Central Statistical Bureau of Latvia

The Report on Quality for Structural Statistics on Earnings

As required By Commission Regulation (EC) No 698/2006 of May 2006 implementing Council Regulation (EC) No 530/1999 as regards quality evaluation of structural statistics on labour costs and earnings

Prepared by
Mr Vladislavs Giruckis
Wages Statistics Section
Social Statistics Department
Central Statistical Bureau of Latvia
Lacplesa Street 1, Riga, LV-1301,
Republic of Latvia
Phone: +371 67366603
E-mail: vladislavs.giruckis@csb.gov.lv

Mrs Ieva Slosa
Mathematical Support Division
Central Statistical Bureau of Latvia
Lacplesa Street 1, Riga, LV-1301,
Republic of Latvia
Phone: +371 67366877
E-mail: ieva.slosa@csb.gov.lv

December 2008

QUALITY REPORT
Structure of Earnings Survey (SES) 2006

1. Relevance.

The main users of the Structure of Earnings Survey data in Latvia and abroad can be classified as follows:

Internal users:
Government institutions – the Ministry of Finance, the Ministry of Economy, the Ministry of Welfare, the Central Bank of Latvia, the Agency of State Social Security, State Employment Agency of Latvia;
Associations – the Latvian Chamber of Trade and Industry, the Latvian Confederation of Employers, trade unions;
Research institutions, education institutions;
Units of the Central Statistical Bureau of Latvia (CSB);
Mass media;
Lecturers, students, independent research institutions and researchers.

International users:
Eurostat, ILO, the International Monetary Fond;
Foreign embassies.

Central government institutions for own needs are using indicators on number of employees in breakdown by regions, occupations, age and educational groups, etc., as well as data on average wages and salaries in various breakdowns.

Trade unions, associations and unions are interested to obtain information on average monthly wages and salaries of selected occupations or occupational groups.

Mass media are using key indicators as average monthly wages and salaries in the country in selected sectors and regions.

User satisfaction has not been studied.

Several potential users may need data at degree of detailed elaboration, which SES 2006 survey may not provide.

Collection of statistical data “Results of the Structure of Earnings Survey 2006” was published on November 6, 2008. Collection of data was issued in 110 copies.
In 2009 selected tables will be published on CSB Web Page.

2. Accuracy

2.1. Sampling errors.

For standard error estimation software "Sudaan release 7.5.2" is used.

Standard error is estimated using Taylor Linearization method without replacement. The standard error is calculated according to the sampling design.

Coefficient of variation (CV) is value, of which the denominator is estimate of indicator and of which the numerator is standard error.

a) Coefficients of variation (CV) concerning total gross earnings in the reference month are broken down by:
Full-time (separately for men and women) and part-time employees;
NACE Rev.1 section;
Occupation (ISCO-88 at the 1-digit level);
Age band (under 20, 20-29, 30-39, 40-49, 50-59, 60 and over);
Level of education (ISCED 0 to 6);
Size band of the enterprise (1-9, 10-49, 50-250, 500-999, 1000+).

Full-time employees	Standard error (1000)	Estimate (1000)	CV (%)
Females	2985.8	123019.8	2.43
Males	2610.5	125065.6	2.09

Full-time and part-time employees	Standard error (1000)	Estimate (1000)	CV (%)
Total	4949.5	266157.4	1.86
Full-time employees	4821.4	248085.4	1.94
Part-times employees	600.9	18072.0	3.32

By NACE sections

NACE Rev.1 section	Standard error (1000)	Estimate (1000)	CV (%)
C	42.5	905.8	4.69
D	749.1	39707.6	1.89
E	526.3	5471.5	9.62
F	655.3	20378.8	3.22
G	1117.0	45080.5	2.48
H	439.7	5858.8	7.50
I	2880.5	24841.4	11.60
J	1821.7	12559.7	14.50
K	704.4	28509.8	2.47
L	1097.7	24796.4	4.43
M	2725.9	31595.8	8.63
N	973.7	15010.8	6.49
O	302.9	11440.7	2.65

By occupation

Occupation	Standard error (1000)	Estimate (1000)	CV (%)
1	1120.7	53999.2	2.08
2	2363.1	61899.2	3.82
3	1124.3	40509.2	2.78
4	643.7	17344.1	3.71
5	839.6	23215.8	3.62
6	107.7	624.9	17.23
7	868.4	28538.6	3.04
8	1180.6	22515.6	5.24
9	475.6	17510.9	2.72

By age

Age band	Standard error (1000)	Estimate (1000)	CV (%)
Under 20	93.5	2290.9	4.08
20-29	1199.7	52621.9	2.28
30-39	1386.7	67386.2	2.06
40-49	1327.1	68959.3	1.92
50-59	1275.6	53613.2	2.38
60 and over	680.3	21286.0	3.20

By level of education

Level of education	Standard error (1000)	Estimate (1000)	CV (%)
1	208.8	917.5	22.76
2	417.8	13008.4	3.21
3	2631.1	125449.0	2.10
4	496.1	13649.8	3.63
5	2857.0	109559.4	2.61
6	1060.2	3573.4	29.67

By size band of enterprise

Size band of the enterprise	Standard error (1000)	Estimate (1000)	CV (%)
1-9	548.0	22900.9	2.39
10-49	780.5	47569.2	1.64
50-249	5350.0	47285.8	11.31
250-499	1352.0	31969.0	4.23
500-999	1272.4	92673.9	1.37
1000+	1828.9	23758.5	7.70

b) Coefficients of variation (CV) concerning average gross hourly earnings in the reference month are broken down by:
   - Full-time (separately for men and women) and part-time employees;
   - NACE Rev.1 section;
   - Occupation (ISCO-88 at the 1-digit level);
   - Age band (under 20, 20-29, 30-39, 40-49, 50-59, 60 and over);
   - Level of education (ISCED 0 to 6);
   - Size band of the enterprise (1-9, 10-49, 50-250, 500-999, 1000+).

Full-time employees	Standard error (1000)	Estimate (1000)	CV (%)
Females	17.0	699.8	2.43
Males	14.5	705.5	2.06

Full-time and part-time employees	Standard error (1000)	Estimate (1000)	CV (%)
Total	29.4	1673.8	1.76
Full-time employees	27.0	1405.3	1.92
Part-times employees	9.2	268.5	3.43

By NACE sections

NACE Rev.1 section	Standard error (1000)	Estimate (1000)	CV (%)
C	0.2	5.5	4.41
D	4.7	239.2	1.98
E	3.0	31.8	9.29
F	4.0	124.1	3.23
G	6.8	281.4	2.41
H	2.6	37.4	7.01
I	15.6	148.2	10.51
J	10.6	73.9	14.35
K	4.3	185.8	2.30
L	6.7	157.6	4.28
M	16.7	210.3	7.92
N	6.4	96.6	6.59
O	5.3	82.0	6.43

By occupation

Occupation	Standard error (1000)	Estimate (1000)	CV (%)
1	7.2	343.9	2.11
2	14.7	411.0	3.57
3	7.0	256.9	2.73
4	3.9	108.2	3.60
5	4.7	141.7	3.31
6	0.6	3.8	16.53
7	5.0	169.4	2.96
8	6.2	129.5	4.75
9	3.0	109.5	2.72

By age

Age band	Standard error (1000)	Estimate (1000)	CV (%)
Under 20	0.7	15.7	4.35
20-29	7.3	325.2	2.24
30-39	8.2	421.8	1.95
40-49	8.2	435.4	1.88
50-59	7.5	335.3	2.23
60 and over	4.3	140.3	3.06

By level of education

Level of education	Standard error (1000)	Estimate (1000)	CV (%)
1	1.2	5.3	22.15
2	2.3	78.6	2.96
3	14.8	766.3	1.93
4	3.2	87.9	3.67
5	17.8	712.4	2.50
6	6.6	23.3	28.49

By size band of enterprise

Size band of the enterprise	Standard error (1000)	Estimate (1000)	CV (%)
1-9	4.1	177.2	2.33
10-49	6.9	308.7	2.23
50-249	31.3	283.9	11.02
250-499	8.6	191.8	4.47
500-999	7.6	571.7	1.34
1000+	10.9	140.4	7.77

The CSB Statistical Business Register holds information on all enterprises registered in the administrative Enterprise Register of Latvia. In the sampling frame only enterprises from the Business Register economically active in 2006 are included.

The sample for the Structure of earnings survey is formed as stratified two-stage random sample. Local units of enterprises are primary sampling units. Employees are secondary sampling units.

This first stage stratified random sampling was done in December 2006.

Sampling plan for the first stage

The information for employees in enterprises with fewer than 10 employees is included.

The micro data cover NACE rev.1.1 section L except armed forces of the Ministry of Defence. Since the Ministry of Defence has not reported the data on armed forces, the employees on military occupations are not included.

The local units of enterprises are stratified according to the 4 size groups of enterprises (1-9, 10-19, 20-49, 50 and over), classification into NACE code groups (2-digit level) and their location into 6 regions of Latvia.

Sampling plan for the second stage

At the second stage employees are chosen from the local units of enterprises selected at the first stage with the following sampling rate: if the number of employees in the local unit ranged from 1 to 50, information had to be rendered on all employees, in case of 51-69 employees – on 50 employees, in case of 70-149 – on 60, in case of 150-249 – on 70, in case of 250-499 – on 80, in case of 500-999 – on 90, in case of 1000 and over – on 8% of the total number of employees in the local unit as of October 31, 2006. If it was more convenient, local units of enterprises could give information on all employees.

But if local unit didn't give the information on all employees, the selection of employees was started with those born on the 15 day following the selection with the next date until the necessary number of employees was reached.

2.2. Non-sampling errors.

2.2.1. Coverage errors.

The sample of the survey was made at the end of year 2006. Time lag between last update of the sampling frame and the moment of the actual sampling was approximately one year. This is the main reason of coverage errors in the survey.

Estimated rate of over-coverage for the reference population is 10,5%.

Rate of under-coverage for the reference population is 13,5%. Weights were calculated using updated frame. Thus the influence of the coverage errors was reduced.

2.2.2. Measurement and processing errors.

In order to reduce the measurement errors questionnaire was supplemented with detailed instructions. In cooperation with the representative from the enterprise validation program for each questionnaire was created. In the validation program for each enterprise the data of the SES survey were compared with the variables from other surveys.

If it was not possible to correct wrong values, or if it was not possible to get the data for not filled-in variables, the questionnaire was considered as not delivered and the weights were recalculated.

One the steps to escape input of erroneous or incomplete information into the database was to develop a data entry program including a wide range of logical controls. Each questionnaire was validated according the validation program. The validation program consists of arithmetical and logical controls. These controls ensure all-round data verification:

- Arithmetical, logical mistakes are reported,
- Comparison with the data of regular survey on labour was done,
- Data input errors were revealed,
- Deviations from average indicators were indicated.

Every statistician has a definite number of enterprises from which the questionnaires have to be collected, entered and verified. If any answer on the questionnaire's variance was missed or (according to the rules implemented in the data entry programme) filled incorrectly, the responsible person could not finish the data entry of the questionnaire. In such situations the responsible person contacted the enterprise once more and made the necessary changes in the questionnaire.

Survey input was conducted both by Central office in Riga and regional offices.
The validation should be done in two ways:

- Individually for each questionnaire (after the data entry),
- For the entered questionnaires of the regional offices and the sections of Central office involved in data collecting.

All expressions not filed could be seen on the monitor or printed.

The variables that have been corrected most often are:

- Share of full-time or part-time normal hours (variable 2.7.1.) worked by employees in PT does not correspond to the share of number of normal hours worked by employee in the local unit in FT;
- Collective pay agreement (variable 1.5) does not correspond any of 5 categories of agreements;
- Wages and salaries in October 2006 included variables on whole year 2006;
- Code of educational attainment;
- Number of years worked in enterprise.

2.2.3. Non-response errors.

The response rate of local units of enterprises is 91%. The design weights are adjusted according non-response in each stratum.

If report from the respondent was not received on time, statistician called him/her and asked to fill in the report or found out the reason of non-response.

During the arrangement of the micro database the following additional procedures were done:

- If employee's earnings in October were calculated at a reduced rates (sick leave certificate A), the employees were excluded from the sample, and the weights were re-calculated so that they could reflect the exclusion of such employees from the sample;
- If during the year or October employees had worked less than one month, they were excluded from the sample, and the weights were re-calculated so that they could reflect the exclusion of such employees from the sample;
- If in October employee's earnings are affected by unpaid absence including the cases when employee joined the enterprise during October, in order to provide an estimate of the employee's earnings for full months, earnings and hours worked in October were adjusted.

2.2.4. Model assumption errors.

The design weights were calculated according to the sample design where all local units within the same stratum having equal design weights. The design weights were adjusted using the data of response level in each stratum. Final weights were calculated using information from updated frame; respondents were grossed up to the updated frame.

In the calculation of the sample weights the first step is to determine the inclusion probabilities π_i of the employees in the sample. The design weights d_i are inversely proportional to the inclusion probability

$$d_i = \frac{1}{\pi_i}.$$

In each stratum primary sampling units are selected using simple random sampling procedure. At the second stage the random sampling of employees is made within each selected local unit.

For each stratum i the inclusion probability of i-th employee in i-th primary sampling unit is

$$\pi_i = \frac{n_i}{N_i} \cdot \frac{k_i}{K_i},$$

n_i – the total number of respondents local units in i-th stratum in the first sampling stage;
N_i – the total number of local units in stratum i in the first sampling stage;
k_i – the number of employees in sample from selected local unit;
K_i – the total number of employees in local unit in the second sampling stage.

3. Punctuality and timeliness.

3.1. Punctuality.

In September 2006 the main statistical units of the survey were set and in October of the same year each respondent received a letter containing information on survey: survey period, term of report submission, report indicators and codes of educational attainment.

Design of the questionnaire was finished in December 2006, and with instructions was sent to respondents in January 2007.

Sample was drawn in December 2006. The last corrections of the weights were made in June 2008.

The deadline for delivering the questionnaires to the statistical office was March 15, 2007.

Part of respondents asked for more time because in March and April enterprises are balancing their accounts and it was allowed to deliver the questionnaires later. Rest of enterprises received reminders. At the same time statisticians also telephoned enterprises.

The validation of the micro data was finished by March 2008. The response rate of units at that time is 91%. The response rate of employees is 98%.

Afterwards the data were tested also on macro levels, and the transmission files for Eurostat were prepared in June 2008. Some corrections were done in July 2008.

Collection of statistical data "Results of the Structure of Earnings Survey 2006" was published on November 6, 2008.

3.2. Timeliness.

The length of time between the release of Collection of statistical data and the reference period in Latvia was 22 months.

4. Accessibility and clarity.

4.1. Accessibility.

The Collection of statistical data "Results of the Structure of Earnings Survey 2006" was sent to all users who ordered the collection of statistical data.

The reporting units can get the information from the bulletin, which is also available in CSB Information Centre.

In 2009 it has been envisaged to publish selected tables of the data collection on the Web Page of the CSB.

4.2. Clarity.

Metadata are published in the Collection of statistical data "Results of the Structure of Earnings Survey 2006".

5. Comparability.

5.1. Geographical comparability.

There are not differences between Latvian and European concepts.

5.2. Comparability over time.

In SES 2002 enterprises as sampling units instead of local units in Latvia were used.

In SES 2006 local units were used as sampling units and indicators (wages and salaries, number of employees) were calculated in breakdown by regions of Latvia.

6. Coherence.

Gross annual earning in the reference year (2006) per employee (code 4.1) constitutes LVL 3882.

Wages and salaries per employee of the NA constitute LVL 4496.

LVL

NACE sections	National Accounts	SES
	Wages and salaries per employee	Gross annual earnings per employee
C	3545	4409
D	3813	3388
E	5087	5141
F	5275	3581
G	3933	3198
H	3633	2392
I	4444	4065
J	7034	8427
K	5903	4266
L	5411	5410
M	3733	3984
N	4460	3882
O	4331	3443
Total	4496	3882

There are the following methodological disparities between “Gross annual earnings in the references year, expressed per employee” (SES) and “Wages and salaries per employee”(NA):

- The SES includes employees on the main job and secondary job, but the NA – only on the main job. NA includes self-employed persons, but SES - not;
- Calculating the wages in SES per employee the wages are adjusted on to a full-time basis. In National Accounts this procedure is not done;
- In National Accounts the adjustments for wages and salaries were made for non-covered employment and for underreported wages.