

**Technical Meeting on Indicators and Operationalisation of the
Conceptual Framework of International Migration Statistics**

**Migration data quality into
question**

Michel POULAIN



TALLINNA ÜLIKOOL
TALLINN UNIVERSITY

Eesti Demograafia Instituut
Estonian Institute for Population Studies

UCL

Université
catholique
de Louvain



MAIN OBJECTIVE OF STATISTICAL DATA COLLECTION

Providing **RELIABLE** data for **policy development**
both at national and international levels
and more specifically for migration-related policies

Why migration data are the weakest among all demographic data ? (1)

Census and survey are more appropriate to describe migrants and population stocks with foreign background.

To better capture international migration flows, the use of administrative databases is recommended.

Nevertheless, producing statistics on the basis of administrative databases is far from the traditional work of statisticians.

Why migration data are the weakest among all demographic data ? (2)

Migration as an demographic event is **less important for life** compared to births and deaths.

Moving from a country to another becomes **common** and perceived as a temporary change.

It is identified and thereafter counted based on

- self-reporting declaration to administration
- answers given to questions in census and surveys.

Why migration data are the weakest among all demographic data ? (3)

Existence of major **advantages** or **disadvantages** linked to the declaration of a change of residence

- for Nationals leaving the country or returning
- for Foreigners immigrating or returning to their home country.

Why migration data are the weakest among all demographic data ? (4)

Immigration is better captured compared to **emigration**

- as the first concerns persons in the observation landscape
- whereas the second exclude them.

Why migration data are the weakest among all demographic data ? (5)

International migration flows and foreign population stock or diaspora are statistics produced by **two different countries**

The use of the so-called '**mirror**' statistics is challenging for several reasons. Among these :

- Some 'receiving' countries do not produce any data or some data without disaggregation by country of origin, country of citizenship and country of birth.
- Some methodological problems linked to country of citizenship (holders of multiple citizenships) and country of birth (change in territory or national borders)

Why migration data are the weakest among all demographic data ? (6)

Some disturbances may emerge when considering the **definition of the usual place of residence** in census recommendations to identify international migrants:

- The double choice for the time criterion:

Either the 12 months' criterion or the 'most of the time during 12 months' which means at least 6 months

- The intention to stay at least 6 or 12 months or the observed duration or presence or absence

To be **reliable**, any statistical data should be

1. Complete
2. Accurate
3. Comparable
4. Timely and regularly produced
5. Accompanied by detailed metadata
6. Largely disseminated and interpreted
7. Accessible for further analysis and research investigations

1. COMPLETENESS

It means covering **exhaustively** the population under concern, either the total population or a well-defined group.

Data related to the whole population may be produced by **adding different sub-populations** (e.g. nationals, foreigners with permanent residence permit, foreigners with temporary residence permit, asylum seekers and other internationally displaced persons).

When adding aggregated data from different population groups, the existence of **duplicate or specific gaps should be carefully checked.**

2. ACCURACY

The sources of **errors** are multiple

- Double records
- Missing records
- No answer
- False answers
- Imputation errors.

Systematic and regular checks are needed. Confrontation with other data sources and cross checking are helpful.

A sounded estimate should be preferred to a precise false figure

3. COMPARABILITY

- On **national level** – comparability between various existing data sources. Statisticians, administrations and scientific researchers may use different definitions to produce statistics that not comparable but using the same terminology.
- On **international level** – How NSO's apply international recommendations ? Is it possible to assess the difference between statistical figures produced on the basis of national rules and other ones following the international recommendations? What mirror data from receiving countries may tell us and what might be the reasons for large differences (different definitions, coverage, accuracy...).

4. Timeliness and regular production

- The timing of the data production and dissemination is a crucial point as far as **international migration flows** are concerned as we are facing quickly changing situations for which a rapid action is needed.
- Census and survey results may take more time to be disseminated as they mostly concern **stocks and characteristics of migrants** and these figures do not change so rapidly.
- The **regularity** of the data production is essential to follow the temporal evolution. For any data that is supporting the development of a policy, similar figures are needed before, during and after the action taken to implement the policy and also for **projecting in future**.

5. METADATA

- **Metadata** are traditionally recommended by international institution and regularly produced by NSO.
- Specific metadata should be available for a better use of administrative data source. **Sharing experiences** and supporting metadata between countries is largely welcome.
- Metadata should also be added to support **understanding** data and trends.

6. DISSEMINATION AND INTERPRETATION SUPPORT

- In addition to a large dissemination, the main features emerging from new released figures should be worded out (for media and public) to avoid **misunderstanding** that often appear if looking only numbers. For policymakers and involved stakeholders, **specific seminar and workshop** might be needed to reach the objective of an effective policy support.
- Any type of dissemination of statistical figures will contribute to **improve data quality** and help to adapt the future data collection in closer agreement with policy needs.

7. ACCESSIBILITY FOR FURTHER INVESTIGATIONS

The use of individual data is essential to conduct more specific analysis. The possibility of **linkage using IDcode** or a specific identifier allowing a strict bi-univocal linkage between different databases.

TO IMPROVE DATA QUALITY AND USEFULNESS

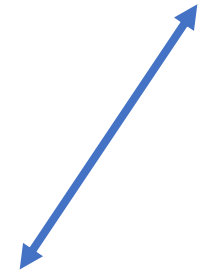
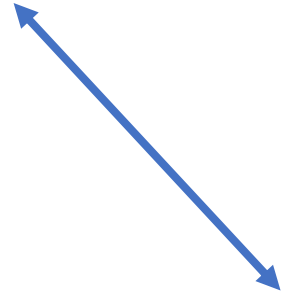
WE NEED

1. To convince stakeholders of the urgent need of statistics
2. To strengthen the role and action of the NSO
3. To cooperate closely with other institutions to optimize the transfer of data
4. To identify policy priorities and key points to be enlightened by an efficient statistical support in close cooperation
5. To force access to individual data from administrative databases solving problems of privacy, national security and data ownership
6. To consider attentively all feedback from users for further improvement
7. To focus on 'hard to count' populations and especially on emigrants and national living abroad.

DATA NEEDED



QUALITY



VALUE