Economic and Social Council

Economic Commission for Europe

Conference of European Statisticians

Group of Experts on Population and Housing Censuses

Twenty-first Meeting
Geneva, 18–20 September 2019
Item 6 of the provisional agenda
Dissemination

The multiannual dissemination programme in the Italian Permanent Census of Population: towards more timely statistics

Note by Italian National Institute of Statistics (ISTAT)*

Summary

> Italy is moving towards a new census strategy, integrating information arising from registers and recurring sample surveys. One of the main targets of such an approach is the timeliness of the data: the delay in release of census data reduces their value and usefulness for many stakeholders. Therefore, a new Italian dissemination programme has been carried out to ensure a subset of census statistics every year. The contents can evolve over time as user needs change and the availability of administrative data improves. Annual data dissemination will be supplemented by a larger decennial publication which will occur for the reference year 2021, the same as is defined at European level.

---

* Prepared by Simona Mastroluca and Mariangela Verrascina.

Please recycle

I. Introduction

1.   The 2011 Italian General Population and Housing Census was the last one carried out in a traditional way—despite the fact that many innovations, such as the mailing-out of questionnaires collected by a multimode system which included Internet, and the short and long forms delivered in the biggest municipalities, were planned to achieve a stable and enduring balance between census costs and benefits.

2.   Considering the huge amount of data from administrative sources now available, the reduced budget allocated for the enumerations and a growing need for more frequent census-like data, Italy has designed an innovative census strategy.

3.   The new Italian Population and Housing Census, which will no longer be decennial but permanent, is based on the Census and Social Surveys Integrated System (CSSIS), a complex statistical process integrating the information arising from registers and surveys. It foresees a two-phased master sample design consisting of a set of coordinated annual sampling surveys. This is planned in order to support the Italian National Institute of Statistics (ISTAT) registers, to increase the amount of statistical information provided and to improve the level of coverage and quality.

4.   The first phase of master sample is based on two different component samples, namely A (Areal) and L (List). Component A aims to satisfy the need to estimate under-coverage and over-coverage rates of the population register. Component L, based on a list sample, has the purpose of thematic integration: that is, estimating the hypercubes which cannot be obtained using information coming from registers.

5.   Every year the A and the L components involve almost 1,500 thousand households and 2,800 municipalities with different population sizes.

6.   The first wave of the new census project took place last October; the next one is foreseen for October 2019.

7.   One of the main goals of the census is the timeliness of the information: decennial data dissemination will be supplemented by an annual dissemination programme in order to meet stakeholders' requirements. The content of the annual collection can continue to evolve as user needs change and the availability of administrative data improves. Over time, the distinction between the annual and decennial data collection is likely to reduce. Some of those data that might now be considered only for the decennial collection may, in the future, be feasibly disseminated every year or at an intermediate frequency.

II. The information content of the Italian Permanent Census

8.   Traditionally the census offers a huge amount of detailed information covering demographic and socio-economic topics. In the new Italian strategy, variables have been classified as totally, partially or not replaceable.

9.   The first group includes variables for which the administrative sources provide the correspondent proxy information. They are considered complete, because they are available for all units in the thematic registers, and accurate, having a good level of coverage and quality.

10.   Partially replaceable variables are considered complete and accurate only for a subset of the target population; for the others, the data are unknown or cannot be considered accurate. This is, for example, the case for the educational attainment. The Italian Ministry of Education does not provide information on post-secondary non-tertiary education (level 4 International Standard Classification of Education (ISCED) 2011) or degrees obtained abroad.

11.   The last group (non-replaceable variables) contains topics not yet available from administrative sources.

12. Therefore, the Italian Permanent Census is a combined census using both the information produced by the statistical registers, for replaceable variables, and the surveys for the remaining topics.

13. For Population and Housing Census purposes, at the moment three statistical registers are under construction: the population register, containing the census variables sex, age, place of birth, legal marital status and citizenship; the employment register with information concerning employment, status in employment, industry, full time and part time employment, fixed term contracts; and the register of buildings, which produces data on buildings and dwellings.

14. The questionnaire arranged for the first phase of the master sample is the same for both the A and L components and includes not only partially or non-replaceable variables, but all the necessary questions to produce hypercubes with the purpose of also using the collected information to test quality and coverage of data already available in registers.

15. The Italian Permanent Census collects variables to be provided to the European Commission (EC, Eurostat); a selection of non-core topics listed within the Recommendations of the United Nations Economic Commission for Europe (UNECE); and data necessary to ensure time series and essential for supplying key information that meets national and local user requirements.

16. In detail, the census topics cover demographic, economic and educational characteristics of persons, internal and international migrations, commuting, households, families and housing.

17. The enumeration form includes about 20 questions on dwellings and buildings and more than 50 variables for each component of the household. Of course, persons must answer a subset of questions in relation to their age, educational attainment, school attendance, current activity status, or migration features that determine the paths of the questionnaire.

18. The first phase of master sample is totally paperless with self-completed web questionnaires (computer-assisted web interviews, CAWI) or computer-assisted personal interviews (CAPI). The electronic questionnaire only provides respondents with pertinent questions, making it easier to fill in.

19. Compared with 2011, some variables, such as the relationship to the household's reference person or the legal marital status, have been modified to ensure increased quality of data (for example now the question on the relationship to the household's reference person has 23 categories instead of 17, to identify family nuclei with a higher level of precision) and to ensure harmonization with social surveys. A few questions relating to variables that were little used by stakeholders have been removed, while others have been added to investigate evolving phenomena.

20. The choice of topics was shared with municipal representatives in order to support social and environmental policies at the local level. In fact, from a statistical point of view, the census is currently the only source able to provide the detail needed for a "functional" study of the territory. Only census data allow to observe the social and economic differences within a restricted zone, analysing small areas.

21. The process of editing and imputation of the data related to the 2018 A and L components is ongoing. The analyses carried out up to now are encouraging and suggest that the information collected is of good quality.

22. Of course, over the years, as a result of the developing availability of high-quality administrative data, some questions will be removed to leave room for other variables not included in the registers but fundamental for the exploration of topics of national interest.

III. Census data dissemination between tradition and innovation

23. The dissemination system of the 2001 census data was innovative compared to the past due to the availability of data via the Internet. In addition to traditional paper

publications, a data warehouse[1] was created, designed to make statistical tables accessible online, permitting users to navigate the tables in a flexible manner and to choose independently the classification levels and regional detail desired.

24.   There was also an interactive cartography system that allowed the display of thematic maps for some tables and a section with data by enumeration areas.

25.   In order to satisfy particular users' needs, requests for information not disseminated online and not present in the published volumes, neither on the computer support attached to them nor among the other products prepared, have been processed through ad hoc queries by the Institute.

26.   During the process of production of the 2001 Italian census data, the preparation of the cross-tabulations for European Union dissemination started. Within the *"Table Programme for the Community Programme of Population and Housing Censuses in 2001"*, Eurostat requested each European Union country to supply a set of 40 statistical tables, at national and municipal level. Later, two more tables were requested at the level of regional detail of urban areas.[2]

27.   The cross-tabulations contained data according to detailed breakdowns, which, in almost all cases, differed from the classifications used at national level. It was therefore necessary to calculate new breakdowns by transforming existing ones. In some cases, the information had not been collected and, in others, the information had been collected differently. Consequently, a complete reconstruction of the required tables was not possible. The difficulties encountered by many countries in the production of data requested by Eurostat, and the differences in definitions, classifications and methods of collection have not always guaranteed the comparability of data.

28.   For the 2001 census, European Union Member States had signed a Gentlemen's Agreement; an experience that, as described above, did not guarantee the comparability of the data. Therefore, for the 2010–2011 census round, the European Commission issued Regulations on population and housing censuses.

29.   Thus, for the first time, the European legislation defined a set of harmonized data. According to an output-oriented approach, the European data needed to fully meet the standards in the legislation (definitions, breakdowns, cross-tabulations and quality reporting) although countries were free to produce data using any appropriate methods and data sources, among those listed in the Framework Regulation (No. 763/2008), as available and preferred in each country.

30.   The Implementing Regulation concerning the dissemination programme of data and metadata (No. 519/2010), provided the list of cross-tabulations which, to various degrees of regional detail (from national to municipal) and at different classification levels, had to be validated and made available. Each Member State prepared a total of 60 hypercubes,[3] corresponding to 175 principal marginal distributions. The data dissemination programme focused mainly on regional detail. The Regulation on hypercubes also defined the textual metadata that should be supplied for the census topics. Data were produced in the form of detailed multi-dimensional cross-tabulations (hypercubes) and accessible via the online Census Hub system, allowing users the flexibility to define tables according to their needs.

31.   The dissemination programme defined by Eurostat did not correspond to the Italian programme. In fact, the hypercubes listed in Regulation No. 519/2010 differed from the cross-tabulations characterizing the Italian dissemination in terms of information (the number of variables requested was smaller than that foreseen in the survey plan of the Italian census), classifications and/or definitions (in some cases Eurostat adopted classifications and/or definitions different from those traditionally used in Italy), classification details (Regulation

---

[1] http://dawinci.istat.it/.

[2] The concept of urban area does not coincide with that of municipality nor with that of the sub-municipal area, but regardless of municipal boundaries. Urban areas had been identified within NUTS2 boundaries; these were areas defined as aggregations of enumeration areas.

[3] 5 at national level, 36 at NUTS2 level, 10 at NUTS3, 5 at LAU2 and 4 per place of work.

on breakdowns and technical specifications No. 1201/2009 provided different levels of classification detail) and regional level.

32.   For the Italian dissemination it was necessary to satisfy the information needs of the national users and to guarantee the time series.

33.   Consequently, for the 2011 census, two different dissemination programmes were produced, one defined on the basis of the European Regulations, the other with tables crossing all the variables contemplated in the Italian survey plan, which, in addition to ensuring continuity with the past, permitted dissemination of information on phenomena of national interest (for example, mobility at municipal level) or investigated for the first time during the census in Italy (for example, the place of birth of parents, variable not mandatory but suggested in the UNECE Recommendations).

34.   There are more than 300 tables. To guarantee the protection of confidentiality, the greater the regional detail and/or the number of variables included in a cross-tabulation, the lower the classification detail.

35.   No paper publications have been foreseen for the census results: the dissemination took place only via the web, through the corporate data warehouse (I.Stat),[4] specifically designed to make statistical tables of aggregated data accessible and navigable on the web. The data are presented in the form of multidimensional tables that users can export.

36.   Data were then disseminated for enumeration areas, thematic maps compared with the previous censuses and data in open format (Linked Open Data).

IV.   The Italian dissemination programme for the Permanent Census

A.   The decennial and annual data collections: topics and cross-tabulations

37.   As already pointed out, the Italian dissemination programme is designed considering the requirements of European Union Census Regulations, to guarantee the time series, to satisfy users' needs, to provide data on phenomena of national interest.

38.   Furthermore, the new strategy, combining data from registers and from annual surveys, ensures not only the traditional decennial collection, but also a yearly availability of detailed census statistics. Therefore, ISTAT has prepared two different proposals for data dissemination: an extended programme and a reduced subset of hypercubes.

39.   The extended dissemination programme, in line with the European Union, will have 2021 as its reference year. This suggestion arises from a critical evaluation of the 2011 Italian programme (complexity of the tables, redundancies, outdated phenomena, requests from stakeholders, etc.), and from the analysis of the use of the information available on I.Stat, and on other dedicated websites and the results of research on information content, tools and dissemination strategies adopted in some foreign countries.

40.   The ten-year programme has been redefined, i.e. cross-tabulations at the regional level have been removed, preferring tables with a municipal or provincial detail; particularly complex cross-tabulations have been simplified but, at the same time, new hypercubes have been introduced to study emerging social phenomena.

41.   The draft decennial programme includes around 250 tables, structured around various thematic areas and with different classification and regional details (over 70 tables at local administrative unit (LAU) 2 level) that, between tradition and innovation, ensure continuity with the past, focusing on topics of particular interest (seniors, children, foreigners, Not in Employment, Education or Training (NEET), housing arrangements of the elderly, etc.).

42.   The thematic areas considered are: Housing, with cross-tabulations on type of living quarters, tenure status, occupancy status of conventional dwellings, type of ownership,

---

[4] http://dati-censimentopopolazione.istat.it/.

number of occupants and other characteristics of conventional dwellings; Demographic features (sex, age, legal marital status, etc.); Presence and Residence one year prior to the Census; International migrations (country of birth, citizenship, citizenship acquisition, country of birth of parents, ever resided abroad and year of arrival in the country, etc.); Education (including school attendance); Economic characteristics; Commuting; plus tables on Household and Family features.

43. The analysis of the demographic changes taking place in Italy (e.g. the ageing of the population) and the structural transformations of households and families has led to new proposals for dissemination that are able to grasp and study these issues. New cross-tabulations have been suggested, both in terms of new statistical units and classifications.

44. As far as education topics are concerned, the adjustments made to the decennial dissemination programme are affected by the changes in the Italian education system during the intercensal decade and by the results relating to the use of the data published on the occasion of the 2011 census. New tables have been designed regarding foreigners and Italian citizens by acquisition, classified by educational attainment, sex, age, place of birth (Italy or abroad) and citizenship.

45. On the subject of foreigners, additional hypercubes have been included on the age of arrival in Italy and on the length of stay in the country, information which is inferable from the variable "ever resided abroad and year of arrival". Further tables concern households with at least one foreign component or with all foreign components. Others, such as those on citizenship already present in the 2011 dissemination programme, have been detailed at LAU2 level as requested by experts in international migration statistics.

46. In relation to the labour market, some cross-tabulations at Nomenclature of Territorial Units for Statistics (NUTS) 3 level were removed as they are already available from the Labour Force Survey (LFS), while it has not been foreseen to focus on persons neither employed nor unemployed, nor on NEETs.

47. Finally, the information supply on housing topics and on the relationship between the population and the living quarters has been reduced. Some variables no longer will be collected because of low demand (e.g. rooms reserved for professional activities or presence and type of kitchen) or because the coverage of some characteristics of conventional dwellings (e.g. water supply system and toilet facilities) in 2011 was close to 100 per cent and, consistent with the European Union Implementing Regulation No. 543/2017, may not be included in the data collection. Moreover, some classifications have been simplified to allow a better interpretation of the data.

48. The next Italian census decennial data will be made available starting from the end of 2023. As usual, the dissemination will be by thematic blocks and it is likely to be completed in 2024.

49. Considering the information included in the registers and the annual survey (master sample) on almost one and a half million households, a minimum set of tables at LAU2 level and possibly sub-municipal (enumeration areas or groups of enumeration areas) for municipalities over a certain demographic threshold has been developed (31 hypercubes, reconfigurable over time) to be disseminated every year starting from 2019. The planned municipal cubes concern 15 cross-tabulations on population and 16 cross-tabulations on households and family nuclei: population structure by sex, age and citizenship, internal migrations, educational attainment, current activity status with a focus on persons outside the labour force, commuting, type of family nucleus and type of private household.

50. The possibility of a uniform base dissemination for all the municipalities and of an additional dissemination depending on the population size is under study. This choice is strictly linked to the sample design (number of households) that the Institute intends to adopt for the master sample in the coming years.

51. The annual collection of census data will evolve over time depending on the incremental availability of administrative data integrated in the registers.

52. From a further analysis carried out on the draft ten-year national dissemination programme proposed for 2021, an additional set of tables to be released to the provincial (and

big cities) level of detail has been identified, according to a periodicity to be defined. These tables contain variables not included in the annual proposal or more detailed classifications than those foreseen for the census statistics at municipal level.

B.   The decennial and annual data collections: dissemination tools

53.   A survey was carried out on what was done in other countries at the national level, in addition to the production of hypercubes under the EU Regulation, with the aim of verifying whether other European contexts had the need to "split" the dissemination, as happened in Italy. Therefore, some foreign countries with different census approaches have been taken into account to explore their 2011 national dissemination of data from population and housing censuses, based on the technique adopted, and to compare it with the Italian census production.

54.   The countries analyzed were: the Nordic countries (Denmark, Norway, Finland, Sweden), Austria, Slovenia and the Netherlands for the register-based census; Spain and Germany for the combined census; the United Kingdom (England and Wales) and France[5] for the traditional census.

55.   In addition to making data available at the European level through the Census Hub, not all countries disseminate national census output. Countries producing census outputs at national level (Norway, Finland, Austria, Spain, Germany, United Kingdom -England and Wales-, France) have numerous and different ways of approaching dissemination.

56.   On the side of information content, excluding information not usually investigated in the census survey in Italy (race, ethnicity, religion, health), few different themes or variables have been identified in the other countries. In general, information traditionally included in the Italian survey plan can be aggregated and disseminated in a different way. There is no new content but new cross-tabulations that could be included in the future Italian dissemination programme and/or topics and classifications that could be analyzed in the context of replaceability (from other surveys or administrative sources).

57.   The analysis of the census data dissemination strategies highlights interesting and diverse methods and tools adopted, alternative to those used in Italy, which represent an opportunity for the improvement of dissemination. Some of these approaches could be adopted for future censuses.

58.   The starting point is the preparation of predefined tables so that the dissemination is made up of tables defined a priori. Then, dissemination differentiated by population size of municipalities should be taken into account, introducing the possibility of publishing tables only for municipalities over a given size, and/or tables with more or less detailed breakdowns according to the population size of municipalities.

59.   Dissemination in tabular form will be integrated with new tools for presenting the results or with the enhancement of tools already used for the 2011 Italian census, or already used in the ISTAT but for the first time to be adopted for the census. In terms of data representation, therefore, ISTAT is considering the possibility of using advanced technologies for the representation of census output and new tools for presenting the results of the census.

60.   Data visualization tools allow transformation of tables of numbers (sometimes difficult for users to understand) into visual tables that can capture attention and explain a phenomenon. It could be useful to complete the dissemination with data representation, meaning using maps and cartography more effectively than in the past.

---

[5] For simplicity, the term "traditional census" includes here the rolling census.

61.   The aim is to increase the representation of data through cartographic tools, i.e. to enhance the use of maps at municipal and sub-municipal detail, where possible, and for 1 km² grids.[6]

62.   Moreover, the focus will be on infographics, i.e. graphical representations of the main results of the analysis, as a synthesis tool for the presentation of data, differentiated according to the content and the channel to be used.

63.   Furthermore, it could be necessary to permit customized analyses, i.e. to create a system that permits customized queries in cases where users cannot find the information they are looking for in the predefined tables, or to prepare a data warehouse that lets users freely build cross-tabulations of variables and classifications according to their needs.

64.   Finally, it will be necessary to implement a system for monitoring the use of data. The information resulting from the monitoring could, in fact, outline the profiles of users who access the data and, for each profile, identify the tables and themes of greatest interest. This would also allow modification of the dissemination "on the fly" according to the aspects more visualized and less visualized and depending on the profile of the users.

65.   By identifying the profiles of the different user segments, customized products and services could be offered and a multi-channel dissemination strategy could be adopted.

V.   Similarities between the future EU censuses and the Italian case

66.   The key aspects of the post-2021 census strategy in the EU are summarized below.

67.   A "complete" ten-year census is expected but supported by annual updates on the relevant topics, mainly on demographic and migration variables. Annual data dissemination should be based essentially on administrative sources, to be introduced in the mid-2020s, while a larger decennial dissemination of results should first be done in 2031.

68.   The annual updates should be disseminated through simpler cross-tabulations than the decennial hypercubes; classifications could be more aggregated, but a greater degree of regional detail must be guaranteed. Compared to the 2011 programme which focused largely on NUTS 2 regions, the annual census results should be on a more detailed regional level (LAU 2 or NUTS 3), and a few data (in particular the annual population count) for a geographical grid (1 km²).

69.   Flexibility must also be ensured, with the possibility of adding information and cross-tabulations from year to year, as high-quality administrative data become available, also taking into account stakeholders' priorities, users' needs and costs.

70.   Another goal is to reduce the time required to disseminate data, so, thanks to the availability of annual updates, it will be possible to achieve a more frequent and timelier census data collection.

71.   Finally (last but not least), to avoid duplication of results and work, efforts will be made to integrate census information with information from social household surveys (e.g. LFS) for the production of combined statistics.

72.   Thus, as described in the previous section, the main features of the new European approach are close to the current strategy of the Italian Permanent Census.

73.   At the European level, for the 2021 census, many countries have already moved away or are moving away from traditional census. Thereafter, for the post-2021, an increased use of data from administrative sources is expected. Annual hypercubes, starting with a core set of demographic and migration topics, are foreseen starting from the mid-2020s. Annual data

---

[6] The characteristic of the grid squares is that they are more stable over time than, for example, municipality boundaries (which is why they are also being invested in heavily at the European level), therefore ensuring comparability over time and space, overcoming the concept of demographic size of municipalities.

dissemination will be supplemented by a larger decennial data dissemination in the form of the recurring decennial census first to be conducted in 2031.

74.   As mentioned above, in Italy, the 2011 census was the last traditional one. The 2021 census will be combined (register-based and sample surveys). A set of 31 tables at municipal level reconfigurable over time will be disseminated starting from 2019 on an annual basis. The decennial dissemination programme (in line with the European Union) will have as 2021 as its reference year.

VI. Conclusion

75.   Thanks to the new Italian census strategy (a combined census based on three registers —the population register, the employment register and the register of buildings — and on sample surveys) a subset of census statistics will be disseminated every year starting from 2019. A core set of demographic, socio-economic and migration topics with limited cross-tabulations and less detailed disaggregations will be promptly made available to stakeholders. The yearly dissemination will be supplemented by a larger decennial publication, for the reference year 2021 (the same as is defined at the European level).

76.   At the same time in Europe the use of data from administrative sources is increasing, and Eurostat is working to define a strategy for post-2021 censuses based on a more frequent data supply that provides relevant and timely census results for data users. The purpose is a multiannual data collection, essentially relying on administrative data sources starting from the mid-2020s, maintaining the traditional decennial data collection that will be done in 2031.

77.   This means that Italy is trying to realize, a few years in advance, the strategy that will be applied at the European level for the post-2021 censuses. This is a big challenge for a country that has to face many innovations in a very short time in order to produce high-quality regional annual census data.

References

Regulation (EC) No 763/2008 of the European Parliament and of the Council of 9 July 2008 on population and housing censuses**,** *Official Journal L 218, 13/08/2008.*

Commission Regulation (EC) No 1201/2009 of 30 November 2009 regards the technical specifications of the topics and of their breakdowns.

Commission Regulation (EU) No 519/2010 of 16 June 2010 adopting the programme of the statistical data and of the metadata for population and housing censuses.

Commission Implementing Regulation (EU) 2017/543 of 22 March 2017 regards the technical specifications of the topics and of their breakdowns.

Commission Regulation (EU) 2017/712 of 20 April 2017 establishing the reference year and the programme of the statistical data and metadata for population and housing censuses.

United Nations Economic Commission for Europe and Statistical Office of the European Communities. 1998. "Recommendations for the 2000 Censuses of Population and Housing in the ECE Region". *Statistical Standards and Studies,* N. 49.

United Nations Economic Commission for Europe and Statistical Office of the European Communities. 2006. Recommendations for the 2010 Censuses of Population and Housing. ECE/CES/STAT/NONE/2006/4, Geneva.

United Nations Economic Commission for Europe, 2015. *Conference of European Statisticians Recommendations for the 2020 Censuses of Population and Housing*, ECE/CES/41, New York and Geneva.

Dardanelli S., Sasso A., Verrascina M., *La diffusione dei dati del censimento della popolazione e delle abitazioni del 2011 alla luce di alcune novità introdotte a livello nazionale e internazionale*, XXXI Conferenza Italiana di Scienze Regionali, Aosta, 20–22 Settembre 2010.

Eurostat, *Future EU census programmes: beyond 2021*. Conference of European Statisticians – Group of Experts on Population and Housing Censuses, Geneva, 4–6 October 2017.

Falorsi S., *Census and Social Surveys Integrated System.* Conference of European Statisticians – Group of Experts on Population and Housing Censuses, Geneva, 4–6 October 2017.

Istat, *Preliminary Experimental: Results on the Italian Population and Housing Census Estimation Methods.* Conference of European Statisticians – Group of Experts on Population and Housing Censuses, Geneva, 26–28 September 2018.

Mastroluca S. (ed.) *I contenuti informativi della rilevazione, la validazione e diffusione dei dati*, Istat 2017.

Mastroluca S., Verrascina M.*, Towards more timely census statistics: the new Italian multiannual dissemination programme.* Book of Short Papers SIS 2018, Palermo, 20–22 June 2018.

Verrascina M., *La diffusione dei dati censuari della popolazione in alcuni stati esteri: contenuti e strategie*, Istat Working Paper (*forthcoming*).

_____