

Working Paper No. 32  
ENGLISH ONLY

**UNITED NATIONS STATISTICAL COMMISSION and  
ECONOMIC COMMISSION FOR EUROPE  
CONFERENCE OF EUROPEAN STATISTICIANS**

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

**Joint ECE/Eurostat work session on statistical data confidentiality**  
(Luxembourg, 7-9 April 2003)

Topic (iii): Emerging legal/regulatory issues

**OPINIONS OF THE CEIES 19<sup>TH</sup> SEMINAR ON “INNOVATIVE SOLUTIONS TO  
PROVIDING ACCESS TO MICRODATA”, LISBON 26-27 SEPTEMBER 2002**

**Contributed paper**

Submitted by Eurostat

**Opinions of the CEIES 19th seminar on "Innovative solutions to providing access to microdata", Lisbon 26-27, September 2002.**

1. Much significant research in the social and economic spheres, both fundamental and of relevance to the formulation and evaluation of public policies, can *only* be undertaken with microdata; it cannot be done using published statistics or aggregate records. The CEIES seminar on September 26/27<sup>th</sup> 2002 provided some excellent illustrations of this point. It is reflected in the increasing demand by public policy makers and private sector organisations (including business enterprises) for good research evidence on which to base and evaluate policy.
2. There is thus an urgent need to improve research access to microdata produced by National Statistical Institutes. Typically data come from surveys, but sometimes from administrative records and include both individual and business data. Protection of confidentiality must remain paramount in all statistical records.
3. National Statistical Institutes do not have the necessary resources or staff in order fully to exploit the data they collect. However, it is wasteful of public money if these data are not analysed more fully. In addition, it is burdensome for the public if new data are collected simply because of inadequate access to existing data.
4. Access to microdata is also needed to allow replication of studies. This is fundamental to good science – not just to ensure that mistakes have not been made in earlier analyses but to test the robustness of findings and their applicability across countries. Increasing access to data should be seen as a priority, but it must be subject to maintaining confidentiality so that the relationship of trust between NSIs and those providing data is guaranteed.
5. Increasingly there is a need for Europe-wide comparisons and National Statistical Offices hold many datasets that support this – for example, the Labour Force Surveys. The need for European-comparative data has led to initiatives like the European Social Survey. If the national statistical institutes themselves do not act to meet this demand, they are in danger of being by-passed and policy-makers will look to research institutes and universities or to commercial sources to provide the required information for Europe-wide comparisons.
6. There are several different routes to providing access to microdata. First, it is important to stress that academic researchers and their institutes are willing to sign contracts and undertake to protect the data. Their livelihoods depend on doing so and the experience of the past decade reveals few if any cases where data have been misused. Second, it is clear that anonymised microdata files are the first and most efficient route of access. Evidence from Canada shows that the availability of anonymised microdata files stimulates research and, once produced, the data files have no additional costs.

Where this is not feasible – either because the data are too confidential or too much detail is needed – alternative modes of access need to be found. Traditionally, safe settings are established. However, these are expensive to install and run and it is often difficult for researchers to stay for extended periods at a location distant from their place of work. They are also elitist by virtue of the fact that not everyone can use them and they disproportionately favour those who are in (or near to) the institution in which they are located.

However, it seems that we are now moving to a situation where there are technological solutions which can produce a ‘virtual safe setting’ over the Internet. We feel that this is an area that needs to be explored as a cheaper and much preferred alternative to safe settings. It is important to stress that, in this virtual setting, the researcher needs to be able to conduct both exploratory and confirmatory analysis.

7. Eurostat, under Regulation 831/2002, has proposed that four surveys should be made available as anonymised microdata files (safe data) and that a microdata laboratory should be set up in Luxembourg to allow access to more detail than is contained in the released files. We greatly welcome this move and acknowledge the work done by Eurostat and the National Statistical Institutes to get the regulation passed. We also welcome the fact that this initiative can proceed on the basis of those Member States who accept it and does not require all Member States to agree.
8. It is vital that there is a timetable for the production of the anonymised microdata files and for implementing the registration procedures. The CEIES subcommittee on dissemination policy will be willing to provide all possible help with achieving this, for example by facilitating consultation with scientific bodies, research institutes and universities in providing some examples of registration procedures used in other countries.

At the same time, it is important to begin considering which additional surveys (or register-based information) should be made available at the microdata level in the near future, in order to expand and update the information needed for Europe-wide research. As a crucial example, attention should be given to EU-Statistics on Income and Living Conditions (EU-SILC) as this will replace the European Community Household Panel, which has been discontinued.

9. CEIES recommends that Eurostat should establish the feasibility of a virtual safe-setting as an alternative to a physical safe-setting. If the virtual setting can be put into place, it will be much more cost effective and provide a preferred means of access for the research community. However, if this does not prove feasible the safe-setting should remain as an alternative.
10. CEIES is aware that its proposals for providing greater access to microdata have resource implications for Eurostat and the National Statistical Institutes. They relate both to the total amount of resources available to these organisations and to their internal allocation. In pursuing the goal of greater access, Eurostat and the National Statistical Institutes are entitled to all the support the CEIES can muster at both EU and national levels in their efforts to get adequate recognition of these implications.