Harmonisation of the Average Earnings Information System (MoLSA) with the Wage Statistics (CZSO)

Libor Malenovský[1] | *TREXIMA, spol. s r.o., Zlín, Czech Republic*
Kateřina Duspivová[2] | *University of Economics, Prague, Czech Republic*

Abstract

In 2011, the methodology of the Average Earnings Information System (ISPV) was harmonized with the methodology applied in the wage statistics of the Czech Statistical Office (CZSO). The benefit of the harmonisation rests in improved quality of the published wage statistics. Within the harmonisation, the ISPV population was extended by economic subjects not monitored before. The extension of the ISPV population allowed to calculate more accurate numbers above all on employees, and thus since 2011, all ISPV publications has newly stated weighed numbers of employees. Due to the harmonisation, the gross monthly wage median for the wage sphere decreased in 2011. Despite the harmonisation, there are still differences between both surveys due to the specifics of the ISPV survey.

Keywords

Sample survey, population, sample design, stratification, harmonisation of statistics, wages and salaries

JEL code

C42, C43, J31, J45

INTRODUCTION

Nowadays, almost all the countries in the world face increasing challenges in the generation, analysis, usage and dissemination of statistics in support of decision-making and policy formulation, monitoring and evaluation. In some cases, there is also the problem of lack of comparability and credibility among data sets due to different methodologies employed to collect data. All the measurement tools and indicators would be meaningless without appropriate and reliable statistical data, without strong capacities for statistical development, including the promotion of the development of improved and integrated statistical databases (UN, 2009).

[1] TREXIMA, spol. s r.o., tř. T. Bati 299, 764 21 Zlín-Louky, Czech Republic.
[2] TREXIMA, spol. s r.o., tř. T. Bati 299, 764 21 Zlín-Louky; University of Economics, nám. W. Churchilla 4, 130 67 Prague, Czech Republic. Corresponding author: e-mail: duspivova@trexima.cz, phone: (+420) 257 220 974.

As for the labour market statistics in the Czech Republic, the challenges might be similar as the above mentioned. Trying to measure labour market changes, one can encounter problems concerning lack of information needed although there are many results from labour market surveys available in the Czech Republic. The most likely disadvantage concerning labour market indicators is the fact that each of the key aspects of the labour market (i.e. employment as well as remuneration) is so far surveyed and evaluated separately, although they are closely related (Duspivová, Spáčil, 2011).

In order to handle these imperfections and to get comprehensive and more accurate results of wage statistics used in the Czech Republic, the Ministry of Labour and Social Affairs (MoLSA) deepened its cooperation with the Czech Statistical Office (CZSO), and these two institutions put collective effort into the project of harmonisation of the Average Earnings Information System (ISPV) conducted by the MoLSA with the Wage Statistics conducted by the CZSO.

The main aim of this paper is to investigate the possibility of the harmonisation of wage statistics in the Czech Republic which enables the compilation of data and results comparable between individual statistics involved. We will focus on the results of the above mentioned project from the point of view of both the methodology and possible implications for the Czech wage statistics.

The structure of the paper is as follows: section 1 presents the project, section 2 presents dataset and methods used, the main empirical results are presented in section 3, section 4 describes differences between the ISPV and the Wage Statistics of the CZSO, and the last section concludes the paper.

1 PROJECT OF WAGE STATISTICS HARMONISATION

The main aim of the statistical system is to provide public with reliable and consistent data on different areas of the economic development. As for the wage statistics in the Czech Republic, the status quo before the harmonization was not in full accordance with this aim. In order to contribute to the development of the Czech statistics on labour market, the harmonization of individual wage statistics was essential.

Status quo ante harmonization

As mentioned above, the most likely disadvantage concerning labour market indicators in the Czech Republic is the fact that each of the key aspects of the labour market (i.e. employment as well as remuneration) is so far surveyed and evaluated separately, although they are closely related. What is more, many used and established labour market indicators are the most common and one would think that are the most well-known as far as their information capability is concerned, as well. This assumption has been questioned during the recent economic crisis because a lot of experts have used these indicators in a misleading way.

As far as remuneration is concerned, there are two different data sources available, namely surveys concerning Labour Statistics conducted by the CZSO and the Average Earnings Information System (ISPV) conducted by the MoLSA. Within the Labour Statistics of the CZSO, there are surveyed the number of employees and sum of earnings in the enterprise, so an average gross monthly wage can be calculated. On the contrary, the ISPV gathered data on individual employees in the enterprise, so – in addition to the average wage – the wage distribution is known. The results of individual surveys were impossible to compare with each other not only because of the methodological issues, but also because of the population covered. The users were often confused and compared e.g. average gross monthly wage issued by the CZSO with the proportion of employees remunerated with lower wage than the average gross monthly wage issued in the ISPV. This status quo did not comply with the generally accepted vision of the Czech statistical system, so the MoLSA and CZSO investigated the possibilities of preparing a mutually interconnected system of wage statistics which would eliminate imperfections of wage statistics handled independently and the idea of harmonisation of wage statistics was brought forth.

Main aim of harmonisation and state of the art

The project of harmonisation of wage statistics in the Czech Republic aimed to provide more accurate and reliable data on remuneration. Within this project, the population of the ISPV survey was extended by the missing economic subjects, so the whole employees' population has been covered. The extension of the ISPV population enables – among others – calculation of more accurate numbers above all on employees and thus all ISPV publications has newly stated weighed numbers of employees since 2011.

To achieve the above mentioned aim, economic subjects classified into the institutional sector of households (i.e. own-account workers) with more than 10 employees has been surveyed since the years 2008 and economic subjects classified into the institutional sector of non-profit institutions serving households since 2010. Economic subjects with less than 10 employees were surveyed during the year 2011 with the reference period 2010. In other words, in the end the ISPV population was extended by the employees of economic subjects previously not surveyed, namely by employees of:

– the legal persons with less than 10 employees and
– the natural persons and non-profit institutions regardless of the number of employees.

Within the project of wage statistics harmonisation, the MoLSA has deepened its cooperation with the CZSO as well. These two institutions started interchanging of data on economic subjects concerning remuneration and number of employees which contributed to a successful achievement of the main aim of the project.

2 METHODOLOGY

The methodology developed within the project of harmonisation of wage statistics was fitted to meet the requirements of both wage statistics producers, i.e. the MoLSA and the CZSO. The current statistical approaches concerning sample surveys (e.g. Fuller, 2009, Levy, Lemeshow, 2008, Elliot et al., 2000) offer wide scope of methods but these methods are in most cases narrowly specialized and do not provide for a complex solution of specific needs. For this reason, the methods used within the harmonisation were drawn on the best international experiences as far as individual methods are concerned, and these individual methods were combined to achieve the best-quality results.

The first part of this chapter will devote to basic methods used within the harmonisation. More detailed description of individual methods will follow the first part.

2.1 Basic methods used within the harmonisation

As for the wage statistics harmonisation, the individual methods used were as follows:

– definition of the population involved and sample design;
– simulation of individual employees' records;
– economic subjects' weights and imputation for non-response;
– post-stratification and grossing-up methods.

The population and an actual sample are defined on the basis of both the population taken over from the Wage Statistics carried out by the CZSO and the data surveyed within the ISPV. The most important part of the procedure is the stratification which precedes simulations, weighting and post-stratification. For more detailed information on the ISPV population and the sample see section 2.2.

Since 2011, detailed information on wages and working hours of individual employees has been surveyed in the ISPV every half-year in order to remove the burden from the respondents in odd quarters so that a simulation of individual employees' records is necessary to quantify statistical characteristics (above all quantiles) in each odd quarter. In a given odd quarter, the half-year individual employees' records and the basic data about wages, working hours and number of employees for the whole economic subject are used for simulation which results in simulated individual employees' records in a given period. The above mentioned quantiles are quantified on the basis of these simulated data. As for the economic

subjects with up to 9 employees, the half-year data are simulated as well because of the four-year survey period. More information on the simulation you can find in section 2.3.

The previous activities are followed by the weighting procedure. The weighting aims to assign such a weights to each economic subject in the sample, so that the grossed-up sample is of the same size and structure as the population. For more detailed information on the weighting see section 2.4.

Last but not least, the post-stratification and grossing-up to the population are carried out. In the first step, number of employees in individual strata is grossed up. In the next step, volume of wages is grossed up according to the individual economic activity (CZ-NACE) sections. More information on the post-stratification can be found in section 2.5.

2.2 Population and sample design

The population of the ISPV survey is taken over from the Wage Statistics carried out by the CZSO. Economic subjects (or employees as the case may be) in the population are classified into the wage or salary sphere according to the form of remuneration[3] which is obtained from the data sources of the Ministry of Finance.[4] To minimize the sampling error and to make the best of the organisational structure of the ISPV survey, the sample is divided into following 5 *sectors in the wage sphere:*

- non-financial corporations and financial corporations with 10 and more employees; general government providing remuneration in the form of wages (not salaries),
- households (own-account workers) with 10 and more employees,
- non-financial corporations and financial corporations with 1–9 employees,
- households (own-account workers) with 1–9 employees,
- non-profit institutions serving households with more than 1 employee

and into one sector of economic subjects classified into the salary sphere regardless of the number of employees.

The population is divided into 5 size categories:

- 1–9 employees,
- 10–49 employees,
- 50–249 employees,
- 250–999 employees,
- 1000 and more employees.

The population of the wage sphere is further divided into 6 basic economic activity groups:

- agriculture, forestry and fishing,
- industry and transportation,
- construction,
- wholesale and retail trade,
- market services,
- other services.[5]

In these economic activity groups, the cluster analysis is used to separate more detailed strata using two variables – the wage level and structure of occupations.

---

[3] The wage sphere includes economic subjects who provide remuneration in the form of wages pursuant to Section 109 (2) of Act No. 262 / 2006 Coll., the Labour Code, as amended. Economic subjects belonging to the salary sphere provide remuneration in the form of salaries pursuant to Section 109 (3) of Act No. 262 / 2006 Coll., the Labour Code, as amended.

[4] The population of the ISPV salary sphere (ISPV-PLS) is a set of economic subjects defined by the Automated Budget Information System (ARIS) administered by the Ministry of Finance. ARIS includes information from the accounting and financial reports of the organization units of the state, state budget chapters, state-funded institutions, state funds, territorial self-administration units and state-funded institutions in the Czech Republic.

[5] For NACE sections classified into the individual sectors of economic activity groups see the Annex.

The sample consists of respondents that participated in the survey in a given period (i.e. without nonresponse). For each odd quarter, the economic subjects classified into the salary sphere, and for each period, the economic subjects with up to 10 employees are added (i.e. the sample is completed). The ISPV is a longitudinal survey, so the problems concerning this fact occur. To be more specific, the institutional sector, size category or economic activity sector can change in the long run so the original stratum differs from the actual one. Other problems occur during the process of linking the ISPV records with the CZSO's structural business survey records because the number of employees may not be identical in both surveys. All these issues are verified before the economic subject is classified into the sector for a given period.

Last but not least, both of the wage surveys complement one another. As mentioned above, the population of the ISPV survey is taken over from the Wage Statistics carried out by the CZSO. On the other hand, the data surveyed within the ISPV are used by the CZSO (e.g. the data on large non-profit institutions), so the cooperation of wage statistics producers disburdens the respondents of individual surveys.

2.3 Simulation of individual employees' records

Since 2011, detailed information on wages and working hours of individual employees has been surveyed every half-year so that a simulation of individual employees' records is necessary to quantify statistical characteristics (above all quantiles including median) in each odd quarter. As for odd quarters, the weighted historical imputation method is used. Data on odd quarters are simulated using current individual employees' records (i.e. half-year data). Specific case is the group of economic subjects with up to 9 employees that is surveyed once in a four-year period. Between individual two surveys, individual employees' records are simulated quarterly using, among others, the data from wage statistics carried out by the CZSO. As far as the algorithm and computational capacity are concerned, this part is the most demanding one.

By way of illustration, the simulation is described for the 1st quarter 2011 in brief.

Process of the simulation of the 1st quarter 2011 data

In the 1st quarter, there is no information on employees (i.e. individual employees' records) available because these data are surveyed in even quarters. To deal with the lack of data needed, the historical data from the same period of the previous year are corrected to correspond to the current aggregate data (from the business statistics) on a given economic subject. This approach guarantees that the development trends will not be interfered.

The data extraction from the employees' records surveyed in the previous year

As for the data structure, the most suitable data on individual employees' records were found the half-year 2010 data. These records concern the period from January up to June so that an expert algorithm was used to split this half-year into the 1st and 2nd quarter. To put it simply, the algorithm split wages or worked hours of individual employees into two quarters according to the proportion of working days in given quarters. In addition to that, information on the last day of being included into the staff and on the number of worked days by individual employees is used to split corresponding items into the right quarter.

Forward movement factor

The very extraction of the quarterly individual employees' records from the 1st half-year 2010 is not sufficient, so the forward movement factor is applied on the extracted 1st quarter 2010 data in order that the sum of gross monthly wages, bonuses, worked hours, hours of illness etc. of employees of a given economic subject corresponds to the actual value reported by the economic subject in the current quar-

ter, i.e. in the 1st quarter 2011. This procedure is applied to all ISPV sectors so the individual employees' records are simulated for the population as a whole for the 1st quarter 2011.

2.4 Economic subjects' weights and imputation for nonresponse

The main aim of the weighting procedure is to assign corresponding weights to each economic subject in the sample, so that the grossed-up sample is of the same size and structure as the population. The individual weights of economic subjects are calculated on the basis of the sample probabilities and modified in case of nonresponse (or mistakes made by respondents as the case may be). These weights represent the current structure of the surveyed population.

In the subpopulation of non-financial corporations and financial corporations with more than 10 employees, individual economic subjects are stratified according to the size (10–49, 50–249, 250–999, 1 000 and more employees), sector of economic activity and region (where two groups are distinguished – Prague and the other NUTS 3). Each of the tridimensional strata gets assigned stratum weight, i.e. proportional share in the population via corresponding weights. Each of the economic subjects gets assigned maximum weight according to the corresponding stratum and the sample probability. If the surveyed respondents classified into the strata in question does not represent all the economic subjects in the corresponding stratum, the second stage of weighting follows. In the second stage, the economic subjects of the same size, region and economic activity group are used as a support for weighting. In case of economic subjects with 250 and more employees, weighting is used just to compensate for nonresponse because this subpopulation is supposed to be surveyed entirely.

As for households (own-account workers) with more than 10 employees, a simplified single stage grossing-up method is used. In case of these economic subjects, their size is not important for stratification (special stratum is made just for economic subjects with more than 250 employees).

Non-financial corporations, financial corporations and households (own-account workers) with up to 9 employees are surveyed once every 4 years. In the current year, the economic subjects' weights are used to gross up the results to the universe. In the following years (i.e. in years when the survey is not conducted), the values in question are multiplied by the forward movement factor (in this case by an average increase in wages of similar units surveyed by the CZSO). In association with post-stratification, the appropriate size and structure of this sector are achieved.

Non-profit institutions serving households (NPISH) are divided into two subpopulations. The subpopulation of NPISHs with more than 65 employees is surveyed entirely. As for the subpopulation of NPISHs with less than 65 employees, the sample survey is conducted once every 4 years as a part of small economic subjects survey (see above). Grossing-up procedures are similar to those used for the small economic subjects.

As for the salary sphere, the entire population is surveyed (i.e. sample probability equals 1), so the grossing-up procedures are of not great importance in this sphere. The only adjustments made in the salary sphere concern treatment of nonresponse (see below).

In practice, even a carefully planned survey suffers from nonresponse of some type (Yung and Rao, 2000). Kalton and Kasprzyk (1986) divided nonresponse types into two groups – unit nonresponse and item nonresponse. As for unit (total) nonresponse, it arises e.g. because of refusals, inability to participate, not-at-homes and untraced elements, so none of the survey responses are available for a sampled element. Item nonresponse arises e.g. because of item refusals, "don't knows" and omissions, so some but not all of the responses are available. The distinction between unit and item nonresponse is useful, because different methods are usually used to compensate for the missing data (e.g. Brown, 1990) or to analyze datasets with missing data (e.g. Little, Rubin, 2002). As for the salary sphere, the unit nonresponse is compensated for by weighting adjustments (i.e. the weights of respondents are increased so that they represent the nonrespondents). As for the economic subjects with more than 1 000 employees in the

wage sphere, weighted historical imputation method is used for unit nonresponse. Weighted historical imputation uses previous value multiplied by the forward movement factor (in this case by an average increase concerning given variable in group of similar units) as a current value.

2.5 Post-stratification and grossing-up methods

The closing step of harmonisation is the post-stratification and grossing-up to the population on the basis of the Wage Statistics carried out by the CZSO. The data on wages and employees from the business statistics are passed on from the CZSO to the MoLSA according to an agreement made by the MoLSA and the CZSO. The business statistics carried out by the CZSO covers more economic subjects and the aggregate results are of lower variability in comparison with the ISPV. The post-stratification is realized in two steps.

In the first step, the sample is post-stratified according to the spheres (wage and salary sphere), 6 basic sectors (5 of them in the wage sphere and 1 in the salary sphere), size groups (0–9, 10–49, 50–249, 250–999, 1 000 and more employees) and economic activity (21 CZNACE sections). These strata are combined within the economic activity sections and spheres so that every stratum is representative in a sense of sufficient number of respondents (at least 10 respondents optimally). In these strata, the number of employees (in full-time equivalent) is get using the corresponding weights of economic subjects and these numbers are grossed up to the current population using post-stratification weights. After the first step of the weighting procedure is finished, the number of employees in individual strata corresponds to that of the population.

In the next step, the sample is post-stratified according to the spheres (wage and salary sphere) and 21 CZ-NACE economic activity sections. In these strata, volume of wages is grossed up using the corrective coefficients that are calculated within the sample using weights of economic subjects and post-stratification weights. As in case of number of employees, the volume of wages in individual strata corresponds to that of the population.

3 RESULTS

The above mentioned changes concerned the wage sphere in particular. For this reason, this part will focus on the results of harmonisation in the wage sphere.

Number of employees and wage level

As for the population of the wage sphere, it was extended by the employees of legal persons with less than 10 employees and of natural persons and non-profit institutions regardless of the number of employees within the harmonisation. The main benefit of the harmonization rests in improved quality of the published wage statistics because the extension of the ISPV population allowed to calculate more accurate numbers above all on employees (and thus all ISPV publications has newly stated weighed numbers of employees since 2011). On the other hand, the gross monthly wage median for the wage sphere decreased in 2011 by about 1 900 CZK due to the above mentioned harmonisation (Malenovský, 2011a). To be more specific, the wage decrease was caused by the extension of the population of wage sphere for economic subjects employing fewer employees. The impact of inclusion of smaller (in terms of number of employees) economic subjects is illustrated by the table 1 which shows the number of employees and average gross monthly wage in institutional sectors of the wage sphere divided according to the size of economic subject.

It is evident that wages of employees in economic subjects with less than 10 employees (i.e. in newly included economic subjects in the population) were lower than wages of those employed by economic subjects with 10 and more employees (i.e. by economic subjects included in the population before). As for non-financial and financial corporations, the difference was 8 611 CZK in favour of employees of big-

ger employers in the first half of 2011. As far as the institutional sectors are concerned, the lowest wages were identified in the institutional sector of households, i.e. the employees of own-account workers were paid on average the lowest wage in the first half of 2011 (15 217 CZK).

Table 1 Number of employees and average gross monthly wage in institutional sectors of the wage sphere divided according to the size of economic subject in the first half-year of 2011

Institutional sector (Wage sphere)	Employees (Full time quivalent, thousand)	Average monthly gross wage (CZK)
Non-financial corporations with 10 and more employees and Financial corporations with 10 and more employees	2 261.9	26 717
Non-financial corporations with 1–9 employees and Financial corporations with 1–9 employees	290.4	18 106
Households (own-account workers) with an employee / employees	220.9	15 217
Non-profit institutions serving households with an employee / employees	44.0	20 350
Total	**2 817.2**	**24 828**

Source: ISPV, own calculation

Of course, this methodological inconsistency affects the time series of majority of statistical characteristics and variables of the ISPV. This inconsistency is an inevitable product of harmonisation. In order to be able to carry out a correct comparison of the wage development, the data for 2010 were revised as well and the year-to-year indices were calculated using the revised data for 2010.[6]

Quality of estimations

As the data come from a survey, all the results are sample-based estimates and therefore subject to differing degrees of sampling variability, i.e. the true value for any measure lies in a differing range about the estimated value. This range or sampling variability increases as the detail in the data increases, for example regional data are subject to higher variability than the whole area data (National Statistics for Wales, 2012).

Structural statistics tend to be more detailed thus information on quality of each estimate is demanded by users. And so further to the above mentioned harmonisation, the quality of an estimate of gross monthly wage median was introduced in the ISPV results for the first time in 2011 (see MoLSA, 2011).

The quality of an estimate of gross monthly wage median in the ISPV was inspired by the structural statistics (the Annual Survey of Hours and Earnings, ASHE) produced by the Office for National Statistics in the United Kingdom. To determine the accuracy of the results from the survey, the ONS uses a measure called the coefficient of variation (CV) which is defined as the standard error of the estimate divided by the estimate itself, expressed as a percentage. ONS subsequently class all estimates with a CV of 5 per cent or less as being 'precise'. Estimates with a CV of more than 5 per cent and up to 10 per cent are classed by ONS as 'reasonably precise'; those with a CV of more than 10 per cent and up to 20 per cent are classed as 'acceptable'; and those with a CV of 20 per cent or more are suppressed for quality reasons (ONS, 2011a). In published tables, the ASHE uses colour coding as a quick reference guide to the CV of the estimates (ONS, 2011b). The colour coding indicates the quality of each estimate (e.g. jobs, median, mean and percentiles) but not the annual percentage change.

[6] In addition to that, there were significant changes made as far as classification of occupations is concerned. The original classification KZAM-R was changed in 2011 for the new classification CZ-ISCO. Due to the change of the classification, results according to the occupation specified in the half-year publications for 2011 for the wage (or salary as the case may be) sphere in the tables MZS-M7 and MZS-M8 (or PLS-M7 and PLS M8 as the case may be) cannot be subject of the year-to-year comparison with the results specified in the tables for the business (or non-business as the case may be) sphere until 2010 (tables PS-M1 and PS-M5, or NS-M1 and NS-M5 as the case may be).

As for the ISPV, quality of an estimated gross monthly wage median by job subgroups and categories of the CZ-ISCO classification is divided into 3 qualitative categories "A", "B" and "C". The categories were defined as follows:

- the category "A" characterizes the best quality estimates of the gross monthly wage median with the average standard error of the estimate ± 1.5%,
- the category "B" characterizes the estimates of the gross monthly wage median with the average standard error of the estimate ± 5.0%,
- the category "C" characterizes the estimates of the gross monthly wage median with average standard error of the estimate ± 10.0%.

In addition to that, symbols "A", "B" or "C" stated in brackets identify those jobs in relation to that there is a natural dominance of one or two economic subjects. To be more specific, this concerns those occupations where more than 80% of employees are employed by the only one or two economic subjects.

4 DIFFERENCES BETWEEN ISPV AND WAGE STATISTICS OF THE CZSO CARRIED OVER THE HARMONISATION

As was mentioned above, the ISPV methodology was harmonized with the methodology applied in the Wage Statistics of the CZSO in 2011. Nevertheless there are still differences between both surveys due to the specifics of the individual surveys.

The Wage Statistics conducted by the CZSO aims to describe the Czech labour market from the macroeconomic point of view. It surveys data on gross wages and registered number of employees that is used for calculation of an average gross monthly wage broken down by characteristics of economic subjects (above all by the sector of economic activity according to the classification CZ-NACE). Contrary to the CZSO's Wage Statistics, the ISPV is aimed at structural statistics of earnings and provides users with results concerning the level as well as the structure of earnings and working period of employees in the Czech Republic. Thanks to the ISPV, the development of wages and salaries can be analysed not only from the macroeconomic (i.e. from the economic subjects' perspective), but also from the social perspective (i.e. from the employees' perspective). The ISPV results are broken down by social-economic characteristics of employees (e.g. occupation according to the classification CZ-ISCO, age, sex, education, etc.) as well as by characteristics of economic subjects (e.g. sector of economic activity according to the classification CZ-NACE).

Detailed information about differences between ISPV and the Wage Statistics of the CZSO is shown in the following tables (Tables 2, 3 and 5). Main aims and concepts of individual wage surveys in the Czech Republic are shown in the table 2, methodological differences in Table 3.

It stems from Table 3, that the ISPV results concerning wage statistics are standardized using the full-time equivalent concept of employment which is further adjusted for the fully paid worked hours – it means that all employees temporarily not paid e.g. because of sickness are excluded from the calculation of statistical characteristics concerning wages. What is more, employees with less than one paid month at their employer or / and less than 30 working hours a week are excluded from the ISPV calculations as well. The latter condition is stipulated by the statistical office of the European Union (Eurostat) concerning the Structure of Earnings Survey (see the Council Regulation (EC) No. 530 / 1999 concerning structural statistics on earnings and labour costs, and the Commission Regulation (EC) No. 1738 / 2005 amending the Commission Regulation (EC) No. 1916 / 2000 as regards the definition and transmission of information on the structure of earnings). The standardisation used in the ISPV makes the international comparison of the results possible.

It is evident, that methods of standardisation of results of individual wage statistics differ because the CZSO's method covers all employees, i.e. also the employees temporarily not paid. The impact of the standardisation method used on the statistical characteristics concerning wages is shown in Table 4.

Table 2 Main aims and concepts of individual wage surveys in the Czech Republic

Comparison from the point of view of the	Wage Statistics conducted by the CZSO	ISPV conducted by the MoLSA
Main aim of the survey	1. Macroeconomic indicators 2. Results of the CZSO wage statistics used as reference values for Czech laws and decrees	1. Level and structure of remuneration according to the occupations (CZ-ISCO classification) 2. Harmonisation with the Structure of Earnings Survey (SES)
The most important classifications	1. National version of the Statistical Classification of Economic Activities in the European Community (CZ-NACE) 2. Institutional sectors (Business and non-business sphere) defined according to the ESA (or possibly to the SNA)	1. National version of the International Standard Classification of Occupations (CZ-ISCO) 2. Education 3. Age 4. Sex 5. Citizenship 6. Wage and salary sphere defined according to the Act No. 262 / 2006 Coll., the Labour Code
Survey periodicity	1. Quarterly	1. Quarterly: basic results according to the sector of economic activity (CZ-NACE) 2. Half-year: detailed results 3. Annual: Regional Earnings Statistics (RSCP)

Note: ESA means the European System of Accounts, SNA means the System of National Accounts. For more detailed information see ESA 2010 (or possibly ESA 1995) and SNA 2008 (or possibly SNA 1993). The SES is conducted under the Council Regulation (EC) No. 530 / 1999 concerning structural statistics on earnings and labour costs, and the Commission Regulation (EC) No. 1738 / 2005 amending the Commission Regulation (EC) No. 1916 / 2000 as regards the definition and transmission of information on the structure of earnings).
Source: Own construction

Table 3 Methodological differences between wage surveys carried out in the Czech Republic

Comparison from the point of view of the	Wage Statistics conducted by the CZSO	ISPV conducted by the MoLSA
Statistical unit	Economic subject (employer)	Individual employee, secondarily economic subject (employer)
Sample	A) Economic subjects with 10 and more employees: entire population surveyed (sample probability equals 1) B) Economic subjects with 1–9 employees: sample survey	A) Economic subjects with 250 and more employees in the wage sphere: entire population surveyed (sample probability equals 1) B) Economic subjects in the salary sphere: entire population surveyed (sample probability equals 1) C) Economic subjects with less than 250 employees: sample survey
Standardisation of earnings	Concept of full-time equivalent (FTE)	1. Concept of full-time equivalent (FTE) adjusted for the fully paid worked hours (e.g. hours of sickness are excluded) 2. Exclusion of employees with less than 1 paid month at their employer or / and less than 30 working hours a week
Average gross monthly wage	Gross monthly wage per FTE	Gross monthly wage per FTE adjusted for the fully paid worked hours

Source: Own construction

Table 4 shows three methods of calculation of number of employees in full-time equivalent and of the average gross monthly wage. The first method corresponds to the method used by the CZSO in processing its wage statistics (i.e. no employees' records are excluded). The second one corresponds to the CZSO's method but only those employees' records were taken into consideration that met the requirements of the ISPV (i.e. the auxilliary variable fHMM equals 1). To be more specific, individual records were excluded because of errors, shorter working hours (less than 30 working hours a week) or shorter paid period (less than 1 paid month at individual employer) than required by the ISPV. The last one corresponds to the ISPV method of standardisation, i.e. in addition to the second method the adjustments for the fully paid worked hours were made.

Table 4 A comparison of methods of standardisation used in the wage statistics applied to the ISPV data (the wage sphere, first half-year of 2011)

Sector of economic activity (CZ-NACE classification)	CZSO method		CZSO method modified for fHMM = 1		ISPV method		
	Employees (Full time equivalent, thousand)	Average monthly gross wage (CZK)	Employees (Full time equivalent, thousand)	Average monthly gross wage (CZK)	Employees (Full time equivalent, thousand)	Average monthly gross wage (CZK)	Monthly gross wage median (CZK)
A Agriculture, forestry and fishing	97.2	17 049	89.9	17 674	86.1	18 447	17 073
B Mining and quarrying	35.1	30 001	34.5	30 298	33.2	31 472	27 890
C Manufacturing	1 061.9	22 952	1 019.2	23 429	980.5	24 354	20 731
D Electricity, gas, steam and air conditioning supply	28.3	41 177	27.9	41 478	27.4	42 288	34 404
E Water supply; sewerage, waste management and remediation activities	46.6	22 600	43.8	23 335	42.5	24 047	21 927
F Construction	230.0	21 126	217.3	21 737	206.4	22 888	19 022
G Wholesale and retail trade; repair of motor vehicles and motorcycles	480.6	21 464	455.1	22 066	441.7	22 732	17 649
H Transportation and storage	235.2	22 907	225.0	23 097	217.5	23 901	21 241
I Accommodation and food service activities	104.2	12 388	95.9	12 892	93.1	13 280	10 434
J Information and communication	94.4	43 523	91.3	44 196	89.5	45 074	34 498
K Financial and insurance activities	68.6	49 707	66.5	49 929	65.0	51 084	34 940
L Real estate activities	45.8	20 368	42.4	21 468	41.1	22 159	18 787
M Professional, scientific and technical activities	142.3	27 660	134.3	28 293	131.1	28 989	22 064
N Administrative and support service activities	139.6	16 202	129.9	16 704	124.1	17 482	14 250
O Public administration and defence; compulsory social security	8.8	31 028	8.5	31 471	8.3	32 289	26 095
P Education	61.1	26 916	59.4	27 215	58.4	27 663	24 122
Q Human health and social work activities	120.8	23 330	114.6	23 778	110.9	24 568	21 450
R Arts, entertainment and recreation	23.6	18 938	22.1	19 380	21.7	19 738	16 980
S Other service activities	41.8	18 972	40.1	19 491	38.6	20 216	17 551
Total	**3 065.8**	**23 412**	**2 917.8**	**23 973**	**2 817.2**	**24 828**	**20 085**

Note: fHMM equals 1 for those employees' records that meet the requirements of the ISPV calculations, i.e. only employees with more than 1 paid month at their employer or / and more than 30 working hours a week are included. The error records are excluded of course.
Source: ISPV, Malenovský (2011a)

As for the definition of the monthly gross wage, it is the same for the Wage Statistics conducted by the CZSO as well as for the ISPV conducted by the MoLSA. Monthly gross wage includes basic wage, bonuses, extra pay, wage compensations, overtime bonuses, bonuses for readiness to work on call and other wage components payable by an employer to an employee in return for work done by the latter during the accounting period. Gross monthly wage does not include wage compensations which employers continue to pay to their employees in case of sickness. Monthly gross wage includes values of any social contributions, income taxes, etc. payable by an employee.

As for other variables and statistical characteristics, the wage surveys in the Czech Republic differ. Main differences between the Wage Statistics conducted by the CZSO and the ISPV conducted by the MoLSA shows Table 5. Differences shown in the table 5 stem from the main aims of individual surveys. The ISPV is aimed at the wage differentiation and structure of earnings, contrary to the Wage Statistics conducted by the CZSO. So it is obvious, that the corresponding variables and characteristics are surveyed solely in the ISPV.

Table 5 Main variables and statistical characteristics used in individual wage surveys in the Czech Republic

Comparison from the point of view of the	Wage Statistics conducted by the CZSO	ISPV conducted by the MoLSA
Wage differentiation	Not surveyed	Median, quartiles, deciles
Structure of earnings		Basic wage, bonuses, extra pay, wage compensations, overtime bonuses, bonus for readiness to work on call
Monthly worked and non-worked hours		Monthly worked hours (close attention paid to overtime hours) and monthly non-worked hours (close attention paid to holidays and sickness)
Hourly earnings		Surveyed every half-year

Source: Own construction

CONCLUSION

The project of harmonization of the Average Earnings Information System (ISPV) conducted by the Ministry of Labour and Social Affairs with the Wage Statistics conducted by the Czech Statistical Office (CZSO) was finished in 2011. The main aim of the harmonization of the above-mentioned statistics was to get comprehensive and more accurate results of wage statistics in the Czech Republic.

Within the process of harmonisation, both wage statistics defined the population of economic subjects in the same way for the first time, so the results of individual wage statistics are consistent as far as the number of employees and the aggregate wages are concerned. The ISPV population was extended above all by the employees of the legal persons with less than 10 employees and of the natural persons and non-profit institutions regardless of the number of employees.

Methods developed within the harmonisation of wage statistics were drawn on the best international experiences. To be more specific, simulation, imputation, clustering, (post)stratification and grossing-up techniques were used to achieve the best results. In pursuit of getting the best results, the quality of an estimate of gross monthly wage median was introduced in the ISPV in 2011 further to the harmonization. The quality of an estimate is measured using the average standard error of an estimate.

The main benefit of the harmonisation rests in improved quality of the published wage statistics because of the extension of the ISPV population. On the other hand, the gross monthly wage median for the wage sphere decreased in 2011 by about 1 900 CZK. The wage decrease was caused by inclusion of smaller (in terms of number of employees) economic subjects where the wage level is lower than of those economic subjects already included into the ISPV.

Despite the harmonisation, the differences between wage surveys remain due to the specifics of individual surveys. As for the ISPV, further adjustments are made to comply with the European Union regulation concerning the Structure of Earnings Survey (e.g. those employees are excluded those working hours were shorter than 30 hours a week) so the published figures do not correspond to those of the CZSO at first sight.

The project of harmonisation of the wage statistics in the Czech Republic showed that the joint effort of two institutions can have synergetic effect. On one hand the better results of wage statistics are produced, so the economic policy can work on improved assumptions. On the other hand, cooperation of wage statistics producers may disburden the respondents of individual surveys.

Even if the harmonisation project of the ISPV and the wage statistics conducted by the CZSO is finished, the further development is still the most challenging one as far as algorithms and computational capacity improvements are concerned.

References

BROWN, C. H. Protecting Against Nonrandomly Missing Data in Longitudinal Studies. *Biometrics*, 1990, 46, 1 (March), pp. 143–155.

DUSPIVOVÁ, K., SPÁČIL, P. The Czech Labour Market and the Current Economic Crisis: What Can the Linked Employer-Employee Data Tell Us? *Statistika, Economy and Statistics Journal*, 2011, 48, 4 (December), pp. 22–34.

EC. Council Regulation (EC) No. 530 / 1999 of 9th March 1999 Concerning Structural Statistics on Earnings and Labour Costs. *Official Journal* L 063, 12.3.1999, pp. 6–10.

EC. Commission Regulation (EC) No. 1738 / 2005 of 21st October 2005 Amending the Commission Regulation (EC) No. 1916 / 2000 as Regards the Definition and Transmission of Information on the Structure of Earnings. *Official Journal of the European Union* L 279, 22.10.2005, pp. 32–46.

EC. Commission Regulation (EC) No. 1916 / 2000 of 8th September 2000 on Implementing Council Regulation (EC) No. 530 / 1999 Concerning Structural Statistics on Earnings and on Labour Costs as Regards the Definition and Transmission of Information on Structure of Eearnings. *Official Journal of the European Communities* L 229, 9.9.2000, pp. 3–13.

ELLIOTT, M. R., LITTLE, R. J. A., LEWITZKY, S. Subsampling Callbacks to Improve Survey Efficiency. *Journal of the American Statistical Association*, 2000, 95, 451 (September), pp. 730–738.

ESA 1995. *European System of Accounts*. Luxembourg: Commission of the European Communities, 1999.

ESA 2010. *European System of Accounts*. Luxembourg: Commission of the European Communities, 2010.

FULLER, W. A. *Sampling Statistics*. New Jersey [US]: John Wiley & Sons Inc., 2009.

KALTON, G., KASPRZYK, D. The Treatment of Missing Data. *Survey Methodology*, 1986, 12, pp. 1–16.

LEVY, P. S., LEMESHOW, S. *Sampling of Populations: Methods and Applications*. New Jersey [US]: John Wiley & Sons Inc., 2008.

LITTLE, R. J., RUBIN, D. B. *Statistical Analysis With Missing Data.* 2nd ed. New York: John Wiley & Sons., 2002.

MALENOVSKÝ, L. Metodika harmonizace (Methodology for Harmonization). *Sborník materiálů pro jednání komise pro řízení a koordinaci prací spojených s ISCP. 8. prosince 2011* (Collection of Materials for the Commission to Manage and Coordinate the Work Associated with the ISCP. 8th December 2011). Prague: Ministry of Labour and Social Affairs (MoLSA), 2011b, pp. VII-1–VII-7.

MALENOVSKÝ, L. Vliv harmonizace na změnu výdělkové úrovně (Effect of Harmonization on the Change of Earnings Level). *Sborník materiálů pro jednání komise pro řízení a koordinaci prací spojených s ISCP. 21. září 2011* (Collection of Materials for the Commission to Manage and Coordinate the Work Associated with the ISCP. 21st September 2011). Prague: Ministry of Labour and Social Affairs (MoLSA), 2011a, pp. III-1–III-6.

MoLSA. *Informační systém o průměrném výdělku. 1. pololetí 2011. Mzdová sféra. Výsledky ke dni 21.9.2011* (Information System on Average Earnings. First Half of the 2011. Wage Sphere. Results by September 21st, 2011). Prague: Ministry of Labour and Social Affairs (MoLSA), 2011.

National Statistics for Wales. *Statistical Bulletin SB1 / 2012 Annual Survey of Hours and Earnings (ASHE) for Wales, 2011*. Cardiff: Welsh Government, Knowledge and Analytical Services, 2012.

ONS. *Summary Quality Report for the Annual Survey of Hours and Earnings (ASHE)*. Newport [UK]: Office for National Statistics, 2011a.

ONS. *Statistical Bulletin 2011 Annual Survey of Hours and Earnings*. Newport [UK]: Office for National Statistics, 2011b.

UN. *A Regional Workshop on the Harmonization of Statistics in Southern Africa. Report. Manzini, Swaziland, 3rd–4th December 2009*. United Nations: Economic Commission for Africa, 2009.

SNA. *System of National Accounts 1993*. Brussels / Luxembourg, Washington, D.C., Paris, New York: Commission of the European Communities, International Monetary Fund, Organisation for Economic Cooperation and Development, United Nations and World Bank, United Nations Publication.

SNA. *System of National Accounts 2008*. Brussels / Luxembourg, Washington, D.C., Paris, New York: Commission of the European Communities, International Monetary Fund, Organisation for Economic Cooperation and Development, United Nations and World Bank, United Nations Publication.

YUNG, W., RAO, J. N. K. Jackknife Variance Estimation under Imputation for Estimators Using Poststratification Information. *Journal of the American Statistical Association*, 2000, 95, 451 (September), pp. 903–915.

ANNEX

CZ-NACE sections classified into the individual economic activity groups in the wage sphere

Economic activity group	CZ-NACE section	Title
Agriculture, forestry and fishing	A	Agriculture, forestry and fishing
Industry and transportation	B	Mining and quarrying
	C	Manufacturing
	D	Electricity, gas, steam and air conditioning supply
	E	Water supply; sewerage, waste management and remediation activities
	H	Transportation and storage
Construction	F	Construction
Wholesale and retail trade	G	Wholesale and retail trade; repair of motor vehicles and motorcycles
Market services	J	Information and communication
	K	Financial and insurance activities
	L	Real estate activities
	M	Professional, scientific and technical activities
Other services	I	Accommodation and food service activities
	N	Administrative and support service activities
	O	Public administration and defence; compulsory social security
	P	Education
	Q	Human health and social work activities
	R	Arts, entertainment and recreation
	S	Other service activities
	T	Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use
	U	Activities of extraterritorial organisations and bodies

Source: Own construction