



**Economic and Social
Council**

Distr.
GENERAL

CES/2001/31
29 March 2001

ORIGINAL : ENGLISH

STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Forty-ninth plenary session
(Geneva, 11-13 June 2001)

**REPORT OF THE MARCH 2001 WORK SESSION ON
STATISTICAL DATA CONFIDENTIALITY**

Note prepared by the secretariat

1. The Joint ECE/Eurostat Work Session on Statistical Data Confidentiality was held in Skopje, The former Yugoslav Republic of Macedonia, from 14 to 16 March 2001. It was attended by participants from: Albania, Armenia, Azerbaijan, Belarus, Bulgaria, Canada, Czech Republic, Denmark, Estonia, Finland, France, Georgia, Germany, Hungary, Israel, Italy, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Netherlands, Norway, Poland, Republic of Moldova, Romania, Russian Federation, Slovenia, Slovak Republic, Spain, Sweden, Tajikistan, The former Yugoslav Republic of Macedonia, Turkmenistan, Ukraine, United Kingdom and United States. The European Union was represented by Eurostat. Mongolia participated at the meeting under the terms of Article 11 of the Terms of Reference of the ECE. A representative from the World Bank was also present. Eight research and academic institutes participated as observers at the invitation of Eurostat.
2. Opening statements were delivered by Mr. Victor Gaber, the Under-Secretary General of the Ministry of Foreign Affairs, Mrs. Svetlana Antonovska, the Director of the State Statistical Office of the Republic of Macedonia and Mr. Heikki Salmi, representative of Eurostat.
3. The provisional agenda was adopted.
4. Mr. Anco Hundepool (Statistics Netherlands) was elected Chair, Ms. Sarah Giessing (Federal Statistical Office of Germany) and Mr. David Thorogood (Eurostat) were elected Vice-Chairs.

5. The following substantive topics were discussed at the meeting:
- (i) Application of statistical disclosure control (SDC) methodology and software in business statistics and social and demographic statistics;
 - (ii) Impact of new technological developments in software, communications and computing on SDC;
 - (iii) Attitudes of respondents towards statistical confidentiality;
 - (iv) Progress in the implementation of SDC methods and techniques in central and eastern Europe.

6. The following participants acted as Discussants: Mr. Lawrence Cox (United States) for topic (i); Mr. Josep Domingo-Ferrer (Spain) for topic (ii); Ms Julia Lane (United States) for topic (iii) and Mr. Eric Schulte Nordholt (Netherlands) for topic (iv).

7. The transition countries attending the Work Session considered the invited paper prepared by the ECE secretariat "Statistical data confidentiality in the transition countries: 2000/2001 survey" (Working Paper no. 43) of great interest. They requested that the secretariat look into the possibility of translating this document into Russian. Eurostat offered to arrange the translation of this document into Russian.

8. The Work Session considered the material on Business Perceptions of Confidentiality presented by the United States to be a very useful methodological material. It was recommended to take it as a basis for further work on UN/ECE guidelines on this topic. The United States volunteered to cooperate with the secretariat and Eurostat on this issue.

9. The meeting recommended that a further Joint UN ECE/Eurostat Work Session on Statistical Data Confidentiality be convened in two years time. It recommended, therefore, that the following text be included in the Integrated Presentation of the Programme of Work of the Conference of European Statisticians:

2.1 Management of information technology infrastructure.

Activities of the ECE

Joint UN ECE/Eurostat Work Session on Statistical Data Confidentiality (2002/2003) to consider:

- (i) Assessment of statistical disclosure limitation methods and software (in conjunction with parallel sub-sessions on different approaches to measurement of disclosure risk and software tools for disclosure of microdata);
- (ii) Impact of confidentiality on new data release techniques (in conjunction with parallel sub-sessions on confidentiality for small area statistics and software tools for disclosure of tabular data);
- (iii) Guidelines for studies on respondent perception of confidentiality.

10. The participants expressed their high appreciation and gratitude to the Statistical Office of the Republic of Macedonia for creating the excellent conditions for the meeting.

11. The main conclusions reached by participants in their discussions on the four substantive agenda items (see para. 5 above) will be prepared by 10 April 2001 and distributed (in English and/or Russian) to participants and other interested parties upon request.

ANNEX

MAIN CONCLUSIONS REACHED AT THE WORK SESSION ON THE SUBSTANTIVE ITEMS OF THE AGENDA

I. Application of statistical disclosure control methodology and software in business statistics and social and demographic statistics

1. The discussion under this topic covered different aspects of disclosure control. Information was provided about plans for further development of software for protecting data in linked tables and microdata. The discussion covered an array of applications, methodology, computational strategies, and empirical findings. Attention was focused on new and stimulating ways to examine six long-standing issues in statistical disclosure limitation: identifying and measuring disclosure, assessing disclosure risks, methods for limiting disclosure, software, computational issues, and the effect of disclosure limitation on data analysis and usefulness.
2. The discussion dealt with the important problems of identifying and measuring statistical disclosure, namely primary disclosure risk assessment, the protection interval, and the problem of common respondents in tabular magnitude data subject to complementary cell suppression. Some widely known disclosure protection techniques were considered in more detail, such as disclosure rules for tabular data and microdata, default values for disclosure rule parameters, procedures based on a p -percent disclosure rule, and the consequences of selecting different values for the threshold parameter in a threshold rule.
3. Another issue considered was assessing and measuring disclosure risks. Disclosure rules and disclosure risk are interrelated. In some situations the disclosure rule is established and risk is based on the likelihood of failing the rule. In other situations, the rule may be set based on analytical or empirical analysis of risk. The discussion focused on measuring re-identification risk based on metrical and nearest-neighbour distances, and the risk from the standpoint of distance-based and probabilistic record linkage.
4. One possible decision theoretic approach to data disclosure problems was presented. Controlling the disclosure risk can be based on two factors: discredit harm to the statistical agency based on intruder claims of achieving disclosure and real increase in intruder knowledge. Such an approach offers a theoretical framework to develop optimality criteria for the choice of the best way of data release. It recognises the different perspectives of the statistical agency and the users in assessing the extent of disclosure and the utility of data. So far, little has been done to measure the impact of SDC techniques on statistical analysis and to assess the trade-off of gains versus risk. Current measures of disclosure are defined as decreasing functions of the intruder's uncertainty. In the reported approach, distinction is made between the intruder's and the agency's perspective and the decision about the best form of data release. It is based on the measure of global risk and global utility. The proposed approach is still theoretical and will be examined in practice.
5. Several papers reported new approaches to methods for limiting statistical data disclosure. These include complementary cell suppression in large and complex tabulation structures that includes table-to-table protection and backtracking strategies, analysing theoretical properties of microaggregation and using mathematical networks for complementary cell suppression in two-dimensional publication tables. Related strategies for organizing table processing were also presented.
6. The effectiveness of several methods for statistical disclosure limitation in microdata was analysed. Based on re-identification experiments, the aim is to optimise the trade-off between information loss and disclosure risk. The compared methods were distortion by probability distribution, microaggregation,

resampling, rank swapping and lossy compression. Sampling methods are suitable for categorical microdata but their adequacy for continuous microdata is less clear. To assess the methods, two measures were used: information loss (how different is the masked data set from the original set) and the disclosure risk. The method that optimises the trade-off between these two measures subject to some user requirements can be considered the best option for SDC.

7. Considerable attention was given to applications based on $\hat{\delta}$ and $\hat{\iota}$ -Argus software. The software has a potential to become a generally applicable standard tool data protection. Facilities will be added to assess the residual disclosure risk, methods for protection of complex hierarchical tables, especially in the context of public use database query systems. With the disclosure risk statement, the user will be able to state how much of certain disclosure risks would be acceptable for him. The software should be able to deal with tables of any size or complexity.

8. Further extensions of $\hat{\delta}$ -Argus are planned to include handling of linked tables, particularly hierarchical tables. More research is needed for the protection of data from consecutive years of the same survey, to have more options available for executing the secondary cell suppressions. As the software can be used with a lot of different protection criteria, it is important to assist the users to understand how different strategies can be executed using the package.

9. Such standard software to be used in many different environments should be flexible and user-friendly. Tools are needed to assist and guide users in the use of the software. The importance of training in this relation was highlighted. Especially the transition countries stressed the need for practical solutions, easy to access and easy to use tools to facilitate implementation of SDC methods in these countries.

10. The possible reasons why the SDC software packages are not developed more widely were considered. There is a sufficient market for such products, and the number of people involved in SDC is increasing. There would be a basis for more cooperation and joint work on software development to stimulate the introduction of good theoretical ideas into statistical practice. However, the software packages are relatively complex and need special training for their development and implementation. Also, they are often based on the specific confidentiality rules implemented in the office which is developing the software. Therefore, in other offices using different approaches the use of the software can be complicated.

11. Many computational issues and challenges surrounding complementary cell suppression and computational complexity issues for microaggregation were addressed. Results of this sort are extremely important in assessing the effectiveness and limitations of practical methods. New computational methods for disclosure audit in a specific class of multi-dimensional tables were discussed.

12. Assessing and limiting the effects of disclosure limitation on data analysis and usefulness is an extremely important but difficult and under-explored area. One possible solution is to provide facilities for researchers to access original data on site (e.g., Netherlands, U.S. National Center for Health Statistics and Eurostat). In this way, researchers can access all the details of original data (except direct identifiers). Export of any information could be possible only with the permission of the responsible statistical officer. It is also possible to submit remotely computer programmes for analysis of original data.

13. Other possibilities to balance the disclosure risk and data utility are through optimal choice of parameters or assigning preferences to candidate cells for complementary suppression. A theoretical framework was provided for assessing interactions between disclosure risk and data utility. The disclosure can be a matter of degree, e.g., divulging confidential information about two respondents is a greater disclosure than divulging the same information about one respondent, or that divulging a narrow

estimate of a confidential quantitative attribute is greater disclosure than divulging a broad estimate. Upon this quantitative framework, one can build a model for disclosure risk incorporating both the degree of disclosure and the likelihood of disclosure.

14. Re-identification studies are extremely important for measuring and assessing disclosure risk that should be conducted in statistical practice. Re-identification risks of register data were considered by matching a sample of register data with survey data. The re-identification studies based on modelling the intruder behaviour are especially important from a practical viewpoint to test the theoretical disclosure solutions in practice.

15. The pros and cons of providing users with the information about exact confidentiality methods used were considered. It was mentioned that in order to be able to model the user's (intruder's) behaviour, it is necessary to give to the user enough information so that their behaviour could be predicted. On the other hand, it was pointed out that the statistical offices' primary interest should be to assist the user and not to prevent the user of making a proper interpretation of the released data. Releasing information about the applied SDC methods can therefore be viewed as fulfilling the promise to provide user with good quality data and with sufficient information to allow proper use of the data.

16. The reorientation of statistics from the stove-pipe approach to the integrated approach has an influence on the use of SDC methods. Different approaches might be required for data warehousing, integrated databases, electronic data reporting, combining data from different sources, etc. The strict delimitation between the two streams of information – data on enterprises and data on individuals can also pose problems. There are generally stricter rules for enterprise data which is often a drawback.

II. Impact of new technological developments in software, communications and computing on SDC

17. The numerous contributions clearly demonstrated that intensive research and development activities are carried out on this topic. The discussion dealt with five thematic blocks, namely access to microdata, protection of microdata, protection of tabular data, (re-) identification and software.

18. Different systems for access to microdata were considered. Three approaches that provide researchers with access to data from complex surveys were reported by Statistics Canada. Public Use Microdata Files (PUMF), remote access and research data centers were explored. Record matching is a very pragmatic approach to assess the disclosure risk for PUMFs. For remote access, automated procedures might be needed. Some initiatives toward automating remote access are currently being considered but more research into the feasibility of this approach is necessary. The possibility to automatically check whether a query is answerable without infringing statistical confidentiality before actually running the query was also discussed.

19. Research in SDC is more necessary than ever. Recent challenges which require more powerful SDC protection include on-line systems offering automated or semi-automated disclosure control. The need to provide customized access to statistical resources appears to be an unavoidable challenge to statistical institutes. The U.S. Bureau of the Census's American FactFinder (AFF) <http://factfinder.census.gov> is a bold initiative toward providing on-line customized access to microdata resources. In its current state, two possible methods of data access are defined: the users can choose from a list of predefined tables or define their own tables.

20. Releasing user-defined tables is a real SDC challenge. The summarized data from user-defined tables will be provided only if they pass disclosure limitation rules. A query filter is supposed to detect queries that will not pass such disclosure rules before they are submitted; the query filter is complemented by a results filter which performs a final check on the resulting table before returning it to

the user. Both the query and the results filter should take into account previous queries by the same user. Otherwise, an intruder can accumulate knowledge through successive queries and eventually succeed in re-identifying an individual. This is the well-known weak point of on-line statistical databases and countermeasures are not obvious.

21. A new methodology for the protection of tabular data was presented. Partial Cell Suppression (PCS) is a possible solution to the Cell Suppression Problem. PCS replaces suppressed cells with intervals rather than “suppressing” such cells. The method can lead to replacing more cells by intervals than complete cell suppression, but the overall information loss can be lower and the method appears to be easier to use from a computational standpoint. The use of modelling languages for quick development of algorithm prototypes for the cell suppression problem was discussed; prototyping was illustrated with a particular network flow method, for which some preliminary computational results were presented.

22. Record matching and data mining are rapidly evolving fields. SDC methods must take those advances into account and provide adequate protection. Ultimately, disclosure risk can be viewed as matching risk. A novel and very realistic re-identification scenario was considered, namely record linkage between two datasets (the original dataset and the SDC-protected dataset) when the variables known by the intruder in both datasets are not exactly the same but similar. Re-identification procedures based on clustering give encouraging results. The concept of “structural information” can be a serious threat to microdata SDC as it is understood today. In fact, the two datasets being matched do not need to share any variables; it suffices that there are variables in both datasets which have a similar structure.

23. Disseminating geographically referenced microdata is a common concern of statistical offices. Current microdata protection methods may not be specially designed for that purpose and often are not usable for such data. Confidentiality protection of small area statistics and geographically referenced data is a promising new area of SDC which should be examined and new methods should be developed if necessary.

24. Experiences on model-based SDC protection were reported. It was discussed whether releasing the model to the user would not undermine the data protection. On the other hand, the model could provide the user with the necessary information to understand the inferential structure of the data. The efficiency of the model-based approach and possible risks of providing the model to the user depend on how the model is built. If the purpose of data use is known, the model can be built in such a way that it would fulfill the user requirements while effectively protecting the data. When the models for different variables are independent of each other they can distort the inferential structure of the data. These questions are usually ignored in using a model-based approach. Generic methods could be used for general release of microdata while model-based approaches could be more suitable for specific research purposes.

25. The objectives and the approach of the main EU-funded initiative on SDC, the CASC project, were explained. The main emphasis of the project is to build practical tools that could be used in real-life situations in official statistics. The primary users of the results are expected to be statistical agencies. Therefore, several statistical offices are active members of the project. The testing stage is considered very important and is included in the project. The side-effects of the project will be that the research community on SDC will work together over Europe. This joint effort will bring the state-of-art to a higher level. The most visible outcome of CASC is an improved version of the Argus software, including protection for large and complex tables and a comprehensive array of microdata masking algorithms. The results of the CASC project will be disseminated through a Website

<http://neon.vb.cbs.nl/casc>

III. Attitudes of respondents towards statistical confidentiality

26. There is a fundamental tension in every statistical agency's mission. Each is charged with collecting high quality data to inform national policy and enable statistical research. This necessitates dissemination of both summary and micro data. Each is also charged with protecting the confidentiality of survey respondents. The trade-off dilemma, which could well be stated as protecting confidentiality (avoiding disclosure) but optimising access, has become more complex as both technological advances and public perceptions have altered in an information age.

27. The discussion addressed a core issue faced in this statistical mission. In particular, statistical agencies enter into an implicit pact with their respondents to protect their data – both because they are mandated to do this by law and because it is necessary to maintain the trust necessary for high quality responses. The papers in this topic shed new light on the need for and ability of statistical agencies to reassure the public that statistical agencies have, in fact, delivered on their confidentiality promises.

28. Statistical agencies have historically relied on two approaches to do this: legal/ criminal penalties for the unauthorized data access and/or disclosure and technical protections for released data. However, these approaches will only enable statistical agencies to maintain high quality responses if there is evidence that the public both knows about and believes in their protection methods. Historical practices may no longer suffice in the wake of new events such as technological change, especially computational progress, and changes in public attitudes to the privacy and confidentiality of their records.

29. The other new events that are likely to influence the perception of confidentiality are the increasing importance of administrative data, the propensity of statistical agencies to disseminate information over the Internet, an increasing demand for micro-data by academic researchers, an increasing demand for timely, relevant information by policy makers, and possibly most important, the increase in data collection by the private sector.

30. It is worth considering in more detail the influence of the private sector because of the complex ways in which it may have changed the landscape of public perception. On the one hand, the vast expansion of private data collection creates a much more sophisticated master file for potential intruders. The public may thus be more aware of and more concerned by the increased potential for re-identification. On the other hand, widespread public awareness of private sector profiling may actually reduce privacy and confidentiality concerns among the public, if they believe that all information about them is already known. Without further research, the net effect of this change is not known.

31. The U.S. Census Bureau has initiated research on household perceptions of the confidentiality protections in conjunction with the 1990 and 2000 Decennial Census. Cognitive interviews were conducted on the wording of confidentiality messages integrated into the Census Bureau's Internet dissemination tool – American Fact Finder - to address identified perception concerns. Although much research has been done on individual attitudes of confidentiality, little is known about the perceptions of businesses in the treatment of their data. This gap is being filled by a new survey and resulting research by the U.S. Census Bureau and the U.S. Internal Revenue Service.

32. Two other papers describe probably the most important new access modality that has been developed in the past decade – that of restricted access sites. These permit statistical agencies to respond to the microdata needs of researchers, avoid the linkage problems posed by the Internet, and address potential perception problems that might be associated with other access modalities. The success of the secure sites may prove to be ground breaking endeavours to demonstrate to other statistical agencies how such sites can be made operational.

33. A third approach is to examine the protection safeguards that are in place. These can be categorized in three ways: technical/legal, organizational and methodological. Some very innovative technical advances were described that are being implemented at Statistics Sweden to automate several of the decision-making processes associated with data release.

34. Statistical agencies should routinely take stock of public perceptions of the way in which they protect different types of data – household/business surveys, censuses and administrative data. The usefulness of this information is potentially great. Confidentiality messages targeted at the respondent can be tailored to address current concerns at the point of data release. If some data are less sensitive than others – either inherently, or as time goes by - then the public might be willing for those data to be less protected, and valuable resources can be used to focus on those data elements that are most sensitive. If the public has become more sensitive to data privacy concerns in response to private sector actions, then statistical agencies need to respond accordingly.

35. The assessment of public perceptions can be an important tool in guiding data dissemination decisions. For example, new access modalities such as restricted access sites, are potentially an important new dissemination mechanism. However, their establishment should be evaluated, not only for the technical protection and data quality issues raised as a result, but also in the context of public perception of these approaches. It may well be that the public perceives these as being much safer than the release of public use files, particularly given the wealth of private data available, and statistical institutes can tailor their response appropriately.

36. The new organizational methods can also be used not only to protect the data but also to assist in public perceptions of the safety mechanisms used by statistical agencies. The use of either automated processing techniques or relatively small, but well structured, confidentiality committees could certainly be used to further reassure the public that their data are fully protected by their national statistical agencies.

37. National statistical agencies should choose disclosure limitation methods based on information on the effect on users' needs and the effect on respondents' perceptions. The best way to do this is to survey both the users of and the respondents to demographic and business surveys. It would be extremely useful to have similar survey instruments applied in different countries, because this would enable statistical agencies to quantify the effects of the different approaches. The lessons learned from such cross-country comparisons would form the basis of informed decision making about SDC methods. One such survey has already been created and fielded on the business side – the aforementioned Census/IRS survey. A demographic survey as an instrument has been developed by a consortium of European institutes but not yet funded.

IV. Progress in the implementation of SDC methods and techniques in central and eastern Europe

38. The numerous contributions from central and east European countries proved that there has been considerable progress in the implementation of SDC techniques in these countries since the previous Work Session on Statistical Data Confidentiality in March 1999 (Thessaloniki, Greece). The progress lies mainly in increasing knowledge about SDC and laying the foundation for its implementation in the statistical office. However, there remains a lot to do concerning the practical implementation of SDC methods.

39. Both successful experiences and problems faced when developing projects on data confidentiality on the national level were described. Several problem areas and ideas for possible solutions were identified that are valid also for the more developed countries. This provides a good possibility for closer

cooperation in SDC between countries. The need for working closer together, learning from each other's experience, identifying good practices and moving towards standard solutions was emphasised.

40. In most of the transition countries, the legal basis for statistical data confidentiality is in place. Several countries have amended the law within the last few years to make it compatible with the EU regulations in this area and to take into account progress in SDC and the impact of technological developments. The basic laws have to be complemented by specific regulations on SDC, and organisational, administrative and technical measures to carry out the stipulations of law in practice.

41. The responsibility for ensuring SDC in the office was discussed. In some offices, the departments disseminating the data are responsible for the correct application of SDC measures and advice can be obtained from legal, technical and statistical advisers. In other cases, the staff who directly work with the data are responsible for guaranteeing confidentiality. However, the final responsibility for SDC lies with the management of the office. The management is obliged to set up the office so that individual departments can carry out their responsibilities concerning data confidentiality. This includes laying down the confidentiality policy for the office, specifying rules and regulations for its implementation, spreading knowledge and organising training. There is an important link between the independent status of a statistical office and the policy on SDC. Only an independent statistical office can thoroughly protect confidential information.

42. Before implementing SDC measures, theoretical knowledge is required. Many statistical offices have some people who know enough about SDC to set up the implementation phase. However, it would be useful to spread the knowledge more widely so as to become less dependent on a few key persons. Training is needed to extend the general level of knowledge on SDC. As soon as the level of the knowledge has reached a reasonable level, a statistical office can initiate the practical implementation of SDC measures. It is not wise to try to change everything at once. The most successful approach could be to improve data confidentiality through several successive projects.

43. It was pointed out that often the implementation of SDC methods is entirely based on personal professional experience and judgement. This can lead to over-protection of data and inefficient application of methods. It would be necessary to outline the official framework for addressing the SDC issues in a systematic way throughout the office. It would also be useful to establish guidelines on how to apply different measures in practice, and on which methods are more suitable to which kinds of surveys.

44. A major problem for most transition countries is the availability of software for statistical disclosure control. Often the confidentiality methods are not used effectively because of a lack of software. Moreover, it takes time to test new software and become an efficient user. Also for this purpose, training is needed. Some countries are testing the ARGUS software. The ARGUS packages are free of charge, so there is no financial problem in testing and applying them to protect microdata and tabular data. (However, in order to obtain an optimal solution, for the time being τ -Argus requires the purchase of a linear programming software which will be replaced by a free software in future.) Training in the use of these packages for official statistics is required by several countries.

45. Many countries have been active in recent years in improving the physical protection of their computer networks. The chain of statistical protection has a weakest link and it is wise to start with an adequate physical protection of the sensitive data before the statistical protection is improved. Physical and statistical protection is not cost-free. Therefore the financial constraints have to be taken into account when defining a reasonable level of protection of statistical data.

46. Confidentiality problems linked with the population census were discussed in more detail. This is an acute problem in most of the transition countries in connection with the recently conducted population censuses. Several organisational, technical, legal and physical protection measures were considered. The

applied methods include strict measures for the physical protection of data and penalties in the case of breach of confidentiality in the data collection, processing and dissemination phase.

47. There are several problems in ensuring the confidentiality of enterprise data. It might be necessary to have slightly different confidentiality policies for enterprises and private persons, and to redefine the rules for enterprise data at the legislative level. Several countries highlighted a dominance problem: often just one or two monopolistic enterprises dominate in certain economic activities. Suppression of such data would imply that the released statistical tables would not meet their goals. A solution could be to make a confidentiality exception concerning monopolistic enterprises on a legal level or to try to get individual consent from these enterprises.

48. An important aspect that should be taken into account when defining the confidentiality policy is the attitude of society towards privacy and confidentiality. An essential and often neglected part of SDC in practice is communication with the citizens. Citizens want more and more information from the statistical offices. Therefore, it is important that they participate in the surveys and give the correct answers. If people have the feeling that protection has not been ensured adequately, the risk is an increased non-response or incorrect answers in the case of obligatory surveys. Therefore, there is a need to explain continuously to the staff as well as to the general public what is data protection and confidentiality, what kind of data are protected and how. Several countries paid special attention to this issue in relation to the 2000/2001 Population Census. In some countries (e.g. Estonia, Latvia, Sweden, Ukraine) citizens have the right to check the data in the statistical office directly relating to them. Exercising this right requires careful organisation but experience shows that it is feasible. This can be a good way to increase public trust in the statistical office.

III. FUTURE WORK

49. The 2003 Work Session will be organised as follows:

DAY ONE

Plenary session

A. Assessment of statistical disclosure limitation (SDL) methods and software: quantification of information loss, level of safety, measures of disclosure risk, record linkage and statistical matching, user utility of released data, impact of SDL techniques on successive analysis, software comparisons and benchmarking.

Parallel workshops

1. Round table discussion on different approaches to measurement of disclosure risk
2. Software tools for disclosure control of microdata: a "hands on" workshop, using real data examples and delegates participation

DAY TWO

Plenary session

B. Impact of confidentiality on new data release techniques: SDL techniques for the release of small area, geographically referenced data, confidentiality of customised data extraction, electronic dissemination.

Parallel workshops

1. Round table discussion on confidentiality for small area statistics

2. Software tools for disclosure control of tabular data: a “hands on” workshop, using real data examples and delegates participation

DAY THREE

Plenary sessions

C. Discussion on methodological guidelines for studies on respondent perception of confidentiality:

A report on the current state of knowledge of respondent perceptions on the business side. A report on the current state of knowledge of respondent perceptions on the demographic side.

Parallel workshops

1. Business Perceptions: Roundtable discussion of key issues and common interests. Comparison of surveys and results
2. Demographic Perceptions: Round table discussion of key issues and common interests. Comparison of surveys and results

D. Reports on results of round tables

E. Future work