CHAPTER IX. DISSEMINATION

A. National reporting

9.1. *Data dissemination*. Data dissemination is one of the key activities in which the national statistical offices are involved. It is a way of providing the policy makers, business community and other users with high quality statistical information and also, it is a way of motivating respondents to participate in statistical surveys. If national statistical offices have the legal power to collect and disseminate statistical information, they also have the obligation to protect confidentiality of respondents.

9.2. *Statistical confidentiality*. Most of the information about individual statistical units classified in section G of ISIC, Rev.4 which is either directly collected by statistical offices or obtained from other sources, is considered to be confidential. Statistical confidentiality is necessary in order to gain and keep the trust of both - those required to respond to statistical surveys and those using the statistical information.

9.3. The sixth United Nations Fundamental Principle of Official Statistics provides the basis for managing the statistical confidentiality.

> "Individual data collected by statistical agencies for statistical compilation, whether or not they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes."

9.4. Legal provisions governing statistical confidentiality at national level are set forth in the countries' Statistical Laws or other supplementary governmental regulations. National definitions of confidentiality and rules for microdata access may differ, but they should be consistent with this fundamental principle. This is especially relevant for countries where the distinction between statistical and non-statistical use of microdata does not have a long tradition, or is not laid down in any legislation.

9.5. Distributive trade data are usually disseminated by the national statistical offices in the form of various statistical tables. The statistical confidentiality is protected if tabulated data do not allow statistical units to be identified either directly or indirectly, thereby disclosing individual information. Direct identification is possible if data of only one statistical unit are reported in a cell, while indirect identification may take place if individual data can be derived from disseminated data (e.g. because there are too few units in a cell, or because of the dominance of one or two units in a cell). To determine whether a statistical unit is identifiable, account shall be taken of all means that might reasonably be used by a third party to identify it[1].

[1] Eurostat definition of confidentiality, Chapter V of the Statistical Law and Council Regulation No 1588/90

9.6. *Statistical Disclosure Control*. Statistical disclosure control techniques are defined as the set of methods to reduce the risk of disclosing information on individual units. While such methods manifest themselves at the dissemination stage, they are pertinent to all stages of the process of statistical production.

9.7. Statistical disclosure control techniques related to the dissemination step are usually based on restricting the amount of data or modifying the data release. Disclosure control methods attempt to find an optimal balance between the improvement in confidentiality protection and the reduction in data quality. Different types of data pose different types of confidentiality problems and inevitably require different solutions. On the basis of available international guidelines[2] in this area, countries are encouraged to develop their own statistical disclosure methods suiting best their specific circumstances. Examples of the most often used methods are presented in para. 10.8.

9.8. *Methods of protecting confidentiality*. As the first step in the statistical disclosure control of tabular data, the sensitive cells need to be identified. The sensitive cells are those that tend to reveal directly or indirectly information about individual statistical units (see para. 10.5). The most common practices to protect the disclosure of confidential data include:

(a) *Aggregation*. A confidential cell in a table is aggregated with another cell and then the information is disseminated for the aggregate and not for the two individual cells. This, for example, often results in grouping of distributive trade data which are confidential at the class level of ISIC with another class and present and disseminate them at the group level of ISIC.

(b) *Suppression*. Suppression means removing records from a database or a table that contains confidential data. This is a method that allows statisticians not to publish the values in sensitive cells while publishing the original values of the others (primary suppression). Suppressing only one cell in a table however, means that the calculation of totals for the higher levels to which that cell belongs cannot be performed. In this case, some other cells must also be suppressed to guarantee the protection of the values under the primary cells, leading to the secondary suppression.

(b) *Other methods*. Controlled rounding and perturbation are more sophisticated techniques for protecting confidentiality of data. Controlled rounding allows statisticians to modify the original value of each cell by rounding it up or down to a near multiple of a base number. Perturbation represents a linear programming variant of the controlled rounding technique.

9.9. In cases when countries prefer suppression as a method for protecting confidentiality of distributive trade data, it is recommended that any data deemed confidential be reported in full detail at the next higher level of classification that

[2] Principles and Guidelines for Managing Statistical Confidentiality and Microdata Access, Statistical Commission, 38th session (http://unstats.un.org/unsd/statcom/sc2007.htm)

adequately protect confidentiality, if data are presented by activities, or higher level of aggregation for any other characteristics. Apart from satisfying the confidentiality protection, this technique must result in a minimum loss of information.

9.10. *Confidentiality of complex enterprises.* Careful measures to respect the confidentiality should be taken in the case of large trade enterprises. In general, large units are more easily identifiable than small ones, have higher probability for being selected in trade surveys and most likely their data will dominate the totals in a number of cells, thus allowing identification of such units.

9.11. *Confidentiality rules for distributive trade data***.** Rules for protecting confidentiality of distributive trade data should be in line with the provisions of countries' national legislation and practice. As a minimum requirement the following two criteria should be taken into account when defining the confidentiality rules: (i) number of units in a tabulation cell; and (ii) dominance of a unit or units' contribution over the total value of a tabulation cell. Decision in respect to the exact definition of the confidentiality criteria, e.g. in terms of number of units per cell and percentage of dominance is left to the national statistical offices. In individual cases confidentiality rules may be relaxed by requesting the permission of the dominating respondent(s) to authorize the statistical office to disclose the data.

9.12. *Internationalization of confidentiality.* Data collected and disseminated by international organizations depend to a large degree on the quality and completeness of the data supplied by the countries. Therefore, the issue of confidentiality has not only a national dimension. It is also becoming an international issue, for the following reasons – (i) high interest in cross-country comparisons; (ii) internationalization of users of statistical data (including international organization); and (iii) increase of data dissemination over the Internet. As a result, there is a growing demand for countries data at very detailed level, even in some cases – demand for countries microdata.

9.13. *Data dissemination timetable.* In producing statistical information there is usually a trade-off between the timeliness with which the information is prepared and the accuracy and level of detail of the published data. A crucial moment, therefore, for the well established relations between national statistical offices as producers of distributive trade statistics and the user community is devising an appropriate compilation and release schedule. Being important for the measurement of timeliness, as one of the quality dimensions of distributive trade statistics (see para. 9.4 (v)) it is recommended that countries develop and announce in advance the precise dates at which distributive trade statistics will be released. The advance release calendar should be posted in the beginning of each year on the country's statistical office website.

9.14. The most important elements that should be taken into account in determining the compilation and release schedule of distributive trade statistics include:

(a) Timing of collection of initial data from major distributive trade surveys;

(b) To what extent data derived from the major data sources are subject to revisions;

(c) Timing of preparation of important national economic policy documents that need distributive trade data as inputs;

(d) Modes of data dissemination (press release, electronic or hard copy).

9.15. Timeliness of release of initial monthly, quarterly and annual distributive trade data varies greatly from country to country, mainly reflecting different perspectives on the timeliness-reliability-accuracy trade-off. In keeping with sound statistical practices, countries are encouraged to release their initial monthly data forty five days after the end of the month in question, quarterly data - three months after the end of the quarter, and their annual data – eighteen months after the end of the year. Monthly and quarterly data should refer to a discrete month or quarter. Most countries use a separate system for compilation of annual distributive trade statistics. In this case the data for the fourth quarter (respectively the twelfth month) need to be published in their own right, and not be derived as a difference between the annual totals and the sum of the first three quarters (or eleven months).

9.16. *Data revisions.* Revisions are an essential part of countries practices on compilation of distributive trade statistics. They occur as a consequence from the trade-off between the timeliness of published data and their reliability, accuracy and comprehensiveness. To solve these issues statistical offices compile provisional data that later are revised when new and more accurate information become available. Although, in general, repeated revisions may be perceived as reflecting negatively on the reliability of official distributive trade data, the attempt to avoid them by producing accurate but very untimely data will result in failing to satisfy the users' needs. It is important to emphasize that the revisions of distributive trade data are conducted for the benefit of users, namely, to provide them with data that are as timely and accurate as possible. The revisions affect both annual and short-term distributive trade statistics but they are more significant for the short-term data.

9.17. *Reasons for revisions of data.* In general, there are two reasons for revisions - (i) revisions due to "normal" statistical procedures (for instance, new information becomes available, change in the methodology, change in data source, change of base year, etc.); and (ii) revisions due to the correction of errors that may occur in source data or in processing. For normal statistical data revisions (also called current revisions) countries should developed revision policy. At any moment of time statistical offices may decide to carry out a special revision in addition to the normal statistical data revisions for reasons of re-assessing the data or investigating in depth some new economic structures. Such revisions are carried out at longer, irregular intervals of time. Often they may require changes in the time series to be taken back as far as the beginning of the series to keep the methodological consistency. It is recommended that these revisions are subject to prior notification from the countries statistical offices to users that covers the reasons and the information on the impact of the revisions on the data.

9.18. *Revision policy*. To deal with the issues surrounding revisions of distributive trade data countries are encouraged to develop a well-designed, carefully managed and coordinated with other areas of statistics revision policy. The development of a revision policy should not aim at impeding revisions but rather it should aim at providing users with the necessary information to cope with revisions in a more systematic manner. The absence of coordination and planning of revisions is considered a quality problem by users. Essential features of a well-established revision policy are its predetermine schedule, reasonable stability from year to year; openness; advance notice of reasons and effects; easy access of users to sufficiently long time series of revised data as well as adequate documentation of revisions included in the statistical publications and databases.

9.19. *Recommended practices for data revisions*. A sound revision policy is recognized as an important aspect of good governance in statistics as it will not only help the national users of the data but it also will promote international consistency. To assist countries which have not yet set out such policy the following good practices are recommended[3]:

(a) Consultations with users elicit views about revisions practices;

(b) A clear, short summary statement of when to expect revisions and why is readily accessible to users;

(c) The current revision cycle is relatively stable from year to year;

(d) Major conceptual and methodological revisions are usually introduced every four to six years, balancing need for change and users' concern;

(e) Revisions are carried back several years to give consistent time series;

(f) Documentation on revisions is readily available to users;

(g) Users are reminded of the size of the likely revisions based on past history;

(h) When a mistake in reporting or processing is made, the revision is made in a transparent and timely manner.

9.20. *Dissemination formats*. A key to the usefulness of distributive trade statistics is the availability of data and hence its extensive dissemination. Data can be disseminated both electronically and in paper publications. It is recommended that countries choose the dissemination format that suits their users' needs best. For example, press releases of distributive trade statistics have to be disseminated in ways that facilitate re-dissemination by mass media; more comprehensive or detailed statistics have to be disseminated in electronic and/or paper formats. If resources permit, current statistics and

[3] For details see *Data and Metadata reporting and presentation Handbook, OECD, 2007,* Chapter 7

longer time series can be organized and accessed (free of charge or for a fee) through the electronic databases maintained by the statistical office. In addition to statistics routinely disseminated, statistical offices can make available to users distributive trade data upon request. For some specific purposes customized tabulations of data (non-standard activity classification, specific types of units etc.) can be provided. It is advisable that countries make well known to users the availability of additional statistics and the procedures for obtaining them.

9.21. *Dissemination of metadata*. Provision of an adequate metadata and quality assessment of distributive trade statistics is as important to users as provision of data. Countries are encouraged to follow the recommendations provided in chapter *IX. Data quality and metadata on distributive trade statistics* and develop and disseminate metadata comprising the following components: (i) data coverage, periodicity and timeliness; (ii) access by the public; (iii) integrity of disseminated data; (iv) data quality; (v) summary methodology; and (vi) dissemination formats It is recommended that countries indicate in the metadata all deviations from internationally accepted statistical standards and guidelines. Distributive trade statistics metadata should be made readily accessible through statistical offices websites and/or publications. Countries may consider development of different levels of metadata detail so to meet the requirements and needs of specialized users[4].

B. International reporting

9.22. Countries are encouraged to make available on their websites or to disseminate distributive trade data internationally as soon as they become available to national users.

9.23. The following tables provide the minimum list of data items on distributive trade statistics recommended for international dissemination, their level of details and periodicity.

Table 9.1. List of data items on distributive trade statistics for international dissemination with annual periodicity

Code	Data item	Level of details	Minimum requirements (in terms of ISIC, Rev.4)	Deadline
A	**Demography**			
1.10	Number of enterprises	Broken down by activity and size class	3-digit level for activity breakdown 1-digit level for size class breakdown	18 months
B	**Employment**			

[4] For more details on data and metadata reporting see *"Data and Metadata reporting and presentation Handbook", OECD, 2007*

2.1	Total number of persons employed	Broken down by activity and size class	3-digit level for activity breakdown 1-digit level for size class breakdown	18 months
2.1.3	Total number of employees	Broken down by activity, and size class	3-digit level for activity breakdown 1-digit level for size class breakdown	18 months
C	**Compensation of employees**			
3.1	Wages and salaries in cash and in kind of employees	Broken down by activity	3-digit level	18 months
H	**Output**			
8.1	Gross output at basic prices	Broken down by activity	3-digit level	18 months
8.1.1	Gross margin	Broken down by activity	3-digit level	18 months
J	**Value added**			
10.1	Total value added at basic prices	Broken down by activity	3-digit level	18 months
K	**Gross Fixed Capital Formation**		1-digit level	18 months

Table 9.2. List of data items on distributive trade statistics for international dissemination with quarterly periodicity

Code	Data item	Level of details	Minimum requirements (in terms of ISIC, Rev.4))	Deadline
B	**Employment**			
2.1	Total number of persons employed	Broken down by activity	2-digit level	3 months
2.1.3	Total number of employees	Broken down by activity	2-digit level	3 months
C	**Compensation of employees**			
3.1	Wages and salaries in cash and in kind of employees	Broken down by activity	2-digit level	3 months
E	**Turnover, sales, shipments, receipts for services and other revenues (excluding property income)**			
5 (a)	Turnover, sales, shipments, receipts for services and other revenues	Broken down by activity	2-digit level	3 months

Table 9.3. List of data items on distributive trade statistics for international dissemination with monthly periodicity

Code	Data item	Level of details	Minimum requirements (in terms of ISIC, Rev.4)	Deadline
	Wholesale and retail trade turnover indices (value and volume)	Broken down by activity	2-digit level	45 days