Forecasting age and sex (and regional) patterns of immigration and emigration

Arkadiusz Wiśniowski
a.wisniowski@manchester.ac.uk

MANCHESTER 1824
The University of Manchester

Social Statistics

Eurostat-UNECE Work Session on Demographic Projections
Belgrade, 25-27 November 2019

Introduction

International migration

- Most difficult demographic component to forecast
- Most influential component of population change

Introduction

International migration

- Most difficult demographic component to forecast
- Most influential component of population change
- Needed for accurate and meaningful population projections

Why is migration difficult to forecast?

- Loosely defined concept / poorly measured
  - Unlike fertility and mortality, no clear observable event
  - Data not always available / consistent over time
- Involves two populations
  - Populations at risk are different for immigrants and emigrants
- Influenced by multiple factors
- Combination of motives

Motivation

- Raymer & Wiśniowski (2018) Applying and testing a forecasting model for age and sex patterns of immigration and emigration. *Population Studies* 72 (3), 339-355.
- Migration exhibits regularities in age profiles
- Can the Lee & Carter model (and Bayesian inference) be applied to forecast migration?

Forecasting migration

- Practices
  - Most countries combine recent trends and argument-based scenarios
  - Net migration

Forecasting migration

- Practices
  - Most countries combine recent trends and argument-based scenarios
  - Net migration
- Regression and time series models
  - De Beer (2008) and Bijak et al. (2019) – different time patterns for various types of migration
  - Cappelen et al. (2015) use econometric model with covariates

Forecasting migration

- Practices
  - Most countries combine recent trends and argument-based scenarios
  - Net migration
- Regression and time series models
  - De Beer (2008) and Bijak et al. (2019) – different time patterns for various types of migration
  - Cappelen et al. (2015) use econometric model with covariates
- Forecasting age patterns of migration
  - Multiexponential model schedule (Rogers & Castro 1981)
  - Multiplicative component (log-linear) model
  - Semi-parametric methods

Probabilistic forecasting of migration

- Reflect the inherent complexity underlying migration decisions, outcomes and data

Probabilistic forecasting of migration

- Reflect the inherent complexity underlying migration decisions, outcomes and data
- Recent developments
  - Hyndman & Booth 2008 (net mig. by age, functional)
  - Bijak & Wiśniowski (2010) (total counts, expert opinion, Bayesian)
  - Bijaks 2010 (total counts and rates, Bayesian)
  - Azose & Raftery 2015 (net mig. totals, Bayesian)
  - Wiśniowski et al. 2015 (immig. / emig. by age, Bayesian)
  - Shang et al. 2016 (immig. / emig. by age, functional)

Lee-Carter model and extensions

- Bilinear model (forecasting mortality: Lee & Carter 1992)
- Has been adopted to fit various settings, including fertility and migration (e.g. Wiśniowski et al. 2015)

Lee-Carter model and extensions

- Bilinear model (forecasting mortality: Lee & Carter 1992)
- Has been adopted to fit various settings, including fertility and migration (e.g. Wiśniowski et al. 2015)
- Decomposes age-time schedule (profile) into average age schedule over time, age-specific changes in schedule in response to changes in time effect (subject to constraints on the model parameters):

$$average\ age\ schedule\ + changes\ in\ age\ schedule \times time\ effect$$

- Time-series models are used to model and forecast time effect

Extending the Lee-Carter model

- Emigration or immigration counts are denoted by $Y_{x,t}$ for single age group x and in year t:

$$Y_{x,t} \sim \text{Poisson}\left(\mu_{x,t}\right) \tag{1}$$

$$\log \mu_{x,t} \sim \text{Normal}\left(\alpha_x + \beta_x \kappa_t, \sigma^2\right), \tag{2}$$

- where a_x and b_x are age-specific model parameters, k_t is time specific parameter and σ^2 is a variance of the expected log-counts
- To ensure identifiability, we constrain all b_x to sum to one for all years of age and the first observation of k_t to be zero

Extending the Lee-Carter model

- The time-specific κ_t captures the pace of these changes for a time period:

$$\kappa_t \sim \mathrm{Normal}\left(\phi_0 + \phi_1 \kappa_{t-1}, \sigma_\kappa^2\right), \tag{3}$$

Extending the Lee-Carter model

- The time-specific κ_t captures the pace of these changes for a time period:

$$\kappa_t \sim \text{Normal}\left(\phi_0 + \phi_1 \kappa_{t-1}, \sigma_\kappa^2\right), \quad (3)$$

- Three main specifications of the time-specific parameter κ_t:
  - random walk: RW; $\phi_0 = 0$ and $\phi_1 = 1$
  - random walk with drift: RW; $\phi_1 = 1$
  - autoregressive of order one: AR(1) with both parameters being estimated

Extending the Lee-Carter model

- The time-specific κ_t captures the pace of these changes for a time period:

$$\kappa_t \sim \text{Normal}\left(\phi_0 + \phi_1 \kappa_{t-1}, \sigma_\kappa^2\right), \quad (3)$$

- Three main specifications of the time-specific parameter κ_t:
  - random walk: RW; $\phi_0 = 0$ and $\phi_1 = 1$
  - random walk with drift: RW; $\phi_1 = 1$
  - autoregressive of order one: AR(1) with both parameters being estimated
  - ‘multivariate mix’ which combines the univariate models that perform best and correlation between males and females

Extending the Lee-Carter model

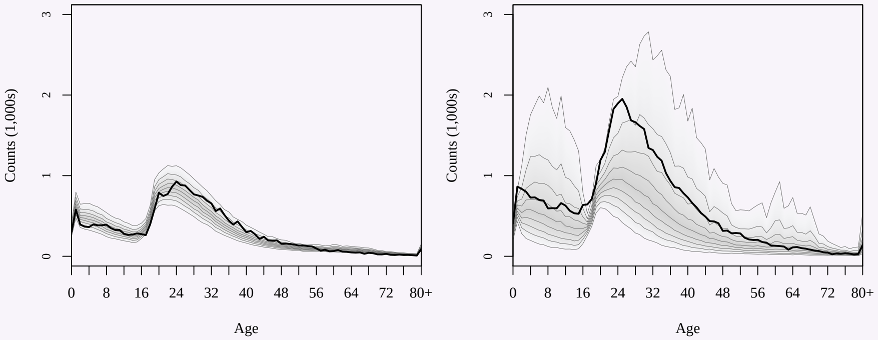
- The time-specific κ_t captures the pace of these changes for a time period:

$$\kappa_t \sim \mathrm{Normal}\left(\phi_0 + \phi_1 \kappa_{t-1}, \sigma_\kappa^2\right), \tag{3}$$

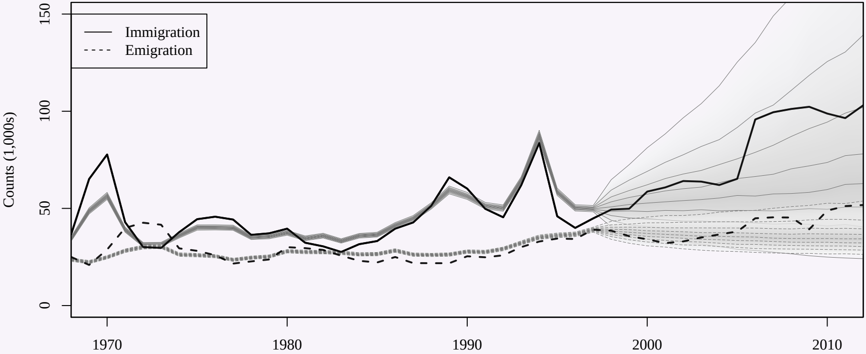
- Three main specifications of the time-specific parameter κ_t:
  - random walk: RW; $\phi_0 = 0$ and $\phi_1 = 1$
  - random walk with drift: RW; $\phi_1 = 1$
  - autoregressive of order one: AR(1) with both parameters being estimated
  - 'multivariate mix' which combines the univariate models that perform best and correlation between males and females

- Divided samples into 2/3 training and 1/3 test and tested forecasting performance

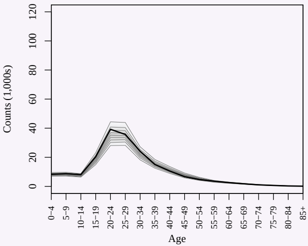
Multivariate model (1998 and 2012):

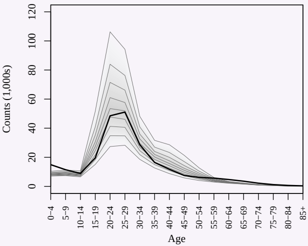


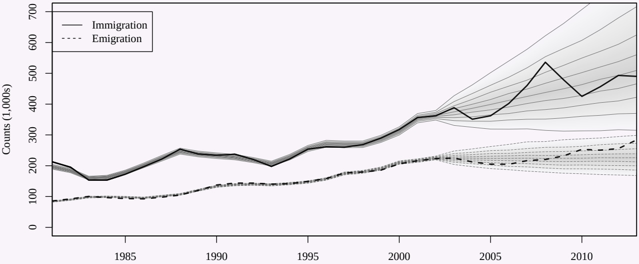
Multivariate model:



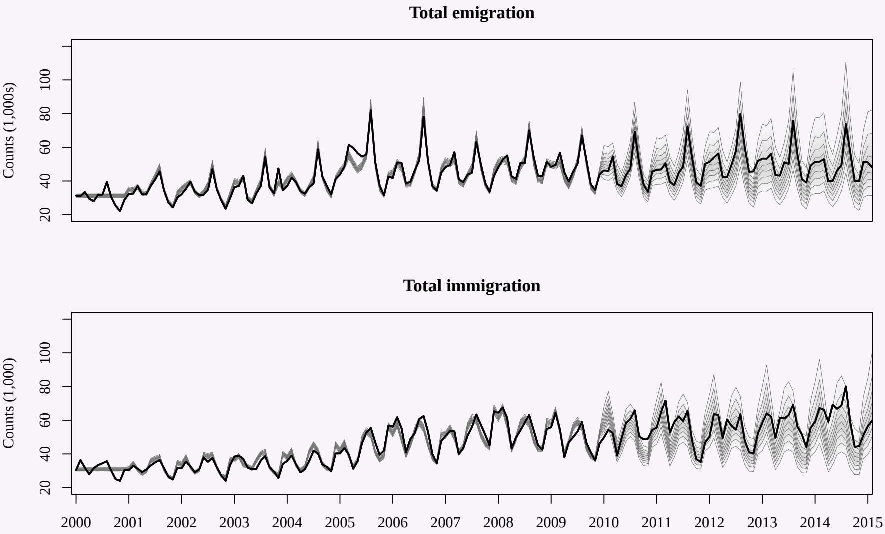
Multivariate model (2003 and 2013):







Korea



Total emigration
Counts (1,000s)
20
40
60
80
100
Total immigration
Counts (1,000)
20
40
60
80
100
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015

Extending the bilinear model

- Testing multivariate models (correlations)

$$\boldsymbol{y}_t \sim \mathcal{MVN}(\boldsymbol{\phi}_0 + \boldsymbol{\Phi}_1 \boldsymbol{y}_{t-1}, \Sigma)$$
$$\boldsymbol{y}_t = (y_t^{FE}, y_t^{ME}, y_t^{FI}, y_t^{MI})$$

Extending the bilinear model

- Testing multivariate models (correlations)

$$\boldsymbol{y}_t \sim \mathcal{MVN}(\boldsymbol{\phi}_0 + \boldsymbol{\Phi}_1 \boldsymbol{y}_{t-1}, \Sigma)$$
$$\boldsymbol{y}_t = (y_t^{FE}, y_t^{ME}, y_t^{FI}, y_t^{MI})$$

- Different trajectories for parts of the age profile (<20, 20-59, $60+$)

$$\log \mu_{xt} \sim \text{Normal}\left(\alpha_x + \beta_x^1 \kappa_t^1 \mathbb{1}_{x \leq A_1} + \beta_x^2 \kappa_t^2 \mathbb{1}_{A_1 < x \leq A_2} + \ldots, \sigma^2\right)$$

Extending the bilinear model

- Testing multivariate models (correlations)

$$\boldsymbol{y}_t \sim \mathcal{MVN}(\boldsymbol{\phi}_0 + \boldsymbol{\Phi}_1 \boldsymbol{y}_{t-1}, \Sigma)$$
$$\boldsymbol{y}_t = (y_t^{FE}, y_t^{ME}, y_t^{FI}, y_t^{MI})$$

- Different trajectories for parts of the age profile (<20, 20-59, $60+$)

$$\log \mu_{xt} \sim \text{Normal}\left(\alpha_x + \beta_x^1 \kappa_t^1 \mathbb{1}_{x \leq A_1} + \beta_x^2 \kappa_t^2 \mathbb{1}_{A_1 < x \leq A_2} + \ldots, \sigma^2\right)$$

- Extending the model to include regions (log-linear)

$$Y_{x,t,r} \sim \text{Poisson}\left(\mu_{xtr}\right)$$
$$\log \mu_{xtr} \sim \text{Normal}\left(\theta^{RA} + \theta^{RS} + \alpha_1^{AS} + \beta_2^{AS} \kappa_t^{AS} + \alpha_1^{R} + \beta_2^{R} \kappa_t^{R}, \sigma^2\right)$$

Extending the bilinear model

- Testing multivariate models (correlations)

$$\boldsymbol{y}_t \sim \mathcal{MVN}(\boldsymbol{\phi}_0 + \boldsymbol{\Phi}_1 \boldsymbol{y}_{t-1}, \Sigma)$$
$$\boldsymbol{y}_t = (y_t^{FE}, y_t^{ME}, y_t^{FI}, y_t^{MI})$$

- Different trajectories for parts of the age profile (<20, 20-59, $60+$)

$$\log \mu_{xt} \sim \text{Normal}\left(\alpha_x + \beta_x^1 \kappa_t^1 \mathbb{1}_{x \leq A_1} + \beta_x^2 \kappa_t^2 \mathbb{1}_{A_1 < x \leq A_2} + \ldots, \sigma^2\right)$$

- Extending the model to include regions (log-linear)

$$Y_{x,t,r} \sim \text{Poisson}\left(\mu_{xtr}\right)$$
$$\log \mu_{xtr} \sim \text{Normal}\left(\theta^{RA} + \theta^{RS} + \alpha_1^{AS} + \beta_2^{AS} \kappa_t^{AS} + \alpha_1^{R} + \beta_2^{R} \kappa_t^{R}, \sigma^2\right)$$

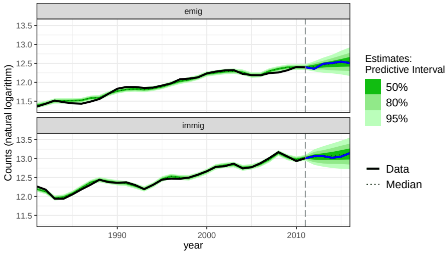
- Tested for Australia 1981-2011, forecasts for 2012-2016

Australia

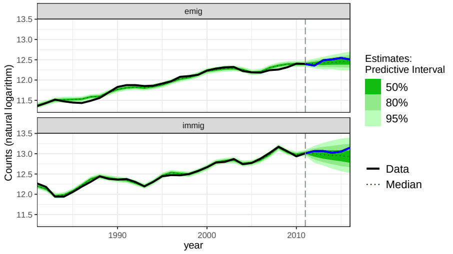
	EF	EM	IF	IM	T	EF	EM	IF	IM	T
	VAR unconstrained					VAR diagonal				
MPE	2	2	8	4	4	9	3	16	12	10
MAE	**763**	**996**	1190	1218	1042	981	1033	2025	1822	1465
RMSE	1049	**1371**	1691	**1777**	**1500**	1599	1433	3687	3060	2625
MAPE	16	17	16	18	17	16	17	18	19	18
cov50	0.48	0.36	0.43	0.41	0.42	0.39	0.28	0.39	0.40	0.36
cov90	0.84	0.82	0.77	0.78	0.80	0.76	0.78	0.70	0.76	0.75
	VRW					VRW different trajectories				
MPE	1	-3	6	1	1	**1**	**-1**	**5**	**1**	**1**
MAE	767	1008	1151	1527	1113	847	1103	**1151**	1282	1096
RMSE	**1018**	1375	1608	2150	1592	1282	1697	**1565**	1868	1618
MAPE	16	17	16	19	17	**14**	**15**	**15**	**16**	**15**
cov50	0.40	0.37	0.38	0.40	0.39	0.39	0.34	**0.46**	**0.53**	**0.43**
cov90	0.81	0.77	0.77	0.79	0.78	**0.89**	**0.84**	**0.91**	**0.87**	**0.88**

Australia

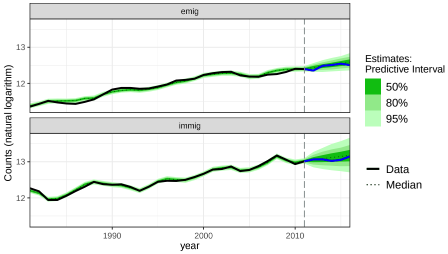
VAR(1) unconstrained



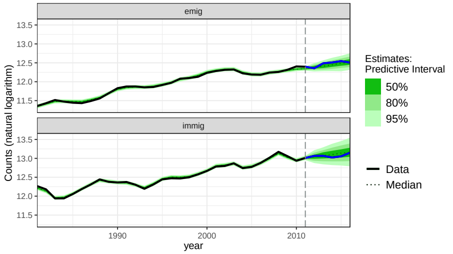
VAR(1) only diagonal parameters



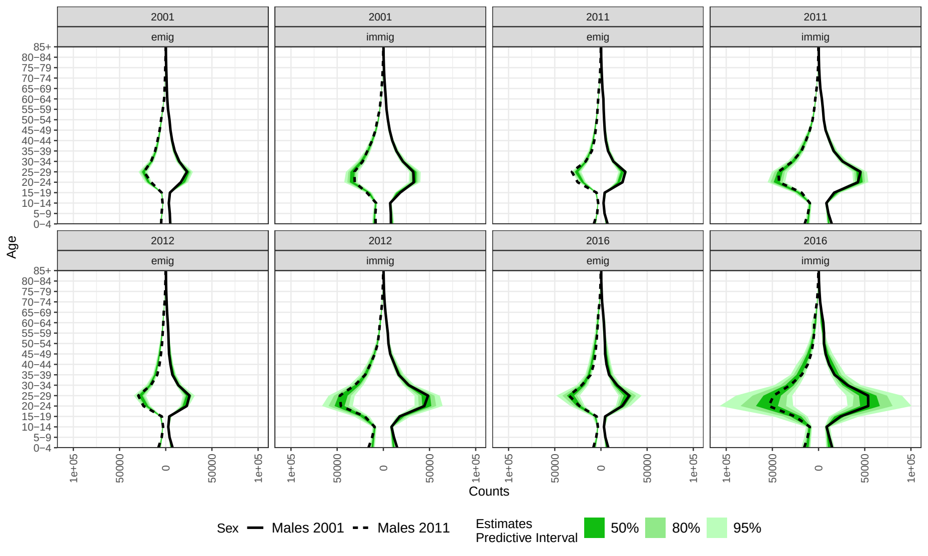
VRW(1)



VRW(1) different trajectories

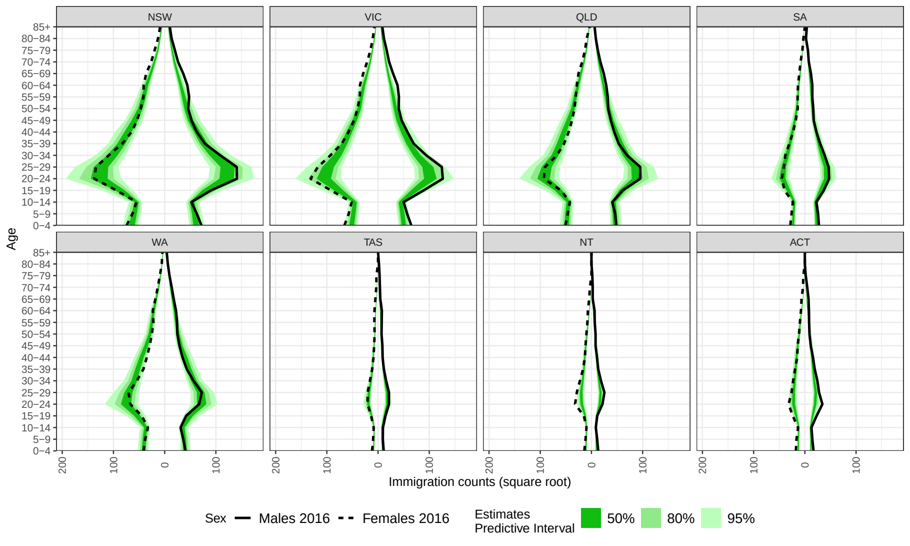


Age profiles: VRW different trajectories



Forecasts by regions - totals

Forecasts by regions - age profiles 2016



Future work

- Limitations
  - Sensitivity to sudden changes in the age profiles
  - Each time series requires a different model
  - Shifts in the migration patterns due to different policy regimes
- Other models
  - Smooth functions (Rogers & Castro, splines, mixture distributions)
  - Compositional models (regions)
- Expert knowledge (prior distributions)

Conclusions

- Bilinear models can capture and forecast age-time patterns of migration

Conclusions

- Bilinear models can capture and forecast age-time patterns of migration
- Realistic uncertainty of migration flows

Conclusions

- Bilinear models can capture and forecast age-time patterns of migration
- Realistic uncertainty of migration flows
- Different time series models required for different flows (Bijak et al. 2019)

Conclusions

- Bilinear models can capture and forecast age-time patterns of migration
- Realistic uncertainty of migration flows
- Different time series models required for different flows (Bijak et al. 2019)

- Future migration patterns are highly uncertain but we need them to produce realistic population forecasts

Acknowledgements

Email: a.wisniowski@manchester.ac.uk

Part of this work has been carried out in the ESRC Centre for Population Change, University of Southampton. Financial support from the National Centre for Research Methods International Visitor Exchange Scheme as well as University of Manchester School of Social Sciences Small Grant is acknowledged. Thanks go to Jakub Bijak, Jon Forster, James Raymer and Peter Smith for their contributions and suggestions.

Bayesian inference

- Weakly informative priors

$$\alpha_x \sim \text{Normal}\left(0, \sigma_\alpha\right), \qquad \sigma_\alpha \sim \text{Cauchy}_+\left(0, 5\right)$$

$$\beta_{1:N-1} \sim \mathcal{MVN}_{N-1}\left(\frac{1}{N}, \frac{1}{N}\Psi\right), \qquad \beta_N = 1 - \sum_{x=1}^{N-1} \beta_x,$$

$$\Psi = J_{N-1} + I_{N-1}$$

$$\sigma \sim \text{Cauchy}_+\left(0, 1\right)$$

$$\phi_0 \sim \text{Normal}\left(0, 2\right)$$

$$\Sigma = [\text{diag}(\sigma_{\kappa c}]^{-1/2} R\, [\text{diag}(\sigma_{\kappa c}]^{-1/2}$$

$$R \sim \text{LKJ}(2), \qquad \sigma_\kappa \sim \text{Normal}(0, 1)$$

- Implemented using Hamiltonian Monte Carlo (HMC) and No-U-Turn Sampler (NUTS; Hoffman and Gelman 2014) in RStan