



# Economic and Social Council

Distr.: General  
9 November 2016

Original: English

---

## Economic Commission for Europe

### Conference of European Statisticians

#### Group of Experts on Population and Housing Censuses

##### **Eighteenth Meeting**

Geneva, 28 - 30 September 2016

Item 9 of the provisional agenda

##### **Other business**

### **Report of the meeting**

#### **Note by the Secretariat**

#### **I. Attendance**

1. The meeting of the joint UNECE/Eurostat Group of Experts on Population and Housing Censuses was held on 28-30 September 2016 in Geneva, at the Palais des Nations, back-to-back with the UNECE Workshop on Population and Housing Censuses for countries of Eastern Europe, Caucasus and Central Asia (26-27 September).

2. The meeting was attended by participants from Armenia, Austria, Azerbaijan, Belarus, Bosnia and Herzegovina, Canada, Chile, Croatia, Czech Republic, Estonia, Finland, Georgia, Germany, Hungary, Israel, Italy, Japan, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Luxembourg, Mexico, Mongolia, Montenegro, Netherlands, Norway, Poland, Portugal, Republic of Korea, Republic of Moldova, Russian Federation, Serbia, Slovakia, Slovenia, Spain, Switzerland, Tajikistan, the former Yugoslav Republic of Macedonia, Turkey, Ukraine, United Kingdom, United States of America and Uzbekistan. The European Union was represented by participants from Eurostat and the Delegation of the European Union to Bosnia Herzegovina. The Food and Agricultural Organization, United Nations Population Fund (UNFPA), United Nations Mission in Kosovo, United Nations Statistics Division (UNSD), Eurasian Economic Commission and Interstate Statistical Committee of the Commonwealth of Independent States were also represented. The meeting was attended by experts invited by the Secretariat from the Inter-American Statistical Institute, and the Integrated Public Use Microdata Series International at the Minnesota Population Center.

3. UNFPA supported financially the attendance of a number of participants.

## II. Organization of the meeting

4. Eric Schulte Nordholt (Netherlands) and Danilo Dolenc (Slovenia) were elected as Chairpersons of the meeting.
5. The following substantive topics were discussed at the meeting:
  - a) Assessing costs and benefits of censuses
  - b) Frameworks and tools for bilateral and multilateral cooperation among NSIs with regard to censuses
  - c) Alternative methods for counting of population, in particular hard-to-count population groups
  - d) Methods for assessing quality and usability of registers and administrative sources
  - e) Innovations in census methodology and technology, and results of testing
  - f) Possible uses of new data sources (e.g. “Big Data”) for censuses.
6. A presentation on “How the 1921 United Kingdom Census anticipated the 2020 Recommendations” was given by Mr Ian White under the item “other business”.
7. The discussion at the meeting was based on 22 papers submitted by the participants. The papers and presentations are available on the UNECE website at the following address: <http://www.unece.org/stats/documents/2016.09.census1.html#/>

## III. Recommendations for future work

8. The Expert Group recommended that its next meeting will be organized over three days on the week of 2-6 October 2017. UNECE will hold a workshop for countries of Eastern Europe, Caucasus and Central Asia back-to-back with the Expert Group meeting.
9. The following topics were suggested for discussion in the 2017 Expert Group meeting, acknowledging that contributions on any other census related topics will also be considered:
  - a) Innovations in census methodology and use of new data sources (including enumeration of hard-to-count groups)
  - b) Innovations in census technology
  - c) Cost-effectiveness of censuses
  - d) Evaluating the census and measuring data quality (including setting targets for data accuracy and other quality dimensions)
  - e) Comparability of census results across time, countries, and census methods
  - f) Meeting users’ needs (including: emerging topics, disseminating anonymised microdata samples, statistical disclosure control)
  - g) Production and dissemination of geo-referenced census data
  - h) Challenges in complying with CES Recommendations for 2020 census round
  - i) Cooperation models among countries
10. Countries of Eastern Europe, Caucasus and Central Asia suggested the following topics for discussion at their workshop in 2017:

- a) Use of electronic devices for censuses
  - b) Measuring census data quality
  - c) How to reduce census costs
  - d) Census publicity and communication campaign
  - e) Tools for modern dissemination of census data
  - f) Country updates on census preparations
11. The Expert Group agreed on the following activities for sharing knowledge on population and housing censuses
- a) Create a new page in the UNECE census wiki for presenting information and materials on the 2020 round of censuses in UNECE countries following the example of similar pages for the 2000 and 2010 rounds.
  - b) Explore additional possibilities for sharing news on censuses, such as a more frequent and detailed update of UNECE wiki pages or an electronic newsletter.
12. The Expert Group agreed with the work plan presented by the UNECE Task Force on Register-based and Combined Censuses..

#### **IV. Adoption of the report of the meeting**

13. The present report was adopted during the closing session.
14. A summary of the discussion in the substantive sessions of the meeting will be presented in an annex to this report, to be prepared by the Secretariat after the meeting.

## Summary of the main issues discussed at the substantive sessions

### A. Assessing costs and benefits of censuses

Documentation: Papers submitted by United Kingdom and IPUMS International.

Discussant: Peteris Vegis (Latvia)

1. In preparation for the 2021 census, the ONS aims to maximize online responses with a planning assumption of 75 per cent online responses. Various other countries consider online response as one of the means - if not the main one – to keep under control or even reduce the census costs. The assumption of 75 per cent online responses (against only 16 per cent in the 2011 census) derives from various factors, including the observation of a general move towards government services being provided online, and people being more connected online. Young people are generally more reluctant to complete paper-and-pencil census questionnaires. However, they are the most connected online, so it is hoped that they will be more willing to fill the questionnaire online.
2. In order to maximize the online response rate, special attention will be given to advertising campaign, sending letters by mail, and increasing cooperation with libraries where it will be made possible to gain some assistance with filling the questionnaire (“Assisted digital”). In the United Kingdom, the online response will be promoted as main enumeration method, whereas the availability of paper questionnaires will not be announced widely. However, it should be taken into account that the possibility of requesting a paper questionnaire could spread rapidly through social media.
3. In the United Kingdom, a test will be conducted in order to distinguish the characteristics of paper first versus online first responders, taking into account different regions or population groups. These tests should show whether there are strong evidences to justify sending paper first to anyone at all.
4. One of the major problems with collecting online responses is the management of peak activity dates, when many people are trying to connect to the server and complete a questionnaire at the same time. For instance, Latvia had to extend the response period by two days due to a massive response peak on the last day. A possible option considered in the United Kingdom is to send the letters in stages. Testing the systems and the various options is very important. Regarding non-response follow up, the United Kingdom adjusts published census data for undercount. Therefore, the response rate needs to be as high as possible. A maximum of 10 visits to non-responding address is being considered. However, other data sources (including potentially big data) could be investigated in order to find the best practices for reducing the number of visits required.
5. While there is potential for cost savings from getting the respondent to self-code, this would need to be thoroughly researched and the quality assessed, taking into account the quality assured by an expert coder.
6. The UNSD is working on a global overview report on how to quantify the costs of censuses. In general, high costs tend to have negative connotation. However, countries need to consider also the cost of not having a census, and lacking information needed for policy making.
7. The IPUMS presentation underlined the importance of making census microdata available for academic purposes and policy research. It was noted that the quality of

harmonized data depends directly on how closely the countries follow the UN census recommendations.

8. Individual data provide better opportunities for academics wishing to create complex models rather than using regional aggregated data. In addition to scientific use, microdata are also used to inform policymaking (including, for instance, by United Nations and World Bank). Although the CES Recommendations include references to the preparation of census microdata samples, for the next census round the recommendations could “push” more explicitly countries to produce microdata samples.

9. For countries which cannot provide microdata samples, other ways of accessing the data could be investigated, including for instance a “census hub”. Remote access could be an alternative way of providing microdata files to IPUMS. Remote access could be especially attractive for countries that already provide this service to researchers and no longer want to invest in preparing protected census microdata for the IPUMS project. There are various implications of the two approaches, regarding for instance access to information vs confidentiality concerns. The global scientific community increasingly requires access to data off the soil of the providing country. Remote access could end up needing countries to do even more preparatory work in-house.

## **B Frameworks and tools for bilateral and multilateral cooperation among NSIs with regard to censuses**

Documentation: Paper submitted by the United States.

10. The paper by the United States presented a new framework proposed for bilateral and multilateral cooperation with other NSOs, aimed at modernizing international cooperation efforts and increasing the collaboration with other countries (not just sharing expertise from the United States to other countries).

11. The Canadian representative addressed the value of the materials published by the US Census Bureau, which are used for learning purposes. Sharing experiences internationally was considered definitely valuable.

12. It was suggested to develop some secretariat-based centralisation of materials to be shared by NSIs, i.e. materials published on the websites, results of testing and modelling of census methodologies (something similar to the big data projects inventory). The UNECE could consider developing the existing UNECE census wiki for this purpose. Distribution of a bulletin or newsletter was suggested as a means of alerting NSIs about the new available materials in the census wikis or databases maintained by UNECE or UNSD.

## **C Alternative methods for counting of population, in particular hard-to-count population groups**

Documentation: Papers submitted by France (not presented), Israel, United Kingdom and United States.

Discussant: Marc Hamel (Canada)

13. The United Kingdom presented some results from research on the suitability of a question on sexual identity. It was clarified that minority categories in gender and sexual identity are ‘hard to count’ in the sense that the attribute is hard to count, even if the people are not so hard to find and collect other data on.

14. Although information on sexual identity can be obtained from large sample surveys, census data would allow for complex multivariate analysis using these characteristics. The

results of this analysis could be useful to provide services to groups with protected characteristics under the Equality Act.

15. With regard to these questions, Canada informed the participants that it is assessing and testing questions on gender identity and working on development of a standard. For that purpose, it is important to have a clear idea of what concepts really need to be measured and what they mean. Based on the Canadian experience, it is hard to anticipate the evolution of public attitude to difficult topics. Political change meant that the public called for inclusion of gender identity questions at the last minute when it was too late. Countries should have answers available to people who ask why this topic is or is not included.

16. The value of including questions about very small minorities was debated. It was noted that when small populations are broken down into small areas, the resulting data may not permit valid statistical analysis. So, these questions need to be examined and tested carefully before deciding about inclusion. Canada underlined that research on these very specific and labour intensive research topics would be valuable for all and should be shared.

17. The United States presented an overview of the current plans for reengineering address canvassing in preparation for the 2020 census. A continuous intensive update of the address list is planned for the period between the censuses. It was estimated that in the 2010 census, approximately 10 million addresses were missed from the frame and another 10 million outdated addresses were not removed from the sample.

18. Israel presented a paper on coverage challenges in its multi-frame integrated census. Immigrants who left the country without registering their exit contribute to the over-coverage issues. With regard to under-coverage, the National Population Register of Israel includes only Israelis.

19. France submitted a paper that was not presented in the meeting. It described the complexities of conducting a census in France, e.g. difficulties with enumerating the population living in Calais, French Guiana and on the island of Mayotte.

#### **D. Methods for assessing quality and usability of registers and administrative sources**

Documentation: Papers submitted by Canada, Estonia, Italy, Portugal, Slovenia, United Kingdom and UNECE Task Force on Register Based and Combined Censuses.

Discussant: Jorge Vega (Spain)

20. Slovenia presented the results of a very detailed survey that was carried out to measure over-coverage and under-coverage in their Central Population Register (CPR).

21. Estonia is currently preparing to conduct its first register-based census in 2021. Registers cover almost all census core topics and are considered to be reliable by the population. However, extra preparatory work is required to ensure the interoperability of information systems. Interoperability helps to produce high quality and easy accessible data at a lower cost for the whole society (not just for the statistical office). For the 2011 census, the team of methodologists consisted of two experts, whereas this number was increased to five for the following census, more big data specialists are currently needed. A general lesson for other countries is that a register-based census cannot be conducted on the basis of contracts between institutions, but a proper Service Level Agreement is required.

22. In Italy, the 2021 census will mark the complete transition from the traditional door-to-door enumeration to a register-based system. The model adopted should be based on a

system of register-based statistics, continuously updated from various administrative sources and linked to the social surveys data by a unique ID.

23. Portugal compared data from various administrative sources and census for the most and the least populated municipalities. They calculated 'equality rates' and estimated some quality indicators for a list of variables. Wherever a source was affected by immigration, its level of similarity was low. An overall high level of consistency was observed, more for demographic than for economic variables. They also compared the population estimates and did not find any major differences. When more administrative data are available and when methodology is more stable, they will look more into regional breakdowns. Owners/producers of administrative data sources now have feedback that encourages them to continue work on improving data quality.

24. Canada has promoted a project to assess the possibility to complement the Canadian Census with a statistical population register. The main element of this project, the Canadian Statistical Database (CSDD) aims to replicate the census population universe with basic socio-demographic variables (age, sex and geography).

25. In the United Kingdom, various options were considered for moving away from the traditional census design, and were object of the "Beyond 2011 Programme". Small area data remains in high demand and a public consultation suggested that the demand would not be entirely fulfilled by surveys alone. The Programme made a recommendation that the census in 2021 will be predominantly online, making increased use of administrative data and surveys. The focus is now on whether a census based on administrative data and surveys can replace a traditional census after 2021. This will require integrating administrative data with two new surveys. A first assessment on the ONS's progress towards an Administrative Data Census was conducted in May 2016. It showed that work to produce population estimates has been quite successful but that integrating surveys and admin data for population characteristics is going to be harder and will require further investigation.

26. In the discussion it was pointed out that owners of "big data" and private sources may eventually come up with alternative information faster or more frequently than the census. The following questions were widely discussed as well: 'Do users value data accuracy more than speed at which the information is obtained? How good is good enough? Do data providers establish a framework in advance by which they will make the call of whether or not something can be used? The United Kingdom has a detailed set of evaluation criteria which they are using to assess their progress towards an Administrative Data Census each year.

27. It was observed that if there is correlation between degree of register correctness and size of urban settlement, and if the allocation of municipal funds is based on statistics, then accepting corrections may be difficult especially when they are applied to different degrees in different places. The United Kingdom may consider whether a different methodological approach is required for more challenging urban areas where high population movement may impact the quality of population estimates. There is necessity for transparency. A balance has to be found between providing high quality and comprehensive data and not being so complex that users cannot understand how estimations are made.

#### **Panel session on "Evaluating the potential for moving away from a traditional census"**

Moderator: Ian White

Panel members: Adelheid Bauer (Statistics Austria); Becky Tinsley (ONS – United Kingdom); Harald Utne (Statistics Norway); Irina Zbarskaya (CIS-Stat).

28. When a country moves away from the traditional census, an important issue is to maintain user confidence. For this purpose, Austria involved census stakeholders in the

discussions from the beginning of the transition period. The data quality indicators developed as well as the results of comparing census and survey data were made public. In the UK, special attention is given to the user consultation. Users are encouraged to provide feedback on research results as well as on data design and methods applied. A biannual research conference is organised to provide necessary consultation to the user working groups. The balance between methods and complexity of methods is considered in order to optimise output and user trust. A group of independent experts are invited to give an overall review of concepts and methods.

29. The control over administrative data is a fundamental issue when this type of data is used for census purposes. Responsibility for the end result and census products must lie on the statistical office disregarding which body produced the initial data. Mutual cooperation is necessary to ensure proper receipt of data. It is important to work to build awareness that what statisticians are doing is mutually beneficial, not only for their own benefit. By definition, administrative data are not collected for statistical purposes. So we will never have full control on them and neither should we, however some influence is valuable. Statistics Norway has the possibility to give feedback when registers are going to be changed. The Statistical Office is influential in that they possess greater knowledge in terms of quality of the sources. Statistic Norway often cannot report regarding specific errors because of confidentiality issues, however, they are able to prepare a technical reports on data quality which are useful to producers.

30. In Austria, the statistical act says that Statistics Austria must use administrative data where it is available. Relevant authorities are responsible for providing administrative data. Statistics Austria should also be allowed a say if registers are changed.

31. In the UK, proposed new legislation would provide the Office for National Statistics (ONS) with greater and easier access to a range of data sources held within the public and private sectors to fulfil its statistical functions. In some cases, the legislation also permits the UK Statistics Authority to set out obligations as part of a formal agreement requiring data suppliers to consult with the Authority before making changes to the data they collect, the way they collect or process these data, as well as any arrangements providing the Authority with access to these data. It will be necessary to build relationships with data suppliers to make the process for sharing data as smooth as possible.

32. In Germany there are concerns about having to accept low quality data from administrative sources. The concern is whether this data will ultimately be of lower quality than data collected for statistical purposes.

33. The participants expressed an overall concern regarding the possible use of big data for censuses. The legal and ethical issues emerging from the attempts to create personalised data from big data sources. There is general agreement that big data has potential to be used for census purposes, but that it will take time before census users would consider this data reliable.

34. There seems to be potential for conducting longitudinal studies in a register-based environment. In Austria, a lot of research is currently being conducted in this field. Norway noted that administrative data have exact dates, often missing in censuses. It also contains information on events that take place between censuses, which allows to conduct longitudinal analysis on a micro level.

35. The UNSD pointed out that countries outside the UNECE region are mostly conducting traditional censuses due to lack of relevant good quality registers and resources. Outside this region, there are many major barriers to overcome to even start considering the possibility of using registers. In many countries, conducting a census is a source of national pride, one of many intangible aspects of traditional census. In some cases even more important than the results themselves.

## **E Innovations in census methodology and technology, and results of testing**

Documentation: Papers submitted by Canada, Hungary, Italy, and Poland.

Discussant: Marco Buscher (Switzerland)

36. Canada presented the approach used for the 2016 census, which is very similar to the one used for the 2011 census (wave collection methodology with response collected mainly via internet). The long form returned to mandatory status for 2016 and was collected on a 25 percent sample. Administrative data were used to support the census operations. In the special case of a city that was heavily affected by a forest fire during the census period, the population count was produced based on administrative data, complemented by long form data collected afterwards. Marketing specialists provided advice in various areas, including for instance the use of social media and other approaches to reach difficult to enumerate populations. Results show that this approach improved self-response for targeted groups.

37. In Hungary, the 2016 microcensus (covering 10% of the population) is being used to test a number of innovations. Data collection included two phases: during the first (lasting 9 days) respondents answered solely via internet, while during the second enumerators visited the households that had not responded and collected data using laptops and tablets. An electronic questionnaire was made accessible for disabled persons (blind and deaf). The importance of communicating data reliability and security of the electronic mode was underlined. Around 2/3 of the enumerators used their own devices (tablets) using the standard application form. Based on the results of this test a decision will be taken on the number of tablets to be purchased for the census. The results of the microcensus will be used to produce estimates of the population size only at a district level

38. For the 2011 census in Italy, administrative data was used only for checking the consistency of the results. Future plans include releasing annual census-type data with increased use of administrative sources and estimation methods. Large work is currently being undertaken in order to prepare a georeferenced integrated system of streets and addresses, which is also a part of Istat's legal responsibility. There is strong cooperation with the cadastre. Similar sources owned by private entities exist already, but the quality is relatively low, compared to the high quality required to produce official statistics.

39. In Poland, work is currently being carried out on the standardization of the geocoding system. There was a call for a common European identification system for addresses: five layers coded with a common system, e.g. ID of admin level, city or town, stat area, street, building number, dwelling number.

40. It was pointed out that merging geospatial and statistical information becomes more and more important, although different units are considered in different countries. In Poland, for instance, the focus is on dwellings. In Italy, the focal points are addresses/street numbers, which then have attributes e.g. residential, commercial. In Hungary, it was specified that the microcensus does not enumerate people without standard addresses, e.g. homeless.

## **F Possible uses of new data sources (e.g. "Big Data") for censuses**

Documentation: Papers submitted by Poland, Russia, and Spain.

Discussant: Fabio Crescenzi (Italy)

41. The presentation from Spain described three new data sources which are being investigated in preparation for its 2021 census. These are tax data, power consumption and

mobile phone information. The 2021 census will be the first in Spain to be mainly based on administrative sources. Statisticians are particularly interested in expanding the amount of data published from these additional sources. Examples of data generated by each data source were presented, with a particular focus on additional geospatial information that could become available through each source by linking these data with census data.

42. Poland provided an overview of the potential and limitations of using Big Data in the 2020 census round. Several examples of possibly beneficial uses of this type of data were described, as were several international initiatives focusing on Big Data (including those from UNECE and ESSNet). In the conclusion, the speaker underlined the importance of further research on the possibilities of implementing innovative data sources, e.g. Big Data. However, due to the technical and legal limitations, implementation into the upcoming census round is not likely.

43. The presentation from the Russian Federation described the modernization of its upcoming 2020 census. Specifically, it focused on the use of electronic questionnaires. An opportunity to participate in the online census will be offered to everyone. The coverage of these questionnaires is expected to range between 25% and 60%. Rosstat is also reviewing the possibility of using existing administrative data to check the census for consistency and also to incorporate Big Data and geoanalytics into these cross-checks. Future efforts will be made to improve these data sources as a part of Rosstat's initiative to modernise its statistics and reduce respondent burden and overall cost.

44. Discussion of this item focused on practical obstacles to implementing Big Data in the upcoming round of censuses. Coverage bias, lack of transparency and the unstructured nature of Big Data need to be addressed before this type of data sources can be implemented into censuses. The differences between public perception of the collection and use of Big Data among private companies and government statistical agencies were also highlighted. Despite these challenges, most participants agreed that Big Data can and should be used already in the upcoming census round as a part of data validation. The 2020 census round should yield additional information on the possibility of further integrating Big Data into future censuses.

---