

**QUALITY REPORT ON THE
STRUCTURE OF EARNINGS SURVEY 2006**

CYPRUS

TABLE OF CONTENTS

	<u>PAGE</u>
INTRODUCTION	1
1. Accuracy	2
2. Punctuality and Timeliness	14
3. Accessibility and Clarity	16
4. Comparability	17
5. Coherence	18
Appendix A	20

INTRODUCTION

The Statistical Service of Cyprus conducted the Structure of Earnings Survey (SES) for the year 2006, during 2007. This was the second time that the SES was conducted in Cyprus. The survey was based on the Council Regulation No. 530/1999 and the Commission Regulation No. 1916/2000 as amended by Commission Regulation 1738/2005. The data collection and verification was completed by August 2008 and raw data were sent to Eurostat by mid-August 2008. The data sent covered enterprises with 1 or more employees, in sections C – O of the NACE Rev 1.1 Classification of Economic Activities System, of the European Union. Data were sent for 1216 enterprises in Cyprus, representing 26476 employees in the sample.

The aim of the Report is to evaluate the quality of the data from the Structure of Earnings Survey 2006 (SES 2006) in Cyprus, focusing on five main areas of interest:

1. Accuracy
2. Punctuality and Timeliness
3. Accessibility and Clarity
4. Comparability
5. Coherence

The report is based on the *Commission Regulation (EC) N° 698/2006* of 5 May 2006, implementing Council Regulation (EC) N° 530/1999 as regards quality evaluation of structural statistics on earnings.

1. ACCURACY

The Structure of Earnings Survey 2006 was a sample survey. The sampling procedure applied was divided into two stages. For the first stage, the sample was drawn using stratified, probability proportional to size sampling and for the second stage, the sample was drawn with the stratified simple random sampling method. Further methodological information relating to the SES 2006, is presented in *APPENDIX A*.

1.1. Sampling errors

Probability Sampling

The tables that follow present the coefficients of variation (C.V.), for the *grossed-up* results of the SES 2006, for the variables:

- **Gross earnings in the reference month**
- **Average gross hourly earnings in the reference month**

The above are broken down by:

- Full-time (separately for men and women) and part-time employees,*
- NACE section,*
- Occupation (ISCO -88 at the 1-digit level),*
- Age band (under 20, 20 – 29, 30 – 39, 40 - 49, 50 – 59, 60 and over),*
- Size band of the enterprise (1 – 9, 10 – 49, 50 – 249, 250 – 499, 500 – 999, 1000+)*

Please note that no breakdown by *NUTS level 1* is provided, since at this level, the whole of Cyprus is considered to be a single region. The estimator used for the coefficients of variation is the sum of gross earnings in the reference month and the average of gross hourly earnings in the reference month.

i) *Full-time Employees (separately for men and women)*

SEX	COEFFICIENTS OF VARIATION % Gross Monthly Earnings
MALES	1.4
FEMALES	1.5
TOTAL	1.0

SEX	COEFFICIENTS OF VARIATION % Average Hourly Earnings
MALES	1.4
FEMALES	1.6
TOTAL	1.0

ii) *Part-time Employees*

	COEFFICIENTS OF VARIATION % Gross Monthly Earnings
TOTAL	5.5

	COEFFICIENTS OF VARIATION % Average Hourly Earnings
TOTAL	5.3

iii) *NACE Section*

NACE REV. 1.1 BY SECTION	COEFFICIENTS OF VARIATION % Gross Monthly Earnings
C	6.9
D	2.1
E	2.7
F	2.8
G	3.7
H	3.8
I	2.7
J	1.8
K	3.8
L	0.9
M	1.2
N	2.2
O	5.4
TOTAL	1.0

NACE REV. 1.1 BY SECTION	COEFFICIENTS OF VARIATION % Average Hourly Earnings
C	6.7
D	2.1
E	2.7
F	2.5
G	3.7
H	3.6
I	2.8
J	1.8
K	3.7
L	0.9
M	1.2
N	2.1
O	5.3
TOTAL	1.1

iv) Occupation (ISCO-88 at the 1-digit level)

OCCUPATION ISCO 88 1-DIGIT LEVEL	COEFFICIENTS OF VARIATION % Gross Monthly Earnings
1	2.2
2	1.0
3	1.5
4	2.1
5	2.5
6	7.3
7	2.0
8	3.0
9	1.5
TOTAL	1.0

OCCUPATION ISCO 88 1-DIGIT LEVEL	COEFFICIENTS OF VARIATION % Average Hourly Earnings
1	2.2
2	1.0
3	1.5
4	2.0
5	2.5
6	6.8
7	1.8
8	2.5
9	1.3
TOTAL	1.1

v) *Age band*

AGE IN BANDS	COEFFICIENTS OF VARIATION % Gross Monthly Earnings
15 to 19 years old	6.0
20 to 29 years old	1.5
30 to 39 years old	1.5
40 to 49 years old	1.9
50 to 59 years old	2.0
60 and over	5.7
TOTAL	1.0

AGE IN BANDS	COEFFICIENTS OF VARIATION % Average Hourly Earnings
15 to 19 years old	4.1
20 to 29 years old	1.6
30 to 39 years old	1.5
40 to 49 years old	1.9
50 to 59 years old	2.0
60 and over	5.7
TOTAL	1.1

vi) *Size band of the Enterprise*

SIZE BAND OF ENTERPRISE	COEFFICIENTS OF VARIATION % Gross Monthly Earnings
1 - 9 Employees	2.9
10 - 49 Employees	2.3
50 - 249 Employees	1.3
250 - 499 Employees	1.1
500 - 999 Employees	1.6
1000 and more	0.5
TOTAL	1.0

SIZE BAND OF ENTERPRISE	COEFFICIENTS OF VARIATION % Average Hourly Earnings
1 - 9 Employees	2.8
10 - 49 Employees	2.3
50 - 249 Employees	1.3
250 - 499 Employees	1.1
500 - 999 Employees	1.6
1000 and more	0.5
TOTAL	1.1

Overall, the coefficients of variation produced for the above breakdowns lie within an acceptable range. Larger coefficients of variation are observed in groups with a smaller number of observations in the sample. However, even in such cases, the number of observations is large enough and all cells are deemed to be sufficiently reliable.

1.1.1. Non-probability sampling

The SES 2006 was based on *probability sampling*. Thus, no lack of precision due to non-probability sampling occurred in this survey.

1.2. Non-sampling errors

1.2.1. Coverage errors

Coverage errors observed in the survey refer to misclassifications, under-coverage and over-coverage errors. The SES 2006 covered enterprises with at least 1 employee, in the areas of economic activity defined by sections C – O of the Nace Rev.1.1 classification system. The *initial sample of the survey was 1856 units*.

1.2.1.1. Misclassifications

Misclassification refers to incorrect classification of units that belong to the target population. All misclassification errors identified in the SES 2006, were corrected in order to obtain a more realistic representation of the labour market in the sampling frame and the sample. Where appropriate, the weights of the affected enterprises were adjusted.

Misclassification with respect to the size-groups:

In the SES 2006, certain misclassifications in the size groups of the enterprises were identified and corrected. In order to correct the misclassification errors, all of the enterprises where the total employment was not consistent with the size group in the sampling frame were identified. Then, they were reclassified into the correct size groups, adjusting the weights of the enterprises accordingly. The table below presents a detailed analysis of the misclassifications.

Table 1.2.1: Number of misclassified units by Size-Group

Initial Size Group (No of Employees)	Final Size Group (No of Employees)	Number of Misclassifications
1 – 9	10 – 49	19
10 – 49	1 – 9	46
10 – 49	50 – 249	16
50 – 249	1 – 9	1
50 – 249	10 – 49	22
50 – 249	250 – 499	3
250 – 499	1 – 9	1
250 – 499	50 – 249	7
250 – 499	500 – 999	2
500 – 999	50 – 249	3
500 – 999	250 – 499	3
TOTAL		123

- Misclassifications in Size-Groups: 123 out of 1856 = **6.63%** of initial sample.

Misclassification with respect to the economic activity:

During the data collection process, misclassifications were identified in the economic activities of enterprises in the sample. In order to correct the misclassification errors with respect to the economic activity, all the enterprises where the economic activity reported was not consistent with that of the sampling frame were identified. Then, they were reclassified into the correct NACE groups, adjusting the weights of the enterprises accordingly. The table below shows misclassifications in the economic activity of enterprises at the NACE 2 digit level.

Table 1.2.2: Number of misclassified units by Economic Activity (NACE 2 digits)

NACE 2 digits in Sample Frame	NACE 2 digits in Final Sample																		Total
	22	23	28	36	45	50	51	52	55	60	63	67	70	72	74	85	91	93	
14		1																	1
15								1											1
16							1												1
17								1											1
18							1	1											2
19							1												1
20			1	2	1														4
21	1																		1
26				1	1		1												3
29			1				1												2
30														1					1
32					1														1
34						1													1
36								1					1						2
45										3			1						4
51			1	1	1														3
52							2												2
60					1						1								2
64								1											1
67						1					1								2
70					2														2
72								1											1
73																1			1
74	1		1						1			2							5
80														1	2				3
85																	1	1	2
90														1					1
92	1							1											2
Total	3	1	4	4	7	2	7	7	1	3	2	2	2	1	2	3	1	1	53

- Misclassifications in Economic Activities (2-digit level): 53 out of 1856 = **2.86%** of initial sample.

1.2.1.2. *Under-coverage errors*

Under-coverage errors refer to errors either due to units not included in the frame (real birth or de-mergers) or to wrongly classified units.

The SES 2006 was conducted under the assumption that the percentage of new enterprises not included in the sampling frame was negligible, since the sampling frame was constructed at a period fairly close to the reference period of the survey as well as the update period of the register of establishments (based on the Census of Establishments 2005).

1.2.1.3. *Over-coverage errors*

Over-coverage errors derive from the following:

- Wrongly classified units that are in fact out of scope
- Duplications in the sampling frame
- Dead or inactive units

In the SES 2006, two types of over-coverage errors were encountered:

- *Dead or inactive units (i.e. temporarily closed)*
- *Misclassified units that are in fact out of scope*

1.2.1.3.1. *Dead or inactive units.*

During data collection, 94 units were identified as dead or inactive units (**5,06% of initial sample**). These units had to be removed from the sample, and the weights of units that remained in the sample were adjusted accordingly.

1.2.1.3.2. *Misclassified units in fact out of scope*

305 units in the sample were identified as in fact out of scope (**16,43% of initial sample**). This was the case in enterprises with no employees (i.e. only the owners working in the enterprise), since the SES 2006 covers only enterprises/establishments with at least 1 employee. These units were removed from both the sample and the frame, and the weights of the remaining units were adjusted accordingly.

- Over-coverage errors: 399 out of 1856 = **21.50%** of initial sample.

1.2.2. Measurement and processing errors

1.2.2.1. Measurement errors

1.2.2.1.1. Survey Instrument Errors

The SES 2006 was the second survey of the series conducted in Cyprus. Therefore, the questionnaire designed, was actually an improvement of the first questionnaire, accommodating all suggestions for improvement in the layout received from the feedback of the first survey.

The questionnaire consisted of four pages and contained all the compulsory variables defined in the regulation for the SES 2006 and a number of optional variables. Some of the data requested were not as straightforward to understand as other. Thus, to prevent any misunderstandings, explanatory notes were prepared, providing detailed explanations on all the variables, as well as guidance on what to include and what to exclude from each variable.

1.2.2.1.2. Mode of Data Collection Errors

Data for the SES 2006 were collected by means of *personal interviews*. The interviewers were trained specifically for the Structure of Earnings Survey and then, they got in touch with the enterprises in the sample, in order to visit them and collect the data.

In most cases an employee of the enterprise would co-operate with the interviewer from the Statistical Service and provide the necessary information. In such cases, the errors were kept to a minimum, since the interviewers were familiar with both the questionnaire and the information needed for the survey.

In other cases the enterprise would provide the interviewer with administrative sources (accounts, payrolls, etc.) and ask them to locate and record the information needed without any further help. In these cases, since the questionnaires were completed by the interviewers the errors were minimised. However, some mistakes occurred in cases where the information provided by the enterprises was not fully understood by the interviewers.

In even fewer cases, the interviewers made the necessary explanations to the contact person from the enterprise and then, the enterprise would take the responsibility of filling out the questionnaires. In such cases the control over the data was even smaller for the Statistical Service, and the probability of errors in the data was larger.

In all of the above cases, the completed questionnaires were sent back to the Statistical Service, where the controlling officers would check the questionnaires for any inconsistencies, uncompleted

fields or other errors, and if any of these occurred, the enterprises were contacted again in order to clarify or correct the information given.

1.2.2.1.3. Respondent Errors

Respondent errors are most common in surveys where questionnaires are filled out by the respondents. Since in the SES 2006 the method used was that of personal interviews, such errors were minimised. Nevertheless, in the few cases where the respondents filled out the questionnaires, the interviewers and the controlling officers were extra careful in order to locate and correct any misleading data or mistakes.

In addition to the above, consistency checks were designed (using a software programme), in order to locate any inconsistencies in the data, which might have resulted from the provision of wrong information by the enterprises.

1.2.2.1.4. Information System Errors

These errors occur when the information system of the enterprise is unable to provide the data required for a specific survey. The information system of most enterprises in the sample could not provide accurate information on the following variables:

1. *Highest completed level of education*
2. *Number of weeks to which the gross annual earnings relate*
3. *Annual payments in kind (optional variable which must be included in the gross annual earnings in the reference year)*

The main source of these problems was that some enterprises did not keep proper records of their employees, especially with regard to their education level and annual payments in kind. This problem was more common in economic activities such as the construction, or industry sector, and mainly in small enterprises (under 10 employees).

Another problem that was quite common was that if the employees had left the enterprise by the time the data were collected, some enterprises usually did not keep their records.

Such problems were dealt with, in the following ways:

- In the cases where the data were not available at the local unit, but it *was* possible to locate the information needed from other administrative sources, then the interviewers would collect the data available from the local units, and then complete the missing data from the administrative sources.

- In cases where the employees selected in the sample had left the enterprise by the time the data were collected, they were replaced by other employees in the same position (occupation). It was therefore possible to obtain information from the records of the newly-selected employees. In cases where the position of employees who had left the enterprise ceased to exist after the employees left, then the employees were completely removed from the sample.
- Concerning information not readily available from the information system of enterprises, such as the “Number of weeks to which the gross annual earnings relate” and the “Annual payments in kind”, a secondary document was prepared requesting auxiliary information (maternity leave taken, sick leave not paid, etc.) in order to estimate the main variables.
- In the cases where it was impossible to obtain data on the above mentioned variables, the enterprises were asked to provide representative estimates.

1.2.2.1.5. Interviewer Errors

Interviewers were hired and trained to collect data specifically for the Structure of Earnings Survey. As was expected, at the beginning of the survey, some of the interviewers had not fully understood some of the concepts of the survey and how to locate the requested data from payrolls, accounts and other sources of information. With the help of the controlling officers who were more experienced in similar types of surveys, the questions and problems encountered were solved.

1.2.2.2. Processing errors

1.2.2.2.1. Data entry

Processing errors due to data entry were limited, and were mainly identified through validation and consistency rules applied after the stage of data entry, using specialised software.

1.2.2.2.2. Coding

The coding of the questionnaires was performed by employees who were specifically trained for this purpose. Their work was also checked by controlling officers, as well as the validation and consistency rules applied to the data file.

1.2.2.2.3. Editing

Editing of the data was either done prior to the data entry (first checks of the questionnaire) or after the first run of the validation and consistency rules. In each case, the edited data were checked/re-checked by the first/second run of the validation and consistency checks.

1.2.3. Non-response errors

1.2.3.1. Unit response rate

The number of enterprises/establishments initially selected in the sample covering NACE sections C–O and enterprises with 1 or more employees was 1856. Out of the 1856 units, 305 were out of scope (misclassified) and 94 units were dead or inactive. Thus, the final sample for the SES 2006 was 1457 units. Data were fully collected for 1216 units.

Thus, if *unit response rate* is defined as *the number of in-scope respondents to the number of questionnaires sent to the population selected*, then:

Unit Response Rate for the SES 2006:

1216 units out of the 1457 in the sample = **83.46 %**.

1.2.3.2. Item imputation rate

For the 1216 respondents of the Structure of Earnings Survey 2006, data for all the compulsory variables were collected as well as the optional variables that were included in the questionnaire. Even in cases where it was difficult for the enterprises to locate the information requested (for optional or secondary variables), realistic estimates were provided, or the data were collected using administrative sources. Therefore, no imputation methods were used.

1.2.4. Model assumption errors

No imputation methods were employed in the SES 2006. However, grossing-up factors were adjusted in order to correct for unit non-response. The level of unit non-response was quite low, 16.54%. Therefore, it could be safely assumed that the adjustment in the grossing-up factors did not significantly affect the estimates of the survey variables.

2. PUNCTUALITY AND TIMELINESS

Timeliness and punctuality refer mainly to pre-established and actual reference periods of data and publication dates.

2.1. Punctuality

2.1.1. The data collection phase

Compiling the questionnaire

The first draft of the questionnaire was designed in year 2006. It was later on revised many times, in order to meet the needs of the controlling officers and the interviewers. The questionnaire was finalised in the spring of 2007.

Methodological and explanatory notes for the questionnaire

Methodological and explanatory notes for the questionnaire were prepared as soon as the questionnaire was finalised, in the spring of 2007.

Drawing the sample

It was decided to use the same sample of enterprises as the Wages and Salaries Survey 2006, in order to be able to use some of the data collected under this survey, in the SES 2006. Thus, the sample of enterprises was drawn in September 2006. The second stage of the sampling procedure (sampling of the employees for each enterprise) was completed in September 2007.

Data collection process

The Structure of Earnings Survey was conducted by means of personal interviews. Thus, no questionnaires were sent out to the enterprises. The data collection began in May-June 2007 and was completed in the summer of 2008 (end of June).

2.1.2. The post-collection phase

Checking of the questionnaires by the controlling officers, and coding

At a first stage the questionnaires were checked by the controlling officers as soon as they were filled out, in order to clarify and correct any misleading or inaccurate data. Then, the coding of the questionnaires occurred. This phase was conducted simultaneously with the data collection phase, and was completed by the beginning of July, 2008.

Keying the data into the computer system

The data of the survey were keyed into the computer system simultaneously with the data collection. As soon as a batch of questionnaires was checked by the controlling officers, it was sent to the IT Department of the Statistical Service to be keyed into the system. The process of entering the data into the computer system was completed in the beginning of July 2008.

Plausibility checks

A software programme was developed in order to perform plausibility checks on the data of the survey, after they were entered into the computer system. The procedure for developing this programme was initiated in January 2008. The plausibility checks were performed simultaneously with the data collection and the data keying phase (as soon as the software programme was developed), and they were completed by the beginning of August 2008.

2.2. Timeliness

Micro-data sent to Eurostat

The micro-data were sent to Eurostat in the requested format, as soon as the plausibility and quality checks on the data were completed, on the 8th of August, 2008.

Dissemination of results

The Structure of Earnings Survey 2006 in Cyprus, covered sectors C – O of the NACE Rev. 1.1 classification system (including sector L), and enterprises/establishments with 1 or more employees. Detailed results of the survey are expected to be published by September 2009.

3. ACCESSIBILITY AND CLARITY

As mentioned in chapter 2 of the report, the results of the SES 2006 have not been released yet. However, the Statistical Service of Cyprus plans to release a report containing a brief analysis of the main results of the survey and detailed tables, concerning all the main and optional variables covered. It is expected that the report will be released by September, 2009.

The report will be available in both the English and the Greek language, on printed and electronic versions. The printed version of the report will be available from both the Statistical Service of Cyprus and the Government Printing Office of the Republic of Cyprus. The electronic version of the report will be available only from the offices of the Statistical Service of Cyprus. Furthermore, the Statistical Service of Cyprus also plans to publish the results of the SES 2006 on its official website, <http://www.mof.gov.cy/cystat>, under the theme of labour statistics.

Even though the results will be published on the internet, the Statistical Service of Cyprus will notify governmental and semi-governmental departments that are interested in the issues covered by the survey, for the publication of these results. Such departments are: the Ministry of Finance, the Ministry of Labour and Social Insurance, the Planning Bureau, the Human Resource Development Authority, etc.

4. COMPARABILITY

4.1. Geographical comparability

The statistical units, economic activities to be covered and the definitions of the variables for the Structure of Earnings Survey 2006 (SES 2006) were based on the Council Regulation No. 530/1999 and the Commission Regulation No. 1738/2005, and were as follows:

4.1.1. Economic Activities covered:

For the SES 2006, NACE Rev.1.1 C – O activities were covered.

4.1.2. Statistical Units:

Data were collected from *enterprises* since the NUTS 1 level refers to the whole of Cyprus. The population of employees targeted for the SES 2006, were those who actually received remuneration during the reference month (October 2006).

4.1.3. Categories of workers included in the survey:

- Employees having a direct employment contract with the enterprise and which received remuneration, irrespective of the type of work performed, the number of hours worked (*full-time or part-time*) and the duration of the contract (*fixed or indefinite*).
- Apprentices and trainees with an employment contract with the reporting unit.
- Seasonal and occasional workers who were working pre-defined hours on a contractual basis.
- Interim or temporary workers employed by/through agencies – providing that the reporting unit was the agency actually employing them.
- Outworkers, but only if they were remunerated on the basis of the amount of hours worked.
- Employees on maternity leave as long as they received remuneration from the employer.

4.1.4. Size Groups of the enterprises:

Size groups were defined according to the regulation, including the size of 1–9 employees.

4.2. Comparability over time

The definitions of variables for the SES 2006 were according to the requirements of the Regulation. The coverage of the survey for 2006 was extended in comparison with that of 2002. More specifically, the survey covers NACE sections C–O (including L) and enterprises of all sizes (1 or more employees). The survey of 2002, covered enterprises with 2 or more employees and NACE sections C–K. Data for NACE sections L, M, N, and O were also collected on a pilot basis, but were not sent to Eurostat.

5. COHERENCE

Coherence refers to comparability of data from different domains and sources. The aim is to inform the users of data about the conceptual differences that exist between several sources of variables that are very similar and to provide information on how to move from one concept to the other. Another objective is to check that statistics, which are in principle coherent conceptually, give comparable results for the same year and reference population.

5.1. Coherence with the “Wages and Salaries”, per employee, of the NA

In the case of the Structure of Earnings Survey 2006, statistics sent to Eurostat should be compared with the variable “Wages and Salaries”, per employee, of the National Accounts (NA).

Differences between the SES and the NA

Before comparing the earnings variables between the two sources and drawing any conclusions, one should take into account the following:

- The source for the Gross Annual Earnings from the SES2006, is the survey, which is enterprise based. The Wages and Salaries variable from the NA, is obtained using a number of different sources, i.e. a number of surveys and administrative sources.
- In several of the groupings constructed, there is only a small number of persons, especially for part-timers. Thus, the distribution of employees among the various groups might also involve wide margins of statistical errors.
- Concerning the government sector, Wages and Salaries from the NA, are taken from the government budget and thus they also include pensions payable to *retired* government employees. This is not the case for the SES 2006, which only includes annual wages, salaries and other regular or irregular bonuses paid to employees who received remuneration in October 2006.
- The classification of economic activities in the SES 2006, was decided at the enterprise level (except for the government where it was based on the activities of the different ministries), while the NA classify economic activities based on the kind of activity unit (KAU) method.
- The “Wages and Salaries” variable from the NA refers to the **total** wages and salaries of employees for the year. In order to convert it to wages and salaries *per employee* to compare with the variable from the SES 2006, the figures provided by the NA were divided by the corresponding employment, which was obtained from a different source. This, of course, may create wider margins for errors when comparing the two sources.

The table that follows presents the Gross Annual Earnings per employee from the SES 2006 and the Wages and Salaries per employee from the NA.

Table 5.1.1: Comparison of Gross Annual Earnings per Employee from the SES 2006 and Annual Wages and Salaries per Employee from the National Accounts

Nace Rev.1.1 Section	Gross Annual Earnings - SES 2006	Wages and Salaries - National Accounts	Difference between the Sources as a percentage of the NA variable (%)
C	13,782	14,120	2.4
D	9,515	8,575	-11.0
E	16,913	18,536	8.8
F	10,568	9,936	-6.4
G	8,840	8,178	-8.1
H	6,878	7,300	5.8
I	14,300	12,742	-12.2
J	19,100	16,675	-14.5
K	11,160	12,889	13.4
L	15,777	24,473	35.5
M	17,639	16,984	-3.9
N	14,498	12,520	-15.8
O	9,360	9,613	2.6
C - O	11,630	12,099	3.9

5.2. Conclusions on coherence

Overall, for NACE sections C – O of the economy, the difference in annual earnings per employee between the two sources is 3.9%. The highest difference is observed in NACE section L (35.5%) and the smallest in NACE section C (2.4%).

The main reason for the difference between the two sources is the different methodology used when recording the wages and salaries for the government. The pension funds figure included in the NA variable is quite large and significantly affects not only NACE section L, but also the overall average.

Differences between the two sources in the distribution of the annual earnings by NACE section can be explained by the different method of classifying the economic activities. As mentioned earlier on, the NA use the economic activity of each unit, while the SES 2006 uses the economic activity of the whole enterprise.

Taking into account the above differences, the percentage difference of annual earnings between the two sources can be considered quite small.

APPENDIX A METHODOLOGICAL NOTES

This appendix provides information on the methodology relating to the statistics provided for the Structure of Earnings Survey 2006 (SES 2006).

1. The survey

The Statistical Service of Cyprus decided to integrate the annual survey on Wages, Salaries and Hours of Work with the SES, since the annual survey was in many ways similar to the SES. This was achieved by extending the existing survey in order to collect data relating to regulatory variables concerning the enterprise and designing an additional questionnaire, to cover any regulatory variables that concerned the employee and were not covered by the annual survey.

Both surveys were conducted by means of personal interviews.

2. The sampling frame

The sampling frame for the Structure of Earnings Survey 2006, covered the NACE Rev. 1.1 sections C – O (including L), and included enterprises with 1 or more employees, in the private, public or semi-government sector.

The sampling frame was constructed, using the Enterprises Register of the Statistical Service.

3. The sample design

The sampling procedure applied was divided into two stages, employing the enterprise as the primary sampling unit (first stage) and the employee as the secondary sampling unit (second stage).

4. Selection procedure - Stratification

First stage

In the first stage of the sampling procedure, the selection of enterprises to be included in the sample was made. The enterprises were selected using the *probability proportional to size* sampling procedure. The population of enterprises was stratified by economic activity (NACE.Rev.1.1 at the two-digit level) and within each NACE.Rev.1.1 division, by employment size class. No regional breakdown was conducted, since the NUTS1 level refers to the country level in the case of Cyprus. This was also the reason why it was preferred to hold the survey at the enterprise level and not at the local unit level.

Thus, the enterprises were stratified into 51 NACE groups (C – O at the 2-digit level) with the following employee size classes per group:

Employee size class	Number of Employees
1	1-9
2	10-49
3	50-249
4	250-499
5	500-999
6	1000+

Within each stratum, the number of enterprises to be included in the sample was decided, in order to achieve adequate representation.

Second stage

The second stage of the sampling process concerned the selection of employees.

The sampling frame for the selection of the employees was actually the set of all employees of the enterprises selected in the first stage of the sampling procedure. This was obtained from the Wages and Salaries Survey 2006, where all employees of the selected enterprises were recorded (occupation, full-time or part-time, monthly paid or weekly paid, etc.).

The employees within each enterprise were stratified according to the following criteria:

1. Full-time or part-time employee
2. Monthly paid or weekly paid
3. Occupation

Then, a simple random sub-sample was selected within each stratum. The number of employees from each stratum to be included in the sample was determined as follows:

Size of stratum	Number of employees included in the sample
Up to 4 employees	1
5 employees or more	1 in every 4 (25%)

The stratified sampling procedure ensured that all occupational groups were adequately represented for each type of employee (full-time or part-time, monthly paid or weekly paid).

5. Sample size

The final sample of enterprises (*respondents*) consisted of **1216 enterprises** from all over Cyprus (**government controlled area**), covering NACE Rev.1.1 sections C to O. These enterprises represented **26476 employees** (excluding non-responses), of whom **26047** were full-timers and **429** were part-time employees.

The 1216 enterprises in the final sample of the SES represent **4,3%** of the total number of enterprises in the sampling frame for the survey (27.971 enterprises), while the number of employees in the final sample, represents **9,5%** of the total number of employees in the sampling frame (272.778 employees in total).