

# Task Force on Data integration for Disaggregated Statistics on International Migration: Status Update

Jason Schachter, Co-Chair  
Chief, International Migration Branch  
U.S. Census Bureau

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

20-22 June 2022, Rabat, Morocco

# Background

- Expert Group on Migration Statistics: Task Force 3
- Contribution to the revision of the 1998 United Nations Recommendations on Statistics of International Migration.
- Technical report that supports countries to produce sufficiently disaggregated data for the measurement of international migration by means of integration of micro- and macro-data techniques.
- The report covers:
  - (a) variety of methodologies to integrate and analyze migration data on the macro- and micro-level
  - (b) legal, policy and technical prerequisites
  - (c) major challenges and opportunities in mainstreaming the use of integrated data in official statistics
  - Includes several illustrative case studies

# Current Task Force Membership

- Canada, Chile, Colombia Georgia (former co-chair), Mexico, Morocco, Norway, Republic of Moldova, (Switzerland), United Kingdom, United States (co-chair), Australian National University, Eurostat, ILO, UNSD/ISWGHS (co-chair), and WHO
- Consultant hired to draft report: Jasmine Trang (Assistant Professor in Sociology, Western University (Canada))

# Current Status of the Report

- Complete first draft finished
- Five chapters, plus an appendix with detailed country case studies
  - Incorporates country examples to illustrate points
- Currently with Task Force members for comments
  - Several questions for TF members and areas for potential expansion
- Once comments are incorporated, will send next version to larger Expert Group for comments/consultation

# Chapter 1: Introduction

- Data Integration in the context of the new guidelines and conceptual framework on international migration statistics
- Definitions of Data Integration
  - Micro-data integration creates new combined data sets, which can produce statistics
  - Macro-data integration creates new statistics (estimates) from outputs from multiple data sets
- How data integration can improve migration statistics
- Use of outputs derived from integrated data in official statistics

# Chapter 1: “to do”/discussion

- Expand discussion of Task Force origin
- Better linkage to new conceptual framework?
- More discussion about how data integration fits into the production of “official statistics” (Academia vs NSO)
- For Task Force discussion:
  - Use of surveys for micro-data integration
  - “Data Integration” vs “Data Harmonization”
  - Reconciliation of multiple (different) estimates of migration from different sources
  - Should the report serve more as a “guide,” as opposed to a literature review/illustrative examples? (wider group question)

# Chapter 2: Macro-Data Integration Methods

- Overview
- Compilation Methods
- Statistical Adjustment Methods
- Statistical Modeling Methods
- New and Hybrid Methods
- Challenges for Macro-Data Integration

# Chapter 2: “to do”/discussion

- More detail needed for some country examples
- Additional country examples?
- Examples of Small Area Estimates methods linked to migration?
- Are there any “big data” examples?
- For TF discussion/expansion:
  - How to compensate for differential coverage in different data sources
  - Is “Data Triangulation” a data compilation or statistical adjustment term?
  - Extent to which countries assess different data sources used as inputs for data integration?
    - Do they accept errors and/or do they inform which data are used?
- How to communicate measures of uncertainty to the general public? (wider group discussion)



# Chapter 3: Micro-Data Integration Methods

- Overview
- Creating/Enabling the Legal Framework
- Creating/Enabling the Technical Framework
- Micro-Data Integration Methodology
- Challenges for Micro-Data Integration

# Chapter 3: “to do”/discussion

- More details needed for some country examples
- Clearer distinction between deterministic and probabilistic linking/parametric and non-parametric matching
- Expand section to focus on variables required to study international migration
- Provide examples of data cleaning
- Include examples of ways to measure emigration (e.g. “signs of life” methodology)
- TF discussion: How to convey work currently in progress? (will report become quickly dated?)

# Chapter 4: Assessing and Communicating Results

- Overview
- Estimate Assessment/Validation
- Communication with Key Stakeholders
- Use of Outputs derived from Integrated Data in Official Statistics

# Chapter 4: “to do”/discussion

- Need to expand this chapter
  - Looking for country examples to help with this
- More focus on dissemination
- **Alternatively, could change this chapter to have more focus on guidelines for data integration**
- Cases where data are used to both evaluate and “correct” results
- Distinction between communicating with partners before and after data integration
- For TF Discussion:
  - Ability of those outside of NSOs to conduct micro-data integration?
  - Lack of open data portal examples (deleted)
  - Interplay between adjusted/modeled statistics and “official” statistics (wider group)

# Chapter 5: Conclusions and Future Work

- Data Integration in the context of the new guidelines on international migration statistics
  - Improve access to and integration of administrative data sources
  - Develop international data exchanges and collaborations
  - Combine micro- and macro-integration
  - Provide better understanding of new data sources, uses, and limitations
  - Apply Machine Learning and other “big data” methods

# Chapter 5: “to do”/discussion

- **Need to discuss recommendations with Task Force**
- Have recommendations stand on their own, not needing to read the rest of the report
- Expand most sections
- Better connect to work done by other Task Forces (e.g., TF 2 data needs for new proposal, T4 on new data sources, etc.)
- Other examples of combined macro-micro data integration?
- TF discussion on possible resistance to use of estimates derived from “integrated” sources

# Appendix: Country Case Studies

- Country case studies summarize information from relevant countries' experiences
- Each example of data integration is organized around six main themes:
  - (1) Data needs,
  - (2) Collaboration framework for data integration,
  - (3) Data sources,
  - (4) Integration methodology,
  - (5) Data evaluation strategy, and
  - (6) Dissemination/ Communication of results.
- Country examples: Canada, Chile, Georgia, Mexico, Moldova, Norway, Switzerland, United Kingdom, United States,

# Case Studies: “to do”

- Additional details needed from several countries
  - “Data evaluation strategy” and “Dissemination” in particular
- Additional country examples?
  - Morocco? South Africa?



# General Questions for EG

- Should there be greater emphasis in the report to serve as a “guide,” as opposed to a literature review/illustrative examples?
  - Would it make sense to convert Chapter 5 into specific guidelines for countries to follow to help them with data integration, or is this beyond the scope of the report?
- How to communicate measures of uncertainty/data quality limitations to the general public?
- Interplay between adjusted/modeled statistics and “official” statistics
- Applicability of report to developing countries?
- Host a wiki platform so new case studies can continue to be added after the report is completed?

# Next Steps

- Task Force met yesterday and will meet again next week to discuss points
- Once incorporate Task Force comments, will send a draft to entire Expert Group for comments/consultation
- Final version complete by end of 2021?
- Submitted a proposal for a session on data integration at the IFMS in Santiago, January 2022.
  - Canada, Chile, Georgia, the UK, and the USA have agreed to present at the session.