Modelling frontier mortality using Bayesian generalised additive models

Jason Hilton, Erengul Dodd, Jakub Bijak, Jon Forster, Peter W.F Smith

University of Southampton

Motivations

- Forecasts should incorporate uncertainty about their projections
- This can aid decision-makers manage risks of under and over-predictions
- In combination with elicited cost functions, probabilistic forecast can facilitate formal decision analysis
- Bayesian statistics provides a natural framework for coherent incorporation of uncertainty from different sources
- Borrowing strength across countries should improve forecast stability

Mortality Forecasting - Previous Work

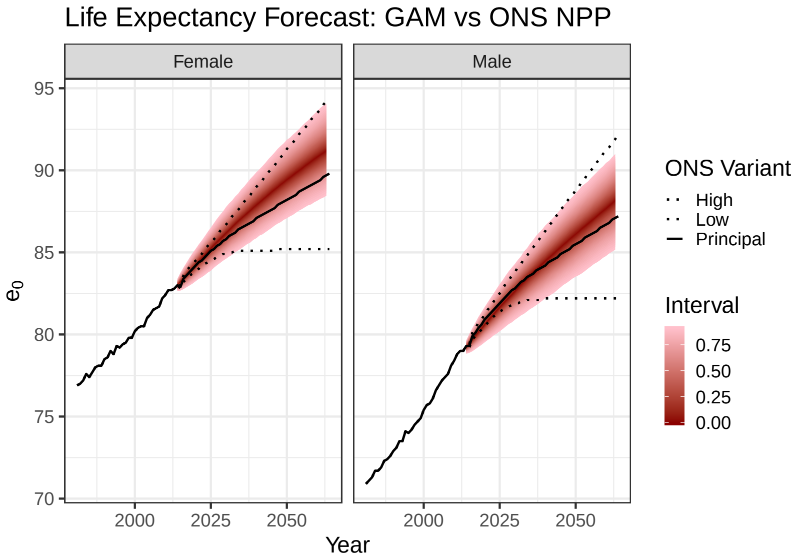
Projecting UK mortality by using Bayesian generalized additive models

Model description

$$log(\mu_{xt}) = s_{\alpha}(x) + s_{\beta}(x)t + s_{\gamma}(t-x) + \kappa_t$$

- Log-mortality μ_{xt} is a sum of smooth functions of age and cohort, a set of smooth age-specific mortality improvements
- Additional period effects κ capturing trend deviations e.g. Flu outbreaks
- Transition from a GAM based model to a parametric logistic model at older ages where data is more sparse
- Smooth functions modeled using B-splines

Life Expectancy



Frontier Mortality

- ‘Frontier’ mortality displays regularities over time (Oeppen and Vaupel 2002)
- Models of frontier mortality rates can aid benchmarking at the national level
- The regularities identified may have utility in forecasting (Bijak 2004, Torri and Vaupel 2012)
- Relies on long-term continuation in mortality rate decline

Definition of *Empirical* mortality frontier:

$$m^*_{x,t} = \min_c(m_{x,t,c})$$

Existing work

Other work has attempted to make use of frontier mortality:

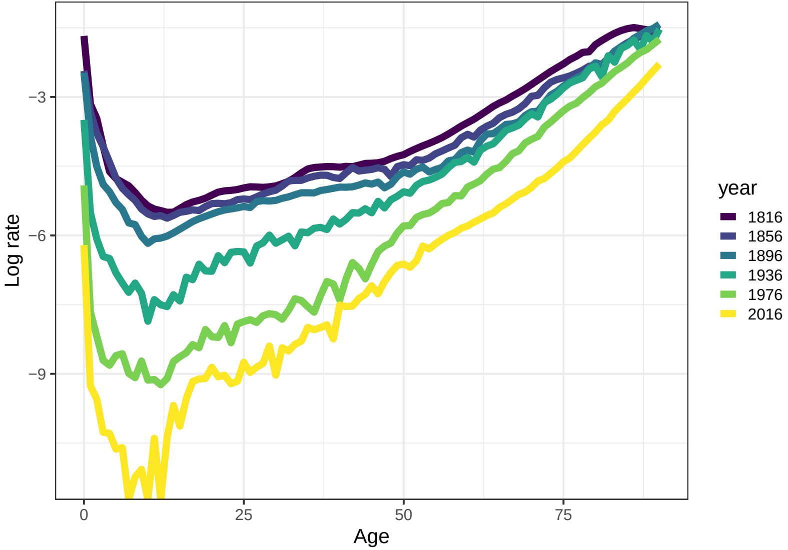
- Bijak (2004) assumed long-term convergence towards frontier life expectancy
- Torri and Vaupel (2012) modelled frontier life expectancy and the gap between this and country values using log-transform
- Pascariu, Canudas-Romo and Vaupel (2018) employed a two-gap model to include males and ensure forecast coherence
- Ebeling (2018) fitted smooth 2-d splines to mortality rate surfaces to identify the location of minimum mortality

Data

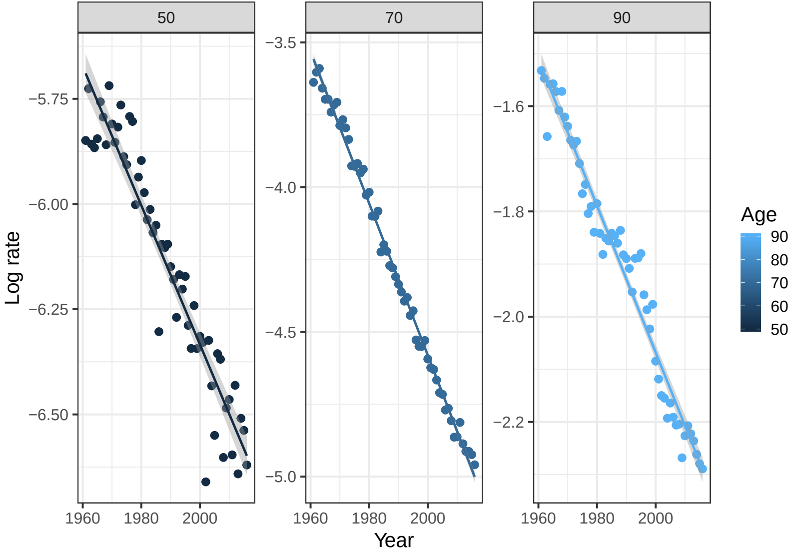
Data from Human Mortality Database from across the developed world provide the opportunity to jointly model the frontier and individual country rates

- Nineteen countries with reasonable population sizes and data going back to 1961 are chosen.
- Female data only
- Men are unlikely to contribute to the frontier given their higher mortality

Frontier Mortality



Frontier Mortality



Model specification

Likelihood

$$D_{xtc} \sim \text{Negative Binomial}(m_{xtc} R_{xtc}, \exp(\phi)).$$

Rate Model

- Frontier model
- Smooth country-specific model
- Country-specific period effects

$$\log(m_{xtc}) = s_\mu(x) + s_\beta(x)t + s_\gamma^c(x)\exp(h(x,t)) + \kappa_{tc}$$

Model specification

The frontier model shares elements with previous work

- Smooth age-specific pattern of mortality
- Smooth age-specific mortality improvements

The country-specific element is constrained to be positive

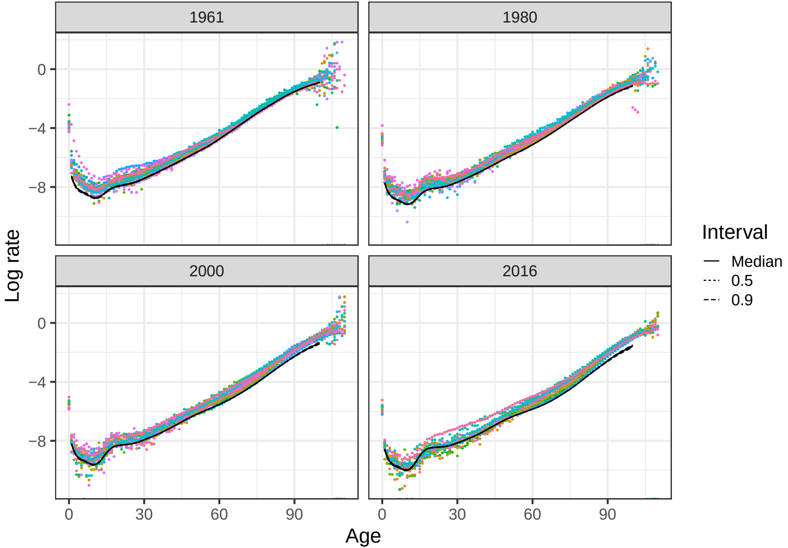
- The coefficients on the age pattern of country-specific deviations $s^c_\gamma(x)$ are forced to positive
- This ensures the frontier lies below all country specific surfaces
- Different choices possible for $h(x, t)$ function describing time evolution of deviations

Model specification

Without further priors and constraints, the model is unidentified, as the frontier could lie anywhere below the country specific rates

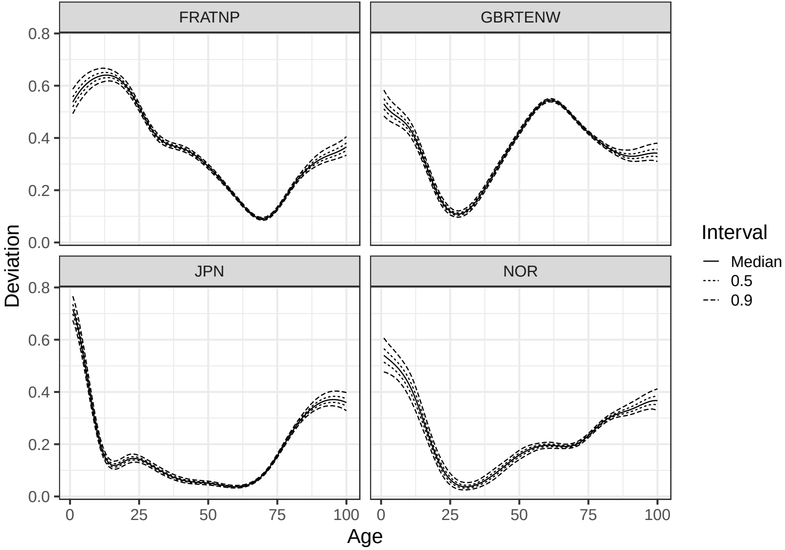
- The size of coefficients on country-specific deviations is penalised so that smaller values are favoured
- Period effects are constrainted to sum to zero and have zero linear and quadratic components
- In addition, standard smoothness penalties are employed for all spline coefficients
- Linear and quadratic age-specific functions are trialled for $h(x, t)$

Results



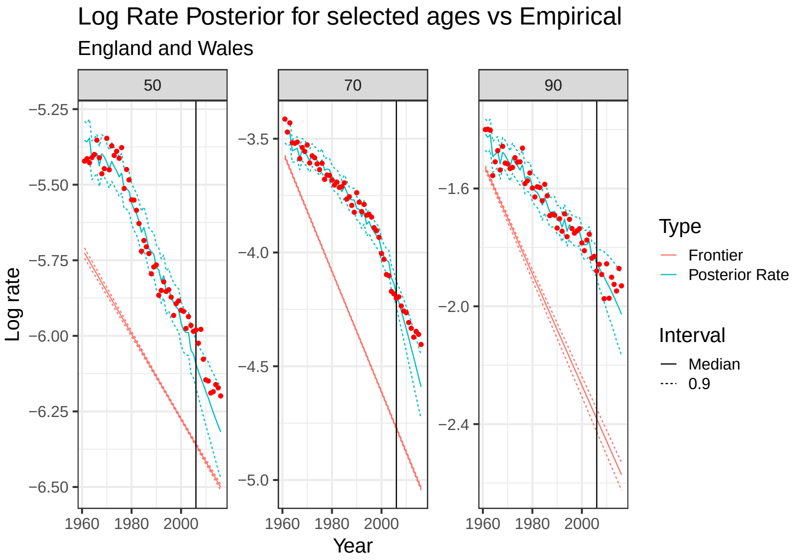
1961
1980
2000
2016
Log rate
0
−4
−8
0
30
60
90
Age
Interval
Median
0.5
0.9

Results



FRATNP
GBRTENW
JPN
NOR
Deviation
Age
0.0
0.2
0.4
0.6
0.8
0
25
50
75
100
Interval
Median
0.5
0.9

Results



Log Rate Posterior for selected ages vs Empirical
England and Wales
50
70
90
Log rate
−5.25
−5.50
−5.75
−6.00
−6.25
−6.50
−3.5
−4.0
−4.5
−5.0
−1.6
−2.0
−2.4
1960
1980
2000
Year
Type
Frontier
Posterior Rate
Interval
Median
0.9

Discussion

- Combining information across countries allows robust estimates of smooth mortality frontier
- Bayesian framework allows posterior estimates incorporating all sources of uncertainty
- Country-specific mortality estimated jointly with the frontier
- Assumptions about frontier avoids extrapolations of recent steep declines indefinitely
- Assessment suggests the frontier model outperforms simpler model fitted to each country separately

Strenths and Limitations

- Quadratic model can lead to problematic extrapolations
- Assumption of linear frontier change
- Frontier estimates dependent on choice of countries in the data

Future work

Country-specific elements

Greater flexibility required in country-specific trends

- Lee-carter model
- Currie model (2-dimensional spline)

Stationary priors to enforce cohorence in forecasts

Frontier elements

Identifying changepoints in development of the frontier (more speculative)

Other

Double-gap two sex model (Pascariu et al. 2018)

Acknowledgements

Code at github.com/jasonhilton/frontier_mort

This work was supported by the ESRC Centre for Population Change - phase II (grant ES/K007394/1) at the University of Southampton. The use of the IRIDIS High Performance Computing Facility, and associated support services at the University of Southampton, in the completion of this work is also acknowledged. All the views presented in this paper are those of the authors only.