Working Paper No. 38 (Summary)
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 (vi): Software tools for statistical disclosure control

THE STATISTICAL PROTECTION OF THE EUROPEAN STRUCTURE OF EARNINGS SURVEY DATA

Contributed paper

Submitted by Statistics Netherlands[1]

[1] Prepared by Eric Schulte Nordholt (else@cbs.nl).

The statistical protection of the European Structure of Earnings Survey data

The twin ARGUS software has been applied to earnings data from several EU countries. In particular, the application of Mu-ARGUS to Dutch microdata from the Structure of Earnings Survey (SES) identified that the collapsing of the regional variable (at the NUTS 1 level) was sufficient to anonymise the microdata to an extremely high degree. Very similar results were obtained in Italy and the UK.

Nevertheless, some individuals could still be identified. In order to deal with the small number of disclosive cells that remained, Tau-ARGUS was applied to the Dutch tabular data in order to remove any remaining risks of disclosure. In this paper the software package Tau-ARGUS is described that can be applied for producing safe tabular data.

The main techniques used to protect sensitive information are global recoding and local suppression. The package Tau-ARGUS is a product developed in the SDC (Statistical Disclosure Control) project under the Fourth Framework Programme of the European Union. A new version of the package (that includes recent research results) has been released in the CASC (Computational Aspects of Statistical Confidentiality) project that is funded under the Fifth Framework Programme of the European Union. In 2003 again a new version of Tau-ARGUS will be released in the CASC project. In this paper, methods are described that have been developed to protect tables, through various means that either alter the data or restrict access to them. After combining categories of the spanning variables of the tables, the remaining sensitive cells are suppressed using Tau-ARGUS. In this way, statistical disclosure control techniques help in keeping the right balance between data confidentiality and data access.