

**Working Paper**  
ENGLISH ONLY

**UNITED NATIONS ECONOMIC COMMISSION  
FOR EUROPE (UNECE)  
CONFERENCE OF EUROPEAN STATISTICIANS**

**EUROPEAN COMMISSION  
STATISTICAL OFFICE OF THE EUROPEAN  
UNION (EUROSTAT)**

**Joint UNECE/Eurostat work session on statistical data confidentiality**  
(Ottawa, Canada, 28-30 October 2013)

Topic (iii): Modes of access to microdata

## **Modes of access to EU microdata in the new legal frameworks**

Prepared by Aleksandra Bujnowska

# Modes of access to EU microdata in the new legal frameworks

Aleksandra Bujnowska

Quality, methodology and research Unit, Eurostat, European Commission, L-2920, Luxembourg, e-mail: [Aleksandra.Bujnowska@ec.europa.eu](mailto:Aleksandra.Bujnowska@ec.europa.eu).

**Abstract:** Researchers may access EU confidential data since 2002. Due to various changes in technical, organisational and legal environment, the legal basis for access to confidential data for scientific purposes has recently been revised. The main objectives of the new regulation establishing conditions of access to EU confidential data were: to offer more datasets, to allow for new modes of access, to improve access procedures and to adapt them better to the broader legal framework.

This paper presents new legal provisions for access to confidential data at the EU level. It focuses on the modes of access enabled in the new Regulation. The advantages and disadvantages of different access systems are analysed from the point of view of their suitability for cross border access to microdata.

## 1 Introduction

According to Regulation (EC) No 223/2009 on European statistics (so called statistical law) the access to confidential EU data is limited to Eurostat staff for the production of European statistics. Eurostat may however grant access to confidential data to researchers carrying out statistical analysis for scientific purposes. The statistical law allows researchers' access under three conditions:

1. The confidential data for scientific purposes can not contain any direct identifiers;
2. The approval of the national statistical authorities that transmitted the data to Eurostat is required for each research project requiring access to microdata;
3. The modalities, rules and conditions for access have to be established by the separate Commission Regulation;

Eurostat in collaboration with national statistical authorities and representatives of research community has been working on the new Commission Regulation since 2009 when the statistical law was adopted.

The final version of the new regulation on access to confidential data for scientific purposes was adopted by the representatives of the EU national statistical institutes

(European Statistical System Committee) in February 2013. The Regulation entered into force in July 2013<sup>1</sup>.

## **2 Principles of the new regulation on access to confidential data for scientific purposes**

The main principles of the new regulation on access to confidential data for scientific purposes are:

- access to EU data not limited to EU universities and research bodies; recognition of research entities based on the assessment of the following criteria: 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 (access facilities);
- new and cost effective procedures; contract is replaced by the licence (confidentiality undertaking) covering all future access requests;
- flexibility in incorporating new standards.

In order to be granted access to EU confidential data the researcher's organisation must be first recognized by Eurostat as a research entity and sign the licence (confidentiality undertaking). Once Eurostat has recognised the researcher's organisation, the application for access (research proposal) can be submitted.

The criteria to be fulfilled by the research projects, and the evaluation made by Eurostat to ensure that those criteria are fulfilled remain the same as in the previous Regulation (EC) No 831/2002. In particular the research proposal must include information on the person requesting access, his or her research entity, the data requested and the mode of access. The criteria require that the research proposal state the legitimate purpose of the research, i.e. scientific purpose, and that the results of the research are to be made public. The planned outputs (articles, presentations, books, etc.) have to be specified in the research proposal. The need to use microdata for the research project should be justified.

---

<sup>1</sup> [Commission Regulation \(EU\) No 557/2013](#) of 17 June 2013 implementing Regulation (EC) No 223/2009 of the European Parliament and of the Council on European Statistics as regards access to confidential data for scientific purposes and repealing Commission Regulation (EC) No 831/2002. OJ L 164, 18.6.2013, p. 16.

The National Statistical Authorities are as before closely involved in, and fully informed of, the provision of access to confidential data for scientific purposes. In particular, and in line with the provisions of Regulation (EC) No 223/2009 on European statistics, the approval of the NSIs or other national authorities which provided the data is required for each access request.

### **3 Types of EU confidential data**

Confidential data for scientific purposes are available in two forms:

- “scientific-use files” partially confidentialised data delivered to researchers on electronic devices (CD-Rom, DVD, etc.);
- “secure-use files” available in Eurostat's "safe centre" in Luxembourg (non-confidentialized data);

Scientific use files are especially prepared to make the identification of survey respondents more difficult. Statistical disclosure control (SDC) methods are applied to this data to reduce to an appropriate level and in accordance with best practices the risk of identification of the statistical unit.

SDC methods are not applied to secure use files. These files can be made available to researchers only in the secure environment ensuring that the results of the statistical analysis are not released prior to output checking.

This arrangement allows for different modes of access:

- Remote execution
- Remote access
- Decentralized access
- On-site access

In addition Regulation permits other parties to be involved in the provision of access to confidential data. These parties (access facilities) have to fulfil the specific criteria and be accredited by Eurostat.

### **4 Micro dataset available for scientific purposes**

All confidential data that national statistical authorities transmit to Eurostat for the purpose of compiling EU statistics can be made available for scientific purposes provided that appropriate protection methods are drawn up and applied.

Methods of protection are decided in collaboration with national statistical authorities, taking into account the mode of access, the probability of re-identification, utility, harmonisation and the impact of unlawful disclosure.

The actual application of the SDC methods is performed by Eurostat (in-house or subcontracted) or by the statistical authority providing the data.

Scientific use files are available for the following data collections:

- European Community Household Panel (ECHP);
- Labour Force Survey (LFS);
- European Union Statistics on Income and Living Conditions (EU-SILC);
- Adult Education Survey (AES);
- Community Innovation Survey (CIS);
- Structure of Earnings Survey (SES);
- European Health Interview Survey (EHIS);
- European Road Freight Transport Survey (ERFT);
- Community Statistics on Information Society (CSIS) – as from November 2013;
- Continuous Vocational Training Survey (CVTS) – as from November 2013;
- Household Budget Survey (HBS) –as from January 2014.

Secure use files are available in the Eurostat safe centre (on-site) in Luxembourg for:

- Community Innovation Survey (CIS);
- Structure of Earnings Survey (SES).

## **5 Modes of access to be implemented**

In addition to the two modes of access available now: transmission of scientific use files on electronic devices and access on site in the safe centre in Luxembourg, Eurostat, in collaboration with the Member States, is investigating the possibility of

establishing **decentralized access to confidential data**. It is an intermediary solution between “full” remote access (where researchers can access remotely confidential data stored in statistical office from their workstations) and on-site access (where researchers access confidential data in the secure environment in statistical office). The idea of the decentralized access is to offer access to EU secure use files not only in Luxembourg but also in other Member States, via safe centres located in statistical offices.

This concept 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.

The data will remain in the secure Eurostat environment and will be available via remote connection in accredited national safe centres.

The solution used in French national statistical authority CASD (Le Centre d’Acces Securise Distant aux Donnees – Centre for secure remote access to the data) was tested by the project team working on the implementation of the decentralized access. The project team includes: Destatis (DE, project leader), Office for National Statistics (UK), Hungarian Central Statistical Office, State Statistical Institute Berlin-Brandenburg Statistics, Instituto Nacional de Estadística (PT) and Eurostat.

CASD provides “full” remote access solution to confidential data in France.

In the longer term **remote access** and **remote execution** will be considered. Two options may be investigated:

- Development of the generic solution by Eurostat
- Use of the access facilities already existing in the EU statistical offices (like e.g. CASD in France, OnSite@Home in the Netherlands or MONA-microdata on-line access in Sweden) or in other data brokers (like e.g. Luxembourg Income Studies).

As agreed with the Member States the development of access facilities will be implemented in a stepwise manner starting with national statistical institutes (institutes coordinating production of statistics at the national level). Based on this experience, access facilities may be extended to other statistical authorities and then other facilities.

## **6 Collaboration in SDC methodology**

Protection of confidential data for scientific purposes is carried out in collaboration with national statistical authorities. In 2014 it is foreseen to establish a network of experts in SDC issues from the Member States. This group will produce guidelines and recommendation on particular SDC problems in the European Statistical System.

## **7 Conclusions**

The new Regulation aims at wider access to EU confidential data. It allows offering more datasets by different modes of access. Its flexibility should permit to accommodate various new ways of releasing confidential data for scientific purposes.