Czech Statistical Office
Na padesatem 81, 100 82 Praha 10, Czech Republic

Quality Report on the 2006 Structure of Earnings Survey

(English version, not edited for language)

Date: 19 December 2008

Contact person:
Mr. Dalibor Holý
Czech Statistical Office (CZSO)
Na padesatem 81
100 82 Praha 10
Czech Republic
Phone: +420 274052694
E-mail: dholy@gw.czso.cz

Preamble

The composition of the 2006 Structure of Earnings Survey (SES) Quality Report is based on the EC Regulation No. 698/2006 of May 2006. The quality of the statistical data as well as the whole course of the survey is assessed according to criteria relevant for users, processors and administrators of the survey (relevance, accuracy, punctuality and timeliness, accessibility and clarity, comparability and coherence).

The main statistical source for the SES data been the "Informacni system o prumernem vydelku" (ISPV, i.e. Information System on the Average Earnings). The course and development of the survey is directed by the Commission on the direction and coordination of works on the ISPV, which is composed by representatives of the Ministry of Labour and Social Affairs, the Ministry of Finance, the Czech Statistical Office, the Czech National Bank, CERGE-EI, High School of Economics in Prague, Czech-Moravian Confederation of Trade Unions, Union of Industry and Transportation. The main goal of ISPV is to show earnings levels broken down by individual occupations, other results include estimations on the earnings parameters of the population of employees in the Czech Republic.

The ISPV consists of two basic parts – surveys of business and non-business spheres of the national economy of the Czech Republic. The business sphere is composed of economical subjects (ES) – enterprises which remunerate their employees with wage according to the Act No 1/1992 Coll.; the non-business sphere is composed of economic subjects (ES) which remunerate their employees with salary according to the Act No 143/1992 Coll. Employees of the household sector and non-profit organizations are not covered by the survey.

The ISPV in the business sphere has been taking place as a quarterly survey in enterprises with 10 or more employees (ISPV-PS). In addition, an ad-hoc survey for ES with less than 10 employees was made in 2007 (micro-subjects).

In the non-business spheres, the source of data has been Information System on Salary (ISP), which covers all ESs of the sphere, i.e. exhaustive survey. ISP is half yearly.

Based on the data from both ISPV-PS and ISP, the Structure of Earnings Survey is annual survey in the Czech Republic. National SES does not cover micro-subjects. The following scheme shows the system of earnings surveys for 2006:
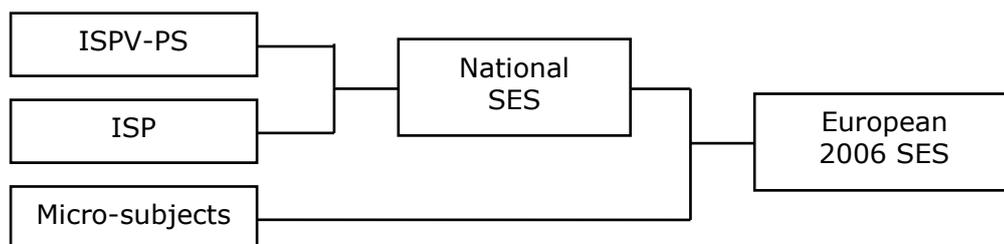
Figure 1 - System of structural earnings surveys in the CR

1 Relevance

Relevance denotes the degree to which statistics meet users' needs. International survey of 2006 SES cannot be judged separately, the results of the national SES are predominantly used in the CR annually, and also results of the ISPV. Therefore, this part refers mainly line to relevance of these surveys and their results for the domestic users.

The users of the structural earnings surveys results

Central government, trade unions and organizations of employers:

Ministry of Labour and Social Affairs, Labour Offices, Ministry of Finance, Czech Statistical Office, Czech National Bank, Czech-Moravian Confederation of Trade Unions, Union of Industry and Transportation. Their requirements and comments are formulated and assessed during meetings of the directing Commission.

Research institutions and universities

CERGE-EI, High School of Economics in Prague, High School of Mining, Aarhus School of Business. Their requirements on the width and quality of the data are formulated during joint research projects with the processing organization.

Labour Offices and Courts

Labour Offices are important users of both national ISPV and ISPV regional results. The Courts lords use the ISPV when solving some problems on labour laws.

Respondents

Respondents of the survey have possibility to get a free hard-copy of the publication of the ISPV-PS results. The contact with them strengthens a feedback with processing organization, their comments help expose drawback of the survey and consequently remove them.

Public

The public has free electronic access to survey results, the web pages are placed both on MLSA and CZSO domains.

2 Accuracy

Generally, accuracy denotes the closeness of computations or estimates to the exact or true values of the variables under consideration. The differences between estimations and true values are called errors, errors caused by the sampling are sampling errors and other errors are called non-sampling errors.

A description of the sampling scheme and the sampling errors is in the first part of this chapter and the second part is devoted to the description of the non-sampling errors: coverage errors, measurement and processing errors, non-response errors and model assumption errors.

2.1. Sampling errors

The 2006 SES stem from three sources: the survey of the business sphere (ISPV-PS), the survey of the non-business sphere (ISP) and the survey of micro-subjects. The data base had been cut before computation with regard to some industries to meet Eurostat requirements. The following description regards the data adapted in this way.

2.1.1. The survey of business sphere

ISPS-PS is a regular quarterly sample statistical survey. The sample of employees can be characterised as a two-stage probability sampling without repetition. In the first stage, enterprises are found using stratification sampling, all employees inside these enterprises are sampled in the second stage. Population frame for the first stage of the sampling is made using the Business Register of the CZSO and the RARIS register which allows to decide whether the enterprise belongs to business or non-business sphere. The population is created by active enterprises with 10 or more employees; it is stratified by:

- Industry
- Size group
- Region.

Industrial strata are set according to the similarity in the labour nature and structure of occupations. There are four size groups: 10-49; 50-249; 250-999; 1000+. The region is defined by Constitution Act no 347/1997,Coll.

The 2006 frame population incorporated 39,218 enterprises, 2,789 of them have been in the sample and sent data worth of processing. The following table show a summary on the sample stratification with the numbers of enterprises in the population and the probability of the sampling.

Industry	Size group			
	10-49	50-249	250-999	1000 +
Non financial enterprises	29,763 4.5 %	7,634 15.0 %	1,409 80.0 %	248 100.0 %
Financial institutions	0 0.0 %	84 100.0 %	23 100.0 %	10 100.0 %
Governmental institutions	8 100.0 %	9 100.0 %	14 100.0 %	16 100.0 %

Table 1 Frame population of the ISPV-PS and the sampling rates

The specific number of enterprises defined in advance in the individual cells of the sampling scheme is selected from the frame population. The enterprises with 1000 or more employees are selected with probability equal 1, it means exhaustively; in stratum 250-999, the enterprise are gathered to survey step-by-step in order to come to exhaustive survey starting from 2007. For other strata, the enterprises are selected by means of systematic random sampling method with non-equal probabilities. The sampling rate for stratum 50-249 was 15% and for stratum 10-49 it was 4.5 %. This part of the sample file is updated by the method of "rotating panel" with a period of nine years, i.e. all enterprises 10-249 employees will be replaced after 9 years.

2.1.2. The survey of the non-business sphere

ISP is a regular half-yearly exhaustive statistical survey. ESs bound to respond by law have been sending the information on all their employees. Consequently, there are no sampling errors in the ISP. The frame population for the 2006 incorporated 18,593 ESs from which 14,444 send data worth of processing.

2.1.3. The survey of micro-subjects

An Ad hoc survey on the enterprises (ESs) with less than ten employees was carried out in 2007. The sample of employees can be characterised as a two-stage probability sampling without revolving. In the first stage, enterprises are found using stratification sampling, while all employees inside these enterprises are sampled in the second stage. The frame population was stratified by:

- Region
- Size class
- Industry

Two size classes have been used, see following table. There have also been fourteen industrial groupings from the business sphere, similarly as in the ISPV-PS.

The frame population for the 2006 incorporated 73,881 ESs from which 826 ESs send data worth of processing. The following table show a summary on the cumulative sampling strata with the numbers of ESs in the frame population and sampling rates:

Size class	
1-5	**6-9**
59,876	14,005
1.3 %	3.1 %

Table 2 The frame population and sampling rates

2.1.4. The calculation of sampling errors

The sampling error is expressed as a coefficient of variation of the mean. The CV is a division of the standard deviation of the mean (numerator) and the mean (denominator). The standard deviation have been, in some problematic cases, calculated by means of non-parametric method of Jackknife. The weighted data have been defined for the calculation adequately as a two-stage stratified probability sample with limited number of elements in every stage.

In contrast with the 2002 situation, the sampling errors can be calculated now from the following reasons:

- There is only approx. 10 % of ESs in the sample selected by quota sampling method
- Gathering of all ESs in the 250+ stratum is almost completed
- All ESs from the non-business sphere have been covered exhaustively as well as all financial institutions with 50+ employees.

A summary on the CVs is attached in Tables 3 – 9. More detailed information on CVs, standard deviations, means and degrees of freedom is in the Annex.

Age group	Gross monthly earnings	Gross hourly earnings
< 20	0.036	0.026
20-29	0.009	0.009
30-39	0.012	0.012
40-49	0.009	0.009
50-59	0.008	0.008
60+	0.016	0.012

Table 3 – CVs by age

Occupation (ISCO-88)	Gross monthly earnings	Gross hourly earnings
0	0.000	0.000
1	0.052	0.051
2	0.013	0.012
3	0.011	0.011
4	0.011	0.011
5	0.011	0.011
6	0.171	0.198
7	0.008	0.008
8	0.007	0.007
9	0.010	0.009

Table 4 – CVs by occupation

Education (ISCED)	Gross monthly earnings	Gross hourly earnings
1	0.042	0.036
2	0.015	0.015
3	0.005	0.005
4	0.025	0.023
5	0.013	0.012
6	0.110	0.109

Table 5 – CVs by education

Size class	Gross monthly earnings	Gross hourly earnings
1-9	0.034	0.033
10-49	0.021	0.021
50-249	0.017	0.017
250-499	0.010	0.010
500-999	0.009	0.009
1 000+	0.003	0.003

Table 6 - CVs by size class

NACE section (NACE rev. 1.1)	Gross monthly earnings	Gross hourly earnings
C	0.015	0.020
D	0.010	0.010
E	0.022	0.023
F	0.015	0.015
G	0.029	0.028
H	0.031	0.028
I	0.021	0.021
J	0.032	0.031
K	0.033	0.032
L	0.003	0.003
M	0.011	0.010
N	0.005	0.004
O	0.018	0.018

Table 7 – CVs by NACE section

FT/PT – sex	Gross monthly earnings	Gross hourly earnings
FT – women	0.006	0.006
FT - men	0.009	0.009
PT	0.021	0.018

Table 8 - CVs by type of contract and sex

Total	Gross monthly earnings	Gross hourly earnings
Czech republic	0.007	0.007

Table 9 - Overall CV

The following table shows the least efficient estimations with rates of 95% confidence interval (CI) to their mean.

	95% CI (to mean)
Occupation ISCO-88: 6	0.676
Education ISCED: 06	0.430
Occupation ISCO-88: 1	0.206
Education ISCED: 01	0.166
Age group: <20	0.142
Size class: 1-9	0.132
NACE section: K	0.129
NACE section: J	0.126
NACE section: H	0.120
NACE section: G	0.113

Table 10 – Efficiency of estimations

The least efficiency of the estimation in the ISCO major group 6, i.e. agriculture workers, is caused by the fact that the whole industry NACE A, B (Agriculture etc.) is not covered in the SES, so the number of these workers is minimal. Other problematic category – ISCED 06 (higher education than university) – refers to the overall high variability of earnings of these employees even in comparison with other categories with similar number of employees. Also category of ISCO 1 (managers) always has very high variability.

2.2. Non-sampling errors

2.2.1 Coverage errors

The differences between target population required by EC Regulation 698/2006 and the reference population referring to the ESs from the CZSO Business register in 2006 have the following reasons:

- Employees of the household sector and of non-profit organizations are not covered in the Czech SES.

- Employees with non-standard labour contracts (temporary jobs) are not covered

There is no assessment of the under-coverage and over-coverage as there is no recent quality report on the Business Register available; there are some problems with the units size.

The dominant source of SES is the ISPV which is permanent survey and its sample file is defined and altering gradually. Because of this fact, some under-coverage errors occur due to the situation when the sampling plan for a specific part of the sampling file is created before the grossing-up plan. An assessment of these errors for individual sources is as follows:

Coverage errors in the ISPV-PS

Sampling of ISPV-PS goes in two stages – updating of the exhaustive part of the sampling file (every year) and progressive changes by the method of rotating panel. Individual regions of the sampling part are updated once in three years. In addition, we must take into account the time needed for the creation of the sampling plan (approx. 2 months) and the time needed for incorporating the new ES into the survey (average time is approx. 6 months). As the survey is carried out just one month after the end of reference period for practical reasons, the new ESs cannot be incorporated at once to the sampling plan.

The following table shows coverage errors that happened because of the sampling plan a technical obstacles, by size classes:

Coverage errors	Size class				
	Total	10-49	50-249	250-999	1 000 +
ESs covered	**33,198**	**24,388**	**7,343**	**1,205**	**262**
ESs covered [%]	**84.7%**	**81.9%**	**95.0%**	**83.5%**	**96.7%**
Under-coverage	**6,017**	**5,383**	**387**	**238**	**9**
Under-coverage [%]	**15.3%**	**18.1%**	**5.0%**	**16.5%**	**3.3%**
of which:					
- born after sampling	1,511	1,352	158	0	1
- under threshold of 10 employees	3,176	3,073	103	0	0
- out of sampling plan	97	70	27	0	0
- ESs under recruitment	162	3	3	148	8
- sampling frame did not catch them	67	58	9	0	0
- sampling frame showed them	5	4	1	0	0
- size unidentified	903	822	81	0	0
- erased, non recruited (exhaustive	96	1	5	90	0
Frame population – total	**39,215**	**29,771**	**7,730**	**1,443**	**271**

Table 11 Coverage errors of ISPV-PS

Based on the high under-coverage because of sampling course in 2002 SES (32.2 %), a new course was adopted in order to minimize the delay at the end of 2006. Consequently, the coverage error in 2006 was cut down by a half – 15.3 %.

Coverage errors in the ISP

The coverage errors have been eliminated by the system of data collection. The "submissioners" have been sending data for themselves as well as for the organizations under their direction (respondents). The submissioners are ministries, regional governments, state funds etc. They have to make a registration to enter the ISP.

The purpose of the registration is to get information on the recent valid contact data, especially the electronic address. In the new campaign, only those submissioners make a registration whose contact data have changed. The respondents do not make a registration, their data are entered by the submissioner.

Coverage errors in the micro-subject survey

Coverage errors in this survey happened because of delay between grossing-up and sampling frame, as the sampling frame as of November 2006 was used.

Coverage errors	Size class		
	Total	1-5	6-9
ESs covered	**67,554**	**54,768**	**12,786**
ESs covered [%]	**92.1%**	**91.6%**	**93.8%**
Under-coverage	**5,834**	**4,995**	**839**
Under-coverage [%]	**7.9%**	**8.4%**	**6.2%**
of which:			
- born after sampling	26	24	2
- over 10 employees in the sampling frame	127	8	119
- sampling frame did not catch them	190	183	7
- size unidentified	5,491	4,780	711
Frame population – total	**73,388**	**59,763**	**13,625**

Table 12 – Coverage errors in micro-subjects

2.2.2. Measurement and processing errors

The data of ISPV-PS are collected from respondents in the form of databases; they include information on employer and employees. In most cases, data stem directly from inter-enterprise information system. It means that Wage and Personal System of the respondent is direct source of data in 90 % cases. About nine out of ten respondents use electronic mail for sending data, less than 9 % use diskettes. Less than 1 % of ESs sent data on paper questionnaires. For data checking and entering respectively, there is "Acquisition software ISPV 2006". It is available on the website of the processing firm www.trexima.cz in the section "Statistics of earnings". There is direct communication with the respondent via phone or e-mail. It could serve for revision of mistakes.

In the ISP, data are send via secured internet page of the Ministry of Finance, where basic check mechanisms are embedded. The data gathered this way are then send to the processor. More difficult communication is in the ISP, because the data are send through a submissioner. On the other hand, this system minimizes coverage errors. All the data are send electronically.

In the survey on micro-subjects, data were send electronically in 25 % cases and almost three fourths of respondents send data on paper questionnaires. The risk of wrong data was the biggest there. Revision was made by the processing firm during phone consultation with the respondent.

All the data files (ISPV-PS, ISP and micro-subjects) have been preserved in the seven data-stores. In the individual data-stores, there are saved original data from the respondents, data before the data-checks and data after the data-checks (both automatic and manual). Grossing-up to the whole population is made by applying weights for individual records on the final database file.

Triple automatic check is made during the data collecting and processing. In addition, a visual check is made after that. Any mistakes found are dealt with in relation to their importance – either by contacting the respondent or directly by the processing company. Some help is obvious with coding of ISCO-88 professions since this task is the most difficult for the respondent; consultations by telephone are provided. After data entering, additional checks are made on the levels of regions and individual professions. Some checks are also made accordingly on the level of ESs. On the aggregated level, we search for changes in time and look for explanations. An example can be earnings level in the individual occupation in region – in case on change more than 20 % y-o-y the enterprise data are analysed.

The very last stage is data check before their sending to Eurostat. 1,970,864 records passed the checks, 54,547 have been erased. The most records erased didn't pass the check on the number of hours paid in month - 27,326; 10,884 records failed in the check on the overtime hours and 4,490 on holidays.

2.2.3. Non-response errors

In the first part of this chapter, reasons for non-response of whole enterprises are described together with procedures for minimising the errors. The item non-response is described in the second part.

Response rate in the ISPV-PS

Sample file of the ISPV-PS is composed of two parts:

- ESs from the kernel sample file. The kernel file is key part of the sample file, ESs inside represent the target population using grossing-up weights. ESs outside have the weight equal one, i.e. they represent only themselves.

- Other ESs, which are not inside kernel sample file, e.g. newly recruited groups of enterprises.

Non-response rate in 2006 was **17.8 %**. The comparison with 2002 SES is difficult due to different circumstances of the surveys – here was dominant part of quota sample in 2002 whereas the quota part of 2006 sample was approx. 10 %.

Sample file - response category (ES)		Size class				
		Total	**10-49**	**50-249**	**250-999**	**1 000 +**
Kernel sample file	**Data processed**	**2,285**	**485**	**738**	**833**	**229**
	Data not processed	**494**	**236**	**195**	**51**	**12**
	Of which:					
	- recruitment	71	34	29	7	1
	- wrong data	124	48	62	9	5
	- data not sent	58	34	18	5	1
	- direct refusal	7	3	2	1	1
	- indirect refusal	14	8	3	3	0
	- no reaction	97	47	32	15	3
	- technical reasons	63	32	23	8	0
	- temporarily off employees	45	23	20	1	1
	- other reasons	15	7	6	2	0
	Non-response rate	**17.8%**	**32.7%**	**20.9%**	**5.8%**	**5.0%**
Other subjects – data processed		**504**	**91**	**141**	**259**	**13**

Table 13 Response rate in the ISPV-PS

Response rate in the ISP

It is important to know the exact frame population of the survey, which comes to existence during the registration and data collecting itself (see chapter Coverage errors).

Statistical file - Response category (ES)	Size class					
	Total	1-9	10-49	50-249	250-999	1 000 +
Data processed	14,444	5,738	6,067	2,338	245	56
Data not processed	4,149	3,434	573	132	10	0
Of which:						
- wrong data	357	300	49	8	0	0
- data not sent	3,792	3,134	524	124	10	0
Non-response rate	22.3%	37.4%	8.6%	5.3 %	3.9%	0.0 %

Table 14 Response rate in the ISP

Response rate in the micro-subject survey

Response rates in this ad hoc survey shows following table:

Statistical file - Response category (ES)	Size class		
	Total	1-5	6-9
Data processed	826	484	342
Data not processed	397	244	94
Of which:			
- refusal	4	1	3
- no response	163	114	49
- address not found	36	30	6
- no withdrawal	13	10	3
- no employees or dead	55	53	2
- after deadline	58	32	26
- wrong data	68	48	20
Non-response rate	32.5 %	37.3 %	24.2 %

Table 15 Response rate in the micro-subject survey

Item imputation rate

Items from Table A marked as A12, A13 and A14 were added from administrative sources.

As for item A15 – type of collective pay agreement – the information was sent by respondent and then matched with administrative source on collective agreements and from the Information system on working conditions (both MLSA data). If the true value was unknown, we supposed that no collective pay agreement existed ("unknown" is not possible in 2006 SES). 14,462 items have been imputed this way, predominantly from the non-business sphere, because ISP does not cover collective bargaining.

The following table shows numbers of imputed values and rates of imputation for items of Table B.

item	Number of imputed values	Rate of imputation [%]
B21 - sex	44	0.0
B22 – age	47	0.0
B23 – occupation	88	0.0
B25 – education	55,476	2.8
B26 – tenure	24,649	1.3
B28 – type of contract	345	0.0
Total - imputed	**80,117**	**4.1**

Table 16 Rates of item imputation in the employees variables

Items B21 and B22 have been imputed only in the micro-subjects survey, the Trade Register was used where statutory heads of the enterprises are filed. If not found, most frequent values in the individual occupations were imputed. For B23 item, it refers to 2-digit level of occupation according to the ISCO-88 in cases where it cannot be done on 3-digit level. B25 item was imputed by probability according to the occupation and age; B26 was supposed as "0" if value was not sent; and " unlimited" value was imputed for the B28 item if wrong code. There was no need to impute other variables. As the extent of the imputations was very small, it has only marginal effect on the accuracy of estimations.

Overall imputation rate

Overall number of imputed variables and overall rate of imputation is lower than the sum of individual cases, because in some cases more than one imputation was made for one employee/record. The overall imputation rate is 4.1 %.

2.2.4. Model assumption errors

Grossing up of non-response

ES non-response was healed by the method of correcting weights by stratification category. Algorithm took a primary weight in the first step (equal 1/x where x is sample rate) and it corrected the weight in the second step by the non-response rate taking into account the number of ESs in the grossing up strata.

Grossing up of new-born ESs

As a result of the longitudinal character of the ISPV, the most recent reference population cannot be covered in any way. The grossing-up procedures are made in a way that ESs that cannot be incorporated (new-born enterprises) are represented by subjects incorporated. If there are differences between characteristics of the new ESs and the old ones, the distortion in the estimation would happen. Comparing 2006 SES with the 2002 survey, there was a large decrease of this kind of under-coverage error.

3 Timeliness and punctuality

3.1. Timeliness

The results of the Structure of Earnings Survey have been processed and published at the national level separately for individual surveys: ISPV-PS, ISP and micro-subjects survey. National SES have been created upon joint database of ISPV-PS and ISP. The most important time limits are summarized in the following table:

Terms	ISPV-PS	ISP	Micro-subjects
Term for sending data	31.1.2007	8.2.2007	15.5.2007
Deadline for data delivery	28.2.2007	14.2.2007	3.8.2007
Preliminary results	13.3.2007	13.3.2007	7.9.2007
Definitive results	16.3.2007	16.3.2007	11.12.2007
SES National version	June 2007		-

Table 17 – Time limits

3.2. Punctuality

Based on the comments from the directing Commission, the ISPV quarterly results have been published. They are available for public with 2 and half months delay after the end of reference period.

As for the annual national SES, the tables have been available since the end of June 2007, i.e. 6 months delay after the end of reference year, and analyses since September 2007, i.e. 9 months.

The results of micro-subjects survey have been made available together with an analysis of the influence on the overall earnings levels since the directing Commission meeting in December 2007.

The EU SES data have been send to Eurostat in June 2008, i.e. the deadline has been met.

4 Accessibility and clarity

4.1. Accessibility

Publications of the ISPV-PS and ISP results in the Excel or PDF form are available on the MLSA website (www.mpsv.cz) in section „Income and Living Conditions". The extent of the list of individual occupations is set by the publication criterion – only those occupations are published which have at least 1 % representation in their respective ISCO Major Groups.

A complete summarization of the occupations under survey is available at www.trexima.cz/vysledky. Respondents had the possibility to get paper results by post.

National SES results for 2006 are available on the website of the Czech Statistical Office: http://www.czso.cz/csu/2007edicniplan.nsf/p/3109-07, in-depth analysis were published only in Czech language, available since September 2007.

Some results are also included in the CZSO publications: "Statistical Yearbook 2003", "Labour Statistics: Time series of Basic Indicators" and "Focused on Women, on Men".

4.2. Clarity

The development of the ISPV methodology is gathered in the internal materials of the directing Commission. Respondents have been obtained with paper Instruction manual how to prepare entering data before the data collection. It is available in electronic form too, on the website of the processing firm: www.trexima.cz, in section Earnings statistics. At the same place, there is available also Quality Report of the ISPV-PS, which has been produced following the European recommendations since 2003. Both publications on ISPV-PS and ISP include Technical annex, with a basic description of the survey and statistical procedures.

The results of the national SES have been presented during a special Press Conference of the Czech Statistical Office every year. The MLSA have published a Press Report on the ISPV-PS and ISP results via internet.

5 Comparability

5.1. Geographical comparability
There is one specific detail in the Czech SES – different concept of representative month. There is no calendar month earnings under survey in the CR, but the average month. The representative month earnings is calculated as:

(Annual Earnings – Irregular Bonuses) / Number of paid month of the employee

The reasons for this procedure is as follows:

- Eurostat requirement is to exclude the influence of irregular bonuses to the monthly earnings
- ISPV collects quarterly data on earnings of employees cumulatively from the beginning of the year to the end of respective quarter. Earnings for the "ideal" month calculated using whole year figures are more precise than any calendar month and are quite comparable with results of other EU countries.
- The whole year earnings (without irregular bonuses) is divided by number of paid months, it means only the time, when the employee had the earnings. By that, gross earnings are not distorted by any unpaid absence, especially illness.

Considering the comparability among EU countries, we must take into account the fact, that our average month has 20.92 working days (whole year presence without unpaid absences), while October 2006 had 22 working days in the Czech Republic.

Other differences of the CZ-SES concept from EU concept:

- Hourly earnings is calculated in analogical way from the monthly earnings as we gather numbers on the hours worked as well as hours paid together with hours of absences.

- Employees of ESA Household sector, i.e. employees of unincorporated self-employed persons (entrepreneurs), are not covered. That sector should be included in the survey for 2010.

- Also employees of ESA sector of non-profit organizations are not covered.

- Only employees with standard employment contract are covered, excluded are workers with temporary work contract, students and apprentices.

- There is no identification of persons; thus, there is no way how to identify whether one employee has only one employment contract with his/her employer or more. One employment contract is one record in the database. In addition, other jobs of the employee in other ES cannot be identified, but whether part-time or full-time, so number of employees with part-time contract in their main job cannot be calculated. Only the share of part-time contracts in all jobs can be assessed.

- Not only employees with earnings for the specific representative month are included in the database, but all employees with one paid month at minimum whenever during the year. This could be taken into account for the calculation of results on the representative month or hourly earnings. Number of paid weeks should be seen (item B31).

- All variable on earnings are collected in the national currency Koruna ceska (CZK), respectively in the units derived from it (CZK/hour; CZK/month; CZK/year).

5.2. Comparability over time

The time comparability of 2006 and 2002 SES are effected by following changes:

- changes of the definition of reference population
- change of data sources and of sampling methods
- changes of the definition of variables and of calculation methodology

5.2.1. Changes of the definition of reference population

- Reference population has been extended to include the employees of ESs with less

than 10 employees. The ad-hoc survey of micro-subjects has been carried out in the business sphere in 2007. In the new ISP (on non-business sphere) are ESs with less than 10 employees instantly covered.

- Non-profit organizations are not covered in 2006 SES. Since 2003, the definition of the framework population of ISPV has been changed in this manner.

- Only employees with scheduled working time of 30 hours or more weekly are covered in 2006 SES.

5.2.2. Changes of the data sources and of sampling methods

- Data source for the business sphere has been altered. In 2002 SES, employees from both business and non-business spheres were covered by one sample survey. In 2004, the separate survey of ISP on non-business sphere has been introduced, and consequently, both spheres are collected separately. The sampling error has remarkably decreased due to this change, but there are new problems with data quality.

- The exhaustive part of ISPV-PS has been extended. In 2002 SES, only ESs with 1000 or more employees were covered exhaustively; in 2006 SES, the threshold is 250 employees. This extension has consequence in sampling error reduction.

- The exhaustive part of ISPV-PS was also extended to all financial institution with 50 or more employees. The reason was huge variability in this sector. This extension results in sampling error reduction as well.

- A rotating panel has been implemented in the ISPV-PS since 2003 together with the first modification in the sampled part. Individual annual samples are made by probability stratified sampling methods. In the end of 2006, the original quota sample has represented only 10 % of the whole survey sample.

5.2.3. Changes of the definition of variables and of calculation methodology

There were changes in the structure of the variables in the ISPV-PS in 2006 because of the 2006 SES requirement. New items (date of ending of employment contract, structure of working hours) help make the calculation of gross earnings for the representative month more precise:

- Regular and irregular bonuses can be distinguished in the input data, it is an improvement in 2006 SES in comparison with 2002 SES.

- The percentage of part-time employment has been directly gathered item in 2006 SES, while in 2002 SES it was estimated by number of weekly hours usually worked in the end of reference period (it could change during the year).

- Absence hours have been gathered in hours in 2006, while in 2002 they were converted from whole days to hours using number of weekly hours usually worked.

All the facts in this chapter should be taken into account in comparison of 2006 SES with 2002 SES.

6   Coherence

Coherence of statistics is the extent to which they can be reliably combined in different ways and for various uses. It is, however, generally easier to identify cases of incoherence than to prove coherence.

Coherence should explain similarities and/or differences between SES and other earnings surveys. In the Czech Republic, it means ISPV and statistics based on enterprise report or SES and LFS.

6.1. Coherence with the Labour Force Survey

The most important methodological inconsistency between ISPV and LFS is the different population coverage. ISPV covers only enterprises with 10 or more employees in the business sphere, while the sectors of households and non-profit organizations are not covered – it means about 10 % of population of employees. On the other hand, LFS covers only individual households not collective ones (hostels etc.). LFS does not distinguish the sizes of enterprises, it covers whole national economy (this approach).

Different methods for coding of industrial and occupational classifications (NACE and ISCO) are other source of inconsistency. In ISPV, NACE codes are gathered straight from the Business Register. In LFS, the NACE codes are coded in the field subjectively by respondent's or interviewer's opinion.

As regards individuals, they are coded by HR specialists inside enterprises for the ISCO in ISPV, whereas in LFS they describe their occupation themselves and are coded with assistence of an interviewer.

6.2. Coherence with Enterprise Report

Enterprise Report is a quarterly survey on the number of employees and earnings sums (among others) for economic subjects:

- with 20 or more employees – sectors of non-financial enterprises and households (i.e. incl. employees of unincorporated private entrepreneurs)

- irrespective of the number of employees, i.e. all subjects, of sectors of financial enterprises, governmental and non-profit organizations (incl. all organizations funded from the state or local budget)

Enterprise Report makes a qualified (mathematical) estimation on ESs with less than 20 employees (not covered with survey) based on the known data from annual surveys and administrative data from social security registers.

The inconsistencies between ISPV and ER are caused by different target population, different calculation of gross monthly earnings (e.g. all unpaid absences are removed in ISPV) as well as different methods of grossing up and weighting. These inconsistencies have been regularly analysed and quantified.

The differences among indicators of single labour market surveys are not shortages of SES or ISPV or Enterprise Report or LFS; they stem from various targets of these surveys. They are a good motive for working towards coordination on the national level.

6.3. Coherence with National Accounts

All the source surveys for 2006 Structure of Earnings Survey were not made with the aim to assess overall number of employees nor to show wages and salaries sum, as is necessary for macro-economic aggregates in the National Account. The intention of all these surveys is to describe differences among various groups of employees as regards levels of their earnings as well as inner distribution.

In order to show the coherence between numbers of employees and gross earnings from 2006 ISPV (both business and non-business sphere) and the 'wages and salaries' of the NA, we made the comparisons for relative figures; the NA figures have been recalculated to meet that approach. The results are portrayed in the pictures below.

The national Structure of Earnings Survey (ISPV) has not been a source for the National Accounts in any way. In the NA, data on the number of employees and on compensation of employees incl. wages and salaries have been based on annual and quarterly Enterprise Report, annual Labour Cost Survey and quarterly LFS (for the former). In addition, the CZSO National Accounts Department make their own estimates on allocation of these variable into NACE sections and sectors using administrative data etc.

There are some differences visible in both graphs that are quite understandable and are obvious consequences of different methodological concepts. In the SES results, enterprises with less than 10 employees are not covered, whereas the non-business sphere is covered exhaustively; also, sectors S13 and S15 are not covered at all. On the other hand, the NA concept covers whole national economy including "shadow economy" etc.
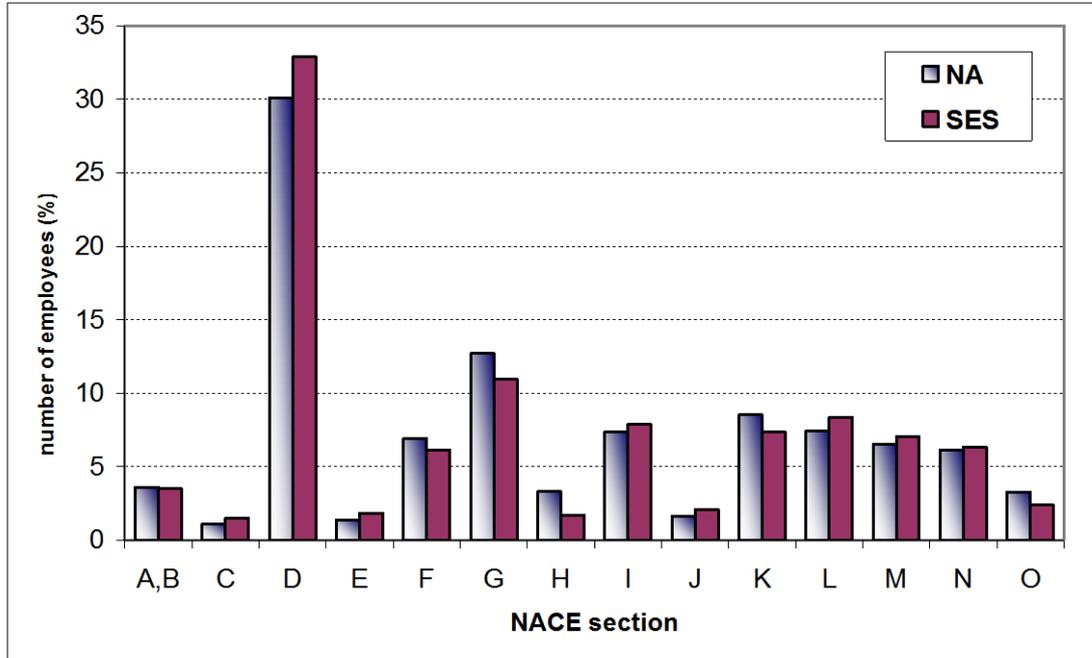
Graph 1 Coherence of NA and ISPV, number of employees, year 2006
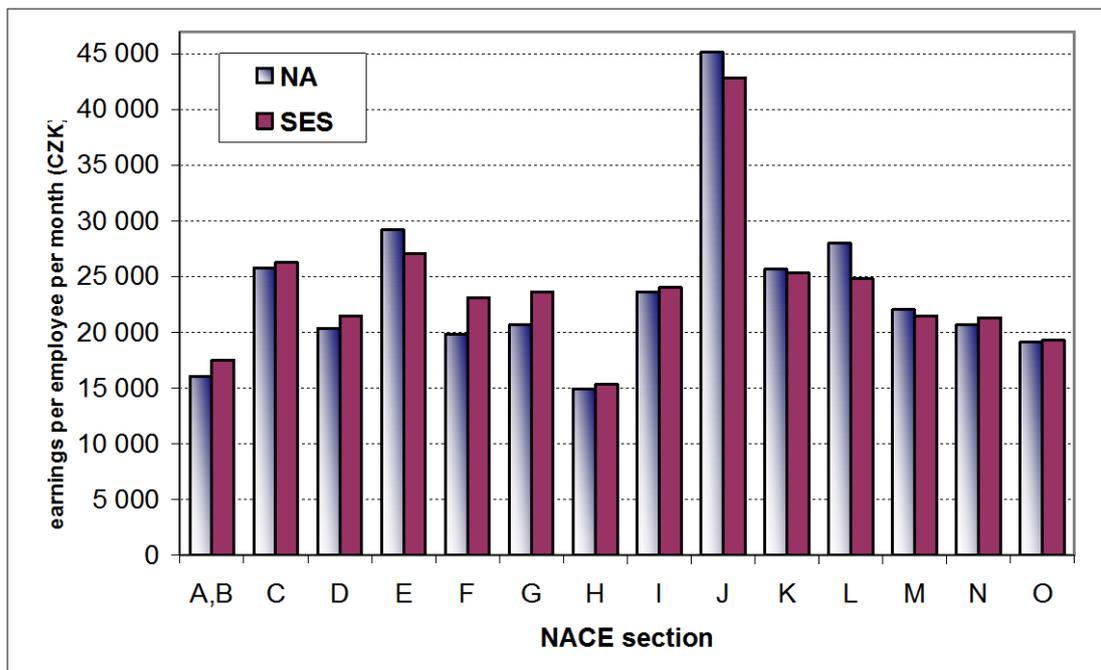
Graph 2 Coherence of NA and ISPV, earnings per employee per month, year 2006