

QUALITY REPORT ON THE 2002 STRUCTURE OF EARNINGS SURVEY

**/Implementation of COMMISSION REGULATION (EC) No 72/2002 of 16 January 2002
concerning quality evaluation of structural statistics on earnings/**

Part A

Structure of earnings survey: grossed-up results: tabular analyses

See the tabular analyses in the annexes, as follows:

- Annex 1 Employees by bands of gross hourly earnings, sex and full-time/part-time employment
- Annex 2 Employees by bands of gross monthly earnings, sex and full-time/part-time employment
- Annex 3 Employees by bands of gross annual earnings, sex and full-time/part-time employment
- Annex 4 Employees by bands of annual days of paid holiday leave, sex and full-time/part-time employment
- Annex 5 Employees by bands of monthly hours paid, sex and full-time/part-time employment
- Annex 6 Employees by NACE Rev. 1 sections (C to O, incl. L), sex and full-time/part-time employment
- Annex 7 Employees by occupation (1-digit of ISCO-88 COM) , sex and full-time/part-time employment
- Annex 8 Employees by education (ISCED 0 to 6), sex and full-time/part-time employment
- Annex 9 Employees by bands of age, sex and full-time/part-time employment
- Annex 10 Employees by length of service, sex and full-time/part-time employment
- Annex 11 Employees by size of the enterprise, sex and full-time/part-time employment

Note: The breakdown by NACE Rev. 1 sections and by NUTS level 1 is not provided as for Bulgaria NUTS 1 level is identical with NUTS 0 – the whole country territory.

Note: The numbers in the tables highlighted in red are either not published or are put in parentheses. The criteria applied are described below in section 2.1.1.2. Variance.

PART B

1. Relevance

The Structure of Earnings Survey (SES) is carried out for the first time in Bulgaria. The survey provides statistical data for an important area for which there has been an increasing demand in the last years from national and international users.

The user groups are defined on the base of the data requests received so far by NSI. The customers of the SES results can be classified as follows:

- Institutions:
 - National – Parliament, Ministry of Labour and Social Policy, Ministry of Finance, National Social Security Institute, Bulgarian National Bank etc.
 - International – Eurostat, ILO
- Social partners: Trade unions and employment associations
- Private institutions and businesses, incl. media
- Researchers and students
- Internal to NSI: other units of NSI, e.g. dealing with LFS, HBS, etc. for purpose of comparison or other.

The information provided by SES is employed by users for policy-making, investment decisions, scientific purposes, human resources management, etc. The most often demanded data are: earnings by occupations, economic activity, sex, full-time/part-time employment etc.; number of employees by level of gross and net earnings; share of persons working on the national minimum wage, etc.

The SES was designed mainly to implement the requirements of CR N530/1999 and CR N1916/2000 therefore to satisfy needs of Eurostat for harmonized at European Union level data. To satisfy Eurostat's needs, the Bulgarian NSI has submitted in due time the micro data form SES 20002 in the required technical format. NSI assumes that a full harmonization has been achieved in regard to the coverage and definitions applied. The survey covers all mandatory variables and a number of optional items. Included are local units that belong to enterprises with 10 and more employees with economic activities within the area of NACE Rev. 1 sections C to O, incl. L. The plausibility checks set in the Eurostat's SES Arrangements were met as well.

2. Accuracy

A *two-stage sample* is employed in Bulgaria. At the first stage, a stratified random sample of local units is drawn from the population of local units that belong to enterprises with 10 or more employees. Stratification criteria used are economic activity (2-digit level of NACE Rev. 1) and size of the local units in terms of the number of employees. Units in each cell defined by economic activity and size class were sorted by geographic location (NUTS3 level) so that an almost proportional sample by regions was drawn. The sample from local units was produced in SAS.

At the second stage, a systematic sample of employees was taken within each of the randomly selected local units. For each of size class was fixed different sampling rates (sampling steps) and a rotating random start digit for each local unit was allocated by computer so that the start digit is rotated continually over the economic activities .

Sample size for local units and employees by size classes

<i>Number of employees</i>	<i>1_49</i>	<i>50_249</i>	<i>250_499</i>	<i>500_999</i>	<i>1000+</i>	<i>Total</i>
Size classes of local units	1	2	3	4	5	
Sampling rates for local units	8%	12,6%	41%	100%	100%	11.5%
Sampling rates for employees within the local units	100%	50%	25%	20%	10%	
Planned total sampling percentage for employees	8%	7%	10%	20%	10%	10%

The total sampling size is distributed proportionally according to the relative share of the number of employees in each stratum of the total population. As the number of employees is distributed unevenly between size classes and enterprises – a large number of employees (25% from the total population) is employed in the very few number of local units (1.5 % from total population) belonging to 4th and 5th size classes – the square root allocation rule is applied to reallocate the sampling units from bigger to smaller size classes.

The size of the sample from local units is calculated on the base of the already reallocated number of employees in each stratum and according to the planned sampling fraction from employees depending on the size of the local unit. All local units from 4th and 5th size classes and stratification cells with fewer than 5 local units are included exhaustively in the sample.

2.1. Sampling errors

2.1.1. Probability sampling

2.1.1.1. Bias (optional item)

2.1.1.2. Variance

The coefficients of variations (CV) are calculated with the kind cooperation of experts from the Head office of the Federal Statistics of Germany in Wiesbaden. A tailor-made national software is employed to produce the CVs using H-T estimator.

Coefficients of variation are calculated by use of the following formulae:

$$CV_{x'_h} = \left(\sqrt{\frac{\sum_{i=1}^{n_h} (f_i * x_i^2) * \sum_{i=1}^{n_h} f_i}{\left(\sum_{i=1}^{n_h} f_i * x_i\right)^2} - 1} \right) * 100 * \left(\sqrt{\frac{\sum_{i=1}^{n_h} f_i - \sum_{i=1}^{n_h} n_i}{\sum_{i=1}^{n_h} f_i * \sum_{i=1}^{n_h} n_i}} \right)$$

x'_h - estimate of the total gross earnings

$$x'_h = \sum_{i=1}^{n_h} f_i * x_i$$

x_i - individual gross earnings of each sampled employee in strata h

n_i - number of sampled employees in strata h

f_i - individual grossing-up factor of each sampled employee in strata h,

$$f_i = \frac{M_h}{m_h} * \frac{N_{kh}}{n_{kh}}$$

M_h - total number of local units in the population in strata h

m_h - number of sampled local units in strata h

N_{kh} - total number of employees in the k-th sampled local unit in strata h

n_{kh} - number of sampled employees in the k-th local unit

CVs are small for most of the relevant items and important classification levels. The highest CVs are calculated in general for part-timers (monthly earnings of part-timers in section C Mining and quarrying shows a CV of 40.73%). High values appeared also in some levels of the more detailed cross tabulations, e.g. the value of 40.16% for monthly earnings of full-time female employees at age over 60 years in section J Financial intermediation.

The following criteria were agreed for data publishing:

- cells with earnings (monthly/hourly/annual) showing CV between 20 and 30% were put in parenthesis
- cells with earnings (monthly/hourly/annual) showing CV higher than 30% were hidden (deleted) and marked¹ with sloped forward dash ‘/’.

Coefficients of variation concerning **total gross earnings**, together with corresponding numerators (V) and denominators (TGE – total gross earnings) from which they result, are provided in the annexes as follows:

a) monthly earnings of employees in October 2002

- | | |
|----------|--|
| Annex 12 | CV by NACE Rev. 1 sections (C to O, incl. L), sex and full-time/part-time employment |
| Annex 13 | CV by occupation (1-digit of ISCO-88 COM), sex and full-time/part-time employment |
| Annex 14 | CV by education (ISCED 0 to 6), sex and full-time/part-time employment |
| Annex 15 | CV by bands of age, sex and full-time/part-time employment |

b) hourly earnings of employees in October 2002

- | | |
|----------|--|
| Annex 16 | CV by NACE Rev. 1 sections (C to O, incl. L), sex and full-time/part-time employment |
| Annex 17 | CV by occupation (1-digit of ISCO-88 COM), sex and full-time/part-time employment |
| Annex 18 | CV by education (ISCED 0 to 6), sex and full-time/part-time employment |
| Annex 19 | CV by bands of age, sex and full-time/part-time employment |

¹ In the same fashion are treated cells where the number of sampled employees is 3 or less.

c) annual earnings of employees in 2002

Annex 20	CV by NACE Rev. 1 sections (C to O, incl. L), sex and full-time/part-time employment
Annex 21	CV by occupation (1-digit of ISCO-88 COM), sex and full-time/part-time employment
Annex 22	CV by education (ISCED 0 to 6), sex and full-time/part-time employment
Annex 23	CV by bands of age, sex and full-time/part-time employment

Coefficients of variation (CV) concerning **total monthly paid hours** of full-time employees, together with corresponding numerators (V) and denominators (TPH – total hours paid) from which they result, are provided in the annexes as follows:

Annex 24	CV by NACE Rev. 1 sections (C to O, incl. L) and sex
Annex 25	CV by occupation (1-digit of ISCO-88 COM) and sex
Annex 26	CV by education (ISCED 0 to 6) and sex
Annex 27	CV by bands of age and sex

Note 1: The breakdown by NACE Rev. 1 sections and by NUTS level 1 is not provided as for Bulgaria NUTS 1 level is identical with NUTS 0 – the whole country territory.

Note 2: In Annexes 15, 19, 23 and 27 the coefficients of variation are calculated by age bands (15 -19, 20 - 29, 30 - 39, 40 - 49, 50 - 59, ≥ 60 years old) used in the national tabulations.

2.1.2. Non-probability sampling

The item is not relevant for the Bulgarian SES 2002.

2.2. Non-sampling errors

2.2.1. Coverage errors

—Description of the main misclassification, under and over coverage problems encountered in collecting the data.

The sample of local units was taken from the local units population as of 31.12.2001. The sampling frame represented the most current situation of the Business Register available at the time of the sampling. In the sampling frame population were included all local units with 1 or more employees that belong to enterprises with 10+ employees within the NACE sections C to O.

To study over- and under-coverage in the whole frame, the frame population of December 2001 was matched with a frame set up for comparison of December 2002 (containing 27 156 local units belonging to enterprises with 10+ employees). The result was that 20 143 local units matched between the years, 4 197 existed in the 2001 frame but not in the 2002 frame, while 7 013 local units were new compared with the 2001 frame.

The table below shows that the absolute majority of the over- and under-coverage local units are in the smallest size class with employees 1 to 49. The main reason for this is the high

variability of the size class of employment in the small units which gives as well an idea of the misclassification concerning inclusion in the sampling frame units that in the reference year had less than 10 employees. Other causes of frame errors are mergers and demergers, restructuring/re-registration of enterprises and the weakness of the register updating procedure to reflect births and deaths that occur in-between.

Size class of local units (number of employees)	Local units only in frame of 2001	Local units in both frames (matching cases)	Local units only in frame of 2002 ²
1_49	3 709	13 797	6 311
50_249	447	5 305	620
250_499	35	657	84
500_999	4	274	7
1000	2	110	-9
All	4 197	20 143	7 013

There were not misclassifications concerning inclusion in the frame of units belonging to NACE activities outside sections C to O.

—Description of the methods used to process these errors.

When the information relating to 2002 target population was available, a correction factor has been used to offset the most significant fluctuations between 2001 and 2002 to achieve results that are more representative for the situation in 2002.

Below is made a comparison between the frame populations of local units and employees as at 31st December of three successive years – 2000, 2001 and 2002. The last row presents the updated 2001 population of local units and the resulting grossed-up number of employees from the survey which shows a very close value to the 2002 employment figure.

It can be noted that the growth rates for employees is far less than these for number of local units. This proves that the main reasons for coverage errors are due to fluctuation of employment in the small units, restructuring and demergers. As the coverage for employees is satisfactory we consider that the coverage errors as regards number of local units would not have had significant impact on the survey results.

Year	No of LU	Growth rates	No of employees	Growth rates
2000 - frame used in the pilot SES	23413		1567634	
2001 - frame used for sampling in the main SES	24340	4.0%	1583272	1.0%
2002 frame	27156	11.6%	1630240	3.0%
2001 - updated with major fluctuations in 2002, used for calculation of grossing-up factors of LU	24798		1609317	

—Rates of misclassification, under-and over-coverage (optional item).

Below is a table showing over-coverage in the sample. The overall over-coverage rate *in the sample* is 3,1 percents of which nearly 77 % are due to change in the employment size class.

² The negative value in the last size class is due to decrease in employment.

Nace	LU population 2001	LU population 2001 updated	LU sample	No of over-covered LU in the sample	Overcoverage rate in the sample
10	26	26	23	5	21.7%
11	5	5	5	0	0.0%
12	2	2	2	0	0.0%
13	36	36	23	1	4.3%
14	99	99	27	0	0.0%
15	1327	1327	98	3	3.1%
16	32	32	20	0	0.0%
17	231	294	50	1	2.0%
18	1409	1409	120	5	4.2%
19	203	203	42	6	14.3%
20	261	261	43	1	2.3%
21	123	123	31	0	0.0%
22	241	241	38	1	2.6%
23	15	15	12	0	0.0%
24	202	202	52	1	1.9%
25	211	211	39	2	5.1%
26	241	241	49	0	0.0%
27	106	106	35	2	5.7%
28	395	425	51	1	2.0%
29	566	566	89	0	0.0%
30	25	25	11	0	0.0%
31	163	163	40	1	2.5%
32	67	67	21	0	0.0%
33	97	97	25	1	4.0%
34	45	45	15	0	0.0%
35	66	66	25	1	4.0%
36	336	336	47	3	6.4%
37	5	5	5	0	0.0%
40	280	296	62	0	0.0%
41	266	266	47	0	0.0%
45	1706	1698	117	4	3.4%
50	759	763	67	4	6.0%
51	2324	2324	119	8	6.7%
52	1677	1677	104	6	5.8%
55	862	1034	76	3	3.9%
60	868	868	100	6	6.0%
61	9	9	9	0	0.0%
62	25	25	13	1	7.7%
63	458	458	59	0	0.0%
64	648	648	77	0	0.0%
65	436	436	61	0	0.0%
66	114	114	29	0	0.0%
67	55	96	19	3	15.8%
70	204	204	32	1	3.1%
71	21	31	13	1	7.7%
72	193	193	34	2	5.9%
73	187	187	35	0	0.0%
74	1037	1066	103	4	3.9%
75	1424	1424	109	0	0.0%
80	1737	1737	185	1	0.5%
85	1356	1356	133	2	1.5%
90	282	343	45	0	0.0%
91	247	276	35	1	2.9%
92	525	528	61	3	4.9%
93	105	113	23	2	8.7%
Total	24340	24798	2805	87	3.1%

The following cases have been regarded as over-coverage in the Bulgarian SES: local units who were sleeping during or died in 2002, local units of enterprises with *considerably* less than 10 employees during and at the end of 2002. When there have been cases of over-coverage, new units have *not* been sampled.

2.2.2. Measurement errors

— Description of the methods used to assess measurement errors.

Survey tools were tested in the pilot SES which was undertaken within the frame of the Eurostat pilot project “Earnings and Labour Costs”. As a result some suitable improvements were done.

Survey tools used for purposes of data collection are as follows:

1. Dispatch note to each respondent about the purposes of the survey
2. Statistical questionnaire: part A collecting information for sampled local units and part B collecting information for each sampled employee.
3. Explanatory notes for sampling of employees within the local units and instructions on information required. Annex with the National Classification of Occupations (4-digit level).

The SES questionnaire and the other accompanying documents are delivered to the respondents either personally by a representative of regional statistical offices of NSI or were sent by post.

The local units itself sampled a certain number of their employees from a given starting number and with a step (from one to the next selected person) following the instructions of NSI and provided individual data for each employee.

Part A was required only on paper. For part B was developed an option for electronic data delivery. At about 10% from the sampled local units used this possibility.

The SES questionnaire used was designed to obtain all the mandatory variables in the Commission Regulation 1916/2000, together with the following optional items:

Var. 1.6. Total number of employees in the local unit

Var. 2.11. Years of career break in the current enterprise

Var. 3.3. Compulsory social security contributions and taxes paid by the employer on behalf of the employee to government authorities during the representative month

Var. 3.3.1. Compulsory Social Security Contributions

Var. 3.1.2. Taxes

This rest of this section gives a brief description of how some of the employee variables were derived from the SES questionnaire:

The employee’s sex (variable 2.1) and age (variable 2.2) are obtained from the personal identification number of employee. The first 6 digits represent the date of birth and the digit before the last gives the employee’s sex. Var. 2.2 is calculated as at 31st October 2002.

Length of service in the enterprise (variable 2.6) is calculated as at 31st October 2002. In the questionnaire we ask for the initial date of entry into the enterprise and for the number of months of career breaks.

Var. 2.6 = 31.10.2002 – Date of entry in to service – Number of long-term career breaks

Share of a full-timer's normal hours (variable 2.7.1). In the questionnaire we ask for “*Normal working hours per day*” and for the “*Normal working days per week*”. For some employees these two variables are averages, as they work different hours per day or different days per week. For a part-time employee with a given occupation, the share of normal hours relates to the normal working hours of the most similar full-time job. When corresponding full-time jobs could not be found, the normal hours of the part-timer was expressed as a % of the average working hours by the full-timers in the unit. In other words, var. 2.7.1. for a part-time employee was calculated as follows:

$$(Hours\ per\ day * Days\ per\ week)_{part-time\ employee} / (Hours\ per\ day * Days\ per\ week)_{full-time\ employee}$$

Average gross hourly earnings in the representative month (new var. 3.0) is not included in the SES questionnaire and no other sources are available to obtain it. Therefore, it was calculated using:

$$Var. 3.0 = var. 3.1 / var. 3.4.$$

Number of weeks to which the gross annual earnings relate (var. 3.2.1) was calculated using three ‘supporting’ variables:

- a) Number of days actually worked in 2002
- b) Number of days actually taken as paid leave in 2002
- c) Normal working days per week

where the value of var. 3.2.1. equals:

$$(Number\ of\ days\ actually\ worked + Number\ of\ days\ taken\ as\ paid\ leave) / (Working\ days\ per\ week)$$

— **Assessment of the bias and a description of the estimators used to correct the bias for a main variable, for instance, monthly earnings.**

2.2.3. Processing errors (optional items)

In Bulgaria, the plausibility checks used for the micro-data validation are incorporated in the computer programme for the data processing. A dialogue procedure is used: immediately after a piece of information is entered (i.e., either a single variable, or complete data for a local unit when electronic data delivery is used), the programme checks the plausibility of the data. The computer operator entering the data sees all data for a given employee on the PC screen, together with the codes and text message of all controls not being met by the data.

In order to ‘automatize’ the data capture process, respondents were given the possibility of returning data in an electronic format. Respondents were given instructions about how to format data extracted from personnel or earnings administration files, so that the dataset received in the statistical office could be automatically added to the survey database.

The controls used can be divided into two basic types: absolute and relative controls.

Absolute controls are those that check impossible/missing values or codes for a single variable. For example:

- the absence of an occupational code; the employee's month of birth cannot lie outside the interval 1 to 12; likewise, the employee's day of birth cannot fall outside the values 1 to 31.
- consistency between values of a variable with other data for a given employee: e.g. an employee working 8 hours per day and a 5-day working week is unlikely to be a part-time employee; an aggregate variable should be greater than each of its components e.g. $\text{var. 3.1} > \text{var. 3.1.1}$, $\text{var. 3.1} > \text{var. 3.1.2}$, etc.
- consistency between parts A and B of the questionnaire: e.g. check if the number of the selected employees in part B corresponds to the minimum sample size defined by the number of employees with earnings in the representative month given in part A and the sampling fraction for the corresponding size class to which the local units belongs.

The presence of absolute controls in the error report means that the value of certain variable(s) is not acceptable and must be corrected. Records that fail the absolute plausibility checks prevent the regional offices from sending their databases to the NSI Head office.

Relative controls are those which indicate if values or codes are either 'improbable', or 'probable' if some very specific conditions are met. For example:

- very low hourly earnings (below the national minimum hourly wage) are 'improbable' but very high hourly earnings (50 times the minimum hourly wage) are 'probable' in banks or in branches of foreign companies.
- Other 'probable' examples include: persons without any annual paid holiday entitlement (this can occur if the job is temporary); persons who do not pay taxes (disabled people are entitled to a special and sizeable tax relief); persons without social security contributions (state officials do not pay social security contributions).

Records that fail the relative plausibility checks allow the regional offices to send their databases to the NSI Head office. A further analysis is carried out in the Head office on the remaining relative checks which are not being met, in order to ascertain their probability.

A number of very demanding absolute controls were initially set when the data processing computer programme was designed. Most of these controls were based on the conditions laid down by the labour, social and tax legislation in force. However, after the survey had started, we were advised by the ROs and by respondents all over the country that many of plausibility checks would not be met because of the variety of working conditions. Put another way, the formal contractual arrangements (or those required under the law) often do not correspond to the actual practices. In the circumstances, it was necessary to revise the original plausibility checks and to convert some of the absolute checks into relative checks.

Treatment of Wrong Data

The territorial units of the Bulgarian NSI undertook the data entry process. Having entered a certain number of questionnaires, an error listing was printed. If the micro-data failed the plausibility checks, the incorrect records were displayed in a prescribed manner.

Incorrect data were treated in different ways. Regional offices received detailed instructions on how to proceed when a plausibility check was not satisfied, together with the possible reasons for failure. Following the production of error reports, the persons responsible in the regional offices checked for data-entry errors, which were then corrected. The other frequent

type of error were systematic errors by the reporting units which appeared in all employees' records – eg due to different bookkeeping methods, or a failure to understand or read the explanatory notes. Examples of systematic errors were: var.3.4.1 was not included in var. 3.4; var. 3.2.2 was not included in var. 3.2; days of paid leave actually taken were included in the number of days worked during the year. After receiving confirmation of such systematic errors by the reporting unit, the regional offices applied the suitable corrections to all records.

When there were significant inconsistencies between variables in a given record, it was deleted. Data was only rarely returned to the respondent, e.g. in the case of totally wrong occupational codes, or the inclusion in the sample of many employees without monthly earnings.

As mentioned above, many of the absolute checks were converted into relative checks due to the great variety of working conditions. So we allowed the ROs to send the data with a number of relative checks having not being satisfied, for analysis in the NSI Head office.

The table below presents the rates of corrected cases for the main variables. In calculations are included all corrections done to the following errors: measurement, respondent, item non-response and data entry errors.

Code	Variable label	% of corrected cases
var. 3.5.1	Annual days of holiday leave	6.1%
var. 3.2	Total gross annual earnings in the reference year	4.4%
var. 3.2.1	Number of weeks to which the gross annual earnings relate	17.0%
var. 3.2.2	Total Annual Bonuses	3.3%
var. 3.1	Total gross earnings in reference month	5.5%
var. 3.1.1	Earnings related to overtime	5.6%
var. 3.4	Number of hours paid during the representative month	8.9%
var. 3.4.1	Number of overtime hours paid in the representative month	3.8%

—**Methodological notes on the estimation of these rates.**

(Section to be completed)

2.2.4. Non-response errors

— **Unit response rate.**

The table below contains of unit-response rates, broken down according to the stratification used for sampling. The overall, *non-weighted*, response rate was 96,4%, *including* 3,1% over-coverage. Units non-responded by reasons are as follows:

- a) dead/sleeping units – 18 LU
- b) merged – 2 LU
- c) not found – 3 LU
- d) refusals - 79 LU

Response rates						
NACE	Number of employees					
	1_49	50_249	250_499	500_999	1000_	Total
10	60%	100%	100%	60%	100%	84%
11	100%	100%	100%			100%
12	100%	100%				100%
13	88%	100%	67%	100%	100%	91%
14	93%	100%		100%		96%
15	92%	100%	100%	100%	100%	96%
16	100%	100%	100%	100%	100%	100%
17	100%	93%	100%	100%	100%	98%
18	90%	94%	93%	100%	100%	93%
19	88%	85%	100%	67%	100%	88%
20	96%	83%	100%	100%		93%
21	94%	100%	100%	100%		97%
22	92%	80%	100%	100%		89%
23	100%	100%	100%	100%	100%	100%
24	95%	100%	100%	100%	100%	98%
25	95%	100%	100%	100%	100%	97%
26	100%	100%	100%	88%	100%	98%
27	100%	100%	100%	83%	100%	97%
28	100%	100%	100%	100%		100%
29	100%	100%	100%	100%	100%	100%
30	100%	100%	100%			100%
31	100%	91%	100%	100%	100%	98%
32	90%	100%	100%			95%
33	100%	100%	50%			96%
34	100%	100%	100%	100%		100%
35	100%	100%	100%	75%	100%	96%
36	100%	100%	100%			100%
37	100%	100%				100%
40	100%	100%	100%	100%	100%	100%
41	96%	92%	83%	100%	100%	94%
45	92%	100%	100%	86%		95%
50	94%	93%	100%			94%
51	98%	96%	100%	100%	100%	97%
52	99%	95%	100%	100%		98%
55	94%	100%	100%	0%		95%
60	98%	100%	100%	100%	100%	99%
61	100%	100%		100%	100%	100%
62	100%	100%	100%		0%	92%
63	100%	93%	100%	100%	100%	98%
64	100%	100%	100%	100%	100%	100%
65	94%	93%	100%	100%	100%	95%
66	57%	71%	100%			62%
67	93%	100%				95%
70	100%	100%	100%			100%
71	91%	100%				92%
72	96%	100%	50%			94%
73	100%	100%	100%			100%
74	94%	96%	93%	100%	100%	95%
75	98%	100%	100%	90%	100%	98%
80	100%	100%	100%	99%	100%	99%
85	100%	100%	100%	100%	91%	99%
90	100%	100%	100%		100%	100%
91	89%	71%				86%
92	94%	95%	100%		100%	95%
93	94%	100%				96%
Total	96%	97%	98%	97%	98%	96%

— **Item response rates for the main variables (for instance, monthly and annual earnings and hours worked).**

Normally, no item non-response (blank or zero values) has been accepted for any of the key variables. Only nine of the collected items could possibly be zero and therefore item non-response could be supposed:

Var. 2.11 – could not be checked as no auxiliary information is available.

Var. 3.2.2 – checked with information available for the same reference period from the quarterly³ and annual⁴ surveys on labour carried out by NSI. Imputations are undertaken when there was significant lack of consistence between the gross monthly (GME) and gross annual earnings (GAE), e.g. $GAE > 15 \text{ times GME}$ and the auxiliary information showed that the given local unit paid irregular annual bonuses to its employees. The rate of imputed cases is insignificant – 0.1%.

Var. 3.1.1 – could only partially be checked with the information available for the whole enterprise (to which the sampled local unit belongs) from the Quarterly Survey on Labour.

Var. 3.1.2 - could not be checked as no auxiliary information is available.

Var. 3.3.1 - missing value is possible only if the person is a state official; in case of zero value the economic activity, type of ownership and the name of the local unit were checked. For those units which are not supposed to employ state officials var. 3.3.1 is estimated using the statutory social security contribution rate for employees in the reference year. The rate of imputed cases is 0.5%.

Var. 3.3.2 – when the zero value had been detected as improbable by the corresponding plausibility check, the item has been suitably estimated applying the monthly tax schedule from the Personal Income Tax in the reference year. The rate of imputed cases is 0.7%.

Var. 3.4.1 – could only partially be checked with the information available for the whole enterprise (to which the sampled local unit belongs) from the Quarterly Survey on Labour.

Var. 3.5.1 – could not be checked with auxiliary information, however some imputations were undertaken for employees with indefinite contract duration and more then 1 year length of service in the enterprise. At about 1% of all cases were imputed.

— **A description of the methods used for imputation and/or re-weighting for non-response.**

(Section to be completed)

— 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.

The plausibility checks applied by NSI did not allowed item non-response for any of the main variables. When item non-response was detected for a main variable local units had been contacted and missing data requested.

2.2.5. Model assumption errors

Report on the use of the following models:

— **to ensure that a representative month is selected,**

³ Quarterly survey on employees, time worked, wages and salaries and other labour costs.

⁴ Annual survey on employed persons, wages and salaries and other labour costs.

The representative month of the Bulgarian SES is October. This month has been selected as least influenced by absences as public holidays and annual paid leave.

— **to adjust the accounting or fiscal year to the calendar year,**

There is no difference between fiscal and calendar year in Bulgaria.

— **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),**

The Bulgarian SES 2002 covers the NACE Rev. 1 sections C to O. The full coverage of these activities depends on the ability of the register to reflect the changes of the target populations.

— **to combine data from administrative sources and surveys.**

No combinations between register and survey data were made.

3. Timeliness and punctuality

— **Key data-collection dates**

— **Key dates for the post-collection phase**

— **Key publication dates.**

The Bulgarian SES for 2002 had the following timetable:

October 2002 – Preparation of the survey organization and training of the staff of regional offices of NSI

November 2002 – Set up of target population from local units and employees

December 2002 – March 2003

Development of the survey tools and computer programming; printing out of the questionnaires explanatory notes, annexes.

April 2003 – Start of the survey at regional level: distribution of questionnaires to the respondents by the 28 regional statistical offices (RSOs) of NSI.

April – June 2003 – Data collection from the respondents. (Collection the last questionnaires from units belonging to strata with high non-response rate or higher significance for final results in given strata completed in October 2003).

July – September 2003

Data processing and data validation at RSOs.

October 2003 – April 2004

Processing, validation and analysis of the data in the Head office of NSI.

May 2004 – Producing of output tables with final results.

June 2004 – Transmission of the micro data files to Eurostat. Data release on Internet.

July 2004 – Sending out of to the respondents a set of survey results

July 2004 - November 2004

Preparation and issue of the printed and electronic publication with detailed survey results.

August - December 2004

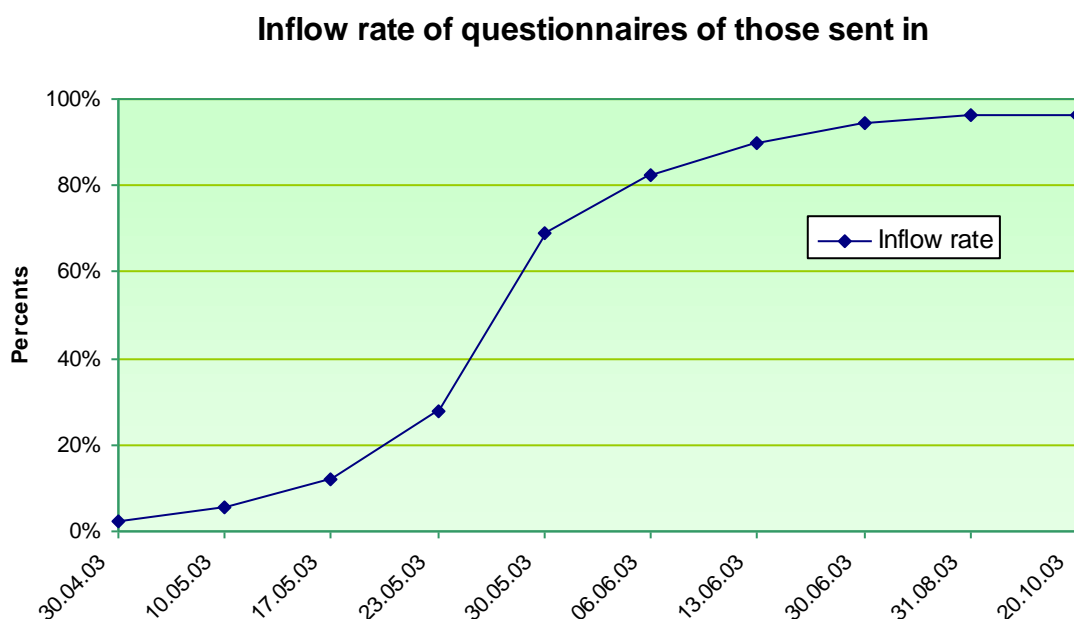
Preparation and 1st transmission of Quality report to Eurostat.

The initially set deadline was 31 May 2003 but as the non-response rate was high it was prolonged to 30 June. As mentioned before the final data collection stopped in October for some activities and size classes.

The follow-up activities are normally done by 28 ROs of NSI. Each regional office has its own practice to send reminders – usually a week before the initially set deadline (30/05/03) and 1 or 2 weeks after this date. Respondents are contacted either by phone or are sent an official letter of reminder. At about 30 letters were prepared in the Head Office and sent to the most important respondents with a signature by the President of NSI.

About 15% of the respondents submitted the completed questionnaires on the second visit from ROs staff and 30 % after reminder phone calls and letters.

It has been possible to draw a timeline over the inflow rate of the sent in questionnaires, since notes were taken about this every Monday throughout the data collection. The inflow rate timeline is useful for planning the SES 2006: how long need to work with survey and during what time, when the optimal time to send out questionnaires and reminders is, etc.



In conclusion, all deadlines posed by Eurostat and EU Regulations are met. However, the planning of the national key dates for the post-collection phase and publication dates for the next survey should be more realistic considering that SES is burdensome and time-consuming one. In the 2002 SES we posed a very ambitious national timetable which was not met – there was a five-month delay from the planned dates for first data release and for the printing publication.

4. Accessibility and clarity

— **A copy of the publication(s), or a reference to where it can be located.**

Specialized publication “Structure of Earnings 2002” containing methodology, review of the main results and detailed table results was prepared on paper and electronic format. The publication is bilingual – in Bulgarian and English language.

A copy of the publication will be sent to Eurostat.

— **Information on what results, if any, are sent to reporting units included in the sample.**

Each of the respondents was informed in the dispatch note for the possibility to receive for free standard results from SES. Approximately 1% of the sampled local units expressed their will to receive results. The information was sent on a diskette accompanied by a personal letter to each respondent and signed by the President of NSI.

Contents of the results sent back to the respondents

Methodology of SES 2002

List of Tables

1. Monthly earnings of employees in October 2002

1.1. Employees, paid hours, gross and net monthly earnings by mode of employment, sex and **economic activity groupings**

1.2. Employees, paid hours, gross and net monthly earnings by mode of employment, sex and **size of the enterprise**

1.3. Full-time employees and gross monthly earnings by sex **economic activity groupings and occupations**

1.4. Full-time employees and gross monthly earnings by sex **economic activity groupings and education**

1.5. Full-time employees and gross monthly earnings by sex **economic activity groupings and age**

1.6. Employees by **bands of gross monthly earnings**, mode of employment and sex, quartile and decile points

2. Hourly earnings of employees in October 2002

2.1. Employees, gross and net hourly earnings, and gross overtime hourly earnings by mode of employment, sex and **economic activity groupings**

2.2. Employees, gross and net hourly earnings, and gross overtime hourly earnings by mode of employment, sex and **size of the enterprise**

2.3. Full-time employees and gross hourly earnings by sex, **economic activity groupings and occupations**

2.4. Full-time employees and gross hourly earnings by sex, **economic activity groupings and education**

2.5. Full-time employees and gross hourly earnings by sex, **economic activity groupings and age**

2.6. Employees by **bands of gross hourly earnings**, mode of employment and sex; quartile and decile points

3. Annual earnings of employees in 2002

3.1. Employees, gross annual earnings and annual days of paid holiday leave by mode of employment, sex and **economic activity groupings**

3.2. Employees, gross annual earnings and annual days of paid holiday leave by mode of employment, sex and **size of the enterprise**

3.3. Full-time employees and gross annual earnings by sex, **economic activity groupings and occupations**

3.4. Full-time employees and gross annual earnings by sex, **economic activity groupings and education**

3.5. Full-time employees and gross annual earnings by sex, **economic activity groupings and age**

3.6. Employees by **bands of gross annual earnings**, mode of employment and sex; quartile and decile points

— **Information on the dissemination scheme for the results (e.g. to whom the results are sent).**

The above listed results were sent to:

- the Parliament of Republic of Bulgaria;
- the Central Bank;
- some ministries and local authorities, incl. few courts of justice;
- few of the leading banks and insurance companies;

- the Confederation of the Independent Syndicates in Bulgaria;
- certain number of private companies.

— **A copy of, or references for, any methodological documents relating to the statistics provided.**

Methodology and the basic results are published in section “Labour market” on the web site of NSI: <http://www.nsi.bg/>

5. Comparability

5.1. Geographical comparability

There are not differences between national concepts and European concepts as regard the definition of statistical units, the reference population, classifications and definitions of variables in the transferred results.

5.2. Comparability over time

The SES is carried out for the first time in Bulgaria.

6. Coherence

The results from the 2002 Structure of Earnings Survey are not fully comparable with the average gross earnings data and data on number of employees from the other surveys carried out by NSI with the same indicators:

1. Quarterly Survey on Number of Employees, Time Worked, Wages and Salaries and Other Labour Costs
2. Annual Survey on Employed Persons, Wages and Salaries and Other Labour Costs
3. Labour Force Survey

Basic conceptual differences of SES from the above-sited surveys:

Target population from enterprises: enterprises with 10 and more employees

Coverage by economic activities: Agricultural sector is excluded (sections A and B)

Target population from employees: employees with earnings during October 2002, having at least 1 working day paid by the employer at a full rate. Excluded are all persons without earnings in the reference month but are nevertheless in labour relationship with their employer.

Sample design: two-stage stratified random sample

Other methodological characteristics:

- individual monthly earnings of employees which have been influenced by periods of unpaid absence (due to sick, maternity, unpaid leave etc.) or because the employee joint or left the firm during the month, are grossed-up onto a full-month level. Total gross earnings in the representative month for a ‘full month’ is estimated by a coefficient obtained from the ratio between the normal monthly working hours and the total number of hours actually paid reduced with the number of overtime hours paid in October 2002.

- from the gross monthly earnings are excluded irregular bonuses and gratuities not paid regularly at each pay period;

- individual annual earnings of employees which have been influenced by periods of unpaid absence (due to sick, maternity, unpaid leave etc.) or because the employee joined or left the firm during the year, are grossed-up onto a full-year level. Total gross earnings in the representative year for a 'full year' is estimated by a coefficient obtained from the ratio between the total number of working weeks (52.2) in the calendar 2002 and the actual number of weeks to which the annual remuneration relates. In the calculation of average gross earnings are included only these employees who have received earnings from the sampled local unit at least for 8 months (36 weeks) in the reference year. In this way the number of employees to which the calculations of average annual gross earnings relates is reduced with 16 %.

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

(Section to be completed)

The LFS results used in the comparison refer to a fixed week in September 2002. This period is the closest to the reference month of SES.

The required cross-tabulations are enclosed in the annexes as follows:

— **sex, age and economic activity (NACE Rev.1, at the section level),**

Annex 28 Coherence between SES and LFS (Sep.2002) structure of full-time employees by NACE Rev.1 sections (C to O)

Annex 29 Coherence between SES and LFS (Sep.2002) structure of part-time employees by NACE Rev.1 sections (C to O)

— **sex, age and educational level (ISCED 0 to 6),**

Annex 30 Coherence between SES and LFS (Sep.2002) structure of full-time employees by education (ISCED'97)

Annex 31 Coherence between SES and LFS (Sep.2002) structure of part-time employees by education (ISCED'97)

Note: ISCED 3 and ISCED 4 levels are not available separately in LFS so for purposes of comparison the SES data are aggregated.

— **sex, age and occupation (ISCO-88 at the 1- digit level).**

Annex 32 Coherence between SES and LFS (Sep.2002) structure of full-time employees by occupation (ISCO-88)

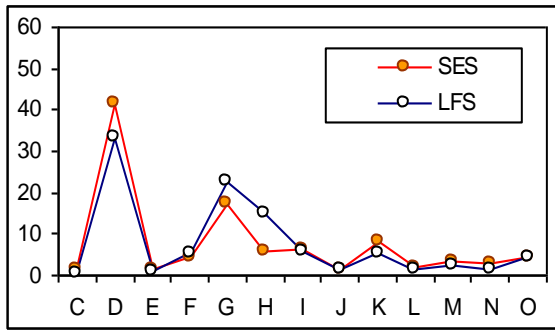
Annex 33 Coherence between SES and LFS (Sep.2002) structure of part-time employees by occupation (ISCO1)

The LFS data relate to all sizes of enterprises in the NACE activities C to O.

In the LFS figures are included the following other categories of employees (different than those covered by SES): employees paid in terms of commissions and fees under non-labour contract; employees working without contract with their employer (the non-registered employment). The graphs below give an illustration on comparability of the employees' structure between SES and LFS.

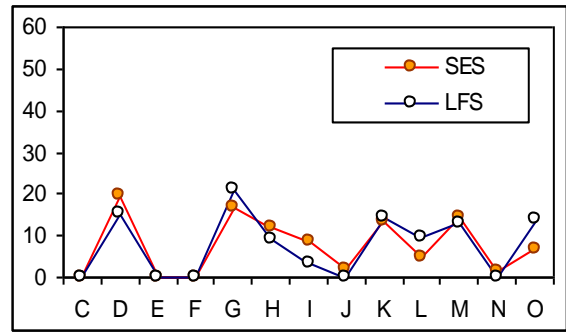
Coherence between SES and LFS (Sep.2002) structure of employees by age and NACE Rev.1 sections (C to O): distribution in percentages

Full-time employees: males and females

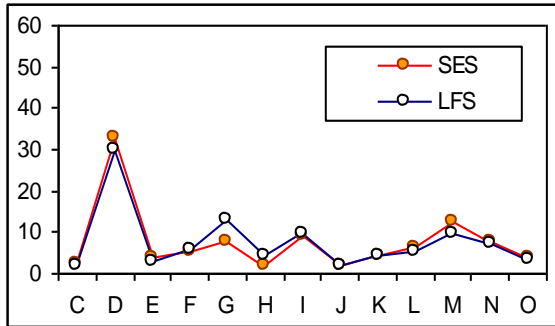


15 – 24 years

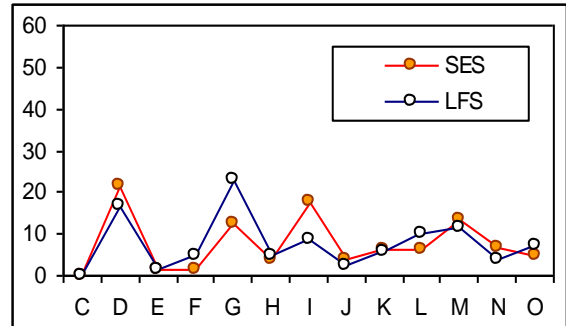
Part-time employees: males and females



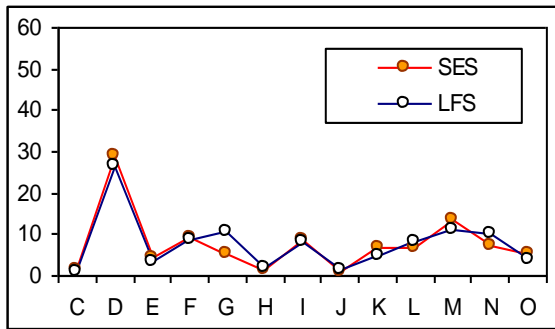
15 – 24 years



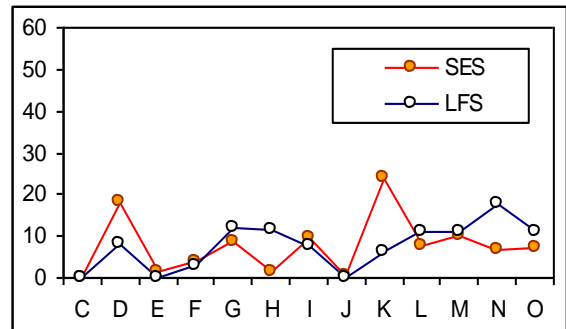
25 -54



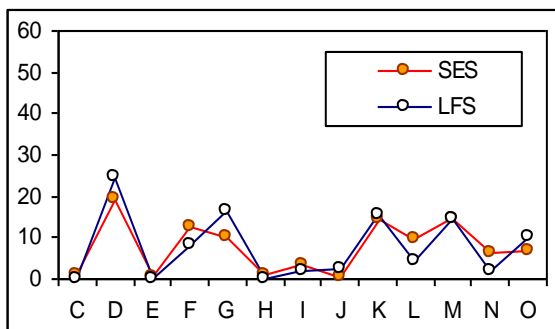
25 -54



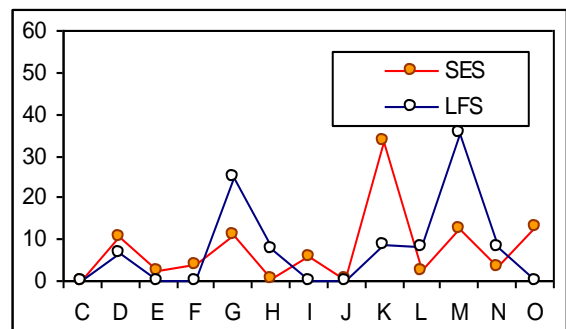
55 - 64



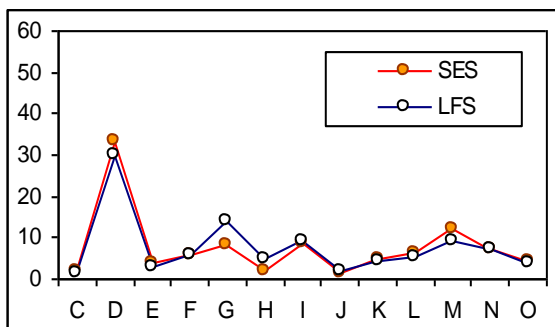
55 - 64



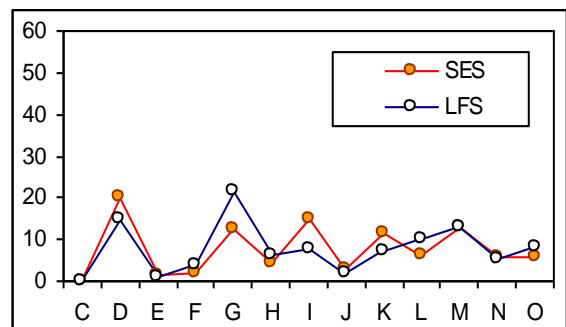
≥ 65



≥ 65



Total

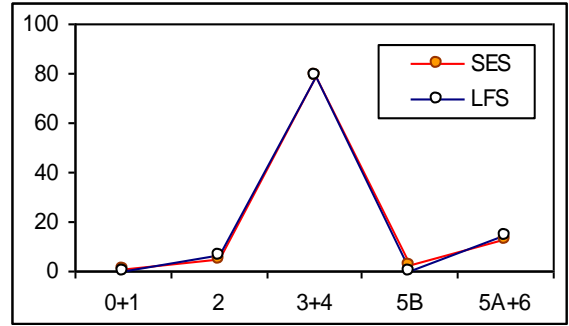
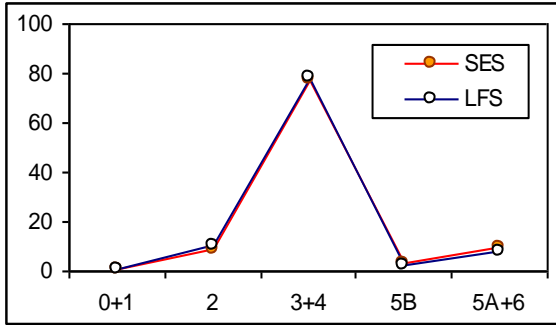


Total

Coherence between SES and LFS (Sep.2002) structure of employees by age and education (ISCED'97): distribution in percentages

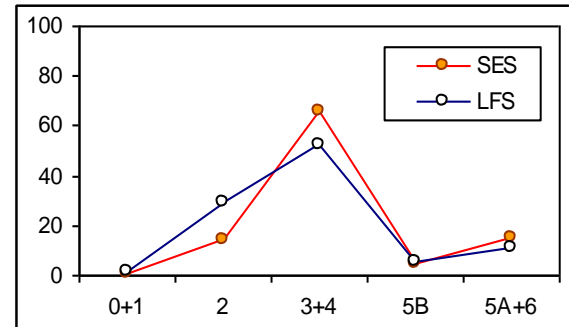
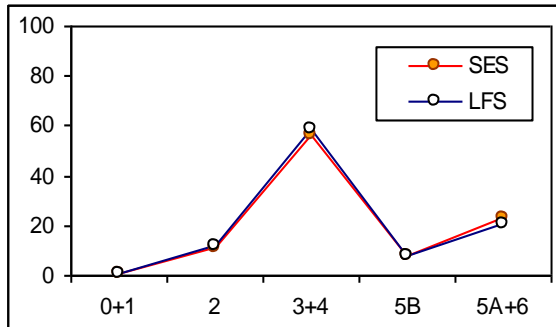
Full-time employees: males and females

Part-time employees: males and females



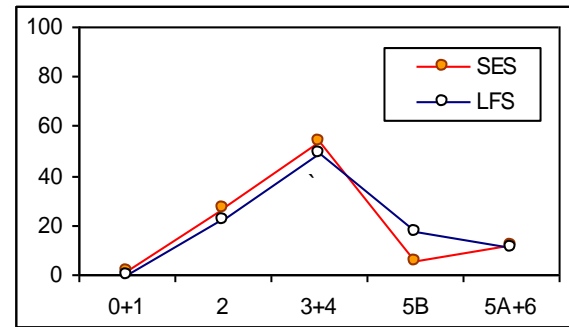
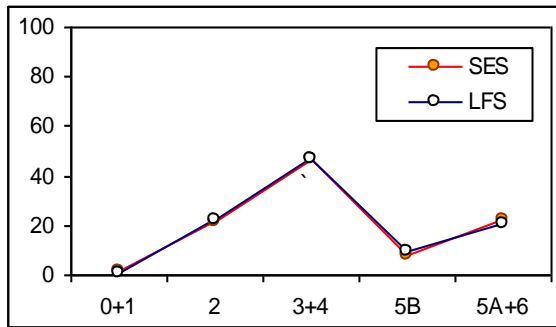
15 - 24 years

15 - 24 years



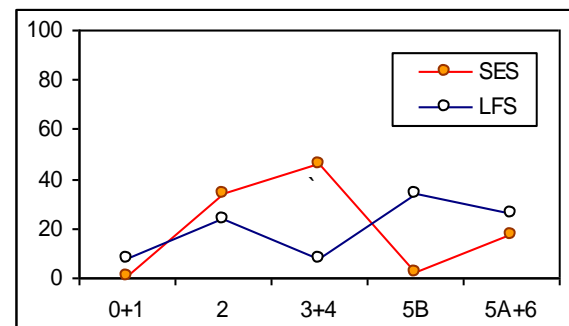
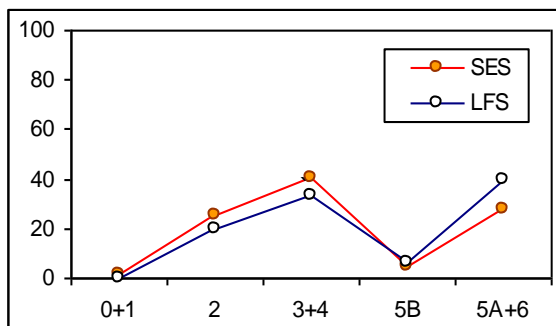
25 -54

25 -54



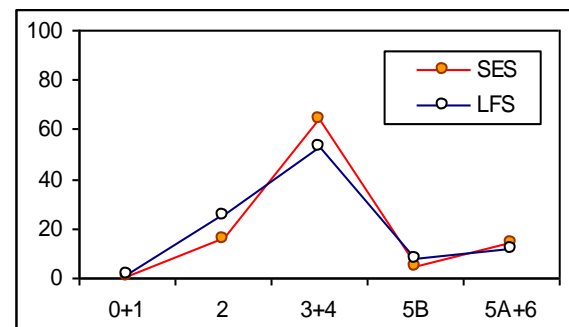
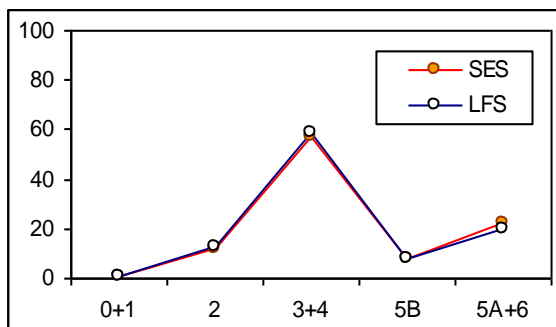
55 - 64

55 - 64



≥ 65

≥ 65

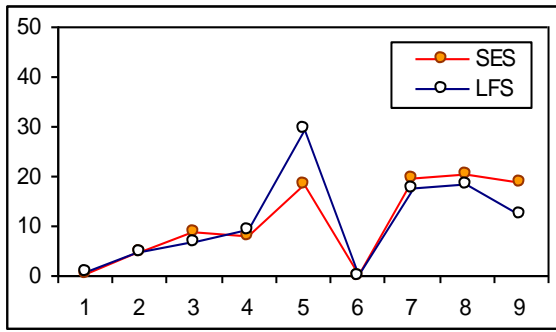


Total

Total

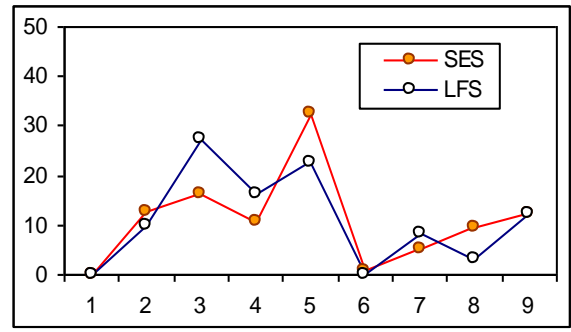
Coherence between SES and LFS (Sep.2002) structure of employees by age and occupation (ISCO-88): distribution in percentages

Full-time employees: males and females

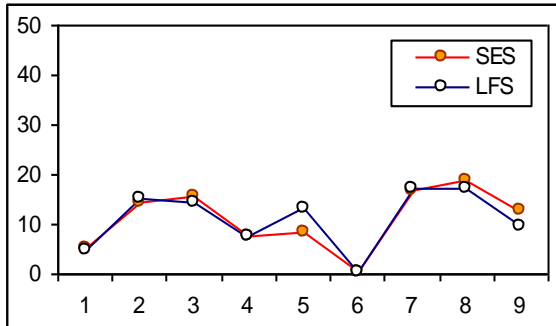


15 – 24 years

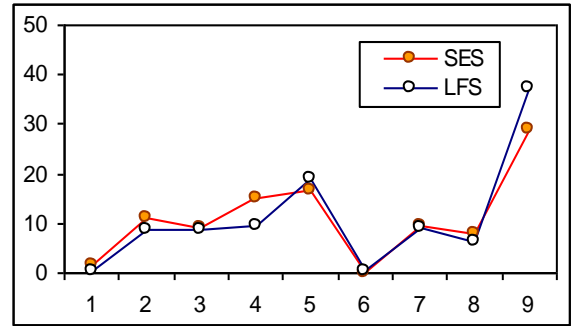
Part-time employees: males and females



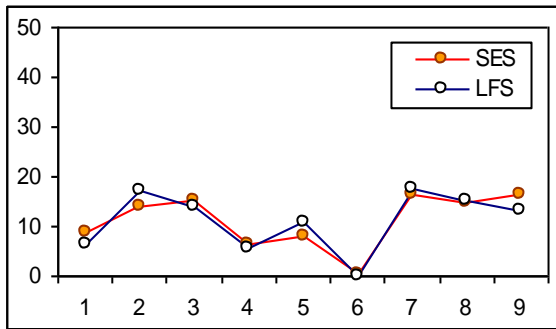
15 – 24 years



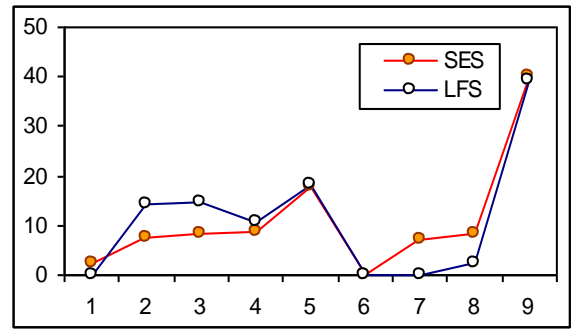
25 -54



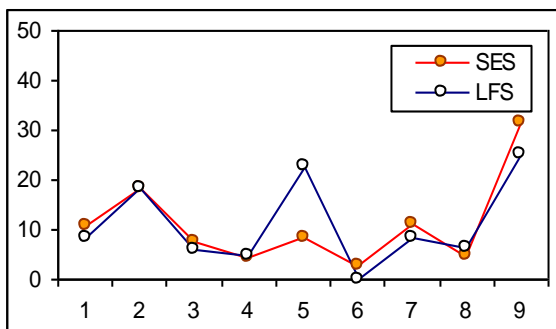
25 -54



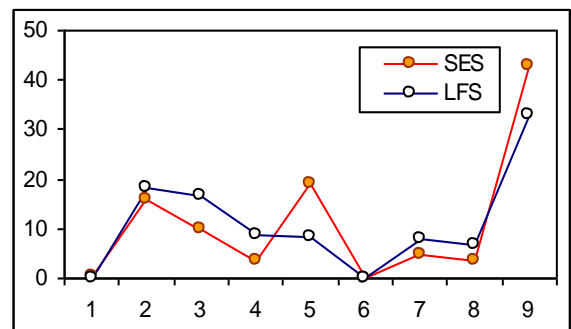
55 - 64



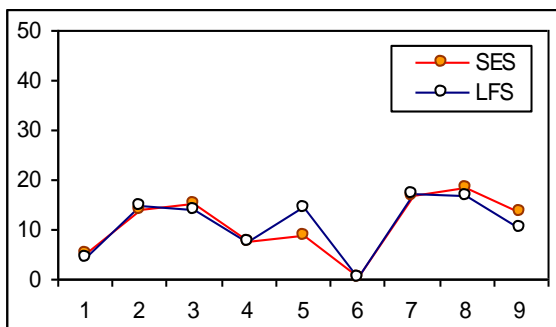
55 - 64



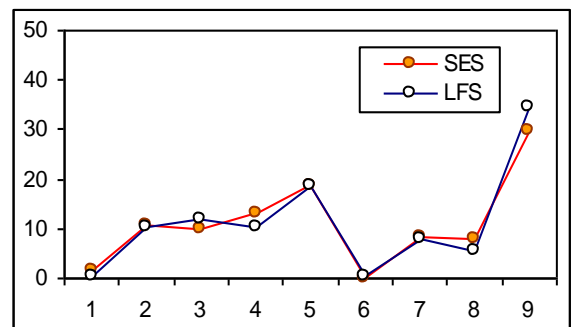
≥ 65



≥ 65



Total



Total

**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)
(Section to be completed)**

The required cross-tabulations are enclosed in the annexes as follows:

— **sex, age and economic activity (NACE Rev.1, at the section level),**

Annex 34 Coherence between SES and LFS (Sep.2002) absolute figures of full-time employees by NACE Rev.1 sections (C to O)

Annex 35 Coherence between SES and LFS (Sep.2002) absolute figures of part-time employees by NACE Rev.1 sections (C to O)

— **sex, age and educational level (ISCED 0 to 6),**

Annex 35 Coherence between SES and LFS (Sep.2002) absolute figures of full-time employees by education (ISCED'97)

Annex 36 Coherence between SES and LFS (Sep.2002) absolute figures of part-time employees by education (ISCED'97)

Note: ISCED 3 and ISCED 4 levels are not available separate in LFS so for purposes of comparison the SES data are aggregated.

— **sex, age and occupation (ISCO-88 at the 1- digit level).**

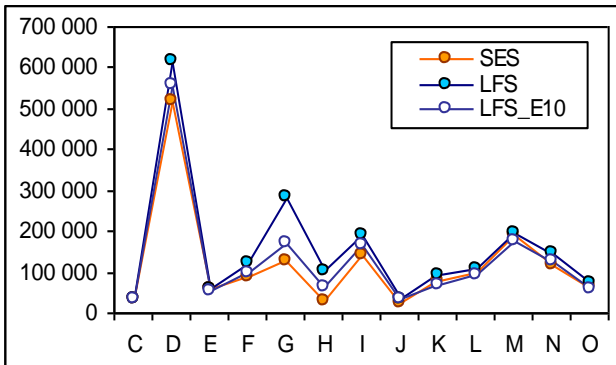
Annex 38 Coherence between SES and LFS (Sep.2002) absolute figures of full-time employees by occupation (ISCO-88)

Annex 39 Coherence between SES and LFS (Sep.2002) absolute figures of part-time employees by occupation (ISCO1)

The graphs below give an illustration on comparability of the employees' figures between SES and LFS. To receive more comparable picture a third set of data are presented **LFS_E10** which refer only to employees under labour contract in enterprises with 10 and more employees. It could be seen that in this way the populations of employees compared are much more similar and the differences observed are lower.

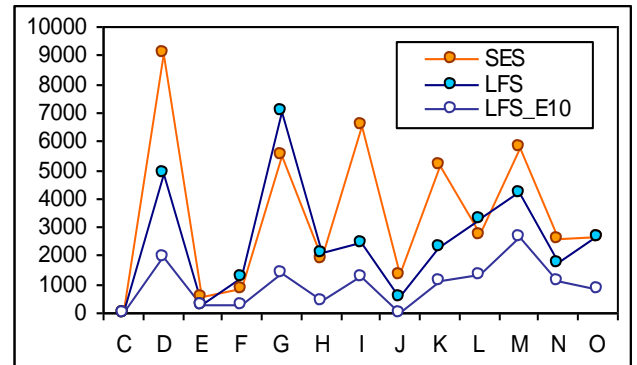
Coherence between SES, LFS (Sep.2002) and LFS_10 structure of employees by NACE Rev.1 sections (C to O): distribution in absolute figures

Full-time employees: males and females



NACE sections

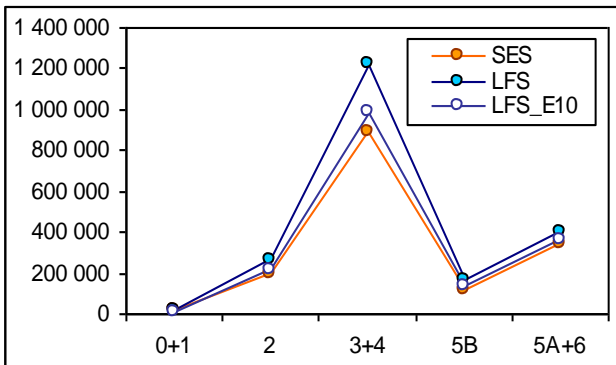
Part-time employees: males and females



NACE sections

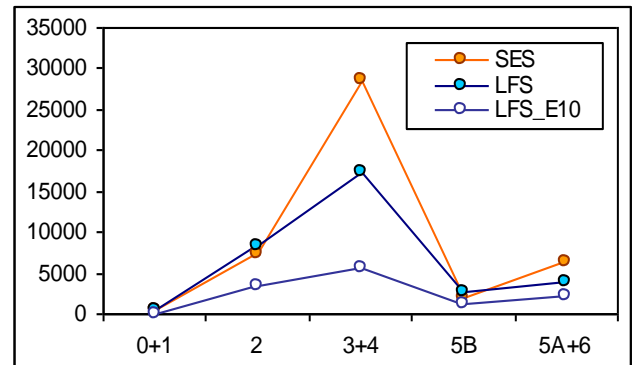
Coherence between SES, LFS (Sep.2002) and LFS_10 structure of employees by education (ISCED'97): distribution in absolute figures

Full-time employees: males and females



ISCED levels

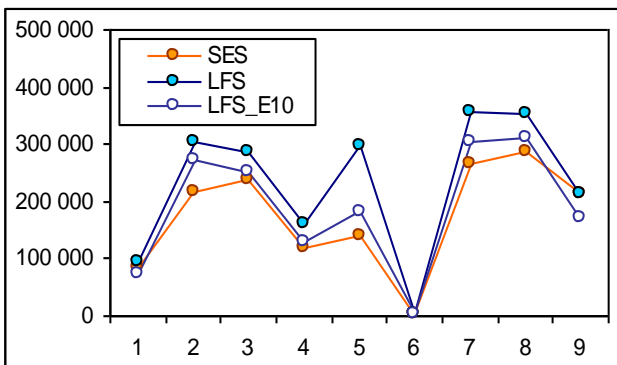
Part-time employees: males and females



ISCED levels

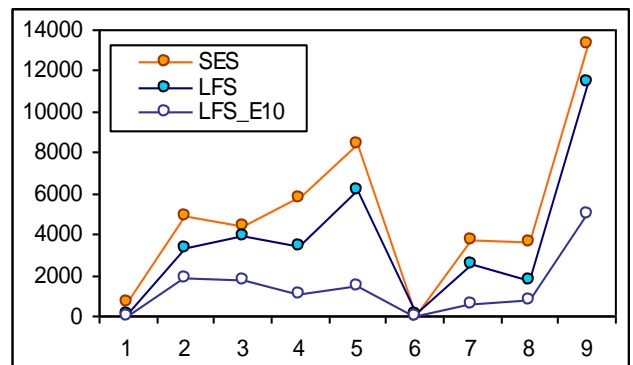
Coherence between SES, LFS (Sep.2002) and LFS_10 structure of employees by occupation (ISCO-88): distribution in percentages

Full-time employees: males and females



ISCO groups

Part-time employees: males and females



ISCO groups

Cross-tabulations concerning 'hours paid' are not provided as only week hours are collected in the Bulgarian questionnaire. The LFS experts in NSI consider that conversion of weekly to monthly hours would not be reliable for the 2002 LFS data.

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

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)

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

7. Completeness

All variables and breakdowns required in CR N1916/2000 are available in the SES 2002.

Contact person: Tanya Alexandrova
e-mail: talexandrova@nsi.bg
(Tanya.Aleksandrova@cec.eu.int)