Bayesian Hierarchical Model for Estimating Age and Sex Patterns of International Migration

Aysha Basheer[a], Arkadiusz Wiśniowski[a], Maciej J. Dańko[b]

[a] *Department of Social Statistics, University of Manchester, Manchester, UK*
[b] *Max Planck Institute for Demographic Research (MPIDR), Rostock, Germany*

Abstract

This study presents a Bayesian hierarchical model to analyze the age and sex patterns of international migration across Europe. The model incorporates age, sex, origin, destination, and time to address the challenges of inconsistent and incomplete migration data. Data from 31 European countries between 2008 and 2022 were used, focusing on 18 five-year age groups (0 to 85+) for both sexes. We implemented three variations of the log-linear model, with the most complex version, including interaction and time terms, showing the best fit according to diagnostic measures. Preliminary results highlight higher migration proportions among males, particularly in the 20-34 age group, with an increase in migration observed between 2009 and 2015. These estimates were compared to reported values and will be further evaluated against the Rogers-Castro migration schedule. Furthermore, we plan to extend the model to add covariates and incorporate splines for smoothing flows. This study provides valuable insights into the age and sex patterns of European migration flows, offering a robust approach to harmonizing inconsistent migration data and potentially improving the accuracy of migration statistics, which are crucial for policy-making and social service planning.

Keywords: Bayesian hierarchical model, International migration, Age-sex demographics

Introduction

Human migration is a complex phenomenon that significantly influences the economic, social, and political dynamics of society. Accurate information regarding the number of migrants entering and leaving a country, along with their demographic characteristics, is crucial for effective planning of future economic development and the expansion of social services. However, the lack of timely and comprehensive data about migrants, combined with the diverse measures and definitions of migration used by various countries, hinders the accurate estimation of international migration. This study addresses the challenges of inconsistent and incomplete migration data by using a Bayesian hierarchical approach with a log-linear model incorporating age, sex, origin, and destination data to provide a comprehensive picture of migration patterns. The model combines age and sex profiles from both sending and receiving countries to produce harmonized estimates of detailed migration proportions with measures of uncertainty.

Methodology

We considered data from 31 European countries, including the European Union (EU) and European Free Trade Association (EFTA) countries, from 2008 to 2022. The migration flows by age (a), sex (s), origin (o), destination (d), and time (t) were estimated using a hierarchical Bayesian framework. The data were organized as a contingency table of migration counts, denoted as z^k_{odast}, representing migration flows from origin o to destination d by sex s and age group a during a specific time t, where k can indicate either the sending country (S) or the receiving country (R). These flows are assumed to follow a Poisson distribution with a mean λ^k_{odast},

$$z^k_{odast} \sim \text{Poisson}\left(\lambda^k_{odast}\right),$$

which is related to the true migration flow m_{odast} through the log-linear model,

$$\log\left(\lambda^k_{odast}\right) \sim N\left\{\log\left(m_{\text{odast}}\right) + \alpha^k_{\text{odt}}, \tau_k\right\}$$

where α^k_{odt} is the interaction parameter and τ_R denotes precision. The true flow m is modeled as a log-normal distribution. We implemented three variations of this model to capture different effects.

Main Effects Model:

$$\log(m_{odast}) \sim N(\alpha_d + \alpha_o + \alpha_a + \alpha_s, \tau_m) \tag{1}$$

Interaction Effects Model:

$$\log(m_{odast}) \sim N(\alpha_{da} + \alpha_{ds} + \alpha_{oa} + \alpha_{os}, \tau_m) \tag{2}$$

Interaction Effects with Time Model:

$$\log(m_{odast}) \sim N(\alpha_{da} + \alpha_{ds} + \alpha_{oa} + \alpha_{os} + \alpha_t, \tau_m) \tag{3}$$

where α_o, α_d, α_a and α_s are parameters capturing the origin, destination, age and sex patterns across all countries, which are then adjusted for a given origin and destination by the interaction parameters. Age and sex proportions of a given origin and destination in year t are then calculated as,

$$\pi_{odast} = \frac{m_{odast}}{m_{od++t}}$$

where m_{od++t} is the summation over age and sex.

Prior distributions are combined with the likelihood of the data to estimate the parameters, with posterior distributions summarized using median estimates and credible intervals (CIs) to communicate the uncertainty around the estimates.

Preliminary Results and Discussions

We conducted an initial analysis on a subset of the data, focusing on four countries—Austria (AT), Finland (FI), Netherlands (NL), and Sweden (SE)—across 18 five-year age groups from 0 to 85+, for both sexes (Female and Male) over the period 2009-2015. Model comparison based on the Deviance Information Criterion (DIC), Bayesian Information Criterion (BIC), and Akaike Information Criterion (AIC) (Table 1) revealed that the model with interaction parameters and time term provided the best fit. This model not only had the lowest DIC, BIC, and AIC values but also closely aligned with observed migration data for both immigration and emigration across different years and age groups (Figure 1).

Model	DIC	BIC	AIC
Interaction Effects Model	17576.46	33483.39	21649.83
Interaction Effects with Time Model	17479.01	33549.25	21350.41
Main Effects Model	18216.25	38626.33	23133.14

Table 1: Model diagnostics

Age and Sex Patterns

A comprehensive set of plots illustrating the age and sex patterns can be accessed here. These plots show a consistent pattern of the age distribution of migrants across all countries and years, with a pronounced peak in the 20-34 age range. This peak is particularly evident for both males and females, suggesting that young adults are the most mobile demographic. The proportion of migrants decreases steadily with age after 30, with a slight increase observed in the 60-70 age range for some countries, possibly indicating retirement migration. Figure 2 show that the migration proportions for different age groups remain relatively stable over the observed years for most countries. However, there are some noticeable fluctuations, particularly in the peak age groups (20-34). Both females and males exhibit higher migration proportions during these years, which could be attributed to career opportunities and other socio-economic factors. Finland shows a slight increase in migration proportions for this age group over time, while Sweden exhibits more variability year-to-year.

While the overall age patterns are similar for both sexes, there are slight differences in migration proportions between males and females (Figure 3). In general, males show slightly higher migration proportions across most age groups and countries. However, the magnitude of this difference varies by country and year. For instance, female migrants are more moving from Finland to other countries, while males dominate migration from all other countries. This is mainly related to education and labour markets [1].

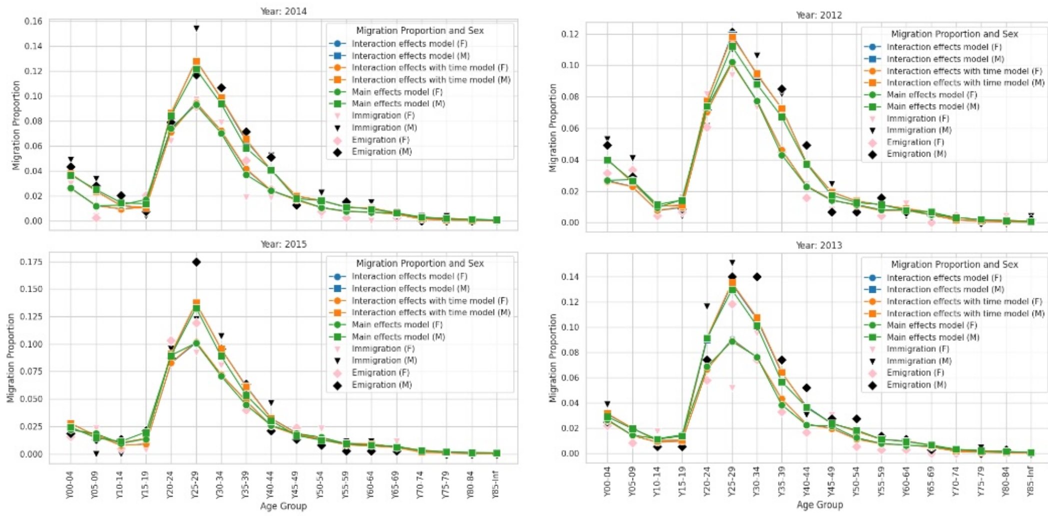
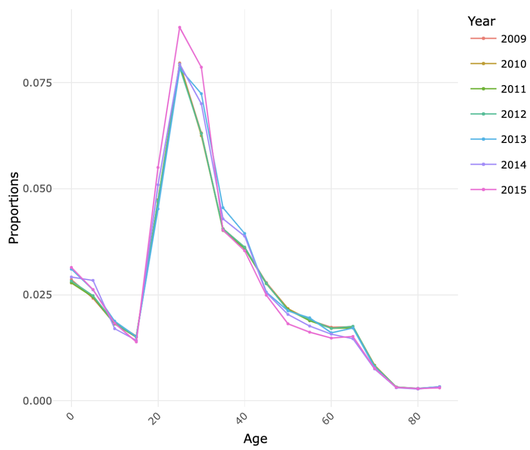
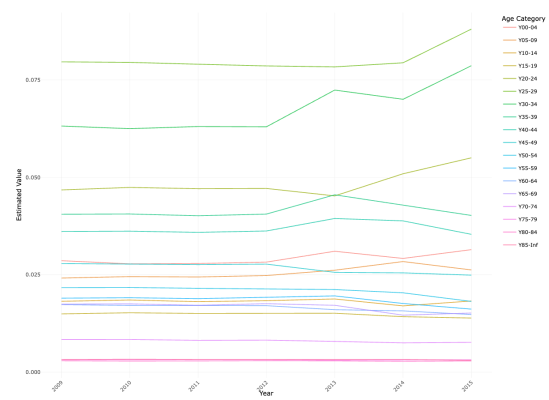


Figure 1: Comparison of model estimates with observed emigration (reported by the Netherlands) and immigration (reported by Finland) for migration flows between the two countries



(a) Age Profile



(b) Age Specific Proportions over Time

Figure 2: Estimated proportions of female migration from Sweden to Austria

Future Work

These preliminary results, based on data from four countries, reveal migration patterns. However, more substantial differences and trends are likely to emerge once the model is applied to the complete dataset covering all 31 European countries. Our model addresses data quality issues and harmonizes migration flow data, offering a comprehensive analysis that will provide deeper insights into the interactions among age, sex, origin, destination, and time.

Future work will involve modifying the model by incorporating splines to smooth the flows and evaluating these models against the Rogers-Castro migration schedule. Additionally, we plan to include parameter-specific covariates, such as life expectancy, education levels, and employment rates, to capture the influence of external factors on migration patterns. This will allow us to account for heterogeneity across demographic groups and improve model fit and accuracy. This expanded scope will improve migration estimates and provide valuable insights for policy-making, assisting in addressing migration challenges and improving planning for economic development and social services.

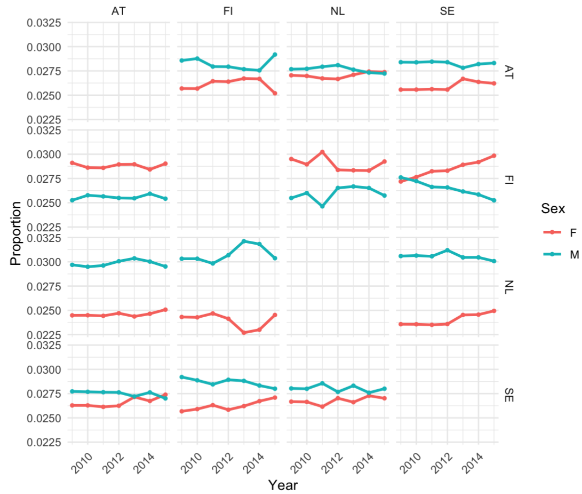


Figure 3: Estimated migration sex patterns, with rows and columns representing origin and destination countries respectively

References

[1] MDI, Women's migration focuses on big cities, and it is not a surprise, MDI Insights (2024). URL: https://www.mdi.fi/en/article-womens-migration-focuses-on-big-cities-and-it-is-not-a-surprise/.

[2] G. J. Abel, Estimation of International Migration Flow Tables in Europe, Journal of the Royal Statistical Society Series A: Statistics in Society 173 (2010) 797–825. URL: https://doi.org/10.1111/j.1467-985X.2009.00636.x. doi:10.1111/j.1467-985X.2009.00636.x. arXiv:https://academic.oup.com/jrsssa/article-pdf/173/4/797/49525384/jrsssa_173_4_797.pdf.

[3] J. Raymer, A. Wiśniowski, J. Forster, P. Smith, J. Bijak, Integrated modeling of european migration, Journal of the American Statistical Association 108 (2013) 801–819. doi:10.1080/01621459.2013.789435.

[4] A. Wiśniowski, J. Bijak, S. Christiansen, J. J. Forster, N. Keilman, J. Raymer, P. W. Smith, Utilising expert opinion to improve the measurement of international migration in europe, Journal of Official Statistics 29 (2013) 583–607. URL: https://doi.org/10.2478/jos-2013-0041. doi:10.2478/jos-2013-0041. arXiv:https://doi.org/10.2478/jos-2013-0041.

[5] T. Shen, J. Raymer, Q. Guan, A. Wiśniowski, The estimation of age and sex profiles for international migration amongst countries in the asia-pacific region, Population, Space and Place 30 (2024) e2716. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/psp.2716. doi:https://doi.org/10.1002/psp.2716. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/psp.2716.

[6] A. Rogers, L. J. Castro, Model Migration Schedules, IIASA Research Report, IIASA, Laxenburg, Austria, 1981. URL: https://pure.iiasa.ac.at/id/eprint/1543/.

[7] A. Wiśniowski, J. J. Forster, P. W. F. Smith, J. Bijak, J. Raymer, Integrated modelling of age and sex patterns of european migration, Journal of the Royal Statistical Society. Series A (Statistics in Society) 179 (2016) 1007–1024. URL: http://www.jstor.org/stable/44682194.

[8] J. Yeung, T. Riffe, M. Alexander, Bayesian implementation of Rogers–Castro model migration schedules: An alternative technique for parameter estimation, Demographic Research 49 (2023) 1201–1228. URL: https://www.demographic-research.org/volumes/vol49/42/. doi:10.4054/DemRes.2023.49.42. arXiv:https://www.demographic-research.org/volumes/vol49/42/49-42.pdf.