Expert Group for International Collaboration on Microdata Access

ADOPTION OF NEW PRACTICES IN MICRON DATA EXCHANGE

OECD Headquarters
2, rue André Pascal
Paris, France

This document was prepared by Ms. Tuulikki Sillajõe, Deputy Director General, Statistics Estonia.

Contact persons: Tuulikki Sillajõe, email: tuulikki.sillajoe@stat.ee;
OECD contact: Mariarosa Lunati, email: mariarosa.lunati@oecd.org

Complete document available on OLIS in its original format
This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.
ADOPTION OF NEW PRACTICES IN MICRODATA EXCHANGE

1. Introduction

1. The aim of the paper is to share the Estonian experience in the adoption of new practices in microdata exchange, supported by the experience in developing new statistical law. This paper, consisting in a description of the policy and implications of innovation in microdata exchange, is an input to the “OECD Expert Group for International Collaboration on Microdata Access” activity “Discovery and reporting best practice in microdata exchange”.

2. The current document tackles only aspects of the dissemination of microdata; other changes in the Estonian law on official statistics are not covered.

3. The State Statistical Central Bureau of the Republic of Estonia was established in 1921 and operated in accordance with its statutes. The first act governing the State Statistical Central Bureau was not adopted until 1938. During the Soviet era, there was no law on official statistics, but a Statistics Act was adopted again in 1990. This was a law typical of the transition period and had to be revised in 1997 according to needs of an independent state.

2. Previous legislation and practice

4. The Official Statistics Act of 1997 was broadly relevant. Its major shortcoming was that it included only 14 sections. Thus, it was also brief about the use of microdata for scientific purposes. The Act stipulated that only data allowing the identification of a natural person could be disseminated for scientific purposes without the person’s consent. In addition there was a government regulation “Procedure for Transmission of Data That Permit Identification of the Data Subject without the Consent of the Data Subject for the Purposes of Scientific Research”. The data of economic entities were not allowed to be disseminated for scientific purposes without the entity’s consent. At the same time, no distinctions were made based on the source of the microdata. Thus, the use of data of natural persons obtained from both surveys and administrative registers was equally allowed, as long as used for scientific purposes.

5. The research institution requiring microdata had to present a simple application to Statistics Estonia, specifying the purpose of using these microdata. The Director General had to answer the request within two weeks and was not allowed to refuse without giving a reason. Microdata were disseminated on the basis of an individual agreement concluded between Statistics Estonia and the research institution. The content of these agreements was not standardised.

6. In practice, there were very few agreements with some major universities. Each agreement had several appendixes concluded over a relatively long time period for the use of different datasets. However, legally, these could not be considered framework agreements.

7. Microdata of the Labour Force Survey, the Household Budget Survey, the Social Survey, the Time Use Survey, etc. were disseminated under these circumstances. Microdata that allowed indirect identification (with light disclosure control applied) were transmitted on CD-ROMs, and always free of charge. Highly sensitive data was used for scientific purposes in the safe centre of Statistics Estonia. These
practices were not publicly known or even communicated, but it were accepted and appreciated by researchers who happened to know about them and were thus able to use the service.

3. Steps of development

Statistics Estonia operates within the area of government of the Ministry of Finance. In 2008 the Ministry of Finance initiated and started to lead an inter-institutional working group for re-drafting the Official Statistics Act. The group included representatives of different ministries, the national central bank and Statistics Estonia as well as researchers.

In order to regulate the dissemination of microdata for research purposes more precisely, the representatives of Statistics Estonia included in the inter-institutional working group studied the corresponding practice of other countries. Based on information gained at international meetings, at conferences, from the websites of statistical offices, but also from the UN/ECE/CES handbook “Managing Statistical Confidentiality & Microdata Access: Principles and Guidelines of Good practice”, the relevant chapter of the Act was drafted. It was decided that the detailed procedures will be described in a separate document to be approved by the Director General of Statistics Estonia. Also, certain institutions were chosen to be visited in order to study their experience and practices more carefully.

In 2010 the representatives of the Ministry of Finance, the Ministry of Education and Research, the Bank of Estonia and Statistics Estonia made study visits to four institutions: Luxembourg Income Study, Statistics Finland, Statistics Denmark and Statistics Netherlands. The working group decided to be realistic in the introduction of a new policy and procedure for dissemination of microdata for scientific purposes.

During the study visits it became clear that it would be possible to take the necessary documentation (e.g. application form, oath of confidentiality, agreement with research institution, etc.) used by other statistical offices, modify it and shortly introduce it at Statistics Estonia. It also appeared that the LISSY software, developed and used by the staff of Luxembourg Income Study, was suitable for the provision of remote execution service to Statistics Estonia’s customers.

4. Current legislation and practice

In August 2010, the fourth law on official statistics in Estonia came into force. By the end of 2010, the Procedure for dissemination of confidential data for scientific purposes had also been approved. So, the new policy of microdata dissemination had been implemented.

Dissemination of microdata is currently governed by sections 35–38 of the Official Statistics Act and the Procedure for dissemination of confidential data for scientific purposes.

All categories of microdata collected for the production of official statistics (social survey data, census data and business data) are available for scientific purposes without the consent of the person, i.e.

---


microdata of both natural persons and all kinds of economic entities. Also, data derived from
administrative records and other databases may be disseminated for scientific purposes.

15. Data permitting direct identification and data permitting indirect identification of a statistical unit
are both allowed to be disseminated for scientific purposes.

16. In case of Statistics Estonia, the following means for access to microdata are available:
   • remote access;
   • remote execution;
   • safe centre;
   • scientific use file on CD-ROM or via FTP-server;
   • public use file on the web;
   • order for information.

17. Remote execution means a service whereby users can use microdata for making statistical
analyses in a manner which precludes the users’ direct access to microdata, i.e. a researcher submits a
query and receives the output later over the Internet. Remote access means a service whereby a researcher
performs the analysis and can immediately see the answer on the screen. Safe centres are located in
the offices of Statistics Estonia in Tallinn and in Tartu. In case of all these three modes of access, the results
are sent back via computer networks after checking for confidentiality. Scientific use file means microdata
to which methods of statistical disclosure control have been applied, in order to reduce the risk of
identification of the statistical unit to the appropriate level in accordance with current best practice. Public
use files are published on the website of Statistics Estonia and do not allow any direct or indirect
identification of a statistical unit, because these data files are prepared by applying statistical disclosure
control methods. Orders for information mean that the staff of Statistics Estonia produce tailor-made
information for a particular user.

18. Microdata are allowed to be used for scientific purposes only by legal persons or agencies, but
not by freelance natural persons. Pursuant to the Estonian law, “a research and development institution” is
an institution specified in section 3 of the Organisation of Research and Development Act, or a university
or another establishment providing higher education of a foreign state or a research institution of a foreign
state, or an institution listed in the relevant Decision of the European Commission. Pursuant to section 3 of
the Organisation of Research and Development Act, a research and development institution is an institution
in case of which: the principal activity is carrying out basic research, applied research or development; the
activity accompanying principal activity is to spread knowledge through teaching, publication or
technology transfer; the results of the principal activity financed from the state budget funds (which do not
involve intellectual property rights) are public information; the membership includes the research staff
necessary for carrying out the principal activity.

19. Students pursuing a Master’s or Doctor’s degree are also considered researchers. The same rules
apply to domestic and foreign research institutions.

20. A legal person or agency in need of confidential data for scientific purposes must submit a
written application to Statistics Estonia. The application must set out the following information: name of
legal person or agency; registration code of legal person or agency; title of the research; objective of the
research; name of the statistical action or a list of data necessary for the research; a list of data which the
applicant has obtained from other sources and which the applicant wishes to link with the data applied for;
a list of persons wishing to use the relevant data during the research (given and surname, personal
identification code, email address); in case the use of personal data is involved, a confirmation issued by
the Data Protection Inspectorate to prove that the applied organisational, physical and information
technology related security measures are sufficient and, if an ethics committee has been founded, also the opinion of such committee; in case the use of sensitive personal data is involved, a confirmation issued by the Data Protection Inspectorate proving that the processing of sensitive personal data has been registered.

21. Statistics Estonia has the obligation to consider each application separately. Applications for the dissemination of confidential data for scientific purposes are reviewed by the Confidentiality Council according to their order of arrival. The Confidentiality Council considers the substance of the application and decides whether the confidential data can be used for scientific purposes, whereas the decision must be made within ten working days from the receipt of all documents necessary for evaluation of the application.

22. The Confidentiality Council consists of the following public servants of Statistics Estonia: the Deputy Director General, the Head of Methodology Department, the Head of Population and Social Statistics Department, the Head of Enterprise Statistics Department, the Head of Price and Wages Statistics Department, the Head of Agricultural Statistics Department, the Head of Information and Marketing Service of the Marketing and Dissemination Department, the person responsible for personal data protection, and the lawyer.

23. The Confidentiality Council meets once a week, if an application has been received. Decisions taken by the Confidentiality Council shall be confirmed by a directive of the Director General.

24. The Confidentiality Council shall consider every dataset requested by the applicant, taking into account two aspects: 1) risk of identifying a statistical unit; 2) the impact that the identification of a statistical unit may have (sensitivity of data). If the identification risk and sensitivity are rated as high, confidential data can be used only in a safe centre, by remote access or remote execution. If both aspects are rated as low, the requested data can also be delivered on removable devices.

25. Statistics Estonia may refuse to disseminate microdata only if:

- it is not convinced that the data will be used only for scientific purposes;
- the applicant wants to use personal data but the conditions provided for in the Personal Data Protection Act are not fulfilled;
- the applicant has previously violated the terms and conditions of an agreement entered into with Statistics Estonia and, in the opinion of the producer of official statistics, the applicant has not implemented sufficient measures to prevent violation of the terms and conditions of an agreement in the future.

26. Before the dissemination of confidential personal data for scientific purposes, Statistics Estonia shall enter into an agreement with the user of microdata stating the purpose of the research, the persons entitled to use the transmitted data for research, the procedure for processing and transmission of data and the obligation to ensure the organisational, physical and information technology related protection of data, and conditions for the destruction of data after completion of the research.

27. All users of microdata mentioned in an agreement shall sign a written confidentiality agreement.

28. The use of microdata is free of charge. A user shall only pay for the tailoring of the data for their needs and purposes (i.e. linking, matching, etc.): the cost is 100 euros for the first set of data per year and 50 euros for every subsequent dataset.

29. Penalties for the violation of relevant regulations are stipulated in section 40 of the Official Statistics Act. The unlawful dissemination of data which have been collected during the production of
official statistics or which enable the identification of a respondent, or the use of data for other than statistical purposes is punishable by a fine of up to 200 fine units (800 euros). The same act, if committed by a legal person, is punishable by a fine of up to 3,200 euros.

30. Since the adoption of the current policy on 19 October 2010, Statistics Estonia has received and processed 63 applications (as of 12 December 2012) leading to the conclusion of 35 agreements on remote access or use of safe centre, and 12 agreements on ftp-services. In case of seven applications, an agreement is being prepared. Five applications have been cancelled and four has been rejected. Two of the agreements have been concluded between Statistics Estonia and the OECD on access to confidential data using remote access.

5. Conclusions and lessons learnt

31. As a result of public debate, the new law on official statistics is quite detailed, including the regulations concerning microdata. The Official Statistics Act includes 62 sections, instead of the fourteen of the previous version.

32. There is no point in reinventing the wheel. This means that by using the best practices already available in statistical offices, a relevant and up-to-date policy for access to microdata can be implemented relatively easily and quickly.

33. Statistics Estonia has observed that supply creates demand; also, a clear and quick procedure enhances access to microdata.

34. Although

• the new policy is one of the most liberal ones in the world,
• the new policy was created in close cooperation with the main stakeholders,
• the microdata collected for production of official statistics are, in practice, more easily accessible for scientific institutions,
• in two years Statistics Estonia has only four times refused to allow access to microdata, because the data were clearly asked for non-scientific purposes;

The change has been perceived as negative by many researchers, and Statistics Estonia has still gained a somewhat negative image as an impeder of research. Therefore, communication with stakeholders and interest groups is an ongoing challenge.