



Economic and Social Council

Distr.: General
20 August 2013

Original: English

Economic Commission for Europe

Conference of European Statisticians

Group of Experts on Population and Housing Censuses

Fifteenth Meeting

Geneva, 30 September – 3 October 2013

Item 6 of the provisional agenda

Operational aspects of censuses

Census quality and coverage: first proposals about the CES Recommendations for the 2020 census round

Note by the UNECE Task Force on census coverage and quality

Summary

In early 2013, UNECE countries took part in a survey on 2010 Census practices for which 51 countries responded. A separate report has been produced detailing the analysis of the survey responses that relate to the questions on quality and coverage (ECE/CES/GE.41/2013/8). This report summarises that analysis and suggests potential implications for the 2020 recommendations.

I. UNECE Survey Results and implications for the CES Recommendations

1. It is encouraging to note that all but one country said they managed accuracy in their 2010 Census round, almost all countries managed at least two dimensions of quality and over 70 per cent managed all six dimensions.
2. Where countries measured the accuracy of their outputs, there was no clear-cut trend in which statistical methods were used most; different methods are clearly best suited to different census collection methods and national contexts. This suggests that it would not be appropriate to mandate any particular choice of method(s) — National Statistical Institutes (NSI) are best placed to choose the methods appropriate to their own domestic circumstances. This view is strengthened by the fact that nearly all countries found their methods effective. The same applies for the editing procedures applied to adjust for errors or inconsistent values; the majority of countries (87 per cent) do this, but the methods used vary.
3. For those that managed accuracy, there was an even split as to whether or not adjustments were subsequently made to the statistics. This suggests that statistics may not be comparable across countries.
4. Given the importance of the statistics (to the European Union (EU) in particular), and the importance of international comparability more generally, it may be appropriate to recommend that a consistent approach is adopted. As a minimum, it might be sensible to recommend that all countries should at least publish an adjusted national total, with flexibility to choose what to publish below that level.
5. Although accuracy is managed by nearly all respondents, nearly three quarters did not set targets. This implies that the majority of countries are not using targets to directly drive the quality of their outputs. However, setting targets is important so that countries know what quality they are aiming for and can explicitly consider the implications for the design of census processes.
6. The survey indicated that there is general support for an international accuracy target, but this view is not unanimous. Support stands at two thirds of respondents (with 8 per cent against and 24 per cent undecided); a distribution that implies that such targets could be considered for inclusion in the revised Recommendations. However, given that setting specific accuracy targets may have implications for the cost of census operations in each country, this should be subject to significant further discussion.
7. Some 62 per cent of countries are in favour of information about accuracy being published and most of these would like to see the measurement methods also included in the publication.

A. Implications for the CES Recommendations

8. If recommendations are made about the measurement and reporting of accuracy, recommendations could also be made about the other five dimensions of quality. However, it is not clear that there would be sufficient benefits internationally from doing so and it might be more appropriate for decisions about the management of other dimensions to be left to individual NSIs.
9. A potential set of recommendations could therefore be along the lines of:
 - (a) Accuracy should be measured by all countries (by whatever method is deemed appropriate by individual NSIs).

- (b) Accuracy measurements and methods should be published by all countries.
- (c) Whichever method is chosen to measure accuracy it should result in a comparable set of output-oriented quantitative indicators, which would require further development.
- (d) All countries should publish a national total population estimate adjusted to account for the findings of their accuracy measurement activities.
10. In addition, we could consider an additional recommendation that sets a national accuracy target. For example, it could be in the form:

(e) At the national level countries should aim to publish statistics that are accurate to within x per cent of the (unknown) true value

11. As a matter for information, in the 2010 Census round, the Office for National Statistics set a target of being within 0.2 per cent of the true value at the national level for England and Wales. A less stringent value may be more appropriate for an international recommendation.

B. Task Force and Steering Group Response

12. Task Force members were fully supportive of the first four recommendations above but there were some concerns about potential cost implications of the last recommendation.

13. The Steering Group agreed that the last recommendation is a difficult issue, but are happy for it to be taken forward and discussed at the UNECE-Eurostat Census meeting in Geneva in October 2013.

14. The Steering Group also suggested creating a new chapter in the Recommendations covering quality management issues, given the importance of quality and its implications for comparability across member states.

II. Proposed text for the CES Recommendations for the 2020 round

15. The CES Recommendations for the 2010 census round already include extensive guidance relating to quality and coverage:

- Part 1, Chapter I (Methodology) includes sections on “Quality assurance”; “Ensuring quality in an outsourcing environment; and “Evaluation of census content and coverage”
- Appendix IV covers “Quality assurance framework and implementation”
- Appendix V covers “Methods of census evaluation”.

16. This text is quite comprehensive and already covers the substance of recommendations 1–3 relatively well (see paragraphs 85 and 86 in chapter I in particular) and also covers recommendation 4 to a degree.

17. However, to give these recommendations greater prominence, it is proposed that the above mentioned sections in chapter 1 should be moved to a completely new chapter in Part 1 of the guidance with a suggested title of “Quality Management”, with appropriate editing to incorporate and emphasise the recommendations.

18. It is proposed that the content of Appendices IV and V should be presented in a unique appendix with appropriate editing to be consistent with the proposed new chapter on quality management.

19. At this point, it is proposed that recommendation 5 should not be incorporated into the revised text, but should be reviewed in the light of discussions by the group of census experts.

20. The text proposed for the new chapter on “Quality management” and for the new appendix in the CES Recommendations for the 2020 census round is presented below.

A. The need for a quality management programme

21. The product of any census of population and housing is information, and confidence in the quality of that information is critical. The management of quality must therefore play a central role within a country’s census.

22. A quality management programme is an essential element in the overall census programme and should touch on all activities during planning, the development period, operations like data collection and processing through to evaluation and dissemination of results.

23. A major goal of any quality management programme is to build in quality from the beginning through the sound application of knowledge and expertise by employees at many levels. It will also include reactive components to detect errors so that remedial actions can be taken during census operations. Without such a programme, the census data when finally produced may contain errors, which might severely diminish the usefulness of the results. If data are of poor quality then decisions based on these data can lead to costly mistakes. Eventually the credibility of the entire census may be called into question.

B. Defining information quality

24. It is generally accepted that there are six dimensions of statistical quality:

(a) The relevance of statistical information reflects the degree to which it meets the needs of users. The challenge for a census programme is to balance conflicting user requirements so as to go as far as possible in satisfying the most important needs within resource constraints. This dimension of quality is particularly important in census content development and in dissemination.

(b) The accuracy of statistical information is the degree to which the information correctly describes the phenomena it was designed to measure. It is usually characterized in terms of error in statistical estimates and is traditionally broken down into bias and variance. In a census context, variance only applies in situations where a longer, more detailed, questionnaire is used for a sample of persons or households, or where only a sample of records is processed. Accuracy can also be described in terms of major sources of error (for example coverage, sampling, non-response, response, data capture, coding).

(c) Timeliness refers to the delay between the time reference point (usually census day) to which the information pertains and the date on which the information becomes available. Often for a census there are several release dates to be considered in a dissemination schedule. Typically there is a trade-off against accuracy. Timeliness can also affect relevance.

(d) The accessibility of statistical information refers to the ease with which it can be obtained. This includes the ease with which the existence of information can be

ascertained, as well as the suitability of the form or medium through which the information can be accessed. Even though censuses are conducted primarily to meet the needs of central government, the data obtained are of great value to many secondary users including local administrations, private organizations and the public at large. To maximize the benefit of the information obtained, it should be widely accessible to all of these potential users. Consequently, censuses often provide a mix of free products, standard for cost products and a user pay service for ad hoc commissioned products. The strategy adopted and the cost of the services also affects accessibility.

(e) The interpretability of statistical information reflects the availability of supplementary information and metadata necessary to interpret and use it. This information usually covers the underlying concepts, definitions, variables and classifications used, the methodology of data collection and processing, and indications of the accuracy of the information.

(f) Coherence reflects the degree to which the census information can be successfully brought together with other statistical information within a broad analytic framework and over time. The use of standard concepts, definitions and classifications — possibly agreed at the international level — promotes coherence. The degree of quality on coherence can be assessed via a programme of certification and validation of the census information as compared to corresponding information from surveys and administrative sources.

C. A quality management framework

25. Quality management has five main components:

- (i) Setting quality targets
- (ii) Quality design
- (iii) Operational quality control
- (iv) Quality assurance and improvement
- (v) Quality evaluation and reporting

1. Setting Quality Targets

26. Setting census quality targets for each of these dimensions at the outset of the census programme enables all involved to know what they are aiming to achieve and, crucially, to determine what it will cost. Early publication of these targets also involves stakeholders, users of the data in particular, to comment and feed in their requirements. In reality, there will be iterations of such targets as initial aspirations may turn out to be unaffordable or unachievable in the time available. Having such a discussion is crucial at the outset to enable realistic, affordable targets to be set and stakeholder expectations to be managed.

27. Simplistically, setting quality targets enables an NSI to answer the question “What does good look like?” and enables a dialogue with stakeholders about “How good is good enough?”.

28. It is easier to set targets for some dimensions than others. It is relatively straightforward to set targets for accuracy, timeliness and accessibility. For example, simple targets could be of the form:

- Accuracy. We will aim to produce national population estimates that are within X per cent of the (unknown) true value with 95 per cent confidence.

- Timeliness. We will aim to publish our first population estimates within one year of census day.
- Accessibility: We will aim to disseminate all outputs online.

29. Setting targets for some of the other dimensions is not so straightforward, and it is sometime helpful to consider setting process related, rather than outcome related, targets. For example:

- Relevance: we will consult with users on the required census content at least two years before finalising the content of the census questionnaire.

30. It is clear that even such simplistic targets will have a significant impact on cost and timetable, hence the necessity of considering such aims early in the planning process.

31. It is recommended that all NSIs should set targets for each dimension of quality at the early stages in the census programmes, and that these should be published to enable stakeholder views to be taken into account. It is particularly important to set targets for accuracy.

2. Quality design

32. Having set quality targets, it is necessary to consider whether the census statistical and operation design is capable of meeting those targets. This can draw of experience of previous censuses or wider international experience.

33. Pre-census tests provide a useful vehicle for planning and developing the actual census. Census tests can be conducted as a national sample (useful for testing content, mail and/or Internet response, and other questionnaire-related features of the census) or as a site test (useful for testing operational procedures). Other pre-census testing could involve cognitive testing of the questionnaire, research and testing of the automated processes for address list development, questionnaire addressing and mail out, data collection, data capture, data processing, conducting innovative research into the use of administrative records, improved cost modelling, and improved methods of coverage measurement.

34. Prior to conducting the actual census, a dress rehearsal provides an opportunity to test the full array of operations, procedures, and questions, much like a play's dress rehearsal provides an opportunity to "fix things" before the real event.

35. Such testing should result in a review of the initial quality targets to confirm their achievability. It may at this point be necessary to change budgets, timetables, or the targets themselves if testing has shown them to be unachievable.

3. Operational quality control

36. Because of the size and complexity of census operations, it is likely that errors of one kind or another may arise at any stage. These errors can easily lead to serious coverage or content errors, cost overruns or major delays in completing the census. If not anticipated and controlled during implementation they can introduce non-sampling error to the point of rendering results useless.

37. To minimize this risk, it is essential to monitor and control errors at all stages of census operations, including pre-enumeration, enumeration, document flow, coding, data capture, editing, tabulation and data dissemination. Every national census organization should establish a system of operational quality control.

38. The dimensions of quality, outlined above are overlapping and interrelated and each must be adequately managed if information is to be fit for use. Each phase in executing a

census may require emphasis on different elements of quality. Again, this requires careful design at the outset to identify:

- the types of errors that may occur at each phase of the operation;
- what information is required to enable such errors to be identified, should they occur;
- how this information will be collected in a timely fashion during live operations; and
- what actions will be taken should the error be found to have occurred (ideally before the phase is complete).

39. Given the speed and scale of census operations, this is no simple task and itself requires careful planning and testing.

40. There is no single standard operational quality control system that can be applied to all censuses or even to all steps within a census. Census designers and administrators must keep in mind that no matter how much effort is expended, complete coverage and accuracy in the census data are unattainable goals. However, clear quality targets should sit at the heart of decision making processes and efforts to first detect and then to control errors should be at a level that is sufficient to produce data of a reasonable quality within the constraints of the budget and time allotted.

4. Quality assurance and improvement

41. Once data collection and processing operations are complete, it is essential that final statistics are quality assured and, where possible, improvement made to the results prior to publication if significant problems are discovered.

42. Quality assurance can be through comparison with statistics from other surveys, through comparison with statistics from administrative data sources, or through analysis of information collected as part of operational quality control.

43. Such quality assurance is challenging, and sufficient time should be allowed from the outset to enable such studies to be completed prior to publication.

44. Quality assurance frameworks and implementation are fully described in Appendix IV.

5. Quality evaluation and reporting

45. It is generally recognized that a population census is not perfect and that, despite rigorous quality control and quality assurance, errors can, and do, occur. Most errors in the census results are classified into two major categories — coverage errors and content errors. Coverage errors are errors that arise due to omissions or duplications of persons or housing units in the census enumeration. Content errors are errors that arise in the incorrect reporting or recording of the characteristics of persons, households, and housing units enumerated in the census. A third type of error is classified as operational errors. These can occur during field data collection or during data processing.

46. Many countries recognize the need to evaluate the overall quality of their census results and employ various methods for evaluating census coverage as well as certain types of content error. In fact, some countries, the United Kingdom for example, include coverage assessment as an integral part of the census process and aim to publish all results after adjustment for coverage error. Most countries, however, undertake coverage assessment as part of their evaluation process, as described here.

47. A comprehensive evaluation program should, however, also include assessments of the success of census operations, in each of its phases. Countries should ensure, therefore,

that their overall census evaluation effort addresses the census process (hereafter referred to as operational assessments), as well as the results (referred to as evaluations). Together, operational assessments and evaluations tell us “How well we did.” A third component of a comprehensive research program includes experiments. Experiments tell us “How we can do better?”

(a) Operational assessments document final volumes, rates, and costs for individual operations or processes, using data from production files and activities; quality assurance files and activities; and information collected from debriefings and lessons learned. Operational assessments can include some discussion of the data, but do not involve explanation of error. The final volumes, rates and costs can be broken out by demographic, geographic level, and housing unit and/or person-level data at intermediate stages of operations or processes. Operational assessments may also document operational errors, although they won’t necessarily include an explanation of how those errors affect the data.

(b) Evaluations analyse, interpret, and synthesize the effectiveness of census components and their impact on data quality and coverage using data collected from census operations, processes, systems, and auxiliary data collections.

(c) Experiments are quantitative or qualitative studies that must occur during a census to have meaningful results to inform planning of future censuses. The census provides the best possible conditions to learn about the value of new or different methodologies or technologies and typically involve national surveys with multiple panels.

48. Evaluation efforts focused on census results should generally be designed to serve one or more of the following main objectives:

(a) To provide users with some measures of the quality of census data to help them interpret the results.

(b) To identify as far as is practical the types and sources of error to assist the planning of future censuses.

(c) To feed into the quality assurance and improvement processes and serve as a basis for constructing a best estimate of census aggregates, such as the total population, or to provide census results adjusted to take into account identified errors.

49. Evaluations of the completeness and accuracy of the data should be made by all countries, and should be issued with the initial census results to the fullest extent possible, including the detail of the methods used. Additional results can be issued after the initial results are published.

50. It is recommended that all NSIs should publish a national population estimate, adjusted in the light of quality assurance and evaluation activities.

51. ***Optional extra paragraph:*** *NSI should aim for this adjusted estimate to be within 1 per cent of the (unknown) true value, with 95 per cent confidence, and should design their quality assurance and evaluation processes to enable this.*

52. More broadly, an assessment of all six dimensions of quality should ideally be made against the initial targets set, with the results published.

53. Such evaluations and measurements can be valuable to indicate priorities and establish quality targets standards for the next census, thus completing the quality cycle.

54. A number of methods exist for carrying out census evaluations and, in practice, many countries use a combination of such methods to fully serve these objectives. These methods are also described in Appendix IV.

D. Ensuring quality in an outsourcing environment

55. Some countries may wish to outsource certain parts of census operations. The motivations and considerations for outsourcing are discussed more fully in Chapter III. In the context of quality management, the outsourcing of components of census operations still requires the census agency to take full responsibility for, and manage the quality of, the census data. This aspect should never be delegated.

56. In setting up outsourcing arrangements, the census agency needs to ensure that it continues to have the ability to both understand and manipulate those elements that contribute to final data quality.

57. Some approaches to outsourcing put an emphasis on “turn key” arrangements — in which contractors deliver systems according to a set of predetermined client specifications with the expectation that the client focuses solely on the outputs and not the internal workings of the system. This assumes that the census agency completely understands and can fully anticipate all data quality issues that might arise during the census and has included these in the specifications. The client is not expected to have any understanding of how these systems work or how they might contribute to the final outputs. Any changes to the system typically require cumbersome processes to determine contract responsibilities and heavy financial costs. This sort of approach effectively hands over the quality of the census data to the contractor, while the risks associated with intervention remain with the census agency. It removes any flexibility and greatly restricts the ability of the census agency to react to quality problems that emerge during processing. This “turn key” approach is not recommended.

58. Suppliers should be made fully aware of the quality targets set at the outset of the census programme, and the quality requirements of the outsourced components that enable the overall census quality targets to be achieved. Operational quality control should apply to outsourced services in the same way as those that are not outsourced.

59. Even when components are outsourced, census agency staff should have an understanding of how such systems work, for example, automatic text recognition engines and coding algorithms, and have the ability to change the tolerances or parameters of these systems at little cost and in a timely manner during processing. Varying these parameters will allow the census agency to determine and manage the appropriate balance between data quality, cost and timeliness as processing progresses.

E. Appendix to the CES Recommendations: Quality management programme implementation

60. As described in chapter III, quality must be managed in an integrated fashion within the broader context of undertaking the entire census programme. Census management will require input and support from all functional areas and it is within this context that trade-offs necessary to ensure an appropriate balance between quality and concerns of cost, response burden and other factors will be made. There needs to be adequate staff with people able to speak with expertise and authority while being sensitive to the need to weigh competing pressures regarding dimensions of quality and other factors to reach consensus. Those responsible for each aspect of census work must be equipped with appropriate expertise. Each of these will develop and implement strategies addressing many aspects of quality. In doing so they must be sensitive not only to their local quality concerns but also to their interactions with quality concerns of others. Strategies to facilitate the necessary information sharing and joint consideration of cross-cutting quality issues are vital.

61. Quality concerns need to receive appropriate attention during design, implementation and assessment. Subject matter experts will bring knowledge of content, client needs, relevance and coherence. Statistical methodologists bring their expertise on statistical methods and data quality trade-offs, especially with respect to accuracy, timeliness and cost. Operations experts bring experience in operational methods, and concerns for practicality, efficiency, field staff, respondents and operational quality control. The systems experts bring knowledge of technology standards and tools that will help facilitate achievement of quality, particularly in the timeliness and accuracy dimensions. In collaboration with subject matter experts, dissemination experts will bring a focus to accessibility and interpretability.

62. This appendix provides further guidance about implementing a quality management programme, building on the ideas that were introduced in chapter III.

63. Firstly, the six dimension of quality are taken in turn, with a description of how the five components of the quality management framework might apply to each.

64. Later sections then provide further detail on:

- operational quality control
- questionnaire design
- management of coverage error
- systems development
- census evaluation.

65. As a reminder, the six dimensions of quality are:

- (i) Relevance
- (ii) Accuracy
- (iii) Timeliness
- (iv) Accessibility
- (v) Interpretability
- (vi) Coherence.

66. The five components of the quality management framework are:

- (i) Setting quality targets
- (ii) Quality design
- (iii) Operational quality control
- (iv) Quality assurance and improvement
- (v) Quality evaluation and reporting.

1. Managing relevance

67. The programs and outputs of a National Statistical Office must reflect the country's most important information needs. Relevance for the census must therefore be managed within this broader context. At the stage of *setting quality targets* it is necessary to discuss how much change in the questionnaire will be contemplated. In cases of severe budget restrictions, some countries have agreed a policy of minimal change or no change. Clearly, this impacts on relevance of the final statistics, but as noted in Chapter III having such discussions at the outset, with stakeholders, is essential.

68. At the stage of *quality design*, relevance is managed through processes to assess the relevance of previous census content and to identify new or emerging information gaps that may be appropriately filled via the census. Major processes to achieve this can be described as: client and stakeholder feedback mechanisms; program review and data analysis. Information from these processes can then be used to ensure the relevance of census content and outputs.

69. Important feedback mechanisms might include consultations with key government departments and agencies, advice from professional advisory committees in major subject matter areas; user feedback and market research; ad hoc consultations with interested groups; and liaison with statistical offices from other countries.

70. While the primary purpose of data analysis is to advance understanding of phenomena, it also provides feedback on the adequacy and completeness of the data used in the analysis. By identifying questions the census data cannot answer it can pinpoint gaps and weaknesses. This must be taken in the context of the analytic potential of other data holdings of the Statistical Office. There is a reduced focus on relevance during *Operational quality control* and *Quality assurance and improvement*, but the emphasis increases again for *Quality evaluation and reporting*, when the published outputs can be reviewed to consider how well they met the originally stated information needs.

2. Managing accuracy

71. Management of accuracy requires attention during all five steps in the quality management framework. Firstly, when *setting quality targets*, as described in Chapter III, accuracy targets should be set as these will fundamentally affect the census costs and design.

72. During *quality design*, parameters and decisions will have a direct impact on accuracy. This includes the design of later components of the quality management framework. The accuracy achieved will depend on the explicit methods put in place for operational quality control and quality assurance and improvement. If these processes are not built in from the outset, including the required data collection processes and feedback loops, it will be much more challenging for them to be implemented effectively.

73. A number of key aspects of design must be considered in every census to ensure that accuracy concerns are given appropriate attention:

(a) Explicit consideration of overall trade-offs between accuracy, cost, timeliness and respondent burden during the design phase.

(b) Adequate justification for each question asked and appropriate pre-testing of questions and questionnaires in each mode of collection, while also ensuring that the set of questions is sufficient to meet requirements.

(c) Assessment of the coverage of the target population. This relates to the adequacy of the geographic infrastructure upon which collection and dissemination geography will be based. It may also relate to the adequacy of address lists to be used in areas where mail out of census questionnaires takes place.

(d) Proper consideration of sampling and estimation options. For example, sampling could be used at the collection stage through the use of short and long form questionnaires in order to reduce respondent burden and collection costs. Alternatively, sampling could be introduced after collection, by processing only a sample of records, at least for a subset of characteristics, in order to produce more timely results or to control processing costs. In either case, careful consideration must be given to the size and design of the sample and to the weighting and other estimation procedures needed.

(e) Adequate measures in place for facilitating and encouraging accurate response, following up non-response and dealing with missing data.

(f) Proper consideration of the need for operational quality control.

(g) Appropriate quality assurance for the final statistics.

74. While individual program managers have considerable flexibility in implementing specific practices and methods, it must be done in an integrated fashion within the overall management of census quality.

75. A good design will always contain protection against implementation errors through for example adequate selection and training of staff; suitable supervisory structures, carefully written and tested procedures and systems and operational quality control procedures.

76. Mechanisms for *operational quality control* should be built into all processes as part of the design. Information is needed to monitor and correct problems arising during implementation. This requires a timely information system that provides managers with the information they need to adjust or correct problems while work is in progress. There is an overlap here with *quality assurance and improvement* and *quality evaluation and reporting* as much of the information collected during operational quality control is also needed to assess whether the design was carried out as planned, identify problem areas and lessons learned from operations to aid design for future censuses.

77. Some examples of activities that could be undertaken to manage and monitor accuracy during implementation and operations are:

(a) Regular reporting and analysis of response rates and completion rates during collection.

(b) Monitoring non-response follow-up rates.

(c) Monitoring interviewer feedback.

(d) Monitoring coverage checks and controls.

(e) Monitoring of edit failure rates and the progress of corrective actions.

(f) Monitoring of results of quality control procedures during collection and processing.

(g) Monitoring of expenditures against progress.

(h) Development, implementation and monitoring of contingency plans.

78. Where applicable, the activities outlined in paragraph 77 should be at different geographic levels or aggregations useful for each level of management, including those suitable for supervising and correcting the actions of groups or individuals involved.

79. Accuracy is multidimensional. Indicators may touch on many aspects of census collection, processing and estimation. Primary areas of assessment include the following:

(a) Assessment of coverage error, both under-coverage and over-coverage. In most countries this is done via a post-censal coverage survey and dual system estimation methods. Comparisons to official population estimates, typically projections from the previous census, are often also used as an assessment tool.

(b) Non-response rates and imputation rates.

(c) Data capture error rates, coding error rates.

(d) Measures of sampling error, where applicable.

(e) Any other serious accuracy or consistency problems with the results. This relates closely to coherence and allows for the possibility that problems were experienced with a particular aspect of the census resulting in a need for caution in using results.

80. Further advice on management of accuracy through management of coverage error and through operational quality control is provided later in this appendix.

3. Managing timeliness

81. Planned timeliness is a decision to be made when *setting quality standards*, refined if necessary during *quality design*, if resources or practicalities indicate that the ideal timeframes are not achievable. There are often important trade-offs to be made with accuracy and relevance. More timely information may be more relevant but less accurate. So, although timeliness is important it is not an unconditional objective. Many of the factors described under accuracy apply equally here. Timeliness is also directly affected by fundamental time requirements to collect and process census data an adequate allowance for *operational quality control* and *quality assurance and improvement*. It might be tempting to aim for challenging timeframes for outputs at early stages in census preparations. These should be tempered by experience gained from *quality evaluation* of previous census operations.

82. Major information releases should have release dates announced well in advance. This helps users plan and provides internal discipline in working towards these important dates.

83. For customized information retrieval services, the appropriate timeliness measure is the elapsed time between the receipt of a clear request and the delivery of the information product to the client. Service standards should be in place for such services.

4. Managing accessibility

84. Census information must be readily accessible to users. Statistical information that users don't know about, can't locate, can't access or can't afford, is of no value to them. In most statistical offices, corporate-wide dissemination policies and delivery systems will determine most aspects of accessibility. Decisions about output dissemination methods and policies are often made late in the census process, as the focus is often on the challenges of data collection and processing. This can lead to later time and resource pressures, to the detriment of accessibility. Explicitly setting aims and policies when *setting quality targets* for the other dimensions of quality can help reduce this impact, as it enables costs and development timescales to be better estimated during *quality design*.

85. In determining information product definition and design, managers must take careful account of client demands. Market research and client liaison will help determine this. The proposed aims and output designs, having been defined, can then be discussed with users, with appropriate modifications, in a controlled way. Whilst not strictly *operational quality control* or *quality assurance and improvement* as defined in chapter III, there are nonetheless parallels to these components in such discussions.

86. In today's world the Internet has the potential to play a role as the primary dissemination vehicle. It should include not only the data released but also information about the data (metadata) such as data quality statements and descriptions of the concepts and methods used. Appropriate links or use should be made of statistical office corporate dissemination vehicles.

87. Finally, as part of *quality evaluation*, client feedback must be monitored on the content of information products and on the mode of dissemination with a view to future improvements.

88. The information needs of the analytic community present some particular requirements. Analysts often need access to microdata records to facilitate analyses. This presents special challenges in order to continue to respect requirements for confidentiality of census data. A number of means could be used to address these needs. Public use microdata files, typically a sample of census records that have been pre-screened to protect confidentiality can be valuable for analysts. Custom retrieval services where specific analyses, designed by external analysts, can be conducted by staff of the statistical office may meet the needs of some analysts.

5. Managing interpretability

89. Managing interpretability is primarily concerned with providing metadata. Information needed by users to understand census information falls under three broad headings: the concepts and classifications that underlie the data; the methods used to collect and process the data; and measures of data quality. The first of these also relates to coherence.

90. A further aid to users is interpretation of census information as it is released. Commentary on the primary messages that the new information contains can assist users in initial understanding of the information. As with accessibility, interpretability can be addressed in *all five components* of the quality management framework.

6. Managing coherence

91. Coherence is multidimensional. Objectives for coherence of census data include: coherence of census data within itself; coherence with data and information from prior censuses; coherence with other statistical information available from the statistical office on the same or related phenomena; coherence with information from censuses of other countries.

92. Aims for coherence should be set when *setting quality standards* as these will drive decisions during *quality design*. For example, there will be trade-offs to be made about the degree with which to standardise across programs within NSI and, for international standards, between countries. Subsequent decisions during quality design will need to be made about the development and use of standard frameworks, concepts, variables, classifications and nomenclature for all subject matters that are measured.

93. The census must ensure that the process of measurement does not introduce inconsistency between its data and that from other sources. Managers for other statistical programs are of course equally responsible for this aspect of coherence.

94. There is less of an emphasis on coherence during *operational quality control*, but the emphasis increases again during *quality assurance and improvement* when the emerging census results can be compared with other available sources (whether published statistics or administrative data for example), which can highlight differences in interpretation of definitions between statistical outputs (or indeed errors in either the census or other surveys — although this is more correctly concerned with managing accuracy, there is an overlap with managing coherence when issues of definition and their interpretation arise).

95. After publication of census results, analysis of census data that focuses on the comparison and integration of information from the census and other sources will give insights to inform *quality evaluation and reporting* and the degree to which quality has been achieved in coherence. The census data should be analysed for domains and aggregations, both large and small that are considered important. Such analysis should consider totals, distributions, relations between variables or sets of variables, relations between domains, growth rates, etc. as appropriate. Comparisons should be made to data from prior censuses and to comparable survey data.

7. Operational quality control

Census activities requiring operational quality control

96. A number of census processes involve massive operations, either manual or automated. Examples of such operations include: dwelling listing operations, preparation of maps, printing of census materials, enumeration procedures, data capture and editing and coding (both manual and automated). Specific operational quality control procedures are particularly relevant and important for each of these.

97. Dwelling listing operations are commonly conducted by enumerators prior to or as questionnaires are dropped off at dwellings. It is particularly important at this stage to minimize both under-coverage and over-coverage of dwellings. To that end, enumerators' procedures must include quality checks to ensure the quality of their work. As well, supervisors should have planned spot checks as listing work starts and planned quality control procedures to be applied as work is completed.

98. When census questionnaires are delivered out, it is usually done on the basis of a list of addresses extracted from an address register. Address register maintenance itself will involve several steps of quality management. Nonetheless, prior to its use, the address list should be validated to confirm that each dwelling is included with correct address and geo-coding information and that no non-dwellings are included. Allowance must be made for dwellings under construction that may be completed prior to the census. This validation is a large operation in the field and is subject to errors. Since this work must be parcelled out to individual employees in batches, acceptance sampling quality control procedures will be appropriate. Again, spot-checking and close communications with supervisors will be important quality assurance steps.

99. Enumeration, whether by interviewing or by collecting completed questionnaires from the dwellings on the list, is similar. Usually one enumerator is responsible for all work in an enumeration area and will be required to implement a number of quality checks on their own work. Further acceptance sampling procedures, implemented by supervisors, will ensure the quality of various aspects of the enumerators' work.

100. Data processing is one of the crucial steps by which raw census data are converted into a complete edited, and coded master file useable for tabulations. In some of these processes the data are being transformed (for example data capture, coding) while in others the data are being corrected (for example edit and imputation). New errors can occur in any of these operations.

Operational quality control methods

101. Clearly a census operational quality control regime comprises a wide variety of mechanisms and processes acting at various levels throughout the census programme. An important technique applicable in many census operations is statistical quality control. It primarily addresses accuracy, although depending on the operation it may also address other elements of quality. What follows is a very brief outline of quality control basics. For a complete explanation of these methods, the reader should refer to a standard text or reference such as Duncan (1986), Hald (1981) or Schilling (1982)¹.

102. The success of any operational quality control programme depends on: laying down quality standards or requirements; determining appropriate verification techniques;

¹ Duncan, A.J. 1986. *Quality Control and Industrial Statistics*. Fifth edition. R.D. Irwin Inc., Illinois.
Hald, A. 1981. *Statistical Theory of Sampling Inspection by Attributes*. Academic Press, New York.
Schilling, for example 1982. *Acceptance Sampling in Quality Control*. Marcel Dekker, New York.

measuring quality; and providing for timely feedback from the results of the programme so that effective corrective action may be taken.

103. Sample verification, complete (or 100 per cent) verification or spot checks are the usual quality control techniques adopted in censuses.

104. Verification can be dependent or independent. In dependent verification, a verifier assesses the work of a census worker by examining that work. However, the verifier may be influenced by the results obtained in the initial operation. In independent verification a job is verified independently by a verifier without reference to the original work. The original results and those of the verifier are compared; if the results agree then the work is considered correct; if not a third, often expert, verifier may resolve the difference.

105. Complete verification theoretically assures a complete check of the work in an operation. However, verifying all items can be time consuming and very costly. In many operations, complete verification is only used as the operation is starting up. Once it is shown that the quality is meeting the required standard, sample verification procedures may be implemented. Usually, this transition is managed on an employee-by-employee basis.

106. Sample verification reduces the cost and can yield results almost as reliable as 100 per cent verification. More experienced and skilled staff often do verification. To be effective the sample must be selected on a scientific basis using probability sampling. It will be designed on the basis of the expected or observed error rates of workers, the outgoing quality to be achieved, the cost of the operation in question and the cost of operating the quality control plan. It will be adaptable to adjust as the quality of work may change. For example, as outgoing quality improves then a reduced rate of quality control sampling may be suitable. Two types of sampling procedures are commonly used: acceptance sampling and continuous sampling.

107. Acceptance sampling is a quality control technique that establishes a sample design and decision rules to determine which batches are acceptable or unacceptable and is usually used in jobs like manual editing, coding, and key entry data capture where work is assembled in lots or batches. Each batch is either accepted or rejected on the basis of the verification of a sample chosen from the batch based on probability methods. The sampling plan is designed so as to provide an outgoing error rate below a certain value, called the average outgoing quality limit.

108. When work is continuous and it may not be possible to group the output into batches for verification, a continuous sampling plan or process control approach may be used. This method is applicable to processes which are fairly predictable in terms of their outputs and which consistently produce output that meets the quality standard — the process is ‘in control’. Statistical process control is a methodology to ensure that such processes stay in control and to provide feedback for corrective action when not in control. Census operations where this may be applicable include: the printing of forms; automated data capture via intelligent character recognition (ICR) or optical mark recognition (OMR); and the scanning of forms for ICR/OMR.

109. Automated data capture, repair and coding systems both increase greatly, and introduce a different set of risks to, data quality compared with traditional census processing approaches. If not properly monitored and managed, data quality problems can remain undetected until late in the process when cost and timing constraints limit the options for any corrective activity. Some methods of measuring data quality from data capture processes, such as substitution rates or measures of key entry errors, are, on their own, inadequate as these forms of monitoring simply measure the overall incidence of errors but not the significance of the errors. Indeed this approach could lead to considerable extra expenditure for the correction of trivial errors that lead to no appreciable gain in quality. For this reason, data quality should be measured at the topic response level rather

than at the individual character/numeral level. This should be done in two ways: independently processing a sample of records using manual processes and comparing the results for each of the records with those obtained through the automated systems; and in aggregate by comparing the overall data for an area with the expected results based on other information for that area (for example from the previous census or other data sources).

110. This process should be undertaken continuously during processing with a focus on early detection of quality problems and an understanding of any systems or processes that have contributed to these. The amount of error that is acceptable and the degree of intervention and systems or process changes undertaken will depend on the assessment by the census agency of the overall fitness of purpose of the output and the overall cost and timeliness impacts. This will vary from topic to topic. For example, it would be expected that there would be a greater focus on the quality of key demographic variables compared with other data items collected on the census form.

8. Questionnaire Design

111. The design of the census questionnaire(s) takes into account the statistical requirements of the data users, administrative requirements of the census, requirements for data processing as well as characteristics of the population. Because censuses often involve multiple collection methods, testing must be performed to ensure that questionnaires will work properly for all applicable methods. The questionnaire should include elements aimed at ensuring accurate coverage of the population (for example who to include, who not to include, where to be enumerated). Qualitative testing is required to check these issues and should cover an adequate variety of situations encountered in the population. In terms of content, quality management approaches for a census are similar to those for a sample-based survey. Qualitative tests and cognitive interviews should be planned to ensure that questions are clear and properly understood not only by the general population but also by special groups to whom certain questions are targeted or for whom there are particular issues of concern (for example the elderly, persons living alone, language difficulties).

112. With the advent of new technologies, introducing web-based questionnaires can provide options not available on their printed counterparts. These options can ensure greater quality in terms of question response and coverage. Such checks serve as opportunities for detecting inconsistencies and presenting them to respondents for correction or confirmation. The design and presentation of a web-based questionnaire to the respondent will differ from the paper version. This means that special care must be taken to minimize any potential mode effects arising from differences between the paper and electronic versions of the questionnaire. Hence, this should be an important topic to be considered in the testing program for the questionnaire.

113. A particular challenge in questionnaire design is to design the questionnaire to be respondent friendly while at the same time, meeting requirements for subsequent processing steps, especially for data capture and coding operations. The testing program must also ensure that these features are thoroughly tested prior to questionnaire finalization.

114. All of these factors should be tested on a small scale (qualitative testing) and then on a large one with a significant number of respondents. A large-scale test can detect a variety of potential issues that qualitative testing cannot. As well such tests make it possible to compare different design and format possibilities via split sample designs. The large-scale test also facilitates assessing how well the questionnaire fits into other census operations (for example collection, data input, coding).

9. Management of coverage error

115. Coverage is a critical element of accuracy. It has a direct influence on the quality of population counts and an indirect impact on the quality of all other data produced by the census. Thus the coverage concerns should be taken into consideration in the design and implementation of most census activities. Enumeration area boundaries must be carefully defined and mapped to ensure no area is omitted or included twice. Instructions and training on dwelling coverage for staff engaged in dwelling listing and enumeration must be clear, explicit and easy to understand. The target population must be well defined and related instructions and questions for both interviewers and respondents need to be carefully developed and thoroughly tested. Clarity and simplicity of instructions concerning place of residence for enumeration is vital to help ensure people are enumerated exactly once and at the correct location. This is particularly important in minimizing overcoverage. Questionnaires should include guidance or questions to assist with situations where it may be unclear whether certain persons should be included or not. Special procedures should be developed for difficult to enumerate population groups (for example remote areas, collectives or group quarters, persons with literacy or language difficulties). Processing procedures should be developed with a view to minimizing the risk of erroneously cancelling, losing or artificially creating households. A well-crafted publicity campaign can play an important role in promoting census awareness and response, thus helping minimizing coverage error.

116. All of these steps, along with appropriate training, supervisory checks and quality control during operations will help minimize coverage error. Nonetheless some coverage error is unavoidable. Hence it is important to measure, analyse and report on coverage error. This is best done via an independent post-census enumeration survey of a sample of census areas or via a Reverse Record Check methodology. Results of coverage studies provide an important evaluation of the current census and can also provide valuable guidance for the next census. Results in conjunction with the census counts themselves are a critical input for population estimation programmes. Analysis of census results vis-à-vis demographic projections of the population from the previous census can also be informative.

10. Systems development

117. Systems development is a cross cutting topic which can have a major impact on quality. In particular the related dimensions of quality are accuracy, timeliness and accessibility. A modern census makes use of numerous automated computer driven systems to operate, manage and control everything from payroll to data capture, edit and imputation, coding, dissemination and others. This pervasive influence makes it very important that an integrated view be taken in the design of the overall architecture as well as the individual design and implementation of systems.

118. A standard methodology for systems development should be implemented and should include steps like: overall system architecture design; design and analysis of individual systems; programming or building of systems; functional testing of components and then of systems; testing of interfaces between systems; volume testing and user acceptance testing; system delivery and implementation; and evaluation. This should be done within a configuration management approach to: manage change; accommodate the reuse of standards and best practices; ensure that all requirements remain clear and valid; communicate each of these to developers and users promptly and precisely; and ensure that results conform to requirements.

119. Specifications must be well written and carefully analysed to produce functional requirements. A standardized approach for change management is required. Ensuring the interoperability of different systems that must communicate with each other is particularly

important. At each stage performance should be evaluated and outputs should be checked to conform to requirements. Many of the systems developed for a census will be used by numerous key entry, coding, editing and other clerical staff. Consequently it is very important that user interfaces be carefully designed and thoroughly tested. More generally, a well-developed standardized testing strategy should be applied throughout in an integrated fashion.

11. Methods of census evaluation

120. Given the importance of the accuracy of census statistics, and the budgets involved, it is common for census programmes to undertake extensive evaluation of the census process. Census coverage and content evaluation are a particular focus of such exercises, but they often also cover wider issues such as the effectiveness of publicity campaigns, project management or contract management.

Coverage evaluation

121. As noted in chapter III, in some countries evaluation of census coverage forms an integral part of the census quality assurance and improvement component of the quality management framework, and feeds directly into the published census results. In other countries, it is a separate exercise, with the results used to adjust the published census results when producing subsequent population estimates.

122. Both gross and net error must be taken into account in developing the overall evaluation plan. Gross coverage error in a census is defined as the total of all persons omitted, duplicated, or erroneously enumerated. Net coverage error takes into account the underestimates due to omissions and the overestimates due to duplications and erroneous inclusions. When omissions exceed the sum of duplications and erroneous inclusions, a net undercount is said to exist; otherwise, a net over-count results. Similarly, both gross and net content errors have to be considered in the evaluation design.

123. The choice of coverage evaluation methods to be used depends upon the evaluation objectives, and the method of undertaking the census (i.e. whether based on field work or population registers). Common methods (that can be applied to both field and register based censuses) include:

- (a) Post-enumeration surveys.
- (b) Comparisons of results with other data sources including previous censuses, current household surveys, and/or administrative records.
- (c) Ethnographic and social network methods to study the effects of mobility on census coverage or to measure census coverage of specific sub-populations.

124. For register-based censuses, it is also possible to compare data from a past traditional census with register data from the same time. Data linkage at the individual level can enabling estimation of both under and over coverage and longitudinal databases would enable such estimates to be carried forward.

Content evaluation

125. Common methods for content evaluation include:

- (a) Post-census surveys designed to measure error in specific census questions.
- (b) Record-checking, in which individual census records are matched against alternative sources and specific data items are checked for accuracy.
- (c) Quality control techniques such as internal consistency checks.

- (d) Surveys to determine customer satisfaction with data collection instruments or questionnaire assistance.
- (e) Focus group interviews to learn how or why respondents behave in a certain way.

Designing an evaluation program

126. The following basic recommendations can be applied to any evaluation program:

(a) Begin planning the evaluation program early in the census cycle. Early planning and design of a structured evaluation program allows appropriate consideration and accommodation of evaluation and experiment needs during the census design.

(b) Decide the high-level scope and focus of research programs before developing research proposals. Define general selection guidelines or criteria, select research topics, and identify high-level research questions before designing the evaluations and experiments. Identify areas to meet the needs of external data users and internal census planners and set evaluation priorities accordingly.

(c) Develop study plans for each evaluation and experiment. These project-level plans become the designated baseline documentation for achieving program research goals.

(d) Develop a standardized Change Control Plan, which describes a protocol to initiate a change process. Recommendations for change (including the reasons for the change and critical implications) are submitted to a Change Control Board. The Change Control Board assesses implications of the change and approves or disapproves it.

(e) Develop a milestone schedule for planning, designing, and implementing the research program. Include in the milestone schedule dates for issuing results of the operational assessments, evaluations, and experiments. Changes to the schedule should also go through the Change Control process.

(f) Anticipate delays or the need to cancel some planned evaluations. During a census, staff may become overburdened with either too much evaluation work or too much of a combination of evaluation and production work. Attrition of project managers is virtually inevitable and can also be a reason to delay or cancel evaluations.

(g) Explore ways to incorporate real-time evaluations during the course of the census.

(h) Develop a Risk Management Plan which identifies potential risk events and their probability of occurring, provides measures of potential impact, offers strategies for dealing with risks if they occur, and identifies the area(s) responsible for addressing each risk event. The Risk Management Plan should be a “living” document where risks can be modified as needed.
