



**Economic and Social
Council**

Distr.
GENERAL

ECE/CES/GE.41/2009/5
17 August 2009

Original: ENGLISH

ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Group of Experts on Population and Housing Censuses

Twelfth Meeting

Geneva, 28-30 October 2009

Item 2 of the provisional agenda

**COMPLIANCE WITH CONFERENCE OF EUROPEAN STATISTICIANS
RECOMMENDATIONS**

**A new strategy for the 2011 Italian Population Census. Product innovations and the
compliance with CES Recommendations**

Note by the National Institute of Statistics, Italy

I. INTRODUCTION

1. A decennial population and housing census has been taken in Italy since 1861, based on the conventional methodology of complete field enumeration. All information was collected and processed on a complete basis. Only a long form was traditionally used and all the information collected was processed without making use of any sampling techniques.

2. For the 2011 census, a number of factors raised questions about the appropriateness of continuing to rely on the conventional methodology. A major change that will apply only to the largest municipalities will be the use of sampling techniques for the collection of socio-economic information. This strategy implies a shift towards the production of estimations concerning the socio-economic set of variables included in EU Regulation. These estimations will be produced by using a long form for a sample of households. Two enumeration forms will be used in municipalities with 20,000 or more inhabitants (and in all provincial capitals): a short form, in order to collect data on the demographic features of the whole resident population, and a long

form, designed for a sample of households and including demographic topics as well as socio-economic variables.

3. The new strategy will be tested in a pilot survey which will be held in October 2009. Some product innovations will be tested as well, namely: the adoption of the household relationship matrix; the collection of information on the country of birth of parents; the collection of information on economic characteristics for the whole economically active population (including unemployed having ever worked before).

II. THE 2011 POPULATION CENSUS STRATEGY

4. A decennial population and housing census has been taken in Italy since 1861, based on the conventional methodology of complete field enumeration. Census forms were delivered and collected by enumerators and self-filled in by respondents. All information was collected and processed on a complete basis. Only a long form was traditionally used and all the information collected was processed without making use of any sampling techniques.

5. In Italy, besides the obvious main goals of determining the legal population and collecting information on the main demographic and socio-economic characteristics of the usually resident population, there is a third main goal to be achieved through census enumeration: the update of municipal population registers (*Anagrafi*) on the basis of the comparison between census data and population registers data. Indeed, the law on population registers (*Regolamento Anagrafico*) provides for the update of population registers on the basis of census results.

6. Italian National Institute of Statistics (Istat) has responsibility for the enumeration and directly performs design and coordination tasks whereas the actual fieldwork is conducted by the municipalities that, in most cases, contribute with their own financial resources to the state allocated budget. They also carry out coverage checks by comparing data coming from the fieldwork and population registers data, perform the first quantitative and qualitative revision of questionnaires and, finally, revise population registers on the basis of census results.

7. The last census was taken in 2001. 22 million private households and more than 30,000 institutional households, accounting for a total of approximately 57 million persons, were enumerated by 100,000 enumerators and 10,000 co-ordinators, organised in a network of 8,101 Municipal Census Offices (MCO) and 103 Provincial Census Offices (PCO).

8. For the 2011 census, a number of factors raised questions about the appropriateness of continuing to rely on the conventional methodology. Among these factors were: a) the huge organizational effort imposed on municipalities, traditionally exposed to a sudden and time-concentrated increase of workload; b) the increasing difficulty by enumerators of finding people at home due to changes in the population life-style and in the population structure (e.g. growing percentage of one-person households or of the so-called *dink* – double income no kids – couples), especially in larger municipalities; c) an increasing concern for confidentiality.

9. In order to identify the main critical points of the fieldwork organization, a number of studies have been conducted during the past years which have allowed the setting up of a new strategy. The need to reduce municipalities' workload and the burden on respondents has called

for a larger use of administrative data and of new collection techniques. Main features of the census enumeration will continue to be the completeness and simultaneity of the population count but the fieldwork will be guided by registers and supported by the use of new data collection techniques and new territorial instruments designed to improve coverage and quality of the enumeration.

10. The historical *register-improved* census will thus become a *register-supported* census, implemented by means of questionnaires' *mail out* to the households registered into the municipal population registers. Self-completed questionnaires will then be collected by a mixed mode system which will include Internet, *mail back*, delivery to Municipal Collection Centres and, finally, intervention on late respondents by enumerators.

11. A major issue concerning the new strategy is represented by over-coverage and under-coverage list errors affecting a *register-supported* enumeration (i.e. a number of households included in the municipality population register might be no longer residing in the municipality and, conversely, a number of households actually residing on the municipality territory might not be included in the population register). Over-coverage will be 'naturally' corrected by a field enumeration that relies on questionnaires *mail out* to units included in the register while specific measures are required in order to manage potential under-coverage. Persons not included in registers but actually residing in the municipality will be enumerated according to one of two possible approaches that are currently being studied in order to make the final choice. The first approach considers the use of traditional fieldwork assisted by auxiliary sources information. The second approach is based on a sample survey to be carried out during the standard data collection phase with the purpose of estimating the number of units missed by the registers by means of a *capture-recapture* analysis.

12. Another basic feature of the new strategy is its modular nature. The necessity to differentiate the census organization according to the needs and capacities of the different actors has required a strategy consisting of a set of modules to be applied flexibly according to the size of the municipality. The organizational impact of census operations is strongly dependent on the municipality population size. Information from the 2001 census monitoring system indicated the biggest municipalities had the largest difficulties in meeting the field operations' deadlines while small municipalities struggled to cope with financial problems. For this reason, municipalities have been divided in four size categories and a different combination of modules has been planned for each of them.

13. A major change which relates only to the largest municipalities will be the shift towards the production of estimations concerning the socio-economic set of variables included in EU Regulation. These estimations will be produced by using a *long form* on a sample of households. Data will be significant at a census area (grouping of contiguous and homogeneous enumeration areas) level. There are five main modules planned for the new census strategy (see Table).

14. Pre-census addresses' enumeration aims to support the census by producing auxiliary information to be used to limit undercounting. The additional costs required by the setting up of an address list for the smallest municipalities compared to the corresponding advantages in terms of accuracy and quality of the count has restricted the planning this operation to the largest municipalities.

15. Similarly, census areas and the *long form* sample enumeration will be applied only to major municipalities. The setting up of census areas is related to the need of producing census data representative at the sub-municipal level while allowing at the same time a consistent reduction of respondent burden by means of a sample strategy. In contrast, for the smallest municipalities (C₂), the sample approach would imply a reduction of response burden not worth the corresponding information loss. Finally, with reference to the intermediate class C₁, a sample strategy could be readily applied only by producing estimates representative at the municipal level, thus implying significant information loss. The use of pre-census lists for supporting field-enumeration and the revision of population registers will thus be applied independently of the size.

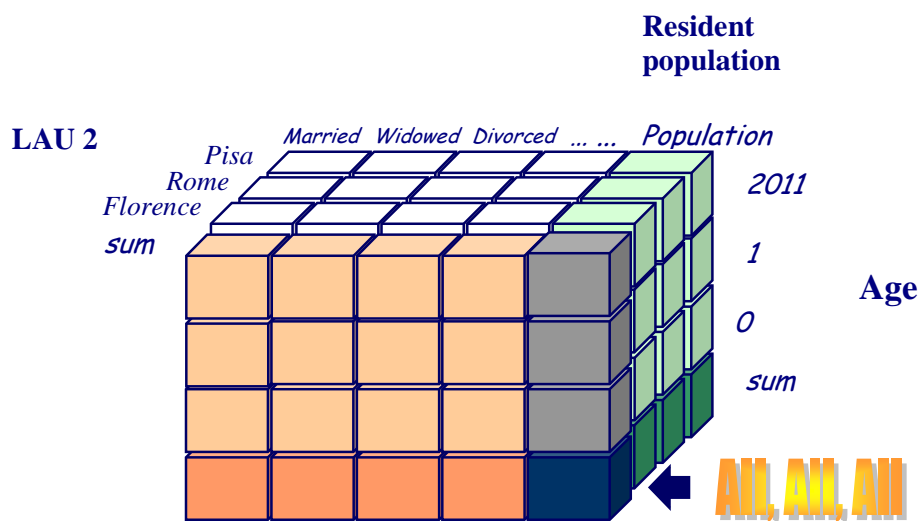
Table
Enumeration strategy by municipality demographic size

Module	Type A (150,000 and over)	Type B (50,000 – 150,000)	Type C ₁ (20,000 – 50,000)	Type C ₂ (< 20,000)
Pre-census addresses' enumeration	X	X		
Setting up of census areas	X	X		
Use of pre-census lists derived from population registers	X	X	X	X
Use of samples of households for <i>long form</i> enumeration	X	X	X	
Revision of population registers	X	X	X	X

III. THE EUROPEAN UNION DISSEMINATION PROGRAMME AND THE SHORT/LONG FORM STRATEGY

16. As already mentioned, one of the main features of the 2011 Italian Census strategy will be the use of sampling techniques for the collection of socio-economic information. Two enumeration forms will be used in the municipalities with 20,000 or more inhabitants (and in all provincial capitals): a short form, in order to collect data on the demographic features of the whole resident population and a long form, designed for a sample of households and including demographic topics as well as socio-economic variables.

Figure 1: Example of Hypercube



17. Attached to the Framework Regulation of the European Parliament and of the Council is a list of obligatory topics (core topics) to be covered in the population and housing censuses, separately for the geographical levels NUTS 3 and LAU 2 (in Italy the level corresponding to municipalities) and for the geographical levels NUTS 1 and NUTS 2.
18. There will be also two Implementing Regulations currently under preparation by Eurostat with the support of the Task Force on the Implementation of Population and Housing Censuses in the EU: 1) the Regulation on the breakdowns of the core topics and the technical specifications and 2) the Regulation on the statistical dissemination programme and the metadata that all the Member States will have to provide to Eurostat by March 2014.
19. The EU census dissemination programme is a complete set of hypercubes clustering essential information that can be derived from the topics listed in the Framework Regulation (see Figure 1).
20. Each Member State shall provide the Commission with a few¹ hypercubes for the level LAU2 and with several hypercubes for the level NUTS 2.
21. The allocation of the topics in the short and in the long form will take into consideration the structure of the hypercubes in terms of variables, breakdowns and geographical level.
22. Given that the topics included only in the long form will be estimated, in order to ensure an high quality of the data to be transmitted to Eurostat, Istat has decided to include in the short form at least all the topics needed to produce the hypercubes for the level LAU 2. In the long form will be included, in addition to the variables of the short form, also the socio-economic topics essential to produce the hypercubes for the level NUTS2. In fact, the wider is the territorial reference, the greater should be the precision of the estimates for long form variables (see Carbonetti and Verrascina, 2009).
23. For the purpose of producing the hypercubes for the geographical level LAU 2, it would be sufficient to include in the short form exclusively the topics needed to produce information on the demographic features of the population such as relationships between household members, sex, age, legal marital status and country of citizenship. Nevertheless, a short form questionnaire design including demographic variables as well as socio-economic topics (at a minimum classification detail) would ensure the availability of data, for example on education and economic characteristics, at a very detailed geographical level (enumeration area) and a high level of accuracy without a too elevated sampling rate². This strategy will allow most information needs to be satisfied. As shown by the results of a survey carried out by Istat in 2006 on the use of the 2001 census data by Organizations belonging to the Italian National Statistical

¹ Up to May 2009 (the implementing regulation on the dissemination programme is still subject to changes) there were 19 hypercubes for the level of LAU 2 and 109 hypercubes for the level of NUTS 2.

² In this way it would be possible to collect more auxiliary information (information collected for the whole population) to be used either for direct estimation methods or for small area estimation methods in order to produce estimates related to topics included only in the long form.

System, the data most used are those related to the demographic topics, to the current activity status, to the educational attainment and to industry at a very detailed geographic level.

24. In the Pilot Survey, which will be held in October 2009, three types of enumeration forms will be tested: a) a short form, including exclusively demographic topics needed to produce Eurostat hypercubes for the LAU 2 level; b) a medium form, to be used in some municipalities as an alternative to the short form, including also topics concerning education, current activity status and commuting; c) a long form including all the variables (the short and medium form topics plus the socio-economic topics).

25. With the new census strategy, in the smallest municipalities (those with less than 5.000 inhabitants) the long form will be administered to the whole population while in the ones with more than 5.000 inhabitants the long form will be adopted together with either the short or the medium form (in order to simplify the fieldwork for the municipalities).

26. At the end of the survey, the spontaneous reply rate (i.e. the number of cases not requiring the enumerator to be sent on the field for the collection of the form) for each of the two shorter questionnaires (short and medium form) will be compared in order to decide which form should be used for the 2011 census. On the basis of evidence from studies conducted in other countries, the hypothesis will be tested that a shorter questionnaire corresponds a higher spontaneous reply rate and vice versa. If the hypothesis is verified, the balance between two opposite needs will need to be determined: that of producing information to a more detailed geography against the need of increasing the spontaneous reply rate to reduce the number of field interventions by enumerators.

27. Studies to verify which topics among those collected for the whole population would also be useful to improve the efficiency of sampling estimates of variables included only in the long form and will also be carried out.

IV. THE 2011 CENSUS PRODUCT INNOVATIONS AND COMPLIANCE WITH CONFERENCE OF EUROPEAN STATISTICIANS RECOMMENDATIONS

28. Besides the newly designed strategy, in the next pilot survey some product innovations will also be tested.

29. In addition to the core topics listed in the Framework Regulation, some of the non core topics mentioned in the Conference of European Statisticians (CES) Recommendations will be included in the questionnaires (especially in the long form). Furthermore, for some variables included in the 2001 enumeration, a new collection strategy will be tested.

30. The main product innovations will be:

- (a) the adoption of the household relationship matrix, even if it is limited to certain members of the households and the introduction of the categories useful to identify the skip generation households;
- (b) two new questions to collect information on the country of birth of parents;

- (c) the collection of information on the topic 'ever resided abroad and year of arrival in the country';
- (d) the enumeration of the topics concerning the characteristics of the job for the whole economically active population, including the unemployed who have ever been employed.

A. The household relationship matrix and the skip generations

31. The private household concept used in Italian censuses is the Population Register is similar to the housekeeping concept of the UNECE Recommendations. Traditionally, information on relationships between household members was collected by means of the question on relationship to the reference person of the household. In 2001 new response choices had been added to the same question (the category child had been split into three: child of both the reference person and his/her partner, child of the reference person only, child of the reference person' partner only) in order to identify reconstituted families (assuming that the majority of step relationships are found between the household reference person or his/her spouse/partner and their children living in the same household). Nevertheless, even if this minimal change provides the opportunity of deriving reliable estimates of reconstituted families by keeping the question as simple as possible, it doesn't solve the problem of identifying all family nuclei in multi-family households.

32. The adoption of a partial household relationship matrix will be tested during the 2009 Pilot Survey. In order to better identify relationships between spouses/partners and their children living in the same household, a table listing all household members will be completed by the respondent and they will be asked to write down the code corresponding to his/her father, mother, spouse/partner (if living in the same household). This will allow the identification of all family nuclei without asking all household members about their relationships with all other members (this differs from the Relationship matrix approach), while at the same time allowing the identification of most relationships between all household members, since the conventional relationship question will be kept as well. Nevertheless, because the questionnaire is self-completed, the method may still be too difficult to be managed by the respondent, who will be required not just to give a simple answer but to perform a coding operation.

33. The feasibility of the use of such a method for collecting information necessary for mapping household structures will also be evaluated by means of a cognitive research survey conducted by the Institute of Cognitive Sciences and Technologies of the National Research Council.

B. Ever resided abroad and country of birth of parents

34. The country of birth of parents is a non core topic. Nevertheless, it is recommended for countries with a significant number of immigrants, in order to identify the group of descendants of immigrants.

35. In Italy, there has been a major increase in the number of foreign citizens since the 2001 Census, while the larger population with migration background (including foreign citizens,

second-generation immigrants and naturalized Italian citizens) will be subject to an even greater increase.

36. An attempt to estimate the number of descendants of foreign-born has been made on the basis of 2001 data (see Bonifazi et alii, 2008), by using the information on the citizenship of parents of all children below 18 years of age (since the question on the citizenship of parents was not included in the questionnaire, this information was available only for persons living in the same household with their parents).

37. The already underestimated number of almost 2 million residents with foreign background (obtained by adding foreigners, naturalized Italians and Italians by birth with at least one foreign or naturalized parent) is expected to continue to rise. Indeed, the group of people with at least one parent with foreign background being Italians by birth will grow due to the increase in the number of mixed couples and to the increasing number of citizenship acquisitions, both by marriage and by ordinary naturalization. Therefore, the collection of information on the country of birth of parents would be of great interest though it obviously increases the statistical burden on respondents. Furthermore, it might be a difficult question to answer to as it is not referred directly to the respondent. Given the relevance of the topic in order to collect information on integration processes, it will be included in the long form tested during the 2009 Pilot Survey.

38. Finally, in order to identify the group of ever-international migrants, a further new topic will be included in the long form. All individuals will be asked whether they have ever had a usual residence abroad and, for those who ever resided abroad, information on the year of last arrival to Italy and on the country of previous usual residence will be asked. This is a core topic (ever resided abroad and year of arrival in the country) and the information is of relevance in a former emigration country like Italy as it allows the identification of former emigrants abroad among Italians by birth who currently reside in the country.

C. The collection of data on characteristics of the job of the unemployed

39. On the basis of the EU Regulation, Member States shall collect information on occupation, industry and status in employment not only for the employed during the reference week but also for unemployed who have ever been employed, concerning their most recent employment.

40. Similarly, 2010 CES Recommendations (as in the 2000 recommendations) underlines that “an unemployed person should be classified by occupation, industry, status in employment and sector on the basis of the last job which he/she had. The collection of data on characteristics of the last job (if any) of the unemployed is particularly important for users to have information on the characteristics of the unemployed in order to identify the specific areas of the economy or particular skills and occupations of unemployed people”.

41. In 2001 Population Census, data on the characteristics of the job were collected only with reference to employed persons. This decision was taken on the basis of the results of the Second Pilot Survey (April 2000), where a very high non response rate was recorded for unemployed to questions on the characteristics of their last job. Generally speaking, the economic characteristics

section of the questionnaire is the most complex to be answered. In addition, in the 2000 Pilot Survey it was possible that a feeling of anger for their current activity status prevented unemployed from providing information related to a job they had lost.

42. According to the implementing Regulations and in compliance with CES Recommendations, the collection of the job's characteristics of the whole economic active population will be tested again during the 2009 Pilot Survey. Both employed and unemployed will be classified by occupation, industry and status in employment as well as by time usually worked³ and part time job.

43. Analysis on the response rates and on the quality of the data collected will be carried out at the end of the survey. Since questions on the characteristics of the job will be included only in the long form, the efficiency of sampling estimates will be evaluated as well.

References

Bonifazi C., Gallo G., Strozza S., Zindato D., (2008) *Foreign and immigrant population: definition, categories and characteristics*, Migration Studies, Volume XLV, n. 171, July-September 2008.

Crescenzi F., Fortini M., Gallo G., Mancini A., (2009) La progettazione dei censimenti generali 2010-2011. 6 – Linee generali di impostazione metodologica, tecnica e organizzativa del 15° Censimento generale della popolazione, *Documenti ISTAT*, n.6/2009, <http://www.istat.it/dati/pubbsci/documenti/documenti2009.html>.

Dardanelli S., Mastroluca S., Sasso A., Verrascina M., (2009) La progettazione dei Censimenti Generali 2010-2011. 5 – Novità di regolamentazione internazionale per il 15° Censimento generale della popolazione e delle abitazioni, *Documenti ISTAT*, n.1/2009, <http://www.istat.it/dati/pubbsci/documenti/documenti2009.html>.

Fortini M., Zindato D., (2008), The 2011 Italian population and housing census: innovations and challenges, 'Les lundis de l'INED', Paris, INED (Institut national d'études démographiques), Salle Sauvy, 17/04/08, http://www.ined.fr/fr/rendez_vous/lundis_ined/bdd/rendez_vous/260/.

Istat, (2009), Accuracy evaluation of Nuts level 2 hypercubes with the adoption of a sampling strategy in the 2011 Italian Population census, *Joint UNECE/Eurostat Meeting on Population and Housing Censuses*, 28-30 October 2009, Geneva, Switzerland.

³ This topic has been included in compliance with CES Recommendations. In 2001 information was collected on the time effectively worked during the reference week.