

Towards Reliable Migration Statistics for the United Kingdom

Response to the House of Commons Public Administration Select Committee Call for Evidence on Migration Statistics

Jakub Bijak^{1,2,3}, George Disney¹, Sarah Lubman^{1,3} and Arkadiusz Wiśniowski²

¹ Division of Social Statistics and Demography, University of Southampton, Southampton SO17 1BJ

² ESRC Centre for Population Change, University of Southampton, Southampton SO17 1BJ

³ British Society for Population Studies (BSPS)

Contact: j.bijak@soton.ac.uk

Executive Summary

- This document discusses main sources of migration and asylum statistics in the United Kingdom in the international context, and makes recommendations considering the needs of data users.
- In order to estimate detailed breakdowns of immigration flows, administrative data sources and statistical modelling should be used to augment the International Passenger Survey estimates.
- To provide better estimates of emigration and immigration in the longer term, alternative sources of data, such as the e-Borders scheme, should be additionally included in the estimation.
- Recording the timing of events during the asylum process would enable detailed policy evaluation and allow asylum seekers to be correctly incorporated into international migration estimates.
- Meeting the requirements of the European law can be improved by combining migration data from other countries, alternative data sources and expert opinion, within a statistical framework.
- Measures of uncertainty in migration statistics should be reported to the users of data, possibly in conjunction with a bespoke decision advice.

Acknowledgements

The views, interpretations and recommendation presented in this document are the ones of the authors and do not necessarily represent the position of the House of Commons Public Administration Select Committee, University of Southampton, ESRC Centre for Population Change, British Society for Population Studies, or any other institution, with which the authors are affiliated. The original evidence is available at <http://www.parliament.uk/business/committees/committees-a-z/commons-select/public-administration-select-committee/inquiries/parliament-2010/statistics/migration-statistics> and is reprinted by a kind permission of the Committee.

We wish to thank Professor Nico Keilman (University of Oslo) and Professor Peter W. F. Smith (University of Southampton) for helpful remarks, which have improved the earlier draft. All the remaining errors are exclusively ours.

I. Introduction

1. Next to the statutory requirements concerning the delivery of high-quality migration statistics for various users in the United Kingdom (UK), the Office for National Statistics (ONS) is now formally required to provide the Eurostat – the statistical office of the European Union (EU) – with reliable data on international migration flows. These data need to comply with the standard definition of long-term international migration, as recommended by the United Nations (UN 1998), whereby a migrant is someone who changes their usual place of residence for at least one year.

2. The requirement concerning Eurostat stems from Article 3 of the Regulation (EC) No. 862/2007 of the European Parliament and the Council on Community Statistics on Migration and International Protection¹. The Regulation introduces a set of definitions and common requirements for statistics on migration flows, stocks, the citizenship and country of birth of persons, as well as administrative procedures concerning immigration, residence permits, asylum and prevention of illegal immigration. To fulfil the requirements, reliable data are required on both immigration and emigration, with further detail on the citizenship and countries of origin and destination of migrants.

3. In this context, the current Response to the Call for Evidence discusses several aspects of the UK international migration statistics. Firstly, in Section II sources of data on the main categories of migrants are examined, followed by an evaluation of the Home Office data on asylum seekers in Section III, and a presentation of the UK migration data in a wider European setting in Section IV. The Response concludes by briefly discussing the issue of uncertainty inherent in migration statistics in Section V, and providing some recommendations for the future in Section VI.

II. Sources of International Migration Statistics in the United Kingdom

4. The main source of official data used to estimate overall levels of immigration to the UK, and the only source of information on total emigration, is the International Passenger Survey (IPS), further adjusted to take into account asylum seekers (see Section III), people who change their migration intentions (“switchers”), and migrants to and from the Republic of Ireland (ONS 2012a). The IPS was originally intended to provide data on travel and tourism; questions about international migration are a later addition.

5. As with all surveys, the IPS is subject to sampling error. International migration, however, is still a rare event in comparison to other international travel such as holidays and work-related journeys. Consequently, when data on international migration are required in a breakdown by age or citizenship, the estimates provided by the IPS are likely to be unreliable.

6. It is worth bearing in mind, though, that the IPS is the only UK source of migration data that uses the UN (1998) definition of an international migrant, and does relatively well estimate large flows of migration based on sufficient numbers of respondents. Empirically, large flows do not exhibit such high degree of random variability as smaller ones. In the short term, these two strengths of the IPS should be utilised fully, with other sources of information benchmarked to these reliable elements.

1 Official Journal OJ L 199, 31.07.2007, pp 23-29; <http://eur-lex.europa.eu>.

7. To obtain reliable international migration flow statistics further data are required to augment or replace the IPS. There are several options available. Firstly, more data can be collected, ideally not from a sample survey. One possible source of such information in the future can be the e-Borders project (ONS n.d.). Secondly, following Willekens (1994), various existing sources of administrative data can be used to enhance the IPS, as is already done by the ONS for distributing the total flows into UK regions (Bijak 2010b).

Such alternative sources are documented in the ONS Conceptual Framework of Population and Migration Statistics (Raymer et al. 2012b; cf. Boden and Rees 2010). Thirdly, statistical modelling can be used to bring different sources together – examples in a multi-national context are discussed in Section IV, but the same principle can be applied to include information from different UK sources in a single statistical model, as set forth below.

8. For immigration to the UK, there are two examples of administrative data sources that provide information on the citizenship of migrants. The first is data of non-UK domiciled students from the Higher Education Statistics Authority (HESA). The second is data on the number of new National Insurance Number (NINo) registrations of foreign nationals, from the Department for Work and Pensions (DWP). Neither source of data includes information on British nationals returning to the UK. They are also not specifically designed to measure immigration, and thus do not conform to the UN definition of a long term migrant used in the IPS. Being administrative records, the HESA and DWP data are not subject to the sampling error, yet still can be biased due to differences in the underlying concepts and definitions, as well as in their coverage (see Section IV).

9. Thus, as mentioned above, one possible method to improve the international migration estimates by citizenship, if the data are available at an aggregate level rather than at an individual level, is to use a statistical model to combine the auxiliary datasets with the IPS. This could take the form of benchmarking the auxiliary data to the reliable parts of the IPS, and modelling various structures – by age, country etc. – from this partial information by using indirect methods (e.g. Rogers et al. 2011). Here the main strength of the IPS – its match to the UN definition of international migration – would be used as a benchmark to adjust the reliable patterns of the auxiliary data. This could provide more reliable and accurate estimates of the immigration in different breakdowns.

10. So far, there are no auxiliary sources of data on emigration, however. To improve detailed estimates of emigration a new source of data is required. With emigration forming one half of international migration, it is vitally important for the UK to collect better data here. It can be expected that in that respect the data situation will improve once the e-Borders scheme is fully operational, however not earlier than before the end of 2017 (ONS n.d.).

11. The statistics on immigration and emigration can be further supplemented by estimates of net migration, based on the census population counts and statistics on vital events (births and deaths), combined through the population accounting equation. However, net migration is being criticised for being an artificial category that conflates two entirely different processes: inflows and outflows (Rogers 1990). In this context, we suggest that the UK Government migration target is revisited. Since the authorities can only extend very limited control over emigration, and some immigration flows (UK and EU nationals), meeting the net migration target may be often a matter of chance, rather than of policy.

III. Home Office Asylum Statistics

12. The total number of asylum seekers is inherently easy to measure given the legal requirement that an application for 'leave to remain' must be lodged with the authorities. It is the transitions and timing between status changes which are more complex – but also more useful – to measure. The amount of data published by the Home Office on asylum applicants has increased in recent years (Home Office 2012). There is more information available on key characteristics of asylum seekers (e.g. age, sex, nationality) as well as a greater temporal breakdown, with many statistics now produced by quarter or month. These developments are clearly to be welcomed as they enable the analysis of trends and more in depth policy evaluation; however, the limited retrospective statistics available continue to hinder the study of past trends and limit comparisons over time.

13. Clarity around the timing of when events occur and are recorded is crucial in the provision of accurate, informative and comparable statistics. Varying lengths of time between each stage in the asylum process, particularly with the backlog of cases awaiting a final decision, mean that clearly identifying stocks and flows over time is problematic. Access to the data on decision status, cross-tabulated with other information, such as geographic area and the level of support received, would increase the analytical potential considerably.

14. As mentioned in Section II, the total number of asylum applicants in a given year (excluding those known to have left the country within a year of their application, and correcting for an overlap with the IPS) is added to the IPS-based estimates, in order to produce international migration statistics for that year (ONS 2012c). However, some of the asylum seekers may already have been in the country for a considerable length of time, and many will leave after a few months. It is not possible to establish how many asylum seekers should be counted as international migrants under the UN definition unless individual-level data on length of time between each stage are used, as it may be the case under the e-Borders scheme.

15. Data on the dispersal of asylum seekers in need of housing support and on the receipt of subsistence-only support are published quarterly, in a breakdown either by Local Authority or by nationality. Access to these data by Local Authority and by nationality would again increase the potential for analysis, although there may be confidentiality considerations which need to be overcome in order to achieve this. Understanding geographic variation is essential to allocate resources and effectively implement policy in an efficient, fair and transparent way. Releasing data from housing providers at a regional or local level could contribute to this.

16. The only data available on asylum seekers after they have received a final decision is a record of the numbers deported or those who have left the country through notified/assisted voluntary return. Clearly, there are no data on how many people remain in the country or leave without informing the authorities. There is also a notable absence of data on the refugee population in the UK, although some effort has been made to improve knowledge about this group through the Survey of New Refugees (Home Office 2010). A lack of data on asylum seekers after they have received their final decision means that it is extremely difficult to support the integration process of individuals who have been granted leave to remain, and to follow up the unsuccessful cases.

IV. UK Migration Statistics in the International Context

17. In general, the availability and quality of migration data in Europe is problematic (Poulain et al. 2006). The data collected by statistical institutes are not reliable and often do not comply with the standard definitions. The main reasons include migration undercount and imperfect data coverage. The undercount concerns sources based on self-declarations. The problem is deemed more severe for emigrants, who usually have fewer incentives to deregister from the system than immigrants, who, after registration, may gain access to certain benefits. The insufficient coverage primarily concerns three population groups: undocumented migrants, asylum seekers and foreign students.

18. The poor reliability of migration data leads to a lack of their international comparability. Besides, there are differences between countries in the adopted definitions of migration and in the data collection mechanisms (Poulain et al. 2006). These differences encompass the temporal dimension of migration, i.e. the minimum duration of stay applied to define a migrant, as well as the populations included in the definition. Different rules often pertain to nationals, foreigners from the other EU countries, as well as third-country (non-EU) nationals.

19. Regulation 862/2007 aims at harmonising the statistics on migration in the EU according to the UN (1998) definition. Harmonisation here refers to the process of reconciling the differences between different measurements of migration data. So far, the implementation of the Regulation has brought about some improvement to the migration statistics (European Commission 2012). The data provided to the Eurostat by the EU Member States are now more complete and are based on a common definition. However, some weaknesses remain and, apart from the Nordic countries, the information on migrants is usually not shared between the statistical offices of the EU countries.

20. According to Regulation 862/2007, Article 9, scientifically based and well documented statistical estimation methods may be used to produce migration statistics. An approach encouraging an integrated use of migration flow data from all EU and European Free Trade Association (EFTA) countries has recently been proposed within the project 'IMEM: Integrated Modelling of European Migration' (Raymer et al. 2012a; Wiśniowski et al. 2012). The approach advocates explicitly accounting for the main aspects of the differences in measurement: intended duration of stay in the UK, undercount, coverage and accuracy of the data collection mechanisms, based on migration flow data from different countries and the covariate information. Additionally, expert knowledge is elicited to learn about the characteristics that cannot be explained by the data, such as the immigration undercount.

21. Another path of improving migration statistics is the Migration Statistics Mainstreaming programme, adopted by Eurostat (Knauth 2011). Here, it is advocated that the migration perspective should be added to all statistics produced by using the available administrative data sources and surveys. Thus, a dimension of migration status should be introduced to all official statistics on income, education, employment or living conditions.

22. In parallel, as part of the mainstreaming, the sampling frames of surveys should be improved in order to capture sufficient number of migrants. For large-scale surveys, such as the LFS and the European Union Statistics on Income and Living Conditions (EU-SILC), sample limitations can be overcome by reviewing the methodologies and the field work approach. First steps into using the large survey data for measuring migration have been undertaken by Rendall et al. (2003), Ródenas and Martí (2007), and Wiśniowski (2012).

V. Uncertainty in Migration Statistics

23. The presence of uncertainty in migration and population estimates is well acknowledged (e.g. ONS 2012b). This uncertainty should ideally be reflected in the statistics in a measurable way, such as through probability distributions, which however may pose challenges to the users of migration data. On the other hand, Lawrence et al. (2006) cited several studies which found that users prefer estimates with some uncertainty measures, such as confidence intervals, as clearly providing more information than point estimates alone. Of course, too wide intervals are hardly informative, while too narrow ones are more likely to miss the reality, and can give the users a false illusion of control. The challenge of prudent decision making is thus in retaining the balance between these opposites.

24. A solution to this challenge might be provided by the statistical decision analysis, whereby the costs of different decisions in the light of under- and over-estimation of migration are taken into account. As noted by Alho and Spencer (2005) and Bijak (2010a), this procedure is specific to the particular decision context – there are no ready answers – and requires communication between the producers and users of the estimates. Should the users share the details of their decision context, the producers of the statistics would be able to offer a bespoke decision advice on that basis.

VI. Conclusions and Recommendations

25. In conclusion, we strongly recommend the use of alternative sources of data to aid the estimates of international migration. We suggest that, in the short term, alternative sources of data and statistical modelling should be utilised to help improve the UK international migration estimates. Administrative data sources have been used by the ONS in their estimates of the local geographical distribution of international migrants (see Bijak 2010b), and this could be extended to estimating the citizenship-detail of immigration flows to the UK.

26. In the long term, we recommend the fullest possible use of information available from the e-Borders scheme. Still, we do not consider it to be a panacea for all challenges of migration statistics, and suggest that it complements, rather than replaces the existing sources of data. In particular, e-Borders data can be very useful for benchmarking the other information to the UN (1998) definition, since it will enable relating migration intentions to the actual length of stay in the country.

27. When statistical modelling is used, we recommend that the results are reported together with a range of the associated uncertainty measures. An interactive engagement of the producers of the statistics with the users is suggested in order to help utilise this information fully, taking into account the decision context. We believe that an honest reporting of the imperfections of knowledge on migration will provide a prudent approach to the migration challenges facing the United Kingdom.

References

- Alho JM and Spencer BD (2005) *Statistical Demography and Forecasting*. New York: Springer.
- Bijak J (2010a) *Forecasting international migration in Europe: A Bayesian view*. Springer Series on Demographic Methods and Population Analysis, vol. 24. Dordrecht: Springer.
- Bijak J (2010b) Independent Review of Methods for Distributing International Immigration Estimates to Regions. Report to the Office for National Statistics. <http://www.ons.gov.uk/ons/guide-method/method-quality/imps/updates-and-reports/historical/updates-and-reports-from-2010/independent-review-of-methods-for-distributing-international-immigration-estimates-to-regions---january-2010.pdf> [14/01/2013]
- Boden P & Rees P (2010) Using administrative data to improve the estimation of immigration to local areas in England. *Journal of the Royal Statistical Society A*, 173(4), 701–731.
- European Commission (2012) Report from the Commission to the European Parliament and the Council on the implementation of Regulation (EC) No 862/2007 on Community statistics on migration and international protection. Brussels, 20 September 2012. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0528:FIN:EN:HTML> [14/01/2013]
- Home Office (2010) UK Border Agency. Analysis, Research and Knowledge Management Survey of New Refugees. <http://www.esds.ac.uk/findingData/snDescription.asp?sn=6556> [14/01/2013]
- Home Office (2012) Data tables - Immigration Statistics July - September 2012. <http://www.homeoffice.gov.uk/publications/science-research-statistics/research-statistics/immigration-asylum-research/immigration-tabs-q3-2012/>
- Knauth B (2011) Migration Statistics Mainstreaming. Paper presented at the 58th ISI Congress, 21–26 August 2011, Dublin. <http://isi2011.congressplanner.eu/pdfs/650162.pdf> [14/01/2013]
- Lawrence M, Goodwin P, O'Connor M & Önkäl D (2006) Judgemental forecasting: A review of progress over the last 25 years. *International Journal of Forecasting*, 22(3): 493–518.
- ONS (2012a) Quality and methodology information for Long-term International Migration estimates. Titchfield: Office for National Statistics. <http://www.ons.gov.uk/ons/guide-method/method-quality/quality/qualityinformation/social-statistics/quality-and-methodology-information-for-long-term-internationalmigration-estimates--ltim-.pdf> [14/01/2013]
- ONS (2012b) Uncertainty in local authority mid-year population estimates. Titchfield: Office for National Statistics. <http://www.ons.gov.uk/ons/guide-method/method-quality/imps/latest-news/uncertainty-in-lamypes/index.html> [14/01/2013]
- ONS (2012c) Long-Term International Migration Estimates. Methodology Document. 1991 Onwards. Nov. 2012. Titchfield: Office for National Statistics. <http://www.ons.gov.uk/ons/guide-method/method-quality/specific/population-and-migration/international-migration-methodology/long-term-international-migration-estimates-methodology.pdf> [14/01/2013]

- ONS (n.d.) Delivering statistical benefits from e-Borders. Titchfield: Office for National Statistics. <http://www.ons.gov.uk/ons/guide-method/method-quality/imps/latest-news/delivering-statistical-benefits-from-e-borders/index.html> [14/01/2013]
- Poulain M, Perrin N & Singleton A (eds.) (2006) *THESIM: Towards Harmonised European Statistics on International Migration*. Louvain: Presses Universitaires de Louvain.
- Raymer J, Forster JJ, Smith PWF, Bijak J & Wiśniowski A (2012a) Integrated Modelling of European Migration: Background, specification and results. NORFACE Migration Discussion Paper 2012-04. http://www.norface-migration.org/publ_uploads/NDP_04_12.pdf [14/01/2013]
- Raymer J, Rees P, Blake A (et al.) (2012b) Conceptual Framework for UK Population and Migration Statistics. Report to the Office for National Statistics. <http://www.ons.gov.uk/ons/guide-method/method-quality/imps/latest-news/conceptual-framework/a-conceptual-framework-for-population-and-migration-statistics---download-file.pdf> [14/01/2013]
- Rendall M, Tomassini C & Elliot DJ (2003) Estimation of annual international migration from the Labour Force Surveys of the United Kingdom and the continental European Union. *Statistical Journal of the United Nations ECE*, 20, 219–234.
- Ródenas C. and M. Martí (2007) Migration Estimation Based on the Labour Force Survey: An EU-15 Perspective. *International Migration Review*, 41(1), 101–126.
- Rogers A (1990) Requiem for the net migrant. *Geographical Analysis*, 22(4), 283–300.
- Rogers A, Little J & Raymer J (2010) *The indirect estimation of migration: Methods for dealing with irregular, inadequate, and missing data*. Springer Series on Demographic Methods and Population Analysis, vol. 26. Dordrecht: Springer.
- UN (1998) *Recommendations on statistics of international migration*. New York: United Nations. http://unstats.un.org/unsd/publication/SeriesM/SeriesM_58rev1e.pdf [14/01/2013]
- Willekens F (1994) Monitoring international migration flows in Europe. Towards a statistical data base combining data from different sources. *European Journal of Population*, 10(1), 1–42.
- Wiśniowski A (2012) Bayesian modelling of international migration with Labour Force Survey data. Paper presented at the European Population Conference, Stockholm, 13–16 June 2012. <http://epc2012.princeton.edu/papers/120192> [14/01/2013]
- Wiśniowski A, Keilman N, Bijak J, Christiansen S, Forster JJ, Smith PWF & Raymer J (2012) Augmenting migration statistics with expert knowledge. NORFACE Migration Discussion Paper 2012-05. http://www.norface-migration.org/publ_uploads/NDP_04_12.pdf [14/01/2013]