



**Economic and Social
Council**

Distr.
GENERAL

ECE/CES/AC.6/2008/SP/4
9 May 2008

Original: ENGLISH

ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Joint UNECE/Eurostat Meeting on Population and Housing Censuses

Eleventh Meeting
Geneva, 13-15 May 2008
Item 2 of the provisional agenda

CENSUS QUALITY ASSURANCE AND EVALUATION

Census Quality Evaluation: Considerations from an international perspective*

Note by UNECE Secretariat

Summary

Regardless of the methodology used and of the measures that can be taken to assure the best quality, the population census cannot be a perfect operation and there will always be some errors affecting the quality of the census results in terms of coverage and content. For this reason, the agencies responsible for the census should always conduct an evaluation of the quality of the census and provide an assessment of the census coverage and content errors. This is particularly important considering the huge resources invested in the census and the key role played by the census results for the national statistical systems and as primary source of statistical data for the countries. Measuring the census quality, however, is not an easy task, in particular when non-traditional census methods are adopted, for which established census evaluation techniques like the Post Enumeration Survey cannot be applied. In this paper, some issues concerning census quality evaluation are discussed for each of the most common census methodologies, considering the implications on the different dimensions of quality. Some considerations are also presented on census quality evaluation from an international perspective, and on the role that International Organisations can play in this field.

* The paper was prepared by Bernard Baffour and Paolo Valente (UNECE). The views expressed in this paper are those of the authors and do not necessarily reflect the views of the United Nations.

I. INTRODUCTION

1. The population census cannot be a perfect operation, and there will always be some errors, regardless of the methodology adopted and of all the measures that can be taken (and should be taken) to assure the best quality in all phases of the census.
2. Different types of errors can affect the census at various extents and at various stages, depending on the methodology. In a traditional census with field operations, for instance, errors can be introduced during data collection by the enumerators/interviewers or by the respondents (voluntarily or involuntarily). When data are processed, errors can be introduced, for instance, during data coding, entering, and editing. For censuses based on registers, on the other hand, there could be errors in the registers used as data source, or errors could be introduced for instance when data from different registers are linked, or when data are estimated for variables that are not available in registers.
3. In order to deal with these errors, an important component of each census is an appropriate quality assurance and improvement system. This system should be planned as an integral part of the census programme with the primary objective “to ensure that quality is appropriately considered in all phases of the census work”¹. A well developed system should allow identifying on time and addressing most of the problems and errors that can affect the census data. However, there will always be some errors that will not be corrected, and that would eventually affect the quality of the census results, in terms of coverage and content.
4. It is, therefore, very important that the agencies responsible for the census are able to evaluate the quality of the census and provide an assessment of the census coverage and, as far as possible, also of the content errors. A comprehensive census evaluation program, however, should not be limited to estimate coverage and content errors, but should include an assessment of all phases of census operations, in order to provide the users with a complete picture of the quality of the census.
5. The next section of this paper is dedicated to what is generally intended by census quality, and what are the different dimensions of quality that are commonly considered. Section 3 is dedicated to the discussion in general terms of the main measurement issues for each of these quality dimensions. In section 4, the most common census methodologies will be reviewed and the implications on quality evaluation will be discussed. Finally, some general results about census evaluation activities carried out by ECE countries in the 2000 census round will be presented, together with some considerations on census evaluation from an international perspective.

¹ Conference of European Statisticians Recommendations for the 2010 Censuses of Population and Housing (United Nations, 2006) para. 72.

II. WHAT IS QUALITY? WHAT IS CENSUS QUALITY?

6. The realization that *information* lies at the helm of any census – it is the end product and the main purpose for which a census is undertaken – is pivotal to the formulation of any census quality evaluative exercise. Thus, an assessment of census quality translates into an assessment of the information that is produced by the census. The definition of quality, on the other hand, is fraught with a lot of difficulties. However, the most widely used definition focuses on the fitness for use of the information. In a census context, quality can be characterized in terms of six dimensions, namely: *relevance*, *accuracy*, *timeliness*, *accessibility*, *interpretability*, and *coherence*².

- a) The *relevance* of information reflects the degree to which census data meet the needs of the population, users and stakeholders. As it is impossible to measure every phenomenon, this dimension looks at whether the country's pertinent informational needs are satisfied by the census. The relevance dimension, hence, looks at achieving a balance between meeting the (sometimes conflicting) user requirements and satisfying the most important needs within the confines of constrained resources, which is admittedly not a straightforward task. A particular category of users of census data is represented by international and supra-national organisations. In the ECE region, for instance, for the 2010 census round countries will be requested to provide census data to UNSD and Eurostat. These data, to be provided in the form of tabulations or hypercubes (multidimensional matrices) will allow presenting comparable country-specific census data that could be used for analysis, policy formulation or other purposes.
- b) The *accuracy* of census results is the degree to which the data describe the phenomenon of interest. It concerns the reliability and precision of the population estimates, and is usually characterized in terms of errors. The precise measurement of most phenomena, although ideal, is unattainable due to the fact that no measuring instrument is perfect; thus there is the distinct likelihood of error. However, a concerted effort to quantify the coverage and other errors leads to the production of data of a reasonable quality within the resources available to the national statistical institute. The accuracy achieved will depend on the explicit methods put in place to identify and control for any errors that occur throughout the entire census, from the inception to its fruition. More so, to achieve the most accurate results in a census there needs to be an extensive effort put into the census design, collection procedures and processing. This dimension is linked to the degree of *coherence* and *timeliness* of the information, and oftentimes key decisions

² These are the six dimensions of quality considered in the Conference of European Statisticians Recommendations for the 2010 Censuses of Population and Housing (United Nations, 2006), para. 76 and Appendix IV. There is likely to be agreement between international statistical agencies on a standard set of quality dimensions later in 2008. The differences between the standard set and the set used here are expected to be minor.

taken to improve accuracy will impact on the delivery timescales, and how coherent the outputs are to users.

- c) The *timeliness* dimension refers to the time frame of the census process, and in particular to the interval between the census reference date and the time when the results become available. Evidently most censuses do take a long time – even as the results are being disseminated, there is planning under way for the subsequent census – but the longer it takes for the national statistical office to release the results the less of a ‘snapshot’ it is, and questions start to arise as regards the data validity and *relevance*. Even so, it must be reiterated that time must be taken in the post-censal processes so as to ensure that the quality of the outputs are of the best standard (there is a trade-off with the *accuracy* dimension).
- d) The *accessibility* of census data refers to the availability and ease with which they are disseminated to the stakeholders and public at large. Census information is usually disseminated through a mix of free and tailor-made products and services by the national statistical institute. The strategy adopted including the costs of the services and products will influence the degree to which the data is accessible. The census is conducted to meet the needs of various categories of users, including the central government (that use the data for instance for the formulation of policies and the distribution of resources), researchers, businesses, non-governmental organizations and the general public. As a result, care needs to be taken to ensure that the census data, albeit widely accessible, is also *coherent* and easily *interpretable* so as to facilitate its effective and appropriate use.
- e) *Interpretability* refers to the degree to which the information is easy to understand and any salient census results are easily found by the user; in essence this dimension focuses on how the data ‘makes sense’ to the users. The interpretability dimension of quality, therefore, measures how the data input and the data output relate, particularly whether the responses given answer the question being asked. A detailed pre-testing programme should ameliorate this aspect of interpretability. The other aspect concerns how available supplementary information is that would make the census output easier to understand (i.e. metadata provision). This metadata should explain the underlying concepts, terms, definitions and classifications. Additionally links to papers and publications that detail the methodology and assumptions behind the data collection and adjustment processes should be made available.
- f) The *coherence* of census information reflects the degree to which the census data can be brought together with other existing statistical information. It also concerns the conceptual integrity of the information, which can be assessed by comparing to existing information either through older censuses, or surveys or administrative data. A detailed coherence strategy will consequently have a programme of certification and validation of the data so that it investigates and explains any deviances from the expected trends.

7. These six elements, although suitably distinct may not be mutually exclusive, and actions taken to address one element of quality may have an effect on the others. Accordingly any actions to improve a dimension need to be weighed against any possible detrimental effect this might have on the other dimensions – realistically an optimum balance may be achieved by careful appraisal of the effect of each of the six attributes.

III. MEASURING CENSUS QUALITY

8. The evaluation of the quality of the census is a very important exercise for a number of reasons. From the organisational point of view, the evaluation should allow verifying whether the very large effort and investments of resources required by the census was worthwhile. Moreover, the evaluation exercise can help in identifying any aspects of the census organisation that could be improved. This could be particularly useful for planning the next censuses, and improving their efficiency and effectiveness.

9. With regard to the use of the census results, the evaluation allows providing the users with some measures of the quality of the data, which would help them to better interpret the results. This is particularly important considering that census results often serve as the benchmark for the national statistical system and as such they are used for a number of other statistical activities and for several years (usually until the next census data are available). The census is often the main (if not the only available) source of information when it comes to measuring some particular social phenomena, particularly with reference to small population groups or data for small geographical areas. In certain cases and for certain purposes, the results of the evaluation can also be used to produce adjusted estimates of the total population or other census aggregates, taking into account identified errors. In the UK, for instance, the final census results are adjusted on the basis of the results of the coverage assessment³.

10. Census quality evaluation is important also in fostering confidence in the information produced and as such needs to be an integral part of the census. As aforementioned, the best quality assessment undertakes a comprehensive evaluation of the various phases of census operations. In the following paragraphs, some general issues will be discussed on the measurement of the different dimensions of quality.

A. Assessment of Relevance

11. The programmes and outputs of a national statistics office must reflect the country's most important and pertinent informational needs. The relevance requirements have to be managed to ensure that not only is the most pertinent information being collected, but the manner in which the information is collected is also considered, in order to minimise the burden on the public when it comes to collecting the census data. This can be achieved through methods to assess the relevance of previous census content and to identify any new and potential informational gaps

³ See: "2011 United Kingdom Census Coverage Assessment and Adjustment Methodology", paper prepared for the UNECE-Eurostat Meeting on Population and Housing Censuses, Geneva 13-15 May 2008 (ECE/CES/AC.6/2008/3)
<http://www.unece.org/stats/documents/ece/ces/ge.41/2008/3.e.pdf>

that may appropriately be filled through the current census. This assessment should also look at other sources especially with regards to surveys and administrative sources. The census loses its effectiveness if the public fail to engage with it and so due care should be employed to make sure that only the most relevant phenomena are measured via the census so as to reduce the public burden. Seemingly, the best way of assessing the relevance is through feedback from users as they can provide comments on the adequacy and completeness of the data used in their analyses. The consultative exercises, in addition to aiding in the planning, allow the census authorities to provide a forum on which to discuss, and be responsive to, the needs of the stakeholders. It can also serve to encourage a greater understanding of the census plans and activities. As such the assessment of relevance should include:

- i. client and stakeholder feedback mechanisms to facilitate their active participation in the census processes;
- ii. user and public consultations during the planning process;
- iii. consultation with professional bodies as to the most relevant mode of measurement of different phenomena;
- iv. consultation with other agencies collecting data to review and co-ordinate data collection;
- v. market research on the efficiency of different response mechanisms.

B. Assessment of Accuracy

12. In general, errors can be classified into three major categories: coverage errors, content errors and operational errors. The assessment of accuracy seeks to identify and quantify these errors.

13. Coverage errors are those affecting the completeness of the census count. In practical terms, they include census omissions and erroneous enumerations of units that should not have been counted (typically, in case of double counting). There are a number of ways of quantifying the proportion of omissions or erroneous enumerations. The most common is to perform a post-enumeration survey (PES) which consists of a re-enumeration conducted in a sample of areas within the country. The results of the survey are then compared to the census results for the corresponding areas. This permits estimates to be made of the completeness (in other words, coverage) of the census population count, specifically identifying the proportion of people missed in the census. It can also allow estimates to be found of the proportion of people who have been erroneously included in the census due to being counted at the wrong address.

14. Another way to assess the completeness of the population count is through demographic analysis. Population estimates are calculated in most countries using current vital statistics on births and deaths and on in-migration and out-migration movements (or estimates of net migration). This can provide a basis for judging the accuracy of the census information. The validity of this approach clearly depends on the quality of the population estimates, in particular the data or the assumptions used to estimate the migration (since data on births and deaths are sufficiently reliable in most countries). Clearly, it is normal that there will be some differences between the population estimates calculated and the actual census results. However, if the difference is considerable, or if there are glaring divergences in the sex ratios or for specific age cohorts, then they could be due to problems in the quality of the census data. An alternative way

to assess the completeness of the census population count relies on comparing the census results to data from other alternative sources, for example administrative registers. The comparison can be conducted at the aggregate or individual level. Also in this case, the validity of the approach clearly depends on the quality of the data used for the comparison, in this case the registers.

15. The second type of errors is content error, which includes the incorrect reporting or recording of the data, and also errors caused by the non-reporting or non-recording of data. The errors may be caused by a variety of factors – non-response, enumerator or interviewer effect, recall effect, mode of measurement effects etc. Many of these errors could be avoided by adequate care in the design and implementation of the methodology used to collect the census data. Additionally, national statistics offices normally develop an edit and imputation strategy with the aim of resolving any inconsistencies and providing estimates for the missing data (at least for basic variables) and as part of the data processing phase. The editing process irons out any inconsistencies while the imputation process fills in some or all of the missing information. Doing this has the advantage of producing a set of consistent complete results.

16. The third type of errors is operational error, which result from the day-to-day processes of the census. Operational error can be difficult to quantify, and more often will be less of an issue when compared to the coverage and content errors. Some types of operational error are data capture error, coding error, tabulation error and classification error. Evidently, the operational error can be minimized if there is a rigorous review of all the processes that make up the census – from the initial pre-census activities of consultation, planning, development and testing of the census methodology, to the actual census processes of enumeration and data processing.

17. All these errors – coverage, content and operational – are interlinked. Clearly, a poorly implemented census will be susceptible to operational error, and will fail to adequately count everyone correctly, leading to coverage error. Further, the impact of a high proportion of coverage error is to lead to a lack of suitable donors with complete data than can be used for the edit and imputation algorithm and this will inadvertently introduce more content error.

C. Assessment of Timeliness

18. The timeliness of the release of census data is an issue of concern to many users. There is often a trade-off between accuracy, coherence and interpretability, but more so in terms of relevance. It goes without saying that the more timely the release of the census, the more realistically pertinent the information is to the phenomena it has been commissioned to measure. Hence, the timeliness can be assessed by various aspects:

- i. the realistic setting of targets relating to the release of the data;
- ii. announcing in advance the release dates so as to manage user expectation;
- iii. setting realistic targets for customized products and ad hoc services provided to users who require non-standard outputs or additional analyses;
- iv. the dissemination schedule should be made clear to users in advance, and they should be made aware of the differences between the planned outputs;
- v. if there are plans for specific census outputs to be released at different time periods (for instance, preliminary and final results, or regular updates of certain variables),

then the time schedule and costs associated with the release of the updates should be detailed.

D. Assessment of Accessibility

19. The benefits of a census are increased exponentially if the results are made widely available. This will also increase awareness of the usefulness of the census, and therefore contribute to improve participation in future censuses. There are a number of issues to be considered with regard to the evaluation of accessibility:

- i. there should be a wide scope of free and priced products and services available;
- ii. the products and services should be in different formats and designs to cater for all the cross-sections of the population, e.g. the blind and partially sighted;
- iii. for chargeable products and services, the prices should be kept at a minimum in particular for general products, that should be affordable to the general public;
- iv. consultancy services and specially commissioned products could be offered to users who want detailed analysis, against payment of the corresponding costs;
- v. sets of census microdata should be available to selected categories of users, taking appropriate measures to ensure data confidentiality;
- vi. the availability of a dedicated website is an important factor for data accessibility;
- vii. the provision of a dedicated customer service is also important to assist the users to find the data they need and to answer any queries, particularly in the period shortly after the release date;
- viii. client feedback on the different products and services should be monitored so as to make improvements and inform future census dissemination.

E. Assessment of Interpretability

20. The assessment of interpretability is closely entwined with coherence, and it can be observed that a meticulous coherence strategy will assist the interpretability of the census information. However, the primary concern of the assessment of the interpretability is the level of provision of additional information (mainly through metadata) to facilitate the understanding and aid usage of the statistical data being disseminated as well as demonstrate transparency. The metadata required for the census is therefore dependent on both the information and the context of its use.

21. Concerning metadata, in accordance with the SDMX Content-Oriented Guidelines⁴ we can consider two main categories: reference and structural metadata. Reference metadata refers to the additional information that provides descriptions of the actual contents (for instance the concepts and methodologies used), the general administration, execution and dissemination of the census (for instance the operations, procedures and other systemic processes). The information provided about the quality assessment activities is also an aspect of reference metadata. The structural metadata meanwhile refers to the additional information that aids the look-up of relevant information, so concerns things like indexes, categorizations and keys that users can consult when looking for particular pieces of census information (bearing in mind that there is often an incredible amount of output disseminated as part of the census).

22. The evaluation of interpretability can be based on:

- i. comprehensive provision of concise and relevant metadata, including information on the methodology used in the data collection and processing, particularly if sampling has been employed;
- ii. clear details of any new elements or changes that may have taken place since the previous census, and that may affect comparability between current and past census results;
- iii. clear and complete explanation of new concepts, in particular if new questions are included in the census for the first time;
- iv. clear statement of any limitations to the released data;
- v. summaries of key findings with directions to where to find relevant data, in consideration to the fact that the census contains a wealth of information, which can prove to be daunting to users;
- vi. availability of a glossary describing underlying concepts, methodology, definitions, variables, classifications;
- vii. details of the quality assessment should be made available to users.

F. Assessment of Coherence

23. There are two aspects to the assessment of the coherence dimension of the data quality; namely internal coherence and external coherence. The data is internally coherent if the whole census results are consistent within themselves. In order to have internally coherent data, a number of verification and validation tests should be carried out prior to the dissemination. Particular care should be paid to the possible impact on internal coherence of the data imputation and editing processes, which have a direct impact on changing the data. For these processes,

⁴ Statistical Data and Metadata Exchange (SDMX) is an initiative by international organizations to facilitate the standardized exchange of statistical information. The SDMX Content-Oriented Guidelines are a key standard for the harmonized dissemination of statistical data. See: http://sdmx.org/wp-content/uploads/2008/02/sdmx_content-oriented-guidelines_draft_february_2008.pdf

verification tests should be complete, in the sense that they should cover all operations. For other processes that may not impact on the data, sample verifications could be sufficient, especially if a continuous quality assessment programme was implemented. Validation tests are evaluative exercises that analyze planned tabulations of the census data to ensure that there is consistency. This is especially important when there are perturbations added to the final tabular counts due to disclosure control. Validation tests should compare the totals, frequencies and distributions produced. Particular care should also be taken in the aggregation of small areas to make sure that relations between variables or sets of variables and relationships between domains are kept.

24. Furthermore, a number of external checks to the current available data will ensure that there is external coherence in the data. The census information, after imputation and editing has been completed, will have to be checked against prior censuses so as to identify any incongruities. Additional checks of external coherency can be carried out based on comparison of census results against other statistical information either from surveys or administrative sources administered by the national statistical office or other external bodies.

25. The assessment of the coherence could include:

- i. the conduction of verification and validation tests undertaken before dissemination of the data;
- ii. clear indication of definitions, concepts, frameworks and classifications to be used for the census in a consistent way;
- iii. for concepts, definitions, classifications for which international standards exist, these should be adhered to in order to foster international comparability;
- iv. an adherence to standard classification of variables and tabular derivations, with explanative text for any reclassification and retabulation;
- v. the use of common question formats so as to make it easier to compare to other surveys – ideally the questions should keep the historical formulation to facilitate longitudinal comparisons;
- vi. details of any unusual trends spotted or inconsistencies in the data.

IV. CENSUS QUALITY EVALUATION FOR DIFFERENT CENSUS METHODOLOGIES

26. The preceding section covered the different dimensions of data quality and presented a general overview of how they can be assessed within a census. However, the methodology adopted to conduct the census clearly influences the assessment of the different dimensions of census quality. This aspect is particularly relevant considering the increasing number of methodological approaches that have been developed in the recent years, particularly in countries in the ECE region. The present section reviews the main methodological approaches to census, and for each of them discusses some issues related to the assessment of the different dimensions of quality.

27. It should be stressed that the methodological approaches considered in this section have been defined in broad terms for the purpose of this paper⁵, and they should not be considered as an exhaustive list of possible census methodologies. Some countries may conduct their census according to variants or combination of these methods, or develop other methods not discussed here.

A. Traditional census - Quality assessment issues

28. The traditional census is intended here as the field enumeration of all individuals at a given moment with exhaustive collection of all characteristics or the collection of selected characteristics on a sample basis (long form/short form). The questionnaire is compiled by the respondents (self-enumeration) or by a census interviewer.

29. The different dimensions of quality and how they come into play in a traditional census are discussed in the next paragraphs.

Relevance

30. In the traditional census, the content and design of the questionnaire plays a central role in determining the relevance of the census. With regard to the selection of topics to be included in the questionnaire, user requirements should be considered taking into account the criteria that census topics should meet, and the limits in terms of costs, number of questions and burden on respondents⁶. Whenever alternative data sources may provide relevant data, their use (instead of the census) should be considered so as to reduce the burden on respondents.

31. With regard to topics that have been included in the questionnaire, the assessment of relevance should take in consideration the following issues:

- i. questionnaire responses should reflect accurately the phenomena under investigation;
- ii. post-census consultations with the stakeholders should inform whether the data received meets their needs;

Accuracy

32. In determining the level of accuracy for a traditional census, the following issues should be considered:

- i. the assessment of coverage is usually conducted through post-enumeration surveys (PES) - an essential requirement is that the PES is undertaken by a different group of field staff than those that participate in the initial census field enumeration;

⁵ A more detailed discussion of the methodologies as well as their relative advantages and disadvantages is presented in the CES Recommendations for the 2010 Censuses, in Appendix II.

⁶ The criteria for selection of census topics are discussed in the CES Recommendations for the 2010 Censuses, paras. 23-28.

- ii. differential undercount should be assessed through a breakdown of the coverage achieved by geography and key demographic variables;
- iii. the question non-response rates and proportion imputed should be reported;
- iv. for countries that use sampling (long form/short form), the sampling error associated should be included in the operational error -; the sampling design and frame should also be assessed, as census incompleteness could be due to frame deficiencies;
- v. the assessment should cover the management of outsourced activities, including adequate contingency provision against failure;
- vi. the assessment should cover errors inherently introduced through the design, collection or processing and corrected for, so that they could be reported.

Timeliness

33. The timely release of the census output is frequently difficult to achieve in a traditional census due to the size of the operation, which may cause unexpected delays for instance in the data entry or data processing phases. Although all efforts should be made to respect the planned release dates, this should not lead to compromised accuracy, and the over-riding factor should be to disseminate data of the utmost quality.

B. Register-based census - Quality assessment issues

34. In a register-based census, data are taken from a number of registers linked to each other, and there is no collection of data specifically for the census. Census results are produced using the method of register estimation, in which registers are used simultaneously to define for each statistical unit the value of the relevant variables. This approach has a number of advantages, including the reduced costs once the system has been set up, the fact that there is no additional burden on the public, and the possibility to produce census data with high frequency (in Finland, for example, register-based census data are produced annually). On the negative side, a register-based census relies solely on the information currently available in registers, and its quality depends largely by the quality of the registers. The universality of the register-based census could also be problematic, unless there is a detailed coverage assessment undertaken to ascertain how well the census enumerates the population in its entirety⁷.

35. There are a number of quality assessment issues that need to be borne in mind as they explicitly affect register-based censuses. These are discussed in the next paragraphs, with reference to the various dimension of quality.

Relevance

36. In a register-based census, the relevance dimension is particularly important and critical for the overall quality of the census. In fact, there is no census questionnaire and the census can

⁷ More information on register-based censuses is available in the publication “Register-based statistics in the Nordic Countries – Review of best practices with focus on population and social statistics”, prepared by the Nordic Countries and published by UNECE (2007)

only provide data on topics that are covered in the registers considered. This may have an impact on the census relevance, for instance if some census variables are not available in the registers, or are available but do not reflect accurately the users' needs, because they may be based on different concepts or definitions.

37. In assessing the relevance of data from a register-based census, the following issues should be considered:

- i. consultation and feedback from users and other stakeholders is important also for register-based censuses, even though the mechanisms may be different from those of a traditional census, since there is no census questionnaire to develop;
- ii. during the consultation with users and public in the planning phase, considerations should be given to how to best meet their information needs using the data from available registers taking into account their characteristics;
- iii. data from registers in most cases are not primarily intended for statistical use, therefore the relevance of the variables should be investigated in detail, in particular when various registers are used for the same variable;
- iv. the interaction and collaboration between the National Statistical Office and other agencies and departments responsible for the various registers to be considered in the census is of primary importance.

Accuracy

38. A properly maintained register can provide very accurate data, but registers in their very nature are susceptible to some errors that need to be taken into account. As such the coverage of the register-based census could be incomplete, and so like the traditional census a coverage assessment should be carried out to ascertain the coverage errors (especially the erroneous enumerations). Moreover, register based statistics obtained from different geographical areas may have varying degrees of accuracy due to the differences in administration, database management and so on.

39. The coverage and quality of the registers used for the census clearly play a fundamental role in determining the accuracy of the census results. However, it is also very important to assess the way data from the different registers are integrated. In this context, the analysis of the matching rates (direct matching, statistical matching, and overall matching rates) can provide useful information to assess accuracy. For topics on which information is taken from different registers, accuracy could be assessed also by indicators measuring the consistency of values from different sources⁸.

⁸ See para. 26-28 in "Quality assessment of the register-based Slovenian census 2011", paper prepared for the UNECE-Eurostat Meeting on Population and Housing Censuses, Geneva 13-15 May 2008 (ECE/CES/AC.6/2008/6)
<http://www.unece.org/stats/documents/ece/ces/ge.41/2008/6.e.pdf>

Timeliness

40. In principle, it could be expected that register-based censuses produce results in a timelier manner compared to traditional censuses because high labour intensive and time consuming operations like field data collection or data entry are not required. However, there are a number of complex procedures that need to be implemented to link the various registers and produce the census data. Moreover, the time needed to produce the census variables may vary from one subject to the other, due to different routines for updating the different registers. This may result in some census results to be released timelier and others to be released less timely compared to traditional censuses.

Interpretability

41. The nature of the metadata is significantly different for register-based censuses, and there could be a considerable amount of supplementary information needed to suitably interpret the data. Additional metadata information could relate to the register(s) used, the creation of units and variables in the register, the different sources that make up the register and any changes to the administrative system. The national statistical office should have access to all the metadata from the various registers, and decide which information should be provided to the users to help them interpret the data.

Coherence

42. Internal coherence of register-based census data should be assured in the phase of data integration. Any inconsistencies or ambiguities that are found at that stage should be resolved, by making sure that there is coherency at a micro-level. At an early stage, there should be a harmonization process aimed at bringing together the different registers under one conceptual and definitional framework. This harmonization could extend to other existing non-administrative sources to facilitate external coherence.

C. Census based on registers combined with sample data - quality assessment issues

43. Some countries have registers that could be used to produce census data, but the registers do not contain all data required for the census, or their quality is not sufficiently good to rely exclusively on them. An option is to integrate register data with results from sample surveys that are weighted to population totals.

44. The Netherlands developed this approach for the 2001 census (known as the Virtual Census) where register data were integrated with results from existing household surveys, namely the Labour Force Survey. Another approach was developed in Israel, where the population register is used as a basis for the enumeration of the population, and an ad-hoc sample survey is conducted to evaluate the accuracy of the information obtained from registers and to collect the traditional long-form census data.

Relevance

45. Compared to pure register-base censuses, this approach may provide more relevant data given the possibility to rely on data from sample surveys, in addition to data from registers.

However, data from sample surveys because of their nature will provide limited geographic and information detail, and this may affect their relevance.

46. There should be a programme review undertaken to assess the relevance of the registers and the sample surveys. Concerning data from sample surveys, in assessing relevance consideration should be given to the fact that they are normally representative only at the national and perhaps regional level, but clearly not for small areas. The sampling design also could have an impact on the results, and possible implications on relevance should be considered.

Accuracy and Coherence

47. Combining data from administrative sources to surveys has the advantage of the promotion of coherence between the different statistical data collection instruments of the national statistical office. For example in the Dutch census a method called *repeated weighting*⁹ ensures that there are no inconsistencies between the table estimates from different sources. An estimate of the accuracy (i.e. the error associated with the production of the re-weighted tables) is given, which facilitates the interpretation of the disseminated output to users. In the Israeli approach, one of the main purposes of the ad-hoc survey is to evaluate the quality of the register data. In this sense, the survey could be considered as a sort of PES for the administrative data.

Timeliness and Accessibility

48. The timely dissemination of census data depends among other factors on the synchronization of the data collection processes of the different surveys used. Once the mechanisms are in place to produce tables by counting from the available register information or weighting up sample information to population totals, the resulting census tables can be made accessible to the general public.

Interpretability

49. Micro-integration¹⁰ could be viewed with suspicion if the methodology adopted is not transparent and no one from outside the national statistical institute will be able to easily reproduce the census results. This may be countervailed by including clear and complete metadata on the characteristics of the different registers and surveys, and the operational processes used to combine and integrate the data.

50. Metadata should also specify which census data are produced from what so that users may be aware of any possible implications to their analyses. For example, data derived from

⁹ Repeated weighting is an imputation procedure, but differs from conventional imputation in that there is not a mass filling of the missing data through the plugging in of attributes of appropriate donors found from the data or other sources.

¹⁰ In micro-integration the data from the different data sources (surveys and registers) are integrated at the individual level with the aim of producing a single set of reliable and consistent results.

sample surveys may not meet the level of statistical significance or geographical detail required by some census users.

D. Census based on registers combined with full enumeration - quality assessment issues

51. Some countries for their census combine the use of register data with a complete enumeration of the population. As for the previous case, there are two main objectives of this type of census. Firstly, to improve accuracy of the population counts. Second, to obtain census variables that are not available from the administrative sources, or cannot be easily collected through a register. The obvious advantage of an intensive field enumeration of the population - instead of a survey as in the previous case - is that there is less need to worry about estimation, modelling, weighting and other sampling errors. It also becomes possible to check the coverage of the population register.

52. Many of the considerations made for the previous methodology (registers combined with sample surveys) are valid also for this case. However, there are some additional issues that are discussed below.

Accuracy

53. By expectation, the population counts in this approach should be more precise than in a traditional census or a fully register-based census. Unlike in the sample survey-based census where there are questions surrounding small area estimation, this type of census permits maximum geographical and conceptual detail. Further, the register-data is probably better served as a benchmark for the data collected through field enumeration, whilst the fieldwork can be used to cross-check register records.

Timeliness

54. The timeliness of the release of the census data can be affected. This is because a full enumeration of the population, although with a shortened questionnaire, will be a pretty huge and time-consuming operation. It must be remembered that the normal traditional censuses operations (preparatory work, testing, enumeration, processing) will have to be carried out. In addition, these data will have to be integrated and made consistent with the data from the registers.

Interpretability

55. Metadata should provide information the different administrative sources used and on the field enumeration procedures, including how the questions that appeared in the questionnaire were chosen. In addition, a glossary of terms, definitions and concepts should be prepared. Documentation should also cover the micro-integration processes, to help the users to understand how the final data were produced.

Coherence

56. Internal coherency is probably facilitated in this type of census as the field enumeration is used to cross-check the register information and vice versa. Albeit any edit and imputation

strategy will be using donors that are expected to be closely linked to the reality because the field enumeration will hopefully yield more donors, this harmonization of the field enumeration to registers is not a simple undertaking.

E. Rolling census - Quality assessment issues

57. The so called “rolling census” methodology consists in a continuous survey covering the whole country over a period of time. In France, where this approach has been implemented since 2005 based on a 5-year cycle, data collection is different for large municipalities, where sample surveys are conducted each year, and small municipalities, where an exhaustive data collection is conducted every five years on a rotation basis. Some issues that arise in the assessment of quality, based on the French approach, are discussed below.

Relevance

58. One of the advantages of the rolling census is that, since data collection is conducted continuously, with annual waves, it is possible to modify the questionnaire in relatively short time in order to provide information on emerging topics. This should contribute to improve the relevance of the census, even though a price may have to be paid in terms of comparability of the results across time.

59. When assessing relevance, it should also be considered that this approach allows producing annual updates of the results. This is particularly important for phenomena that may evolve relatively quickly in the society, and that would be difficult to monitor through a traditional census that would produce data every ten or five years.

Accuracy

60. There are a number of issues on the accuracy of the rolling census which have been raised. Firstly, in large municipalities the census relies on sample data. In this case, the coverage assessment should cover also the sampling errors associated with the design. Secondly, although care is taken to ensure that the dwellings register used as a sampling frame is kept as up-to-date as possible, some properties will unavoidably be missed, so there needs to be a regular assessment of the frame coverage. In small municipalities, where an exhaustive data collection is conducted, the same issues discussed with regard to traditional censuses should apply.

Timeliness

61. One of the main advantages of this approach is that data collection and processing are distributed over time. This, in combination with the adoption of sampling for large municipalities, should allow producing results in a timely manner compared to other approaches.

Interpretability

62. The concept of a rolling census is complicated - even to statisticians. Therefore, it is particularly important that clear and complete documentation be provided at different levels of detail, for experienced users and the general public. The documentation should explain how population figures that are universal and refer to the same time frame are produced on the basis of the data collected annually.

Coherence

63. External coherence with other censuses could be an issue in a rolling census, in particular in terms of how it meets the different census criteria set in the CES Recommendations of individual enumeration, simultaneity, universality, small area data and defined periodicity. For some of these criteria, like individual enumeration and simultaneity, an assessment of the compliance with the criteria could contribute to a comprehensive evaluation of the external coherence of the census results.

F. Census based on traditional enumeration and yearly updates - Quality assessment issues

64. This method, developed for the first time in the USA, is a variation of the traditional census method with sampling (long form/short form). Here a very large annual household survey (the American Community Survey) provides the detailed characteristics of the population, replacing the census long-form. However, the basic demographic characteristics are still collected every decade through an exhaustive traditional field collection, with the dual purpose of benchmarking and general coverage improvement.

65. The quality assessment issues that affect a census of this kind are pretty similar to those of a rolling census. However, since once every decade there is benchmarked information collected on every individual, the issues of individual enumeration and simultaneity that surround the undertaking of a rolling census are somewhat avoided, at least for the basic short-form information collected.

Relevance

66. As for other methods based on sampling, the recourse to a sample survey (although a very large one) to provide information on the detailed characteristics of the population can affect the relevance of the results, which would not be available for small areas. On the other hand, the annual survey gives the possibility to include in the census new emerging topic, and follow annually the evolution of phenomena that may change rapidly.

Accuracy

67. An issue that could potentially affect accuracy is response rate for the annual survey. In fact, given its very large size (250,000 households are sampled every month), there is the risk that in the long term the response rate could fall. Consideration should be given to what would happen to the accuracy of the continuous estimates if the response falls below the projected levels. Similarly, the risk that households will be sampled more frequently could lead in future to apathy and scarce cooperation by the respondents. Moreover, the estimates of the population produced by the American Community Survey are liable to significantly larger margins of error as compared with the long-form sample in the decennial census. The main reason for this is that the sample used in the continuous survey is much smaller than the decennial census, leading to lower precision of the small area estimates.

Timeliness

68. The distribution over time of data collection and processing, in addition to the adoption of sampling, should allow this method to produce timelier results compared to most other methods. In particular, for detailed characteristics the American Community Survey should represent an improvement over the traditional decennial census in terms of timeliness and frequency.

Interpretability

69. Besides the metadata associated with a traditional census there needs to be information on the weighting scheme used to produce the population estimates, including the sampling mechanism and statistical model assumption. The move from long-form sample data to the continuous survey will have to be supervised properly; particularly the user expectation of what type of data is produced by the American Community Survey will have to be managed. Users will have to be made aware, especially in the initial stages, of the challenges and the impact of producing continuous estimates instead of decennial estimates will have. Clearly, there is an associated learning curve and connected resources required by users to make the transition from the once in a decade long form sample to the continuous sample of the American Continuous Survey.

Coherence

70. External coherence should be assessed with results from previous censuses. This would allow assessing the possible effect of the methodology on the census results.

V. CENSUS QUALITY EVALUATION ACTIVITIES CONDUCTED IN ECE COUNTRIES IN THE 2000 ROUND

71. As mentioned in section 1, the most common approaches adopted for census quality evaluation are PES, demographic analysis and comparison with data from other sources. Given that the validity of all these approaches depend on the quality of the data used for the comparison (and on the modality with which the PES is conducted), countries usually select more than one method in order to have a more complete and reliable assessment of census quality. In the 2000 census round, more than 60% of the countries in the ECE region conducted two or more evaluation activities (see table 1). The data collected by ECE show regional differences: most countries in the European Union and some in South East Europe conducted coverage PES. In Eastern Europe, Caucasus and Central Asia (EECCA), instead, census evaluation is conducted mainly through field re-interviews.

Table 1. Number of evaluation methods used in ECE countries in the 2000 census round

Number of evaluation methods	Number of countries
Only one method	9
Two methods	13
Three methods	6
Four methods	6
Five methods	3
No evaluation activities	7
Total	44

Source: UNECE survey, 2004.

72. Demographic analysis and comparison with other sources were the most common evaluation approaches in the 2000 round, adopted by about half of the countries (23 out of the 44 that reported this information to ECE). PES was adopted by 20 countries to evaluate coverage, and by 12 countries to evaluate quality (see table 2)¹¹.

Table 2. Evaluation methods used in ECE countries in the 2000 census round

Evaluation method	Number of countries (out of 44)
Quality PES	12
Coverage PES	20
Demographic analysis	23
Field re-interviews	14
Comparison with external data	23
Other method	2

Source: UNECE survey, 2004.

73. The countries that took in 2000 a pure register-based census did not conduct any evaluation, or conducted a quality PES or comparison with external data. Among the countries that took in 2000 a census based on registers combined with full enumeration, in general several evaluation methods were adopted, including coverage PES, demographic analysis and comparison with external sources.

¹¹ Source: UNECE survey on national practices in the 2000 census round (conducted in 2004). More information is available in "Measuring population and housing – Practices on UNECE countries in the 2000 round of censuses" (forthcoming publication, United Nations), also available at: http://www.unece.org/stats/publications/Publication_on_2000_censuses.pdf

VI. CENSUS QUALITY EVALUATION: CONSIDERATIONS FROM AN INTERNATIONAL PERSPECTIVE

74. As the results presented in the previous section show, there is a wide variety of approaches to census evaluation across the various countries. Several countries in Western and Central Europe and North America have a long tradition of census evaluation, which is often conducted using a number of different methods including a PES aimed at evaluating coverage and quality. This group includes mostly countries with traditional census, with an increasing number of countries that are moving to alternative approaches based on the use of register and other sources. The countries with register-based censuses, in general, have limited evaluation programs, often restricted at evaluating the quality, while evaluating coverage seems to be more complex. The EECCA countries, with traditional censuses, have limited evaluation programs based mainly on field re-interviews, a method that does not provide accurate information on census coverage and quality.

75. International organisations have been increasingly promoting among member countries the importance of census evaluation and the adoption of adequate evaluation programs. The new CES Recommendations for the 2010 round of censuses cover census evaluation in detail for the first time (see paras. 83-88 and Appendix V).

76. In the EU programme for the 2011 censuses, census evaluation is given an important role. The EU Framework Regulation on Population and Housing Censuses (in the version currently being discussed by the European Parliament and Council) includes an article explicitly dedicated to census quality. The Eurostat Task Force on Census Legislation, among other tasks, is working to identify criteria to assess the quality of the next round of censuses.

77. The work on the quality reports being conducted at EU level in view of the 2010 round of censuses is particularly important because it is probably the first time that there is an attempt to assess the quality of census data at the international level. On the other hand, this initiative is challenging from the methodological point of view, considering that a large number of countries is involved in the exercise, including a very broad range of census methodologies, including traditional census, register-based census, rolling census etc. The methodological work that is being conducted at the EU level could be beneficial also to other countries, because it could provide a useful standard for census quality assessment.

78. There is another role – potentially very important - that could be played by international organisations in the field of census evaluation: produce and disseminate guidelines and methodological material on evaluation methods, and possibly collect information from countries on how evaluation was conducted (methodology, results, etc.). As shown in this paper, in many countries where the census is conducted in a traditional way there is limited experience with census evaluation, and clear demand to increase knowledge and build internal capacity. Manuals, guidelines and possibly other tools could be developed at the international level to share knowledge on census evaluation methods, in particular Post-Enumeration Surveys. The United Nations Statistical Division (UNSD) has plans to prepare a handbook on evaluation and PES. This is an important initiative to be supported, but other initiatives could also be promoted to develop methodological material on census evaluation methods applicable to alternative census methodologies, including those discussed in the present paper.

VII. CONCLUSIONS

79. No census is perfect and, regardless of the methodology adopted, the results will always contain some errors, caused by a number of factors. These errors may affect in different ways the coverage or the quality of the census data. For this reason, a comprehensive evaluation program should be planned as an integral part of the census, aimed at assessing the quality of the census process as well as the results. In this way, users will be provided with some measures of the quality of the census data, which will help them in interpreting the results. The results of the evaluation could also be used to adjust the census results, as it is done in some countries, in order to provide the best estimates on the size and characteristics of the population.

80. Measuring census quality, however, is not an easy task. There are a number of methodological issues and challenges in the measurement of the different dimensions of quality. Moreover, many of these measurement issues depend on the census methodology adopted, and in the ECE region there is a wide and growing variety of methodologies.

81. This paper tried to review and discuss some of the main issues in the measurement of census quality, considering the specificities of different census methodologies. The discussion, clearly limited in scope, showed how complex evaluating the census can be, especially for countries that adopted census methodologies for which traditional evaluation methods do not easily apply.

82. There is clearly need to develop and disseminate methodological material on census quality evaluation. On one side, for some established evaluation methods - like PES - guidelines should be developed to facilitate the implementation in countries that have no or limited experience. On the other side, work should be promoted to develop evaluation methods applicable to the new census methodological approaches. This could be done by adapting or improving existing evaluation methods, or by developing new ones if necessary. In this context, the international organisations may play a coordination role, to promote methodological developments and facilitate the transfer of knowledge among countries.
