



Economic and Social Council

Distr.: General
30 April 2012

English only

Economic Commission for Europe

Conference of European Statisticians

Sixtieth plenary session

Paris, 6-8 June 2012

Item 3(a) of the provisional agenda

2010 round of censuses – innovations and lessons learned

Lessons learned from the Population and Housing Census in Poland

Note by the Central Statistical Office of Poland

Summary

The paper presents the experience of the Central Statistical Office of Poland in conducting the National Census of Population and Housing in 2011.

Carrying out censuses usually involves high costs and considerable citizen engagement. The paper explains how the Central Statistical Office managed to reduce expenses of census operations and to lower the burden of respondents, while maintaining the high quality of results.

This paper describes the technical and software solutions used for combining multiple data sources, including administrative registers, replacing paper forms with electronic questionnaires and making use of the latest information technology, such as Geographic Information Systems.

I. The National Census of Population and Housing in Poland

1. The cyclical organisation of population and housing censuses is of great importance and significance and censuses need to be carried out in various economic and social conditions. However, regardless of the country's economic potential or the degree of social awareness, census implementation usually involves high costs and considerable citizen engagement. Therefore, the aim is to reduce expenses of census operations, and to lower the burden of respondents, while maintaining the high quality of results.

2. It should be noted that the changing modes of statistical production, not only national but also global, are reflected in the quality and effectiveness of census taking. Furthermore, the continuously raised issue of minimising the burden of respondents is one of the organisational guidelines for the statistical surveys conducted. The census is influenced by innovative forms of information technology (IT) and telecommunication technologies, as well as constantly enhanced alternative sources of information which enable reorganizing the way censuses are carried out.

3. The organisation and implementation of the National Census of Population and Housing 2011 (NSP 2011) in Poland greatly benefitted of the development of the information society evidenced in the course of census taking. The latest census round illustrated that technological innovativeness in statistics is not only possible, but also leads to an improved quality of results, and minimises the burden of respondents and reduces costs.

4. The NSP 2011 was designed and carried out for the first time applying a mixed model, i.e. employing data from administrative registers and data obtained from respondents, exclusively by using electronic questionnaires. As a result, paper questionnaires were eliminated altogether.

5. Use of administrative sources, already existing within the State administration structures was chosen as a starting point for NSP 2011. Such a decision was made for NSP 2011 and it turned out to be one of the key elements of success. Use of the existing and increasingly improved system of registers, synchronised with statistical systems and supported with on-going surveys, enabled the use of the data that had been obtained earlier through other surveys and for record keeping purposes.

6. Census implementation called for integrated statistical processes, efficient cooperation in methodology, infrastructure/IT tools and metadata.

7. Apart from the use of IT systems of public administration, three new data collection channels were introduced (known by a common name of CAxI):

(a) Computer Assisted Internet Interview (CAII) – an online self-administered questionnaire, which entails checking the data obtained from administrative sources by the respondent, within a specified time frame, and, if need be, correcting the same and providing missing information;

(b) Computer Assisted Telephone Interview (CATI) – a computer assisted telephone interview, conducted by a statistical interviewer;

(c) Computer Assisted Personal Interview (CAPI) – an interview conducted by a census enumerator, registered on a hand-held device.

8. All three channels made use of adaptive electronic questionnaires ensuring high quality of data at the collection stage. The electronic questionnaires were filled in and

adjusted by means of CAxI technology. . The questionnaire application (available at a mobile terminal or Internet browser) verified if the questionnaire had been filled in accurately, among other things, through logical and accounting control.

9. Yet, the essential input data source for NSP 2011 was provided by administrative systems.

10. In order to use data coming from administrative sources and non-administrative sources in NSP 2011, statisticians worked hard to make these data suitable for statistics production. To achieve the required quality, they had to perform intensive data editing and integration measures. The most extensive part was the data cleaning process. The data cleaning procedure was based on algorithms developed by programmers and database administrators and included, among other forms of verification, standardization, de-duplication and imputation of data. These methods enabled improvement in data quality and usability of the Census 2011.

11. In accordance with the National Census Acts, all entities maintaining IT systems in the public administration shall deliver data to be used in census operations. The Acts specify the scope and time line for delivering such data.

12. In the framework of NSP 2011, a total of 27 registers maintained by 15 public administration institutions were used. Additionally, 3 non-administrative datasets were also utilized.

13. The reasons for using data from administrative systems in Polish statistics were as follows:

- (a) Economic reasons – demand for effectiveness: minimisation of the costs of statistics production, including administrative burden on respondents;
- (b) The risk of increased non-response in statistical surveys, including censuses;
- (c) An intensive development of IT systems in the public administration, based on advanced technologies.

14. Making use of the administrative systems has brought numerous benefits including, but not limited to:

- (a) Effective use of administrative systems;
- (b) Reduced census costs;
- (c) Reduced social burden connected with data transfer;
- (d) Improvements in data safety;
- (e) Further harmonisation of surveys;
- (f) The availability of census information based on annual registers;
- (g) The availability of data from administrative registers for any level of territorial disaggregation;
- (h) The possibility to identify double entry errors;
- (i) The creation of a micro-database supporting indirect estimation – modelling at the unit level;
- (j) Improvement of the estimation methods for small areas;
- (k) Improvement in the coherence and reliability of statistical data.

15. Data from administrative systems were used in the census:

(a) For compilation related to buildings, dwellings and persons; an address-residence register; a sampling frame;

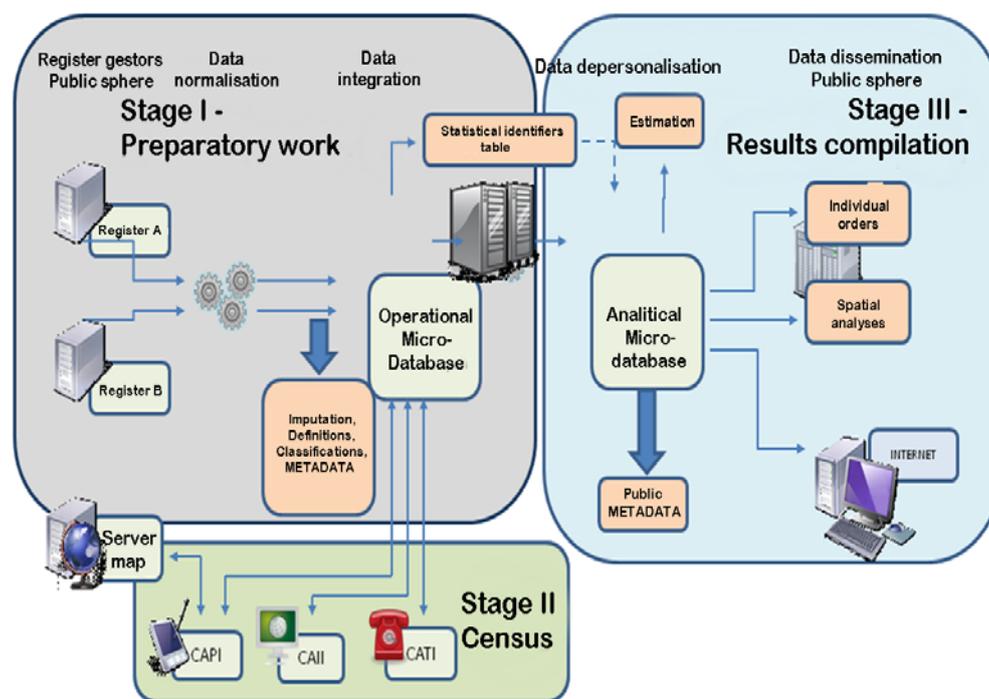
(b) As a direct source of census data (personalisation of questionnaires).

16. To enable the administrators to transfer data from dispersed systems via tele-transmission, the Central Statistical Office of Poland (CSO) constructed an electronic platform for data collection and processing, together with a net-based application for a direct data transfer via electronic means in a secure connection. These solutions were also applied when collecting data on buildings from real estate managers.

17. The register-based census was carried out through a set of IT systems comprising extensive data on the entities included in the census. The systems comprising data on the entire (or almost entire) population included in the census played the key role. In the production of census results, multiple systems were used with various ranks of importance, from the perspective of adequate quality. The results were obtained by transferring the value from an administrative system or estimating based on the largest possible combination of systems, including data collected via the CAXI method.

18. To enable the optimal application of advanced IT and telecommunication technologies in censuses, new kind of census architecture had to be constructed. For the purposes of census design and conduction, the CSO of Poland developed an IT Census System (ISS). The system consisted of more than 10 components and was implemented by different contractors. It provided IT tools for all operations within the census process. The ISS integrated various technologies (from mobile terminal applications, through telephone interview applications, to specialist data bases, data warehouses and analytical and reporting tools).

Architecture of the Information Technology Census System



19. The ISS employed various solutions for ensuring a high level of security for processed census data. Certain organisational means were implemented, obliging everyone taking part in census operations to ensure statistical confidentiality and thus guarantee

protection of personal data. As part of the processing of census data, including data from administrative registers, numerous rules and theoretical solutions were established in the CSO, which were later implemented in practice.

20. The mobile application served as an auxiliary tool for census enumerators. The tool was useful at all regional levels of census operations: at the central level (the CSO), voivodship level (Voivodship Census Offices) and gmina level (Gmina Census Offices). The central dispatcher (a member of the Census Management Centre, responsible for the monitoring and coordination of NSP 2011 at the central level) and a voivodship dispatcher (a voivodship coordinator of the data collection process in the census, responsible for the census monitoring and coordination via the CAPI method, conducted by census enumerators) received assistance in the form of a dispatching application. A similar application was used by Gmina leaders (a member of the Gmina Census Office who provided constant support to census enumerators working in his region. He was equipped with a mobile terminal facilitating communication with the census enumerator; he functioned as the first-level support for the enumerator), and the proper functioning of the applications was overseen by the administrators of the census systems, who worked in every statistical office as well as the CSO.

21. An unquestionable breakthrough was the replacement of paper questionnaires by electronic ones. It enabled a more effective collection of data. Enumerators were provided with mobile hand-held devices equipped with electronic questionnaire applications and digital maps. The application of digital maps and the global positioning satellite (GPS) technologies changed the planning and managing of census operations, both prior to and during the census.

22. Data obtained from the State geodetic, cartographic and ortophotos resources (processed aerial photographs), were combined with statistical data to develop sampling frames for censuses. The frames comprised statistical address points and their spatial reference.

23. Digital maps facilitated the work of census enumerators in moving around the area, verifying the sampling frame, etc. Gmina leaders, and voivodship and central dispatchers, could verify on a map the progress of the census and, for example, the route or location of an enumerator using the Geographic Information Systems (GIS).

24. The introduction of coordinates and address points in the statistical data enabled changing the previous system based on area assignment (census districts) to point assignment. The change allowed for a more flexible grouping of collected data for even the smallest areas. It also facilitated the creation of a spatially-oriented micro database and enabled conducting geo-statistical analyses.

25. In accordance with the organisational principles adopted, pre-census rounds were made prior to launching the census. During the pre-census rounds, census enumerators verified the existence of buildings and supplemented the register with missing address points. Moreover, they examined the area where the census was to take place in order to resolve any possible ambiguities which had arisen during data revision on the gmina level.

26. Mobile terminals were equipped with the GIS application, which enabled revisions and showed on the map, among other things, the current location of the census enumerator (GPS) and address points assigned to him. Using the mobile application, the enumerators could change the location of an address point, delete an address point, or add an address point not included in the register. During the pre-census stage, the enumerators were also responsible for controlling the entire area of the census districts assigned to them. It was particularly important when the revision in the gmina district was performed only on the basis of the registers kept, and the census enumerator was the first, and in many cases the only, person directly involved in the field work. The GIS application was also actively used

during the census – to manage its course. It enabled monitoring and control of the enumerators' work, as well as tracing their movement in the area (among other things, to ensure safety).

II. Operational micro database 2011

27. In accordance with the Act on NSP 2011, a database known as Operational Micro Database was developed in the CSO. The system comprised an equipment-system-tool infrastructure (hardware, system software, computer tools) and application software (computer programs developed through programming).

28. The database allowed both the inclusion of data transferred in an electronic format by the entities obliged under the Act to do so, including entities maintaining IT systems. Data on persons collected by means of the Internet, telephone interviews conducted by statistical interviewers or direct interviews conducted by census enumerators were stored in the data base. It also provided a platform for further data processing.

III. Census meta-information sub-system

29. The Meta-Information Sub-System stored the necessary metadata describing data and census operations, including any operations necessary to prepare quality reports.

IV. Analytical micro database

30. An Analytical Micro-Data Base (AMB) was also created to store impersonalised census data in its final version. The AMB is currently used, among other things, to conduct statistical analyses to be disseminated to the public. The AMB enables a prompt production of aggregated data. It comprises data in the field of population migration, the labour market, work commuting, education, disabilities, nationality and religion, as well as housing.

31. It should be noted that the effectiveness of census implementation was due to both the methodological as well as the organisational and logistic preparations.

32. The time frame proved right, since the deadlines set for the data collection stage (specified in the Act on NSP) did not have to be extended, and neither did the census budget. The detailed schedule for the implementation of NSP 2011 was regularly updated. The framework schedule comprised over 250 items. The detailed schedules for tasks, for example for carrying out a control census and providing support for the census systems, were kept in separate files. These schedules comprised a total of several thousand tasks.

33. The census frame was organized according to a hierarchical structure, corresponding to the territorial division of the country and the systematics of data collected in NSP 2011. It consisted of the central level (the CSO), voivodship level (Voivodship Census Offices) and gmina level (Gmina Census Offices). Some parts of the census structure were permanent and some of the census frame (voivodship and gmina levels) was established only for the time of preparation and implementation of NSP 2011 in the field.

34. A comprehensive analysis of census conduction, accounting for all its participants including the members of the field census frame, allows one to draw conclusions about further implementation of new technology. The technology was easily applicable to questionnaire-based surveys. It makes the process less expensive, employs up-to-date control mechanisms, enhances the quality of the material collected, and, in consequence, reduces the burden of respondents.

35. Replacing the paper forms by electronic questionnaires and introducing four new data collection channels reduced census costs in Poland. Higher costs would have incurred in a traditionally conducted census.

36. Successful use of advanced data collection methods led to a decision in the CSO of Poland to systematically implement the architecture created for the census in other questionnaire-based statistical surveys in Poland.

V. Lessons and suggestions

37. There is another round of censuses ahead. Thanks to the current experience, in the next round of censuses Poland is going to use newer technology making it more effective and innovative, which is still delivered and modified by the IT world.

38. Until then, census methods should be developed using the experience gathered in 2011 as a starting point. Considerable efforts are needed in developing a new census strategy, so as to guarantee further improvements. Attempts should be made at:

- (a) Reducing census costs;
- (b) Using administrative sources in an effective way;
- (c) Reducing social burdens connected with data transfer;
- (d) Improving the safety of transferred data;
- (e) Improving the coherence and reliability of statistical data.

39. Poland will further develop the census methods used during NSP 2011. The CSO will pay attention to the technical, software and telecommunication development, but first of all, will assess the legal aspects on census data collection and protection of confidentiality (especially in the context of access to individual data). Methodological development in linking data coming from different sources and from different surveys (full-scale and sample surveys) will be essential.

40. More information on the web: e-census.stat.gov.pl
