

**Economic and Social Council**Distr.: General
27 May 2013

English only

Economic Commission for Europe

Conference of European Statisticians

Sixty-first plenary session

Geneva, 10-12 June 2013

Item 4(a) of the provisional agenda

Drivers for micro-data access**Business case for sharing confidential data in a future
European Statistical System: vision and ongoing initiatives****Note by Eurostat****Summary*

This paper presents the vision for sharing confidential data in the European Statistical System and highlights the ongoing initiatives in the field. Provision of access to confidential data for scientific purposes has evolved notably in recent years. Whereas national rules and procedures for access to confidential data for researchers exist, there are no protocols for sharing confidential data between statistical agencies for statistical purposes.

Protection of confidential data is a Fundamental Principle of official statistics. While ensuring trust and privacy of respondents, it may also build barriers against wider integration of statistical data. Under a Schengen-type approach, all the data collected under the European legislation should be considered common good of the European Statistical System and, consequently, each national authority should be authorised to access and grant access to all European data for suitable purposes.

* This document was submitted late due to delayed inputs from other sources.

I. Introduction

1. Regulation (EC) No 223/2009 of the European Parliament and of the Council on European Statistics establishes a legal framework for the development, production and dissemination of European statistics. The European statistics are defined in the statistical programme. Eurostat and national statistical institutes (NSIs) and other national authorities responsible in each Member State for the development, production and dissemination of European statistics constitute a European Statistical System (ESS).
2. The regulation on European statistics allows “transmission of confidential data from an ESS authority that collected the data to another ESS authority provided this transmission is necessary for the efficient development, production and dissemination of European statistics or for increasing the quality of European statistics”.
3. The regulation on European statistics specifies also the conditions for access to confidential data for scientific purposes. Researchers may be granted access to confidential data which only allow for indirect identification of the statistical units. The approval of the NSI or other national authority which provided the data is required for each submitted research proposal.

II. “Vision” for production of European Union statistics

4. European statistics are developed, produced and disseminated on the basis of uniform standards and harmonized methods. The National Statistical Institutes (NSIs) of the Member States collect and produce harmonized data which are compiled by Eurostat in order to construct European statistics at European Union (EU) level.
5. The implementation of the new Regulation on European statistics (March 2009) and the reinforcement of the European Statistical System triggered further discussion on the ultimate model for production of EU statistics.
6. In the so called “vision” document¹ published soon after the entry into force of the new Regulation on European statistics, the weaknesses of the current production method of statistics are pointed out and changes aiming at efficiency gains recommended. The “vision” suggests replacing the current model by a new one in which statistics for specific domains would no longer be produced independently from each other but instead as integrated parts of comprehensive production systems for clusters of statistics.
7. An initiative aiming at this direction is called the Single Market Statistics (SIMSTAT) project concerning statistics on trade between EU countries. The idea of SIMSTAT is to simplify the reporting requirements on intra-EU imports and to make the exchange of micro-data on EU intra-exports compulsory for the ESS members concerned so that the export data can be used to compute the import data of other countries. It requires not only access to all (including confidential) data of the other providers, but also implies that the quality of statistics depends on the sources provided by other members of the system. The expected advantages of this approach are, however, substantial: better consistency of the import/export data (symmetry of the intra-EU trade statistics), efficiency gains, reduced burden on respondents and cost effectiveness.

¹ Communication from the Commission to the European Parliament and the Council on the production method of EU statistics: a vision for the next decade (COM/2009/0404).

8. In order to be implemented this project requires full collaboration between ESS members and adequate legal provisions. It also assumes the existence of the information technology solutions enabling secure access to confidential data for other partners in the system.

III. Sharing of and providing access to micro-data

9. In accordance with the "vision", all the ESS members shall be allowed to have access to the data collected under European legislation if this contributes to the improvement of the data quality and efficiency of the system. Micro-data sharing would be particularly useful in those statistical domains where exchange of micro-data would enable statistical offices to collect data otherwise difficult to obtain, for example information about foreign affiliates of multinational companies.

10. National statistical authorities should also be empowered to grant access to European data. In the "Schengen-like" micro-data access policy, it would be possible to access the whole sets of European data from any accredited access point, with the authorization, management and control of access by the appropriate national/international authority. The rights and obligations of the accredited access points (NSIs in the first place) shall be clearly defined and be adapted to the different types of data and access modes. The Schengen-like micro-data access in the EU would consist of:

- (a) A unique recognition mechanism of institutions and researchers accessing EU datasets;
- (b) A central database and access maintained by Eurostat;
- (c) Access to the network via a terminal server solution (remote access technology).

11. Sharing of and providing access to micro-data shall not compromise statistical confidentiality which is a Fundamental Principle of official statistics, and essential for ensuring trust and privacy of respondents. The security and integrity of the systems and of the data should be an overarching principle at every stage of the process.

12. In the next few years, it might also be necessary to develop a model for sharing data with private companies holding "big data". This data, too large and complex to process it with traditional data processing applications, constitute a great public value and a huge potential for official statistics. In order to receive access to "big data", statistical institutes will need not only to guarantee privacy of "big data" subjects but also to persuade the data holders that their commercial interests will not be endangered by the access. The model for sharing of "big data" may recommend creation of an intermediary body to cope with related legal, administrative and technical aspects² and to manage an appropriate platform at European level.

IV. Ongoing initiatives

13. The new Commission Regulation on access to confidential data for scientific purposes, which has been adopted by the European Statistical System Committee in February 2013 and should enter into force in July 2013, is an important step towards the

² See more: *Official Statistics and Big Data*, Keynote address by Robert M. Groves, Provost, Georgetown University; NTTTS 2013 conference <http://www.cros-portal.eu/content/ntts-2013-programme>.

long term vision (see annex 1 for more details on the new Regulation). It introduces the concept of accredited access facilities which will provide access to EU micro-data at the national level. In order to be accredited the access facilities have to comply with specific legal, administrative, technical and security requirements.

14. The concept of access facilities is currently being studied by the ESSnet on “Decentralized and remote access to confidential data in the ESS” described in a separate document for the Conference of European Statisticians (CES). The ESSnet foresees the creation of a network of accredited safe centres located in NSIs where access to EU confidential data would be provided. In the short term Eurostat will remain the central point of the system but more distributed model is considered in the longer term with NSIs checking the results of the scientific analysis carried out in their safe centre, regardless of the source of the data and on behalf of the countries whose data are used.

15. The 7th Framework Programme project “Data without Boundaries (DWB)” works on the broader infrastructure, which includes not only NSIs but also data archives and research bodies. It covers also broader range of datasets, i.e. not only European statistics but also statistics available at national level only.

16. The purpose of the DWB project is to enhance researchers' access to national and European official micro-data. Its aim is to bring together NSIs, data archives and researchers, to agree on standards (for eligibility of researchers and research projects, for metadata, for statistical disclosure control methods etc.) and common views, and to build mutual trust. The project enables researchers to test the access infrastructure and to gain access to other countries' data. One of the interesting ideas brought forward by the DWB project is to establish a Service Centre for Official Statistics, a single entry point providing information on the micro-data (with appropriate metadata) available at the national and EU level. The information on the data provided by other agencies, such as Integrated Public Use Micro-data Series (IPUMS), could also be available through the Service Centre.

17. The involvement of third parties acting in between “official” statistics and the researchers and fulfilling specialized tasks like metadata preparation, trainings, provision of dedicated information about the data, is increasing. In some countries these “third” bodies (e.g. data archives) are not only providing information about the data but also providing the actual data on the basis of the specific agreements with statistical authorities.

V. Conclusions

18. Provision of access to EU micro-data sets for research purposes is one of Eurostat's core activities. At present, the system is managed centrally in Luxembourg. The trend towards a more decentralized model, in which data providers at the national level are directly involved in granting access to EU confidential data, is ongoing. Changes in legal, technological and methodological frameworks are being made in order to develop a modern system for the provision of access to micro-data in the European Statistical System.

19. It is expected that the changes in the ESS business architecture triggered by the “vision” will enhance sharing of micro-data between ESS members and therefore improve the way European data are produced and made available to the users.

20. The role of “third” bodies acting as intermediaries between data providers and data users and contributing to the improvement of micro-data access will probably increase in the next few years. The establishment of a new model for sharing “big data” may become necessary.

Annex 1

Revision of Regulation 831/2002 on access to confidential data for scientific purposes

Access to EU confidential data for scientific purposes began in 2002 with the entry into force of Commission Regulation 831/2002, which established the procedures for access, i.e. the conditions under which access to confidential data transmitted to the Commission (Eurostat) may be granted and the rules of cooperation between the Commission and national authorities in order to facilitate such access.

This Regulation has now been revised in order to take account of the new advanced technologies, changes in the legal framework and various other developments since 2002.

The weak points of the micro-data access conditions established by the Regulation 831/2002:

1. limited list of bodies allowed to get access – in practice access is restricted to European universities and research bodies; other research bodies must go through the long process of accreditation;
2. only anonymised data released and access to original data provided only at Eurostat's SAFE Centre;
3. limited scope of datasets available – currently 6 surveys available as anonymised micro-data files and 2 ready to be used in the Eurostat's SAFE Centre.

Eurostat in collaboration with delegates from NSIs has been working on the revised version of the Regulation. The new regulation is in the process of adoption and should enter into force in July 2013. Its main principles are:

- access to EU data not limited to EU universities and research bodies; recognition of research entities based on the assessment of following conditions: purpose of the entity, publication of the results of research, appropriate organisation structure, safekeeping of confidential data (appropriate security measures in place);
- new modes of access (including remote access) enabled;
- involvement of external partners (NSIs in the first place) in the provision of access to researchers;
- new and cost effective procedures; contract will be replaced by the licence (confidentiality undertaking) covering all future access requests;
- flexibility in incorporating new standards.