

Italian Structure of Earnings Survey 2002

Microdata file for scientific purposes

Survey methodology

Summary

Introduction	3
General Information	3
Survey strategy	5
Quality Evaluation	8
Dissemination of the survey results	9
Glossary	10
References	11
Contacts	11

1. Introduction

According to Council Regulation No. 322/97 and Commission Regulation 831/2002, microdata files for SES 2002 may be released for scientific purposes. Microdata protection ought to take into account both levels: employers and employees.

The purpose of this document is to provide the user with a general description of the microdata file production process. The main features of the Italian Structure of Earnings Survey (SES) 2002 will be presented.

2. General Information¹

Introduction, legal framework and objectives of the survey

In the framework of the European System on Earnings and Labour Cost statistics, the Structure of Earnings Survey is a four-yearly survey whose objective is to provide accurate and harmonised data on earnings in the EU Member States for policy-making and research purposes. It is covered by several legal acts, the main one being the Council Regulation 530/99², followed by the implementation of Commission Regulations 1738/2005 (amending Commission Regulation 1916/2000) on definition and transmission of information and Commission Regulation 698/2006 on quality evaluation. More recently SES has been included into Commission Regulation 831/02 on microdata release for scientific purposes.

The objective of the SES is to monitor the structure and distribution of earnings. The survey provides information on earnings taking into account job-related factors, such as working hours, principal economic activity, size of the enterprise and location, and individual characteristics of employees such as gender, age, occupation, education, length of stay in service and others. In order to improve the performance of the entire survey process in terms of quality, timeliness and enterprise statistical burden reduction, in 1999 Istat coordinated a Commission of Experts to study the information value of the SES (2002); the output of the Commission of Experts consisted of a tailor-made questionnaire containing information on flexibility of the labour market and collective pay agreements.

The SES for the reference year 2002 (conducted in 2003-2004) covered NACE rev.1 Sections C to K for enterprises with at least ten employees.

In derogation to Council reg.530/99 for the **Italian 2002 wave, the observation unit was the enterprise instead of the local unit.**

The Structure of Earning Survey refers to year 2002 for annual variables and to October 2002 for monthly data in line with Commission Regulation 1916/2000.

According to the Regulation specifications, the following variables³ (Table 1) were observed in the Italian 2002 SES survey and subject to statistical disclosure control. The abbreviations used in this document are given in brackets.

¹ The paragraph has been written by S.Cardinaleschi

² Concerning structural statistics on earnings and labour cost

³ As already mentioned above a wider set of variables is collected in the Italian questionnaire

Table 1 – List of Commission Regulation 1916/2000 variables collected in the Italian SES 2002

A. Variables⁴ related to enterprise

- A.1.1 Geographical location of the observation unit⁵ [*Nuts*]
- A.1.2 Size of enterprise to which the local unit belongs to⁶. [*SizeEnt*]
- A.1.3 Principal economic activity of the observation unit⁷. [*Nace*]
- A.1.4 Form of economic and financial control⁸.
- A.1.5 Existence of collective pay agreements covering the majority of employees in the observation unit⁹.
- A.1.6 (optional) Total number of employees in the local unit¹⁰. [*Emp*]
- A.4.1. Enterprise sample weight¹¹.

B. Variables related to employee

- B.2.1 Gender. [*Gender*]
- B.2.2 Employee's age. [*Age*]
- B.2.3 Occupation¹². [*Occup*]
- B.2.4 Management position or supervisory position. [*ManPos*]
- B.2.5 Highest completed level of education and training¹³.
- B.2.6 Length of service in the enterprise.
- B.2.7 Is the employee full-time or part-time? [*FtPt*]
- B.2.7.1 Share of a full-timer's normal hours.
- B.2.8 Type of contract of employment.
- B.3.0 Average gross hourly earnings in the representative month.
- B.3.1 Total gross earnings for a representative month. [*MontlyEarnings*]
- B.3.1.1 Earnings related to overtime. [*Earnings related to overtime*]
- B.3.1.2 Special payment for shift work. [*Special payment for shift work*]
- B.3.2 Total gross annual earnings in the reference year. [*AnnualEarnings*]
- B.3.2.1 Number of weeks to which the gross annual earnings relate.
- B.3.2.2 Total annual bonuses. [*Total annual bonuses*]
- B.3.2.2.2 (optional) Annual bonuses based on productivity. [*Annual bonuses based on productivity*]
- B.3.4 Number of paid hours during the representative month. [*PaidHrsMonth*]
- B.3.4.1 Number of overtime hours paid in the reference month.
- B.3.5 Annual days of absence.
- B.3.5.1 Annual days of holiday leave excluding days of sick leave¹⁴.
- B.3.5.1.1 Holiday entitlement or number of holidays actually taken.
- B.4.2 Employee sample weight¹⁵.

⁴ The labels follow the Commission Regulation 1916/2000.

⁵ nomenclature of territorial statistical units NUTS level 1

⁶ the number of employees is classified in: 10-49, 50-249, 250-499, 500-999, 1000 and more employees

⁷ 2-digit level of NACE Rev. 1 for sections C-K

⁸ Public Control, Private Control, Shared Control.

⁹ national level or interconfederal agreement; industry agreement; agreement for individual industries in individual regions; enterprise agreement; single observation unit (local unit) agreement; other type of agreement; no collective agreement exists.

¹⁰ a simple head count of the total number of employees in the reference month is required, covering all employees

¹¹ calibrated with respect to *Nace*, *Nuts* and *SizeEnt* categories (see section 1.2 for more details)

¹² ISCO-88 (COM) classification at the two digit level

¹³ the information required concerns the level of general, professional or higher education, which the employee has completed. Completed implies the successful completion of the training and is normally (but not invariably) accompanied by an appropriate paper qualification. It is only necessary to code the highest level reached

¹⁴ total number of paid annual holidays (expressed in days) actually taken by the employee, excluding sick leave and public holidays. Depending on B3.5.1.1, it can be also be the number of paid entitled annual holidays

¹⁵ calibrated with respect to the total number of employees (see section 1.2 for more details)

3. Survey Strategy¹⁶

Sample design, response rate and handling of non-response, grossing-up factors

Sample design

The Italian Structure of Earnings Survey has a two stage sampling design: a sample of employees in a sample of enterprises. The target enterprise population has been divided into two strata: small-medium size enterprises (from 10 to 249 employees) and large size enterprises (with more than 250 employees).

As far as small-medium enterprises are concerned, the first stage units consisted of a stratified sample of enterprises from 10 to 249 employees, while the second stage units consisted of simple random sampling of employees. For large enterprises, a total survey was carried out at the first stage, followed by a simple random sampling of employees at second stage.

The first level sample was drawn from ASIA (Archivio Statistico delle Imprese – an ISTAT internal business register) with reference period 2001: the target population (the frame) contains about 152000 units (of which 2438 enterprises have more than 250) with about 7000000 employees.

A stratified sample has been studied for enterprises with 10 to 249 employees, which fulfils the 3% bound for the coefficient of variation of monthly gross earning per employee.

The strata were the combinations of 47 categories of NACE 2-digit, 4 employees size-classes, and 11 categories of NUTS1. A probabilistic proportional to size selection has been performed over a total of about 1.900 strata. Multivariate allocation procedures fulfilling the predefined accuracy and precision requirements for each planned domain provided the number of units to be selected in each strata.

Where possible, an increase of 30% of sample was considered in order to prevent the negative effect of total non-response. A total survey was carried out for the enterprises with more than 250 (2438 enterprises).

The final first level sample included 19032 enterprises.

The number of second level units (employees) to be selected in each enterprise was proportional to the enterprise size (from a minimum of 5 up to 100 employees). Table 2 shows the number of employees intended to be sampled by size-class of enterprises. Firms were trained on how to randomly select the required number of employees. The expected numbers of second level units was 160330.

Table 2 - SES 2002 Theoretical number of employees to be selected by size class of enterprises

	size class	number of employees to be selected
Sampling survey	from 10 to 49	5
	from 50 to 99	6
	from 100 to 249	7
Total survey	from 250 to 499	15
	from 500 to 999	20
	from 1000 to 1999	60
	from 2000 to 9999	65
	10000 and more	100

¹⁶The paragraph has been written by A.Palmieri

Response rate

The Italian respondent sample resulted in 8817 enterprises and 81975 employees. The effective response rate taking into account only the valid returning questionnaire is about 46.3%, representing 54.1% of the expected total number of employees. Table 3 summarises the main features of the sampling design.

Besides the already mentioned over-sampling of 30%, other actions to prevent from unit non responses were taken: the set up of special telephone lines (to help respondents in answering the questions), written reminder together with a new copy of the questionnaire sent out two months later, and special follow-up of major enterprises “critical cells”.

Non-response

As far as item non response was concerned, variables on monthly and annual earnings have been estimated according to a deterministic model based on auxiliary variables available from Social Security data base; the auxiliary variables were the average of monthly and annual earnings by position in employment.

Grossing-up factors

Responding units have been post stratified to take into account the non response pattern and the final weight resulted from the standard procedures implemented in Istat according to the calibration estimators. Number of business and employees in the frame provided by ASIA register reference year 2002 have been used as auxiliary variables.

The final weights guarantee the convergence of the results to the known population totals of the auxiliary variables of the following estimation domains:

- a. Subsection of economic activity (Nace rev.1) and employees classes (10-19, 20-49, 50-99, 100-249, 250-499, 500 and more)
- b. Subsection of economic activity (Nace rev.1) and 5 macroregions (NUTS1 year 2003)

Table 3 - SES 2002 – Summary table: Main features of sampling design.

Main sampling procedure	Sample survey						
Records used for sampling	Business register.						
Observed unit	Enterprise						
Characteristics of stratification: (Number of regions, branches of economic activity and size class of enterprise used)	4 employees classes 47 economic activities (NACE 2-digit) 11 NUTS1 regions						
Sampling fraction in each size class (in terms of the number of employees)	<i>100% for units with more than 249 employees</i> <table><tr> <th><u>Size class:</u></th><th><u>Sampling fraction</u></th></tr> <tr> <td>10-49</td><td>9.3</td></tr> <tr> <td>50-249</td><td>34.0</td></tr> </table>	<u>Size class:</u>	<u>Sampling fraction</u>	10-49	9.3	50-249	34.0
<u>Size class:</u>	<u>Sampling fraction</u>						
10-49	9.3						
50-249	34.0						
Overall sampling rate ¹⁷ (including those units exhaustively covered)	12.5%						
Response rate ¹⁸	46.3%						
Overall final sample size (number of observed units actually used)	8817						
Coverage rate ¹⁹	28.3%						

¹⁷ represents the number of units in the sample survey expressed as a percentage of the total number of units in scope (i.e. in the population)

¹⁸ represents the number of enterprises whose responses have actually been used, expressed as a percentage of the number of enterprises in the sample

¹⁹ represents the total number of employees actually covered by the survey expressed as a percentage of the number of employees in scope (i.e. in the population)

4. Quality Evaluation²⁰

Non sampling errors – Coverage, list, measurement and non response errors

The target population on which the survey was designed, was derived from ASIA, reference year: 2001. ASIA consists of a statistical Business Register resulting from the logical and physical integration of data included in several statistical and administrative sources (Tax Register, Register of Enterprises and Local Units provided by the Chambers of Commerce, Social Security Register, Work Accident Insurance Register, Register of the Electric Power Board). The data integration is performed by means of sound statistical methodologies. As a consequence, this frame could be the only possible misclassifications origin. Out of scope units (2% of the sample), associations to wrong addresses (5.9% of the sample) and to change of status (1.5% of the sample) sum up to a final list error of 9.4% of total responding units. No replacement was performed.

Measurement errors have been detected with selective editing procedures including scatter plot and histograms of monthly hours, gross monthly and annual earnings. Finally a deterministic imputation procedure was used. Total measurement errors affected only 0.2% of the records, resulting to be negligible.

More detailed information on first level unit (enterprises) response rates (by principal economic activity, size classes and geographical detail) can be found in the Quality report available on Eurostat Metadata Web site²¹.

Item response rate was calculated for the variables monthly hours, monthly earnings, and annual earnings. Some aggregated results are presented in Table 4.

About 1800 (2.36%) records had a missing value in at least one of those variables: 1.39% of the records had a non response on monthly hours (1141 occurrences), while for data on earnings the percentage was less then 0.9%.

Table 4 - Variables monthly hours, monthly earnings, and annual earnings – Item non response: absolute values and percentages

Number of item non response	Number of occurrences	Percentage
0	80122	97.74
1	1307	1.59
2	354	0.43
3	192	0.23
total number of records	81975	100

²⁰ The paragraph has been written by A.Palmieri

²¹ See section References

5. Disseminations of the Survey Results

Besides the microdata file for scientific research, several statistical information products are available.

The Italian SES2002 aggregated indicators are disseminated mainly through the ISTAT web-site: see

1. www.istat.it and
2. http://www.istat.it/salastampa/comunicati/non_calendario/20050311_00/retribuzionidef5.pdf

The aggregated indicators of the European survey may be found at the New Chronos web-site. The New Chronos data bank allows elaboration on the non-perturbed file for the following aggregations:

- by economic activity, gender, age
- by size of the enterprise, gender, occupation
- by economic activity, gender, level of educational attainment
- by economic activity, gender, length of service with the enterprise
- by economic activity, economic control, gender

Researchers might analyse the non-perturbed microdata file through the Italian or European Safe Data Centers. More information is available at the ISTAT and EUROSTAT web-sites, <http://www.istat.it/servizi/infodati/adele.html> (e-mail: adele@istat.it) and <http://ec.europa.eu/eurostat>, respectively.

6. Glossary

DEFINITION OF INDICATORS

Employees are all persons who have a direct employment contract with the enterprise or local unit and receive remuneration, irrespective of the type of work performed or the number of hours worked.

Gross earnings cover remuneration in cash paid directly by the employer, before deductions of tax and social security contributions.

Gross Monthly earnings are restricted to gross earnings which are paid in each pay period.

Gross Annual earnings also include allowances and bonuses which are not paid in each pay period, such as 13th month payments or holiday bonuses. Severance payments and payment in kind are not included

Gross hourly earnings are defined as gross monthly earnings divided by the number of hours paid in the same month.

Hours paid cover normal and overtime hours. Hours not worked but nevertheless paid are also counted as hours paid. Examples are annual holidays or sick leave.

CLASSIFICATION SYSTEM AND CONFORMITY WITH OFFICIAL STANDARDS

Economic activity is broken down at the division level of the General Industrial Classification of Economic Activities (NACE) Rev.1.1 for Sections C to K.

Five **size categories** are distinguished: 10 to 49 employees, 50 to 249 employees, 250 to 499 employees, 500 to 999 employees and units having at least 1 000 employees.

The **regional** breakdown is based on the Nomenclature of Territorial Units for Statistics (NUTS) which was valid when the 2002 survey data were communicated mid-2004. Any regional breakdown is shown only at NUTS-1 level.

7. References

1. Statistica in Breve- Structure of Earnings Survey 2002: S.Cardinaleschi “Dipendenti, ore lavorate e retribuzioni nelle imprese dell’industria e dei servizi – Anno 2002”,
2. http://www.istat.it/salastampa/comunicati/non_calendario/20050311_00/
3. A.Palmieri “Structure of Earnings Survey for the year 2002 – Quality report” – http://circa.europa.eu/Members/irc/dsis/wages/library?l=/4snationalsmetadatasands/structure_earnings/10_italy&vm=detailed&sb=Title
4. Chronos data base
5. http://epp.eurostat.ec.europa.eu/portal/page?_pageid=1996,45323734&_dad=portal&_schema=PORTAL&screen=welcomeref&close=/labour/lc&language=en&product=EU_MASTER_labour_market&root=EU_MASTER_labour_market&scrollto=0
6. Chronos SES Metadata
7. http://europa.eu.int/estatref/info/sdds/en/earn/earn_ses_sm.htm

Eurostat's Metadata Server (RAMON)

[/http://ec.europa.eu/eurostat/ramon/index.cfm?TargetUrl=DSP_PUB_WELC](http://ec.europa.eu/eurostat/ramon/index.cfm?TargetUrl=DSP_PUB_WELC)

8. Contacts

The Italian SES responsible may be contacted at cardinal@istat.it.