Raymer, James

Article

Measuring flows of international migration

IZA World of Labor

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: Raymer, James (2017) : Measuring flows of international migration, IZA World of Labor, ISSN 2054-9571, Institute for the Study of Labor (IZA), Bonn, https://doi.org/10.15185/izawol.354

This Version is available at:
https://hdl.handle.net/10419/162364

JAMES RAYMER
Australian National University, Australia, and IZA, Germany

I Z A
World of Labor
Evidence-based policy making

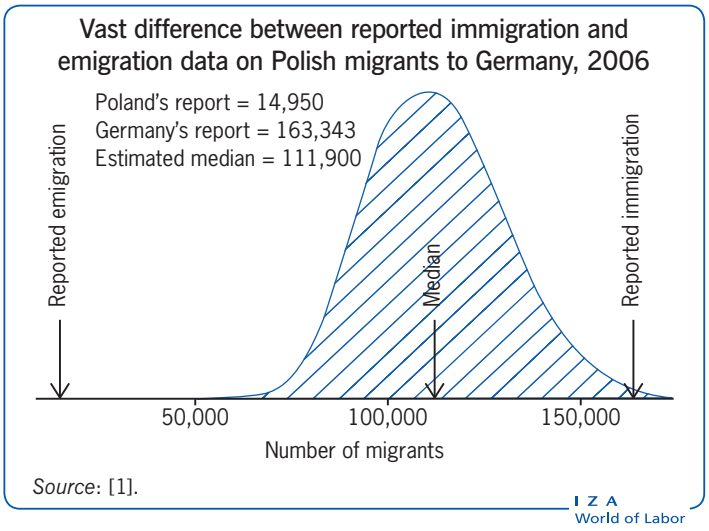
Measuring flows of international migration

Consistent measures of migration are needed to understand patterns and impacts on labor market outcomes

Keywords: international migration, measurement, data

ELEVATOR PITCH

International migration alters the socio-economic conditions of the individuals and families migrating as well as the host and sending countries. The data to study and to track these movements, however, are largely inadequate or missing. Understanding the reasons for these data limitations and recently developed methods for overcoming them is crucial for implementing effective policies. Improving the available information on global migration patterns will result in numerous and wide-ranging benefits, including improved population estimations and providing a clearer picture of why certain migrants choose certain destinations.



KEY FINDINGS

Pros

- Migration is important for understanding population and societal changes.
- Data on international migration flows are becoming increasingly available, especially in Europe.
- Countries can improve their migration flow reports by sharing data with each other.
- Statistical modeling can be used to harmonize and estimate missing and conflicting international migration flows.
- Measures of uncertainty improve researchers' understanding of the quality of migration data and estimates.

Cons

- International migration data are highly inconsistent and incomplete due to different measurements and collection methods.
- The effects of incorrect measurement on the levels of migration are poorly understood.
- Even the best available data sources likely undercount flows of immigration and emigration.
- Most national statistical offices do not share information on cross-border movements.
- It is unrealistic to expect countries to change their data collection practices in the next ten years.

AUTHOR'S MAIN MESSAGE

Migration flow data are deficient due to differences in measurement and collection systems. However, most analysts ignore this and instead design policies based on incomparable or inaccurate data. In the best case, effective information sharing and standardized migration flow measuring practices would be adopted on a global scale. However, this is a lofty and perhaps unrealistic short-term goal. In the interim, recent research using statistical modeling techniques to produce synthetic data holds great potential to provide more reliable and consistent information on international migration and its impacts over time.

MOTIVATION

International migration has become an increasingly important global issue. Despite long-term efforts by the UN to provide clear guidelines on how to measure migration, very little is known about the actual number of annual migrants throughout the world and the scarce information available is contradictory [2], [3]. Today, countries typically rely on their own definitions of what constitutes a "migration." This creates inconsistencies among international data and makes it challenging to understand the process by which people move across national borders.

Consistent and reliable data on international migration flows are needed so that governments know where their populations are moving; this knowledge would enable governments and policymakers to recruit the appropriate types of workers needed in increasingly specialized markets, or to develop policies for providing effective services for migrants.

DISCUSSION OF PROS AND CONS

Different types of migration data

There are many types of migration data to consider, with migrant population stocks and migration flows representing the two main categories used for analysis [4]. Populations disaggregated by place of birth represent the most abundant migration data available because they can be consistently collected from census data. Most countries conduct censuses about once every ten years with a question on country of birth. The number of people born abroad as measured in nearly all population censuses represents the net cumulative effect of immigration and emigration over time. This information is important for understanding the long-term effects of migration and the characteristics of migrant populations, but it does not reveal when migrants arrived in a specific country or how many people have exited.

Migration flows, on the other hand, are much less abundant and rarely collected in a consistent manner. Flows capture the number of people moving within a specified period of time, usually a one-year period. These data are needed to study the push and pull factors of migration between origins and destinations and the deterrence effects of distance, costs, and cultural differences. Annual flows are also required for assessing migration's contribution to demographic change in relation to natural increase (i.e. the number of births minus deaths) occurring in a population. They are also essential for understanding policies designed to regulate migration and for developing any improvements to these polices.

A critical issue for determining consistent and reliable data on this topic involves the measurement of migration flows. Ideally, consistent information would be available on both flows and stocks of migration, which would allow people to understand how migration flows are changing populations' characteristics throughout the world and which groups are contributing most to that change. It is important, for example, to know if some migrants are more likely to stay than others, or, if they are likely to bring their families with them or not. It might also be useful to know whether migrants bring existing skills to their host countries, or whether they seek to gain skills after arrival.

There are three main types of migration flows; they refer to the numbers of people moving by place of birth, citizenship, and previous/next country of residence. Country-of-birth flows are

particularly valuable for understanding how birthplace-specific migrant population stocks change over time and throughout the world. The UN Population Division, for example, provides international migrant population stock data through its Global Migration Database. Some countries and many policymakers are only interested in keeping track of those who need permission to enter and remain in the country (i.e. citizenship-related issues). These countries focus on entries, visas, and citizenship status that are directly related to the legal status of the migrants and the services they receive. To understand where people are moving to, how these patterns differ from other countries, and how they contribute to population redistribution, information is needed on the origins and destinations of migration. All three types of migration flow data are important in their own right; however, this article focuses primarily on the third main flow type, often referred to as the "change in usual country of residence" notion of migration.

The UN defines an international migrant as "a person who moves to a country other than that of his or her usual residence for a period of at least a year (12 months), so that the country of destination effectively becomes his or her new country of usual residence" [3]. This particular definition has been in place for nearly 20 years, yet hardly any noticeable changes to the measurement of migration data have occurred in response to its formation. Actually, the problem of measuring migration goes back much further in time [4], [5], [6], [7]. Considering the enormity of the issues associated with migration, it is surprising that this problem still exists today. Current practices of migration data collection are predominately driven by administrative requirements to process the entry of foreigners rather than for the purposes of demographic accounting or cross-national comparison.

To ensure consistency in the measurement of migration flows, countries need to collaborate with each other. This work has begun in the EU over the last ten years [2]. Starting in 2007, the European Parliament passed a regulation requiring countries to provide harmonized migration flow statistics to Eurostat in accordance with Regulation 862/2007. The regulation specifies a set of tables that each EU member state must provide for comparison across Europe. Countries are required to use the best available data; however, they are not required to change their existing administrative systems or to collect new data. To help achieve the goal of consistent and complete data, Article 9 of the Regulation states that: "As part of the statistics process, scientifically based and well documented statistical estimation methods may be used." While this has not resulted in major differences in the collection of migration data in Europe, it has led to some methodological improvements to estimate and harmonize flows, and has certainly promoted research into assessing the data's quality [8], [9], [10]. Elsewhere in the world, there has been little progress on these issues, but the lessons learned from Europe may serve as a guide for future data collection and statistical estimation practices in other regions.

Why do migration data differ?

Migration data are primarily gathered to keep track of foreign nationals entering countries and to ensure that they are legally allowed to work, study, join family members, or seek refuge. As a result, migrants are often categorized into different groups, such as international students, asylum seekers, laborers, or people joining family already in the country. Data are rarely gathered for the purposes of measuring demographic change or to compare the levels of migration across countries [2]. Moreover, the emphasis is usually placed on those entering and not on those leaving. Outside the EU, there are no legally binding incentives for countries to provide internationally comparable data. Countries

gather their own data to meet their own requirements, meaning that migration data are measured in a vast number of different ways. Therefore, to understand movements across countries and over time, one must have both a detailed understanding of each country's particular collection method and the means to reconcile different data measurements. Obviously, this implies significant challenges.

The availability of statistics on international migration flows is conditioned by the existence of country-specific data collection systems that provide meaningful information. Most countries simply do not have the capability or motivation to provide such data. For those that do, the major types of data sources used to produce statistics on international migration flows are: (i) population registration systems; (ii) other administrative registers related to foreigners, alien registers, residence permit databases, or asylum seeker databases; (iii) statistical forms filled in for all changes of residence; and (iv) border crossing data collection and other surveys [2].

International migration flows can also be obtained from population censuses, but these are rarely included for a number of reasons: First, intervals between census dates are long. Second, censuses are only able to capture an individual's current residence and residence at particular points in the past (e.g. one year or five years ago), and finally, they are only able to identify immigrants, as emigrants are no longer present to be counted. Sample surveys face similar difficulties, but on top of that, their main obstacle is that most samples are not large enough to capture the relatively small number of people who make an international move. As such, migration flow data obtained from censuses and surveys are usually not considered when reporting international migration flows.

In situations where migration flow data are published by national statistical offices (i.e. the data are typically produced for purely administrative purposes), they may not be reliable for understanding migration patterns [2]. First, migration flows are likely to be under-counted in countries where data collection relies on self-declarations. Second, subgroups of the total migration flow may be excluded, particularly if different types of migrants are treated separately. For example, asylum seekers may only be included once they have been granted refugee status and have received temporary or permanent residence permits. In many developed countries, it is common for students not to be included in the population registers and to not deregister after they leave. For students originating from outside the EU, the situation is considered more reliable, at least for those entering, as all of them are required to obtain a specific residence permit. Finally, immigration statistics are usually considered more reliable than emigration statistics. This is because people have more incentives to report their arrival to a country (e.g. to receive access to health care services), than their departure. Consequently, countries may overstate their net immigration totals if they do not have an accurate account of people that have left.

The measurement of international migrants may also differ within countries depending on how the various sources of statistical information are measured and if they have changed over time. Furthermore, there is the possibility that flows of immigration are measured differently from emigration and that foreigners are measured differently from nationals. For example, in Finland all emigrants are measured according to a 12-month duration definition (i.e. only people who have been out of the country for 12 months are considered to have emigrated). Likewise, foreigners coming into Finland are only considered migrants after 12 months. However, nationals returning from abroad are registered as soon as they arrive home, thus bypassing the 12-month qualification period applied to other

migrant groups. Depending on the measurement criteria of the other country involved in the emigration/immigration situation, one person may be counted as having multiple residence statuses. Ideally, all countries would adopt a unified definition for all types of flows and base their measurement of migration either on a change of usual residence or a 12-month definition. This would guarantee that a person can only have one country as their usual residence at any one time.

It should be noted that a further issue concerns the notion of "usual residence," which can be measured in different ways for different population groups. Nationals, for example, have an unconditional right of residence in their country of citizenship. Among foreigners, there are those with the right to access the labor market and social services and potentially achieve citizenship, while there are others, such as international students or asylum seekers, who are more restricted in the types of activities they can do and the duration of their stay. The difficulty with nationals is that they may still be counted as part of the population even after they have been living abroad for a number of years if there is no clear mechanism or incentive for the emigrant or receiving country to inform the origin country's population register.

Challenges involved in handling unstandardized migration data

Without unified definitions of usual residence, it is impossible to know the real levels of international migration. At present, some countries measure a change in the country of residence according to a minimum duration of being present in the country or being away from the country. Others use a vague notion of change of residence, such as the idea of "permanent" residency, without specifying a precise duration. When a precise period is used, another problem arises related to the distinction between intended and actual duration [2]. The use of the *actual* duration results in the reported statistics being delayed for publication. As a result, most countries that specify a precise period use the *intended* duration under the assumption that the intended duration will become the actual one. In reality, the two measures may differ considerably, depending on the origin country and the economic situations or relative successes of the migrants.

Other problems with migration measurements are associated with the timing of data collection [2]. For immigration this might be the date of issuing a permit, the date of arrival, or the date of registration. For emigration, the date of expiry of a permit, the date of reporting the departure, or the date of departure are variously used. When a very short (or no) duration of stay criterion is employed, it is possible that an individual could migrate more than once during the reference period, which may inflate the migration numbers relative to the real amount. To achieve comparable statistics, only one migration should be allowed per migrant within the measured period.

Available data on international migration flows are still a long way from being comparable across countries. This is evident when comparing data on flows between pairs of countries that are reported by countries of origin and countries of destination, using a so-called double-entry matrix [8], [9]. This is exemplified by the illustration on page 1, which compares official statistics from Poland showing the number of Polish citizens reported as emigrating to Germany with the German statistics showing the number of Polish citizens they reported as immigrating into Germany in the year 2006. As seen, the figures vary dramatically, with a more than ten-fold difference between the two countries' reported

data. In an ideal world, the emigration figures produced by sending countries and the immigration figures collected by receiving countries would be the same; however, this requires that the two countries' data collection systems use identical definitions and the data are reliable and complete.

How to overcome migration data limitations?

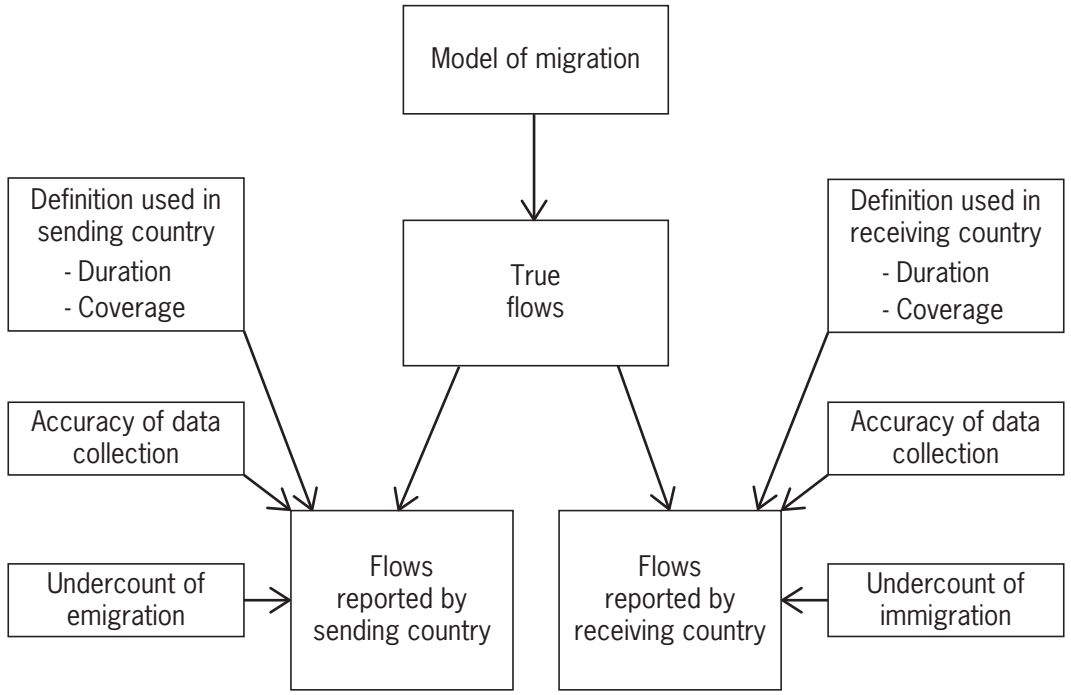
There are two ways to overcome the current problems of inconsistent and incomplete migration data [9]. The first requires that national statistical offices in different countries share information about their migrants. This already happens amongst the Nordic countries of Denmark, Finland, Iceland, Norway, and Sweden—thus providing the world with a best-practice example of how reliable migration statistics may be produced. They exchange information with each other by notifying the sending country when someone from that country has been registered in their system. This creates a situation whereby one person can be included on only one population register at a time (within the Nordic system).

Statistical models provide the second method; they help to harmonize the variations between different countries' reported figures on migration and to estimate missing data [10]. Since 2007, there have been two international and interdisciplinary research projects focused on modeling migration flows in Europe [9]: Migration Modelling for Statistical Analyses (MIMOSA) conducted from January 2007 to December 2009 [2], [10] and Integrated Modelling of European Migration (IMEM) conducted from November 2009 to April 2012 [1], [11], [12]. The IMEM project differed from the MIMOSA project by modelling the measurement aspects of data, incorporating expert information, and including measures of uncertainty.

The conceptual model framework developed by the IMEM project is presented in Figure 1. The objective of this model is to estimate an unobserved set of true flows of migration based on (i) flows reported by the sending country, (ii) flows reported by the receiving country, (iii) covariate information, and (iv) expert judgments. The measurement of interest is the "change in country of usual residence," consistent with the UN recommendation [3]. The IMEM measurement models distort the true migration flows by taking into account duration definitions used in various countries (e.g. the 12-month rule), relative accuracy of the data collection mechanisms (e.g. population register, survey), the overall undercount of migration, and coverage (e.g. includes international student arrivals, asylum seekers). Expert judgments are included in the IMEM project to augment the measurement model, especially to include information on the level of unobserved undercounts in both immigration and emigration.

Coming back to the estimated flow of migrants from Poland to Germany presented in the illustration on page 1: Vastly different reports were provided by Poland and Germany, and both countries' reports differed substantially from the estimated posterior distribution provided by the IMEM model. Two important points are worth highlighting about this case. First, the single reports provided by Poland and Germany are misleading. It is difficult to imagine how their reports for the same flow could be so different: Poland reported about 15,000 people migrating to Germany in 2006, while Germany reported nearly 164,000 Polish immigrants in the same year. Closer examination of how the flows were measured reveals that Poland used a very restrictive "permanent" definition of migration, which implies that migrants never intend to return. Germany, on the other

Figure 1. Conceptual framework for modeling migration flows



Note: The (unobserved) true flows of migration are estimated by using data from the sending and receiving countries, adjusted for measurement differences, and augmented with a spatial interaction model of migration.

Source: Raymer, J., A. Wiśniowski, J. J. Forster, P. W. F. Smith, and J. Bijak. "Integrated modeling of European migration." *Journal of the American Statistical Association* 108:503 (2013): 801–819 [1].

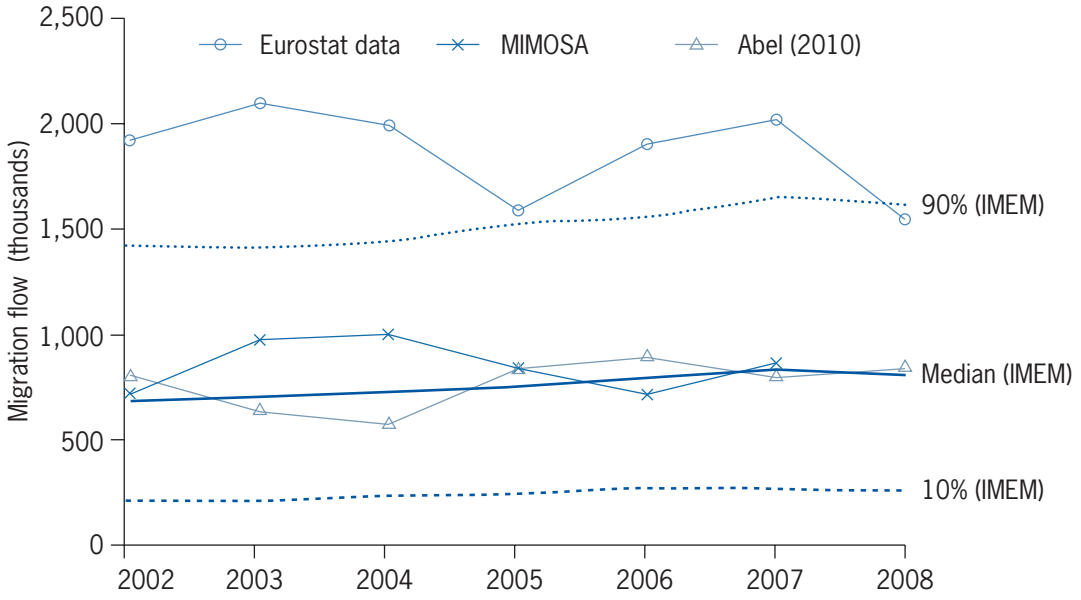
I Z A
World of Labor

hand, used a relatively loose "instant" definition of migration, implying that migration occurs upon registration with local authorities, regardless of intended stay.

The IMEM estimates in the illustration on page 1 are presented in the form of a distribution from which any measure can be drawn. For example, the median flow estimate was 112,000 with an interquartile range of 24,000. Alternatively, one could say that 90% of the estimates fell between 87,000 and 143,000. The IMEM estimates thus provide a range of plausible flows and a sense of accuracy. The distribution of the estimates is based on the differences in the reported figures, data measurement, data collection system, and expert information regarding these aspects and levels of undercount.

One thing that has become clear from the IMEM research [1] and earlier efforts by the MIMOSA team [2], and a third relevant model developed by Abel [13], is that the reported population totals for countries in the EU and European Free Trade Association (EFTA) are overstated. This is because they do not adjust for different measurements of migration data, nor do they account for underreporting of emigration. Figure 2 shows Eurostat's reported net migration flows in the EU/EFTA area from 2002 to 2008 along with estimates from the three above models. As seen, there are substantial differences in the official data and the modeled net migration totals. For example, the overall population gain in Europe due to migration from outside Europe in 2008 was around 800,000 people according to the statistical models. Eurostat data, by contrast, lists a net migration gain of approximately 1.5 million people. Consistency is obtained in the MIMOSA, IMEM, and Abel models because the full origin–destination matrix of migration is modeled, rather than simply summing up each country's flows, independently of the others, as is recorded

Figure 2. Net migration flows for EU/EFTA countries



Note: IMEM (Integrated Modelling of European Migration) = predictive intervals (10%, median, 90%). The median/mean values from the statistical modeling results are all fairly similar and much less than the official Eurostat data.

Source: Raymer, J., A. Wiśniowski, J. J. Forster, P. W. F. Smith, and J. Bijak. "Integrated modeling of European migration." *Journal of the American Statistical Association* 108:503 (2013): 801–819 [1].

I Z A
World of Labor

in the Eurostat data. Thus, the advantages of combining migration data are reductions in the biases and inaccuracies inherent in the reported data.

LIMITATIONS AND GAPS

As discussed above, migration flow data suffer from serious inaccuracy and incompleteness. Research that compares different reports of international migration statistics could provide the basis for improving the consistency and quality of data. Comparing international reports could also improve understanding of existing data and, ultimately, the migration processes themselves. Ideally, this process would be conducted by an independent organization that receives data inputs directly from national statistical offices. However, there are many short-term obstacles to consider before this type of structure could be implemented. The exact procedures and methodologies for sharing migration data across countries would take time to organize and develop, though the current Nordic system does provide a working example.

In the absence of a functioning international system for sharing migration data, statistical modeling is really the only feasible option for harmonizing and estimating flows. Using such synthetic estimates as a means of understanding migration represents a radically different practice than what currently exists. However, recent research on European migration has shown the advantages of doing so [2], [11], namely a consistent and comparable set of flows that can be used to understand migration processes and to assess migration policies across countries. Considerable efforts will be needed to apply these tools to other regions in the world, where international migration flow data are even sparser. While many developed countries, such as the US and Australia, provide

somewhat reliable information on international migration flows, there is a real and urgent need to understand how people are moving amongst less developed countries in Africa, Asia, and Latin America—areas that are experiencing rapid economic development and urbanization—where nearly no reliable data exist.

SUMMARY AND POLICY ADVICE

Consistent migration flow data are needed for a range of reasons that are critical to the development of a reliable evidence base for migration policies and research. One of the main obstacles is the absence of a standard definition for measuring migration flows. There also needs to be communication and sharing of data between countries of origin and destination, and procedures to assess the reported migration flows. Problems associated with poorly measured migration data are not new and have been well documented; however, only recently have alternatives been proposed. The IMEM model for combining differently measured migration data provides a mechanism to overcome current migration data limitations and has the potential to lead to better migration policy and research.

To fully comprehend international migration, researchers and policymakers must first overcome the inadequacy of existing data. Having a reliable database on migration flows would provide a better understanding of the mechanisms driving migration patterns and population changes. To start with, countries should agree to send information about the number of migrants they receive from individual countries to the national statistical office in each different origin country. Second, more effort is needed to ensure that reported migration flows add up to the stocks of migrants measured by censuses. As population stocks are easier to measure consistently across countries, this could help verify the quality of the migration flow data. If countries are unable to collect migration flow data, then estimates from models that utilize information from other countries' reports and covariates could be considered. These two activities in combination with ongoing developments in statistical modeling of migration flows should greatly improve understanding of migration and its impacts on societies at large, thus enabling policymakers to devise better targeted and more effective migration policies.

Acknowledgments

The author thanks two anonymous referees and the IZA World of Labor editors for many helpful suggestions on earlier drafts. Previous work of the author together with Jakub Bijak, Joop de Beer, Jon Forster, Peter Smith, Rob van der Erf, and Arkadiusz Wiśniowski contains a larger number of background references for the material presented here and has been used intensively in all major parts of this article [1], [2], [10], [11].

Competing interests

The IZA World of Labor project is committed to the *IZA Guiding Principles of Research Integrity*. The author declares to have observed these principles.

REFERENCES

Further reading

Poulain, M., N. Perrin, and A. Singleton (eds). *THESIM: Towards Harmonised European Statistics on International Migration*. Louvain-la-Neuve: UCL Presses, 2006.

White, M. J. (ed.). *International Handbook of Migration and Population Distribution*. Dordrecht: Springer, 2016.

Key references

[1] Raymer, J., A. Wiśniowski, J. J. Forster, P. W. F. Smith, and J. Bijak. "Integrated modeling of European migration." *Journal of the American Statistical Association* 108:503 (2013): 801–819.

[2] Raymer, J., J. de Beer, and R. van der Erf. "Putting the pieces of the puzzle together: Age and sex-specific estimates of migration amongst countries in the EU/EFTA, 2002–2007." *European Journal of Population* 27:2 (2011): 185–215.

[3] UN. *Recommendations on Statistics of International Migration*. UN Department of Economic and Social Affairs, Statistics Division, Statistical Papers Series M, No. 58, Rev.1, 1998.

[4] Santo Tomas, P. A., L. H. Summers, and M. Clemens. "Migrants count: Five steps toward better migration data." Report of the Commission on International Migration Data for Development Research and Policy, Centre for Global Development, Washington DC, 2009.

[5] UN. *Problems of Migration Statistics*. UN Department of Social Affairs Population Studies Report No. 5, 1949.

[6] Kelly, J. J. "Improving the comparability of international migration statistics: Contributions by the Conference of European Statisticians from 1971 to date." *International Migration Review* (1987): 1017–1037.

[7] Bilsborrow, R. E., G. Hugo, A. S. Oberai, and H. Zlotnik. *International Migration Statistics: Guidelines for Improving Data Collection Systems*. Geneva: ILO, 1997.

[8] Nowok, B., D. Kupiszewska, and M. Poulain. "Statistics on international migration flows." In: Poulain, M., N. Perrin, and A. Singleton (eds). *THESIM: Towards Harmonised European Statistics on International Migration*. Louvain-la-Neuve: UCL Presses, 2006; pp. 203–231.

[9] Kupiszewska, D., and B. Nowok. "Comparability of statistics on international migration flows in the European Union." In: Raymer, J., and F. Willekens (eds). *International Migration in Europe: Data, Models and Estimates*. Chichester: Wiley, 2008; pp. 41–71.

[10] De Beer, J., J. Raymer, R. van der Erf, and L. van Wissen. "Overcoming the problems of inconsistent international migration data: A new method applied to flows in Europe." *European Journal of Population* 26 (2010): 459–481.

[11] Wiśniowski, A., J. J. Forster, P. W. F. Smith, J. Bijak, and J. Raymer. "Integrated modelling of age and sex patterns of European migration." *Journal of the Royal Statistical Society Series A (Statistics in Society)* 179:4 (2016): 1007–1024.

[12] Raymer, J. "Migration in Europe." In: White, M. J. (ed.). *International Handbook of Migration and Population Distribution*. Dordrecht: Springer, 2016; pp. 371–388.

[13] Abel, G. J. "Estimation of international migration flow tables in Europe." *Journal of the Royal Statistical Society: Series A* (*Statistics in Society*) 173:4 (2010): 797–825.

Online extras

The **full reference list** for this article is available from:

http://wol.iza.org/articles/measuring-flows-of-international-migration

View the **evidence map** for this article:

http://wol.iza.org/articles/measuring-flows-of-international-migration/map