

# QuantMig Migration Estimates: A New, Harmonised Set of Probabilistic Migration Flow Estimates for Europe, 2009-19

Georgios Aristotelous<sup>1</sup>, Peter WF Smith<sup>1</sup>, Jakub Bijak<sup>1,✉</sup>

<sup>1</sup>Department of Social Statistics and Demography, University of Southampton, Southampton, UK

## 1. Introduction: Towards Harmonised Migration Estimates for Europe

Problems with the availability, quality and comparability of official statistics concerning migration flows across Europe are well documented (Poulain et al. 2006). This recognition has led to the adoption in 2007 of the Regulation of the European Parliament and the Council No. 862/2007 on community statistics on migration and asylum statistics<sup>1</sup>. The provisions of the regulation required that the official migration statistics in the EU should be reported to a common standard, which in the case of flows should correspond to the duration of stay of 12 months or longer. The provisions of the regulation started being implemented across Europe from 2009 onwards.

At the same time, there have been many research efforts to harmonise the existing data through modelling, so that the resulting estimates based on disparate data can conform to standard definitions (e.g. De Beer et al. 2010; Raymer et al. 2013; Wiśniowski et al. 2016; Wiśniowski 2017). Following this tradition, in this paper, we present harmonised probabilistic estimates of migration flows among 32 countries in the European Union (EU), the United Kingdom, the European Free Trade Association (EFTA), and North Macedonia, as well as to and from the rest of the world, for 2009–2019, based on publicly-available Eurostat data.

The estimates, developed as a part of the QuantMig project ([www.quantmig.eu](http://www.quantmig.eu)) extend and update the previous work on the Integrated Modelling of European Migration (IMEM, Raymer et al. 2013, which produced estimates for 2002–2008), additionally allowing us to reflect on the state of European migration statistics following the adoption of Regulation 862/2007. In this extended abstract, we first summarise the architecture of the model underlying the IMEM and QuantMig estimates, before presenting selected results, and offering a discussion of implication for collection and analysis of migration data across Europe. Both the IMEM and QuantMig estimates are now available from the QuantMig Migration Estimates Explorer, at [bit.ly/quantmig-estimates](http://bit.ly/quantmig-estimates).

## 2. Model-Based Estimates: From IMEM to QuantMig

The estimation model is based the framework originally designed in the IMEM project (Raymer et al. 2013). The estimates are obtained by using a hierarchical Bayesian model and are therefore naturally accompanied by measures of uncertainty, for example based on the quantiles from the appropriate posterior distributions. The model hierarchy is summarised in Figure 1, and has two levels, separated by the horizontal dashed line: the higher-level *migration model*, based on a range of socio-economic drivers of migration (upper black dots), imputes the missing data, and the lower-level *measurement model* corrects the reported official statistics (lower black dots) according to their quality characteristics, mainly bias and variance (white dots). The estimated migration flows (grey dot) relate to long-term migrants, moving for 12 months or longer, as defined in the EC Regulation 862/2007.

---

<sup>1</sup> Regulation (EC) No 862/2007 of the European Parliament and of the Council of 11 July 2007 on Community statistics on migration and international protection, OJ L 199, 31.7.2007, p. 23–29, with subsequent changes. (<http://data.europa.eu/eli/reg/2007/862/oj>).

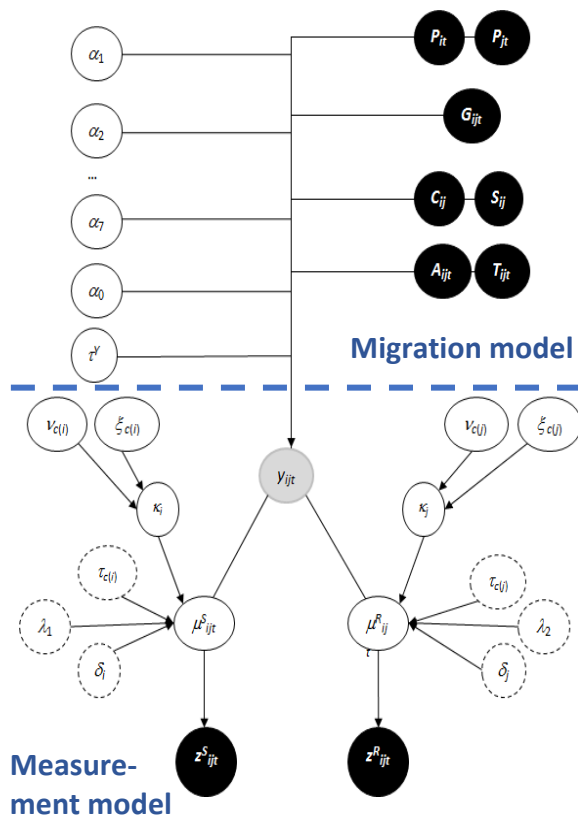


Figure 1. Basic Architecture of the IMEM Model  
Source: Raymer et al. (2013: 804)

Where possible, model parameters related to data quality were informed by an analysis of meta-information available from Eurostat or similar open sources, the results of which are reported in Mooyaart et al. (2021). This was particularly important, as since the IMEM project, the European migration data landscape has changed, notably as a result of adopting Regulation 862 on migration and asylum statistics in 2007.

The analysis focused on data availability, undercount of data and known issues with coverage, duration-of-stay definitions used in individual countries, accuracy (errors), and disaggregation by key dimensions of interest (age, sex, as well as region of birth). Overall, since the adoption of Regulation 862/2007, comparability of data across Europe has improved, but at the expense of reduced availability. Notable examples include Germany and Poland, two of the largest migration-sending and receiving countries in the EU, whose detailed data on migration flows by origin and destination are not currently available from Eurostat.

Another important element of the original IMEM model was expert opinion on those aspects of data quality that could not be identified solely from the available meta-information, and was introduced into the model as subjective, expert-based prior distribution. To inform the process for the updated estimates, we have updated the IMEM expert elicitation study of Wiśniowski et al. (2013). The results, reported in Keilman and Aristotelous (2021), indicate that expert-based information on migration data remains too uncertain to provide useful input for the part of the model dealing with data quality. Instead, we just used expert-based ratings of relative quality characteristics of migration data in individual countries, to inform some of the parameters (white dots) of the measurement model. The technical details and formal specification of the updated QuantMig version of the IMEM model is provided in Aristotelous et al. (2022).

### 3. QuantMig Migration Estimates Explorer: Selected Results

The update and refinement of the IMEM methodology enabled estimating the migration flows within the system of 32 European countries, as well as from and to eight other regions of the world, for 2009–2019 (Aristotelous et al., 2022). The estimates and their main statistical features (selected uncertainty bounds) are available in a queryable format from a dedicated online tool, the QuantMig Migration Estimates Explorer ([bit.ly/quantmig-estimates](https://bit.ly/quantmig-estimates)), broken down by countries of origin and destination, sex, five-year age groups, and by a broad region of birth (EU vs non-EU). An example of such estimates is shown in Figure 2 for the total immigration to Europe, juxtaposing the earlier IMEM estimates for 2002–2008 with the updated QuantMig ones for the later period. The break in series between the two sets of estimates reflects both changes in data harmonisation and availability, following the adoption of Regulation 862/2007, as well as an update and fine-tuning of the original IMEM methodology.

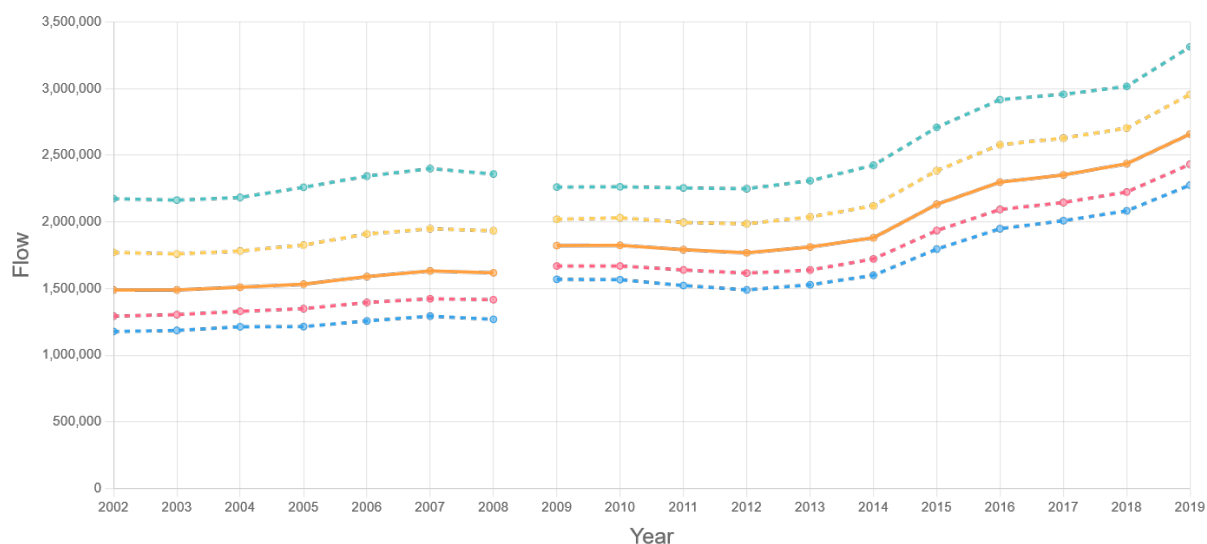


Figure 2. Estimated immigration to EU-27 countries in 2002–2019. Note changes in legislation, methodology and data availability in 2009. Source: QuantMig Migration Estimates Explorer, [bit.ly/quantmig-estimates](https://bit.ly/quantmig-estimates)

#### 4. Discussion: Reflection on European Migration Statistics

The QuantMig estimates, and their comparison with the earlier series obtained from the IMEM model for 2002–2008, facilitate an up-to-date reflection on data quality, illuminating the various trade-offs between data harmonization and availability since the Regulation 862/2007 came into force. In particular, we can reflect on possible reasons of discontinuities between the IMEM and QuantMig figures, both in terms of the central tendencies (median estimates), as well as their measures of uncertainty. Some issues clearly stand out here: while there was some welcome progress in terms of data harmonisation, significant improvements were hampered by lack of availability of data from some major European migration destination and origin countries, such as Germany and Poland (Mooyaart et al. 2021). From the modelling point of view, it is clearly preferable to have imperfect information, which may not conform to standard definitions or requirements, but which can still offer useful insights into relative distributions of quantities of interest, than to have no information. To help improve the data availability situation, barriers to reporting and therefore to fuller harmonisation of definitions across the EU need to be examined in more detail.

In methodological terms, the IMEM framework for estimating migration by combining different macro-level (aggregated) data – the ‘mirror statistics’ reported by sending and receiving countries wherever data are available – proved very versatile and useful. The presented approach can be naturally extended for other types of different data sources describing the same flow. One recent example of an application includes combining non-traditional data sources, such as digital traces, with traditional data, such as surveys, taking advantage of both the greater timeliness of the former and the known features of the latter (Rampazzo et al. 2021).

On the whole, the uncertainty of the harmonised migration estimates is large, but is reducible with greater data availability. At the moment, the information on migration contained in multiple data sources alone is insufficient to provide reliable assessment of the true flows and needs supplementing with expert assessment of data quality aspects. Still, information from experts and meta-information on data quality, coupled with available statistics on migration flows across Europe, enable the application of a very flexible framework of the IMEM model. Given the political prominence of migration, this is definitely an area worth investing in at the European level and with partner countries.

## References

- Aristotelous G, Smith PWF and Bijak J (2022) Technical report: Estimation Methodology. QuantMig Project Deliverable D6.3. University of Southampton.
- De Beer J, Raymer J and van der Erf R (2010) Overcoming the Problems of Inconsistent International Migration Data: A New Method Applied to Flows in Europe. *European Journal of Population*, 26, 459–481.
- Keilman N and Aristotelous G (2021) Expert Opinion on Migration Data. QuantMig Project Deliverable D6.1. University of Oslo and University of Southampton.
- Mooyaart JE, Daňko MJ, Costa R and Boissonneault M (2021) Quality Assessment of European Migration Data. QuantMig Project Deliverable D6.2. NIDI and MPIDR.
- Poulain M, Perrin N and Singleton A (eds.) (2006) *THESIM: Towards Harmonised European Statistics on International Migration*. Louvain-la-Neuve: Presses Universitaires de Louvain.
- Rampazzo F, Bijak J, Vitali A, Weber I and Zagheni E (2021) A framework for estimating migrant stocks using digital traces and survey data: an application in the United Kingdom. *Demography*, 58 (6), 2193–2218.
- Raymer J, Wiśniowski A, Forster JJ, Smith PWF and Bijak J (2013) Integrated Modeling of European Migration. *Journal of the American Statistical Association*, 108(503), 801–819.
- Wiśniowski A (2017) Combining Labour Force Survey data to estimate migration flows: the case of migration from Poland to the UK. *Journal of the Royal Statistical Society A*, 180(1), 185–202.
- Wiśniowski A, Bijak J, Christiansen S, Forster JJ, Keilman N, Raymer J and Smith PWF (2013). Utilising expert opinion to improve the measurement of international migration in Europe. *Journal of Official Statistics*, 29(4): 583–607.
- Wiśniowski A, Forster JJ, Smith PWF, Bijak J and Raymer J (2016) Integrated modelling of age and sex patterns of European migration. *Journal of the Royal Statistical Society A*, 179(4), 1007–1024.

Work funded from the EU's Horizon 2020 programme, grant No. 870299 *QuantMig: Quantifying Migration Scenarios for Better Policy*. Contributions by Michael Boissonneault, Rafael Costa, Maciej Daňko, Nico Keilman and Jarl Mooyaart, and comments by Jon Forster and Michaela Potančoková are gratefully acknowledged. All the interpretations are ours.

