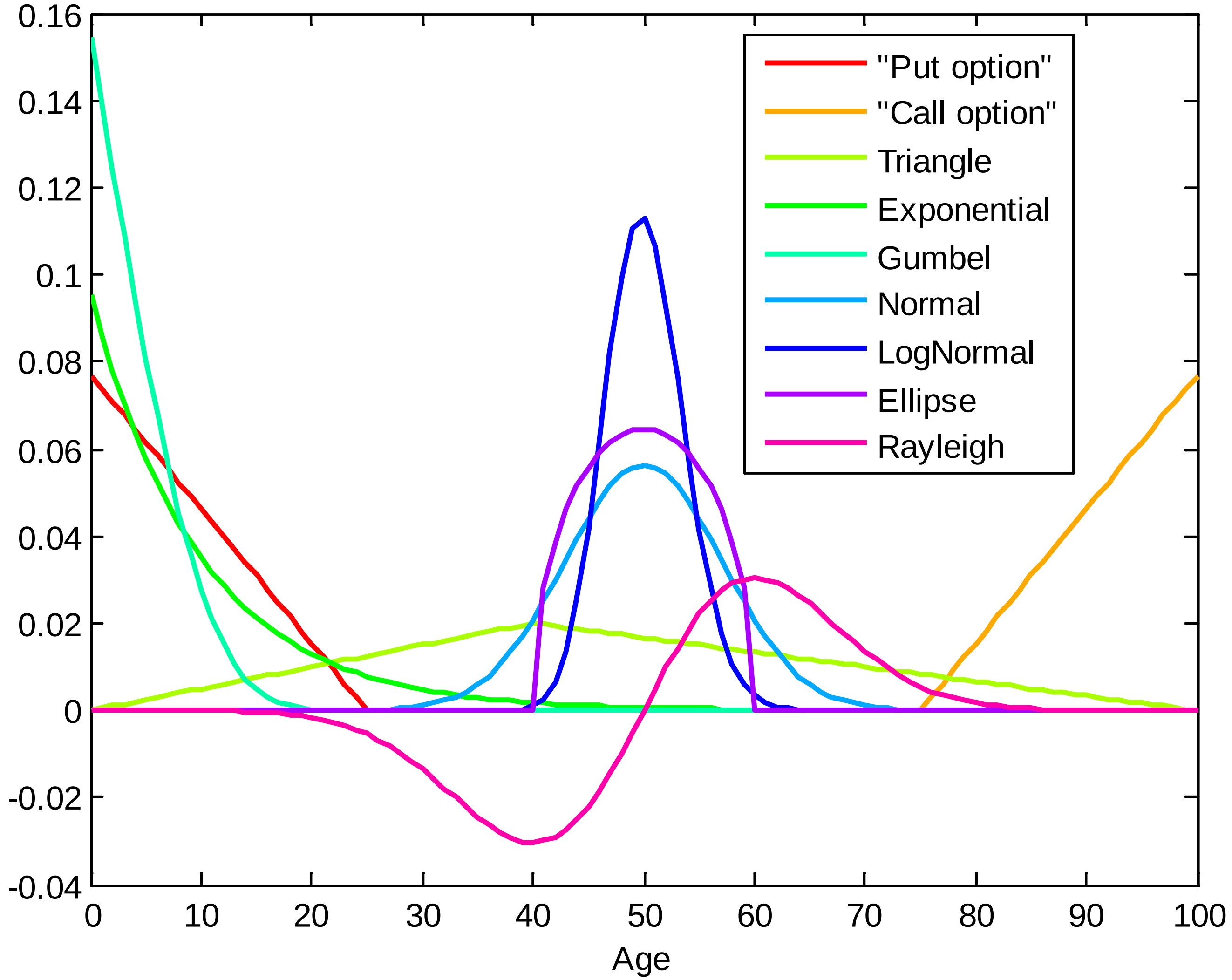
"Rectangularisation" of mortality curve with time

Stage 2 – Effect of 2nd Term

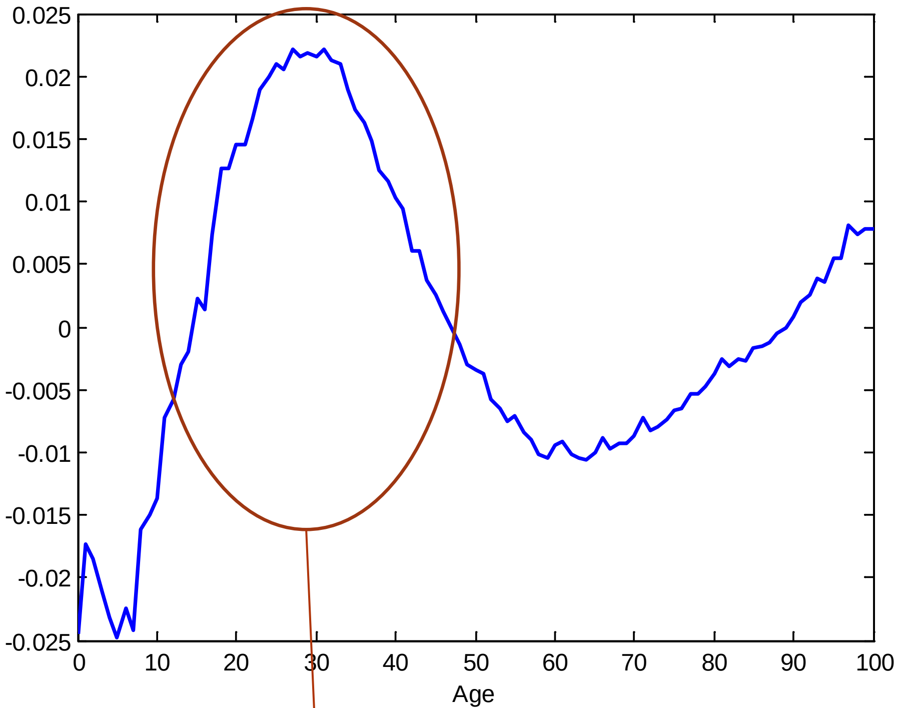


Stage 2 - Test Goodness of Fit

Toolkit of Age Functions

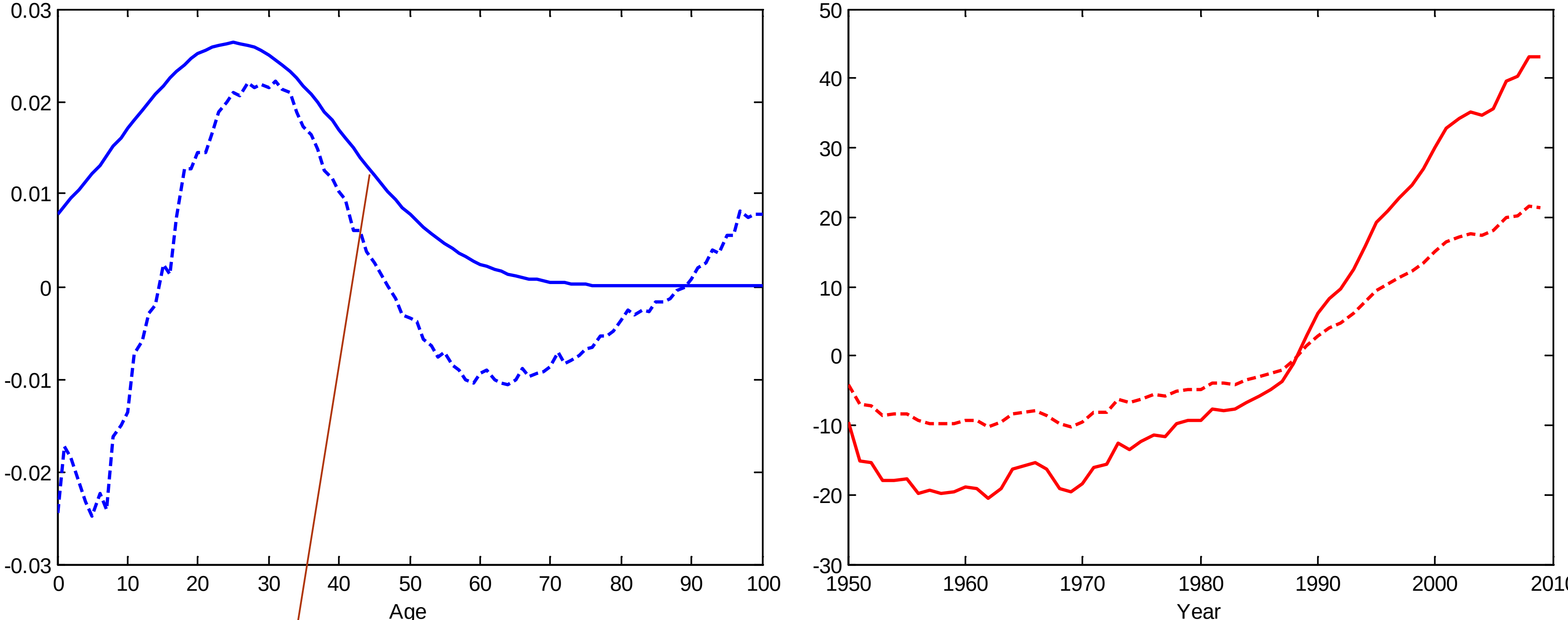


Stage 3 - Add Non-Parametric Age/Period Complementary Pair to Find 3rd Dominant Trend



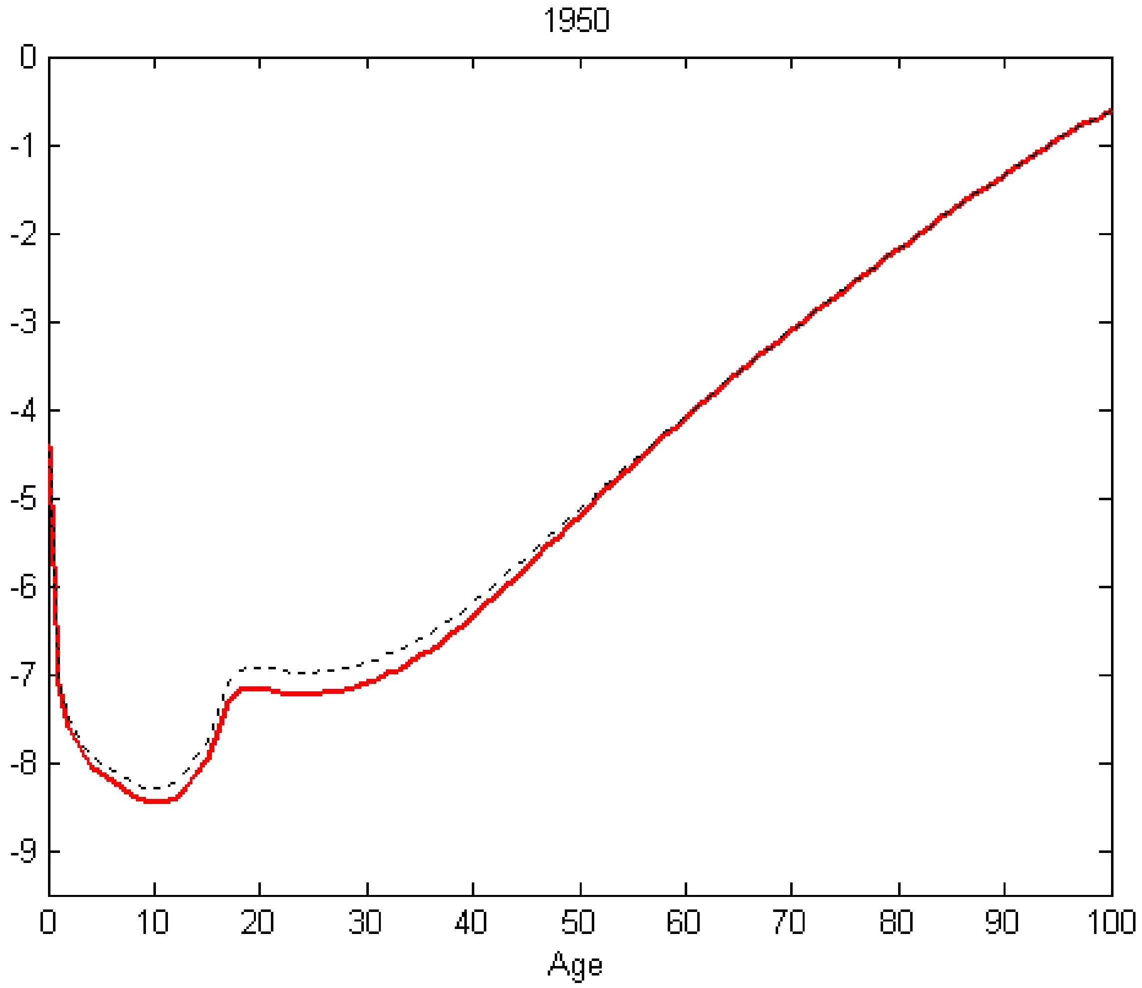
"Hump" for low ages

Stage 3 – Find Simple Parametric Term That Does the Same Job



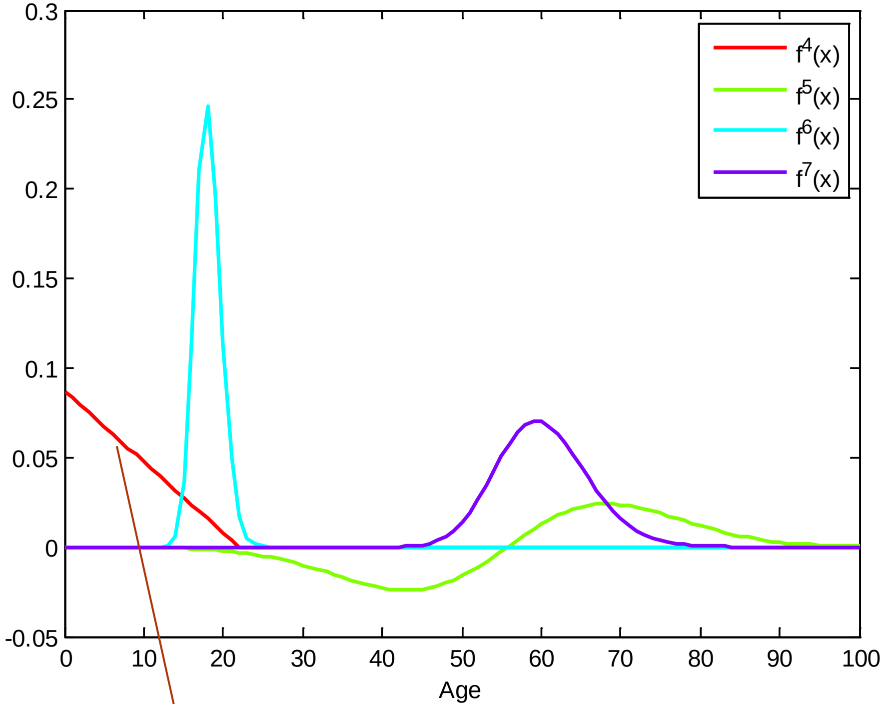
Gaussian "hump" governs changes in mortality rates for young adults

Stage 3 – Effect of 3rd Term



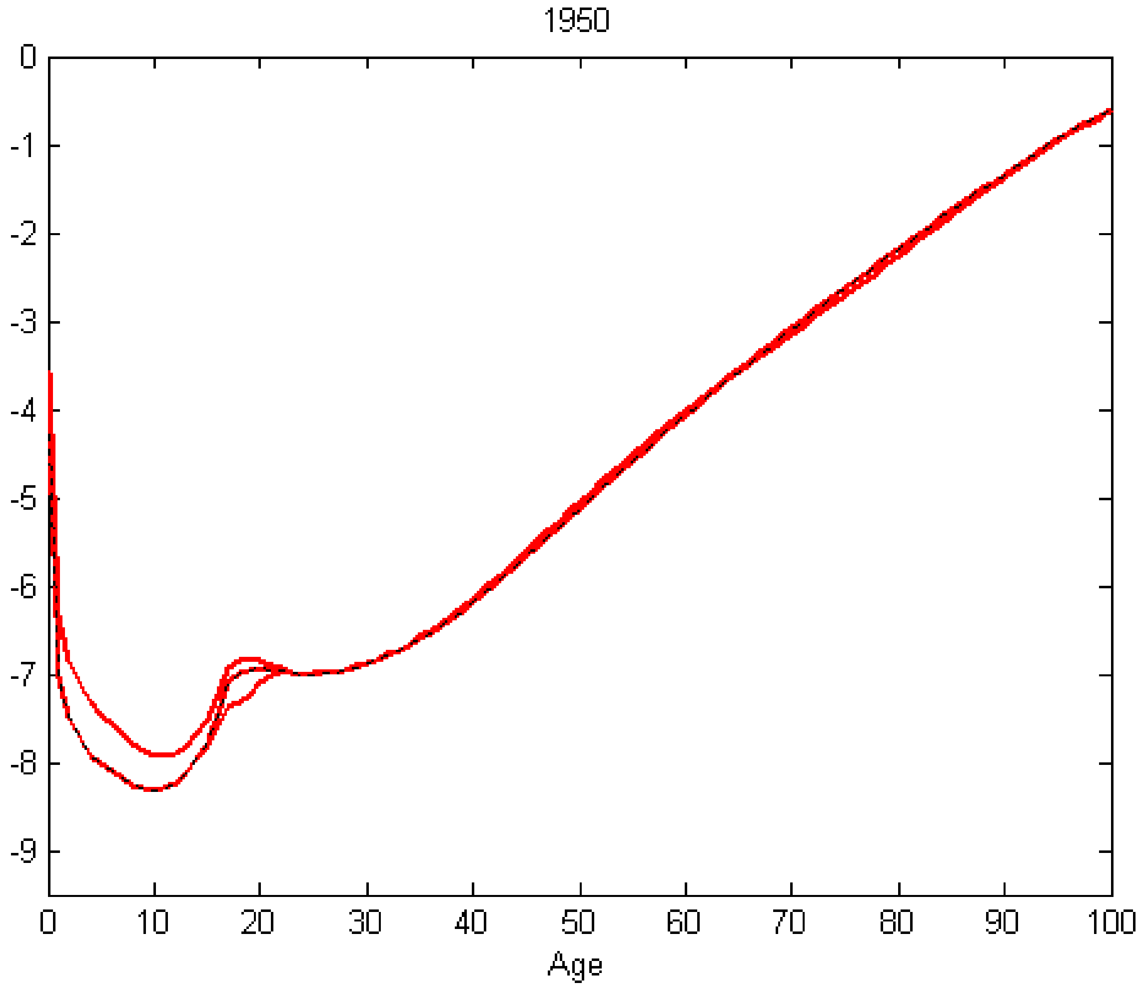
Stage 3 - Test Goodness of Fit

Stage 4 Onwards



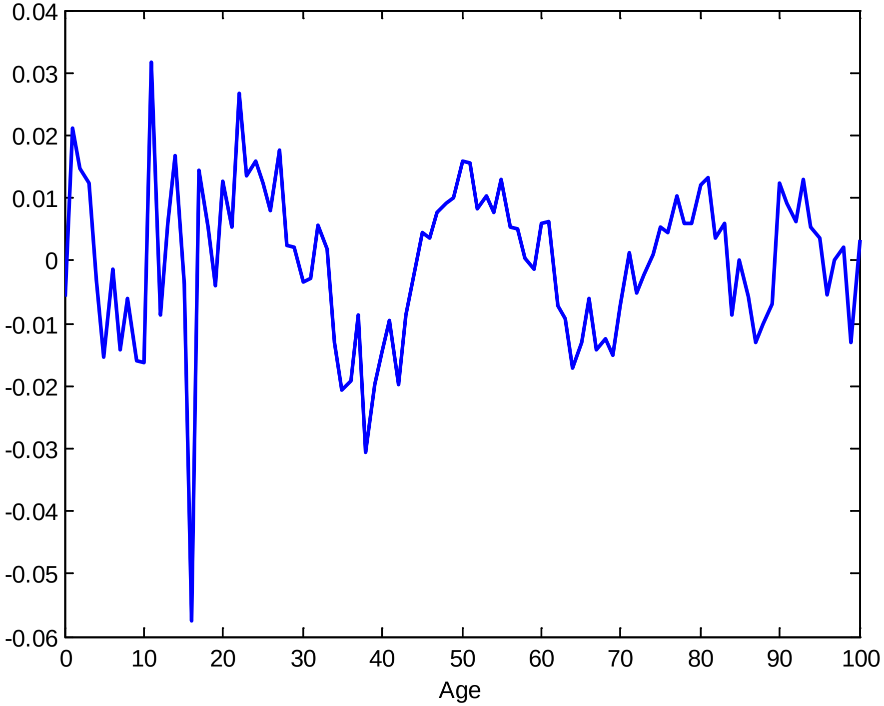
"Call option" function – childhood mortality

Stage 4 Onwards – Effect of Terms

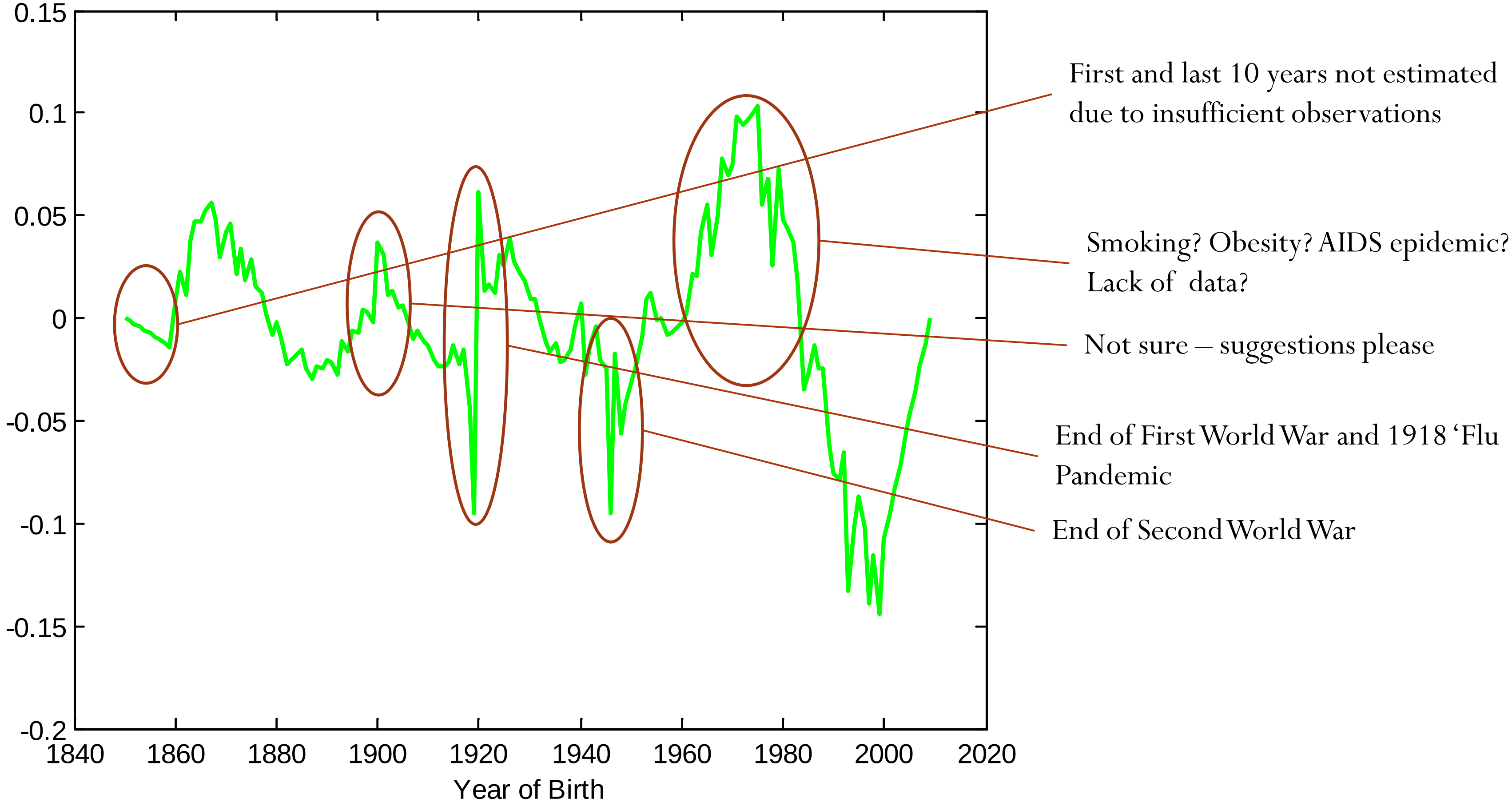


Stage 7 - Test Goodness of Fit

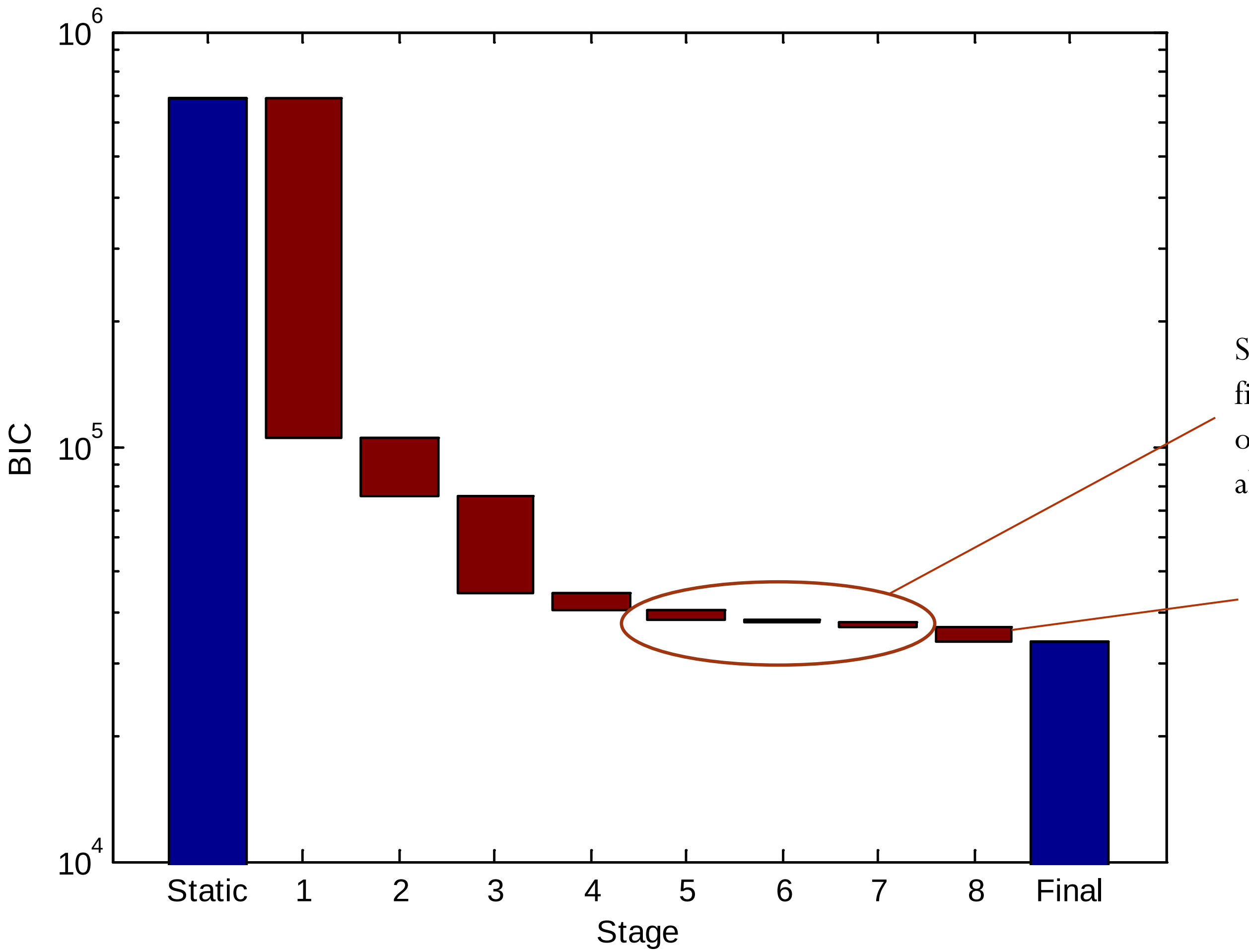
Stopping the Loop



Stage 8 – Add Cohort Parameters



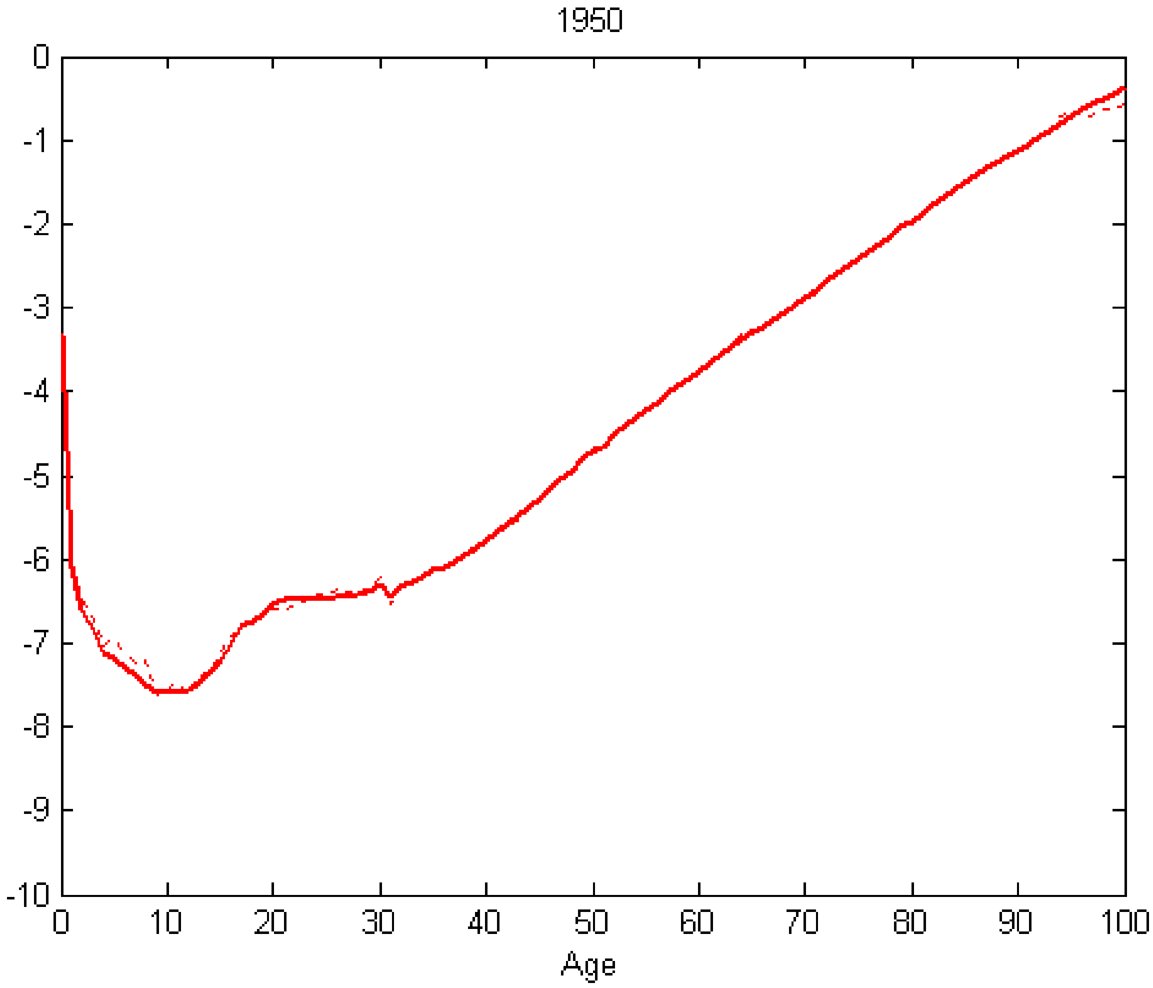
Improvements in Goodness of Fit



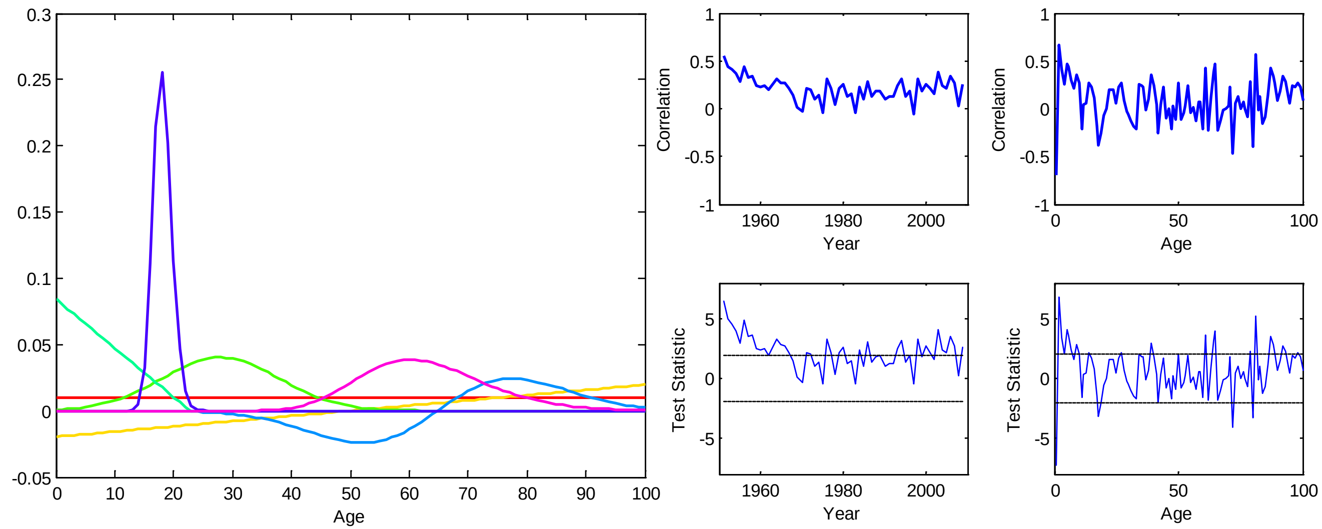
Small impact on goodness of fit but should not be ignored otherwise structure wrongly allocated to cohort term

Impact of adding cohort parameters

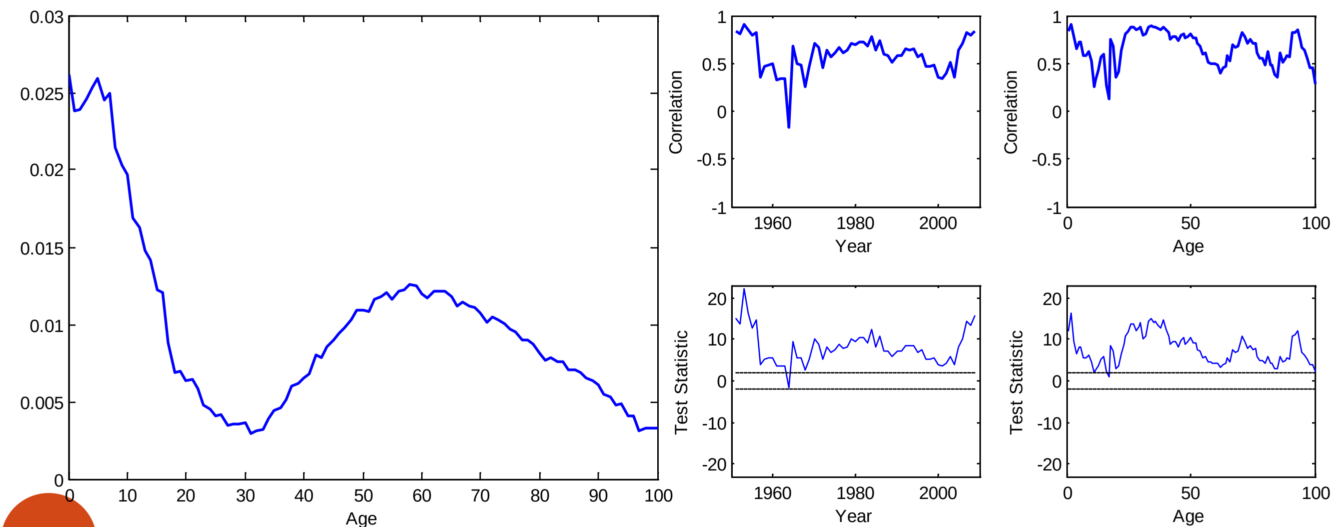
Fitted vs. Observed Mortality



Comparison with Lee-Carter

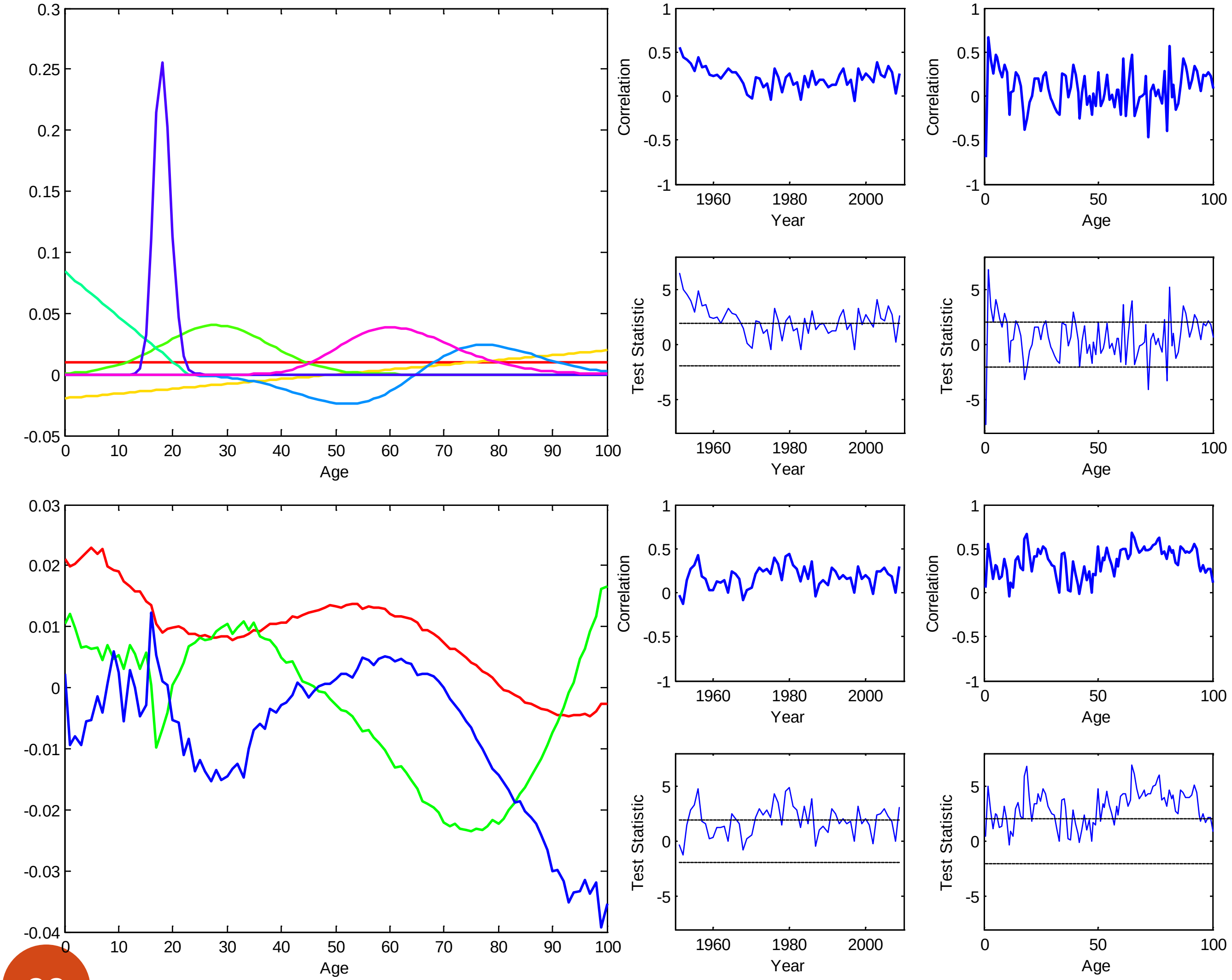


Model BIC	-3.38e+04
Res Mean	-0.01
Res SD	0.94
Res Skew	-0.03
Res Kurt	3.38



Model BIC	-5.25e+04
Res Mean	-0.02
Res SD	0.98
Res Skew	0.47
Res Kurt	9.75

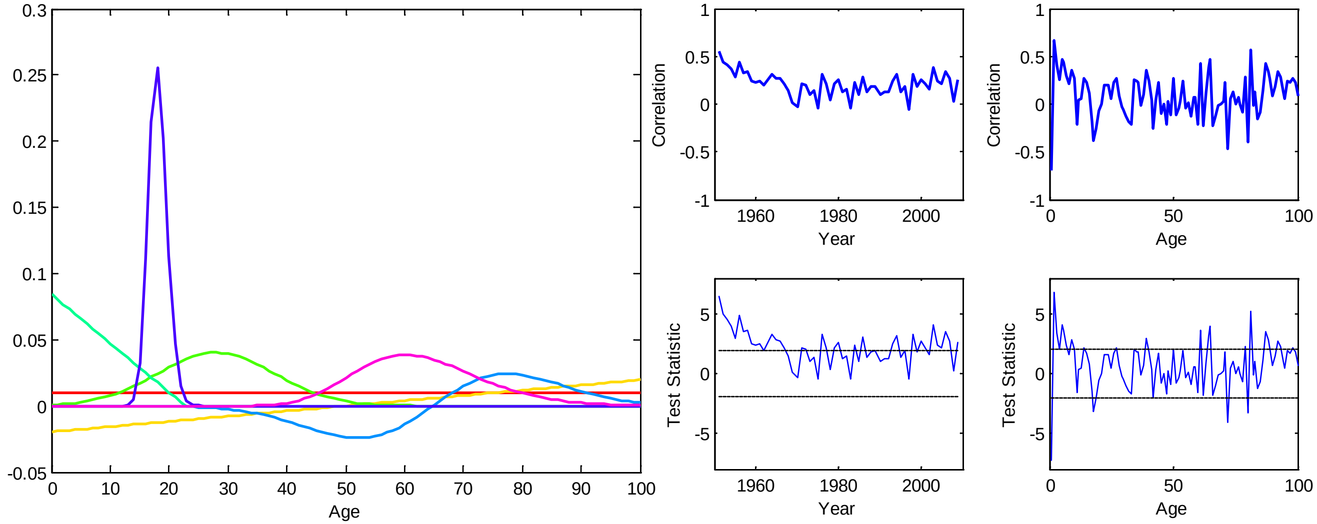
Comparison with PCA



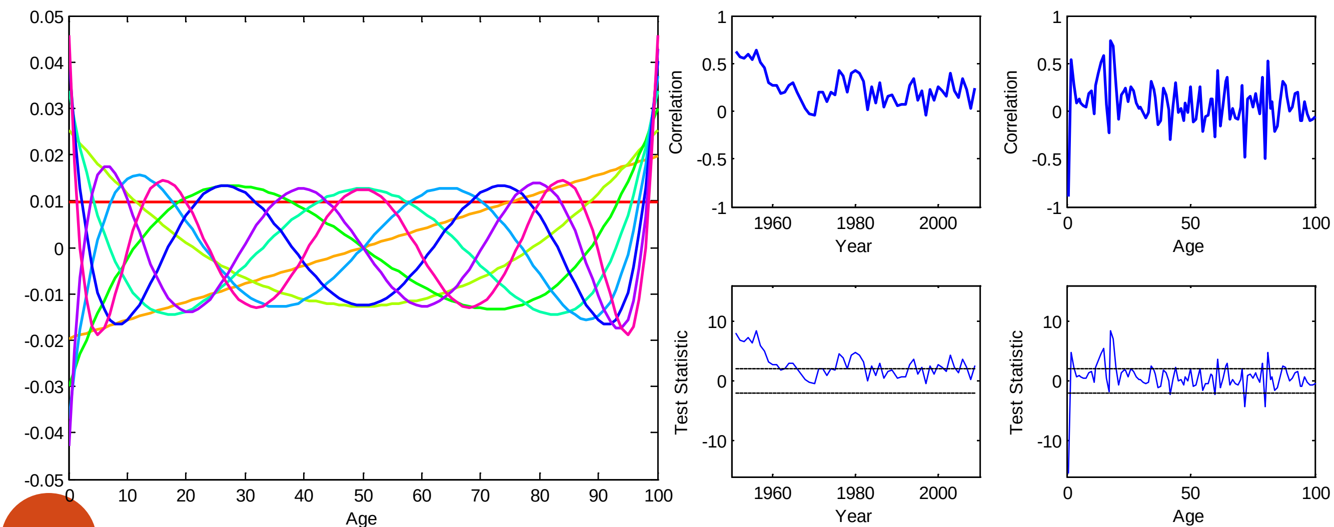
Model BIC	-3.38e+04
Res Mean	-0.01
Res SD	0.94
Res Skew	-0.03
Res Kurt	3.38

Model BIC	-3.39e+04
Res Mean	0.00
Res SD	0.94
Res Skew	0.02
Res Kurt	3.20

Comparison with Legendre Polynomials



Model BIC	-3.38e+04
Res Mean	-0.01
Res SD	0.94
Res Skew	-0.03
Res Kurt	3.38



Model BIC	-3.44e+04
Res Mean	-0.00
Res SD	0.93
Res Skew	-0.17
Res Kurt	3.69

Assessing the Model Selection Criteria

- Adequacy –
  - Each age/period term has been justified statistically as improving the fit to data
- Parsimony –
  - Fewest terms needed to capture information in the data
  - Each age function requires no more than two free parameters
  - Fewer parameters than alternative PCA and Legendre polynomial procedures
- Demographic Significance –
  - Each term we can demonstrate demographic significance in terms of causes of death at relevant ages
- Completeness –
  - Models spans entire age range
  - Models includes allowance for cohort effects and we can link features of this to history of the population

Other Data Sets and Robustness

- Have followed procedure for a female UK data since 1950
- Also obtain mode with seven age/period functions
  - Shape of age/period functions similar to those for male data
  - Sequence of selected functions different (same processes, different importance for men and women?)
- Female data slightly less easy to apply procedure for as some of the trends of comparable size and highly correlated

- Have also tested for parameter uncertainty by changing range of data, using residual bootstrapping technique (Koissi et al (2006)) and by systematically removing ages/years from dataset
- Final model gives parameter estimates which are robust under all approaches

Next Steps

- Remaining structure in the residuals:
  - Geostatistical techniques to analyse correlation structures - Débon et al. (2010)
- Projections:
  - No reliable way of projecting model with multiple time series consistently
  - Nielsen and Nielsen (2010) show that suitable time series depend upon identification constraints
  - Intuition on cohort parameters implies mean-reverting structure
  - Structural breaks and changes of trend need to be allowed for
  - Modelling two populations (e.g. UK men and women) coherently with linkages between similar age/period terms

Selected References

- Brouhns, N., Denuit, M., Vermunt, J., 2002a. A Poisson log-bilinear regression approach to the construction of projected lifetables. Insurance: Mathematics and Economics 31 (3), 373-393.
- Cairns, A., Blake, D., Dowd, K., 2006a. A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. Journal of Risk and Insurance 73 (4), 687-718.
- Cairns, A., Blake, D., Dowd, K., Coughlan, G., Epstein, D., Ong, A., Balevich, I., 2009. A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. North American Actuarial Journal 13 (1), 1-35.
- Débon, A., Martinez-Ruiz, F., Montes, F., 2010. A geostatistical approach for dynamic life tables: The effect of mortality on remaining lifetime and annuities. Insurance: Mathematics and Economics 47 (3), 327-336.
- Dowd, K., Cairns, A., Blake, D., Coughlan, G., 2011. A gravity model of mortality rates for two related populations. North American Actuarial Journal 15 (2), 334-356.
- Haberman, S., Renshaw, A., 2009b. On age-period-cohort parametric mortality rate projections. Insurance: Mathematics and Economics 45 (2), 255-270.
- Lee, R. D., Carter, L. R., 1992. Modeling and forecasting U.S. mortality. American Statistical Association 87 (419), 659- 671.
- Nielsen, B., Nielsen, J. P., 2010. Identification and forecasting in the Lee-Carter model.
- O'Hare, C., Li, Y., Jan. 2012. Explaining young mortality. Insurance: Mathematics and Economics 50 (1), 12–25.
- Plat, R., 2009a. On stochastic mortality modeling. Insurance: Mathematics and Economics 45 (3), 393-404.
- Renshaw, A., Haberman, S., 2006. A cohort-based extension to the Lee-Carter model for mortality reduction factors. Insurance: Mathematics and Economics 38 (3), 556-570.
- Tuljapurkar, S., Li, N., Boe, C., 2000. A universal pattern of mortality decline in the G7 countries. Nature 405 (6788), 789-792.
- Wilmoth, J. R., 1990. Variation in vital rates by age, period and cohort. Sociological Methodology 20, 295–335.

Questions?

- Thank you very much for your attention and your feedback