



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ



ΥΠΟΥΡΓΕΙΟ ΟΙΚΟΝΟΜΙΑΣ & ΟΙΚΟΝΟΜΙΚΩΝ



GENERAL SECRETARIAT OF
THE NATIONAL STATISTICAL SERVICE OF
GREECE

GENERAL DIRECTORATE OF STATISTICAL SURVEYS
DIVISION OF POPULATION AND LABOUR MARKET STATISTICS
WAGES AND SALARIES STATISTICS SECTION

“QUALITY EVALUATION ON LABOUR COSTS STATISTICS IN GREECE FOR THE YEAR 2004”

by

**Nikolaos Nembaris¹, Sofoklis Chrysanthopoulos²,
Ioannis Nikolaidis³ and Ilias Virras⁴**

Piraeus, May 2007

¹ Head of the Wages and Salaries Statistics Section, e-mail: nicneb@statistics.gr

² Officer of the Wages and Salaries Statistics Section, e-mail: schry@statistics.gr

³ Head of the Methodology, Analysis and Research Section, e-mail: giannikol@statistics.gr

⁴ Officer of the Methodology, Analysis and Research Section, e-mail: ivirras@statistics.gr

Introduction

This paper summarizes the collection of structural data on labour cost in Greece, and it presents the quality of the produced statistics, according to the quality evaluation criteria included in the Commission Regulation (EC) No 698/2006/5.5.2006 (Official Journal of the European Communities L 121, 6.5.2006, page 30).

1. Relevance

The relevance is the degree to which statistics meet current and potential users' needs. As the relevance is not an inherent characteristic of the statistical data, it can be measured only with the help of user satisfaction survey. The NSSG conducts a user satisfaction survey twice a year for the selection of information on the relevance of the produced statistics. This survey is limited to the customers visiting the library of the NSSG.

The results of the 'user satisfaction survey' of the year 2004 conducted by the NSSG, showed that the labour market statistics (employment and labour cost) met the users' needs as stated by the following scale of satisfaction:

- Completely by 78,3%
- Partially by 16,7%.
- Not at all by 5%

The fields of the labour market statistics that did not meet the users' needs were the following:

- The variable "wages and salaries in shares of company"
- The regional statistics (NUTS II and NUTS III)

As the questionnaire is the indispensable tool for the data collection, the survey questionnaire of the labour cost survey was designed in accordance with the European Union requirements as well as the national ones. The Community obligations represent the 95% of the questionnaire, and the rest questions were formed, after consultation with the *main national core users*, which are:

- The academic and research community,
- The banks and business,
- The government agencies,
- The national Parliament and
- The media.

The main purposes for which the users need the labour cost statistics are:

- Analysis of current developments for short-term decision making
- Analysis of trends for longer-term decision making
- Forecasting
- Research purposes

2. Accuracy

2.1 Sampling errors

2.1.1 Sample design

The labour costs survey covers the sections C-K and M-O of the NACE Rev.1 and the enterprises with average annual employment equal to or greater than 10 employees. The single stratified random sampling method was applied, employing the enterprise as a surveyed unit and obtaining statistical information from *each separate local unit* of the enterprises included in the sample.

The sampling frame used for the sample design was based on the Business Register (BR) of the NSSG. This BR is based on the VAT Register of the Ministry of Economy and Finance and it is updated through the statistical surveys of the NSSG and the register of the Social Insurance Foundation.

The statistical data for the public services of the sections M (Ministry of education and public schools) and N (Ministry of Health, public hospitals and public health centres) were collected from the Ministry of National Education and Religion Affairs, and the Ministry of Health and Social Solidarity, respectively.

2.1.2 Stratification

The enterprises with 10 or more employees included in the survey were stratified as following:

- a. By geographical region – NUTS I,
- b. By Division (two – digit NACE Rev.1 code) within each geographical region (Geography x Economic activity = Major stratum), and
- c. By size class of the enterprise. In each of the major strata, the enterprises were stratified into H=7 size classes, according to their size, determined by their average annual number of employees in the business register, as follows.

Class 1	10-19	Employees
Class 2	20-49	"
Class 3	50-99	"
Class 4	100-249	"
Class 5	250-499	"
Class 6	500-999	"
Class 7	1000 or more	"

The enterprises that belong to the 7th size class were surveyed exhaustively.

2.1.3 Sample size

The sample size is 4.313 enterprises (sampling fraction=20,1%) and the response rate =61,2%. The sample size of the enterprises was defined, so that the relevant standard error (co-efficient of variation CV) of the variables “annual labour costs” and “hourly labour costs” at 2-digit code level of economic activity at the whole country does not exceed 5%. The sampling units (enterprises) were distributed to size strata applying the method of optimal (Neyman) allocation.

The population (N) and the sample size (n), broken down by section and by size class of enterprises, are presented in the following table.

Table 1. The population (N) and sample size (n) by section and by size class

NACE Rev.1	Total		Size classes									
			10 –49		50-249		250 - 499		500 - 999		1000+	
	N	n	N	N	N	n	N	n	N	n	N	n
Total	21506	4313	17865	976	3080	2847	327	273	145	128	89	89
C	148	47	124	24	19	17	4	5	0	0	1	1
D	5955	1465	4.754	401	981	880	142	117	52	41	26	26
E	75	41	46	24	23	13	3	1	1	1	2	2
F	1112	221	861	22	222	178	15	9	12	10	2	2
G	5897	836	5.163	112	644	641	51	46	21	19	18	18
H	3302	390	2.947	37	322	318	21	19	9	13	3	3
I	1325	312	1.096	93	189	184	18	14	9	8	13	13
J	134	73	65	18	43	32	9	5	3	4	14	14
K	1689	441	1.323	103	314	293	35	29	13	12	4	4
M	191	58	152	25	31	28	6	3	2	2	0	0
N	367	144	211	19	124	99	16	12	12	10	4	4
O	1310	285	1.123	98	168	164	7	13	10	8	2	2

2.1.4 Selection of the sampling units (enterprises)

In each of the final strata (let h), a sample of n_h enterprises was selected. The enterprises to be surveyed were selected from the total of the N_h enterprises with equal probabilities and by applying systematic sampling. The sampling units (enterprises) were selected from the sample frame based on data from the Business Register of the NSSG.

2.1.5 Survey characteristics estimation

a. Symbols

Defining with index i the selection order of an enterprise from the sampling frame in the stratum h and symbolizing with the y one of the survey characteristics, we can define the following:

y_{hi} : the value of the survey characteristic y of the enterprise of order i in the stratum h

Y_h : the sum of the values of the characteristic y for all enterprises falling into the survey and belonging to the stratum h

Y : the sum of the values of the characteristic y of all enterprises under survey belonging to one economic activity with two digit code . That is:

$$Y = \sum_h Y_h \quad (2.1)$$

b. Estimation process

The estimations of the magnitudes Y_h and Y come from the following relations:

$$\widehat{Y}_h = \frac{N_h}{n_h} \cdot \sum_{i=1}^{n_h} y_{hi} \quad (2.2)$$

$$\hat{Y} = \sum_h \hat{Y}_h \quad (2.3)$$

Generally, in order the estimations of the survey characteristics to be produced at any level, we add up the estimations of the (final) strata, which form the level under survey. The estimates of totals are produced using the Horvitz-Thompson estimator of the relation (2.2), which is *unbiased*.

There are quantities being produced through the ratio of two variables (as the annual labour costs per employee or the hourly labour costs). The estimations are produced using the ratio estimator, which is usually *slightly biased*.

We assume that the population parameter to be estimated is the ratio:

$$R = \frac{\sum_{i=1}^N y_i}{\sum_{i=1}^N x_i} = \frac{Y}{X} = \frac{\bar{Y}}{\bar{X}}$$

where y_i and x_i are the values for the each unit of order i in the population of size N (e.g. the variable y is the total labour cost and the variable x is the number of hours actually worked). If the stratified random sampling scheme is applied, then the combined estimation of R is:

$$\hat{R} = \frac{\sum_h \frac{N_h}{n_h} \sum_{i=1}^{n_h} y_{hi}}{\sum_h \frac{N_h}{n_h} \sum_{i=1}^{n_h} x_{hi}} = \frac{\hat{Y}}{\hat{X}} \quad (2.4)$$

The ratio estimator \hat{R} is biased. In general, the ratio estimation has a bias of order $1/n$. Since the standard error (s.e.) of the estimation \hat{R} is of order $1/\sqrt{n}$, the quantity $\text{Bias}/\text{s.e.}$ is also of order $1/\sqrt{n}$ and it becomes negligible, as the sample size n becomes large. In practice, this technical bias is usually unimportant in samples of moderate and large size.

As the technical bias of \hat{R} occurs because the denominators x of $R = y/x$ are random variables, one can use the $CV(\hat{X}) < 20\%$ (CV: Coefficient of variation of \hat{X}), as an indicator examining if the effect of bias on the accuracy of \hat{R} to be neglected. Thus, the $CV(\hat{X})$ serves as a critical control on the validity of combined ratio estimations and it is a useful and safe-check on the bias of ratio statistics.

c. Variance estimation

The estimations of the variances of \hat{Y}_h and \hat{Y} come from the following relations:

$$Var(\widehat{Y}_h) = \frac{N_h(N_h - n_h)}{n_h(n_h - 1)} \cdot \left[\sum_{i=1}^{n_h} y_{hi}^2 - \frac{\left(\sum_{i=1}^{n_h} y_{hi} \right)^2}{n_h} \right] \quad (2.5)$$

$$Var(\widehat{Y}) = \sum_h V(\widehat{Y}_h) \quad (2.6)$$

The coefficient of variation (%) of the \widehat{Y} is given by the following relation:

$$CV(\widehat{Y}) = \frac{\sqrt{Var(\widehat{Y})}}{\widehat{Y}} \cdot 100 \quad (2.7)$$

The estimation of the variance of \widehat{R} is calculated from the following relation:

$$Var(\widehat{R}) = \frac{1}{\widehat{X}^2} \cdot \sum_h \frac{N_h \cdot (N_h - n_h)}{n_h} \cdot [S_{yh}^2 + R^2 \cdot S_{xh}^2 - 2 \cdot R \cdot Cov(y_h, x_h)] \quad (2.8)$$

$$\text{where: } S_{yh}^2 = \frac{1}{n_h - 1} \cdot \left[\sum_{i=1}^{n_h} y_{hi}^2 - \frac{\left(\sum_{i=1}^{n_h} y_{hi} \right)^2}{n_h} \right], \quad S_{xh}^2 = \frac{1}{n_h - 1} \cdot \left[\sum_{i=1}^{n_h} x_{hi}^2 - \frac{\left(\sum_{i=1}^{n_h} x_{hi} \right)^2}{n_h} \right],$$

$$Cov(y_h, x_h) = \frac{1}{n_h - 1} \cdot \left[\sum_{i=1}^{n_h} y_{hi} \cdot x_{hi} - \frac{\left(\sum_{i=1}^{n_h} y_{hi} \right) \cdot \left(\sum_{i=1}^{n_h} x_{hi} \right)}{n_h} \right]$$

The coefficient of variation of \widehat{R} is calculated from the following relation:

$$CV(\widehat{R}) = \frac{\sqrt{Var(\widehat{R})}}{\widehat{R}} = \frac{1}{\widehat{Y}} \cdot \sqrt{\sum_h \frac{N_h \cdot (N_h - n_h)}{n_h} \cdot [S_{yh}^2 + R^2 \cdot S_{xh}^2 - 2 \cdot R \cdot Cov(y_h, x_h)]} \quad (2.9)$$

The coefficient of variation of the variables “annual labour costs” and “hourly labour costs” are shown in the *Annex I*, according to the structure of tables A (national data), B (regional data) and C (national data by size class of enterprise).

In the section with code E, the coefficient of variations of the variables “annual labour costs” and “hourly labour costs” are equal to 0,1%, because 90% of the total statistical information was collected from two large enterprises (average annual employment higher than 1000 persons) belonging to the census (take –all) stratum.

In the sections M and N, the coefficient of variations of the variables “annual labour costs” and “hourly labour costs” are less than 1%, because the statistical information for the public services was collected from administrative sources (Ministry of National Education and Religion Affairs, Ministry of Health and Social Solidarity).

In the sections with codes D, I and J, the coefficient of variations of the variables “annual labour costs” and “hourly labour costs” are less than 1,6%.

In the sections with codes C, G, K and O, the coefficient of variations of the variables “annual labour costs” and “hourly labour costs” are ranged between 3% and 5%.

In the sections with codes F and H, the coefficient of variation *only* of the variable “annual labour costs” is higher than 5%, because in these sections strong seasonality is appeared. As a result, two different types of enterprises belong to the same size classes, as follows:

- Enterprises operating all the year
- Enterprises operating only a time of period less than one year (approximately, half a year)

The enterprises of the first type have annual labour costs higher than the enterprises belong the second type, and as a result, in the same size stratum the annual labour costs of the enterprises are not homogeneous due to the different types of enterprises. Thus, although in the same size stratum, internally homogenous enterprises exist, according to their number of employees, however internally heterogeneous enterprises are appeared, according to their values of the “total annual labour costs”. This increases the variance of the total annual labour costs of the enterprises, reducing the gain in the precision from the stratification that was introduced initially in the sample selection.

Concerning, the problem of the internal heterogeneity in the size strata was not appeared in the variable “hourly labour costs”, because the variables “annual labour costs” and “hours actually worked” are strong correlated. As a result, there was not any high variability in the hourly labour costs due to the different types of enterprises.

2.2 Non-sampling errors

2.2.1 Coverage Errors

There were problems of over-coverage, under-coverage and miss-classification. The *over-coverage* problems mainly have to do with enterprises that were included in the business register, they were selected in the sample, but they were not actually existed at the time of the survey (closed enterprises). These enterprises actually reduced the initial sample size of primary units, n_h . The decrease of the number of sampling units from n_h to m_h in each stratum inflates the variance of the estimated statistics. In this case the estimator is unbiased under the condition that the death rate of enterprises is equal to their birth rate.

The *under-coverage* refers to units missing from the sampling frame. The probability of selection of each missing unit of order i is equal to zero ($P_i = 0$) and thus, the

extrapolation factor w_i of the missing unit cannot be defined ($1/P_i = 1/0$). As a result, the under-coverage problem underestimates the produced statistics. Corrections and weighting for *non-coverage* is difficult, because the under-coverage rates cannot be obtained from the sample itself, but only from external sources.

Due to *miss-classification* problems of the register, some sampling units changed design strata after data collection. These units were allocated to the new strata, *retaining* their initial probabilities of selection. This event changes the initial element variance, destroys the initial allocation of the enterprises of the sample and as a result inflates the variance of the estimations. Consequently, the co-efficient of variation of the produced statistics is *higher* than the co-efficient of variation based on the initial sample design.

2.2.2 Measurement and processing errors.

The data collection method used was face-to-face interview completing paper questionnaires. The collection method applied ensured the high quality of the information gathered, since the interviewers assisted the respondents, and carefully checked the filled in questionnaires, before leaving the enterprise.

The interviewers participated in the survey were experienced permanent staff of the National Statistical Service of Greece (NSSG), as well as private collaborators. Before launching of the survey, the interviewers attended a one-day training seminar. The scope of the seminar was to enable the interviewers to: a) fully understand the definitions of the survey characteristics in order to avoid the respondent bias, (b) correctly fill in the questionnaire, and (c) efficiently check for errors by applying logical controls.

The structure and the size of the questionnaire were designed to be user-friendly for the interviewers and the questions were formulated in a clear and simple language, using appropriate vocabulary. Additionally, documents containing useful instructions were compiled, analyzing all the questions of the questionnaire. This activity targeted at collecting fully completed questionnaires, with no missing variables.

The support and supervision of the data collection and the data processing were decentralized in the regional offices of our Service. In regional offices were carried out coding, checking for the detection of measurement errors, logical controls and comparisons with other sources of statistical information.

After performing all final controls for discovering non-sampling errors, the database was ready for the extrapolation weighting process and the plausibility checks after tabulation. These checks included comparisons of data with relevant data of previous years and other surveys.

2.2.3 Non-response errors

The following table shows the unit response rates (%), total and broken down by section and size classes of enterprises.

Table 2. Unit response rates (%) by section and size classes

NACE Rev.1	Total	Size classes				
		10-49	50-249	250-499	500-999	1000+
	%	%	%	%	%	%
Total	61,0	74,2	52,6	80,1	84,4	100,0
C	68,1	66,7	64,7	80,0	100,0	100,0
D	67,3	76,6	59,8	80,3	100,0	100,0
E	100,0	100,0	100,0	100,0	100,0	100,0
F	47,1	100,0	33,1	100,0	80,0	100,0
G	66,5	100,0	58,0	84,8	94,7	100,0
H	61,3	100,0	54,4	84,2	30,8	100,0
I	63,1	72,0	51,6	92,9	100,0	100,0
J	54,8	61,1	53,1	40,0	50,0	100,0
K	41,5	47,6	31,4	82,8	100,0	100,0
M	53,4	32,0	42,9	100,0	100,0	100,0
N	52,8	63,2	46,5	58,3	60,0	100,0
O	49,8	55,1	45,1	38,5	75,0	100,0

In the census (take-all strata), in which all population units are included in the sample, the unit response rate is equal to 100%. In the sampling strata, in which only a part of population is included in the sample, the re-weighting method was applied for statistical adjustments of the produced statistics.

The re-weighting method amends suitably the extrapolation factors taking into account the response rates in all final strata. This method compensates for non-responses, and reduces the absolute bias in the estimation of \hat{Y} . If $\bar{Y}_{rh} = \bar{Y}_{mh}$ (where \bar{Y}_{rh} and \bar{Y}_{mh} are the means for respondents and non-respondents in stratum h for the variable y), as it occurs in expectation when the non-respondents are missing at random, then in stratum h the bias of non-response is equal to zero. Generally, the total bias due to the non-response is approximately equal to zero, if either the response rates or the respondent means do not vary between strata.

Any imputation method was not applied for the item non-response, as the item non-response was not appeared in the enterprises included in the sample.

2.2.4 Model assumption errors

Not any model was used in the structural survey.

3. Timeliness and punctuality

3.1 Punctuality

The multiple operations of the Labour Cost Survey were carried out in four phases, as detailed below:

Phase 1: Organization and preparation of the survey

The first phase was carried out from 1st January 2005 to 31st March 2005, and it comprised the organization activities and the preparatory work for the survey. More precise the following actions were carried out:

- Issuing of a special decision of the Ministry of Finance and Economy, which sets out the time schedule, the organization and the cost of the survey
- Sample design
- Design and printing of the questionnaire (paper and pencil),
- Printing of the manual with the instructions for the data collection
- Creation of software program for the data entry and automatic controls
- Programming for the creation of database files
- Selection and appointment of the interviewers for the conduct of the survey
- Training seminar of the interviewers for the effective data collection
- Delivery to the regions (prefectures) of the questionnaires and the questionnaires' instructions
- Transmission of information letters to the enterprises belonging to the sample

Phase 2: Data collection

The second phase was carried out from 1st June 2005 to 31st December 2005. In this phase the following operations were carried out:

- Distribution to the interviewers of the questionnaires and the lists with the sample units and other necessary documents
- Collection of the statistical data
- Monitoring and supervision of the operation from the beginning to the end by the supervisors (heads of the regional offices and the head of the competent department of the Central Office)
- Delivery by the interviewers of the questionnaires to their supervisors

Phase 3: Data processing

The third phase was carried out from 1st January 30th June 2006

- The following operations were carried out:
- Checking for the completeness of the questionnaire
- Logical and consistency controls of the data
- Coding
- Data entry and automatic data editing
- Creation of a database with the survey data
- Qualitative controls of the data in the database
- Calculation of the extrapolation factors
- Estimation of the survey characteristics
- Tabulation of the estimated statistics for qualitative analysis

Phase 4: Evaluation of the results-Publication and Dissemination

This phase was conducted in July 2006 and the following operations were carried out:

- Qualitative analysis and documentation of the results
- Production of national tables with the final results
- Transmission to Eurostat tabular data through the appropriate technical format for the transmission of the results

The *punctuality* of data transmission to Eurostat is evaluated according to delays stated in Council Regulation (EC) No 530/1999, in which the results are forwarded to Eurostat within a period of 18 months from the end of the reference year.

The data elaboration had to be completed, not later than 15th June 2006, in order the produced statistics to have been transmitted to Eurostat not later than the end of June. The qualitative controls of the data in the database were out of schedule and as a result the whole survey was extended by one month.

3.2 Timeliness

The length of time between the release of data and the reference period of data is equal to 19 months.

4. Accessibility and clarity

4.1 Accessibility

a. There is a publication in Greek containing:

- A short description of the methodology applied for the data collection
- Tables with the results of the survey

The publication with data of the year 2004 will be available before the end of October 2007.

b. Tables with the results of the survey are available in the website of the NSSG.

Before the end of September 2007, a short description of the methodology applied for the data collection in Greek and English will be available in the website of the NSSG.

In the case that the users need more detailed information, they can ask for it and special tables can be produced on request. Moreover, in some cases (ie for research purposes) anonymised individual data can also be provided to the users. The format of the anonymised data is so that the confidentiality to be protected and the respondents not to be revealed.

5. Comparability

5.1 Geographical comparability

The definition of the statistical units, the reference population, the classifications and definitions of the observed variables in the transferred results to Eurostat were determined according to the Council Regulation (EC) No 530/1999/9.3.1999 (Official Journal of the European Communities L 63, 12.3.1999, page 6). Thus, the produced statistics are comparable between the member-states of the European Union.

5.2 Comparability over time

The labour cost surveys with reference periods the calendar years 1969, 1973, 1974, 1976, 1977 and 1978 produced statistics only for section “Manufacturing” and the reference population was limited to establishments with 10 or more employees.

The surveys of the years 1981, 1988 and 1992 apart from manufacturing, were extended to the sections “Mining and quarrying” and “Electricity, gas and water supply”. Moreover, the surveys were designed and conducted, in order the produced statistics to be harmonized and comparable with the corresponding data of the rest member-states.

The survey of the year 1996 widened the economic activity coverage to include the sections “Construction”, “Wholesale and retail trade; repair of motor vehicles, motorcycles and personal household goods”, “Hotel and restaurants”, “Transport, storage and communication” and “Financial intermediation”. However, only enterprises with 10 or more employees participated in the survey. Furthermore, the classification applied for economic activities corresponded to NACE Rev.1.

The survey of the year 2000 widened the economic activity coverage to include the section “Real estate, renting and business activities”. However, only enterprises with 10 or more employees participated in the survey.

The survey of the year 2004 widened the economic activity coverage to include the sections with NACE Rev.1 codes M, N and O. However, only enterprises with 10 or more employees participated in the survey.

6. Coherence

a. Coherence with statistics from the labour force survey

The number of hours actually worked per employee of the labour cost survey (LCS) and the labour force survey (LFS) are appeared in the following table.

Table 3. Hours actually worked per employee by section and survey

NACE Rev.1	LCS	LFS	Diference (%)
Total	1.583	1.957	-19,1
C	1.722	2.095	-17,8
D	1.708	2.067	-17,4
E	2.096	2.052	2,1
F	1.655	2.081	-20,5
G	1.756	2.093	-16,1
H	1.466	2.202	-33,4
I	1.898	2.197	-13,6
J	1.589	1.963	-19,1
K	1.630	2.002	-18,6
M	929	1.136	-18,3
N	1.881	1.937	-2,9
O	1.665	1.950	-14,6

Some significant differences in both surveys are appeared due to the fact that, the micro-enterprises (1-9 persons employed) were not included in the surveyed population of the labour cost survey. The micro-enterprises in Greece represent a significant share of the production structure and have a relatively high impact on the overall employment.

b. Coherence with structural business statistics

The variable “wages and salaries” per employee of the labour cost survey (LCS) and the Structural Business Survey (SBS) are appeared in the following table.

Table 4. Wages and salaries per employee by section and survey

NACE Rev.1	SBS	LCS	Difference (%)
Total	17.030	18.056	6,0
C	30528	21.315	-30,2
D	17981	17.305	-3,8
E	31963	42.482	32,9
F	12894	15.306	18,7
G	13419	15.198	13,3
H	11359	12.501	10,0
I	21569	26.022	20,6
J	30.349	29.261	-3,6
K	22847	17.624	-22,9

The deficiencies in coherence between SBS and LCS are due to the following reasons:

- In the SBS, the surveyed unit is the enterprise, whilst in the LCS the local unit. So, in the SBS, one enterprise may contain local units belonging to different economic activities and to different geographical regions
- In the SBS, the values of D11 do not contain the values of variables “payments to employees saving schemes” and “wages and salaries in kind”.

c. Coherence with Labour Cost Index

Table 5. Hourly Labour Costs of the LCS for the years 2000 and 2004 by economic activity

NACE Rev.1	YEARS		Difference (%)
	2000	2004	
Total	10,8	15,3	41,1
C	12,7	16,9	33,4
D	10,3	14,0	36,8
E	17,0	28,0	64,5
F	7,6	13,1	72,3
G	8,2	12,0	46,5
H	7,7	11,3	45,7
I	13,8	19,7	43,4
J	17,8	26,9	51,2
K	11,6	14,9	28,4

Table 6. Average annual LCI for the years 2000 and 2004 by economic activity

NACE Rev.1	YEARS		Difference (%)
	2000	2004	
Total	100,0	127,9	27,9
C	100,0	120,7	20,7
D	100,0	129,5	29,5
E	100,0	125,1	25,1
F	100,0	127,9	27,9
G	100,0	126,5	26,5
H	100,0	127,8	27,8
I	100,0	125,4	25,4
J	100,0	129,1	29,1
K	100,0	131,5	31,5

Table 7. Annual growths of hourly labour costs of the LCI and the LCS by economic activity

NACE Rev.1	Growth Rates		(LCS/LCI)-1 %
	LCI	LCS	
Total	27,9	41,1	47,2
C	20,7	33,4	61,6
D	29,5	36,8	25,0
E	25,1	64,5	157,0
F	27,9	72,3	159,6
G	26,5	46,5	75,6
H	27,8	45,7	64,6
I	25,4	43,4	70,6
J	29,1	51,2	75,8
K	31,5	28,4	-9,7

The reason for differences between the two sets of statistics (growth rates from LCI and LCS) is the different time schemes of the production of statistics for a given year. The annual statistics are collected after the year, whilst the short-term statistics are collected during the year. The population being surveyed changes during the year (births and deaths, mergers and break-ups etc). Such changes are better known when producing the annual than the short-term statistics. Hence, even if the target population is the same, the frames may be different for the two surveys.

6.3 Coherence with national accounts (NA)

In the following table the variable “compensation of employees” expressed per employee are appeared by section and source.

Table 8. “Compensation of employees” per employee by section and source.

NACE Rev.1	NA	LCS	Difference (%)
Total	17.547	18.056	2,9
C	24.857	21.315	-14,2
D	15.818	17.305	9,4
E	23.195	42.482	83,2
F	14.047	15.306	9,0
G	14.739	15.198	3,1
H	15083	12.501	-17,1
I	18.811	26.022	38,3
J	25.462	29.261	14,9
K	13.899	17.624	26,8
M	22.027	14.699	-33,3
N	20.512	19.880	-3,1
O	17.146	16.846	-1,7

The National Accounts build first on the short-term statistics and later on annual statistics, when the annual statistics are available. The values coming from NA are provisional data and they have been produced using the changes over time from the values of the Labour Cost Index. As a result, some significant differences are appeared between the compensations of employees between the NA and the LCS. Additionally, in national accounts the variables D1 and D11 do not contain the values of variables “payments to employees saving schemes” and “wages and salaries in kind”.

References

1. Bellhouse (1988). Systematic sampling. In Handbook of Statistics, Vol. 6, (Eds. P.R. Krishnaiah and C.R. Rao). Amsterdam: Elsevier Science, 125-145
2. Cochran, W.G. (1977). Sampling Techniques, New York: John Wiley and Sons
3. Dalenious T., and Hodges, J.L (1959). Minimum variance stratification. JASA, 54,88-101
4. Deming, W.E. (1953). On a probability mechanism and the bias of non-response. JASA, 48, 743-772
5. Evans, W.D. (1951). On stratification and optimum allocation. JASA, 46, 95-104
6. Hansen, M.H., Hurwitz, W.N., Madow, W.G. (1953). Sample Survey Methods and Theory. Vol. I, New York: John Wiley and Sons
7. Hess, I, Sethi, V.K., and Balakrishnan,T.R (1966). Stratification: A practical investigation. JASA, 61, 74-90
8. Holt, D., and Elliot, D. (1991). Methods of weighting for unit non-response. The Statistician, 40, 333-342
9. Kalton, G. and Kasprzyk, D. (1986). The Treatment of Missing Survey Data. Survey Methodology, 12, 1-16.
10. Kalton, G. and Flores – Gervantes, I. (2003). Weighting Methods. Journal of Official Statistics, 19, 81-97.
11. Kish, L., (1965). Survey Sampling, New York: John Wiley and Sons
12. Kish, L., and Frankel, M.R. (1974). Inference from complex samples. Journal of the Royal Statistical Society, A, 139,80-95
13. Kish, L., and Anderson, D.W. (1978). Multivariate and multipurpose stratification. JASA, 73, 24-34
14. Kish, L., (1987). Statistical Design and Research, New York: John Wiley and Sons
15. Kish, L., (1988). Multipurpose Sample Designs. Survey Methodology, 14, 19-32
16. Kish, L., (1992). Weighting for Unequal Pj. Journal of Official Statistics, 8, 183-200
17. Little, R.J.A. (1982). Models for non-response in sample surveys. JASA, 77, 237-250
18. Little, R.J.A. (1986). Survey non-response adjustments for estimates of means. International Statistical Review, 54, 139-157
19. Madow, L. H (1946). Systematic sampling and its relation to other sampling designs. JASA, 41, 207-214
20. Murthy, M.N. and Rao, J.T.(1988). Systematic sampling with illustrative examples. In Handbook of Statistics, Vol. 6, (Eds. P.R. Krishnaiah and C.R. Rao). Amsterdam: Elsevier Science, 147-185
21. Nikolaidis B. I.(2000). Variance Estimation in Business Statistics. Invited paper for the Eurostat working group on assessment of quality in statistics (4th and 5th April 2000). Doc. Eurostat/A4/Quality/00/General/Variance Estimation/NSSG
22. Royall, R.M and Herson, H.J. (1973). Robust Estimation in Finite Populations I. JASA, 68, 880-889
23. Royall, R.M and Herson, H.J. (1973). Robust Estimation in Finite Populations II: Stratification on a Size Variable. JASA, 68, 890-893
24. Sarndal, Swensson, and Wretman (1992). Model Assisted Survey Sampling. New York: Springer-Verlag

ANNEX I

The coefficients of variations (%) of the quantities “annual labour costs” and “hourly labour costs” are appeared broken down according to the structure of the tables A, B and C.

Table A. Coefficient of variations (%) by economic activity (Sections)

NACE Rev. 1	Annual labour costs	Hourly labour costs
Total	0,8	0,7
C	4,5	3,8
D	1,4	1,0
E	0,1	0,1
F	6,1	4,1
G	3,3	3,2
H	7,5	5,0
I	1,5	1,2
J	1,0	1,0
K	4,8	4,5
M	0,2	0,2
N	0,6	0,4
O	4,9	4,2

Table B. Coefficient of variations (%) by NUTS I

NUTS1	Annual labour costs	Hourly labour costs
Total	0,8	0,7
North Greece	1,7	1,4
Central Greece	1,9	1,4
Attica	0,9	0,8
Islands of Aegean and Crete	4,4	2,8

Table C. Coefficient of variations (%) by size class of enterprises

Size classes	Annual labour costs	Hourly labour costs
Total	0,8	0,7
10-49	3,3	2,4
50-249	1,2	1,0
250-499	2,1	1,7
500-999	2,7	1,9
1.000+	0,0	0,0