

COMMISSION REGULATION (EC) No 72/2002
of 16 January 2002
implementing Council Regulation (EC) No 530/1999 as regards quality evaluation of structural
statistics on earnings
(Text with EEA relevance)

THE COMMISSION OF THE EUROPEAN COMMUNITIES,

Having regard to the Treaty establishing the European Community,

Having regard to Council Regulation (EC) No 530/1999 of 9 March 1999 concerning structural statistics on earnings and labour costs ⁽¹⁾, and in particular Article 11 thereof,

Whereas:

- (1) Implementing measures for Regulation (EC) No 530/1999 are necessary concerning the content and evaluation criteria of the report on quality to be forwarded to the European Commission (Eurostat) after each reference period.
- (2) The information delivered in that report should refer to the variables defined in Commission Regulation (EC) No 1916/2000 of 8 September 2000 on implementing Council Regulation (EC) No 530/1999 concerning structural statistics on earnings and on labour costs as regards the definition and transmission of information on the structure of earnings ⁽²⁾.
- (3) The feasibility and the relevance of some optional items provided for the first quality report should be reviewed by Eurostat and the national statistics institutes in the light of the information delivered by the Member States.
- (4) The measures provided for in this Regulation are in accordance with the opinion of the Statistical Programme Committee,

HAS ADOPTED THIS REGULATION:

Article 1

1. The content and evaluation criteria of the report on quality referred to in Article 10(2) of Regulation (EC) No

530/1999 are laid down in the Annex to the present Regulation, Parts A and B.

The variables specified are those defined in Annex II to Regulation (EC) No 1916/2000.

2. The information and the optional characteristics or breakdowns, laid down in the Annex, are requested, in so far as that is made possible by the derogations in Community legislation concerning structural statistics on labour costs and earnings, the labour force survey, the structural business statistics and national accounts.

Article 2

The first quality report shall be delivered for the reference year 2002 structure of earnings survey.

It shall be transmitted to Eurostat together with the data delivery and at the latest 24 months after the end of the reference period for which the data was collected.

Article 3

The feasibility and relevance of the optional items laid down in the Annex, Part B, shall be reviewed in the light of the information actually delivered by the Member States.

This review shall be carried out by Eurostat and the national statistics institutes.

Article 4

This Regulation shall enter into force on the 20th day following its publication in the *Official Journal of the European Communities*.

This Regulation shall be binding in its entirety and directly applicable in all Member States.

Done at Brussels, 16 January 2002.

For the Commission
Pedro SOLBES MIRA
Member of the Commission

⁽¹⁾ OJ L 63, 12.3.1999, p. 6.

⁽²⁾ OJ L 229, 9.9.2000, p. 3.

ANNEX

CONTENT AND EVALUATION CRITERIA OF THE REPORT ON QUALITY FOR STRUCTURAL STATISTICS
ON EARNINGS

PART A

Structure of earnings survey: grossed up results: tabular analyses

Frequency distributions and the associated arithmetic means and medians ⁽¹⁾ should be provided for at least:

(a) The number of full-time employees, and

(b) the number of part-time employees,

broken down for each of the following variables:

— band of hourly gross earnings and by sex,

— band of monthly gross earnings and by sex,

— band of annual gross earnings and by sex;

— band of annual holidays (< 10, 10-19, 20-24, 25-29, 30-34, ≥ 35 days) and by sex,

— band of monthly hours paid or worked (< 130, 130-139, 140-149, 150-159, 160-169, 170-179, ≥ 180 hours) and by sex,

— NACE Rev.1 section and by NUTS level 1,

— NACE Rev.1 section and by sex,

— occupation (ISCO-88 at the 1-digit level) and by sex,

— education (ISCED 0 to 6) and by sex,

— age band (15-24, 25-54, 55-64, ≥ 65 years old) and by sex,

— length of service (< 10 years, 10-19, 20-29, 30-39, ≥ 40 years) and by sex,

— size of enterprise in terms of the number of employees ⁽²⁾.

For example, for hourly gross earnings and by sex:

(a) number of full-time employees

Band of hourly earnings	Men + women	Men only	Women only
	Frequency (%)		
Under EUR 5	<i>These bands of hourly earnings are illustrative. Each country should use bands which evenly distribute the frequencies over the range of hourly, monthly and annual earnings. If necessary, bands other than those specified above should also be used for annual holidays and for monthly hours paid or worked</i>		
EUR 5 to under 10			
EUR 10 to under 15			
EUR 15 to under 20			
EUR 20 to under 30			
EUR 30 to under 40			
EUR 40 to under 50			
EUR 50 and over			
Overall frequency	100 %	100 %	100 %
Total number of employees			
Overall mean (euro)			
Median value (euro)			

PART B

1. **Relevance** (optional item)

A summary including users' description, a description of users' needs (by main groups of users), and an assessment of the satisfaction of users' needs.

⁽¹⁾ For each specified distribution, the following should be provided: the total number of employees, the relative frequencies (%) for each band, the overall mean and the median value. (Arithmetic means and medians are not relevant for the variables NACE Rev.1, NUTS 1, occupation or educational level.)

⁽²⁾ Six bands in terms of the total numbers of employees: 1-9, 10-49, 50-249, 250-499, 500-999, 1 000 and over. The first of these six bands is optional for the 2002 structural earnings.

2. Accuracy

2.1. Sampling errors

2.1.1. Probability sampling

2.1.1.1. Bias (optional item)

Estimation of the biases due to the estimation method, if any.

2.1.1.2. Variance

- Coefficients of variation ⁽¹⁾ concerning total gross earnings, giving separate data for at least monthly earnings (data for hourly and annual earnings are optional), for:

(a) full-time employees, and

(b) part-time employees,

broken down by:

- NACE Rev.1 section and by sex,
- occupation (ISCO-88 at the 1-digit level) and by sex,
- age band (15-24, 25-54, 55-64, ≥ 65 years old) and by sex,
- NACE Rev.1 section and by NUTS level 1 (optional item),
- education (ISCED 0 to 6) and by sex (optional item).

- Coefficients of variation, giving data for monthly hours, for:

full-time employees, broken down by:

- NACE Rev.1 section and by sex,
- occupation (ISCO-88 at the 1-digit level) and by sex,
- age band (15-24, 25-54, 55-64, ≥ 65 years old) and by sex,
- NACE Rev.1 section and by NUTS level 1 (optional item),
- education (ISCED 0 to 6) and by sex (optional item).

2.1.2. Non-probability sampling

If non-probability sampling is used, a description of the possible sources of a lack of precision should be provided.

2.2. Non-sampling errors

2.2.1. Coverage errors

- Description of the main misclassification, under and over coverage problems ⁽²⁾ encountered in collecting the data.
- Description of the methods used to process these errors.
- Rates of misclassification, under- and over-coverage (optional item).

Note: Where individual administrative data are used, a similar analysis has to be reported based on the administrative reference file.

2.2.2. Measurement errors

- Description of the methods used to assess measurement errors ⁽³⁾.
- Assessment of the bias and a description of the estimators used to correct the bias for a main variable, for instance, monthly earnings.

⁽¹⁾ The 'coefficient of variation' is the ratio of the square root of the variance of the estimator to the expected value. It is estimated by the ratio of the square root of the estimate of the sampling variance to the estimated value. Both numerator and denominator should be provided, together with the resulting coefficient of variation. The estimation of the sampling variance must take into account the sampling design.

⁽²⁾ 'Misclassification' refers to incorrect classification of units that belong to the target population. 'Under-coverage' refers to (new) units not included in the frame, either through real birth or demergers, and to wrongly classified units. The black economy is not concerned here. 'Over-coverage' relates either to wrongly classified units that are in fact out of scope (e.g. actual local unit activity not in C-K NACE Rev.1) or to units that do not exist in practice.

⁽³⁾ Measurement errors are errors that occur at the time of data collection. There are a number of sources of measurement errors, including the survey instrument (the form or questionnaire), the respondent, the information system, the mode of data collection and the interviewer.

2.2.3. Processing errors (optional items)

Processing errors are errors in post-data-collection processes such as data entry, coding, keying, editing, weighting and tabulating.

- The error rates of data entry or coding for the main variables, for instance:
 - total gross earnings for the reference year,
 - total gross earnings for the representative month,
 - number of hours paid or worked in the representative month.
- Methodological notes on the estimation of these rates ⁽¹⁾.
- Assessment of the bias and variance due to processing errors.

2.2.4. Non-response errors

- Unit response rate ⁽²⁾.
- Item response rates for the main variables (for instance, monthly and annual earnings and hours worked). The rate is the ratio of the number of item responses to the number of in-scope respondents.
- A description of the methods used for imputation and/or re-weighting for non-response.

Note: Where individual administrative data are used, non-availability of the administrative record or item of data replaces non-response.

- A description of the reasons for non-response and an assessment of non-response biases for one of the major questions in the questionnaire, for example, concerning monthly or annual earnings, or for hours worked.

2.2.5. Model assumption errors

Report ⁽³⁾ on the use of the following models:

- to ensure that a representative month is selected,
- to adjust the accounting or fiscal year to the calendar year,
- to ensure that NACE Rev.1 Sections C to K are fully covered for all enterprises (as a minimum, for enterprises with 10 or more employees),
- to combine data from administrative sources and surveys.

Note: Where individual administrative data are used, there should be comments on the correspondence between the administrative concepts and the theoretical statistical concepts.

3. Timeliness and punctuality

- Key data-collection dates: e.g. the deadline imposed on respondents in the Member State, when the questionnaires and recalls and follow-ups were sent out and when the field work took place.
- Key dates for the post-collection phase: e.g. starting and finishing dates for completeness, coding and plausibility checks, date of the quality check (congruency of results) and non-disclosure measures.
- Key publication dates: e.g. when the advance and detailed results were calculated and disseminated.

Note: Punctuality of data transmission to Eurostat will be evaluated according to the regulation specifying periodicity and delays for data transmission.

4. Accessibility and clarity

- A copy of the publication(s), or a reference to where it can be located.
- Information on what results, if any, are sent to reporting units included in the sample.
- Information on the dissemination scheme for the results (e.g. to whom the results are sent).
- A copy of, or references for, any methodological documents relating to the statistics provided.

⁽¹⁾ The measures of the error rates can be achieved through standard techniques of quality control, for instance by verifying the quality of a subsample of the processed questionnaires (to check the level of errors either during the keying phase or during the processing of edits by National Statistics Institutes (NSI) staff).

⁽²⁾ The rate is the ratio of the number of in-scope respondents to the number of questionnaires sent to the population selected.

⁽³⁾ Comments, for example, should cover: the selection procedure of these models (i.e. why a given model has been chosen in preference to alternative ones); if relevant, the associated estimation error of the corresponding estimates; elements on the verification of the assumptions underlying the model; the test of the predictive power of the model by using historical data; the comparison of the results generated by the model with other related sources of data; the use of screening and cross-validation studies; the tests of sensitivity of the model to parameters' estimation; and the validation of the data inputs to the model.

5. Comparability

5.1. Geographical comparability

A comparison between national concepts and European concepts has to be reported if there are differences especially on the definition of statistical units, the reference population, classifications and definitions of variables in the transferred results. The differences should be quantified.

Note: When classifications and units come from the register, the quality of this information should come from the report on the quality of the register.

5.2. Comparability over time

Details of changes in definitions, coverage or methods compared with the previous structure of earnings survey and an evaluation of the consequences of non-negligible changes.

6. Coherence

Note: This item has two objectives. The first aim is to inform the users about the conceptual differences that exists between several sources for variables that are very similar and which usually have the same name in statistical publications, and to provide them with information on how to move from one concept to the other. The second objective is to check that statistics, which are in principle coherent conceptually, give comparable results for the same year and reference population. With these aims, statistics on the structure of earnings should be compared with other statistics sent to Eurostat taking into account, for example, that the structure of earnings survey (SES) is based on local units belonging to enterprises with 10 or more employees.

6.1. Coherence with the structure employees in the labour force survey for the same reference period

The structure of the SES should be compared with the structure of the labour force survey (LFS) as both surveys have several variables in common. In particular, cross-analyses of the distribution of employees in the SES and LFS should be expressed in percentages and made separately for full-time and part-time employees. The cross-tabulations should use the following variables:

- sex, age and economic activity (NACE Rev.1, at the section level),
- sex, age and educational level (ISCED 0 to 6),
- sex, age and occupation (ISCO-88 at the 1-digit level).

The following bands should be used for age (15-24, 25-54, 55-64, 65 or more years old).

6.2. Coherence with absolute figures from the labour force survey data for the same reference period (optional item for the 2002 structural earnings survey)

The following table summarises what is common to both sources.

	Structure of earnings survey (SES)	Labour force survey (LFS)
Statistical unit	Local unit	Household
Activity coverage	NACE Rev.1 Sections C-K, M-O	NACE Rev.1 Sections C-K, M-O
Variables to be compared as between the two sources	Number of employees, separately for full- and part-time, and by sex Number of hours worked during the representative month (or in a standard working month)	Number of employees, separately for full- and part-time, and by sex Number of hours per week usually worked converted to hours per month
Required breakdown by occupation	ISCO-88 at the 1-digit level	ISCO-88 at the 1-digit level
Required activity breakdown	NACE Rev.1, at the section level	NACE Rev.1, at the section level
Required regional breakdown	NUTS level 1	NUTS level 1

The variable 'number of employees' (separately distinguishing full-time and part-time employees and by sex) and the variable 'number of hours worked' should each be broken down by occupation, by region and by NACE activity (it is recognised that NACE Sections M-O are optional for the 2002 SES). Cross-tabulations between occupation, region and economic activity are not required.

- 6.3. *Coherence with structural business statistics for the same year: regional data* (optional item for the 2002 structural earnings survey)

The structural business statistics regulation⁽¹⁾ covers all market activities in Sections C to K of NACE Rev.1, except Section J. The following table summarises what is common to both sources.

	Structure of earnings survey (SES)	Structural business survey (SBS) (annual regional statistics)
Statistical unit	Local unit	Local unit
Activity coverage	NACE Rev.1 Sections C-K	NACE Rev.1 Sections C-K, except J
Variables for comparison as between the two sources	Number of local units in the population Total gross annual earnings in the reference year	11 21 0: Number of local units 13 32 0: Wages and salaries
Required activity breakdown	NACE Rev.1, at the section level	NACE Rev.1, at the section level
Required regional breakdown	NUTS level 1	NUTS level 1

The variables 'number of local units' and 'total gross annual earnings/wages and salaries' should each be broken down by NACE activity and by region. Explanations of the main differences for these variables should take account of the differences in concepts, definitions, accuracy and coverage for each source. Cross-tabulations should be attempted by NACE activity and by region.

- 6.4. *Coherence with structural business statistics for the same year: national data by size class of enterprise* (optional item for the 2002 structural earnings survey)

	Structure of earnings survey (SES)	Structural business survey (SBS) (annual enterprise statistics by size class)
Statistical unit	Local unit	Enterprise
Activity coverage	NACE Rev.1 Sections C-K	NACE Rev.1 Sections C-K
Variables for comparison between the two sources	Total number of enterprises in the population Total number of employees in the population Total gross annual earnings in the reference year	11 11 0: Number of enterprises 16 13 0: Number of employees 13 32 0: Wages and salaries
Required activity breakdown	NACE Rev.1, at the section level	NACE Rev.1, at the section level
Required breakdown by size of the enterprise	Number of employees: 1-9 (*), 10-49, 50-249, 250-499, 500-999 and 1 000 +	Number of persons employed 1-9, 10-49, 50-249, 250-499, 500-999 and 1 000 +

(*) Optional for the 2002 structure of earnings survey (SES).

The variables 'number of enterprises', 'number of employees' and 'total gross annual earnings/wages and salaries' should each be simultaneously broken down by NACE activity and by size of enterprise. Explanations of the main differences for these variables should take account of the differences in concepts, definitions, accuracy and coverage for each source. For example, for the SBS, the number of enterprises (11 11 0), the number of employees (variable 16 13 0), and wages and salaries (variable 13 32 0) are not available for all NACE Sections C to K.

⁽¹⁾ OJ L 14, 17.1.1997, p. 1.

- 6.5. *Coherence with national accounts for the same year: national data* (optional item for the 2002 structural earnings survey)

The following table summarises what is common to both sources.

	Structure of earnings survey (SES)	National accounts (Table 3: Tables by industry — annual exercise (*))
Activity coverage	NACE Rev.1 Sections C-K, M-O	NACE Rev.1 Sections C-K, M-O
Variables for comparison between the two sources	Total number of employees Total gross annual earnings in the reference year	Total number of employees D11 Wages and salaries
Required activity breakdown	NACE Rev.1, at the section level	NACE Rev.1, at the section level

(*) European System of Accounts, ESA 95, transmission programme of data, European Communities 1997 (also see ESA 95 questionnaire).

The variables 'number of employees' and 'total gross annual earnings/wages and salaries' should each be broken down by NACE activity. Explanations of the main differences in the variables between the two sources should take into account differences in concepts and coverage and, if available, the accuracy of each statistic. For the 2002 SES, NACE Rev.1 Sections M to O are optional.

7. **Completeness**

A list of variables and/or breakdowns required in the Regulation that are not available and the improvements envisaged in order to rectify these deficiencies.