

**Joint ECE/Eurostat Work Session on
Statistical Data Confidentiality**

(Skopje, The former Yugoslav Republic of Macedonia,
14-16 March 2001)

Working Paper No. 29
English only

Topic III: Attitudes of respondents towards statistical confidentiality

**CONTROL OF STATISTICAL DISCLOSURE VERSUS NEEDS OF DATA USERS IN ISRAEL:
A DELICATE BALANCE**

Contributed paper

Submitted by the Central Bureau of Statistics, Israel¹

I. INTRODUCTION

1. National Statistical Offices (NSO) which release microdata files for outside use must balance data needs of users against confidentiality requirements. Expansion of NSO data-gathering activities, increased utilization of administrative data files for statistical purposes, and continual improvement in data-processing capabilities and technologies means that more data will be available, and using it will become easier. The growth in demand for data is both accompanied and fueled by an increase in data availability and accessibility. NSOs must consider not only the attitudes of respondents about the chance that information they provide in confidence may become publicly available. They are caught between their obligations to respondents, on the one hand, and their obligations to data users, on the other. Tension between user needs and confidentiality restrictions is unavoidable, because data users will always want greater detail than the NSO will be willing to provide. Such tension will only grow as a result of the inevitable increase in demand for detailed data that can be used for research, planning and evaluation. NSOs have developed principles and procedures to minimize the risk of disclosure in microdata files released to the public. It is worth examining how these principles and procedures are actually applied in practice. In this paper I will discuss some of the problems which arise as the Israel Central Bureau of Statistics (ICBS) attempts to implement the principles and procedures it has developed for the dissemination of microdata files, and how we deal with them.

2. Holvast (1999) provides a useful classification of methods employed to protect confidentiality, based on a survey of EU practices. He distinguishes among legal and ethical safeguards, organizational and administrative safeguards, and technical safeguards, noting that "it is striking that most safeguards appear to be technical and legal; organizational methods are rarely mentioned" (p. 197). One reason for this apparent neglect may be the difference between technical and legal safeguards, on the one hand, and administrative and organizational procedures, on the other. Laws for the protection of confidentiality state general rules, but these must be implemented by means of operational procedures. Technical safeguards specify how a particular data set is to be treated, but their content depends on substantive decisions regarding acceptable levels of risk. The organizational and administrative procedures established to implement the legal requirements mediate between them and the technical operations carried out on particular data sets.

3. A second reason that organizational methods for protecting confidentiality may receive less attention than legal or technical safeguards, lies in the kind of information required to evaluate them properly. It isn't very difficult to describe the formal structure established by an NSO to handle requests

¹ Prepared by Charles S. Kamen (charles@cbs.gov.il).

for the release of microdata files. But description of the formal structure is just the first step in an adequate evaluation. What we also need to know is how the procedures actually operate in practice. Information on the effectiveness of legal safeguards can be obtained from an analysis of court cases, while knowledge regarding the effectiveness of technical safeguards can be obtained by an evaluation of whether the procedures successfully identify cases of disclosure risk at the predetermined level of certainty. In order to evaluate the administrative and organizational mechanisms it isn't sufficient to count the number of requests submitted, approved and denied, and to characterize them. It is necessary to follow the process of deliberation in order to see how decisions are actually made and how the legal principles are interpreted in particular cases. Such an undertaking is time consuming, and requires access to the internal workings of the administrative bodies which make these decisions. Access may be difficult to obtain, for practical reasons – time, manpower, the possible disruption of orderly procedure. The NSO, or the relevant department, may also be reluctant to provide access to internal deliberations.

4. The inherent tension between demands for data and the desire to preserve confidentiality appears in various guises and reflects different competing interests. The internal deliberations in response to requests for microdata files create, through the resolution of issues raised by particular requests, the actual rules governing their release. There is variation among NSOs in the degree to which rules and procedures allow room for case-by-case reviews of requests for files, both with regard to content and with regard to who is entitled to receive them. The legal rules must be operationalized in administrative and organizational mechanisms in order for them to have an effect. Administrative and organizational procedures provide the framework and the setting within which conflicts between competing confidentiality interests are manifested and resolved.

II. ISRAEL CENTRAL BUREAU OF STATISTICS POLICY REGARDING RELEASE OF MICRODATA FILES

5. ICBS policy is to maximize outside access to microdata files, to a degree limited only by the need to preserve confidentiality. Fundamental changes have occurred over the past two decades in the Bureau's attitude to releasing microdata files. It was once thought that data files should be released only with great care, both out of concerns for confidentiality and because of worries that recipients of the files would analyze the data "incorrectly" or "inappropriately." It was felt that the Bureau would be held responsible for such errors, and would be put in the position of explaining findings or tabulations which it had no hand in creating. Restricting access to microdata files was one way of avoiding such dangers. There was some pressure to release microdata files, primarily from university researchers, but the limited availability of computing facilities and the relatively small number of persons with the skills required to analyze large data sets, or an interest in doing so, meant that demand for such files remained small.

6. The situation today is completely different. Computer capability, access and expertise are widespread in Israel, and the number of potential users of data files has greatly increased. The Bureau's paternalistic approach has also changed. It no longer sees itself responsible for the mistakes of others, but limits its obligation to providing accurate and sufficient documentation of data files, as well as answering questions from users.

7. The policy and procedures of the ICBS regarding statistical confidentiality are detailed in the proceedings of the first UNECE Work Session on Statistical Data Confidentiality (Burshtein, 1999). Burshtein described the Bureau's policy on confidentiality, and the mechanisms through which it is implemented. In this paper I expand his discussion to consider the issues which arise in trying to translate policy into practice. I refer only to microdata files on households and individuals, not on enterprises. Given the small size of the Israeli economy, it would be relatively simple to identify individual businesses in certain economic branches. An advisory committee established to recommend policies for microdata dissemination concluded that such data should not be released. In order to meet user demand for such data, however, access to it is granted in the framework of a Research Room (described below). I will summarize the main components of our approach under the three headings of legal safeguards, organizational safeguards and technical safeguards.

Legal safeguards

8. The Statistics Ordinance empowers the ICBS to collect information, to require that respondents provide information, and to release information to the public. It obligates the ICBS to maintain the confidentiality of information it collects under the Ordinance in order to prevent disclosure. Penalties may be imposed for violations of the respondent's obligation to provide information and the ICBS's obligation to prevent disclosure. Only when respondents have been specifically informed that the information requested is not being collected under the provisions of the Ordinance may the Bureau release individual records. This would occur only in the very few cases where we conduct a survey together with a government ministry that wanted information on individual records – for example, a survey of educational institutions which collected detailed information on each school that would also be made available to the Ministry of Education. In such a case, however, the respondent – here, the school administration - is not legally obligated to respond.

Organizational safeguards

9. Department heads are responsible to insure that cell frequencies in aggregate data tables exceed a predetermined minimum. Microdata files released to the public must be cleared by a Statistical Confidentiality Committee. Requests for approval are submitted to the committee by the head of the department responsible for the subject area covered by the file. Two levels of release are permitted: Public Use Files (PUF), available under license to anyone, and Microdata Under Contract (MUC), available only to researchers at approved institutions. The level of detail provided in Public Use Files makes extremely unlikely the identification of individual units. MUC files contain slightly more detail, but are released subject to contractual restrictions on permitted use. These restrictions reduce the likelihood of disclosure, and are meant to compensate for the greater detail. Institutions whose researchers are entitled to MUC files include degree-granting universities and colleges, and non-university research departments which have been certified by the Government Statistician as eligible to receive them. At present, these include the Brookdale Institute (a social welfare research organization) and the research departments of the National Insurance Institute (the Israeli social security program) and the Bank of Israel. In no case are names, addresses, population registry numbers or other direct identifiers included in the released file. Researchers requiring a greater level of detail than available in an MUC file may apply to the ICBS's Chief Scientist for approval to work on such files (without identifiers) in a Research Room on ICBS premises. Such approval is based on an evaluation of the scientific or public worth of the project, not on criteria relating to confidentiality. Output of these projects, on the other hand, is examined to insure that disclosure limitations are not breached.

10. The Bureau considered, and rejected, a proposal to condition the amount of detail that would be provided in the microdata file according to the "trustworthiness" of the recipient – the more s/he could be trusted, the more detail we would provide. It was clear that there was no way to establish publicly defensible standards of trustworthiness, and the proposal failed. In a sense, the distinction between PUF and MUC files represents a variation on the earlier idea, but one that is applied to categories of applicants based on an institutional affiliation, rather than on an individual basis. We assume that the researchers who are entitled to receive MUC files have no desire to identify individual records, and can therefore be "trusted" with files containing more detail. Our willingness to release of microdata files under MUC is doubly grounded, first on the legal constraints provided by the contract, and second on our belief that applicants entitled to receive MUC files are unlikely, by virtue of their category membership, to misuse them.

Technical safeguards

11. The ICBS does not apply automated data processing methods to microdata files in order to reduce disclosure risks. We neither introduce perturbations in the individual records nor employ a computer program to identify combinations of variables that may raise the chance of identifying individual records. We commissioned a preliminary study to examine the feasibility of developing software for evaluating disclosure risks, and undertook an analysis of i-ARGUS software to determine whether we might adopt it. The preliminary feasibility study did not yield usable results, and our evaluation of i-ARGUS concluded that its adoption would be difficult for us.

III. ACCESS TO ICBS MICRODATA FILES

12. Microdata files are released for use in two ways: in response to requests from individuals who wish to receive a file, and to the Social Science Data Archive (SSDA) at the Hebrew University in Jerusalem. "Individual" requests are of two main types: those coming from individual researchers who wish to obtain data files for their personal use, and those coming from individuals as representatives of public or private organizations or government agencies that seek the files for purposes connected with the activities of the organization. The SSDA, on the other hand, does not itself carry out analysis of data contained in the files it holds. It both serves as a repository for, and disseminates, files received from the ICBS and from other sources. The primary purpose of the SSDA is to serve faculty members of Israeli universities that subscribe to the archive through payment of an annual subscription fee. It may sometimes alter the structure of a file to make it more user-friendly. ICBS microdata files disseminated through the SSDA are subject to the same conditions as would apply to those files were they obtained directly from the Bureau. ICBS files deposited in the SSDA have already been cleared for release as PUF or MUC by the Bureau's Statistical Confidentiality Committee, and are re-disseminated by the SSDA according to the restrictions applying to each level. Recipients of MUC files from the SSDA must sign the same agreement required of those who obtain an MUC file directly from the Bureau.

13. Use of files in the ICBS Research Room is restricted to approved projects. Once approved, researchers have access to detailed microdata files from which only direct identifiers – name, specific address, Population Registry ID number – have been removed. Results of work carried out under in the Research Room auspices must be cleared for release in order to insure that confidentiality is not breached in the material removed from its premises. For example: a geographer working in the Research Room on an analysis of ethnic segregation in selected Tel Aviv neighborhoods was provided with a census file showing the ethnic distribution of residents by building. The initial maps he prepared using this data showed portions of the street grid, thereby creating the danger that characteristics of individual residents could be identified. After discussion with members of the Statistical Confidentiality Committee, the map was redesigned to eliminate disclosure risk while continuing to show patterns of segregation.

IV. HOW THE STATISTICAL CONFIDENTIALITY COMMITTEE WORKS

14. The committee has five members. Two represent subject areas, one comes from the census department, one from Information Services (data processing) and the fifth member is the Bureau's legal advisor. Requests to release files are submitted to the committee as they come in, and the committee schedules meetings according to the number of requests pending. Each request must include a detailed description of the file and record structure - fields and categories. The committee evaluates the disclosure risk by taking into account the level of geographic detail, the degree of content detail, and the balance between them.

15. Requested files are of two types: "standard" files, and "tailor-made" files. Standard files are designed to serve the general user who requires a basic data set. The files deposited in the Hebrew University SSDA may be considered standard files, since their record structure cannot be altered. Results of the 1995 Census of Population have also been released in the form of a standard PUF file, available to all. If the standard file does not meet the user's needs, the current policy of the confidentiality committee is to approve the release of a tailor-made file that does satisfy them, within disclosure limitations. The tailor-made file will be released as a PUF or as an MUC, according to the level of detail it contains. That same file will then be available to other users whose needs it meets. Creating tailor-made files is consistent with ICBS policy to maximize use of data we collect.

16. The committee makes considerable effort to meet the data needs of persons seeking microdata files. That is the justification for "tailoring" files to their specifications. Of course, any such tailoring involves a judgement as to the additional disclosure risk posed by the changes that are to be made. When such changes are judged to increase disclosure risks beyond an acceptable level, the committee will require additional changes that are intended to reduce them. Thus, for example, if a researcher needs greater geographical detail than is provided by the standard file, he or she will be required to accept less detail in subject fields that carry a disclosure risk.

17. In considering requests for microdata files the committee asks three questions. First, is the level of detail requested too great to permit release of the data file? If the answer is “yes,” the committee will either recommend that the applicant resubmit their request with a reduced level of detail, or suggest that the researcher apply to carry out the project in the framework of the Research Room. In that case, the decision whether to approve the project would be made, on substantive grounds rather than based on confidentiality restrictions, by the Bureau’s Chief Scientist.

18. Second, if the level of detail does not exceed that which would in principle allow release of the data file, the committee will determine whether the file should be released as PUF or MUC. Even if the applicant is entitled to receive an MUC file, the committee will determine whether the level of detail requested would in any case permit release as a PUF. If the answer is “yes,” the file will be approved as a PUF and be available to any future user. If the answer is “no,” it will be approved as an MUC and be available to others entitled to obtain MUC files.

19. Third, the committee considers the degree of overlap between the record structure – fields and categories – of the file under review and other files released, but usually does not use this as a criterion in deciding whether to approve a file for release. The issue of overlap arises precisely because we tailor data files to user needs. Imagine two data files. The first includes detailed geographic information, but the values of variables such as age, years of education, country of birth and others are collapsed into broad categories. The second file contains only broad geographical categories, but greater detail in demographic and social fields. Each file, on its own, may well be releasable as a PUF. Both files contain many fields in addition to the geographic and basic socio-demographic information. The level of detail in these fields would ordinarily not be altered in tailoring the file to the user’s needs, because they don’t add to the disclosure risk. But these common fields, if there are enough of them, can be used as links to merge the two original files into a new one containing both geographic detail and socio-demographic detail. Using common fields to link data files undermines disclosure controls.

20. Disclosure risk from file linkage is greater when both original files are released to the same researcher. Since all requests for data files come through the confidentiality committee, and the number of such requests is thus far less than ten per month, it should not be difficult to keep track of requests from the same researcher for different versions of the same microdata file. The committee would, in principle, evaluate later requests in light of what was released earlier. In practice, we have not yet been confronted with the need to make such determinations. Requests from different researchers for different versions of the same data file present additional problems. Though perhaps unlikely, the possibility exists of collusion among applicants to avoid disclosure restrictions. Assume, however, that the likelihood of collusion is negligible. Different researchers may still want alternative versions of the same data file. It would be very difficult to refuse approval of a later request only because someone else already received a previous version of the same file. The obvious solution to these problems is, of course, to cease releasing tailor-made microdata files, and restrict dissemination to standard products whose content can be controlled and which can’t be linked with others.

V. ICBS EXPERIENCE IN ESTABLISHING STANDARDS TO CONTROL DISCLOSURE RISK

21. In 1995, an internal committee of the ICBS, and an external committee set up by the Public Council for Statistics (PCS: a public advisory committee which oversees government statistical activities), each submitted reports recommending standards and procedures to control disclosure risk. The ICBS committee began with the premise that to eliminate all chance of disclosure would make it impossible to release files, which was unacceptable. It was sufficient to minimize the chance that an active search would be successful. The committee briefly reviewed the state of the art regarding automated procedures for evaluating disclosure risk, concluded that they were complicated, expensive and little used, and did not recommend adopting any of them. It rejected introducing perturbations in the data as one of the control techniques. It recommended reducing the number of variables and/or the level of detail in their categories, on the one hand, and expanding access mechanisms, on the other. Access to enterprise microdata files was rejected, except within the framework of the Research room. The PCS report accepted these recommendations.

22. These two committee reports referred to the situation as it was more than five years ago. The same advances in computer technology, greater sophistication in using data and the spread of computer skills among users that I referred to at the beginning of this paper as increasing the demand for data, also increase the risk of disclosure. At the same time, while the principles that the ICBS committee established have been adhered to, some of its specific recommendations have not been implemented. Thus, for example, the committee thought that released microdata files should not contain variables which the recipient didn't "need." It is difficult to identify such variables a priori. In the course of analysis new questions may arise that were not foreseen at its outset, requiring access to variables that may earlier have been thought of as "unnecessary." We remove only those variables that directly increase disclosure risk. While the committee recommended not releasing ungrouped data on age or other "year of ..." variables (such as marriage, immigration, settling in locality), often the analysis requires just such detail. Nor is the recommendation always followed to prevent identification of localities whose population is less than 100,000. Each such request is evaluated in the context of the user's needs.

23. Here are some examples of problems that arise as we try to apply the recommended standards:

- Traffic accidents in Israel number about 25,000 per year, of which some 450 involve fatalities. The ICBS receives a microdata file of accidents from the police. This file includes the Population Register ID Number of persons involved in the accidents. We use this ID number to pull demographic data from the Population Registry and add it to the file of persons involved in accidents. That file contains, of course, specific information about the location of the accident, including locality, street name and house number in towns and cities, or kilometer on non-urban roads. This "improved" file, minus identifiers, is made available to the Ministry of Transport research department, and to academic transport research institutions. The Ministry, in turn, makes it available to traffic safety engineers who conduct analyses under contract. The file we release does not contain the police accident file number, because that would allow someone to gain access to the police record for that accident (which is public information) and link it to the information we added from the Population Registry. But information from a news report of an accident, if it contains the name of one of the casualties, would be enough to enable someone to identify that person in the data file.
- The Bureau contracted to carry out fieldwork for a utilization survey of well-baby services by women who had recently given birth. The sample was drawn from March births in specified localities. Because health services are organized on a district basis, the researchers needed a district identifier in the file. Because the researchers were interested in different size localities, they wanted a "locality type" variable. But the combination of district identifier and locality type led to unique outcomes that enabled identification of individual localities. Since the sample was drawn only from March births in each locality the total population was small, and identifying the locality type would unacceptably raise the disclosure risk. We solved the problem by limiting geographic detail to prevent identification of individual localities.
- Annual microdata files of the Family Expenditure Survey are deposited in the Hebrew University SSDA, and released to the Research Departments of the National Insurance Institute (NII) and of the Bank of Israel as PUF files (available to all applicants). The 1997 file released to the SSDA included information only on household members aged 15 and older, even though information on all household members was collected in the survey. The file released that year to the SSDA and to the Bank of Israel didn't identify localities having fewer than 100,000 inhabitants, but in response to its request the file released to the NII identified localities having 50,000 inhabitants or more. When the 1998 microdata file was ready for release, the NII and the Bank of Israel requested that information be included on household members aged 0-14, so they could calculate household expenditure on education by type of educational institution. That request was granted, and data on persons aged 0-14 was also included in the file released to the SSDA. When the 1999 microdata file became available, the NII asked to obtain more detailed information about income and expenditure. That request was also granted. By this point committee members were becoming

increasingly uncomfortable with the unintentional "salami tactics" which may have had the effect of increasing disclosure risk. One simple solution would have been to tighten the conditions of release, and define the files as MUC, available only to research institutions. The problem with that solution was that the Bank of Israel's Research Department had not been certified as a research institution entitled to receive MUC files. Upgrading the classification would have meant refusing to give the Bank files previous versions of which it had already received, and which others would continue to receive. We saw no way to justify such discrimination, and recommended to the Government Statistician that he examine whether the Research Department of the Bank of Israel was entitled to be certified as a research institution. He determined that it was, thus enabling the files to be released as MUC to all three organizations.

VI. THE "POLITICS" OF MICRODATA FILE DISSEMINATION

24. The solution suggested in the previous section of this paper – limiting release of microdata files to standard products, instead of tailoring files to user needs – would create a number of problems for the Bureau. An examination of these problems must consider the context in which the ICBS operates, as well as how data is used, by whom it is used, and the possible consequences of restricting its use more than we do today.

25. The generation and publication of official statistics in Israel is divided between the Central Bureau of Statistics and other government agencies. The Bureau has a solid reputation for accuracy and neutrality. Much of the data it publishes is based on administrative files obtained from other government agencies. In some cases, the agencies that collect the data do not have the capacity needed to carry out the tabulations necessary to prepare the needed statistics. Even when they have the technical and professional capacity to do so, they may prefer that this work be carried out by the Bureau in order to avoid suspicion that the numbers published are inaccurate or biased.

26. If the Bureau has obtained an administrative data file from another organization, and the file is to be returned, it must be returned unaltered. The results of work carried out by the Bureau to improve the file, such as corrections based on logic checks, imputation or filling missing fields with information from other sources in the Bureau's possession, cannot be included in the file which is returned to the originating organization. This prohibition stems from the intent of the Statistics Ordinance. We interpret it as preventing us from returning "improved" records because such improvements may have been based on other information obtained under the Ordinance. Rather than attempting to determine whether such is true for a particular alteration, we refrain from releasing any individual records that we have altered.

27. The data files that the Bureau prepares are in increasing demand by the agencies from which they were obtained, by other organizations and by researchers. Many of the surveys conducted by the Bureau are carried out for a "client" which is a government agency. Some surveys are planned and funded jointly by the Bureau and other organizations. In such cases, the client or partner may expect to receive a copy of the data file in order to carry out analyses additional to those prepared by the Bureau. It has happened that the client or partner is surprised or dismayed to discover that they cannot have access to the detailed individual records, and that release to them of a microdata file must be approved by the Confidentiality Committee. In order to prevent such surprises in the future, contracts between the Bureau and the client or partner will include a clause specifying that release of microdata files must be approved by that committee.

28. Current Bureau policy, of supporting and encouraging government agencies to use data for planning and evaluation, imposes on us a moral obligation to provide them with the data they need. But our goal of educating public officials to use data may conflict with our obligation to prevent disclosure. It is relatively easy to explain why we can't release identified records, but more difficult to gain the agreement of frustrated users that our caution in releasing detailed information is justified. Too much frustration may even weaken public support for the Bureau in circles whose support is very important – the government agencies who, as clients, finance a portion of our work and make it possible for us to carry out activities which are not covered in the Bureau's internal budget. There seems to be no

immediate danger of this occurring but if, as we hope, data become more widely used, so may pressure to obtain greater detail.

29. The demands of the academic research community also create pressures on the Bureau to provide more detail in microdata files. Academic researchers represent a second important source of support for the Bureau's activities. Some of the pressures we feel from them are similar to those we experience from government users – a need for greater detail because of data needs in specific research projects. Others pressures take a form that stems from the particular social structure of the Israeli research community. Israel is a small country, and the number of academic researchers using the kind of data produced by the CBS is relatively limited. Many university researchers have had close ties with research departments in government agencies. Some have moved back and forth between academic and government positions. These career paths create personal and professional relationships among senior personnel in the academic and non-academic social science research sectors. This is also the case with senior CBS personnel, who themselves have held, or continue to hold, university appointments. These ties may create an atmosphere in which, in the absence of an automated procedure for evaluating disclosure risks, special efforts are made to meet data needs.

30. Introduction of a computer-based procedure for evaluating disclosure risks will not necessarily lead to standardization of criteria for releasing files. Such procedures operate on the basis of parameters entered by the users. The function of these parameters is to specify the criteria defining disclosure risks in order that they may be identified. Only if the application of the parameters is standardized will the results also be. An example of such a process seems to be the American Fact Finder established by the United States Bureau of the Census. That service provides tabulations based on microdata files, and evaluates requests according to the likelihood that the results will contain cells that provide information that must be kept confidential. Requests failing the test are not processed. But there is a difference between providing tabulations to anonymous users by means of a standard procedure implemented via computer, and providing tailor-made microdata files to identified, known applicants through personal interaction.

VII. STRIKING THE RIGHT BALANCE

31. At present, ICBS microdata dissemination policy balances level of detail against user "trustworthiness" in order to establish access criteria which preserve confidentiality, in the framework of organizational policy encouraging data dissemination. We may sometimes be too flexible in our willingness to tailor data to user needs, or in granting off-site access. Instituting changes in our policies is made more difficult precisely because users have become used to relative ease of access. It is harder to cut back on access than to refuse to expand it.

32. Imposing content restrictions doesn't seem to be the best solution, because they will only raise opposition among those who will be "penalized." It also raises problems of equity, for unless we recall files already released – a procedure whose implementation is difficult to imagine – later recipients will be given less useful material than was obtained by their predecessors. Greater strictness in approving requests would also invite opposition.

33. Preventing disclosure by restrictions on content is a policy difficult to defend. It undermines support for the Bureau, leads to dissatisfaction, and by increasing the salience of the inherent conflict between our obligations to those who provide data and to those who use it may cause both parties to lose confidence in our commitment to these obligations. A better approach may be to develop mechanisms that facilitate access to data under conditions that also strengthen our ability to protect it. Rather than responding by limiting what is available, we could improve access while also increasing our control. The user, after all, is primarily interested in obtaining the best data possible. If we can insure that he continues to do so, we will also be justified in asking him to accept somewhat greater inconvenience.

34. One way of meeting requests for increasing detail in released microdata files, while protecting confidentiality, would be through the establishment of satellite data centers, under ICBS control, where more detailed data files could be accessed. Such data centers could be located in ICBS branch offices and at universities. Dispersing them would provide wider access, and solve some of the restrictions on when

files are available only through the single “research room” on Bureau premises. Expanding research room facilities would allow us to put more emphasis on developing standard files, and reduce the amount of effort we devote to tailor-making them. The category of MUC file might eventually be eliminated. We would have standard public use files which would not be tailor made, and research room facilities for all needs which could not be met by the standard files.

35. Assuming we could make the new arrangements sufficiently attractive to be acceptable to users, the plan could be expensive for the Bureau. While “host” institutions could be asked to provide appropriately secured facilities and computers, their operation would have to be overseen by staff responsible to the Bureau. Perhaps the host institution could also cover staffing costs. The Bureau would still have to find funds to cover the expense of research facilities established in its regional offices. Another expense involves the preparation of the files for use in the research room. Easier access to microdata files through the research room might increase the number of requests for them and raise total costs. It is unclear whether the savings resulting from eliminating tailor-made files would balance the cost of preparing files for the research room.

VIII. CONCLUDING REMARKS

36. Rules, procedures and structures established by NSOs to minimize disclosure risks in the release of microdata files represent one aspect of the mechanism intended to preserve confidentiality of individual data. Equally important is the way in which these rules, procedures and structures operate in practice, and respond to changing user needs. Unlike Odysseus, who was only once compelled to navigate the passage between Scylla and Charybdis, NSOs endlessly traverse the straits flanked by respondent concerns, on the one hand, and user needs, on the other. We can expect public concern to grow regarding the security of personal information collected by statistical agencies, while demand by government, business and academic researchers increases for access to such data. NSOs have to satisfy the conflicting demands of both publics in order to fulfill their mandate, retain public confidence and continue to receive support. The way they do so is reflected in the decisions they make in specific cases. I have used the case of Israel to discuss some of the difficulties that arise in making such decisions.

References

Burshtein, Gideon, “Administration and policy of statistical data confidentiality in Israel,” in *Statistical Data Confidentiality*. Proceedings of the first Joint Eurostat/UNECE Work Session on Statistical Data Confidentiality, Thessaloniki, March, 1999, Eurostat, Luxembourg, 1999, pp. 245-250

Holvast, Jan, “Statistical dissemination, confidentiality and disclosure,” *ibid.*, pp. 191-208.