Economic and Social Council

Economic Commission for Europe

Conference of European Statisticians

Sixty-sixth plenary session
Geneva, 18-20 June 2018
Item 4 (b) of the provisional agenda
Data integration for measuring migration

Guidance on data integration for measuring migration

Addendum

Results of the consultation on the Guidance on data integration for measuring migration

Note by the secretariat

Summary

  The document summarizes the comments by members of the Conference of European Statisticians (CES) on the *Guidance on data integration for measuring migration* (ECE/CES/2018/2). The UNECE Secretariat carried out the electronic consultation in March-April 2018.

  A total of 44 countries and 2 international organizations replied to the consultation. All of them supported the endorsement of the Guidance, subject to the comments provided in the consultation. This note presents the substantive comments received, together with the response of the UNECE Task Force on Data Integration for Measuring Migration, including suggestions for amendments to the Guidance to address the comments.

  In view of the support received, the 2018 Conference of European Statisticians plenary session will be invited to endorse the *Guidance on data integration for measuring migration*, subject to the amendments presented in this document.

I.   Introduction

1.   This note summarizes comments made by the members of the Conference of European Statisticians (CES) on the *Guidance on data integration for measuring migration*.

2.   The Guidance was developed by the Task Force on Data Integration for Measuring Migration (chaired by Spain) that was established by the CES Bureau in October 2015. The Bureau reviewed the draft document in February 2018 and requested the UNECE Secretariat to send it to all CES members for electronic consultation. The UNECE Secretariat conducted an electronic consultation on the Guidance in March-April 2018.

3.   The following 44 countries and 2 international organizations replied to the consultation: Albania, Armenia, Australia, Austria, Belarus, Belgium, Bosnia and Herzegovina, Canada, Chile, Colombia, Croatia, Cyprus, Czechia, Denmark, Finland, Georgia, Germany, Hungary, Ireland, Israel, Italy, Japan, Latvia, Lithuania, Malta, Mexico, Mongolia, Netherlands, New Zealand, Norway, Poland, Portugal, Republic of Moldova, Romania, Russian Federation, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, United States, Interstate Statistical Committee of the Commonwealth of Independent States (CIS-Stat), and United Nations Statistics Division (UNSD).

4.   The comments and the Task Force's responses are summarized in sections II and III. Several countries provided detailed editorial comments and suggestions, which are not presented here but will be taken into account in finalising the document. Comments on the specific situation in individual countries that have no direct bearing on the content of the Guidance are not presented in this document.

II.   General comments

5.   All responding countries and organizations supported the endorsement of the document by CES, subject to the amendments resulting from the comments provided in the consultation and to final editing of the text.

6.   All responding countries and organizations also indicated that the document provides useful guidance on data integration for measuring migration, and agreed with the conclusions. Many countries acknowledged the value of the Guidance. For example:

(a)   Armenia: The Guidance is very useful and informative, as the experiences and case studies of integrating data on migration give detailed description of situations, which is very important for measuring migration;

(b)   Belgium: We have thoroughly scrutinised the document, since data integration for migration statistics is one of the priorities of our demography department. It is an excellent point of departure for future work in this field;

(c)   Hungary: Hungary welcomes the UNECE initiative of gathering information on data integration strategies for measuring migration. The collection of experiences of other NSIs on how to overcome challenges of producing high-quality migration statistics are very useful;

(d)   Poland: The document presents in a clear and synthetic way the essence of data integration in the context of measuring migration. It was good to provide working definition of data integration at the very beginning of the document to clarify the main topic. Variety of national approaches presented as case studies shows how complicated it is to provide reliable and internationally comparable migration data.

III. Specific comments

7.   Several countries also provided specific comments on various chapters of the guidance. The substantive comments are presented in this section, together with the responses by the Task Force. Some countries (including Australia, Canada and Hungary) provided an updated version or amendments to their case studies.

A. Chapter I. Introduction

8.   Mexico: It would be important to mention the work of UNSD concerning the definition of migratory status.

Response and changes proposed by the Task Force

9.   This work will be reflected in the final version of the text.

B. Chapter II. Definition of "data integration"

10.   Australia: As data integration for statistical purposes has matured, when people talk about data integration they almost always are referring to micro integration – linking corresponding record pairs. In addition, the difference between macro integration and other processes such as accounts compilation are often unclear. We consider necessary to adopt the micro definition.

11.   Referring to paras. 14, 27 and 28 of the Guidance, Hungary pointed out the need to clarify definitions of record linkage and statistical matching, considering that the availability of a unique identifier determines whether deterministic or probabilistic matching is meant. Hungary further commented: "Statements on statistical matching should be revised, because it can be divided in micro- and macro level approaches. It is possible that the authors erroneously use the term statistical matching instead of probabilistic matching."

12.   On the working definition of data integration (para. 26), Hungary expressed preference for the wide definition that the Guidance proposes, and stated, "However, in our view, the definition is not consistent with previous para. 22 – there are statistical operations which are border cases with data integration, such as data compilation and data comparison." Referring to the sentence "Data compilation can result in *enlarged*, and data comparison (for validating) can result in a *higher quality dataset"*, Hungary asked for clarification on whether data compilation and data comparison are understood as data integration or not, in an additional sentence after the definition in para. 26.

13.   Mexico asked for a scheme on data integration at macro and micro level, the methods used and the results to illustrate paras. 23, 27 and 31.

Response and changes proposed by the Task Force

14.   The Task Force agrees that micro integration is the most common practice of data integration. However, the Guidance should show also other practices, such as macro integration, which is illustrated by the United Kingdom case study.

15.   On record linkage and statistical matching, the text will be edited to reflect precisely the definitions of statistical matching, probabilistic and deterministic record linkage, and the distinctions between them. It will also be clarified that both data compilation and comparison are not examples of data integration.

16.	A new diagram will be added to help readers understand examples of data integration.

## C.	Chapter III. Survey on data integration practices

17.	No substantive comments received on this chapter.

## D.	Chapter IV. Case studies

18.	Chile noted that some case studies go into more detail on the methods than others, and asked to know more about the methods used in the Austrian case study, especially on specific limitations or strengths of the methods.

19.	Latvia suggested adding a reference to a source where more information is available, for example in the case of Hungary, Israel and United States.

20.	New Zealand provided an update on their case study concerning implementing the 12-16 month rule to measure migration that is expected to be introduced in a few months' time.

21.	Portugal noted that some case studies do not encompass significant data integration within the production of migration statistics. On the other hand, examples of international exchange of micro-data, namely among Nordic countries, or a reference to relevant and publicly available documents should be included (namely, in para. 305). Some of the case studies' summaries could be more detailed on aspects for which there is available information in the underlying methodological documents (for example, paras. 93, 211 and 232). Finally, it would be important to include case studies in the table of contents.

22.	UNSD considered the case studies helpful to understand how countries use data integration to produce migration statistics. The level of detail and clarity of explanation vary from country to country. For example, the coverage of different data sources is not always clear. Each administrative source has its rules for registering/de-registering persons, which have a major impact on the coverage of migrants. It would be useful to spell out those rules as metadata. Furthermore, most case studies cover the issue of integration at the micro level. It would be useful to have more examples on how integration can be done at the macro-level, especially when there is a difference in aggregated numbers obtained from multiple data sources. Finally, the use of terminology is not always consistent in the country case studies. The same term is used with different meanings (e.g. emigration, out-emigration, return emigrant, long-term emigration) which may lead to confusion.

Responses and changes proposed by the Task Force

23.	The case studies were intended to follow a common structure based on four sections, but some countries preferred to provide shorter reports, which could not be presented in this manner. In contributing case studies, countries did not have to strictly follow the proposed structure.

24.	Updated versions of the case studies have been provided by Australia, Canada, New Zealand and other countries. They will be incorporated in the document. The case studies will be also listed in the table of contents. The Task Force recognizes that some case studies refer to migration flows whereas others address stocks of foreigners. This is consistent with the scope of the guide, which covers both flow and stock statistics.

25.	The Guidance reflects the situation in terms of actual practice with data integration for measuring migration. Micro integration is clearly more frequent and the Task Force

could collect information on macro integration to the extent countries could provide it. The terminology will be harmonised throughout the Guidance.

E. Chapter V. Metadata

26. Germany raised a question whether a standardized description of the data sources (para 308) and procedures (para 309) used when integrating can contribute to a complete understanding of the work actually done in a country. This holds true even more to the standardized description of the value added (para 311).

27. Hungary asked to add definitions of relevant variables and values (as determined by the legal bases of data collection, if available) as well as the definition of target populations in para. 308. Additionally, the quality of datasets to be integrated, first of all issues of coverage and accuracy, are of crucial importance: knowledge on the quality of data sources should be included in metadata. It is recommended to add in para. 309 the aim of the data integration exercise: e.g. comparison/validation, summing data on different sub-populations, etc.

Responses and changes proposed by the Task Force

28. The metadata chapter should be considered a first contribution in the direction of providing metadata about the process of integrating sources to measure migration. The list of items in paras. 308 and 309 is not intended to be exhaustive. This will be made clear in the text. The considerations about target population and quality issues are covered in para. 312.

F. Chapter VI. Conclusions

29. CIS-Stat considered that the document does not pay enough attention to population censuses as a possible source of migration data. The block of questions on migration in the census questionnaire occupies a significant place in countries. This makes it possible to obtain a variety of data on migrants. However, the census is not mentioned in the Conclusions and Recommendations.

30. Finland commented that the report mentions that the main weakness of the population registers is their difficulty in detecting emigrations (para 320). But in countries like Finland, where population registers are used, it is nevertheless possible to find out about emigrations that have not been notified with the register data and signs of life analysis. Especially the Nordic registers are of very high quality and offer a better data source than using only the traditional censuses.

31. Germany found that the suggested exchange of individual level data may prove to be difficult with respect to data protection and statistical secrecy issues in various countries. The chance of introducing such mechanisms would substantially rise, if they didn't require exchanging names, dates of birth and addresses. Germany raised a question whether such 'anonymised' forms of data exchange are possible.

Responses and changes proposed by the Task Force

32. The Guidance focuses on integrating different existing sources and not on those sources themselves. The Guidance recognises that in many countries the main data source on migrants and foreigners is the population census, on which CES recommendations exist.

33. In paras 323 and 333, the Guidance addresses exchange of identified individual level data that can be matched across the exchanged data sets. The Task Force did not have

knowledge or experience to consider exchanges of anonymised data where such matching would be possible.

G. Chapter VII. Recommendations

34. Latvia noted that increasing use of data at the individual level with personal identifiers (PIN) should comply with personal data protection requirements. It is highly recommended to elaborate proposals and IT and confidentiality solutions to enable safe processing of personal data in countries using registers. Not only the quality of data should be evaluated but also the quality of data sources – there could be linkage with recommendations and examples from the *Guidelines for the use of registers and administrative data for population and housing censuses*.

35. UNSD emphasised that improving access to administrative data for national statistical offices is extremely important for migration data, and asked for providing some examples of how countries could accomplish this. For instance, it would be useful to provide: (a) examples of legal frameworks adopted by countries, granting access to administrative data; (b) process followed and steps taken by countries to get access to administrative migration data, including an estimation of the required time frame.

Responses and changes proposed by the Task Force

36. Paragraph 330 will be edited to introduce recommendations on data protection requirements.

37. Legal issues of access to administrative data are outside the scope of this Guidance and are covered elsewhere. The Guidance provides a general recommendation to increase access to administrative data because of its potential for improving measurement of migration.

H. Chapter VIII. Further work

38. Colombia considered that it would be important to reflect in chapter VIII the progress made by the Max Planck Institute for Demographic Research (MPIDR), which may imply the incorporation of new research lines on human mobility and migration.

39. Germany noted that the potential of 'big data' as a means for better migration statistics may be overstated – even when combined with traditional data sources. Again, the data protection and statistical secrecy issues associated with the use of 'big data' are not addressed.

Responses and changes proposed by the Task Force

40. The Task Force will add information about the mentioned initiatives among possible areas of further work.

41. Para. 335 will be edited to mention confidentiality issues.

IV. Conclusion

42. All responding countries and organizations supported the endorsement of the document.

V. Proposal to the Conference

43.   In view of the strong support expressed by countries and organizations, the Conference is invited to endorse the *Guidance on data integration for measuring migration*, subject to amendments presented in this document and final editing of the text.

―――――――