



## Editorial: Modelling migration flows

Migration is a loosely defined process that represents the relocation of people during a period of time that causes them to relinquish the ties with their previous locality. The key factors that separate migration from general mobility are distance travelled and length of time spent in the destination; together they work to alter the economic and social networks of the migrant. Migration can involve people moving within a country, as well as across international borders.

To obtain estimates, and to understand the determinants of national or international population movements, researchers need reliable and comprehensive data, and appropriate modelling techniques. However, analysts are often faced with the situation of inadequate or missing data, as well as various levels of uncertainty in the data, models used and future trends. Furthermore, the available data are most often collected for general purposes and do not necessarily coincide with theoretical or contextual definitions of migration (Bell *et al.*, 2002). The five papers in this themed issue of *Statistics in Society* address various aspects of these problems. These papers were originally presented in September 2008, either at the contributed session on 'Migration' at the Royal Statistical Society conference in Nottingham, or at the 'Quantitative methods in social sciences' workshop on 'Estimation and projection of international migration' which was hosted by the Southampton Statistical Sciences Research Institute and funded by the European Science Foundation.

Although all papers in this issue are concerned with the modelling of migration flows, there are several themes that deserve mention here. These are the types of migrants, methods for combining and harmonizing data, modelling approaches and incorporating expert opinion and uncertainty.

The first theme concerns the type of migrants who are being modelled. In recent decades, migration has gained importance in driving both subnational and national population change. In developed countries, internal migration is driven by people seeking employment, or a better physical or social environment to live in, depending on what stage in the life course the migrant is in. The drivers of international migration, however, differ by direction (de Beer, 2008). Immigration is often motivated by capitalist economies requiring supplies of (cheap) labour, the need to fill specific (high skilled) labour gaps, people seeking to reconnect their families, nationals returning home after spending a period of time working abroad and people seeking asylum. Emigration is driven by (high skilled) nationals seeking better employment opportunities abroad and by foreigners returning home.

It matters whether one is modelling internal migration, where the data may come from several national level sources (e.g. censuses, surveys or registers), or international migration, where the data are specific to the sending or receiving country. Here, the measurement of emigration and inconsistencies in measurement across countries are particularly problematic (Bilsborrow *et al.*, 1997; Poulain *et al.*, 2006). The papers by Peter W. F. Smith, James Raymer and Corrado Giulietti, and Peter Congdon focus on internal movements in England, whereas the papers by Peter Boden and Phil Rees, Jakub Bijak and Arkadiusz Wiśniowski, and Guy J. Abel focus on international movements in Europe.

The second theme concerns methods for combining and harmonizing available data to improve the evidence base. Boden and Rees argue that better local authority estimates of

immigration in England can be obtained by integrating administrative data, such as national insurance number registrations by workers from abroad or the registration of foreign students for higher education, with national level estimates obtained from the International Passenger Survey. Tracking local levels of immigration is very important for countries receiving large numbers of immigrants. However, some countries, such as the UK, do not have population registers or other means to measure the changes over time directly. Smith, Raymer and Giuliatti apply iterative proportional fitting to combine registration, census and survey data to estimate a series of intercounty migration flows by age, sex and economic activity. This allows them to overcome the strict reliance on census data to understand how internal migration patterns are evolving over time. Finally, Abel estimates a complete and harmonized set of international migration flows between 15 European Union countries by combining information provided by the countries separately and covariate information. Both Smith, Raymer and Giuliatti, and Abel use a statistical model, whereas Boden and Rees use a data accounting framework.

The third theme concerns the modelling approach itself. Congdon, Bijak and Wiśniowski, and Abel explicitly fit regression-type statistical models, which include random effects, lagged values of the response variable and explanatory variables motivated by migration theory, whereas Boden and Rees, and Smith, Raymer and Giuliatti rely on previously observed structures to guide the (implicit) modelling process. When including covariate information, a key distinction is made between population level variables that describe the overall levels and drivers of migration and variables that capture the interactions between places.

Congdon, for example, applies a Bayesian hierarchical random-effects approach to extend the well-known gravity model of migration to study migration attractivity and retentivity between 354 English local authorities during the 2000–2001 period. Random effects are used to capture the correlations between and within areas. He argues that the Bayesian approach is more flexible in general, and is less complicated for random-effects modelling in particular. It also allows one to assess the distributional properties of the parameters of interest (e.g. migration attractivity). In this case, he assumes a negative binomial model to account for both the high proportion of zero flows and overdispersion in the data. Bijak and Wiśniowski also rely on the Bayesian model to combine standard time series models with expert judgements to forecast immigration flows.

The final theme that is covered in the papers in this issue is expert opinion and uncertainty. To understand reported migration data better, both expert opinion and measures of uncertainty are important. Bijak and Wiśniowski incorporate opinions obtained from country-specific experts in their forecasts of immigration for seven European countries (i.e. Austria, the Czech Republic, France, Hungary, Italy, Poland and Portugal) with the aim to augment the relatively short time series of available data. To do so, they rely on a two-round Delphi survey, allowing the experts to revise their judgements after seeing the summarized results of the other experts. As for measuring uncertainty, two approaches are utilized in this issue. Abel derives uncertainty for the estimates of missing data by using a supplemented EM-type algorithm. Congdon, and Bijak and Wiśniowski, in contrast, rely on Bayesian methods, which combine expert information in distributional form (the priors) with empirical data, to obtain posterior distributions which quantify the uncertainty.

The papers that are presented in this themed issue capture some of the new and exciting developments being made in terms of modelling migration flows. Future work should focus on integrating the various aspects of migration data, theory and uncertainty into the modelling framework. One such integration is the aim of a project that has recently been funded by New Opportunities for Research Funding Agency Co-operation in Europe, entitled 'IMEM—Integrated Modelling of European Migration', and involving several of the researchers

contributing to this issue (Raymer *et al.*, 2010). This project has two main objectives. First, a Bayesian statistical model for international migration count data is developed that allows for flows to be measured to different accuracies and that can incorporate auxiliary information on the associations between origins and destinations of migration (e.g. language, borders and distance) to estimate missing flows. Second, the model accommodates the various complexities in the reported international migration data in the modelling framework. This involves the reconciliation of reported flows based on various definitions used by countries to measure international migration and the use of multiple sources of data on particular flows, including the use of qualitative judgements. Ultimately, the modelling approach will include multiple time periods and disaggregation by age and sex.

The contributions in this themed issue reflect a variety of approaches to modelling migration flows, whether they involve distributing immigration totals to local areas, combining multi-dimensional internal migration data obtained from multiple sources, measuring attractiveness and retentiveness of areas, forecasting immigration or modelling migration between countries in the context of inconsistent and missing data. The papers also set out practical strategies and innovative methods for improving the evidence for both policy and population planning, and our understanding of current patterns. This is important because migration is currently, and increasingly, the major factor contributing to population change in developed countries.

### References

- de Beer, J. (2008) Forecasting international migration: time series projections vs argument-based forecasts. In *International Migration in Europe: Data, Models and Estimates* (eds J. Raymer and F. Willekens), pp. 283–306. Chichester: Wiley.
- Bell, M., Blake, M., Boyle, P., Duke-Williams, O., Rees, P., Stillwell, J. and Hugo, G. (2002) Cross-national comparison of internal migration: issues and measures. *J. R. Statist. Soc. A*, **165**, 435–464.
- Bilsborrow, R. E., Hugo, G., Oberai, A. S. and Zlotnik, H. (1997) *International Migration Statistics: Guidelines for Improving Data Collection Systems*. Geneva: International Labour Office.
- Poulain, M., Perrin, N. and Singleton, A. (eds) (2006) *Towards Harmonised European Statistics on International Migration (THESIM)*. Louvain-la-Neuve: Université Catholique de Louvain–Presses Universitaires de Louvain.
- Raymer, J., Forster, J. J., Smith, P. W. F., Bijak, J., Wiśniowski, A. and Abel, G. J. (2010) The IMEM model for estimating international migration flows in the European Union. *Working Paper 14*. United Nations Economic Commission for Europe, Geneva. (Available from <http://www.unece.org/stats/documents/2010.04.migration.htm>.)

James Raymer and Peter W. F. Smith  
*University of Southampton*