Project Meeting on Measuring Human Mobility
27 - 29 March 2019, Tbilisi, Georgia

UNECE Guidance on data integration for measuring migration

Paolo Valente, UNECE

STATISTICS

Background and process

- ## 2014 UNECE-Eurostat Work session on migration statistics
  - Discussion on data integration for migration
  - Need to promote methodological work

- ## 2015 UNECE Task force created by CES Bureau
  - Members: Spain (chair Antonio Argueso, INE), Austria, Canada, Italy, New Zealand, Switzerland, UK, USA, UNECE, Eurostat

- ## 2016-2018 Task force developed the document
  - September 2016: Survey on national practices
  - October 2017: Draft version presented at Work session
  - April 2018: Electronic consultation with CES members
  - June 2018: Endorsement by CES

- ## February 2019: Publication on UNECE website:
  https://www.unece.org/index.php?id=51143

Purpose and content of the publication

- Purpose: To provide a general overview of data integration for measuring migration, and of the approaches adopted in various countries

- Content:
  - Definition of "data integration" in the context of migration
  - Review of country experiences on the basis of a survey covering over 50 countries
  - Detailed case studies for 13 countries
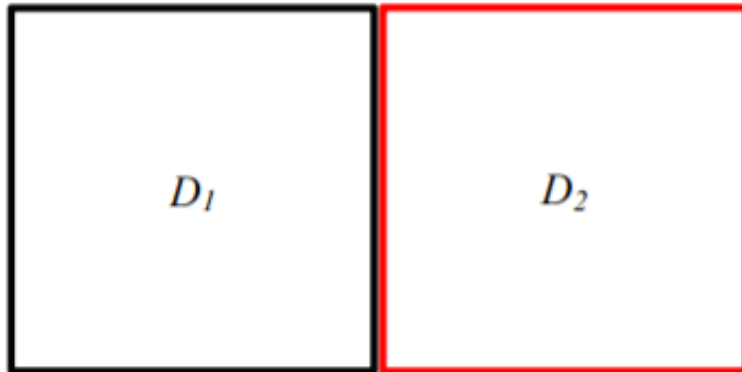 - Metadata
  - Recommendations
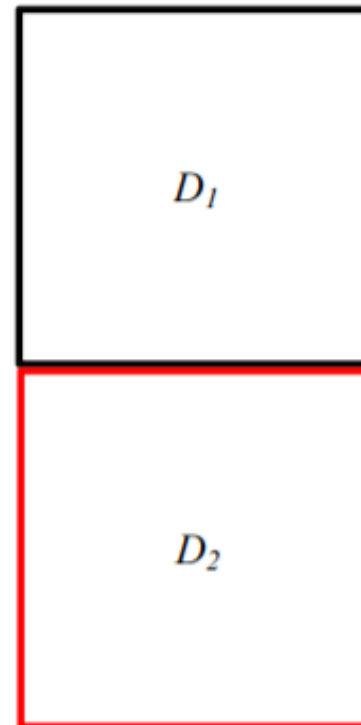
UNECE

Working definition of "data integration"

- "A statistical activity on two or more datasets resulting in a single enlarged and/or higher quality dataset"

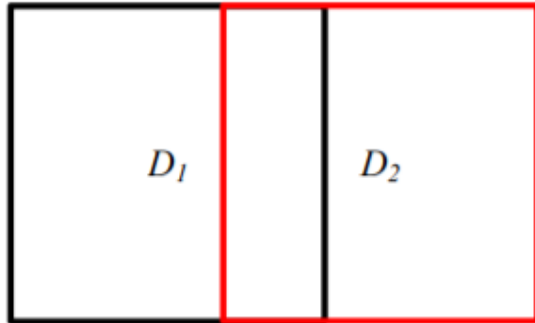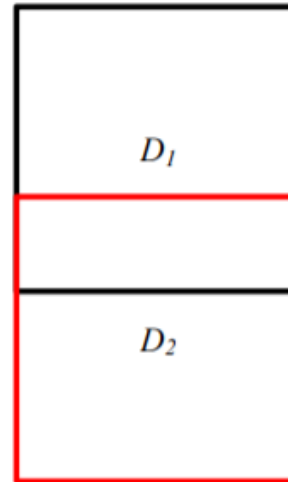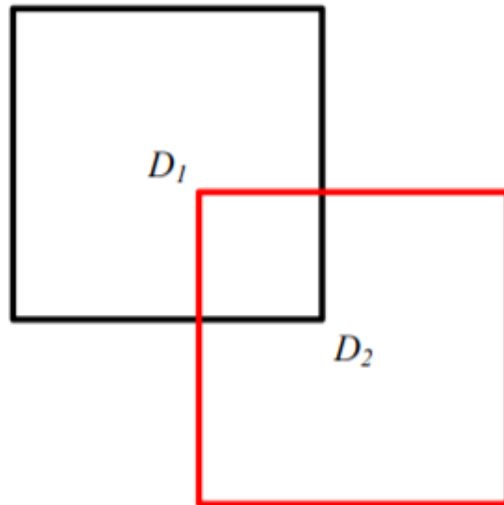Case 2: same variables, different units

Case 1: same units, different variables

D_1 D_2

D_1

D_2

Case 3: same units, partially overlapping variables

D_1 D_2

Case 4: same variables, partially overlapping units

D_1

D_2

Case 5: partially overlapping units and variables

D_1

D_2

Other cases: calibration to external sources, mirror statistics
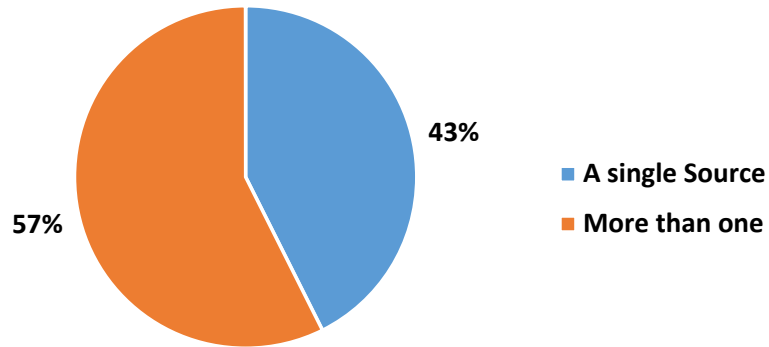
D1
(Survey)

D2
(Pop.
Register,
external
source…)
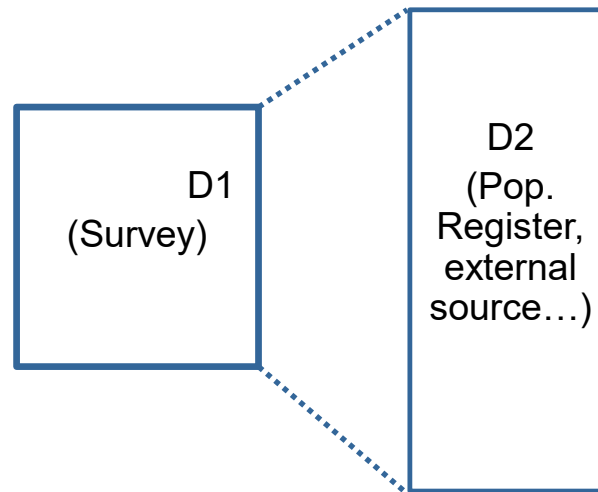
Survey on data integration practices

Data on international flows based on more than one source?

43%

57%

- A single Source
- More than one

How are these sources integrated?

8

11

2

10

- Macro-level (combining aggregated data)
- Micro-level (combining individual data)
- Both macro and micro
- Not combined
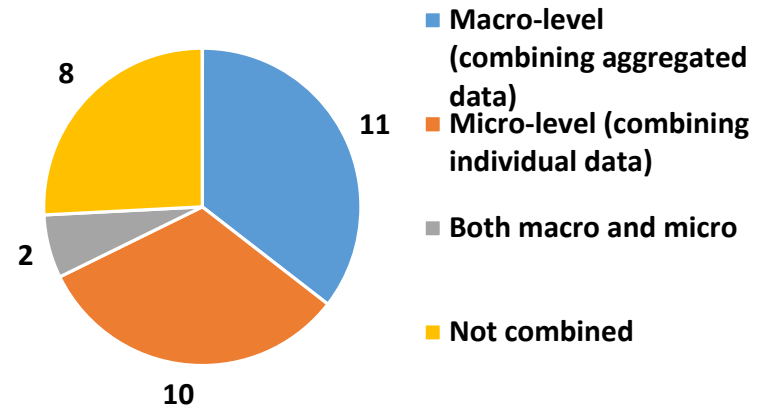
Data on foreign population or population born abroad (stocks)

46%

54%

- A single Source
- More than one

How are these sources integrated?

8

7

1

12
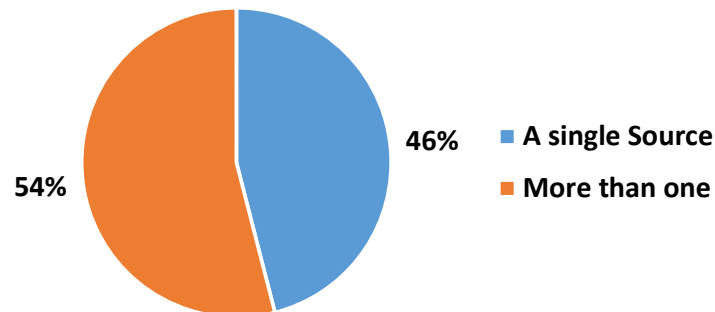
- Macro-level (combining aggregated data)
- Micro-level (combining individual data)
- Both macro and micro
- Not combined
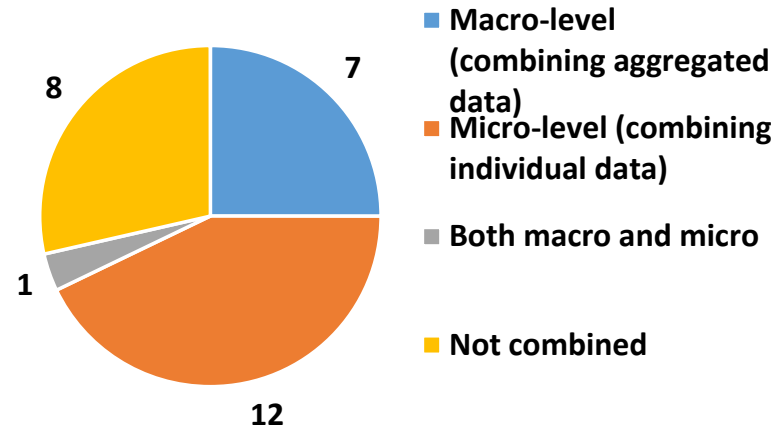
Case studies – main outcomes (1/3)

STATISTICS

- In countries with a population register, this is often (but not always) the main source of migration data
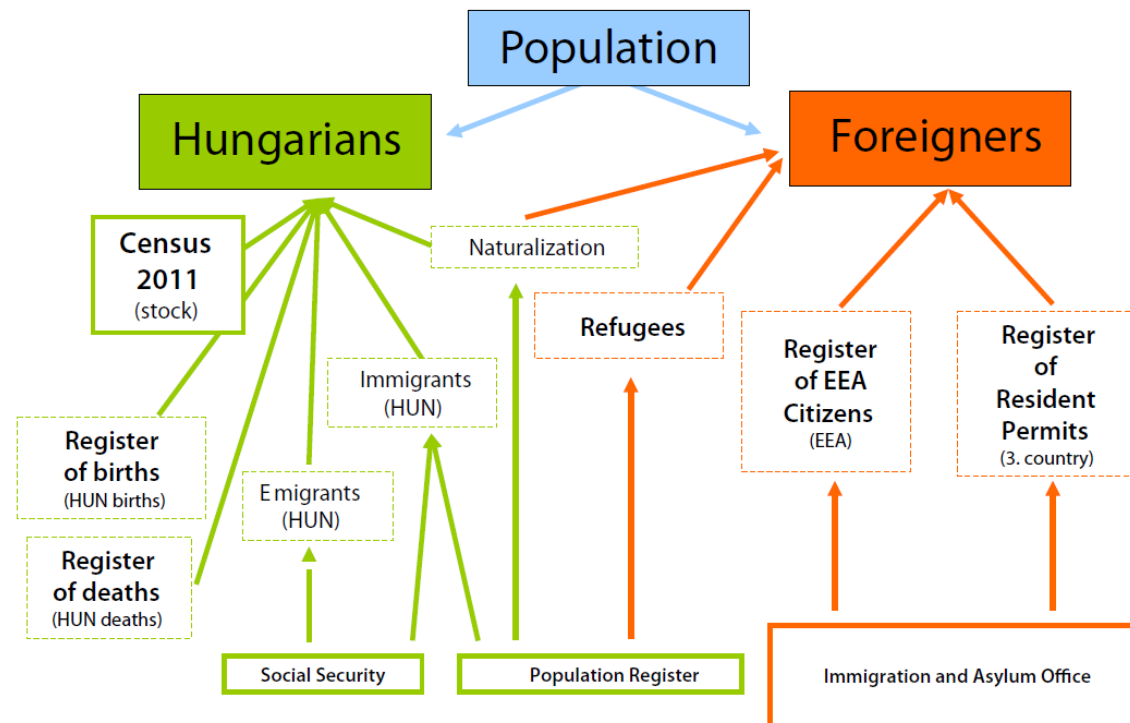 - data integration is carried out to:
    - improve data quality
    - adjust for missing de-registrations of emigrants
    - identify deceased persons
    - provide additional information on variables not included in primary sources
  - practices based on 'presence signals' (or 'signs of life') of individuals

UNECE

Case studies – main outcomes (2/3)

- In some countries, the population register is one of several sources, not necessarily the most important

Data sources for annual Hungarian population calculation

Case studies – main outcomes (3/3)

- In countries without a population register
  - sources of migration data include:
    - data on passports, visas and stay permits
    - passenger cards filled at borders
  - data integration is carried out to:
    - improve data quality
    - adjust for missing de-registrations of emigrants
    - identify deceased persons
    - provide additional information on variables not included in primary sources
    - → improve information on actual place and date of settlement of immigrants

UNECE

Metadata

- Comprehensive metadata are necessary to assess the comparability of migration data produced by different countries

- Three main types of metadata:
  1. Metadata specific to each data source, describing main features and how data are transformed before integration
  2. Metadata for each step of data integration (i.e., for each pair of data sets being integrated)
  3. Quantitative measures of data quality, to assess the effects of data integration

Recommendations

1. Improve access to administrative data for national statistical offices

2. Use administrative data for migration statistics

3. Combine data from different sources using 'presence signals'

4. Pay attention to the quality of integrated data

5. Be transparent about data integration methods used and develop standards

6. Promote international comparison and exchange of migration data

UNECE