

# Bilateral International Migrant Stocks: An Inquiry of Measurement Discrepancies (1960-2020)

Cris Beauchemin<sup>1</sup>, Arnaud Bringé<sup>1</sup>, Bruno Schoumaker<sup>2</sup>

**EPC – Bologna 2026**

## Short abstract (max 250 words)

At the turn of the millennium that several international organizations (the World Bank, the United Nations, and the Organization for Economic Cooperation and Development) published new bilateral migration databases that made it possible to measure not only international *immigrant* stocks in all destination countries of the world, but also *emigrants* stocks of all origin countries. Accessible in open access, these databases became major sources of academic and institutional publications. They are used extensively for descriptive purposes at the global and national levels, for estimating international migration flows, or for studies on the determinants and consequences of migration. All these databases are based on the same principle: using mainly census data, they are migration matrices that cross all origin and destination countries in the world. They should therefore show similar measures of international migrant stocks. However, we discovered that estimates of bilateral stocks (ie. number of migrants from country A living in country B) are quite frequently radically different from a source to another. In this paper, our objectives are following: (1) to map the dozen existing databases, examining their conceptual and technical underpinnings, (2) to compare the bilateral migrant stocks provided by the diverse databases in order to systematically identify outliers, (3) to model the factors that explain them, and (4) to assess the impact of inconsistent estimates on the measurement of international migration across the world.

## Context

It was only at the turn of the millennium that several international organizations (the World Bank, the United Nations, and the Organization for Economic Cooperation and Development) published new bilateral migration databases that made it possible to measure not only international *immigrant* stocks in all destination countries of the world, but also *emigrants* stocks of all origin countries. These new databases truly revolutionized international migration studies. Accessible in open access, they became major sources of academic and institutional publications. They are used extensively, by international organizations and researchers, for descriptive purposes at the global (World Bank 2023; UN-DESA 2024; IOM 2024; Czaika et de Haas 2014; Özden et al. 2011) and national levels (for instance, in IOM migration profiles<sup>3</sup>), for estimating international migration flows (Abel et Cohen

---

<sup>1</sup> Institut national d'études démographiques (INED), F-93300 Aubervilliers, France.

<sup>2</sup> UCLouvain. Université catholique de Louvain. Centre de recherche en démographie, 1-1 Place Montesquieu, L2.08.03, 1348 Louvain-la-Neuve - Belgium

<sup>3</sup> See the IOM National and regional migration profile repository: <https://gmdac.iom.int/national-and-regional-migration-profile-repository>

2019), or for studies on the determinants and consequences of migration (Beine and Parsons 2015; Clemens 2020; Benček and Schneiderheinze 2024; Docquier et al. 2014; Beine et al. 2013).

All these databases, called “**Bilateral international migrant stocks databases (BIMS-DB)**” in our project, are based on the same principle: they combine, at the world level, variables on the origin and destination of international migrants registered all over the world. The destination is the country of residence where demographic information is collected. And the origin is known from available information on the country of birth or nationality in demographic sources. In other terms, these BIMS-DB are migration matrices that cross all origin and destination countries in the world, providing the number of lifetime migrants originating from a given country in each of the other countries in the world. For the most part, these databases rely on the same primary sources (essentially country censuses) and should therefore show similar measures of international migrant stocks. However, in the course of our previous research, accidental checks showed us that estimates of bilateral stocks (ie. number of migrants from country A living in country B) are sometimes radically different.

## Objectives

Considering the extraordinary impact of these databases on research and information to policymakers, it is crucial to measure the extent of inconsistencies and to understand them in order to improve data quality in matter of international migration. This is the objective of this paper, in which we analyze systematically discrepancies between international databases that provide bilateral estimates of migrant stocks. The specific objectives are the following: (1) to map the dozen existing BIMS-DB, examining their conceptual and technical underpinnings, (2) to compare the bilateral migrant stocks provided by the diverse BIMS-DB in order to systematically identify outliers, (3) to model the factors that explain them, and (4) to assess the impact of inconsistent estimates on the measurement of international migration across the world.

## A mapping of existing Bilateral international migrant stocks databases (BIMS-DB)

We restrict the project to the dozen databases that were available online in 2023 and 2025. Altogether, they cover the 1960-2020 period. While some databases provide disaggregated estimates by gender, age or education level<sup>4</sup>, we restrict our analyses to the global migrant stocks, without any sort of disaggregation by personal characteristics.

In our paper, we will provide a synthetic comparison of all BIMS-DB characteristics, based on an analysis of (1) their conceptual choices (eg. definitions of international migrants, of countries' delimitations), and (2) methodological foundations (estimation techniques), as a first step to theoretically understand the observed differences between sources. Box 1 summarizes the possible theoretical explanations for the observed differences between the BIMS-DB

### *Box 1. Hypotheses regarding the reasons to explain BIMS-DB discrepancies*

- **Migrant definition.** For a same destination country, the migrant definition may vary across BIMS-DB (birth country vs. citizenship; inclusion or not of refugees).
- **Primary data sources** may differ by BIMS-DB (censuses or others), some BIMS-DB rely on previously published BIMS-DB.
- **Geographical coverage** may differ across BIMS-DB: overseas territories are identified as specific entities in some BIMS-DB; the methodology to treat territorial recompositions (e.g., break out of the USSR, decolonization...) varies by data sources.
- **Imputation techniques.** Missing information is a common issue for BIMS-DB either because some destination countries

<sup>4</sup> For instance, the DIOC databases, produced by OECD, and the companion database of the 2023 World Development Report, produced by the World Bank, provide estimates by gender, age and education level. The UN International migrant stock database provides separate estimates for men and women.

simply lack sources counting immigrants, or because national statistical offices (NSOs) refuse to provide detailed information on migrants' origin. Imputation techniques may vary by BIMS-DB.

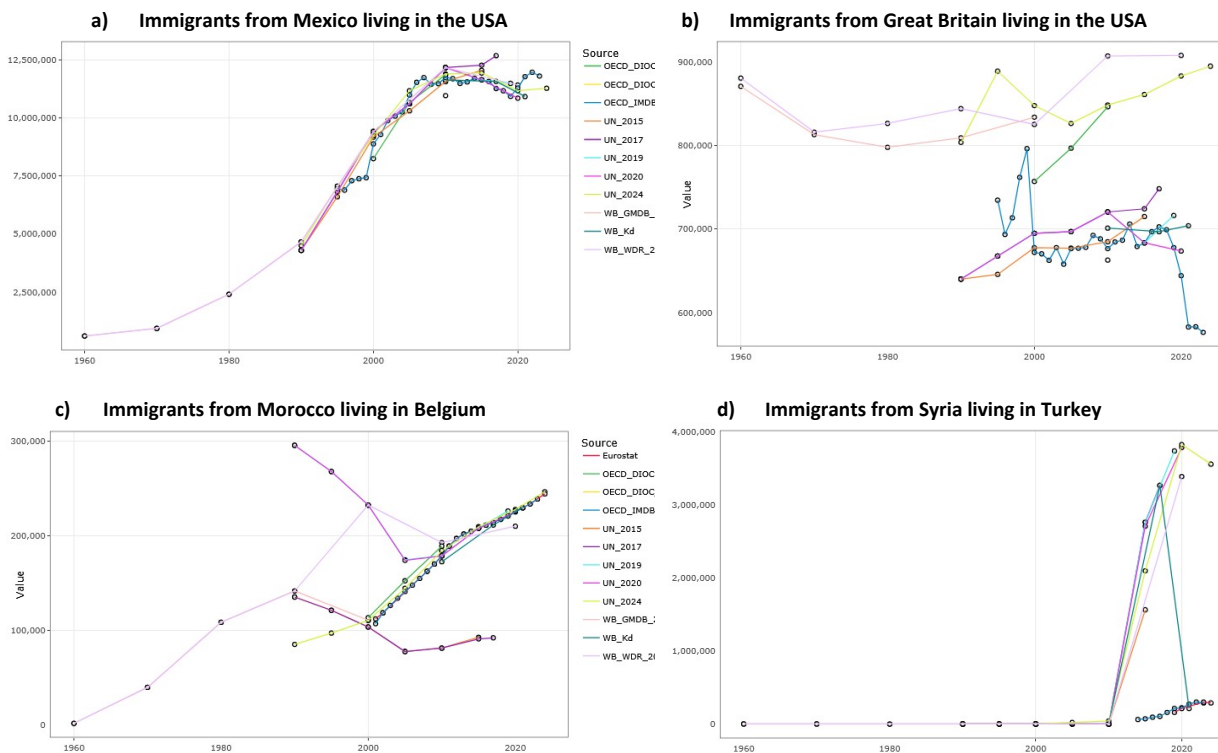
- **Statistical adjustments to round dates** (1960, 1970, 1980...) are computed or not.
- **Data collection process.** Assembling data from all countries in the world requires the contribution of many people (including officers from each country NSO), leaving room for potential human mistakes.

## A systematic comparison of existing Bilateral international migrant stocks databases (BIMS-DB)

To compare all existing BIMS-DB, we started to aggregate their estimates in a single database, the so-called ABIMS-DB (Aggregated Base of International Migrant Stocks Database). We developed a prototype of R-Shiny application that allows visualizing migrant stocks for all origin-destination couples in the world, with 3 features: (a) as in Figure 1, it shows estimates of all BIMS-DB for all possible couples; (b) it will include statistical tools designed for systematic identification of outliers; (c) it will provide contextual information on the content of each BIMS-DB at the origin-destination couple level in relation to the possible explanations of discrepancies (see Box 1).

Figure 1 provides some examples of significant origin-destination couples. The case of Mexican migrants in the USA shows consistent trends, although one-off estimates may vary significantly (e.g., by more than 2 million migrants in 2010). The other figures show much more erratic trends with strikingly diverse estimates in different contexts, including in countries that are supposed to provide high-quality data. According to our preliminary analyses, these examples are far from being exceptional. Our objective is twofold: (1) identify outliers systematically, and (2) measure their extent.

**Figure 1. Immigrants number according to BIMSD-DB (1960-2023)**



Note: Each curve represents estimates from a specific database; a same producer may have published several databases (eg. the United Nations provides regular revisions since 2015). Eurostat does not produce a migration matrix but provides us with benchmark data for EU members and a selection of countries.

## Modeling the outliers

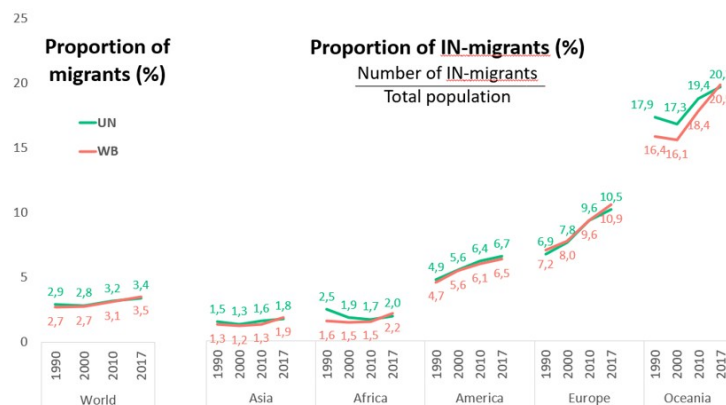
Once outliers are identified, our objective is to explain them. Our methodology is the following.

Each estimate of the migrant stock of a given origin in a given destination at a precise date will be considered as an observation. Among them some are tagged as outliers thanks to our analyses at the previous step. Our objective is to study the factors explaining that an observation is considered as an outlier (our dependent variable). To do so, we build a database containing information that relate to the hypotheses presented in Box 1 (our set of independent variables). Our preliminary analyses suggest that the observed discrepancies are largely explained by the migrant stocks size: outliers tend to be much more frequent where migrant stocks of a specific origin are small.

## Assessing the impact of inconsistent estimates on the measurement of international migration

The different bilateral migrant stock databases are commonly used to generate global, regional and national estimates of international immigration and outmigration rates. In this paper, we assess to what extent these measures vary depending on the data source. Our first results show that inconsistencies between the databases vary from year to year and that their scale depends on the level of observation. They are relatively small on a global scale (figure 2). For example, in 1990, the United Nations database estimated that international migrants represented 2.9 percent of the world population, compared with 2.7 percent for the World Bank (153 million vs. 142 million). The gap widens at the regional level. At the same date (1990), in Africa, the estimated immigrant population varied from almost 10 million according to the World Bank up to 16 million according to the United Nations. Even greater differences can be observed at the national level.

**Figure 2. Proportion of migrants by data source (1990-2017)**



## References

- Abel, Guy J., at Joel E. Cohen. 2019. « Bilateral International Migration Flow Estimates for 200 Countries ». *Scientific Data* 6 (1): 1. <https://doi.org/10.1038/s41597-019-0089-3>.
- Beine, Michel, Frédéric Docquier, et Maurice Schiff. 2013. « International Migration, Transfer of Norms and Home Country Fertility ». *Canadian Journal of Economics/Revue Canadienne d'économie* 46 (4): 1406-30. <https://doi.org/10.1111/caje.12062>.
- Beine, Michel, et Christopher Parsons. 2015. « Climatic Factors as Determinants of International Migration ». *The Scandinavian Journal of Economics* 117 (2): 723-67. <https://doi.org/10.1111/sjoe.12098>.

- Benček, David, et Claas Schneiderheinze. 2024. « Higher economic growth in poor countries, lower migration flows to the OECD – Revisiting the migration hump with panel data ». *World Development* 182 (octobre): 106655. <https://doi.org/10.1016/j.worlddev.2024.106655>.
- Clemens, Michael Andrew. 2020. « The Emigration Life Cycle: How Development Shapes Emigration from Poor Countries ». *SSRN Electronic Journal*, publication en ligne anticipée. <https://doi.org/10.2139/ssrn.3679020>.
- Czaika, Mathias, et Hein de Haas. 2014. « The Globalization of Migration: Has the World Become More Migratory? » *International Migration Review* 48 (2): 283-323. <https://doi.org/10.1111/imre.12095>.
- Docquier, Frédéric, Giovanni Peri, et Ilse Ruysen. 2014. « The Cross-Country Determinants of Potential and Actual Migration<sup>1</sup> ». *International Migration Review* 48 (1\_suppl): 37-99. <https://doi.org/10.1111/imre.12137>.
- IOM. 2024. *World Migration Report 2024*. United Nations.
- Özden, Çağlar, Christopher R. Parsons, Maurice Schiff, et Terrie L. Walmsley. 2011. « Where on Earth is Everybody? The Evolution of Global Bilateral Migration 1960–2000 ». *The World Bank Economic Review* 25 (1): 12-56.
- UN-DESA. 2024. *International Migration and Sustainable Development*. United Nations. <https://doi.org/10.18356/9789213588024>.
- World Bank. 2023. *World Development Report 2023: Migrants, Refugees, and Societies*. World Bank.