

**Economic and Social Council**Distr.: General
1 June 2011

English only

Economic Commission for Europe

Conference of European Statisticians

Fifty-ninth plenary session

Geneva, 14-16 June 2011

Item 9 of the provisional agenda

Statistical Confidentiality**Results of the consultation on the implementation of the principles and guidelines on confidentiality aspects of data integration undertaken for statistical or related research purposes****Note by the secretariat***Summary*

The Conference, at its 2006 plenary session, requested that a task force be appointed to examine the confidentiality and privacy concerns related to integrated data sets and to develop common principles (ECE/CES/70). The Task Force on Confidentiality and Privacy Aspects of Statistical Data Integration was set up in February 2007.

The *Principles and Guidelines on Confidentiality Aspects of Statistical Data Integration Undertaken for Statistical or Related Research Purposes* were presented to the 2009 plenary session of the Conference for endorsement. The Conference endorsed the Principles and Guidelines (presented in Annex of the document ECE/CES/2011/10), and decided that they should be tested over the next two years, and should be reviewed by the Conference in 2011. The Secretariat launched a consultation process in March 2011 to gather feedback. This note summarises the response to that consultation.

I. Introduction

1. At its 57th Plenary Session in June 2009, the Conference of European Statisticians (CES) endorsed the Principles and Guidelines on Confidentiality Aspects of Data Integration Undertaken for Statistical or Related Research Purposes (presented in Annex of the document ECE/CES/2011/10). The Principles and Guidelines were developed by a Task Force chaired by Australia. The Conference decided that the Guidelines should be tested over the next two years, and should be reviewed by the Conference in 2011. Countries were invited to inform the secretariat about any problems encountered with the practical implementation of the Guidelines.
2. At its November 2010 meeting, the Bureau of the CES decided that the secretariat should ask countries to update the case studies published in “Managing Statistical Confidentiality and Microdata Access: Principles and Guidelines of Good Practice”, or provide new case studies.
3. A consultation was launched in March 2011, where countries were invited to send their observations and report any problems with the practical implementation of these Principles and Guidelines. The results of this consultation are presented in this paper.

II. Summary of feedback

4. Twenty-six countries and two international organisations responded to the consultation.
5. Eight countries had no comments and endorsed the principles (Armenia, Belarus, Bulgaria, Finland, Latvia, Netherlands, Sweden and Switzerland).
6. Fourteen countries provided comments on national implementation of the Principles and Guidelines, often referring to the relationship with national legislation or practices, but did not make any specific proposals for change (Australia, Azerbaijan, Canada, Croatia, Czech Republic, Italy, Lithuania, Mexico, New Zealand, Slovakia, Slovenia, Tajikistan, Turkey and Ukraine).
7. Four countries (Denmark, Greece, Romania and the United States), and the two international organisations (Eurostat and the IMF) made proposals for changes or clarifications. These proposals are set out below in the order of the paragraphs of the Principles and Guidelines to which they refer.

A. Introduction, paragraph 3

“These principles and guidelines apply to data integration work carried out in national statistical organizations (NSOs). In some cases international statistical organisations combine micro-data sets from different countries, but as there are unlikely to be any units in common between the national data files, no confidentiality issues arise here.”

8. **United States:** It is overstated to say that there are not commonalities between national data files, thus no confidentiality issues arise. This assumption depends on what statistical information is going to be released. People move all the time (e.g., retired and military persons) and if data from several years are integrated, people could show up on more than one nation’s registry. Even if micro-data sets from different countries did not contain common units, there could be confidentiality issues involved.

9. **Eurostat:** It has to be noted that international organisations may also match microdata from different sources, e.g. different surveys (such work is being actually carried out in Eurostat) and confidentiality issues may arise in such situations.

B. Introduction, paragraph 5(d)

“Data Matching - the linkage of micro-data from different sources based on common features present in those sources.”

10. **IMF:** Clarify whether data matching also refers to statistical matching in paragraph 6.

C. Principle 2

“NSOs should only undertake data integration activities consistent with their official statistics mandate and after completing a standard approval process (for example, a business case).”

11. **Denmark** agrees with the principles mentioned. Concern is expressed that approval of possible data integration as described in principle 2 and 5 is too heavy. It should be more general and not necessarily an approval process for each project.

D. Principle 2, guideline (a)

“Where a NSO has a mandate(s) that goes beyond statistical and related research purposes, such as involving the use of data for administrative or regulatory purposes related to natural persons, it should abstain from any data integration activities for statistical or related research purposes pertaining to these units, unless this is specifically authorized by law.”

12. **United States:** There is uncertainty regarding what this guideline meant by “units”. Perhaps it is better to say that if an NSO has a mandate to perform non-statistical work, it should be kept functionally separate from the record linkage for statistical purposes to ensure the integrity of the statistical activities and the confidentiality of the statistical data.

E. Principle 2, guideline (c)

“A standard approval process should be followed for any new data integration proposal. This may take the form of a formal business case. An example of a business case outline is given in the Annex, but each country should establish their own template for the process of endorsing data integration projects. The approval process should identify how the integration work will produce or improve official statistics or contribute to related research.”

13. **United States:** Suggest that an additional sentence be added to this guideline that addresses what policies and procedures should be followed to protect the privacy of individuals and to protect the confidentiality of the data. For example something like: The approval process should also identify the policies and procedures set forth to protect the privacy of individuals’ data and to protect the confidentiality of the integrated data.

F. Principle 3

“The public benefits of any data integration project should be sufficient to outweigh any privacy or confidentiality concerns about the use of data and/or risks to the integrity of the official statistics system.”

14. **Greece:** Which is the criterion for public benefits to be characterized as sufficient?

G. Principle 3, guideline (a)

“Data integration should occur in a secure environment and in a manner that does not pose risks to the integrity of the official statistical system.”

15. **United States:** Replace “should” with “must”.

H. Principle 3, guideline (c)

“Where appropriate, bodies with responsibility for ensuring that all benefits, privacy concerns and risks are identified and properly considered by the NSO as part of their standard approval process, should be consulted. The list of benefits should include those resulting from any intended long term retention of, or planned extension over time to, the integrated dataset.”

16. **United States:** Confidentiality of the data also needs to be ensured.

I. Principle 3, guideline (e)

“Where reasonable and practicable, consent should be obtained from the data provider(s).”

17. **United States:** The phrase “where reasonable and practicable” suggests an option to obtain consent from individuals. It may be helpful to acknowledge that informed consent is a necessary condition to data integration in some countries.

J. Principle 3, guideline (f)

“The notions of privacy and confidentiality also require careful management of the risks of indirect identification (typically for units with unusual characteristics), and the increased sensitivity of integrated data sets, which may contain a wider range of variables than any of their sources.”

18. **United States:** Proposes the guidelines should also suggest conducting disclosure analysis on the data.

K. Principle 4

“Data should not be integrated where any commitment has been given to respondents that would specifically preclude such action.”

19. **United States:** An exception could be written in this principle stating if data integration is deemed necessary, the respondents could be re-contacted to acquire their consent.

L. Principle 5

“Integrated data should only be used for approved statistical or research purposes and any significant variation in the originally approved purposes should result in the submission of a new standard approval process.”

20. **United States:** A new approval should be obtained if unanticipated events occur under the currently approved process—such as a data breach or change in policies and procedures affecting the security of or confidentiality of the data. This principle as written suggests that a new approval process is required. We think the intent of this principle is to require a new approval when significant variation of the originally approved statistical or research purpose has occurred.

21. **Greece:** The characterization "significant" for the variation is subjective. The variations in the approved purposes for which the submission of a new standard approval process is needed should be clear and specific.

M. Principle 7, guideline (b)

“The main statistical results of any data integration work should be made publicly available. When data integration work is used to improve the production of official statistics (e.g. through improving quality), the publication of that official statistic meets this requirement. Metadata of statistics published from composite databases should contain information about the original data sources used for data integration.”

22. **Romania:** Proposes the insertion of the following wording at the end of the first sentence: *The main statistical results of any data integration work should be made publicly available with respect for the legal basis for confidential data protection.*

N. Principle 8

“Access to composite unit record data resulting from data integration, but not containing any identifiers, should generally be limited to authorized staff of the NSO. As for other statistical microdata, any proposal to grant access to an external person(s) should have a clear legal basis and be consistent with the purposes of use of data for official statistics. Any person(s) granted such access should provide a legally enforceable institutional and logistical guarantee that their use will be consistent with the approved proposal and that non-authorized persons will have no access to the dataset.”

23. **United States:** It may be helpful if this principle provides greater clarity concerning access to microdata by internal NSO staff and authorized external persons. For example, files that include direct identifiers should only be accessed by a limited number of internal NSO staff. Files that include indirect identifiers only, and no direct identifiers, could be accessed by approved external persons in a restricted environment since individual unit records can still be identified. Also suggest that the phrase, “or research purposes”, should be added at the end of the second sentence. It would then read “...for official statistics or research purposes.”

O. Annex, point K

“A privacy impact assessment should be completed unless a country’s legislative and/or relevant NSO policy provides an exemption. It should also be noted that

although privacy generally relates to natural persons, it may also relate to a legal person in the case of some businesses or industries. For example, certain unincorporated businesses, such as farms, may generate privacy considerations in some countries.”

24. **Greece:** The content of a privacy impact assessment should be fully described.

III. Conclusion

25. In the view of the secretariat, the comments presented above are mainly seeking further clarification of the Principles and Guidelines, and do not indicate any fundamental problems with their implementation or use.

26. The secretariat proposes that the Conference note the comments made and request the secretariat to retain the comments from this consultation for use in a future review of the Principles and Guidelines.
