

**Models for migration age schedules: a Bayesian perspective with
an application to flows between Scotland and England.**

**P. Congdon, Dept of Geography, Queen Mary University of London,
Mile End Rd, London E1 4NS; p.congdon@qmul.ac.uk**

Abstract

This chapter will consider alternative models for migration age schedules, using as an example data on migration from Scotland to England. Migration exchanges between Scotland and England exhibit features of international migration flows such as young adult labour migration to England and retirement age return migration to Scotland (Gordon, 1982). A comparison is made between alternative models under a fully Bayesian inference and estimation approach. The data consists of five year age group migration flows but the models are for individual age rates, allowing imputation of the missing single year of age migration data. One model considered is the multiexponential scheme, as in Rogers and Castro (1981). An alternative parametric regression approach retains the Rogers-Castro parameterisation but with gamma mixing on the model rates. The gamma mixing is often necessary to model overdispersion. A contrasting approach is by nonparametric regression involving an unknown smooth function. Inclusion of random effects under either nonparametric or parametric approaches facilitates the pooling of inference strength in joint estimation for similar migration schedules (e.g. when male and female migration rate schedules from the same origin region are likely to be similarly shaped) as opposed to separate estimation of models for each schedule. How the different model schemes perform in terms of fit and model complexity is based on a Bayesian version of the AIC, that adjusts fit to take account of model complexity. A model check based on predictive sampling is also used.

1. Introduction

Schedules of migration rates tend to show regular features such as a peak in the young adult ages and declining migration propensities in old age, though there may be elevated migration around retirement age. These features characterise flows between countries when the young adult flow is associated with labour migration but return migration in later life may reflect enduring cultural or familial affiliations. One way to capture these regularities, whether present in internal or international migration, involves nonlinear regression whereby the rate of migration is represented by a sum of exponential or shifted exponential terms; this is known as the multi-exponential migration schedule (Rogers & Castro, 1981).

This paper considers a Bayesian modelling approach to migration schedules, starting with a purely parametric model but considering also a model that avoids a nonlinear regression specification but instead uses a dynamic general linear

model (DGLM) approach. For related work on mortality graduation see Neves & Migon (2004). The Bayesian approach used here employs iterative (Markov Chain Monte Carlo) sampling to update information about the density of parameters θ from the observed data, so revising existing knowledge summarised in prior densities on the components of θ .

2. Parametric vs DGLM Approaches

Let Y_x be migration flows at individual years of ages ($x=1,..,\omega$), N_x be mid year populations, and assume Poisson sampling with $Y_x \sim \text{Po}(N_x r_x)$ where r_x are migration rates. The multi-exponential model for r_x with both labour and retirement effects can be specified with the form (a 'purely parametric' model)

$$r_x = a_0 + a_1 \exp[-\alpha_1 x] + a_2 \exp[-\alpha_2(x - \mu_2) - \exp\{-\lambda_2(x - \mu_2)\}] + a_3 \exp[-\alpha_3(x - \mu_3) - \exp\{-\lambda_3(x - \mu_3)\}]. \quad (1)$$

The component with parameters a_1 and α_1 represents child migration in the company of young migrant parents, the component with parameters $\{a_2, \alpha_2, \mu_2, \lambda_2\}$ represents young adult migration which is mainly for labour reasons when the migration flows are international, and the retirement age effect is represented by the shifted exponential term with parameters $\{a_3, \alpha_3, \mu_3, \lambda_3\}$. One advantage of the parametric model is that one obtains a "standard" predicted age profile of migration with a labor force peak, even if the data are irregular.

The presence of such irregularities in the observed schedule and/or some degree of overdispersion (excess heterogeneity) indicate that one should often specify a hierarchical model allowing for uncertainty in the r_x after allowing for the parametric form in (1). The conjugate option is to take the r_x as subject to gamma mixing, where a gamma with parameters α and β has mean α/β and variance α/β^2 . Substantively there is often likely to be heterogeneity in migration behaviours at a given age between population subgroups. This leads to the specification

$$y_x \sim \text{Po}(N_x r_x) \quad (2) \\ r_x \sim \text{Ga}(\kappa, \kappa/m_x) \\ m_x = a_0 + a_1 \exp[-\alpha_1 x] + a_2 \exp[-\alpha_2(x - \mu_2) - \exp\{-\lambda_2(x - \mu_2)\}] + a_3 \exp[-\alpha_3(x - \mu_3) - \exp\{-\lambda_3(x - \mu_3)\}]$$

where κ is an additional positive parameter. The variance m_x^2/κ of r_x about the purely parametric model represented by m_x declines as κ increases. An additive (e.g. normal) error in a model for $\log(r_x)$ is another possibility.

While the form of the multiexponential model is based on accumulated substantive evidence it is a preconceived form that presumes a regularity that may not be present in all datasets. There may be virtues in letting the data choose the most appropriate nonlinear regression (Stone, 1985). In particular, one may avoid parametric regression assumptions by using non-parametric regression which for migration events y_x with means $N_x r_x$ has a generic form

$$\log(r_x) = S(x)$$

where $s_x = S(x)$ is an unknown smooth function in age x . For example, one

might consider age effects s_x following a low order random walk (e.g. RW1 or RW2) or an autoregressive model. For example, an RW1 model could take the form

$$\begin{aligned} \log(r_x) &= s_x \\ s_x &\sim N(s_{x-1}, 1/\tau_s), \quad x = 2, \dots, \omega \end{aligned} \tag{3}$$

where τ_s is a precision parameter and s_1 is typically assigned a diffuse prior, such as $s_1 \sim N(0, V_1)$ where V_1 is large. For τ_s one may assume a Gamma or lognormal prior with known parameters; for example, Besag et al (1995) use the prior $\tau_s \sim \text{Ga}(1, b)$ with b small, e.g. $b=0.001$. The initial values of the series are assigned fixed effects priors. The RW2 prior penalizes deviations from a linear trend $s_x = 2s_{x-1} - s_{x-2}$ with

$$s_x \sim N(2s_{x-1} - s_{x-2}, 1/\tau_s), \quad x = 3, \dots, \omega \tag{4}$$

and flat priors on s_1 and s_2 . This will tend to produce greater smoothing than an RW1 model and so may be more sensible if the data are irregular due to a small sample size.

Random walk priors allow flexibility in the face of possible nonstationarity and also generalize easily to multivariate situations. An alternative is to introduce an autoregression parameter. For instance, a lag 1 autoregressive model could be specified as

$$s_x \sim N(\rho s_{x-1}, 1/\tau_s), \quad x = 2, \dots, \omega$$

where one may assume stationarity a priori, or assess it from the posterior probability $Pr(|\rho| < 1|y)$. The latter may be estimated as the proportion of MCMC iterations $t = 1, \dots, T$ subsequent to convergence where the condition $|\rho^{(t)}| < 1$ on sampled autocorrelations $\rho^{(t)}$ obtains.

Nominally the models in (3) and (4) involve $(\omega + 1)$ parameters, as compared to eleven in (1). However, the fact that the s_x are drawn from a single hyperdensity means that their effective dimension is typically considerably less than this. A model such as (3) may yield a better unpenalised fit (e.g., deviance or error sum of squares) than (1) while its effective dimension (or complexity) is not that much greater. Using the approach of Spiegelhalter et al (2002) one may estimate an effective parameter total (d_e) by comparing the deviance at the parameter mean $D(\bar{\theta})$ with the mean deviance \bar{D} , where both are obtained from an MCMC run. So $d_e = \bar{D} - D(\bar{\theta})$. Then a summary measure of fit that penalizes for complexity (in a similar way to the AIC) is the Deviance Information Criterion (DIC), with $DIC = \bar{D} + d_e$. Smaller values of DIC indicate better fit. As a predictive model check (to ensure the model is reproducing the data satisfactorily) one may sample new migration flows and assess whether the 95% interval for such predictions includes the actual data; approximately 95% of the observed data points should lie within the 95% prediction intervals (Gelfand, 1996).

3. Pooling Strength over Different Schedules

One advantage of the model structure exemplified in equation (3) is its simple

extension to situations where there are several schedules migration schedules, for example, migration totals Y_{xg} from Scotland to England by age x and gender g , or migration schedules Y_{xgc} by age, gender and social class c . One may expect modelling assumptions about the underlying rate structure to be exchangeable between schedules specific to socio-demographic groups (for example, male and female young adult migrants from Scotland to England will both be predominantly labour migrants). In particular, one may generalise the models in (3) and (4) to multivariate random walks in age, for example multivariate normal effects with covariance $\Sigma = \Upsilon^{-1}$ expressing the intercorrelation between the migration schedules of say, males and females. Thus the male and female labour force peaks and low elderly propensities are likely to be closely similar.

So, for two schedules Y_{xg} by gender, one may generalise model in equation (3) to become

$$\begin{aligned} \log(r_{xg}) &= s_{xg} & g &= 1, 2 \\ \mathbf{s}_x &\sim N_2(\mathbf{s}_{x-1}, \Upsilon^{-1}), x = 2, \dots, \omega \end{aligned} \quad (4)$$

where $\mathbf{s}_x = (s_{x1}, s_{x2})$ and the precision matrix Υ may be assigned a Wishart prior. By contrast, there is no clear way that model in equation 1 can be estimated in such a way as to pool information over the schedules for different groups, except possibly by a constrained model with some parameters equated between models for different schedules.

The above models are for single year of age migration flows, but migration data are often available only for aggregated age bands, denoted $[x]$, where in the Scotland-England migration application, data are available for five year age bands denoted $[x]=0,1,\dots,17$, for bands 0-4,5-9,...80-84,85+. One may then seek a model to use aggregated migration data $Y_{[x]}$ but with the likelihood based on an individual age model. This amounts to a missing data problem and Bayesian methods provide a relatively straightforward approach to imputing the single year of age flows. However, inferences about the parameters in models such as those above can proceed without imputing the missing single year of age flows.

4 Case Study: Scotland to England Migration in 1990-91

As a case study, data relating to migration from Scotland to England in the year preceding the 1991 UK Census are analyzed. Separate schedules for males and females are available. However, an initial analysis comparing the three models above considers only male flows from Scotland to England. The data are for eighteen five year age groups but a single year of age migration model is used to define the likelihood. Thus, define five year age groups $[x]$, $[x]=0,\dots,17$, centred at single year of age $5x+2.5$ and spanning single years of age $5x+0.5$ to $5x+4.5$. For example, the five year age group [5] consists of single years of age 25,26,...29. Mid-year estimate male populations for 1990 for Scotland (ages 0,1,...89,90+) define the single year of age populations $N_{x+0.5}$.

Then assume the five year migration totals are Poisson with means defined by summed products of one year age group populations and rates:

$$Y_{[x]} \sim Po(N_{5x+0.5}r_{5x+0.5} + N_{5x+1.5}r_{5x+1.5} + N_{5x+2.5}r_{5x+2.5} + N_{5x+3.5}r_{5x+3.5} + N_{5x+4.5}r_{5x+4.5}) \quad (5)$$

with r_x as defined in equations 1 to 3. Notice that this is not an assumption of multinomial sampling in which the five single year unknown migration flows sum to a known total $Y_{[x]}$. The model form in (5) is consistent with the likelihood assumptions for single year of age migration data as outlined in section 2, since if $Y_{5x+0.5} \sim Po(N_{5x+0.5}r_{5x+0.5})$, $Y_{5x+1.5} \sim Po(N_{5x+1.5}r_{5x+1.5})$, etc, then the total $Y_{[x]}$ will be Poisson as in (5). Since $N_{x+0.5}$ are input data (mid year official population estimates by single year of age) and single year of age rates $r_{x+0.5}$ define the model, predictions of the missing one year migration data can be sampled. However, one may estimate the parameters assumed to generate the $r_{x+0.5}$ without sampling the missing one year migration data. A model check is based on sampling new five year migration totals or 'predictions', namely $Y_{new,[x]}$, obtaining their 95% intervals, and checking whether the observation $Y_{[x]}$ is included in the 95% interval of the prediction.

For the Scotland to all England flows it is assumed that no retirement element is present, so providing a seven parameter version of equation 1, namely (Single Outcome Model 1, or S1 for short)

$$r_x = a_0 + a_1 \exp[-\alpha_1 x] + a_2 \exp[-\alpha_2(x - \mu_2) - \exp\{-\lambda_2(x - \mu_2)\}]$$

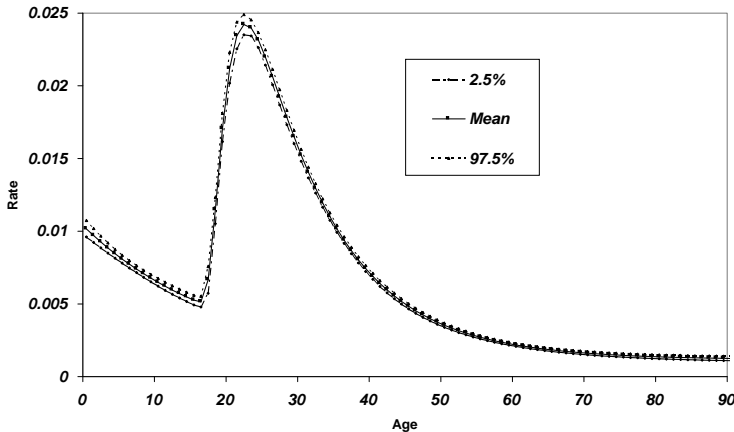
for ages $x=0.5, 1.5, \dots, 90.5$. Model 1 is estimated using priors on $\theta = \{a_0, a_1, \alpha_1, a_2, \alpha_2, \mu_2, \lambda_2\}$ based on results about typical parameter values presented in Rogers and Raymer (1999) and Rogers et al (2004). The baseline (theoretical minimum) rate a_0 is constrained to have a maximum of 0.005. Furthermore $a_1 \sim Ga(0.05, 1)$, $a_2 \sim Ga(0.075, 1)$, $\mu_2 \sim Ga(20, 1)$, $\lambda_2 \sim Ga(0.425, 1)$, $\alpha_1 \sim Ga(0.05, 1)$, $\alpha_2 \sim Ga(0.15, 1)$ and $a_0 \sim Ga(0.001, 1) I(0, 0.005)$.

A two chain run of 10,000 iterations shows early convergence using Gelman-Rubin diagnostics (Gelman et al, 1995). The resulting estimated r_x is highly stylized (Figure 1) with precisely defined estimates of rates. The rate plot shows a disjunction at young adult ages, reflecting the much larger Scotland-England migration total for ages 15-19 than for ages 10-14. One might consider some form of moving average to reduce this effect. Table 1 contains the estimated parameters and their 95% credible intervals. It may be noted that only 13 of the 18 actual flows are included in the 95% intervals of $Y_{new,[x]}$ so the model has some inconsistencies with the data - in particular, the model rates are too precise to reproduce the actual observations. The DIC is 219 with $d_e=7$.

Table 1 here

Alternatively consider the gamma mixture form in equation (2). This constitutes model S2 for male Scotland-England flows. Priors are as in model 1, with additionally $\kappa \sim Ga(1, 0.01)$. The resulting model schedule is smoother in the late teens though less precise (i.e. having wider 95% intervals) than under

Figure 1 Multi-exponential Model S1, Scotland to England, Males



model 1 (see Figure 2). The model dimension increases from 7 to 15.4, but the fit is improved with $D(\bar{\theta})=156.8$ rather than 206. The DIC falls from 219 to 187.6. This model has better predictive accuracy (checks better against the data) with 17 of the 18 actual flows now included in the 95% intervals of $Y_{new,[x]}$. The exception is the oldest age group for whom migration is underpredicted.

By contrast, models S3 and S4 are nonparametric regressions involving respectively first and second order random walks in the log scale; see equation (3). A Gamma prior on the precision τ_s is assumed with shape parameter 1 and scale parameter 0.1. The posterior mean of the precision parameter under model S3 is 24.5 with 95% interval (12,42). The DIC for this model improves on that for model 2, namely 176.7 ($d_e=12.8$, $D(\bar{\theta})=151.1$). As might be expected the resulting migration schedule is less smooth than under models 1 and 2 (see Figure 3), but in its central features very similar in form. The plot shows a slight upturn in migration rates among the very old, and now all 18 of the actual flows are included in the 95% intervals of $Y_{new,[x]}$.

A slightly worse DIC of 181 (with $d_e=17$) is obtained with an RW2 model. However, the resulting schedule of rates (Figure 4) is preferable on substantive grounds in showing a smoothly changing migration propensity over successive ages. Figure 4 shows features that are not reproduced by the multi-exponential model, such as a relatively flat propensity in young child rates rather than the decline at these ages shown in Figure 1; Figure 4 also shows a sharper peak in the young adult propensity.

Figure 2 Multi-exponential Model with gamma mixing, Scotland to England, Males

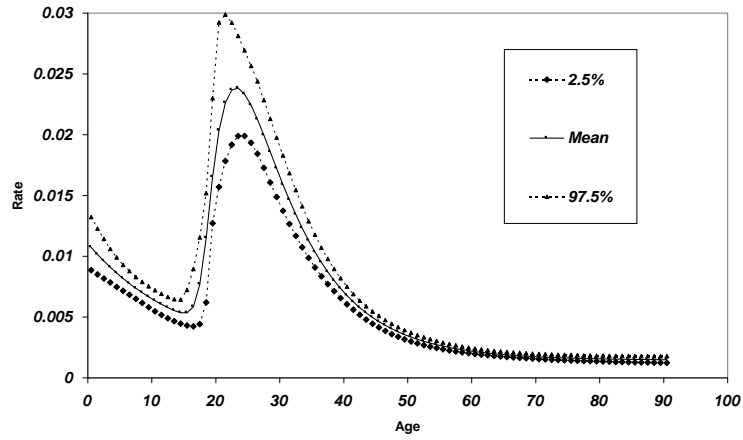


Figure 3 DGLM with RW1 Age effects (Model S3), Scotland-England, Males

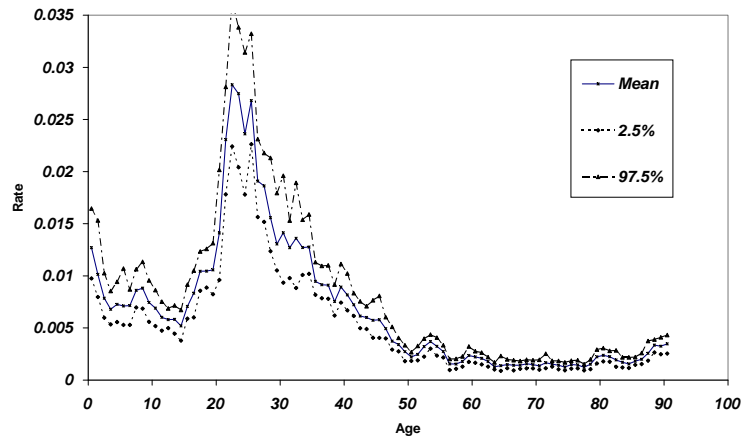
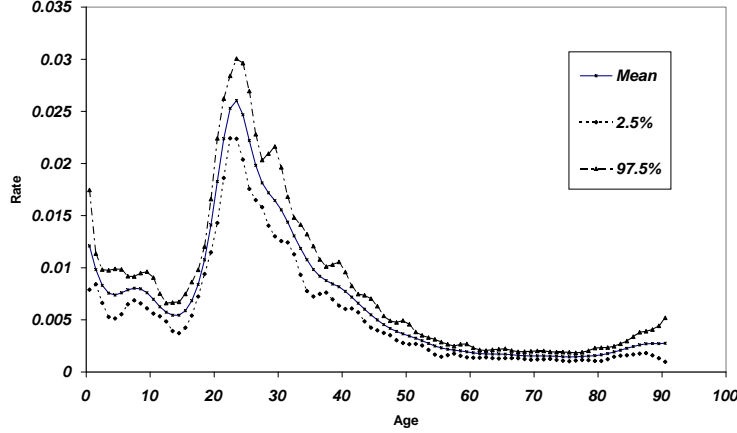


Figure 4 DGLM with RW2 Age effects (Model S4), Scotland-England, Males



5 Multivariate (Multiple Schedule) Model Estimates

Migration schedules are often likely to show similar features which suggests benefits from modelling two or more schedules jointly. For a large collection of schedules one might consider random variation in the parameters of a multi-exponential model. However, this section consider a relatively limited multivariate problem, namely joint modelling of male and female out-migration schedules from Scotland to England. The input data now includes single year of age mid-year population estimates $\{N_{xg}, g = 1, 2\}$ for both males and females for 1990. In the pure parametric model for these data a distinct parameterisations for males and females is assumed (Joint Model 1, or J1 for short)

$$r_{x1} = a_{01} + a_{11} \exp[-\alpha_{11}x] + a_{21} \exp[-\alpha_{21}(x - \mu_{21}) - \exp\{-\lambda_{21}(x - \mu_{21})\}]$$

$$r_{x2} = a_{02} + a_{12} \exp[-\alpha_{12}x] + a_{22} \exp[-\alpha_{22}(x - \mu_{22}) - \exp\{-\lambda_{22}(x - \mu_{22})\}]$$

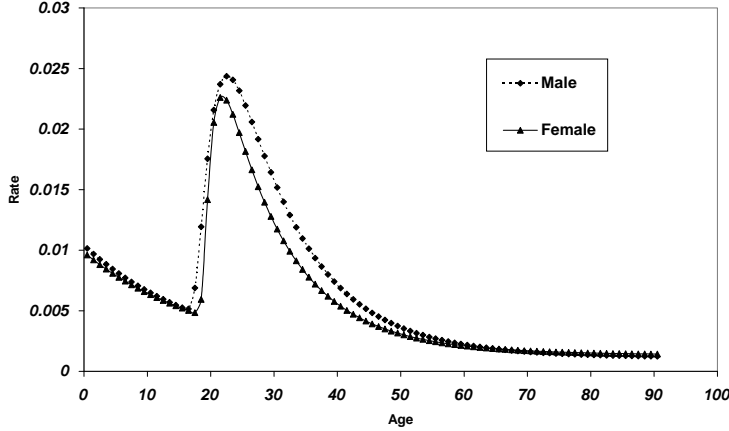
Let θ_1 be the set of male parameters and θ_2 be the female parameters, each of dimension 7. A feature of MCMC estimation by repeated sampling is that one can assess whether particular parameters in the male and female models (say the k^{th}) are effectively equal by monitoring the proportion of iterations where θ_{1k} exceeds θ_{2k} .

If the posterior probability $\Pr(\theta_{1k} > \theta_{2k} | Y)$ is inconclusive (e.g. between 0.1 and 0.9) then there may be grounds for equating the parameters between the two models. One may similarly monitor model outputs of substantive importance, such as the gross migraproduction rate (GMPR) by gender

$$G_g = \sum_{x=1}^{\omega} r_{xg}$$

and assess (for example) the probability $\Pr(G_1 > G_2 | Y)$ that the male GMPR exceeds the female one.

Figure 5 Joint Outcomes, Estimated Schedules by Gender, Purely Parametric Model

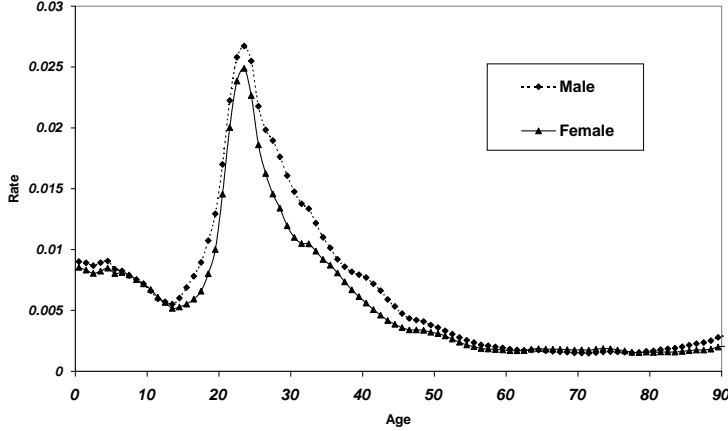


In estimation, the same priors as used in Model S1 are assumed, and the 2nd half of a two chain run of 10000 iterations used for inferences. Note that the model does not check adequately against the data with only 27 of the 36 observations lying within the 95% credible intervals of $Y_{new[x]}$. Bearing this limitation in mind, the results shows that the $\Pr(\theta_{1k} > \theta_{2k}|Y)$, $k = 1, \dots, 7$ rule out equating parameters with the clear exception of the decay parameters of the child component, namely $\{\alpha_{11}, \alpha_{12}\}$ where $\Pr(\theta_{1k} > \theta_{2k}|Y) = 0.62$, and the posterior means on the two parameters are 0.052 and 0.051 respectively. There are marginal findings (in terms of whether parameters are significantly different) with regard to the baseline constants $\{a_{01}, a_{02}\}$ where $\Pr(\theta_{1k} > \theta_{2k}|Y) = 0.07$, and the young adult exponential decay parameters $\{\alpha_{21}, \alpha_{22}\}$ where $\Pr(\theta_{1k} > \theta_{2k}) = 0.10$. The evidence that the male GMPR exceeds the female one is conclusive with $\Pr(G_1 > G_2|Y) = 1$ (see also Figure 5). Because of the similarity of the α_1 parameters the pure parametric model is rerun with $\alpha_{11} = \alpha_{12}$; this leads to a slight fall in DIC, from 432.5 to 430.9, with d_e falling from 13.7 to 12.7.

The next stage is to apply the gamma mixture to the multivariate analysis with the equality constraint on α_1 . However, monitoring $\Pr(\theta_{1k} > \theta_{2k}|Y)$, where θ_1 and θ_2 now include the precision parameters κ_1 and κ_2 , suggests that when the gamma mixing on r_{xg} is introduced, most parameters (including the κ_g) could be equated. The exception is the rate parameters $\{a_{21}, a_{22}\}$ of the young adult migration component, reflecting greater young adult migration intensity for males. So a reduced model is applied (model J2) with

$$\begin{aligned} r_{xg} &\sim Ga(\kappa, \kappa/m_{xg}) \\ m_{x1} &= a_0 + a_1 \exp[-\alpha_1 x] + a_{21} \exp[-\alpha_2(x - \mu_2) - \exp\{-\lambda_2(x - \mu_2)\}] \\ m_{x2} &= a_0 + a_1 \exp[-\alpha_1 x] + a_{22} \exp[-\alpha_2(x - \mu_2) - \exp\{-\lambda_2(x - \mu_2)\}]. \end{aligned}$$

Figure 6 Joint Outcomes, Estimated Schedules (Posterior mean rates) by Gender, Multivariate Random Walk



The complexity of this model exceeds that of the pure parametric model with $d_e = 28.1$ but $D(\bar{\theta})$ is reduced to 309.1 with a DIC of 365.4. Now 35 of the 36 data values are contained within the 95% intervals of predictions from the model, the exception being the oldest male group (over 85s).

Finally a multivariate generalisation of the first order random walk model is fitted (model J3). Thus

$$\begin{aligned} \log(r_{xg}) &= s_{xg} & g &= 1, 2 \\ \mathbf{s}_x &\sim N_2(\mathbf{s}_{x-1}, \Upsilon^{-1}), & x &= 2, \dots, \omega \end{aligned}$$

where $\mathbf{s}_x = (s_{x1}, s_{x2})$ and the precision matrix $\Upsilon = \Sigma^{-1}$ is assumed to be Wishart with scale matrix $0.1I$ and 2 degrees of freedom. The complexity is similar to that for model 2 ($d_e = 29$) but $D(\bar{\theta})$ is now 300, giving a DIC of 358, better than for model J2. All 36 observations are within the 95% predictive intervals under this model. The 95% interval for males over 85 is (26,58) with median 41 as compared to the actual flow of 41.

Figure 6 confirms the similar shape of the two schedules and the higher GMPR of males, confirmed by a posterior mean 0.618 with 95% interval (0.607,0.629) compared to females, 0.532 with 95% interval (0.522,0.541). The interrelation of schedules is apparent in posterior means of Σ_{11} , Σ_{22} , and Σ_{12} of 0.039, 0.045 and 0.031, and a correlation of 0.75 (with 95% interval 0.42, 0.90) between the male and female migration rate schedules r_{x1} and r_{x2} . Compared to Figure 5 the young adult propensity is more peaked and the child rates less steeply declining.

6 Discussion and Conclusions

This paper has discussed and implemented model options that go beyond purely parametric forms developed by Rogers and Castro (1981). The MCMC sampling perspective to estimate model parameters has illustrated the additional benefits of a modern Bayesian approach, such as the assessment of hypotheses (e.g. on the relative sizes of GMPRs) that might be difficult in classical estimation. Model choice and comparison for models involving random effects is also difficult under the classical perspective. The analysis of this paper is based on the freeware package WINBUGS which facilitates Bayesian estimation without the need to specify the full conditional densities or develop application specific sampling algorithms. The case study exemplifies the benefits of a modelling approach in terms of providing a perspective on the underlying single year of age migrant propensities even though the observations are for five year migration totals.

The analysis has showed the benefits of random effects approaches both in estimating single age schedules and in joint estimation of several schedules - for example, in terms of overall fit and in terms of model predictions replicating the actual data. Models that introduce random effects into a fully parametric (multi-exponential) model may be competitive in fit terms with purely nonparametric approaches. However, the latter may be useful for a preliminary smoothing analysis to assess which features (e.g. retirement humps or not) actually seem to be present in the data. The case study has shown that the nonparametric regression approach detects features in the migration data that may not be detected by the multi-exponential model.

7 References

- Besag, J, Green, P, Higdon, D, Mengersen, K (1995) Bayesian computation and stochastic systems, *Statistical Science*, 10, 3-66
- Gelfand, A (1996) Model determination using sampling based methods, Chapter 9 in *Markov Chain Monte Carlo in Practice*, eds W. Gilks, S. Richardson, and D. Spiegelhalter, Chapman and Hall/CRC, Boca Raton
- Gelman, A, Carlin, J, Stern, H, Rubin, D (1995). *Bayesian Data Analysis*. London: Chapman and Hall.
- Gordon, I (1982) Return migration to Scotland: some direct evidence. *Reg Stud.* 16, 129-36.
- Neves, C, Migon, H (2004) Graduação Bayesiana de taxas de mortalidade: uma aplicação na cobertura de sobrevivência e na avaliação da provisão matemática. *Revista Brasileira de Risco e Seguro*, 1

Rogers, A, Castro, L. (1981) Model Migration Schedules, Research Report. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rogers, A, Raymer, J (1999) Fitting observed demographic rates with the multi-exponential model schedule: an assessment of two estimation programs. Review of Urban and Regional Development Studies, 11, 1-10.

Rogers, A, Castro, L, Lea, M (2004) Model Migration Schedules: Three Alternative Linear Parameter Estimation Methods. Institute Of Behavioral Science, University of Colorado. Working Paper POP2004-0004

Spiegelhalter, D, Best, N, Carlin, B, van der Linde, A (2002) Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society, Series B, 64, 583-639

Stone, C (1985) Additive regression and other nonparametric models, Annals of Statistics, 13, 689-705.

Appendix: WINBUGS code

Example codes are provided for the for the gamma-mixture univariate model and for the bivariate RW model. Data inputs for the first model are eighteen five year age group migration flows $Y[]$, and 91 single year of age mid year populations $N[]$; for the second model the data are male and female flows $Y1[]$, $Y2[]$ and single year of age populations $N1[]$ and $N2[]$.

Univariate Gamma Mixture

```
model { for ( i in 1:18) {Y[i] ~dpois(m.Y[i]); Ynew[i] ~dpois(m.Y[i])}
  for (j in 1:91) {r[j] ~dgamma(kappa,r1[j]); r1[j] <- kappa/m[j] }
  for (i in 1:17) { m.Y[i] <- sum(q[5*i-4:5*i])}
  m.Y[18] <- sum(q[86:91])
  for (i in 1:91) {m[i] <- C1[i]+C2[i]+a0; q[i] <- r[i]*N[i]
  x[i] <- i-0.5; C1[i] <- a[1]*exp(-alph[1]*x[i])
  C2[i] <- a[2]*exp(-alph[2]*d[i]-exp(shift[i]))
  d[i] <- x[i]-mu2; shift[i] <- exp(-lam2*d[i])}
  # priors
  a[1] ~dgamma(0.05,1); kappa ~dgamma(1,0.01)
  a[2] ~dgamma(0.075,1); mu2 ~dgamma(20,1)
  lam2 ~dgamma(0.425,1); alph[1] ~dgamma(0.05,1)
  alph[2] ~dgamma(0.15,1); a0 ~dgamma(0.001,1) I(0,0.005)}
```

Bivariate RW model

```
model { for ( i in 1:18) {# model for male & female five year age flow data
  Y1[i] ~dpois(m1[i]); Y2[i] ~dpois(m2[i])
  # sample replicate data
  Ynew[1,i] ~dpois(m1[i]); Ynew[2,i] ~dpois(m2[i])}
  # gross migraproduction rates and inequality test
```

```

G[1] <- sum(r[1,]); G[2] <- sum(r[2,]); testG <- step(G[1]-G[2])
# total over 1 year age rates to get five year rates
for (i in 1:17) { m1[i] <- sum(q1[5*i-4:5*i]); m2[i] <- sum(q2[5*i-4:5*i])}
                m1[18] <- sum(q1[86:91]); m2[18] <- sum(q2[86:91])
#Rates
for (i in 1:91) {log(r[1,i]) <- beta[1]+a[1,i]; log(r[2,i]) <- beta[2]+a[2,i]
                q1[i] <- r[1,i]*N1[i]; q2[i] <- r[2,i]*N2[i]}
# prior on precision matrix
omega[1:2,1:2] ~ dwish(Q[,],2)
for (j in 1:2) {for (k in 1:2) {Q[j,k] <- equals(j,k)*0.1}}
# covariance matrix
Sig[1:2,1:2] <- inverse(omega[,]); rho <- Sig[1,2]/sqrt(Sig[1,1]*Sig[2,2])
for (j in 1:2) {beta[j] ~ dnorm(0,0.001)}
# RW1 bivariate prior
a[1:2, 1 : Nage] ~ mv.car(adj[], w[], num[], omega[ , ])
# Weight & adjacency matrix for RW(1) prior
adj[1] <- 2; w[1] <- 1; num[1] <- 1; w[(Nage-1)*2] <- 1;
adj[(Nage-1)*2] <- Nage-1; num[Nage] <- 1
for (j in 2:(Nage-1)) { w[2+(j-2)*2] <- 1; w[3+(j-2)*2] <- 1;
adj[2+(j-2)*2] <- j-1; adj[3+(j-2)*2] <- j+1; num[j] <- 2}}

```