

REPORT OF THE WORKSHOP

I. Organization and attendance

1. The United Nations Economic Commission for Europe (UNECE) organized the Workshop on Population and Housing Censuses for countries of Eastern Europe, Caucasus and Central Asia (EECCA), in partnership with the Interstate Statistical Committee of the Commonwealth of Independent States (CIS-STAT) and with financial support from the United Nations Population Fund (UNFPA). The Workshop was held in Geneva on 26-27 September 2016, back-to-back with the meeting of the UNECE-Eurostat Group of Experts on Population and Housing Censuses (28 September – 30 September 2016).

2. The Workshop was attended by 48 census experts and managers representing the following countries and organizations: Armenia, Azerbaijan, Belarus, Chile, Estonia, Georgia, Israel, Kazakhstan, Kyrgyzstan, Lithuania, Mexico, Mongolia, Republic of Korea, Republic of Moldova, Russian Federation, Tajikistan, Ukraine, United States of America, Uzbekistan, United Nations Statistics Division (UNSD), UNFPA. The meeting was also attended by representatives from the Eurasian Economic Commission, CIS-Stat, Inter-American Statistical Institute (IASI), IPUMS International census dissemination partnership, Minnesota Population Center, University of Minnesota. Ms. Irina Zbarskaya (CIS-Stat) facilitated the discussion.

3. The present report presents the summary of the discussion and the recommendations of the Workshop. The presentations made at the Workshop are available at the UNECE website¹.

II. Purpose

4. The purpose of the Workshop was to provide an opportunity to census managers and experts from countries of Eastern Europe, Caucasus and Central Asia to discuss their plans for the next census of the 2020 round, and possible issues related to the implementation of the *Conference of European Statisticians (CES) Recommendations for the 2020 Censuses of Population and Housing*.

¹ <http://www.unece.org/index.php?id=41294#/>

III. Summary of the discussion

A. Technological innovations for the 2020 census round

5. The session started with an overview presented by **UNECE** of the use of different census methods and internet response in the 2010 census round and the plans for the 2020 round, at the UNECE level and for EECCA countries in particular.

6. **CIS-Stat** presented the plans for methodologies and technologies to be used in CIS countries in the 2020 round. Several CIS countries are moving towards the use of multiple data collections channels, including internet response, administrative data (mostly to support and check the data collection) and possibly also “Big data” in Russian. Belarus is preparing to use tablets for the first time in the CIS region, and not use paper questionnaires at all. New data entry methods will reduce the need for manual entry, which is subject to human error. There will be new modes of dissemination because people access data in new ways. New applications based on geospatial data will allow monitoring the interview processes, and make possible new forms of presentation of the results.

7. Estonia, United States, Belarus, Russian Federation and Lithuania gave some presentations on new information technologies used in the 2010 round or planned to be used in the 2020 round. The presentations and the following discussion focused on various issues encountered by countries on the implementation of new technologies, and on lessons learned.

8. In **Estonia** new technology was used at all stages of the 2011 census: preparation, enumeration (laptops with GPS used for interviews), workflow (optimising enumerators’ working lists), monitoring, post-census period. New IT solutions reduce manual work but some manual processes are still needed for checking. IT use reduced time taken in all stages and costs. It is estimated that the online census allowed a 20% reduction of costs.

9. During the discussion, it was explained that about 20% of the addresses required visits, and about 50% of the addresses had to be reviewed and corrected during the data processing stage. Georeferencing was done on the street to avoid bothering respondents too much. Working lists were prepared using data from registers, but the clearly the situation may be different by the time of the census (e.g. people may change house, dwellings might burn down).

10. Pre-filling was used for some variables that do not change easily, including for instance citizenship, place of birth, education, legal marital status. There was a discussion about data protection implications of using pre-filled data.

11. Various methods were used to convince people to respond to the census. The public information campaign began one year before the census. Then, one month before the enumeration various methods were used to promote the participation: email, letters, media channels, tv, radio, website announcements, social media acting as support centre and as a means of contacting the statistical office (e.g. if a household had not been enumerated).

12. For the internet response one of the identification methods used was the digital signature (Estonian id cards can be used for this purpose).

13. An under-enumeration of 3% was estimated, using registers and mathematical modelling.

14. The representatives from the **United States** shared experience and lessons learned on the planned use of handheld devices in the 2010 Census, and on the technology planned to be used in the 2020 census. It was noted that technological change is very fast and technology becoming obsolete rapidly considering the 10 year gap between censuses.

15. For the 2010 census there was an increase in costs compared to previous census, because the population is less willing to participate, harder to contact and because of higher employment costs of staff. In the US there is expectation of 100% coverage with perfect data, and the final few percent are the most expensive to count. If the US government could accept 98% coverage, then a large percentage of costs could be saved.

16. The possibility of census staff using their own devices was tested but it did not work due to different operating systems and hardware versions. It was decided to use off the shelf devices (and not custom made) that will be rented, not purchased. Identification of dwelling units may in future happen in partnership with commercial entities (e.g. Google).

17. **Belarus** shared experience on the use of geoinformation technology in statistics. There will be an assessment of the quality of the population database as a source of information on the number of people in each dwelling. It was discussed whether the 2020 census should be based on the actual place of living of persons, or the place where they are registered. There will be obligation to participate in the census. A process has been developed to check the address lists against the official registry of real estate. Every six months the real estate registering agency is required to send an update to the NSO (more frequent updates will take place closer to the time of the census). The country is starting to follow the Lithuanian example, where users can click on map images to see census data, and there is an exchange data between geospatial, census and real estate databases.

18. The expert of the **Russian Federation** introduced the paperless technology used in the Federal statistical survey "Socio-demographic survey (microcensus of the population) 2015". For the 2020 Census it is planned that paper questionnaires will be printed for approximately 55% of population plus a reserve. Tablets will be used in big federal cities, paper in the rest. The logistics of distribution of census materials differs for the two formats since it is expensive to move piles of paper around the country.

19. During the discussion it was clarified that although the automated system for tablets was developed in-house, the procurement deal includes an expensive servicing contract. The internet collection phase will take 2 weeks, then 4 days will be taken to clean the data and determine the premises to be visited, and then 10-13 days to visit and enumerate. Overall it will be about one month for the field

enumeration. There are plans to provide open access to microdata on the website, without authorisation or registration.

20. The representative of **Lithuania** spoke about dissemination of census data using two GIS applications: Interactive Atlas, and Detailed Statistics. Since these tools can present census results at very high level of detail, there was a discussion about the confidentiality of the data presented.

21. In the discussion that followed the presentations **CIS-stat** proposed to have in future a special session on census quality assessment. The participants supported the proposal.

B. Implementation of the CES Recommendations for the 2020 round

22. **CIS-Stat** informed about the Resolution of the CIS Council of Heads of State to conduct population censuses in CIS countries as close as possible to 2020, preferably in the period from October 2019 to October 2020. The representative of CIS-Stat gave a presentation on key socio-demographic indicators for comparability of census results across CIS countries. The Heads of NSIs of CIS member countries approved a list of indicators to be included in the censuses of the 2020 round, based on the CES Recommendations. The majority of the indicators refer to migration topics. It was stressed that for CIS countries migration is a topic of main importance.

23. It was also observed that the census can provide a large quantity of disaggregated data that are needed for SDG monitoring. CIS countries have identified 103 SDG indicators that are most relevant to them. They are now working to find the best way to provide information on them.

C. Update on plans for next censuses of 2020 round

24. In this session, countries provided a short update on the preparations for the census of the 2020 round, informing about progress made and highlighting the major challenges.

25. In **Armenia**, a traditional census is planned in 2020, with data from the State Population Register planned to be used to compile respondents' lists and to control the completeness of coverage. Scanning devices will be used to input data from paper questionnaires if funds will be available. Otherwise, data entry will be done manually.

26. In **Azerbaijan** the next census will take place in 2019. There are plans to make use of data from the population registers to support the census operations. Moreover, the use of tablets and other handheld devices will be tested during the pilot census.

27. In **Belarus**, like in Russia, for the next census there will be a combination of online data collection and interviews, using paper questionnaires. It is estimated that the online census could save about 30% of the costs. The legislation will be revised to better define the role of state authorities with regard to the census.

28. **Kazakhstan** plans to do the census in 2020, making use of data from registers, internet response, tablets for field data collection, and GIS technologies. It is expected that 20-25% of data will be derived from registers. Work is under way to develop the registers for this purpose. Kazakhstan intends to procure 'industrial type' tablets, and pre-load them with data from registers. It was not yet decided whether security arrangements for tablets will be provided by the supplier or in-house. It was noted that if data protection measures will be provided by the suppliers, the costs could be very high.

29. In **Kyrgyzstan** the next census is planned in 2020. In principle it will be a traditional census, although the possibility of providing internet response and to use data from the State Population Register will be considered.

30. In the **Republic of Moldova** the last census was conducted in May 2014, and manual data processing is still in progress. For the 2020 census, there are plans to use internet response (CAWI) and CAPI, and use administrative sources for listing.

31. In the discussion that followed the presentations, it was observed that several countries are planning to use registers, but information is often limited on the quality of those registries. There is need to understand their potential shortcomings, including completeness and data protection issues. For countries planning to use tablets, the protection of data on the devices against possible abuse by enumerators is an important issue. Usual security measures such as closed rooms, disconnection from internet, etc. can't be applied to mobile tablets. With regard to pre-loading tablets with existing data, it is important to ensure that enumerators actually ask questions rather than just relying on what is pre-uploaded from registers.

32. The **Russian Federation** plans to conduct the 2020 census using a combination of internet response and interviews. In order to log in to the census portal and complete the online questionnaires respondents will use a code assigned by state pension system. In principle anyone over 14 can have one. The number of codes assigned is growing by 25% each year. If in the internet collection there will be mistakes or missing data, then interviewers will be able to add the missing data. The payment of census field staff will be done according to the amount of data collected regardless of the source (internet response, interview). In case of refusals or households not found, administrative data can be used for selected variables, but the fact that the person(s) actually live in the place must be verified with neighbors, utility providers, housing managers, etc.

33. **Tajikistan** plans to carry out the next census in 2020. Many innovations are planned. The possible use of tablets for the first time will be tested in 2018. It is expected to get 70% of population responses using tablets.

34. **Ukraine** will hold a census in 2020. Cartographic and program-technological material are being prepared for use in the census, and GIS technologies will be used. There will be need to buy the equipment.

35. **Uzbekistan** is planning to conduct a census in 2020, with a traditional data collection using scanners, and possible internet response. Some equipment plans to be bought.

36. In the discussion that followed the presentation, the issue of the population count used was raised. In fact, when national ID numbers or similar data are used as a basis for the census frame, then countries are inadvertently switching from usual residence to legal residence (i.e. people who are there illegally don't have a number). High migration countries face big problem with concept of usual residence. It is important to think of these issues very carefully because they affect very large numbers of people.

37. For some census questions, like those on disability, there are issues with data quality and speed of data processing. Whether or not enumerators understand the questions they are asking has a big effect on the quality of the data collected.

D. Experience from the 2010 census round

38. The representative from **Georgia** shared some considerations about the experience of conducting at the same time the population census and the agricultural census (in 2014). There are trade-offs between respondent burden, ability to ask many questions and being able to do it all at once. Georgia will consider doing the two censuses together again in the future, because the outcome was much better than expected, and it is unlikely that the quality would be significantly better if they were done separately. Belarus is planning to merge the two censuses as well. The percentage of response to the questions on agriculture in Tbilisi was about 15 percent, and in all urban areas about 50 percent. Many respondents in cities considered the questions sensitive or wondered why they were asked.

39. UNFPA and FAO have considered the positive and negative sides of merging. Possibilities range from identification of households with agriculture to using it for a later survey, or asking the full range of questions. Some guidelines for linking population censuses and agricultural census were prepared by UNFPA and FAO (a reference is included in the CES Census Recommendations).

40. In the discussion that followed, it was noted that this approach is particularly suitable in Georgia where subsistence farming is a very common phenomenon. In other countries where most agricultural enterprises are not household based, this approach may not be suitable. The workload for census field staff had increased from the previous population census, with each enumerator responsible for 100-120 households in a period of 15 days.

41. In Georgia the local administration recruited people to conduct the census on a contractual basis (not as civil servants). It was reported that the relation with local authorities can be challenging because the census results may impact them, so they have vested interests.

E. Review of the availability of administrative sources and possible use to support field collection

42. The representative from CIS-Stat spoke about the use of administrative data sources for supporting field collection. Admin data has always been used in preparatory stages, to prepare frames, maps and working lists. A major change is

that administrative data is now being used in the collection itself, although quality is still a big issue. The UNECE representative noted that it can be expected that most but not all countries will follow the trend towards using admin data as a direct data source for census in medium to long term.

43. Countries were warned against over-optimism in using registers for the census too soon, before they are well developed. Based on the experience of the Nordic countries, a long time is needed to develop the registers to the level required to use them as a replacement of the traditional census. Moreover, the registered population is not the same as the usually resident population, and it is important to define clearly what is to be counted. It was stressed that the move towards using registers affects the whole statistical system of a country, not only the census. Long-term effects of this move in register-based NSIs is that there is decreasing need for methodologists and survey designers, and increasing need for data miners.

44. The Estonian representative presented a review of the availability of administrative sources in Estonia and their possible use to support field collection. Registers were used also to estimate – through statistical modeling – the under-coverage of the 2011 census (2.3%) and the structure of the population that was not enumerated. There were questions about the way the undercount was calculated. When there is a discrepancy between register and census enumeration in some cases register data, and not the enumeration, are not correct. For instance, people counted in the register might have left prior to the moment of the census and therefore should not be included. The relationship between the census moment and the use of registers updated at later moments were also discussed.

45. The representative from the Russian Federation shared experience on using the resources of the Common Authorization and Identification System (CAIS) for the 2020 population census. The inclusion of people from hard to reach regions and nomadic groups was discussed. The population should be allowed to choose the mode of enumeration, and register in one specific place. For the groups in remote areas, it is easier to reach them through internet, because most often they have internet access, rather than sending enumerators in person (e.g. by helicopter).

46. In concluding the session, the representative of CIS-Stat summarized that a number of issues are relevant for CIS countries in view of the 2020 census round, including: census quality control, census costs and ways to reduce them, educational campaigns for the public.

F. Work plan for future activities to support the 2020 census round

47. UNECE reminded that the Task Force on Register-Based and Combined Censuses is working to prepare new guidelines in this field, which are expected to be available in draft form by the summer 2017. Moreover, the “UNECE Census Week”, with a workshop and an expert meeting, is planned to take place in Geneva in the week 2-6 October 2017.

48. UNFPA informed that they have plans to support countries on data dissemination. One or more experts are expected to received full training on the dissemination software “Redatam”, and then provide support to countries.

49. UNSD presented the main activities promoted to support countries in preparations for the 2020 census round. A revised Handbook on census management is in the final stages of preparation, and will be available in English shortly. A collection of best practices in the use of electronic devices for census is being prepared, that will be presented as a technical report annex to UNSD’s Principles and Recommendations. Two reports on quality control and assessment, and on methodologies for calculating the costs of censuses are also planned and expected to be published during 2017. Some global guidelines on the legal framework for censuses (covering also confidentiality issues) are also envisaged. Initial research shows that some countries do not have a specific law on censuses as distinct from general statistical law. New methodological work on measuring disability in censuses could also be developed.

IV. Conclusions

50. Participants expressed appreciation of the Workshop as an opportunity to discuss census issues that are of specific interest to countries in Eastern Europe, Caucasus and Central Asia.

51. It was proposed to schedule a similar workshop for countries in Eastern Europe, Caucasus and Central Asia back-to-back to the next UNECE-Eurostat expert meeting on censuses in Geneva, in the week 2-6 October 2017. The 2017 workshop could be dedicated to the following topics:

- 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
-