

Short and long form: planning the questionnaires in the new census strategy

Simona Mastroluca

Istat, via A.Ravà 150 – 00142 Roma, mastrolu@istat.it

Abstract: The 15° Italian General Population and Housing Censuses will be based on a combination of register-based census, conventional census and sample survey. In this context, different kinds of questionnaires have been foreseen: a short form, in order to collect information on the demographic features of all the resident population and a long form, designed for a sample of households and including demographic topics as well as socio-economic variables. The choice of the informative contents of the questionnaires, that will be tested on the occasion of the next rehearsal (October 2009), feels the effects of the Eurostat dissemination programme, of the breakdowns and technical specifications of the core topics, of the studies concerning sampling variance and the effects of the results of a survey (2006) on the use of the 2001 census data.

Keywords: new census strategy, short/long form, sample

1. The new census strategy

The next census round, that is in an advanced planning stage, will be characterized by a larger use of existing administrative registers, in order to reduce the production costs, the time to disseminate data and the statistics burden on households, with a positive effect on data quality.

The population and housing censuses' production process will be completely changed: for the first time the Municipal population registers will be used in the survey, in an explicit and organized way. Therefore, the questionnaires will be no more distributed by the enumerators but sent by mail. The respondents will be able to give back the census forms in different ways: by Internet, by mail or delivering them c/o the collection centre, with an improvement of the response rate. The missing will be retrieved by the enumerators using the information given by the Municipal census offices.

In order to reduce the burden on the respondents, in the Municipalities with more than 20.000 inhabitants and in the provincial capitals two different kinds of questionnaires will be used: long form will be submitted only to a sample of households while short form will be administered to the remaining households.

The expected consequence of the innovations proposed is that the use of less enumerators, due to the postal services distribution of the census forms, and the possibility of a multichannel return will bring a clear benefit, improving the efficiency and effectiveness of the process.

The use of new web based data collection technologies will provide a rationalization of the field operations and opportunities to improve the management of monitoring and checking and thus the quality of the census itself.

2. The regulation of the European Parliament, breakdowns and hypercubes

The Framework Regulation of the European Parliament and of the Council, adopted in order to ensure the comparability of the data provided by the Member States, follows an approach based on the harmonization of the output rather than on the method of data collection.

It means that Member States can choose whichever data source (conventional censuses, register-based censuses, a combination of conventional censuses and sample surveys, a combination of register-based censuses and sample surveys, a combination of register-based censuses and conventional censuses, a combination of register-based censuses, sample surveys and conventional censuses and rolling censuses) but they should take into account a wide range of issues such as quality of the data, data protection and, above all, the essential features of a census (individual enumeration, simultaneity, universality within a defined territory, availability of small-area data and defined periodicity). Attached to the Regulation there is the list of the obligatory topics (core topics) to be covered in Population and Housing Censuses separately for the geographical levels NUTS 3 and LAU2 (municipalities in Italy) and for the geographical levels national, NUTS 1 and NUTS 2.

Planning the enumeration forms it is necessary to consider also other two Implementing Regulations: a Regulation on the breakdowns of the core topics (from 1 to 3 for each variable) and the technical specifications, that help to interpret the categories of the breakdowns listed, and a Regulation on the reference year, the statistical dissemination programme and the metadata that Italy and the other Member States have to report to Eurostat by march 2014.

The EU census dissemination programme is a complete set of hypercubes presenting essential information that can be derived from the topics listed in the Framework Regulation on censuses.

At the moment (the Implementing Regulations are still subject to changes) each Member State shall provide the Commission with 19 hypercubes for the level of LAU 2 (the maximum number of obligatory cells would be below 600 cells¹) and with 109 hypercubes for the level of NUTS 2 (the maximum number of obligatory cells would be below relevant 6000 cells²).

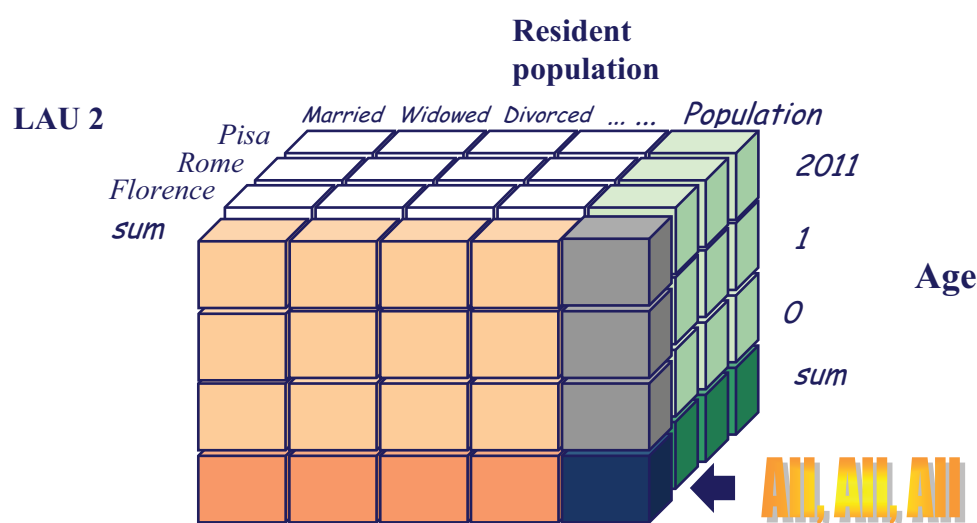
The distribution of the topics in the short and in the long form shall take into account the structure of the hypercubes in terms of variables, breakdowns and geographical level.

Considered that the topics foreseen only in the long form will be estimated, in order to ensure an high quality of the data to be transmitted to Eurostat, for the next pilot survey the Italian General Census Division has decided to test the new strategy including in the short form at least all the topics needed to produce the hypercubes for the level LAU 2; in the long form, in addition to the variables of the short form, all the other topics foreseen in the enumeration project will be included.

⁽¹⁾ they includes all cells, except the totals, subtotals and the “not applicable” category

⁽²⁾ they includes all cells, except the totals, subtotals and the “not applicable” category

Figure 1: *Example of Hypercube*



In this way, as far as the cells of the hypercubes at LAU 2 level are concerned, we will have only non sampling errors. In fact, we will enumerate the information needed for the hypercubes at LAU2 level for all the Italian population.

On the contrary, in order to produce the hypercubes for the level NUTS2, we will use also topics including only in the long form and enumerated on a sample of households. So, in the hypercubes at NUTS 2 level there will be an estimated component too.

3. The enumeration forms

In order to produce the hypercubes for the geographical level LAU 2, it would be enough to include in the short form exclusively topics to collect information on the demographic features of all the resident population such as sex, age, legal marital status and so on. Nevertheless, a short form questionnaire design including as well as demographic variables also 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 an high level of accuracy without an elevated sampling rate ⁽³⁾.

Thanks to a survey carried out in Istat in 2006 on the use of the 2001 census data from Organizations belonging to the Italian National Statistical System, we know that the variables mainly used pertain to demographic topics, to the current activity status, to the educational attainment and to industry.

Moreover, referring to the geographical level, the analysis carried out by the same Organizations have concerned first of all data at LAU 2 level, following by data at NUTS 3 level and by data for enumeration area. For example, information on the educational attainment have been used for the 65% at LAU 2 level, for the 18.4% at

⁽³⁾ In this way we could have more auxiliary topics (enumerated for all the population) to be used either for direct estimation methods or for small area estimation methods

NUTS 3 level and for the 10.0% at enumeration area level. Data on current activity status have been used for the 60.7% at LAU 2 level, for the 18.5% at NUTS 3 level and for the 11.1% at the enumeration area level.

In substance the use of data at national, NUTS 1 or NUTS 2 level reaches percentages from 2.4% to 5.6%.

So, on the occasion of the next census rehearsal (October 2009) three types of enumeration forms will be tested: a short form, including exclusively demographic topics needed to produce the Eurostat hypercubes for the level of LAU 2, a medium form that is an alternative to the short form and that includes also topics concerning education, current activity status and commuting and a long form with all the variables, including those already foreseen in the short and in the medium forms. At the end of the survey, the percentages of missing pertinent to the two types of questionnaires (short and medium form) will be analysed in order to evaluate, for the 2011 census, the opportunity of submit a short form more or less rich of questions.

Studies to verify how much topics, like educational attainment and current activity status, enumerated for all the population, are useful to improve the efficiency of sampling estimates relating to variables including only in the long form, will be carried out too.

A functional questionnaire design is very important to manage, after the enumeration, the huge quantity of information coming from a population census and to ensure an high level of data quality.

Compared with last census, on the occasion of the next pilot survey, product innovations such as the collection of information about the country of birth of parents will be tested. This topic permits the identification of the group of descendants of “foreign-born”, persons born in the country whose parents were born abroad (“second generation”).

Complying with the Framework Regulation, new questions on the topic “ever resided abroad and year of arrival in the country” will be included in the long form too.

References

- S.Dardanelli, S.Mastroluca, A.Sasso, M.Verrascina (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
- F.Crescenzi, M.Fortini, G.Gallo, A.Mancini, (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