Comparing Fertility Forecasting Methods

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

CPC centre for population change

University of Southampton

Motivations

- Fertility forecasting is vital for policy and planning
- Quantitification of uncertainty allows proper management of risks
- Quantitative assessments of the utility of fertility forecasting methods are invaluable
- Theses assessments should be conducted at the level of the underlying Age-Specific Fertility Rates
- Assess the efficacy of Bayesian Parametric methods across a range of countries

Fertility Forecasting Methods

Booth (2006) and Bohk-Ewald et al. (2018) provide reviews.

Several types of (overlapping) approach

- Lee-carter / principal component / functional data models.
  - Lee (1993), Hyndman and Ullah (2007), Wisniowski et al. (2016)
- Other extrapolative approaches (ARIMA, linear extrapolation)
  - de Beer (1985), Myrskylä et al. (2013)
- Methods which complete cohort fertility schedules.
  - Hoem (1981), Evans(1986), Li and Wu (2003), Peristera and Kostaki (2007)...
- Bayesian methods which borrow strength across countries
  - Schmertmann (2014), Ševčíková (2016)

Method Comparison

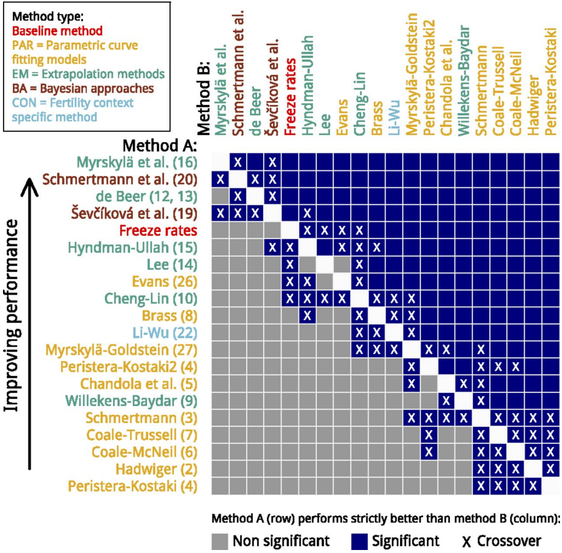
Recent comparison of methods (Bokh-Ewald et al. 2018) shows that few perform better than naive 'freeze rates'.

Parametric models perform poorly, perhaps because they allow no dependence between cohorts.

Bayesian approaches that borrow strength across countries and simple extrapolative methods perform best.

- **Myrskylä et al. (2013)**: Projects recent slope of each Age-Specific Rate forward for 5 years
- **Schmertman et al. (2014)**: Makes projections based on priors that penalise deviation from historically plausible schedules and time series profiles

Method Comparison

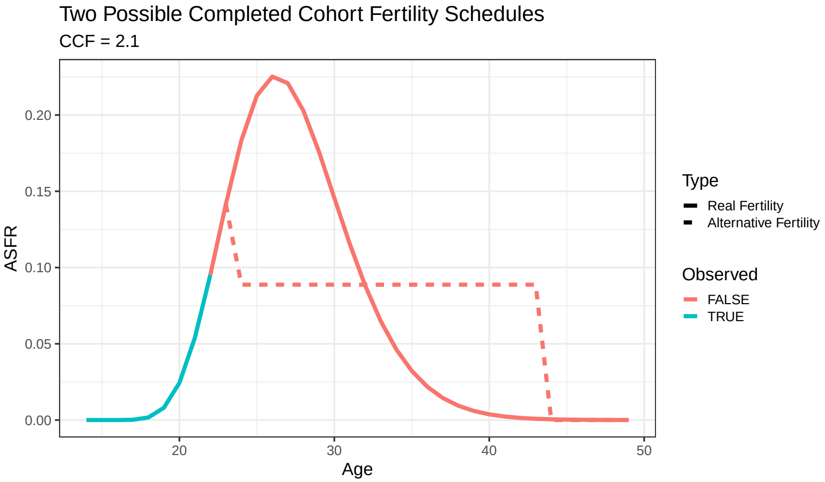


{ Bokh-Ewald et al. 2018, PNAS, 201722364 Image reproduced under CC-BY-NC-ND license.] }

Limitations of Bokh-Ewald et al.

- This paper only assesses the ability of forecasting methods to complete cohort fertility
- Planners need forecasts of numbers of births in particular years depend on age-specific fertility
- Cohort component projections will also require accurate age-specific rates and forecasts of new cohorts
- The parametric models tested do not include dependences between cohorts

Limitations



Objectives

- Compare the best performing forecasting models of fertility on their ability to predict ASFR (not completed fertility)
- Examine the performance of different specifications of a Bayesian parametric mixture model of fertility drawing on work by Hoem et. al 1981, Chandola 1999, and others.
- Provide coverage assessments for these models across a range of countries.

Bayesian parametric mixture models

1. We forecast the overall *level* of fertility using time series methods.
2. The shape of the age-specific fertility curve, whether uni-modal or bi-modal, can be modelled using parametric functions.
3. A *mixture* of parametric functions is therefore used to *decompose* total fertility across ages.
4. The locations and scales of these functions, together with the mixture weighting parameter, are also modelled using standard time series methods.
5. *Either period or cohort* can be used as the forecasting axis

Specification (1)

$$B_{xt} \sim \text{Negative Binomial}(R_{xt}\lambda_{xt},\ \exp(\phi_x))$$
$$\lambda_{xt} = \theta_t \tau_{xt}$$
$$\sum_{x=14}^{X} \tau_{xt} = 1 \ , \forall t$$

Total Fertility θ for a given period (or cohort) is *decomposed* across age according to age specific proportions $\boldsymbol{\tau}$ to obtain the age-specific fertility rates $\boldsymbol{\lambda}$.

Births are modelled as following a Negative Binomial distribution. This allows for age-specific over-dispersion relative to the Poisson distribution.

Specification (2)

$$\lambda_{xt} = \theta_t \tau_{xt}$$
$$\lambda_{xt} = \theta_t \left[\psi_t \; f_1(x;\; \mu_t^{(1)}, \tau_t^{(1)}) + (1 - \psi_t) \; f_2(x;\; \mu_t^{(2)}, \tau_t^{(2)})\right]$$
$$\psi \in [0, 1]$$

Here, total fertility is decomposed using a *mixture* of two parametric functions f_1 and f_2, with the mixture parameter ψ controlling the relative weight of the functions.

The functions f must sum (integrate) to one over age, and are parameterised by location and spread parameters μ and τ.

Gamma and Weibull functions are tested - 4 possible configurations

Estimation

Data

- Births and Exposures for 25 countries from the Human Fertility Database.
- Most recent 10 years of data held back (i.e. not used for fitting) for each assessment

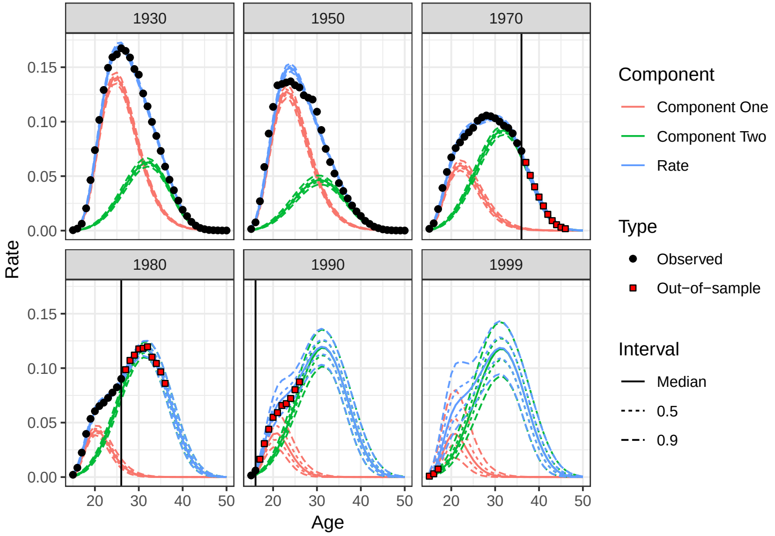
Estimation of parametric mixture models

- Posterior samples obtained using the `stan` software package.
- Convergence to stable distribution is relatively fast.

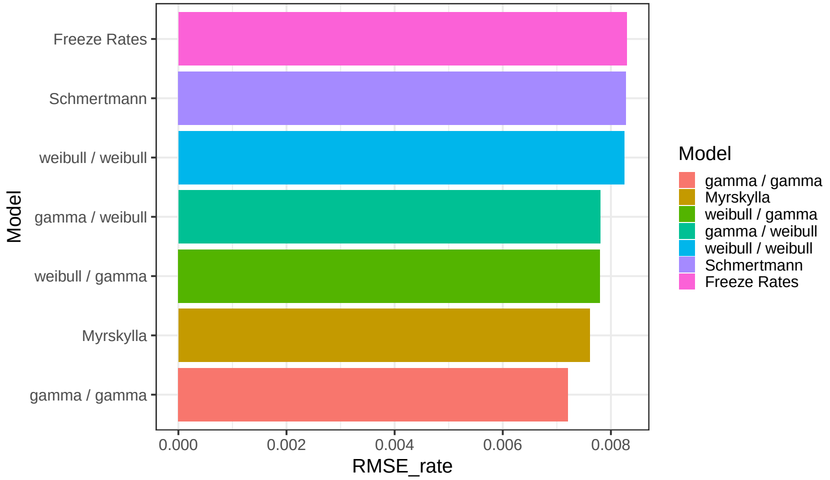
Evaluation

- All models assessed based on average RMSE over the out-of-sample data accross 25 countries.
- Coverage also reported - the proportion of held-back observations falling within particular intervals (e.g. 90%, 50%)

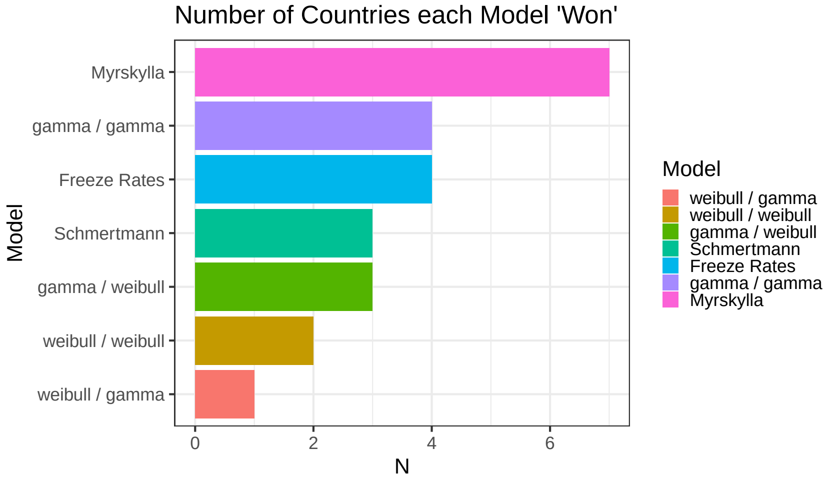
Example - England and Wales



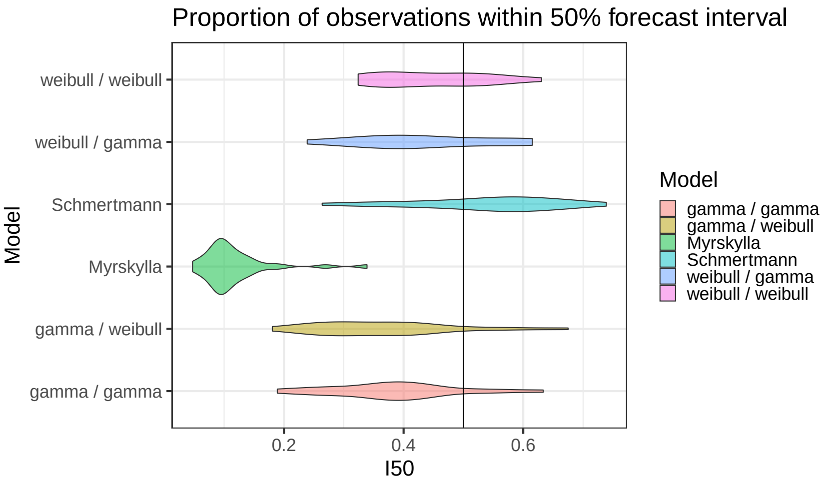
Out of Sample Comparisons - RMSE



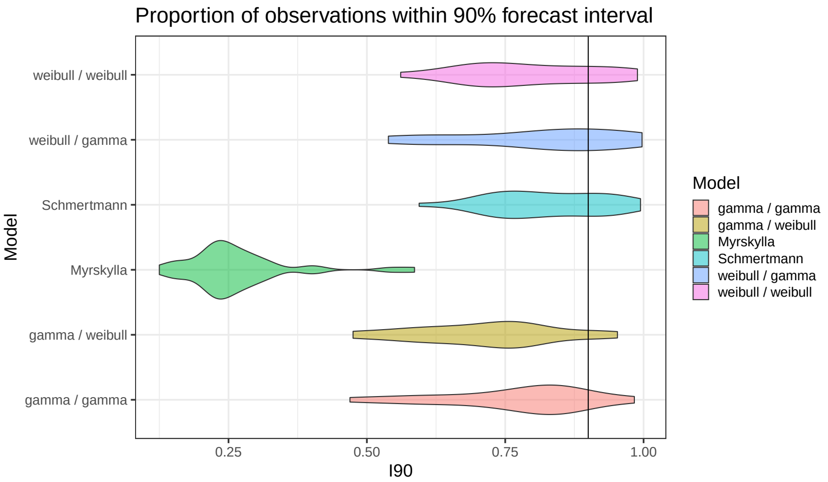
Out of Sample Comparisons - Winners



Out of Sample Comparisons - Coverage 50%



Out of Sample Comparisons - Coverage 90%



Conclusion

- Assessments of fertility forecasting models should consider errors and coverage in Age-Specific-Fertility-Rates as well as completed fertilities
- Bayesian Parametric Mixture Models are a plausible alternative to the best performing fertility forecasting models
- Gamma-gamma mixtures and the model of Myrskylla et al. appear best for forecasting

Future work

- Period based forecasting
- ARIMA models of most important components, for instance (1,1,0)

Acknowledgements

Code at github.com/jasonhilton/fertmix

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.

Key References

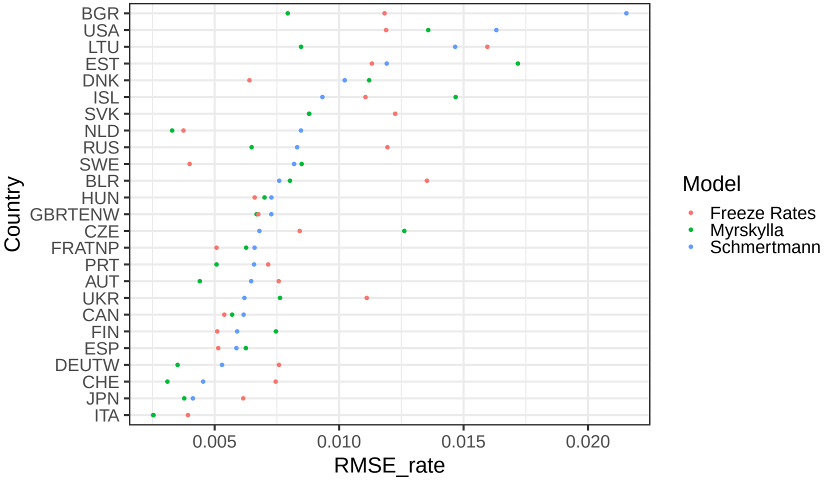
Bohk-Ewald, C., Li, P., & Myrskylä, M. (2018). Forecast accuracy hardly improves with method complexity when completing cohort fertility. Proceedings of the National Academy of Sciences, 201722364. Chandola, T., Coleman, D. A., & Hiorns, R. W. (1999). Recent European fertility patterns: fitting curves to "distorted" distributions. Population Studies, 53(3), 317–329.

Hoem, J. M., Madsen, D., Nielsen, J. L., Ohlsen, E.-M., Hansen, H. O., & Rennermalm, B. (1981). Experiments in Modelling Recent Danish Fertility Curves. Demography, 18(2), 231–244.

Myrskylä, M., Goldstein, J. R., & Cheng, Y.-H. A. (2013). New Cohort Fertility Forecasts for the Developed World: Rises, Falls, and Reversals. Population and Development Review, 39(1), 31–56.

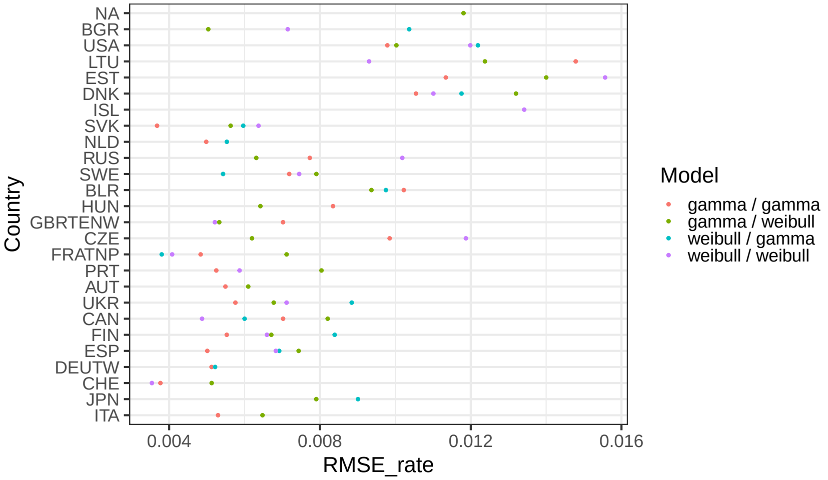
Schmertmann, C., Zagheni, E., Goldstein, J. R., & Myrskylä, M. (2014). Bayesian Forecasting of Cohort Fertility. Journal of the American Statistical Association, 109(506), 500–513.

Extras



BGR
USA
LTU
EST
DNK
ISL
SVK
NLD
RUS
SWE
BLR
HUN
GBRTENW
CZE
FRATNP
PRT
AUT
UKR
CAN
FIN
ESP
DEUTW
CHE
JPN
ITA
Country
0.005
0.010
0.015
0.020
RMSE_rate
Model
Freeze Rates
Myrskylla
Schmertmann

Extras


NA
BGR
USA
LTU
EST
DNK
ISL
SVK
NLD
RUS
SWE
BLR
HUN
GBRTENW
CZE
FRATNP
PRT
AUT
UKR
CAN
FIN
ESP
DEUTW
CHE
JPN
ITA
Country
0.004
0.008
0.012
0.016
RMSE_rate
Model
gamma / gamma
gamma / weibull
weibull / gamma
weibull / weibull