



Economic and Social Council

Distr.
GENERAL

ECE/CES/GE.41/2009/8
20 August 2009

ENGLISH ONLY

ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Group of Experts on Population and Housing Censuses

Twelfth Meeting
Geneva, 28-30 October 2009
Item 4 of the provisional agenda

CENSUS QUALITY AND DISCLOSURE CONTROL

A quantitative approach to the European Union census quality reporting

Note by the Statistical Office of the European Communities

I. INTRODUCTION

1. The Council of the European Union and the European Parliament adopted in 2008 a regulation¹ on population and housing censuses. For the first time², a European Union (EU) regulation defines the concepts to be used in the census exercises, the topics to be covered and the data sources, and fixes the reference year for the first round on 2011. Further legislative measures – currently under preparation – are foreseen to define the technical specifications and breakdowns of the topics, the programme of statistical data and metadata, the technical format for data transmission, the modalities and structure of the quality reports. The methodological elements of this Regulation acknowledge the Conference of European Statistician (CES) Recommendations for 2010 Round of Population and Housing Censuses.

¹ Regulation (EC) No 763/2008 of the European Parliament and of the Council of 9 July 2008 on population and housing censuses (Text with EEA relevance). OJ L 218, 13.8.2008, p.14.

² Since 1973, the European Union has issued Directives on population censuses, mainly aiming at the synchronisation of the exercises across Member States. For the 2000 round, only a Gentlemen Agreement – but with wider scope - was adopted.

2. The Regulation (EC) No 763/2008 attributes great importance to quality issues. In particular, it devotes one entire article to the quality assessment, defining its dimensions, prescribing the submission of quality reports from the Member States to Eurostat and stating that "The Commission (Eurostat) shall assess the quality of the data transmitted." Hence, for the first time, the population and housing censuses in the EU will have a formal, structured and comparable quality reporting. Further, quality shall be assessed and this is a significant step forward in comparison to simple reporting.

3. As for the preparation of the Regulation (EC) No 763/2008, the work of Eurostat on the preparation of the implementing measures is supported by a Task Force composed of experts from Germany, Ireland, Italy, Austria, Portugal, Sweden and the United Kingdom. This Task Force is now focusing its efforts on a regulation concerning the report on the quality of the transmitted data. So far, the experts have discussed only once a first draft in a meeting, and that discussion indicated how challenging it can be to structure a quality report when different methodologies are used in the censuses exercises. Moreover, the discussion among experts showed that also basic concepts can be interpreted differently, thus making necessary the inclusion in the implementing regulation of a large paragraph on definitions of quality and statistical concepts. Given that the discussion at the Task Force is still at a very early stage, this paper will not report on the provisional outcomes of its work.

II. QUALITY DIMENSIONS AT INTERNATIONAL LEVEL

4. For the last decade, quality has been a topic widely discussed and analyzed at EU level. From 1998, a specific Working Group of Eurostat on Quality has developed several aspects of quality, among which an agreed (in the European Statistical System) definition of quality and standard quality reporting documents, the latest version of which was released in January 2009³. The importance of quality has been clearly acknowledged by its inclusion in the latest EU "Statistical Law"⁴, where Article 12 lists the quality dimensions for European statistics. As this Regulation is to be applied in all statistical domains, the importance of quality in statistics has certainly been reinforced.

5. The quality components listed in the EU regulations (both in the EU Statistical Law and in the EU census regulation) are:

- (a) 'relevance': it refers to the degree to which statistics meet current and potential needs of users;
- (b) 'accuracy': it refers to the closeness of estimates to the unknown true value;
- (c) 'timeliness': it refers to the period between the availability of information and the event or phenomenon it describes;

³ See Eurostat (2009): "ESS Standard for Quality Reports" and Eurostat (2009): "ESS Handbook for Quality Reports", both issued in the series Eurostat Methodologies and Working Papers.

⁴ Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European statistics. OJ L 87, 31.3.2009, p.164.

- (d) ‘punctuality’: it refers to the delay between the date of release of the data and the date by which they should have been delivered;
- (e) ‘accessibility’ and ‘clarity’: they refer to the conditions and modalities by which users can obtain, use and interpret the data;
- (f) ‘comparability’: it refers to the measurement of the impact of differences in applied statistical concepts, measurement tools and procedures where statistics are compared between geographical areas, sectoral domains or over time;
- (g) ‘coherence’: it refers to the adequacy of the data to be reliably combined in different ways and for various uses.

6. The EU components listed above are not the only ones provided at international level. The UNECE dimensions of quality are presented in para. 76 of the CES Recommendations and further described and analysed by Baffour and Valente (2008)⁵. These six dimensions are: relevance, accuracy, timeliness, accessibility, interpretability and coherence. The CES Recommendations devote two appendixes to quality, although more from the point of view of quality assurance and methods for evaluation. At the worldwide level, UNSD focuses its attention more on the quality assurance rather than on quality evaluation. However, the UNSD’s Principles and Recommendations for Population and Housing Censuses, Revision 2, do identify and describe (paragraphs from 1.228 to 1.238) a set of ten quality dimensions: relevance, completeness, accuracy, comparability, coherence, timeliness, punctuality, clarity, accessibility, and metadata. An approximate correspondence among these sets of quality dimensions is shown in the following table (the remainder of this paper refers only to the EU concepts):

Table 1

Correspondence between quality components as defined by international organisations

EU	UNECE	UNSD
Relevance	Relevance	Relevance Completeness
Accuracy	Accuracy	Accuracy
Timeliness Punctuality	Timeliness	Timeliness Punctuality
Accessibility Clarity	Accessibility Interpretability	Accessibility Clarity Metadata
Comparability Coherence	Coherence	Comparability Coherence

⁵ B. Baffour and P. Valente (2008): “Census Quality Evaluation: Considerations from an international perspective”. Paper ECE/CES/AC.6/2008/SP/4 of 9 May 2008, prepared for the Joint UNECE/Eurostat Meeting on Population and Housing Censuses, Geneva, May 2008.

III. A CRITICAL VIEW OF QUALITY REPORTING AND ASSESSMENT

7. The implementation of the EU quality concepts in the population and housing censuses is not straightforward. First of all, these generic dimensions have to be translated into components which are appropriate to the census domain. Further, in order to make them operational, it is necessary to define appropriate tools for their measurement (indicators). In most of the cases, it is necessary to define more than one indicator for each quality dimension. Although it is possible to find many references, such as the above-mentioned ESS Handbook for Quality Reports, the peculiarities of each domain have to be properly taken into account. Therefore, there are two aspects that need to be dealt with in the context of the EU population and housing censuses. First, which of the quality dimensions listed above are actually significant; second, how to define a quality indicator, i.e. if they should meet specific requirements.

8. It may be useful to start from what a quality report is not, instead of from what it is. First of all, a quality report is not exactly a methodological report. In rather simplistic terms, while the latter describes how something has been done, the former should inform us about how good is the result. Certainly the area of overlap between these two kinds of reporting is large, and in some cases a methodological report is actually proposed as quality report. Nonetheless, conceptually they refer to two different frameworks. For instance, in case of a sample survey, the methodological report should inform the reader about the sampling design and estimators, while the quality report should focus on the accuracy of the estimates. Obviously, the methodological description can include the estimates of the sampling variance (or whatever other indicator of accuracy), as well as the quality report can inform about the way the sampling was made, but the purpose of the two reports should – at least conceptually – be kept separate. In a few words, a methodological report can be drafted ex-ante, a quality report only ex-post.

9. Secondly, quality reporting is not only about errors. For certain dimensions, quality can be expressed as different degrees of achievement of a positive result. Missing a full target, meaning an only partial achievement of an objective of “perfection”, does not necessarily imply that an error has been made, but simply that the outcome could have been (even) better. Therefore, quality reports concern also efficacy/efficiency, and they could include elements on performance, costs and respondents burden.

10. Thirdly, quality reports are not equal to metadata. These latter should help the user to better understand the content of the data, but they do not necessarily represent a key to assess their quality. Even in the case where the information reported in the metadata is significant from the point of view of the quality, this needs to be “translated” into quality measurements in order to be properly taken into account in the quality assessment. Metadata tell us (more) about what it is, while quality tells us about how it is. Again, metadata can be incorporated in quality reporting, as quality indicators can be part of the metadata, but the two reporting systems serve different purposes.

11. Fourth, quality is not an intrinsic feature of the data. Its assessment is dependent on the perspective taken. Even using the same set of indicators – or, more generally, the same criteria – two users could come to different conclusions on the quality of a given set of data. Therefore, for instance, the quality of a census can be optimal for national purposes, and at the same time to be

less optimal from an international point of view. This consideration points to the need to establish non-subjective and transparent criteria to assess the quality.

12. From the considerations above it emerges that, in setting a framework for the quality reporting, the first decision is on which perspective to take. In the case of the EU population and housing census, and taking into account the sense and the provisions of the Regulation (EC) No 763/2008, it is clear that the assessment of the national census exercises of the EU Member States should be made from an international point of view, and in particular from the European Commission standpoint. This has some consequences in terms of the set of quality components to be taken into account. It goes without saying that this would concern only the quality reporting as prescribed by the EU regulation and not any other quality evaluation the Member State wish to make for whatever purpose.

13. If the EU perspective is the one adopted in the quality reporting, not all the above-listed EU dimensions of quality have the same importance. In fact, the EU regulation lists the topics that have to be covered by the censuses in EU, and the Member States have the freedom to choose the data source of their preference from a given list. It is clear that the EU list of topics is not exhaustive, meaning that the countries are free to add to their census exercise whatever other topic they consider relevant for national purposes. These additional topics are not subject to the EU quality reporting and assessment. Further, as the country is free to choose the census methodology (among conventional approach, registers-based, sampling, rolling census), it derives that the same EU quality dimension may apply differently in the Member States. Therefore, the EU quality reporting focuses on a common but not exhaustive list of census topics, which can be collected by means of different methods.

14. Another peculiarity is related to the system of data dissemination. As the current planning is to disseminate all national census data via a Eurostat Census Hub, the census data dissemination is in fact – from an EU point of view – summarised in the transmission to this IT interface. This is clearly a different perspective from the one usually considered for the dissemination of census data. Therefore, in very general terms, in defining the quality reporting for the EU population and housing censuses, it has to be taken into account that not all the six EU quality dimensions are equally important (or even relevant) and, within a defined dimension, not all the quality indicators are significant across countries.

IV. A QUANTITATIVE APPROACH TO THE QUALITY REPORTING

15. It is necessary to define the suitable characteristics of a quality indicator. So far, nothing has been written about the nature – qualitative or quantitative – of it. For several scholars, a quality assessment is mostly based on qualitative information, such as the description of the methodology. Although qualitative reporting may bring relevant added value, in absence of well defined rules of interpretation of the qualitative information its potential weakness is the risk of non consistent assessments. For instance, supposing that it is commonly recognised that the methodology “A” produces better quality than methodology “B” for a given quality dimension, then this should be clearly defined as rule of assessment, in order to avoid different reviewers drawing different conclusions. Unfortunately, such a clear understanding and preferences setting is rarely the case. In practice, the qualitative information is left to the judgment of the single reviewer, who will derive from it his/her own personal opinion about the quality.

16. In addition, such an assessment based on qualitative information would most probably require a much deeper technical knowledge. Recalling the example above, not all the common users may be aware that methodology “A” is better than “B”. It is therefore necessary to identify information easy to understand and interpret. The task of the interpretation may be further complicated by the volume of the information provided to the reader. Summarising a very detailed quality report in a global assessment of the kind “good/bad quality” – operation that at the very end any user will make in his/her mind - may prove to be rather challenging. From this perspective, an excess of information can be as negative as a lack of information.

17. On the basis of the considerations above, it can be concluded that a suitable quality indicator should be transparent and easy to interpret. From this point of view, quantitative measurements will provide a better result than a qualitative measure. To meet these requirements, the quantitative indicators should preferably:

- (a) be in a range from 0 to 1;
- (b) become bigger when the quality is higher;
- (c) be a dimensionless number, i.e. without unit of measure.

18. The first criterion would help to understand the size of the phenomenon under consideration. Together with the second criterion, intuitively, it becomes easier to get an idea of the level of quality for a specific dimension/indicator. Finally, the third criterion allows for comparability and aggregation.

19. From a census management perspective, quality quantitative measurements provide a benchmark to monitor improvements in the next census rounds. Comparability over time of the indicators can be a powerful tool to observe how real the enhancements are, and may also be helpful from the quality assurance point of view, meaning the process aiming to ensure the highest quality during the census exercise.

20. In order to make clearer what has been proposed in para.170, an example of application is developed below. This example is provided for illustrative purposes, and it is not intended to substitute for the contribution made by experts in various fora.

21. The first step is to assess the importance of each EU quality dimension in the context of the EU population and housing census:

- (a) **Relevance.** This dimension concerns the topics covered by the censuses in the Member States. As the list of topics is provided in the EU regulation, which has been developed through a significant consultation process with Member States, it can be assumed that the relevance is fully met by definition.
- (b) **Accuracy.** While a very important dimension, accuracy is a complicated dimension to assess, given the range of methodologies used by different countries and its inherent multifaceted nature.

- (c) Timeliness and punctuality. The Regulation (EC) No 763/2008 defines the latest date by which the countries have to provide the census data (end of March 2014). Being an obligation from an EU regulation, in principle all countries are expected to respect the deadline. Punctuality will thus probably be less varying than timeliness, as countries are free to transmit the data earlier if possible.
- (d) Accessibility and clarity. These dimensions apply only partially. Accessibility is closely related to the dissemination system planned for the census data (Eurostat Census Hub) and, as such, it is independent of national performance. Clarity is more dependent on the information provided with the data (metadata) and which is going to be defined by EU regulation.
- (e) Comparability and coherence. These two dimensions are complicated to implement and for the sake of simplicity are not pursued here.

22. The next step is the definition of quantitative indicators for each quality dimension.

23. With regard to the relevance dimension, although the list of topics is mandatory and therefore countries are expected to fully comply, for illustrative purposes it is assumed that the countries might transmit only a part of these core topics⁶. An indicator of relevance could then be:

$$I_{RL} = \frac{t}{T}$$

where t is the number of topics for which data are provided and T the total number of topics as listed in the EU regulation. This indicator would assume values ranging from zero (if no data on any topic are transmitted) to one (if all data related to EU topics are provided)

24. With regard to the dimension of accuracy, it is possible to find in the literature a large number of indicators. In order to keep the example as simple as possible, and to further show how to build indicators according to their characteristics listed above, only two basic indicators of accuracy will be considered. The first is the indicator of under-coverage, meaning the missed enumeration of units belonging to the target population; the second is an indicator of sampling error⁷, meaning the error due to surveying only a subset of the population rather than conducting a complete enumeration of all the units in the target population. The former indicator can be defined as:

$$I_{AR,uc} = 1 - \frac{\hat{N}_{uc}}{\hat{N}}$$

where \hat{N}_{uc} is the estimated⁸ number of missed units and \hat{N} the estimated total number of units of the target population⁹. The higher is the under-coverage, the lower the value of the indicator,

⁶ The consequences of an infringement of an EU regulation are not considered here.

⁷ The Regulation (EC) No 763/2008 includes the sampling as possible data source for the population and housing census.

⁸ The way how this estimation can be made is not a subject of this paper.

thus the lower the quality in terms of accuracy. The second indicator chosen for the example is the well-known coefficient of variation; however, in order to meet the requirements listed above, this indicator could be defined as follows:

$$I_{AR,cv} = \begin{cases} 1 - \frac{\hat{\sigma}_s(\hat{Y})}{\hat{Y}} & \text{if } \hat{\sigma}_s(\hat{Y}) \leq \hat{Y} \\ 0 & \text{otherwise} \end{cases}$$

and thus if the standard error $\hat{\sigma}_s$ is more than 100 per cent of the value of the estimate \hat{Y} , the quality (in terms of accuracy) of the data is considered totally bad. The assessment on accuracy may be complicated by the adoption of different census methodologies, which could make it not relevant for the chosen indicators, or (like any other quality dimension) for missing topics. Decisions have to be taken in advance on how to deal with these special cases. For instance, the indicator $I_{AR,cv}$ can be set equal to one whether no sampling is used for a given topic, as there is no sampling error by definition¹⁰.

25. As reported in para. 21(c), the deadline for census data transmission to Eurostat is defined by the Regulation (EC) No 763/2008. The punctuality would then refer to the theoretical case of a delay by a country to meet a legal obligation. Neglecting the legal aspects of an infringement of the EU regulation, a basic indicator of punctuality satisfying the properties in para.17 could be as follows:

$$I_{PT} = \begin{cases} 1 - \frac{d}{365} & \text{if } d \leq 365 \\ 0 & \text{otherwise} \end{cases}$$

meaning that any delay d (expressed in days) equal or greater than one year would make the data quality to be very bad (actually of level zero) in terms of punctuality. More interesting is the case of timeliness, because the Member States are free to transmit the data earlier than the deadline. For this quality dimension, a basic indicator could be expressed as:

$$I_{TL} = \begin{cases} 1 - \frac{d}{821} & \text{if } d \leq 821 \\ 0 & \text{otherwise} \end{cases}$$

where d is the time expressed in days from the end of the reference year of the census (2011) until the data transmission, and the denominator is the number of days from the end of the reference year to the official deadline; if the data are transmitted after the deadline, they are not considered timely anymore. This indicator could be made even more precise taking into account

⁹ The example can be further developed considering the different types of enumeration units in a population and housing census.

¹⁰ The coefficient of variation could be extensively requested, also at level of cross-tabulations. However, it should be made a distinction between the measures strictly used for the quality assessment and those requested as a kind of metadata, thus more for purposes of information to the user.

the case of a data release within the reference year and thus adapting accordingly its definition. For instance, d would then mean the number of days from the reference day of the census and not from the end of reference year. However, by doing so, the possibility to have a perfect timeliness is purely theoretical, because for this to happen the data should be released on the same day to which they refer. Other options are possible, such as considering timeliness starting not earlier than 3 months from the reference day, including or not revisions, etc.

26. Following the remarks expressed in para.210, no indicator of accessibility will be proposed as an example. For purposes of clarity, as this quality dimension is mostly of a descriptive nature, defining an appropriate quantitative indicator is even more challenging. Focussing on the metadata, if the EU regulation foresees the transmission of information on a number T_x of topics, then a basic indicator could be:

$$I_c = \frac{t_x}{T_x}$$

where t_x is the number of topics for which the country has transmitted the information. This indicator can be further developed to take into account the quality of the information¹¹ provided.

27. As the logic to be followed in defining the indicators should now be clear, there is no need to propose further indicators for coherence and comparability.

28. In general, it may be that some quality dimensions are considered of minor importance in the overall assessment of census data or neglected because of the excessive burden such assessment would cause on the data producers.

V. CONCLUSIONS

29. From the point of view of the data producer, the transmission of the required indicators may well be a much lighter task than the drafting of a very comprehensive report on quality. It is important that all the information provided in a quality report is effectively used for the quality assessment. Even if there is a lot of information which would be interesting to know, it has to be kept in mind that this is in fact an additional burden for the data provider, and fairness demands that maximum use is made of the information provided.

30. From the point of view of the evaluator(s), this approach would simplify the analysis and reduce the time for the assessment, which would be a substantial advantage when dealing with a large number of countries. Moreover, currently in the EU there are 23 official languages and potentially each of them could be used to draft the quality report. Focussing on quantitative indicators and reducing to the greatest possible extent the assessment based on qualitative information could then be an asset when the evaluator operates in a multilingual context.

¹¹ For instance, this could be by means of variables expressing the quality of the information for each topic, ranging from zero (very bad quality) to one (optimal quality).

31. From the point of view of dissemination, the users can always complement the quantitative information with the national reports on quality (which will have a wider scope) or any other source of information¹² that may be available.

32. Therefore, the EU standard quality report to be transmitted by the country could be based on a descriptive part on the life cycle of the census and on a set of indicators, significant from the EU perspective. Explanatory notes are especially helpful to understand how certain measures have been derived¹³ and should enrich the content of the reports. This EU reporting is by no means intended to replace the quality evaluation made at national level or by any other institution.

Table 2

Structure of an EU Quality Report adopting the proposed approach

Part I	Description of the life-cycle of the census
Part II	Methods of estimation of the indicators
Part III	Values of the EU quality indicators
Annex (voluntary)	National quality report/evaluation

¹² In particular, there may be indicators which could well be applied more in detail and provided as information to the users (e.g., the coefficient of variation for tables). However, this kind of additional information should better be included in the metadata and not in the quality reports, because it is not directly used for the quality assessment.

¹³ For instance, a description of the method used to measure the under-coverage, or the sampling design and the formula used for the sampling variance, etc.