

Seeing the Hidden Part of the Iceberg: Gauging the Real Dimension of International Migration

Adrian Oțoiu¹ | *Bucharest Academy of Economic Studies, Bucharest, Romania*

Emilia Țițan | *Bucharest Academy of Economic Studies, Bucharest, Romania*

Abstract

The reliability and comparability of international migration statistics belong to the most important statistical issues due to the importance of correct dimensioning of the migration flows and stocks for effective and timely design of effective policy measures. This paper presents an assessment of the migration statistics provided by Eurostat, reveals the most prominent discrepancies between stock and flow data, prepares a summary of vital issues affecting both quality and completeness of the migration data, and identifies certain solutions in order to improve data comparability, reliability and completeness. There is no one-size-fits-all solution, but an eclectic mix which extends the use of administrative and private data, matches data coming from distinct sources, harmonizes the way in which data is compiled and reported by different countries, matches observed flows with (demographic) stock-based estimates, provides consistent estimates of the bilateral migration flows between countries, and improves the measurement of temporary and illegal/undeclared migration.

Keywords

Official statistics methodology, international migration, migration measures, data harmonization, data dissemination

JEL code

F22, C81, C18, C46

INTRODUCTION

International migration is one of the most important social phenomena, which affects both the origin and destination countries, and gives birth to fairly large changes in their societal structure. An estimate by Vasileva (2011) shows that “6.5% of the EU population are foreigners and 9.4% are born abroad”, with about 2/3rds of foreigners born in a country outside the EU.

Besides the fact that it raises significant cultural changes triggered by the increase of the multicultural character of the destination countries, international migration has far-reaching economic implications. Immigration is acknowledged as a solution to population ageing, as most immigrants are of working

¹ Department of Statistics and Econometrics. Corresponding author, e-mail: otou.adrian@gmail.com.

age. Immigration is a solution to skills shortages in destination countries, caused by both massive retirements in certain industries, and by an increasing demand for health care and assistance services workers.

The other component of migration, emigration, casts more light on the importance of the phenomenon. Thus, while emigration can be detrimental to the originating jurisdiction, by aggravating shortages of professionals (e.g. health care professionals), it can also alleviate labour surpluses occurring in declining industries. For example, in developing countries, productivity advances in agriculture can lead to widespread unemployment that can be reduced by mass emigration.

Additionally, in multinational blocks such as the EU, migration, which usually refers to movements of people into and outside the Union, and movements within the EU defined as mobility, (Boswell and Geddes, 2010) can act as a means for a better redistribution of labor resources, and contribute to both an improved personal fulfillment of their citizens and to an increased competitiveness of the economy. Migration in this paper will also refer to cross-border mobility within the EU unless otherwise stated.

However, despite its importance and far-reaching economic and social implications, it is the consensus that international migration is not adequately measured in official statistics. This is due to many facts: the volatility of various forms of migration that is seasonal (and often undeclared), the methodology based on self-reporting of the migrants that can be either restrictive or not relevant in case of illegal migration, etc. This leads to significant misestimations of migrant flows, with direct results on the effectiveness of some policies that address successful integration of the migrants in the host societies, planning for the occupational deficits, adequate dimensioning of specific social services, etc.

Therefore, this paper attempts to explore the underestimation of migration flows, and seeks to identify the key directions to address it. The first deals with discrepancies between different migration data available from official statistics. Then the main reasons behind the differences are examined. The final part summarizes and suggests potential solutions for improving the quality of migration data.

1 WHERE ARE WE NOW? A CRITICAL LOOK AT THE DATA

The migration phenomenon cannot be properly measured in some countries. One of the examples is Romania, where the National Statistical Institute warns that administrative sources for external migration do not cover the entire phenomenon, especially the emigration flows from Romania. As such, there is a severe underestimate of the phenomenon which leads to an overstatement of the Romanian population (National Institute of Statistics Romania, 2014). A closer examination shows that, while the National Institute of Statistics' flows' data counts 129 000 officially declared net migrants between 1992 and 2002, census statistics show 697 thousand migrants after taking the natural decline of the population and the officially recorded net migration out of the total population decline (Ghețu, 2007).

The situation affects many European countries. While migration statistics revolve around migration flows and stocks of migrants within a country, there are often major discrepancies between the two, which cannot be satisfactorily explained by statistical adjustments or other related procedures.

A comparison between the (demographic-based) net migration plus statistical adjustment data reported in the Eurostat population tables, and the net international migration flows, computed as the difference between the immigration and emigration, shows that for the EU 27 during 2002–2012 period, the latter represents only 71% of the former. Further examination of the data for 29 European countries for which data covering more than five years were available, shown in Appendix 1, reveals that for 14 of them, stocks and flow measures differed by over 10%, even after removing the highest and the lowest annual discrepancies between the two migration measures. While for most of them flow data seems to cover only part of the migratory flows, in several cases (Italy, Hungary, Poland and Portugal), the stock measure appears to be much lower than the recorded flows, which raises a question of reliability of the flow measures.

The results obtained led to the question of how well statistics about the number of foreigners within a county match its statistics on immigration. To this end, the annual change in the stock of foreigners

based on the population statistics by country of birth was computed, and compared with immigration flows net of return migration, where available. Return migration consists of emigrants whose place of birth differs from the country of emigration, and immigrants born in the country of immigration.

Results from the Eurostat data from 2010 to 2012, show that the change in foreign-born population in the EU-27 is 2.5 times lower than annual immigration numbers. The annual change in the numbers of foreign-born reveals potential data issues since for Hungary and Slovenia, a decrease in foreign-born population is accompanied by an increase in immigration (see Appendix 2). The above, plus the fact that annual changes in the stocks of the foreign-born population are, as a rule, much lower than (net) immigration numbers, and that variations in the stocks of the foreign-born individuals are considerably larger than those observed in the immigration numbers, can point out to serious reliability problems.

Major discrepancies between countries can be found with respect to bilateral flows data. Thus, in the absence of any reported data quality issues and disclaimers, the EUROSTAT data on annual emigration flows from the UK to Spain sums up to 58 thousand for the 2009–2012 period, whereas immigration to Spain from the UK is 33% higher than that, amounting to 77 thousand.

2 THE NATURE OF DIFFERENCES

In order to improve understanding of differences, there is the need to inquire about their sources.

The main cause behind the lack of reliability of migration statistics is its coverage, which is incomplete due the methodology. Thus, a person is considered to be a permanent migrant if he/she changes its country of usual residence for a period of at least one year (EUROSTAT, 2013, following the UN definition). This is mainly based on the self-reporting of individuals, and excludes the number of persons that have not reported, or avoided reporting themselves as migrants due to the illegal nature of their move. The same applies to short-term migrants, which are considered as such if they change the place of usual residence for more than three months, but under a year.

The process of statistical data collection itself leads to fluctuations in the number of migrants within the country. Flows data may be subject to two-year delays as it is available only after a given calendar year has passed. To overcome this, some countries have chosen to report as migrants those who stayed for over three months (Beer et al., 2010). In other cases, immigration counts are based on the intended period of stay (Fassmann, 2009, cited by Beer et al., 2010).

Differences in country-specific methodologies also account for many differences in the data. While many countries impose a time limit for the intention to stay in another country except for temporary purposes (visiting relatives, medical treatments, etc.), some (e.g. Germany) consider the intention to migrate regardless of the duration of stay (Beer et al, 2010). In other cases (Lemaitre, 2005), countries which issue permanent residence permits tend to exclude non-holders from the statistics (e.g. international students enrolled in degree programmes in Canada and the US).

Transitions from temporary to permanent migration could also induce a bias in the migration flows data. While transitions can be straightforward for highly trained professionals that are legally employed (Lemaitre, 2005), the same cannot be said about short-term seasonal workers or undocumented migrants. In order to solve these issues and provide a clean break for the undocumented migrants, some countries have performed one-time regularizations for undocumented workers, e.g. Italy in 2002, and Spain in 2005 (Finotelli and Arango, 2011). A summary of findings by Finotelli and Arango (2011) report that, following regularization, immigration irregularity seems to have increased again.

Census under-coverage affects migration stock data, which originates from demographic data, which is, in turn, based on population census data. However, censuses are conducted based on administrative records, which can leave out migrants who were not recorded at the time of establishing the census sampling frame. Likewise, emigrants are likely to be included in the sample frame, even if they have left the country. These issues can be compounded at the time the data collection is carried out through

significant non-response rates. Clark (2009) shows that the permanent and temporary migrants have the lowest response rates for the 2006 Canadian census undercoverage survey.

In addition to undercoverage and non-response issues, stock-based methods can preserve the inertial character of demographic statistics, based on the past census estimates extrapolated to the future using fertility, mortality and migration assumptions. Where the first two components are fairly stable, the known volatility of the migration component (Hatton, 2010) may invalidate stock measures, even though they are useful in providing migration data that compensates for some undercoverage from the incomplete reporting of migration. The issues of temporal comparability, resulting from the impact that economic conditions have on migration, do affect the reliability of stock measures that are not able to factor in the observed volatility of migration flows (Bell et al., 2002).

Short-term movement of qualified workers is often not captured in the statistics. While EUROSTAT publishes data about movement of students, teachers, and Ph.D. holders, there is no regular data series that cover movements of some highly trained professionals such as medical doctors, IT specialists, etc. This data could significantly minimize the gaps between data reported by receiving and originating country, and would improve the measurement of labor migration (UNSTATS, 2004).

3 GETTING TO SEE MOST OF THE ICEBERG. SOLUTIONS TO IMPROVE INTERNATIONAL MIGRATION DATA

One way to improve the consistency and reliability of statistics is to increase international cooperation and harmonization (ILO, 1995). To this end, the example of UK and Spain is relevant; as both are developed countries, and migratory movements are established, it is not acceptable to see that 25% of the UK arrivals in Spain are not reported by the UK statistical office. If all bilateral flows were consistent, there could be significant gains in the reliability of the data.

Promoting an extended use of administrative sources can improve the reliability of the data. In addition to population records, tax and social security records can help assess de-facto movements to another jurisdiction, bar the illegal migrants who do not officially work, and some senior retirees that do not need to comply with the local requirements, as long as they do not exceed certain length-of-stay provisions. Other administrative sources, immigration data, border crossing data, travel records, establishment and labor force surveys, may provide data on foreign nationals who entered the country, and how long they stay.

The major caveat of this data is that it does not catch the entire migration phenomenon. Leaving aside compatibility of statistical sources, delays in data availability and processing, and lack of coverage for migrants, this data also may fail to cover internal movements within Schengen-like areas, for which administrative and border crossing controls are laxer.

Where public data may show its limitations, private data sources may be the key to complement the undercoverage of migratory movements. Transportation tickets issued by private companies may help document migrant routes that are less affluent and more prone to go illegal in the destination country (e.g. coach transportation records). (Travel) Insurance data may also tell about the international movements of policyholders. Credit bureau data, electronic payment data, and phone records may, too, provide information on the whereabouts of migrating individuals, and reveal their location relative to their social network. Remittance information can help identify anyone who regularly sends money to one's relatives working and living abroad, even if this person's property taxes are paid in his country of origin where remittance recipients live.

Private data usually contains information on the customers, such as national ID's passport numbers, date and place of birth, address, etc, which may help link it with administrative sources on one or more attributes.

Recruitment agencies can help record migration of skilled workers by providing additional information insights about their intended occupation, their level of education, the industry they will work in, along with age, gender and prospective salary (Hoffmann and Lawrence, 1994).

Application for residence, refugee, asylum and crime data are useful in recording undocumented migration. While other data discussed so far has a rather neutral stance with respect to the migrants by merely recording facts concerning particular persons, this data focuses on migration-specific intentions and foreign residence and stay issues. This type of data may be useful in covering at least partially the recent massive inflows of refugees from Africa and Middle East, coverage that should increase within the next months after migrants will come in contact with immigration and law enforcement authorities.

Finding ways to account for the transitions from temporary to permanent migration should be another priority for improving the quality of statistics, which partly addresses the issues of measuring irregular migration, avoiding overestimation of migration for countries that report migration over three months as permanent migration, avoiding double counting of migrants, and improve the estimates for both types of migration.

Incorporating information of census undercoverage and/or overcoverage data, and on response rates, into demographic projections, can improve estimates of the migrant stocks and bring them closer to reality. In this respect, other survey frames, especially those focusing on migrants, can be useful through comparisons with the census frames.

Apart from solutions that attempt to harmonize methodologies across countries, calculation methods are useful quantitative tools applied for improvement of migration data consistency and comparability.

Some of the most common methods are based on minimized sum of squares of the bilateral matching flows, similar to the one proposed by Beer et al., (2010).

Other methods propose to derive migration flows from migrant stock data (Abel, 2013). While this method has its benefits, we consider that its merits merely lie in matching stable but rather inertial demographic-based with the volatile and undercoverage-affected flows data, rather than in correctly estimating migration from stock estimates.

Another method proposed by Raymer et al. (2013) attempts to calculate migration data by factoring in country differences in migration reporting, estimates of the undercount of the flows, expert-based prior distributions, and variables showing impact on the levels of migration: population, gross national income per capita, migrant stocks, along with indicators of cultural affinity, labor market and immigration openness, contiguity, etc. This model is one of the first attempts to incorporate several types of information about the caveats of the migration data into a complex model.

DISCUSSION, CONCLUSIONS AND PERSPECTIVES

The lack of reliability and comparability affecting international statistics of migration is not new and has attracted significant efforts with the aim to improve them.

The analysis of the caveats of migration measures and the solutions proposed to improve them clearly shows that there is no unique strategy to yield a satisfactory solution of migration flows. Rather, a multi-pillar strategy that achieves a synergy between the suggested solutions could bring the migration statistics in line with other demographic statistics and allow preparing reliable estimates and analyses to inform public policy makers and other researchers.

A first major direction is the improvement in using and matching several data sources to produce migration data. While many researchers favor the use of migration stocks, their reliance on rather static demographic estimates, which leave out almost entirely the fluctuations of the economic conditions that drive most of the international migration, may be unjustified. An optimal solution should consider the advantages and disadvantages of both data.

Another major direction should involve harmonization of data coming from different national sources. Harmonization must take into account the national characteristics of international migration through the establishment of common or comparable definitions of migration, to be used in the harmonized data.

Computation of migration data must appropriately take into account its known caveats (undercoverage, reliability, etc.) and if necessary, make adjustments to bring it in line with what should be their closest-to-truth values.

Aggregation across several data sources could overcome the caveats of administrative data and specific migration data (e.g. holders of permanent resident permits, temporary visa holders, etc.), and provide an indirect assessment of someone's migration status as reflected by his or her actions rather than declarations of intent and/or official records in population registers.

Aggregation can prove useful in the context of renouncing of border controls and free movements of people between countries, and especially in mobility cases involving cross-border and regional commuting, which can go unassessed. Student records and employment information can document such short-term movements.

And, last but not least, the use of specific statistics can assist in improving the coverage of certain groups of migrants. For example, UK health and pension statistics can improve coverage of migrant pensioners and women (Evans et al., 2007).

A proper distinction between temporary and permanent migration and mobility should be a priority in improvement of international migration data. This distinction is increasingly required by various aspects of mobility (e.g. cross-border and regional commuting, short-term student exchange programmes), which do not involve formal changes of residency, and by the ease of transition between the two types which can lead to unreliable data. An increased ability to gauge temporary migration and mobility can lead to an improved ability to estimate illegal migration.

All the above solutions should be complemented by the use of statistical methods. Some statistical techniques are needed in order to harmonize stocks and flows data, on one hand, and adjust bilateral migration flows on the other hand, so that they match one another.

And, last but not least, given the complexity of the issues pertaining to international migration, it may be appropriate to generate a set of comparable, harmonized migration data, which should not replace, but merely offer an alternative to the existing migration data. The latter may be better suited to the economic and social reality of the country that disseminates it and thus be more effective in the appraisal and policy-making of that country. This would also increase the acceptance of the new data, which, given the actual state of affairs and current practices, may take a while until it provides the much-needed gauge of international migration.

As a final remark, while solutions presented in this paper can constitute valid directions for improving international migration statistics, it should be acknowledged that there is still a long way to go until they will be implemented by the National Statistical Offices and be formalized within the frame of European Statistical System. There is definitely need for further research on how to transform these directions into actionable solutions.

ACKNOWLEDGEMENT

This work was co-financed from the European Social Fund through Sectoral Operational Programme Human Resources Development 2007–2013, project number POSDRU/159/1.5/S/134197 „Performance and excellence in doctoral and postdoctoral research in Romanian economics science domain“.

References

-
- ABEL, G. J. Estimating global migration flow tables using place of birth data. *Demographic Research*, 2013, 28(March), pp. 505–546.
- BEER, J. DE, RAYMER, J., ERF, R. VAN DER, WISSEN, L. VAN. Overcoming the Problems of Inconsistent International Migration data: A New Method Applied to Flows in Europe. *Eur J Popul.* 2010, Nov. 26(4), pp. 459–481.
- BOSWELL, C., GEDDES A. *Migration and Mobility in the European Union*. Palgrave Macmillan, 2010.

- BELL, A. M., BLAKE, M., BOYLE, P., REES, P., STILLWELL, J., HUGO, G., BLAKE, M. Cross-national comparison of internal migration: issues and measures. *J. R. Statist. Soc. A*, 2002, Vol. 65, Part 3, pp. 435–464.
- CLARK, C. *Census Population Coverage Error: Results and Methods, Federal Committee on Statistical Methodology Research Conference*. Washington D.C., 2009.
- EUROSTAT. *International migration flows (migr flow) Reference Metadata in Euro SDMX Metadata Structure (ESMS)* [online]. 2013. [cit. 8.7.2014] <http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/EN/migr_flow_esms.htm>.
- EVANS, H., VICKERS, L., WRIGHT, E. Using administrative data sources in the estimation of emigration [online]. *Population Trends*, 2007 (128), pp. 33–40. [cit. 7.7.2014] <<http://www.ncbi.nlm.nih.gov/pubmed/17691539>>.
- FINOTELLI, C., ARANGO, J. Regularisation of unauthorised immigrants in Italy and Spain: determinants and effects. *Documents d'Anàlisi Geogràfica*, 2011, 57, pp. 495–515.
- GHEȚĂU, V. Declinul demografic și viitorul populației României. *Editura ALPHA MDN*, Buzău, 2007.
- HATTON, T. J. The Cliometrics of International Migration: a Survey. *Journal of Economic Surveys*, 2010, 24(5), pp. 941–969.
- HOFFMANN E., LAWRENCE, S. *Statistics on international labour migration. A review of sources and methodological issues*. ILO Geneva, Bureau of statistics, Interdepartmental Project on Migrant Workers, 1994.
- LEMAITRE, B. G. *The Comparability of International Migration Statistics*. OECD Statistics Brief No. 9, 2005.
- NATIONAL INSTITUTE OF STATISTICS – ROMANIA. POP309A – *Emigrants by sex, macroregions, development regions and counties of departure* [online]. 2014. [cit. 8.7.2014] <<https://statistici.insse.ro/shop/index.jsp?page=tem-po3&lang=en&ind=POP309A>>.
- RAYMER, J., WIŚNIEWSKI, A., FORSTER, J. J., SMITH, P. W. F., BIJAK, J. Integrated Modeling of European Migration. *Journal of the American Statistical Association*, 2013, 108(503), pp. 801–819.
- UNSTATS. *Improving the Compilation of International Migration Statistics. Results of a United Nations Workshop*. 2004.
- VASILEVA, K. 6.5 of the EU population are foreigners and 9.4 are born abroad. *Statistics in focus*, 34/2011, Eurostat.

APPENDIX 1

Coverage of migration stocks by annual flows data, 2002–2012

| Country | Belgium | Czech Rep. | Denmark | Germany | Estonia | Ireland | Spain | France | Croatia | Italy | Cyprus | Latvia | Lithuania | Luxembourg | Hungary |
|--------------|---------|------------|---------|---------|---------|---------|-------|--------|---------|-------|--------|--------|-----------|------------|---------|
| Coverage (%) | 99 | 61 | 87 | 96 | 53 | 100 | 63 | 100 | 88 | 131 | 106 | 100 | 100 | 100 | 120 |

| Country | Malta | Netherlands | Austria | Poland | Portugal | Slovenia | Slovakia | Finland | Sweden | UK | Iceland | Norway | Switzerland |
|--------------|-------|-------------|---------|--------|----------|----------|----------|---------|--------|----|---------|--------|-------------|
| Coverage (%) | 88 | 51 | 86 | 783 | 147 | 100 | -393 | 99 | 100 | 74 | 80 | 101 | 103 |

Note: Coverage is computed as emigration less immigration flow data, divided by the annual change in the population due to net migration plus statistical adjustment. Calculations cover the period 2002–2012 remove the highest and the lowest values, and are based on at least 6 annual data points.

Source: Authors' calculations based on Eurostat data <<http://ec.europa.eu/eurostat>>

APPENDIX 2

Coverage of changes in foreign-born population by net immigration, 2010–2012

| Country | Belgium | Denmark | Estonia | Ireland | Spain | Italy | Lithuania | Luxembourg | Hungary |
|--------------|---------|---------|---------|---------|-------|-------|-----------|------------|---------|
| Coverage (%) | 120 | 147 | 121 | 253 | 1 123 | 89 | 554 | 83 | -157 |

| Country | Netherlands | Austria | Slovenia | Finland | Sweden | Iceland | Liechtenstein | Norway |
|--------------|-------------|---------|----------|---------|--------|---------|---------------|--------|
| Coverage (%) | 129 | 277 | -179 | 78 | 120 | 70 | 195 | 128 |

Note: Coverage is computed as net immigration divided by the annual change in the foreign-born population for the entire 2010–2012 period.
Source: Authors' calculations based on Eurostat data <<http://ec.europa.eu/eurostat>>